China plays a global game in its bid for COVID-19 vaccines p. 1263

Poverty and mental health p. 1289

Evolution of a contagious cancer in Tasmanian devils p. 1293

Science

$15
11 DECEMBER 2020
sciencemag.org

AAAS

PUSHED TO FAILURE

Probing the limits of battery cathode crystals p. 1313

CONTENTS

11 DECEMBER 2020 • VOLUME 370 • ISSUE 6522

NEWS

IN BRIEF

1252 News at a glance

IN DEPTH

1254 Tweaking genes with CRISPR or viruses fixes blood disorders
People with sickle cell disease or beta-thalassemia could be freed of debilitating symptoms for a lifetime
By J. Kaiser

1255 Spare a thought for the teeming ecosystem beneath your feet
Global review of soil biodiversity calls for protecting often overlooked subterranean life *By E. Pennisi*

1256 HIV researcher named to lead battered CDC
Biden also picks California attorney general for health and human services
By J. Kaiser

1257 Tasmanian devils claw their way back from extinction
Genomes show spread of deadly cancer is slowing *By E. Pennisi*

1258 Scientists rally around plan for fusion power plant
Road map calls for Department of Energy to fund applied as well as basic research
By A. Cho

1259 Europe ramps up defense R&D
A €7 billion fund is launched to bolster meager military spending by EU nations
By N. Wallace

1260 Ecologists push for more reliable research
Psychology's replication crisis inspires efforts to expand samples and stick to a research plan *By C. O'Grady*
PODCAST

1261 Mexico's controversial coronavirus czar
Hugo López-Gatell Ramírez's response has drawn criticism and praise *By R. Pérez Ortega*

FEATURES

1263 China's vaccine gambit
With its global campaign to test and promote COVID-19 vaccines, China aims to win friends and cut deals *By J. Cohen*

INSIGHTS

PERSPECTIVES

1268 Diversity hotspots: Coldspots of speciation?
Phylogeny of suboscine passerine birds finds lower speciation rates in biodiverse regions *By H. Morlon*
REPORT p. 1343

1269 Using genetics to reveal protein structure
Measurements of genetic interactions in vivo can be used to determine protein structure *By D. Wang*
RESEARCH ARTICLE p. 1294

1271 The glassiness of hardening protein droplets
Protein condensates can age to form glasses that increase in viscosity but retain elasticity *By H. Zhang*
REPORT p. 1317

1272 Remembering seasonal coronaviruses
Antibodies against seasonal coronaviruses react with SARS-CoV-2
By J. J. Guthmiller and P. C. Wilson
REPORT p. 1339

POLICY FORUM

1274 Ensuring vaccine safety
Comprehensive safety testing is based on experience with prior vaccines
By D. M. Knipe et al.

1276 Angelika Amon (1967–2020)
Trailblazing cell cycle biologist
By J. M. Sheltzer and R. Visintin

1277 COVID-19 vaccine trial ethics once we have efficacious vaccines
Some placebo-controlled trials can continue ethically after a candidate vaccine is found to be safe and efficacious
By D. Wendler et al.

PHOTO: REUTERS/THOMAS PETER

Published by AAAS

BOOKS *ET AL.*

1280 An outside-the-box take on time
A physicist's provocative theory offers an optimistic view of our cosmic destiny
By M. Johnson

1281 Social capital's value in the courtroom
Inequalities are reinforced during court proceedings and deliberations, finds a sociologist *By L. Seamster*

LETTERS

1282 Renewables in Spain threaten biodiversity
By D. Serrano et al.

1283 Small Aral Sea brings hope for Lake Balkhash
By N. V. Aladin et al.

1283 Brazil's areas of not-so-permanent preservation
By M. A. U. de Andrade Junior and W. C. Fonseca

1284 Technical Comment abstracts

RESEARCH

IN BRIEF

1286 From *Science* and other journals

REVIEW

1289 Economics
Poverty, depression, and anxiety: Causal evidence and mechanisms *M. Ridley* et al.
REVIEW SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY0214

PHOTO: MURRAY COOPER/MINDEN PICTURES

RESEARCH ARTICLES

1290 Molecular biology
CiBER-seq dissects genetic networks by quantitative CRISPRi profiling of expression phenotypes *R. Muller* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB9662

1291 Developmental biology
Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites *J. V. Veenvliet* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABA4937

1292 Developmental biology
Developmental clock and mechanism of de novo polarization of the mouse embryo *M. Zhu* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABD2703

1293 Disease dynamics
A transmissible cancer shifts from emergence to endemism in Tasmanian devils *A. H. Patton* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB9772
PODCAST

1294 Systems biology
Genetic interaction mapping informs integrative structure determination of protein complexes *H. Braberg* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAZ4910
PERSPECTIVE p. 1269

1295 Climate change
Recent global decline of CO_2 fertilization effects on vegetation photosynthesis
S. Wang et al.

1300 Solar cells
Monolithic perovskite/silicon tandem solar cell with >29% efficiency by enhanced hole extraction *A. Al-Ashouri* et al.

REPORTS

1309 Quantum information
Optically addressable molecular spins for quantum information processing
S. L. Bayliss et al.

1313 Batteries
Reversible planar gliding and microcracking in a single-crystalline Ni-rich cathode
Y. Bi et al.

1317 Biophysics
Protein condensates as aging Maxwell fluids
L. Jawerth et al.
PERSPECTIVE p. 1271

1324 Ultracold chemistry
Resonant collisional shielding of reactive molecules using electric fields
K. Matsuda et al.

1328 Cancer immunotherapy
Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer
S. Krishna et al.

1335 Indian monsoon
Indian monsoon derailed by a North Atlantic wavetrain *P. J. Borah* et al.

1339 Coronavirus
Preexisting and de novo humoral immunity to SARS-CoV-2 in humans
K. W. Ng et al.
PERSPECTIVE p. 1272

1343 Evolution
The evolution of a tropical biodiversity hotspot *M. G. Harvey* et al.
PERSPECTIVE p. 1268

1348 Paleoclimate
Southern Ocean upwelling, Earth's obliquity, and glacial-interglacial atmospheric CO_2 change *X. E. Ai* et al.

DEPARTMENTS

1251 Editorial
An opportunity to improve innovation
By H. Holden Thorp

1366 Working Life
Dare to be different
By Chee Kiang Ewe

ON THE COVER

Schematic representation of the gliding phenomenon in a micrometer-sized cathode crystal charged to high voltages. Crystal gliding, induced by localized stress, is partially reversible and eventually evolves into microcracks after extensive battery cycling. A critical crystal size has been predicted to prevent such gliding in battery cathode crystals. See page 1313. *Illustration: C. Bickel/*Science

Science Staff 1250
New Products 1353
Science Careers 1354

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2020 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Published by AAAS

Science AAAS

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Monica M. Bradford

Editors, Research Valda Vinson, Jake S. Yeston **Editor, Insights** Lisa D. Chong

DEPUTY EDITORS Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Di Jiang, Marc S. Lavine (Canada), Yevgeniya Nusinovich, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Katris, Nida Masiulis, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronmel Navas, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Lucy Hicks, Cathleen O'Grady **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Shanghai), Elisabeth Pain (Careers), Charles Piller, Michael Price, Tania Rabesandratana (Barcelona), Joshua Sokol, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Xing Liu **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **WEB DESIGNER** Jennie Pajerowski

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman, Meron Kebede **FINANCIAL ANALYST** Alexander Lee **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SENIOR SALES COORDINATOR** Shirley Young **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Jacob Hedrick, Antoinette Hodal, Lori Murphy **PRODUCTION SPECIALIST** Kristin Wowk **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Tori Velasquez, Mike Romano, Ashley Hylton **DIGITAL MARKETING SPECIALIST** Asleigh Rojanavongse **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **SPJ ASSOCIATE** Samantha Bruno Fuller

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **MARKETING MANAGER** Kess Knight **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer **SALES COORDINATOR** Josh Haverlock

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US WEST COAST SALES** Lynne Stickrod **US SALES MANAGER, SCIENCE CAREERS** Claudia Paulsen-Young **US SALES REP, SCIENCE CAREERS** Tracy Anderson **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Bryony Cousins **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Kaoru Sasaki (Tokyo), Miyuki Tani (Osaka) **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **CONTRACTS AND LICENSING ASSOCIATE** Lili Catlett

MAIN HEADQUARTERS
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
sciencemag.org/authors/science-information-authors

REPRINTS AND PERMISSIONS
sciencemag.org/help/reprints-and-permissions

MEDIA CONTACTS
scipak@aaas.org

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.sciencemag.org/products-services

science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.sciencemag.org/science-careers

advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org

support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

AAAS BOARD OF DIRECTORS
CHAIR Steven Chu
PRESIDENT Claire M. Fraser
PRESIDENT-ELECT Susan G. Amara
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Rosina M. Bierbaum
Ann Bostrom
Stephen P.A. Fodor
S. James Gates, Jr.
Laura H. Greene
Kaye Husbands Fealing
Maria M. Klawe
Robert B. Millard
Alondra Nelson
William D. Provine

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Adriano Aguzzi, *U. Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
Sebastian Amigorena, *Institut Curie*
James Analytis, *UC Berkeley*
Trevor Archer, *NIEHS, NIH*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Clare Baker, *U. of Cambridge*
Nenad Ban, *ETH Zürich*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *UT Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Gabriele Bergers, *VIB*
Kiros T. Berhane, *Columbia U.*
Bradley Bernstein, *Mass. General Hospital*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med.*
Peer Bork, *EMBL*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
György Buzsáki, *NYU School of Med.*
Blanche Capel, *Duke U.*
Annmarie Carlton, *UC Irvine*
Nick Chater, *U. of Warwick*
Ling-Ling Chen, *SIBCB, CAS*
M. Keith Chen, *UCLA*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Carolyn Coyne, *U. of Pitt.*
Roberta Croce, *VU Amsterdam*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *UNC*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Christian Davenport, *U. of Michigan*
Frans de Waal, *Emory U.*
Claude Desplan, *NYU*
Sandra Díaz, *U. Nacional de Córdoba*
Ulrike Diebold, *TU Wien*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Christopher Dye, *U. of Oxford*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Jennifer Elisseeff, *Johns Hopkins U.*
Tim Elston, *U. of North Carolina*
Andrea Encalada, *U. San Francisco de Quito*
Nader Engheta, *U. of Penn.*
Karen Ersche, *U. of Cambridge*
Beate Escher, *UFZ & U. of Tübingen*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Michael Feuer, *GWU*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzl, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Eileen Furlong, *EMBL*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *UCLA*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *U. of South Florida*
Sandra González-Bailón, *UPenn*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Kip Guy, *UK Coll. of Pharm.*
Taekjip Ha, *Johns Hopkins U.*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. Coll. London*
Jian He, *Clemson U.*
Carl-Philipp Heisenberg, *IST Austria*
Ykä Helariutta, *U. of Cambridge*
Janet G. Hering, *Eawag*
Hans Hilgenkamp, *U. of Twente*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Lora Hooper, *UT Southwestern Med. Ctr.*
Fred Hughson, *Princeton U.*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS & U. of Arizona*
Erich Jarvis, *Rockefeller U.*
Kai Johnsson, *EPFL*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Wash.*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *UC Berkeley*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med.*
Nancy Knowlton, *Smithsonian Institution*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Thomas Langer, *Max Planck Inst. Cologne*
Mitchell A. Lazar, *UPenn*
Wendell Lim, *UCSF*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *Wash. U. in St. Louis*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Asifa Majid, *U. of York*
Oscar Marín, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
Jason Matheny, *Georgetown U.*
C. Robertson McClung, *Dartmouth*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Graham Medley, *LSHTM*
Jane Memmott, *U. of Bristol*
C. Jessica Metcalf, *Princeton U.*
Baoxia Mi, *UC Berkeley*
Edward Miguel, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Yasushi Miyashita, *U. of Tokyo*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci., Res. & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Harry Orr, *U. of Minnesota*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Daniel Pauly, *U. of British Columbia*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Institut Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Lei Stanley Qi, *Stanford U.*
Félix A. Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial Coll. London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *UT Austin*
Shimon Sakaguchi, *Osaka U.*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Jürgen Sandkühler, *Med. U. of Vienna*
Erica Ollmann Saphire, *La Jolla Inst.*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Rebecca Sear, *LSHTM*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UCSF*
Robert Siliciano, *JHU School of Med.*
Lucia Sivilotti, *U. Coll. London*
Alison Smith, *John Innes Centre*
Richard Smith, *UNC* **(S)**
Mark Smyth, *QIMR Berghofer*
Pam Soltis, *U. of Florida*
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Dartmouth*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Kimani Toussaint, *Brown U.*
Wim van der Putten, *Netherlands Inst. of Ecology*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Kathleen Vohs, *U. of Minnesota*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Jonathan Weissman, *UCSF*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *UPenn School of Med.*
Jonathan Zehr, *UC Santa Cruz*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

Published by AAAS

EDITORIAL

An opportunity to improve innovation

H. Holden Thorp
Editor-in-Chief,
Science journals.
hthorp@aaas.org;
@hholdenthorp

The rapid development of COVID-19 vaccines is the outcome of Big Pharma's R&D prowess, billions of dollars in federal investment, and the dedication and ingenuity of scores of scientists. But it also illustrates the logic of the Bayh-Dole Act, passed 40 years ago in the United States, which governs how universities can work with industry to reap the benefits of federally funded research. The act permits universities to collaborate with private companies to license and commercialize these technologies under the rationale that the payoff for the government's investment is increased economic activity for the country. It is unlikely that the act will ever be brought back to the floor of Congress for revision, and its staying power suggests that it is a permanent fixture of the U.S. innovation ecosystem. Nevertheless, there are legitimate philosophical and logistical objections that the incoming administration should work to address.

Over the years, the federal government has funded most of the basic research that underlies the COVID-19 vaccines. But the government itself lacks the capacity to carry out massive clinical trials or to manufacture and distribute the vaccines on its own. The Moderna vaccine, for example, relies on patents that are licensed under Bayh-Dole to the biotechnology company. The fact that vaccines could be available to health care workers as soon as this week is a testament to the effectiveness of the arrangement.

When Bayh-Dole first came along, there were notable concerns among university faculties about the conflicts that would arise. Would scientists be objective about their published research if they also stood to gain financially? Would students and postdocs see their careers stalled out because results were held back while patents were filed or—even worse—results were kept secret to protect financial interests?

Universities set up ways to monitor and correct such conflicts, and though there have been problems, the system has held up well and contributed to important innovation. Still, the maintenance of technology transfer offices and conflict monitoring have introduced costs to conducting research that are not fully compensated for by the federal government—costs that have taken resources away from other important university priorities.

When a faculty member holds equity in a startup company, their interests are not completely aligned with those of the university, which can make negotiating licenses cumbersome and strained. When I was a university administrator, we addressed this problem by creating a boilerplate license with standardized terms that could be automatically agreed to when a company was formed. At the University of North Carolina, we called this the "Carolina Express License" and at Washington University, we called it the "Quick-Start License." A few other universities have such policies—notably, Carnegie Mellon and the University of California, San Diego—but unfortunately, this idea has not been widely adopted. The rationale against the boilerplate license is that every deal is different and needs to be separately negotiated, but that leaves the problem of those additional administrative costs, as well as the risk of damaging the university's relationship with the faculty entrepreneur. A number of faculty departures over the years have resulted from sour relations caused by these negotiations.

> "...there are major drawbacks to depending on the marketplace to spur...research ..."

Although Bayh-Dole has produced much economic success and progress on important fronts, there are major drawbacks to depending on the marketplace to spur the kind of research that benefits society—a stated rationale for passing the act. This disconnect provides the strongest argument to create a more public system that doesn't rely on the financial short-sightedness of industry collaboration; however, most attempts at public solutions to this problem have not led to innovations applied outside the public-private model. Thus, the current system does not address what to do when there is insufficient financial interest to attract solutions to problems like antibiotic resistance or unrealized pandemics. As the new Biden administration forms in the United States, a productive effort might be seeking a means of working within the framework of Bayh-Dole to address compelling needs that are not market-driven.

–H. Holden Thorp*

PHOTO: CAMERON DAVIDSON

*H.H.T. declares a number of competing interests that remain updated on his biography on the *Science* website. He is a venture partner at Hatteras Venture Partners, on the board of directors of Artizan Biosciences, and a consultant to Ancora.

10.1126/science.abg0169

Published by AAAS

8848.86 meters

The height of Mount Everest, which straddles China and Nepal and is the world's tallest mountain, as determined by geological survey teams from both countries. It's the first time the nations have agreed on the measurement, which is slightly higher than previous ones.

IN BRIEF

Edited by **Cathleen O'Grady**

A Japan Aerospace Exploration Agency worker retrieves asteroid samples brought back to Earth.

PLANETARY SCIENCE

Hayabusa2 brings back asteroid bits

After lighting up the night sky during its fiery descent through the atmosphere, the return capsule of Japan's Hayabusa2 mission parachuted to a soft landing on 6 December in Woomera, Australia, bearing soil and rock samples from the asteroid Ryugu. Hayabusa2's remote observations of Ryugu indicated it contained organic molecules and water, supporting the theory that similar asteroids may have delivered the building blocks of life to Earth billions of years ago. The capsule will be opened in clean rooms this week in Tokyo to see whether it contains 100 milligrams of material; samples will be shared internationally for further analysis. In 2010, the original Hayabusa mission was the first to return material from an asteroid to Earth. Launched in 2014, its successor, Hayabusa2, has already traveled nearly 5.3 billion kilometers. After dropping off the return capsule, the craft set off for a 2031 rendezvous with the asteroid 1998 KY26.

Journal will require preprints

PUBLISHING | The open-access biology journal *eLife* will only review manuscripts that have been posted on preprint servers—and will publish all peer reviews even for rejected papers, it announced last week. Most of its submissions are already posted as preprints, so this is the next logical step, says longtime open-access publishing proponent Michael Eisen, a biologist at the University of California, Berkeley, and the journal's editor-in-chief. Until July 2021, authors will be able to opt out of the preprint requirement but will be asked to explain why. Authors of a rejected manuscript may postpone publication of the reviews only until the paper is accepted by another journal. "We're pulling back the curtain on a process we all know is happening," Eisen says.

Origins of record dust storm

CLIMATE SCIENCE | A record-setting dust storm that in June swept nearly 15 million tons of dust from Africa's Sahara Desert to the Americas might be tied to a warming Arctic, according to a paper in *Geophysical Research Letters* last week. Formed after an unusual high-pressure weather system parked over Africa's northwestern edge, the storm's winds kicked up the dust. The pressure system arose from low-latitude meanderings in the winds of the polar jet stream that normally circle the Arctic. The cause of such waverings is debated, but some scientists say Arctic warming is to blame for the far-reaching effects.

Congress: Disclose all funding

U.S. SCIENCE POLICY | Scientists applying for federal grants will need to list all existing funding sources, according to a bill nearing final approval in the U.S. Congress. The new rules would unify the current hodge-podge of requirements across a dozen research agencies. The legislation requires all key personnel on a grant application to disclose both public and private sources of support, foreign and domestic. "We don't want to stifle research or discourage partnerships, but we want to know about them," says an aide to the House of

PHOTO: JAPAN AEROSPACE EXPLORATION AGENCY

Published by AAAS

PHOTO: AHMED HASAN/AFP/GETTY IMAGES

The dead do tell tales

Archaeologists recently unearthed more than 100 elaborately decorated wooden coffins at the Saqqara necropolis, a burial complex located 20 kilometers south of Cairo, stacked two and three deep at the bottom of deep underground shafts. The finds date back roughly 2500 years to Egypt's late period, an era separated from the reigns of more familiar pharaohs like Tutankhamun and Ramesses II by more than 700 years of turmoil, civil war, and decline. *Science* spoke to University of Tübingen archaeologist Ramadan Hussein, who was not involved with the latest dig, about what the necropolis reveals about the ancient Egyptians.

Q: Why were there so many coffins at this site?
A: By 650 B.C., Egypt was starting to get back on its feet and become a power in the Mediterranean again. Saqqara once more becomes the main cemetery for a thriving, wealthy city full of temples. Meanwhile, there's an intellectual movement to look back at Egyptian history and revive its traditions. You can see a nostalgia for what was good in Egyptian history in the cemeteries at Saqqara, like inscriptions on the coffins replicating religious texts from the walls of nearby pyramids.

Q: What do the burials tell you about Egypt at the time?
A: In the late period, Egypt has started becoming an international power again, and as a result it's becoming a real mosaic of ethnicities: There are Phoenicians and Greeks and Libyans, and you can see their influence in the grave goods. Trade connections with Greece are intensifying. Many of the coffins at Saqqara are made from expensive wood brought in from southern Europe and elsewhere around the Mediterranean.

Q: Why were they not found in pyramids?
A: Death has become a big business. Discoveries like this tell us about how you administer a cemetery and run the business of death. Priests and undertakers at Saqqara are selling everything from mummification services to burial plots. The ideology of death had shifted. People weren't focused on the size of their tomb; they were happy to be buried in a sacred precinct and a nice coffin.

Representatives Armed Services Committee, which drafted the language as part of a massive bill providing policy guidance to the Department of Defense. Legislators hope the increased transparency will curb China's moves to obtain U.S. technologies.

New verdict on diplomats' illness

SECURITY | A top U.S. science body has concluded that pulsed radio waves may have sickened diplomats in Cuba, lending credence to the controversial notion that an unknown adversary targeted them with a microwave weapon. Starting in 2016, U.S. diplomats in Havana and later in China reported sudden, loud sounds and pressure in the head, followed by symptoms such as dizziness, nausea, and tinnitus. Proposed causes have included the Zika virus, insecticides, or the effects of stress. But a report released last week by a panel convened by the National Academies of Sciences, Engineering, and Medicine describes the symptoms as "unlike any disorder." It asserts that directed, pulsed radio energy is "the most plausible mechanism" for acute symptoms in some of the diplomats. The Department of State "needs a strike team that can move in quickly when a sudden or unusual health event occurs," adds panel member Linda Birnbaum, former director of the National Institute of Environmental Health Sciences.

Cell culture meat wins approval

BIOTECHNOLOGY | California-based food startup Eat Just will be the first to market meat grown from cultured animal cells, the company announced last week. Following approval last month from the Singapore Food Agency, the company plans to sell its cultured chicken nuggets at an unnamed restaurant. The meat is produced by extracting muscle cells from a live chicken and growing them in large quantities in a nutrient-rich bioreactor. Proponents argue that cultured meat is more humane than farmed meat and could reduce global greenhouse gas emissions, 14.5% of which comes from livestock.

Published by AAAS

IN DEPTH

Victoria Gray (right), shown with researcher Haydar Frangoul, was the first patient to be treated with the gene-editing tool CRISPR for sickle cell disease.

PHOTO: AMANDA STULTS/SARAH CANNON RESEARCH INSTITUTE/THE CHILDREN'S HOSPITAL AT TRISTAR CENTENNIAL

BIOMEDICINE

Tweaking genes with CRISPR or viruses fixes blood disorders

People with sickle cell disease or beta-thalassemia could be freed of debilitating symptoms for a lifetime

By **Jocelyn Kaiser**

It is a double milestone: new evidence that cures are possible for many people born with sickle cell disease and another serious blood disorder, beta-thalassemia, and a first for the genome editor CRISPR.

Two strategies for directly fixing malfunctioning blood cells have dramatically improved the health of a handful of people with these genetic diseases, teams reported online on 5 December in studies in *The New England Journal of Medicine* (*NEJM*) and at the American Society of Hematology (ASH) meeting. One relies on CRISPR, marking the first inherited disease clearly helped by the powerful tool created just 8 years ago. And both treatments are among a wave of genetic strategies poised to expand who can get durable relief from the blood disorders. The only current cure, a bone marrow transplant, is risky, and matched donors are often scarce.

The novel genetic treatments have many of the same safety issues as those transplants, and may also be extraordinarily expensive, but there is hope those risks can be eliminated and the costs pared down. "This is an amazing time, and it's exciting because it's happening all at once," says hematologist Alexis Thompson of Northwestern University, who with a company called Bluebird Bio continues to test yet another genetic strategy that first demonstrated a sickle cell fix several years ago.

People born with sickle cell disease have mutations in their two copies of a gene for hemoglobin, the oxygen-carrying protein in red blood cells. The altered proteins stiffen the normally flexible cells into a sickle shape. The cells can clog blood vessels, triggering severe pain and raising the risk of organ damage and strokes. Sickle cell disease is among the most common inherited diseases, affecting 100,000 Black people in the United States alone. (The sickling mutations became widespread in Africa, as one copy protects blood cells from malaria parasites.)

People with beta-thalassemia make little or no functioning hemoglobin because of mutations that affect the same subunit of the protein as in sickle cell. About 60,000 babies are born each year globally with symptoms of the disease, largely of Mediterranean, Middle Eastern, and South Asian ancestry. Blood transfusions are standard treatment for both diseases, relieving the severe anemia they can cause, and drugs can somewhat reduce the debilitating "crises" that often hospitalize sickle cell patients.

In the two new treatments, investigators have tinkered with genes to counter the malfunctioning hemoglobin. They remove a patient's blood stem cells and, in the lab, disable a genetic switch called BCL11A that shuts off the gene for a fetal form of hemoglobin early in life. The patient then receives chemotherapy to wipe out their diseased cells, and the altered stem cells are infused. With the fetal gene now active, the fetal protein restores missing hemoglobin in thalassemia. In sickle cell disease it replaces some of the flawed adult hemoglobin, and also blocks any that remains from forming sticky polymers.

"It's enough to dilute the effect," says Samarth Kulkarni, CEO of CRISPR Therapeutics, which partnered with Vertex Pharmaceuticals on using the genome editor. The companies engineered CRISPR's DNA-cutting enzyme and "guide RNA" to home in on and break the *BCL11A* gene. A team led by gene therapy researcher David Williams of Boston Children's Hospital achieved the same goal with a harmless virus that pasted into the genomes of harvested blood stem cells a stretch of DNA encoding RNA that silences the fetal hemoglobin off switch.

Patients treated in both trials have begun to make sufficiently high levels of fetal hemoglobin and no longer have sickle cell crises or, except in one case, a need for transfusions. In one *NEJM* paper, the Boston Children's team described the success of its virus gene therapy in six sickle cell patients treated for at least 6 months. They include a teenager who can now swim without pain, and a young man who once needed transfusions but has gone without them for nearly 2.5 years, says Erica Esrick of Boston Children's. "He feels perfectly normal."

The first sickle cell patient to receive CRISPR, 17 months ago, a Mississippi mother of four named Victoria Gray, has called the results "wonderful." "We have ameliorated the symptoms," says Haydar Frangoul, a hematologist at the Sarah Cannon Research Institute who treated Gray as part of the CRISPR

Published by AAAS

trial. A beta-thalassemia patient treated with the genome editor 22 months ago is also doing well, the companies report in *NEJM*. Frangoul described positive results for 10 patients with these diseases at the ASH meeting.

The CRISPR results "are really very impressive," says Boston Children's stem cell biologist Stuart Orkin, whose lab discovered the BCL11A switch that led to both trials. (He is not involved with either.)

The results are comparable to those of the Bluebird strategy, which relies on adding a gene for an adult hemoglobin that has been tweaked so it reduces polymerization of the sickling form. Among 32 sickle cell disease patients who received the treatment within the past 3 years, 19 who previously experienced severe pain crises have had none for 6 months or more, Thompson reported at the ASH meeting. The Bluebird treatment was approved in Europe in 2019 for certain beta-thalassemia patients, and the firm expects to seek U.S. approval for both diseases in the next few years. Bluebird Chief Scientific Officer Philip Gregory says the long-term data for the firm's methods is an advantage over other approaches. "We've set a very high bar."

Others who treat these diseases say it's too early to crown a specific genetic treatment the winner. For example, reversing the fetal hemoglobin off switch, as the new CRISPR and RNA-based gene therapy strategies do, allows blood cells to make natural levels of the protein, unlike Bluebird's approach. And although a virus-carrying gene can land in the wrong place and trigger cancer, CRISPR could in theory similarly make harmful off-target edits. "We need long-term follow-up" for all the strategies, says the National Institutes of Health's (NIH's) John Tisdale, a co-leader of the Bluebird study.

These genetic treatments, as currently practiced, seem unlikely to help the many patients worldwide who don't have access to sophisticated health care. Bluebird expects to charge $1.8 million for its beta-thalassemia product in Europe—a sum it bases on a patient's gains in life span and quality of life—and the other genetic treatments are likely to be similarly expensive.

Bluebird and other groups are exploring whether antibodies, instead of harsh chemotherapy, can wipe out a patient's mutant cells. In a bolder effort to reduce costs and increase safety, NIH and the Bill & Melinda Gates Foundation last year announced a plan to each put at least $100 million into developing technologies that would modify blood stem cells directly in a patient's blood marrow. "It's a big hairy goal, but it's an engineering challenge," says gene therapy researcher Donald Kohn of the University of California, Los Angeles, who leads another sickle cell treatment trial. "We'll get there." ■

BIODIVERSITY

Spare a thought for the teeming ecosystem beneath your feet

Global review of soil biodiversity calls for protecting often overlooked subterranean life

By **Elizabeth Pennisi**

Reach down and scoop up some soil. Cupped in your hands may be 5000 different kinds of creatures—and as many individual cells as there are humans on the globe. That random handful might hold microscopic fungi, decomposing plant matter, a whisker-size nematode munching on the fungi, and a predatory, pinhead-size mite about to pounce on the nematode. One bacterium may fend off another with a potent antibiotic. It's a whole world of often overlooked biodiversity.

Last week, on the eve of World Soil Day, the Food and Agriculture Organization of the United Nations released its first ever global assessment of the biodiversity in this underground world. Some 300 experts have pooled their knowledge and data to describe the diversity of these organisms, the roles they play in both natural and agricultural environments, and the threats they face.

A complex dance among microbes, plants, and arthropods and other animals sustains soil ecosystems.

PHOTO: © MATTEO SALA

"The organisms below ground are arguably just as important, if not more important, than what's above ground," says Noah Fierer, a soil ecologist at the University of Colorado, Boulder, who did not contribute to the report. It details how life in the soil boosts crop growth and purifies soil and water. Together with plant root systems, soil organisms store more carbon, potentially for longer, than the aboveground parts of trees do. "Depending on how we handle soil, it could become a help or a burden to face the crisis of biodiversity or climate change," says Francisco Pugnaire, a soil and plant ecologist at the Spanish National Research Council's Experimental Station of Arid Zones.

Yet with each pass of the bulldozer or tractor, each forest fire, each oil spill, even the constant traffic of hikers along a popular trail, more and more soil organisms are being killed off. By compiling research on these subterranean ecosystems and how they affect visible ones, the report's authors hope to convince scientists, policymakers, and the general public to take steps to slow this loss.

"You just can't have a Mars-like soil and expect to maintain the food supply and forests," warns Diana Wall, an ecologist at Colorado State University who contributed to the report. Current conservation efforts are not helping much, she adds. For example, soil biodiversity hot spots aren't necessarily in the same place as the biodiversity hot spots that conservationists focus on. "We are managing [conservation] by what we see above ground, which doesn't necessarily match what's below ground." In contrast, Fierer says, "If you preserve the soil, you will likely preserve the whole ecosystem."

Soil is a mix of organic material, minerals, gases, and other components that provide the substrate for plants to grow. About 40% of all animals find food, shelter, or refuge in soil during part of their life cycle.

Scientists have mostly focused on the largest and smallest soil creatures. For centuries, natural historians have observed the moles, earthworms, ants, and termites that chew, wiggle, and dig their

Published by AAAS

way among soil's particles, feasting on decaying leaves and other debris or on each other. Those ecosystem engineers aerate the soil and create underground passageways that make soil more hospitable for other life. And over the past few decades, microbiologists sequencing soil DNA have discovered an astonishing diversity of bacteria and fungi, which process that litter into organic material.

But in between the scales of macroscopic animals and microbes lie thousands of long-overlooked tiny creatures—the micro- and meso-fauna. Microscopic protists, nematodes, and tardigrades inhabit the watery films around soil particles. Slightly larger animals up to 2 millimeters in size, such as mites, springtails, and insect larvae, live in the airy pores between those particles, helping make soil one of the most diverse habitats on Earth. "How little we know [about these creatures] is a bit overwhelming," Fierer says.

This diversity creates a rich, complex ecosystem that boosts crop growth, breaks down pollutants, and can serve as a nearly inexhaustible sink for carbon. Some soil organisms promote plant diversity and many have yielded important compounds, from antibiotics to natural pesticides. "Without soil organisms and the activities they carry out, it would be impossible for other organisms to survive," says Stephen Wood, a soil ecologist at the Nature Conservancy.

Hidden below ground, these ecosystems seemed immune to surface disturbance, Wood says. "For a long time, soil scientists thought soil microorganisms were so well spread around the world that land management would not harm them," he explains. "We now know that soil microorganisms can be very specific to very specific habitats and species," habitats that are disappearing as farms and cities expand.

The report lists a dozen human activities taking a major toll on soil organisms. They include deforestation, intense agriculture, acidification due to pollutants, salinization from improper irrigation, soil compaction, surface sealing, fire, and erosion. "If you pave over a site, you are sealing off an entire belowground ecosystem," Fierer says. "And that's happening all over the globe."

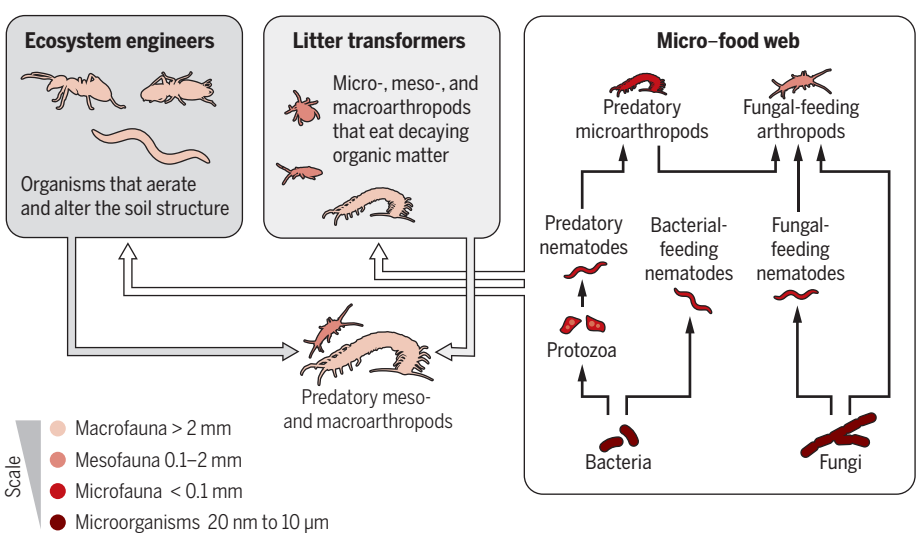
A few governments and companies are making some progress. Several states are considering legislation to help protect soils. In China, the Agricultural Green Development program works to conserve soil by avoiding tilling and by interweaving different crops to preserve biodiversity. "Most organizations want to protect soil biodiversity as a means to an end," Wood points out, to benefit people or aboveground ecosystems.

Some researchers hope the report will encourage protecting soil organisms for their own sake. "Soil biodiversity is huge, and we must not destroy it without knowing what potential there is for improving sustainability," says Mary Scholes, a biogeochemist at the University of the Witwatersrand.

Fierer thinks the new assessment will also awaken a sense of wonder. "My hope is people will look at this document and say, 'Huh, I never thought about [soil organisms] and all the things they do for me.'" ■

The soil is alive

Organisms including ants and termites physically structure soil, while leaf litter creatures and microbes break down organic matter. All these organisms, including nematodes, mites, and springtails, are linked in an intricate food web (arrows show the flow of nutrients).



PEOPLE

HIV researcher named to lead battered CDC

Biden also picks California attorney general for health and human services

By **Jocelyn Kaiser**

President-elect Joe Biden moved this week to fill two top health positions. California Attorney General Xavier Becerra will be nominated as secretary of health and human services, and HIV/AIDS researcher Rochelle Walensky of Harvard Medical School and Massachusetts General Hospital has been selected to direct the Centers for Disease Control and Prevention (CDC).

In Congress, Becerra served on a House of Representatives committee that oversaw Medicare and Medicaid. As California attorney general, he has led efforts to defend the Affordable Care Act and women's reproductive rights.

Biden's selection of Walensky drew praise. "She's an outstanding choice. It's a bright future for the CDC," says James Curran, dean of the Rollins School of Public Health at Emory University, a former director of CDC's HIV/AIDS division. A physician and epidemiologist, Walensky's specialties include promoting HIV testing and modeling the cost effectiveness of using drugs to prevent HIV infections in high-risk groups. This year, she has advocated strong action to control and monitor the spread of the pandemic coronavirus.

Walensky is "an incredible leader in infectious diseases globally and particularly in HIV," as well as a "clear communicator" on COVID-19, says Jen Kates, director of global health and HIV policy at the Henry J. Kaiser Family Foundation.

CDC has struggled to issue clear public health messages about COVID-19 amid pressure from the Trump White House to downplay the pandemic. Walensky will need to help oversee an unprecedented vaccination campaign, and "she will face enormous challenges in upgrading the scientific capabilities and the credibility at CDC," says Barry Bloom, a researcher and former dean of the Harvard T.H. Chan School of Public Health. "I have no doubt she will provide the leadership required," he adds. ■

CREDITS: (GRAPHIC) C. BICKEL/*SCIENCE*; (DATA) U.N. FOOD AND AGRICULTURE ORGANIZATION

Published by AAAS

Biting during mating can transmit a deadly cancer from one Tasmanian devil to another.

PHOTO: ANOEK DE GROOT/AFP/GETTY IMAGES

BIODIVERSITY

Tasmanian devils claw their way back from extinction

Genomes show spread of deadly cancer is slowing

By **Elizabeth Pennisi**

For decades a ghastly facial cancer has been decimating Tasmanian devils. Spreading from animal to animal when the stocky, raccoon-size marsupials bite each other, the transmissible cancer has killed up to 80% of the devils in Tasmania, their only home for millennia. Some researchers saw extinction as inevitable. Now, a new study, reported on p. 1293, suggests the remaining 15,000 devils have reached a détente with the cancer. Until recently it was spreading exponentially, like the pandemic coronavirus among humans in many parts of the world. But geneticists calculate that each infected devil now transmits tumor cells to just one—or fewer—other devils. That could mean the disease may disappear over time.

"It is a promising sign for the future," says Gregory Woods, an immunologist at the University of Tasmania, Hobart, who was not involved with the work. Along with two other recent studies, the findings hint that changes in devil behavior—and possibly the emergence of less virulent tumor cells—may be taming the cancer's spread, and that desperate efforts to breed the devils in captivity may not be needed. "[This] deeper understanding of the links between host behavior and infectious disease may help reveal new insights that can help both devils and other wildlife facing emerging disease threats," says Vanessa Ezenwa, a disease ecologist at the University of Georgia, Athens.

Transmissible cancers are rare in mammals, and devils—whose nocturnal screams and growls earned them their name—are mostly solitary. But starting in 1996, researchers began to notice more and more devils with tumors. Sick animals infected others with cancer cells during mating season fights and scuffles over scavenged carcasses, triggering rapid spread of devil facial tumor disease (DFTD). Tens of thousands of devils died, and conservationists established captive breeding programs to create a reserve population for reintroduction.

The dynamics of the disease are complex, according to a recent study led by cancer geneticist Elizabeth Murchison of the University of Cambridge. Her team genetically analyzed more than 600 tumor samples collected between 2003 and 2018 and found five genomic versions, three of them widespread, with some devils contracting multiple types. That complexity could hamper efforts to develop vaccines to conquer the cancer, according to their 24 November study in *PLOS Biology*.

To further investigate the spread of the rogue tumor cells, Washington State University, Pullman, geneticist Andrew Storfer and his graduate student Austin Patton examined differences in tumor genomes through time, an approach routinely used to trace the spread of viruses including SARS-CoV-2, the pandemic coronavirus. A mammalian tumor genome is much larger than that of a virus, so Patton and colleagues had to work out ways to analyze their data, gleaned from 51 tumors collected from 2003 to 2018.

Focusing on 28 genes that seemed to be evolving at a consistent rate, they traced how specific mutations spread through the tumor samples over time. That enabled them to infer the rate at which the cancer itself was spreading among devils. "The application of these methods for transmissible cancer is very clever," says Michael Metzger, a molecular biologist at the Pacific Northwest Research Institute.

At the disease's peak in the early 2000s, each infected devil spread the disease to at least 3.5 others, the team reports in *Science*. But transmission has slowed recently, with some infected animals not passing DFTD on at all. The reduced density of devils accounts for much of the decline, Patton suggests, as animals come into contact with fewer of their fellows. Remaining devils may also have better immune systems or altered behavior, speculates Patton, now a postdoc at the University of California, Berkeley.

A study out on 9 December in the *Proceedings of the Royal Society B* supports the idea that an animal's behavior can slow transmission. Disease ecologist Rodrigo Hamede and behavioral ecologist David Hamilton from the University of Tasmania, Sandy Bay, did devil contact tracing: For 6 months they put radio collars on 22 devils that revealed when an animal came into close contact with another. The tracking showed that once infected, even dominant, aggressive devils withdrew from others as they became sicker. These individuals were "superspreaders" only early in the mating season, Hamilton, Hamede, and their colleagues report. "The fact that they behave in this way is likely to have a big impact on disease dynamics," Hamilton says.

The *Science* study authors argue against plans to introduce captive-bred devils into remaining wild populations. Beefing up devil populations may increase their density and rev up transmission again, and captive-bred animals may lack resistance built up in wild populations, Storfer speculates.

Though this week's news is good, "devils are still not out of the woods," warns conservationist Max Jackson of Aussie Ark, which helps breed captive devils. Indeed, researchers detected a second transmissible facial cancer in devils in 2014. But the new findings offer hope, Hamilton says. "It looks extremely unlikely that we'll be losing them any time soon." ■

Published by AAAS

FUSION

Scientists rally around plan for fusion power plant

Road map calls for Department of Energy to fund applied as well as basic research

By **Adrian Cho**

U.S. fusion scientists, notorious for squabbling over which projects to fund with their field's limited budget, have coalesced around an audacious goal. A 10-year plan presented last week to the federal Fusion Energy Sciences Advisory Committee is the first since the community tried to formulate such a road map in 2014 and failed spectacularly. It calls for the Department of Energy (DOE), the main sponsor of U.S. fusion research, to prepare to build a prototype power plant in the 2040s that would produce carbon-free electricity by harnessing the nuclear process that powers the Sun.

The plan formalizes a goal set out 2 years ago by the National Academies of Sciences, Engineering, and Medicine (*Science*, 21 December 2018, p. 1343) and embraced in a March report from a 15-month-long fusion community planning process. It also represents a subtle but crucial shift from the basic research that officials in DOE's Office of Science have favored. "The community urgently wants to move forward with fusion on a time scale that can impact climate change," says Troy Carter, a fusion physicist at the University of California, Los Angeles, who chaired the planning committee. "We have to get started."

Fusion scientists and DOE officials strived to avoid the sort of meltdown they suffered during their last planning exercise. Six years ago, the fractious community was already reeling from budget cuts that forced DOE's Fusion Energy Sciences (FES) program to shutter one of three major experiments. Then, the associate director for FES decided to write the plan himself, with limited input. Many researchers rejected the road map (*Science*, 19 December 2014, p. 1436).

This time, DOE wants no infighting. "We've been told in no uncertain terms that either you guys get in line, or you're going to get nothing," says Nathan Howard, a fusion physicist at the Massachusetts Institute of Technology. For the first time, FES leaders let researchers hash out consensus in a series of workshops and meetings. Howard and other leaders of that process used anonymous polling and even hired a facilitator to ensure the "loudest voices in the room" couldn't dominate deliberations.

The process was also comprehensive, says Carolyn Kuranz, a plasma physicist at the University of Michigan, Ann Arbor. FES mainly funds research on magnetically confined fusion, in which an ionized gas or plasma is squeezed and heated until atomic nuclei fuse and release energy. But it also supports smaller efforts in plasma physics, such as using high-power lasers to re-create plasmas like those in stars. The consensus building did not neglect them. "This was the first time we included the whole portfolio and the entire community," Kuranz says.

Fusion wish list

U.S. researchers have agreed on the need for projects that would aid a future power plant (gray) and advance basic plasma science. However, funding limits could curtail plans.

PROJECT	FLAT BUDGETS	2% INCREASES	UNCONSTRAINED
Neutron source to test materials for fusion power plant	Yes, but highly delayed	Yes, but delayed	Yes
Tokamak to test integrated systems for fusion power plant	No	Yes, but highly delayed	Yes
Facility to test "blanket" that would surround reactor and absorb neutrons	No	No	Yes
Matter in Extreme Conditions Upgrade	No, but develop further	No, but develop further	Yes
Solar wind facility	No	No	Yes
Multipetawatt laser	No	No	Yes

The plan that emerged does not call for a crash effort to build the prototype power plant. During the next decade, fusion researchers around the world will likely have their hands full completing and running ITER, the international fusion reactor under construction in southern France. ITER, a huge doughnut-shaped device called a tokamak, aims to show in the late 2030s that fusion can produce more energy than goes into heating and squeezing the plasma.

ITER will teach valuable lessons about a "burning plasma," researchers say. But they add that its cost of more than $20 billion is far too steep for an actual power plant. So, after ITER, U.S. fusion researchers want to build a much smaller, cheaper power plant, leveraging recent advances such as supercomputer simulations of entire tokamaks, 3D printing, and magnet coils made of high-temperature superconductors.

The new fusion road map identifies technological gaps and nearer-term facilities to fill them (see partial list, below). "By identifying [a power plant] as a goal, that can trigger more research in those areas that support that mission," says Stephanie Diem, a fusion physicist at the University of Wisconsin, Madison. For example, in a fusion power plant a barrage of energetic neutrons would degrade materials, so the report calls for developing a particle-accelerator–based neutron source to test new ones.

Such technology development pushes a sensitive boundary for the fusion program. Fusion investigators have long complained that DOE's Office of Science has limited them to basic research. Now, DOE leaders are more receptive to a practical approach, says James Van Dam, DOE's associate director for FES. "There's been much more openness and interest in fusion moving ahead."

To realize their ambitions, fusion scientists will need more funding from Congress. The planning committee considered three scenarios: flat budgets, increases of 2% per year, and unconstrained budgets. Only the most generous scenario would allow DOE to build new facilities, the report says. FES's annual budget is now $671 million, including $247 million for ITER.

Tighter budgets might strain the newfound consensus. Plasma physicists want several new facilities, such as one to simulate the solar wind. But without a funding boost, they won't even be able to build a project DOE has already said it wants: the Matter in Extreme Conditions Upgrade, which would improve a petawatt laser at the SLAC National Accelerator Laboratory to create energetic plasmas so they can be probed with the lab's x-ray laser.

No matter how things play out, the fusion plan expresses the will of younger scientists who led the community exercise, says Scott Baalrud, a plasma theorist at the University of Iowa. "People don't get into this career just to study the science that may one day, long after they're dead, lead to a fusion reactor," he says. "They want to get going and change the world." ■

DATA: *POWERING THE FUTURE FUSION AND PLASMAS*, FUSION ENERGY SCIENCES ADVISORY COMMITTEE (2020)

Published by AAAS

Drones will be one of the research topics supported by a growing European Defense Fund.

MILITARY RESEARCH

Europe ramps up defense R&D

A €7 billion fund is launched to bolster meager military spending by EU nations

By **Nicholas Wallace**

This summer, in a leafy, wooded area near Utrecht, the Netherlands, scientists were testing out battlefield haute couture: adaptive camouflage. Researchers with the Netherlands Organisation for Applied Scientific Research mounted a swath of fabric on a stand and watched as its pattern shifted to match the greens and browns of the foliage. Cameras connected to the fabric picked up the scenery and hundreds of embedded light-emitting diodes mimicked it, like the skin of a chameleon. The team is testing other materials to weave into the futuristic camouflage, including polymers that absorb body heat and radio waves, making soldiers harder to detect with thermal imagers and radars.

Just as striking as the fabrics is the project's funding source: the European Union, better known for trade rules and farm subsidies than military maneuvers. The camouflage work is part of a Swedish-led, six-country project that received a €2.6 million grant from the union's Preparatory Action on Defence Research (PADR). The 3-year fund, worth €90 million, also supports research in artificial intelligence (AI) for bomb detection, laser weapons, railguns, and drones. It is a sign of much bigger things to come: Next year, PADR will be rolled into the new European Defense Fund (EDF), with a budget of €7 billion over 7 years, split between early-stage research and late-stage development.

That's tiny compared with the $80 billion per year the United States spends on defense R&D. And it's even small compared with the combined €5 billion or so spent on defense research each year by EU nations. But the European Union, which has no military resources of its own, hopes the EDF, by topping up joint national investments, will encourage its members to strengthen their modest defense capabilities. For European researchers, the spending is opening new opportunities—and stirring some qualms.

European governments slashed defense budgets in the 1990s, believing the danger of major conflict in Europe had ended with the Cold War. "It was a bit of a paradox, because of course we had the Balkan wars," says Zdzisław Krasnodębski, a member of the European Parliament for Poland's governing Law and Justice Party who steered the EDF through Parliament. But recent events—particularly Russia's annexation of Crimea in 2014 and the subsequent war in Ukraine—have changed many minds, he says.

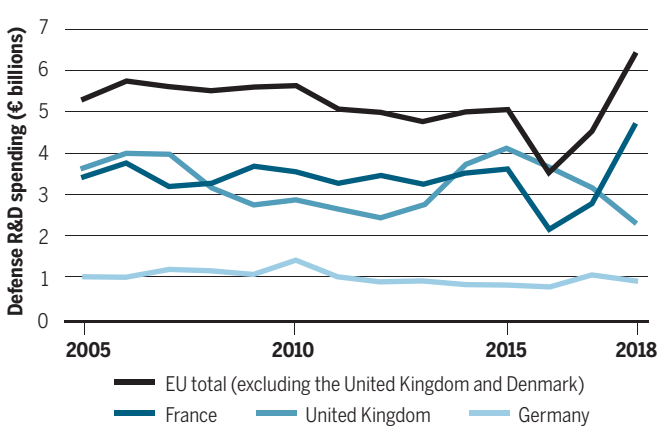
At the same time, the United States is withdrawing from its role as guarantor of European security, says Julia Muravska, a researcher at the RAND Corporation, a global defense policy think tank. The Obama administration shifted U.S. defense resources to the Far East, and both it and the Trump administration have announced troop withdrawals from Europe. Successive U.S. presidents have urged the 21 EU countries that are also North Atlantic Treaty Organization members to honor commitments to spend at least 2% of their gross domestic product on defense. But only three actually do: Latvia, Estonia, and Greece.

To address these gaps, European leaders are discussing ways to pool military resources, and French President Emmanuel Macron has even called for "a true European army." The EDF aims to be another stimulant for collaboration. Every EDF project must include participants from at least three nations, and mandatory cofinancing for late-stage development projects will give the budget a "lever effect," says Frederic Mauro, a Brussels-based lawyer specializing in defense. Even though the EDF is minuscule compared with U.S. spending, it "can make a huge difference at the European scale," he says.

One unanswered question is whether researchers in non-EU countries such as the United Kingdom and Switzerland can apply for EDF funding, and on what terms. If legislators allow

Going great guns

Most EU nations, with the exception of France, spend very little on defense R&D. EU officials hope a new defense R&D fund, worth about €1 billion a year, will boost national spending.



CREDITS: (PHOTO) FINNBARR WEBSTER/GETTY IMAGES; (GRAPHIC) C. BICKEL/*SCIENCE*; (DATA) EUROPEAN DEFENCE AGENCY

Published by AAAS

their participation, the rules will likely be stricter than for Horizon 2020, the European Union's civilian research program, which already includes non-EU nations like Switzerland and Israel. Jean-François Ripoche, R&D chief for the European Defence Agency, which runs PADR, says the argument boils down to whether a foreign firm or institution can reliably contribute to EU military projects without interference from its home country. "In the defense business, it's hard to do without talking to your own government," he says.

PADR funding has gone to a mix of research institutes, including Germany's Fraunhofer Society; engineering companies such as French giant Thales; weaponsmakers like MBDA; and smaller businesses, such as Space Applications Services, a Belgian research firm. Universities also participate: For example, the University of Siena leads another camouflage project.

Future EDF research topics will be specified in annual calls run by the European Commission, the EU executive branch, and approved by a committee of national delegates. AI will be a big topic, Ripoche says. He says EDF funding will also go to new materials, such as discreet metamaterial antennas that can be engineered into the surfaces of vehicles and weapons. Muravska says she expects "a healthy take-up" in the EDF by European academic researchers, "provided they are aware of it."

But some scientists are uneasy. "Military research can feed into technologies which are then exported to countries with poor human rights records," says Stuart Parkinson, director of Scientists for Global Responsibility, a U.K. advocacy organization. He adds that nuclear arms races could be stoked by development of ostensibly nonnuclear technologies, such as hypersonic weapons.

Reiner Braun, a board member of the Berlin-based International Network of Engineers and Scientists for Global Responsibility, says he wishes EDF money could be used for research into understanding and defending against poverty, climate change, and disease, which can also lead to conflict. "We need much more research for peaceful activities for supporting the sustainable development goals, for climate research, and for many other purposes, including the fight against the pandemic."

European security should come through dialogue with Russia and multilateral disarmament, Parkinson says. But he recognizes that a weapons-free world remains a far-off dream. "I'm not saying we should throw the doors open and throw our weapons on the floor." ■

Nicholas Wallace is a journalist in Brussels.

RESEARCH INTEGRITY

Ecologists push for more reliable research

Psychology's replication crisis inspires efforts to expand samples and stick to a research plan

By **Cathleen O'Grady**

Ecologists love to study blue tits. The birds readily nest in boxes in the wild and have striking plumage that seems ideal for testing ideas about the evolutionary point of the ornamentation. Dozens of studies have reported that male coloring is substantially different from that of females, that females choose mates based on differences in that coloring, and that male plumage is a signal of mate quality.

But Tim Parker, an ecologist at Whitman College, wasn't so sure. In a 2013 meta-analysis of 48 studies on blue tit plumage, Parker found many researchers had cherry-picked the strongest findings from data they had sliced and diced. They had worked backward from results to form hypotheses that fit the data. And reams of boring, negative results were missing from the published picture. There was no reason to think these problems were limited to blue tits, Parker says: "I just became convinced that there was a lot of unreliable stuff out there."

Parker soon found an ally in Shinichi Nakagawa, an ecologist at the University of New South Wales with similar concerns. "It's an existential crisis for us," Nakagawa says. The two began to publish on the issue and gathered more collaborators. That has culminated with the launch last week of the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE), dedicated to connecting ecologists who want to add rigor to their field. SORTEE draws inspiration from the Society for the Improvement of Psychological Science (SIPS)—another discipline that has wrestled with reliability. SORTEE plans to host satellite events at ecology conferences and, eventually, its own meetings.

Although SORTEE's agenda will be set by its members, Parker says that, like SIPS, it could offer statistics training, build collaborations, and support metaresearch on the health of the discipline. Yolanda Wiersma, a landscape ecologist at the Memorial University of Newfoundland who is not involved with SORTEE, is eager to see whether the society makes a difference. Research credibility is something "we haven't wrapped our heads around completely as ecologists," she says.

Ecology suffers from many of the same underlying problems as psychology. Surveys of the ecology literature have found that small sample sizes are common, often driven by high cost or limited access to a species or other study system. In landscape ecology, each landscape is unique, meaning the sample size is one, Wiersma says. "There is one Yellowstone park," she says. "There's one Lake District." Small samples lead to erratic results that sometimes miss the effects researchers are looking for and other times hit on noise that looks like a real signal.

Worsening those problems are "questionable research practices," says Fiona Fidler, a metascientist at the University of Melbourne. In a 2018 study published in *PLOS ONE*, Parker, Fidler, and colleagues reported on a survey of more than 800 ecologists and evolutionary biologists. About half of the respondents said they sometimes presented unexpected findings as if they confirmed a hypothesis they'd had all along, and about two-thirds said they sometimes reported only significant results, leaving out negative

PHOTO: NHPA/PHOTOSHOT/SCIENCE SOURCE

Published by AAAS

Flaws in studies about blue tit plumage motivated a push for more rigor in ecology.

ones. Together, these forces mean a literature overflowing with potentially dubious results, Parker says. It's a "house of cards."

But unlike psychology, in which researchers have tried to replicate famous studies and failed in about half the cases, ecology has no smoking gun. A 2019 *PeerJ* study found only 11 replication studies among nearly 40,000 ecology and evolution biology papers—and only four of these 11 studies managed to replicate the original finding. It's hard to replicate ecology studies, Parker says, because it often entails expensive and difficult data gathering in remote places or over long time frames. And ecosystems are so complex that any number of variables could affect the outcome of a repeat experiment—like the age of the organisms in the study, the temperatures at the time, or the presence or absence of pollutants. "No man can step into the same river twice because it's not the same man and it's not the same river," says Phillip Williamson, an ecologist at the University of East Anglia who has criticized a high-profile effort to replicate ocean acidification research.

Yet Williamson doesn't think ecology as a whole is at risk just because some experiments fail to replicate. "Biology isn't physics," he says. "I think that the consensus of science gets there eventually." Parker takes a harder line. "If we don't expect anything to replicate, why do we bother doing any of this?" he asks.

Even before they set up SORTEE, Parker and his corevolutionaries were pushing for change. They worked with journal editors to create checklists for details that papers should include—like whether researchers were blinded to the conditions of different subject groups. They've also set up a preprint server that Nakagawa hopes will help preserve results that never make it into journals. Julia Jones, a conservation scientist at Bangor University who is not involved with SORTEE, is advocating for preregistration, which forces a researcher to commit to a data collection plan and hypothesis before the study begins. Some journals offer registered reports—peer-reviewed preregistrations with a commitment to publish the results, however dull or dazzling. Preregistration isn't always possible, because the vagaries of fieldwork often force researchers to change plans. But she says it can help scientists avoid the "siren song" of looking for a clean story in messy data.

In April, Jones and her colleagues published the first registered report for the journal *Conservation Biology*. She analyzed extra data from a randomized controlled trial in Bolivia's highlands that had already found that paying farmers to keep their cattle out of rivers did not improve water quality. Jones found other interesting behavior changes—farmers kept their cattle on their farms rather than roaming the forests—but many results were statistically insignificant. In a normal review process, "we would have been forced to cherry-pick and tell a much simpler story," she says.

Others are working to address the sample size problems, by gathering massive amounts of data using consistent methods. They hope the data sets will make it easier to see which findings apply beyond a single ecosystem. The U.S. National Ecological Observatory Network (NEON), a continentwide program of more than 100 heavily instrumented field sites, became fully operational in 2019, and the first studies drawing on its data are now underway.

The Nutrient Network (NutNet), cofounded by University of Minnesota, Twin Cities, ecologist Elizabeth Borer, also pulls in large amounts of standardized data, to explore how changes in nutrients and herbivores affect grassland plant diversity. Rather than building infrastructure like NEON, NutNet gets research teams around the world to perform the same experiments—in return for access to a huge data set. Unifying experiments is hard, Borer says. For instance, the team discovered that the fertilizer brand Micromax had slightly different micronutrient mixes on different continents, forcing researchers to import or mix their own.

Borer, Wiersma, and Jones are all sympathetic to SORTEE's aims—and curious to see whether it takes off. Like the systems they study, ecologists can be fragmented, and developing sound research principles sometimes seems impossible, Wiersma says. "But I think maybe we could," she says. "We just need to try harder." ■

VOICES OF THE PANDEMIC

Mexico's controversial coronavirus czar

Hugo López-Gatell Ramírez's response has drawn criticism and praise

By **Rodrigo Pérez Ortega**

There's hardly a Mexican who doesn't know Hugo López-Gatell Ramírez by now. Mexico's undersecretary of prevention and health promotion has sat across from reporters at 7 p.m. sharp almost every single night since late February to update them, and the country, on the toll of the coronavirus pandemic. His firm demeanor, careful speech, and courteous personality have made his televised coronavirus press briefings even more popular than those of the country's president.

But as COVID-19 deaths in Mexico continue to soar—surpassed only by the United States, Brazil, and India—many have questioned López-Gatell Ramírez's leadership. Critics accuse him of undercounting the true numbers and mishandling the nation's response. In early August, the governors of nine Mexican states demanded his resignation. His defenders, though, say he's making sound decisions based on science and doing the best he can with the resources at his disposal.

López-Gatell Ramírez says the country's chronically underfunded public health system has complicated efforts to track the disease and coordinate the response. But he insists the government's approach—which includes forgoing widespread testing in favor of a more targeted approach—has been based on science. Many critics, he says, are politically motivated. He's been sleeping 3 to 4 hours a night since the pandemic started, he says, and he regrets the disruption to his family life. "The mission calls me and until I deliver results—I hope favorable—I cannot stop."

Fame might be new to López-Gatell Ramírez, but pandemics are not. Just 2 years after finishing his Ph.D. in epidemiology in 2006 at Johns Hopkins University, he became head of epidemiology at the Mexican health ministry. One year later, in 2009, the swine

Published by AAAS

flu pandemic struck Mexico, and López-Gatell Ramírez helped lead the response.

The H1N1 virus, which caused panic around the world, likely originated on a pig farm in Mexico. Mexico City was shut down for 2 weeks, and the country scrambled to curb the virus' spread. López-Gatell Ramírez says the experience taught him many lessons, including the dangers of poor coordination among government institutions and "political opportunism" by officials trying to use the crisis to their advantage. This time, he says, President Andrés Manuel López Obrador has given his team decision-making power and protected them from political meddling.

Faced with the new pandemic, López-Gatell Ramírez placed his trust in the country's "sentinel" system, originally designed for flu outbreaks. In contrast to the massive testing and contact tracing efforts in many other countries, the system relies on a small, strategic, nationwide sample of tests, which are then used to predict the spread of a disease. As a result, even though Mexico's flagship diagnostics laboratory—the Institute of Epidemiological Diagnosis and Reference (InDRE)—developed the world's first polymerase chain reaction protocol to detect SARS-CoV-2, the country ranks among the world's lowest in per-capita testing. Mexico has performed about 17 tests per thousand people in total, whereas the United States has done more than 30 times that many.

Reliance on the sentinel model may have prevented López-Gatell Ramírez and his team from making accurate predictions. They forecast the pandemic would peak around 8 May, with 4500 daily cases, aided by a national voluntary lockdown in effect from late March until June. (The lockdown was voluntary because many Mexicans live in poverty and rely on daily work to survive.) Instead, daily cases kept climbing, to a peak of 9556 on 3 August. Since then, the number has fluctuated between 3000 and 12,000, and the pandemic is worsening. The official death count now exceeds 110,000, nearly twice the model's original worst case.

Researchers, politicians, and national and international media have suggested the true toll is even greater. As in almost every country, not every COVID-19 death is registered. But the rift between statistics and reality may be particularly wide in Mexico. In late November, officials released a report of excess deaths from all causes, suggesting nearly 156,000 of them could be attributed to COVID-19 in people who were never tested. The rate of positive tests recently hit 47%, evidence of a rampaging epidemic.

Some observers think scarce funding was the real reason for the limited testing. Alejandro Macías, an infectious disease physician and Mexico's former H1N1 czar, doubts López-Gatell Ramírez truly believed the approach would be successful. "He got trapped in a system in which the InDRE was not given sufficient resources or all the money that was going to be needed for testing," Macías says.

"You can buy thousands of tests, but if you don't have the personnel [to perform them], you can't do much," adds Celia Mercedes Alpuche Aranda, an infectious disease researcher at the National Institute of Public Health and former director of InDRE. López-Gatell Ramírez's strategy, she says, is adapted to the country's reality: too many people to test and not enough infrastructure.

Testing isn't the only area where López-Gatell Ramírez has drawn criticism. As Europe and the United States were dealing with their first waves of infections, he advised against closing Mexico's borders. Doing so would harm the economy, he said, but would not prevent the pandemic from arriving—a view shared by the World Health Organization and others at the time. Many epidemiologists now concede that closing borders can help stem viral spread.

His stance on face masks is perhaps his most controversial: Despite growing pressure from the public, politicians, and scientists citing evidence that masks can reduce viral transmission, López-Gatell Ramírez still hasn't mandated their use nationwide. Masks would give a "false sense of security" and lead people to relax other measures such as hand washing and keeping a healthy distance, he said early in the pandemic. Even as evidence of the benefits of wearing masks became clear, he still has not fully come around, critics say. Some Mexican states, however, have imposed their own mask mandates.

Macías suspects López-Gatell Ramírez's statements on masks have been influenced by the views of his boss, López Obrador, who refuses to wear a face mask in public and constantly downplays the pandemic. "He had to defend the indefensible," Macías says.

Despite the adversities, the Mexican health system has scored some victories under López-Gatell Ramírez. Starting with a deficit of more than 310,000 health care workers, he says, the system hired more than 47,000 people in a matter of months and tripled the number of intensive care beds with ventilators.

"I think he's the right person to be in that position," says infectious disease physician Gustavo Reyes Terán, who leads the country's National Institutes of Health and its affiliated hospital network. The institutes and hospitals Terán coordinates avoided becoming overwhelmed, he says, thanks to López-Gatell Ramírez. "That, to me, has been one of the most important successes here," he says. (Media have reported, however, that many Mexicans died at home without setting foot in a hospital.)

López-Gatell Ramírez has also taken action against one cause of the virus' toll in Mexico: high rates of hypertension, obesity, and diabetes, which can make severe disease more likely. In response, López-Gatell Ramírez has renewed his long-standing campaign against junk food and sugary drinks, calling them "bottled poison."

"He is a person with good intentions to do things well, to be based on science," Alpuche Aranda says. "He's doing the best he can," Macías adds.

How history judges López-Gatell Ramírez, however, may ultimately depend on the pandemic's toll on his nation. ■

"The mission calls me and until I deliver results—I hope favorable—I cannot stop."

Hugo López-Gatell Ramírez, Mexico's undersecretary of prevention and health promotion

Science's COVID-19 reporting is supported by the Pulitzer Center and the Heising-Simons Foundation.

Rodrigo Pérez Ortega is a journalist in Mexico City.

ILLUSTRATION: KATTY HUERTAS

Published by AAAS

Beijing-based Sinovac, which makes a COVID-19 vaccine from inactivated virus, has efficacy trials underway in Brazil, Turkey, and Indonesia—but not in China.

PHOTO: REUTERS/THOMAS PETER

CHINA'S VACCINE GAMBIT

With its global campaign to test and promote COVID-19 vaccines, China aims to win friends and cut deals

By **Jon Cohen**

The first people in the world to receive a COVID-19 vaccine were not part of a clinical trial. No TV stations or newspapers covered the historic event. No company issued a statement.

On 29 February, less than 2 months after the world awakened to the threat of the new disease, virologist Chen Wei, a major general in China's army, and six military scientists on her team stood in front of a Chinese Communist Party flag and received injections of an experimental COVID-19 vaccine. Chen, a national hero for her work on Ebola vaccines, had come to the initial center of the pandemic, Wuhan, with her group from the Academy of Medical Military Sciences, in part to help make the candidate vaccine with pharmaceutical company CanSino Biologics. Commentators inside and outside of China later questioned whether the event, which received wide play on social media, was real. No less than *People's Daily*, the Communist Party's main newspaper, labeled a photo of Chen receiving the vaccine as "#FAKE-NEWS." But Hou Li-Hua, a researcher at the academy who works on the vaccine project, says it was "true news"—an attempt to protect the scientists in the hard-hit city.

In the United States, the Trump administration's $10.8 billion Operation Warp Speed accelerated vaccine R&D faster than many researchers thought possible, specifically for the U.S. population. But an equally massive effort has unfolded in China. CanSino and two other Chinese companies—one owned

Science's COVID-19 reporting is supported by the Pulitzer Center and the Heising-Simons Foundation.

Published by AAAS

by the government, the other working closely with its regulatory agency—are investing substantial resources, testing four candidates in tens of thousands of volunteers around the world. They are likely only days or weeks away from announcing the outcomes of efficacy trials, just behind the encouraging early results recently announced by a brace of companies and institutions outside China.

But the low profile of those historic first injections, the military collaboration with a "private" company, and the ethically fraught decision to start with vaccinations outside of a clinical trial telegraphed that aside from the similar scale and speed, China's vaccine effort is following a very different course from those in the United States and Europe. Most leading Western vaccines rely on sexy technologies such as genetically engineered viral vectors, designer proteins, and snippets of RNA. Three of China's leading vaccine candidates use an unfashionable stalwart: the whole inactivated virus, an approach that dates back to the first successful flu vaccine in the 1930s.

And China's vaccine effort is cursed by the country's dramatic success with aggressive public health measures to stop the spread of the virus, SARS-CoV-2, including forced isolation of cases and testing of entire cities. Whereas the raging pandemic in the United States has enabled trials there to quickly deliver signals of efficacy, "China crushed the coronavirus epidemic early, so they lost the opportunity to test the efficacy of their vaccines there," says epidemiologist Ray Yip, who closely follows COVID-19 vaccine development as an adviser to Bill Gates. "If they had plenty of cases in China, they could have finished an efficacy trial ahead of other people."

So China's vaccine developers have gone abroad. Although the United States has shut them out of Operation Warp Speed, they have brokered deals with 14 other countries on five continents. They have mounted massive trials in the Arab world—and given candidate vaccines to top government officials there—and navigated toxic politics in Brazil, where the pandemic is raging fiercely, to test a vaccine and explore producing it there.

But China isn't just seeking promising venues for clinical trials. Not urgently needing the vaccines at home to fight a virus it has largely quashed, it is playing a global game by pledging to send any proven vaccine to countries that are conducting trials for its candidates, or to share the technologies behind them. "They know they don't need a vaccine to contain the epidemic in China," Yip says. "They can take their sweet time."

Yanzhong Huang, a global health specialist at both Seton Hall University and the Council on Foreign Relations, says the country is "actually using the vaccine to promote the diplomacy of foreign policy objectives." This "vaccine diplomacy" he says, contrasts starkly with Warp Speed's "vaccine nationalism" and aims to "fill in the void left by the United States."

"It is a very carefully executed and carefully thought out strategy," says Stephen Morrison, who directs the Global Health Policy Center at the Center for Strategic & International Studies. "A strategic goal of the Chinese government is to achieve hegemonic influence in the bioeconomy within the next decade."

At home, too, attitudes toward vaccines contrast with those in the United States and Europe, where mistrust is high, Morrison says. To the consternation of vaccine experts overseas, hundreds of thousands of people in China have already lined up to receive the experimental vaccines—even before their value and safety have been proved. "There has not been a collapse of faith and trust in science and in the state," Morrison says. "There's less fear about where this is all going."

THE SPEED AT WHICH Chen and her colleagues were able to get those first shots is all the more remarkable given that CanSino was arguably slow off the mark.

Although some COVID-19 vaccinemakers launched their projects the day after the sequence for SARS-CoV-2 became public on 10 January, CanSino CEO Yu Xuefeng had reservations. "We started to look into it in the middle of January, but there was a hesitation," he says. COVID-19, Yu worried, might be a blip, like severe acute respiratory syndrome (SARS), another coronavirus-caused disease, which alarmed the world in 2003. Companies and governments poured resources into developing vaccines, but the disease disappeared a year later.

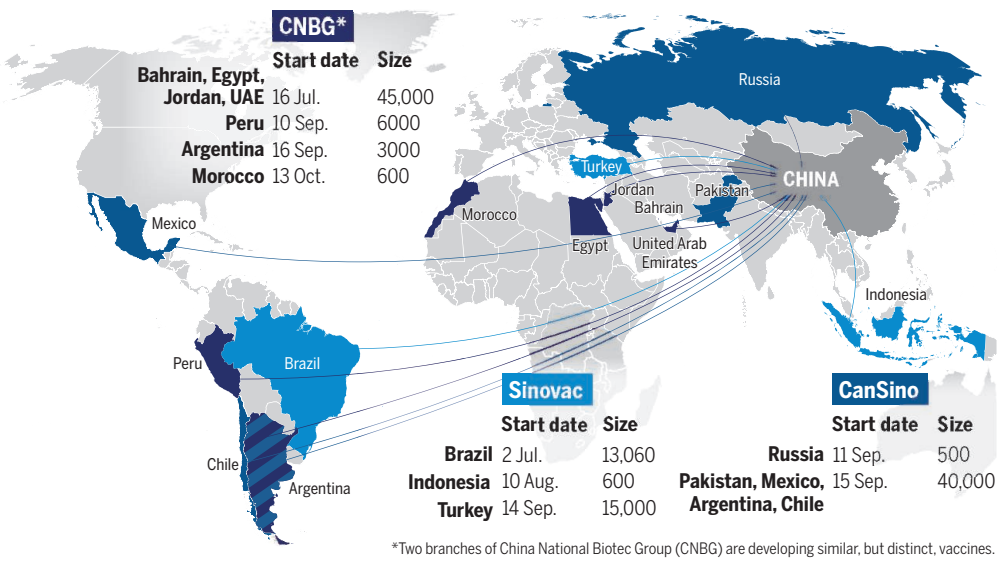
Originally from China, Yu completed his Ph.D. in microbiology at McGill University in Canada in 1997, and then stayed, working on vaccines for nearly 9 years at a Sanofi Pasteur branch there. He co-founded CanSino—a portmanteau of Canada and China—in 2009. A team led by Major General Chen back in China helped develop its only previous product to receive approval: an Ebola vaccine based on a widespread and largely harmless virus known as adenovirus 5 (Ad5), into which they stitched a gene for the surface protein of the Ebola virus.

Yu and his team considered making a COVID-19 vaccine with messenger RNA (mRNA) for the new coronavirus' surface protein, called spike—the innovative approach taken by Pfizer and its partner BioNTech, the "winner" of the race to report preliminary efficacy data. But CanSino decided to go with what it knew, using the Ad5 vector to carry the spike gene. "I thought that's the quickest and the mature way of developing a new vaccine," Yu says.

In just 1 month, CanSino's candidate was ready to be given to Chen and her team, and on 16 March the company launched the world's first COVID-19 vaccine trial, in Wuhan, to test its safety and ability to provoke immune responses. CanSino had beaten Moderna, which has an mRNA candidate, by 8 hours—though a world transfixed by

Vaccine road trip

With few COVID-19 cases at home, Chinese vaccinemakers have had to test the worth of their candidates abroad. Four are in efficacy trials in 14 countries.



*Two branches of China National Biotec Group (CNBG) are developing similar, but distinct, vaccines.

CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) WORLD HEALTH ORGANIZATION; CLINICALTRIALS.GOV; CHINESE CLINICAL TRIAL REGISTRY; WCG CENTERWATCH

Published by AAAS

The Chinese company CanSino Biologics had the first COVID-19 vaccine to enter clinical trials, and by April had already progressed to a phase II study in Wuhan.

the vaccine race among Western companies paid little attention.

Several U.S. and European contenders, including AstraZeneca, have also adopted adenoviruses to carry the spike protein, some opting for an Ad5 vector similar to CanSino's, despite several concerns about the approach. In 2007, two disastrous efficacy trials of an Ad5-based AIDS vaccine found that—for still-debated reasons—it actually raised the risk of HIV infection. The other worry is that preexisting immunity to Ad5 can attack the vector, which could explain why, in early trials, the CanSino vaccine elicited a weaker-than-expected antibody response. "We do see there's some impact," Yu concedes, "but it's not black and white." (The AstraZeneca vaccine's preliminary efficacy data suggest immunity against its adenovirus vector may have compromised that candidate's performance, as well.)

The two other Chinese players, Sinovac Biotech and China National Biotec Group (CNBG)—a subsidiary of one of the world's largest vaccinemakers, the state-owned Sinopharm—are taking a different approach: vaccinating people with the whole, "killed" virus. This requires no sophisticated protein or RNA design or genetic engineering: Scientists simply inactivate the virus with a chemical (beta propiolactone) and mix it with an adjuvant (alum) that effectively puts the immune system on full alert by irritating it. In theory, such vaccines can produce broader antibody and T cell responses, because they contain the full set of viral proteins, rather than a single one such as spike. And unlike mRNA vaccines, which have to be stored at subzero temperatures, inactivated viruses require no more than ordinary refrigeration.

But many scientists view inactivated virus vaccines as outmoded, difficult to make in high volume, and potentially dangerous. Warp Speed outright rejected the approach. "I really don't think the inactivated vaccine is a good idea," says Moncef Slaoui, scientific head of Warp Speed.

A major worry is that inactivated SARS-CoV-2 vaccines might trigger more severe illness, known as "enhanced respiratory disease," in immunized people who do get infected. Basically, if a vaccine triggers ineffective antibodies, they can form immune complexes that clog the lungs. This occurred with a vaccine against respiratory syncytial virus given to children in the 1960s, and in animal experiments with vaccines against SARS and another coronavirus disease, Middle East respiratory syndrome. The prospect of growing large batches of virus before killing it also poses challenges; twice in the past 5 years, live poliovirus has escaped from European plants making inactivated virus vaccines for that disease.

But inactivated virus vaccines, unlike mRNA and other technologies handsomely supported by Warp Speed, have a solid track record. "There are lots of different ways that vaccines are made, and it's great that innovation is occurring alongside tried-and-true approaches," says Nicole Lurie, a strategic adviser to the Coalition for Epidemic Preparedness Innovations (CEPI) who formerly served as U.S. assistant secretary for preparedness and response. "Inactivated vaccines are one of several tried and true approaches." Meng Weining, a senior director at Sinovac, says the company compared the inactivated approach—which it already uses to make six vaccines—with two other strategies in animal models. "The inactivated whole virus vaccine gave a much, much better result," Meng says.

Although it is easier, in theory, to produce mRNA in vast quantities than it is to grow the virus on a similar scale, vaccine experts say producing the inactivated virus vac-

Mohammed bin Rashid Al Maktoum (left), prime minister of the United Arab Emirates, on 3 November received a COVID-19 vaccine from CanSino Biologics.

PHOTOS: (TOP TO BOTTOM) TPG/GETTY IMAGES; WAM/HANDOUT VIA REUTERS

Published by AAAS

In Brazil, Sinovac's vaccine is caught in a battle between President Jair Bolsonaro and São Paulo's governor, João Doria. The mask says, "No vaccine! Oust Doria."

cines is unlikely to be a hurdle. CNBG, for example, has "enormous resources: 10,000 employees and scientists, huge manufacturing capability," says Nicholas Jackson, who heads CEPI's China office and previously worked on vaccine R&D at Pfizer. "They are a very competent beast." And, crucially for China's vaccine diplomacy, many other countries have manufacturers that have produced inactivated virus vaccines for decades.

If China's COVID-19 vaccines work, manufacturers say they could turn out 1.5 billion doses in total next year. And countries that cannot access vaccines bankrolled by Warp Speed—especially countries that hosted China's efficacy trials—might have a more secure vaccine supply.

THE PRIME MINISTER of the United Arab Emirates (UAE), Sheikh Mohammed bin Rashid Al Maktoum, on 3 November tweeted a photo of himself in Dubai, the right sleeve of his kandura rolled high, being injected with a CNBG COVID-19 vaccine. "We wish everyone safety and great health, and we are proud of our teams who have worked relentlessly to make the vaccine available in the UAE," Al Maktoum wrote. Two of the country's top ministers had received the vaccine 3 weeks earlier.

UAE has become the cornerstone of CNBG's efficacy trials and is following China's controversial lead in allowing people to receive the vaccine outside of clinical trials.

In a video conference on 23 June that linked UAE and China, health officials, ambassadors, and CNBG executives sat at long tables in rooms decorated with each country's flags and celebrated their decision to stage an efficacy trial together. The trial has since expanded to Bahrain, Egypt, and Jordan and hopes to recruit 45,000 people.

CNBG says it came to UAE to test its two whole virus vaccines—similar inactivated preparations made by two independent, and even competitive, laboratories—because the high SARS-CoV-2 infection rate there should speed an efficacy signal. But diplomacy and commerce also drove the decision. UAE's enormous foreign workforce means trial participants come from 125 different countries. "If you can prove these vaccines work in UAE," Huang says, "that means everybody in the world would think that the vaccine would work in their countries, too."

China may be hoping for a public relations (PR) benefit as well: UAE and many of the other collaborating countries have large Muslim populations, which Huang says could help mitigate human rights complaints about China's treatment of Uyghur Muslims in Xinjiang province. "They certainly don't want to have more enemies overseas," he says.

Huang adds that through its array of overseas trials, China hopes to build goodwill for its Belt and Road Initiative (BRI), a massive investment in infrastructure in more than 100 countries to increase trade. Critics have charged that the BRI is "debt-trap" diplomacy that's a form of neocolonialism. "China wants to work with these countries and prioritize them to have the vaccine because I think they believe this is going to facilitate the implementation of the BRI," he says.

China's vaccine diplomacy has not always gone smoothly. On 9 November, after Brazil suspended a trial of Sinovac's vaccine following the death of a participant, President Jair Bolsonaro took to Facebook. "*Morte, invalidez, anomalia*," he wrote, quoting from a Brazilian health agency that had listed possible reasons for the suspension: death, disability, genetic anomalies. Bolsonaro's message was clear: This Chinese vaccine, CoronaVac, was dangerous.

"Many people were really taken aback because of that post," says Esper Kallas, who heads the vaccine trial site at the University of São Paulo that the participant had joined. "He was celebrating the failure of a vaccine." For Bolsonaro, it was a PR victory over his arch political rival, the governor of São Paulo, who backed the CoronaVac trial. The president was also delighting in an apparent setback for China, which Bolsonaro, like his ally, U.S. President Donald Trump, has criticized relentlessly.

It turned out the participant died from a drug overdose. His death had nothing to do with CoronaVac, and the efficacy trial quickly resumed.

China chose to navigate Brazil's daunting politics because with an out-of-control pandemic—it is third in the world in total infections, with more than 100,000 new cases every week—the country is a magnet for vaccine testing and is desperate for vaccines. São Paulo state in September committed $90 million to Sinovac for 46 million doses. (This, notably, is 10 times cheaper than what the U.S. government is paying for the Pfizer-BioNTech and Moderna mRNA vaccines, which appear to be powerfully effective.) And Brazil could augment the supply by making vaccine itself: Sinovac says it may transfer its technology to the Butantan Institute, a major vaccine manufacturer in São Paulo, a collaboration

PHOTO: REUTERS/AMANDA PEROBELLI

Published by AAAS

that Meng describes as a "win-win."

China has had warmer receptions in other countries. Turkey in September launched a 13,000-person efficacy trial of Sinovac's vaccine. Serhat Ünal, who heads the Hacettepe University Vaccine Institute—which is similar to Butantan—and is on the scientific board of the Ministry of Health, says Turkey has "a good infrastructure for the phase III studies" and, unlike the United States and much of Europe, welcomed a Chinese vaccinemaker.

The three Chinese manufacturers also have large efficacy trials planned or underway in Indonesia, Pakistan, Saudi Arabia, Mexico, Argentina, and Chile (see map, p. 1264). It's a good strategy, Ünal says. "When you do the phase III in different countries, it's more transparent, it's more trustworthy," he says.

As much as vaccine diplomacy and the "soft power game" influence where the Chinese vaccinemakers cut deals for efficacy trials, they are also driven by capitalism, says Yip, who for 4 years headed the China office of the U.S. Centers for Disease Control and Prevention. "Everybody's clamoring for some COVID vaccine," he says. "They all want to tell their people we have secured some vaccine for you." And Chinese companies aim to profit by supplying it.

IT'S A SAFE BET that one or more of China's overseas trials will announce efficacy data any day. The results so far for other vaccines have fed a growing sense that many of the candidates will wallop what is, from a vaccine's point of view, a somewhat wimpy virus. But China is not waiting for the phase III results before widely using its vaccines at home. Its regulators appear to be satisfied with animal studies, combined with the minimal safety and immune response data from phase I and II trials. In June, CanSino received an emergency use authorization to vaccinate the military, and since then both Sinovac and CNBG have received green lights to vaccinate large populations outside of clinical trials.

With the pandemic vanquished at home, China is vaccinating its people as insurance—often, against a dangerously infected world. CanSino's Yu says "thousands" of troops on peacekeeping missions have received his company's vaccine before traveling to places that have a high burden of COVID-19. CNBG says "hundreds of thousands" of people in China have received its vaccines. "By doing this, we are able to build an immune barrier among specific groups of people like healthcare workers, pandemic prevention personnel, and border inspection personnel," the company explained in its written replies to *Science*. Vaccination is "*completely voluntary with informed consent*," CNBG stresses. What's more, "We did not receive a single case report of severe adverse reaction, and no infections reported for vaccinees working in high risk areas."

Sinovac's Meng says "more than 90%" of the company's employees have received its vaccine because they are considered a high-risk group; he received it because he travels overseas. (According to China's Ministry of Culture and Tourism, 155 million Chinese people traveled overseas in 2019, and 145 million tourists visited the country.) In October, the company began to sell its vaccine—$60 for two doses—in Yiwu, a city in Zhejiang province.

Yip says the government was even considering vaccinating all of Beijing after a COVID-19 outbreak there in June. Yip says officials "already had written the guidelines"; if more than 500 cases had surfaced, "they would shoot up everybody in Beijing with the vaccine." In the end, contact tracing, testing, and isolation of infected people limited the outbreak to 335 cases.

> **"They know they don't need a vaccine to contain the epidemic. ... They can take their sweet time."**
>
> **Ray Yip**, vaccine expert

Morrison says the Chinese government has clearly "decided at the highest levels" that it's worth the gamble to create "facts on the ground" and gain a global marketing advantage by having the first COVID-19 vaccines in wide use. "It's high risk and it's potentially high gain," he says.

But what if harm does occur? "You shouldn't apply the rules of peacetime during the war. Our lives are turned upside down," says Yip, who lives part-time in Beijing.

IF ITS VACCINE GAMBLE succeeds, China's image will gain a boost both at home and abroad. "They have reputational problems, internal and external," Morrison says.

In May, China's President Xi Jinping told the World Health Assembly, which governs the World Health Organization (WHO), that the country would make its COVID-19 vaccines "a global public good," a somewhat vague declaration that had many China watchers scratching their heads. But then China followed up on this commitment in October by joining the COVID-19 Vaccines Global Access (COVAX) Facility, an effort led by WHO, CEPI, and Gavi, the Vaccine Alliance, to make sure that any products proved safe and effective quickly reach rich and poor countries alike.

Although joining COVAX arguably gives China an insurance policy to obtain vaccines if its own candidates fail, Morrison says it's primarily a diplomatic move. COVAX hasn't received support from the United States or Russia, and China sees that it "could have a controlling influence over a major international mechanism." In addition, says Alexandra Phelan, a lawyer at Georgetown University's Center for Global Health Science and Security who specializes in China, "It is a good act of a global citizen to support this effort."

If a Chinese-made vaccine proves safe and effective, it could help people forget that the pandemic started there and how badly the government responded at first, Morrison says. And at home, it could brighten the image of China's vaccinemakers. Chinese citizens have reeled from a series of scandals over the past decade that include ineffective diphtheria, pertussis, and tetanus vaccines; improper records for a rabies vaccine; and sales of an expired polio vaccine.

In something of a twist, Yip says China's middle class may prefer a vaccine from a reputable foreign company. "Their level of confidence in Chinese-made vaccines is quite low because of all the repeated scandals," he says. AstraZeneca and Pfizer have agreements to produce their products with Chinese manufacturers. "They will outsell the CNBG, Sinovac, and CanSino [vaccines] by 10 to one—and they will charge 10 times more," Yip predicts.

A successful Chinese-made COVID-19 vaccine that has been scrutinized by outside regulators would reassure the domestic market, Phelan says. "There's a lot of domestic ground to make up."

In Brazil, Kallas says a similar dilemma could play out if Butantan, as hoped, starts to make Sinovac's CoronaVac. "There is a saying here that the neighbor's chicken is always the most tasty," Kallas says. "We have this perception that everything we do is not as good as the imported thing."

But for now, Brazil is embracing the Chinese vaccine. With cases surging, the arrival of a mere 120,000 doses of CoronaVac on 19 November became a big news story. The bias against China is little more than a far-right political "contamination," Kallas says, and most Brazilians see CoronaVac as "a viable option."

"I'd take it, no questions asked—this is a no-brainer," he adds. "The Pfizer and Moderna news was taken as a relief, but the problem is that both these vaccines are not in Brazil's grasp."

In Brazil, as in much of the world, China's warp speed vaccines may still, in the most meaningful way, come in first. ■

Published by AAAS

PERSPECTIVES

GENOMICS

Diversity hotspots: Coldspots of speciation?

Phylogeny of suboscine passerine birds finds lower speciation rates in biodiverse regions

By **Hélène Morlon**

From the tropical Andes to the Mediterranean Basin, some regions of Earth host an extraordinary number of species compared to less diverse regions. Robust species-level phylogenies (which map evolutionary relatedness) obtained from genetic sequencing are key for understanding the ecological and evolutionary processes at the origin of such species richness gradients. Yet, they are still incomplete for species-rich groups, even among the most studied organisms such as birds. On page 1343 of this issue, Harvey *et al.* (*1*) assemble an impressive phylogenomic dataset for the largest Neotropical bird group, the suboscine passerine radiation. Their analysis of these data suggests that species origination is slower, rather than faster, in hotspots of suboscine diversity and that time, rather than speciation rates, explains suboscine diversity gradients.

Species-level phylogenies have been instrumental in understanding the origin of biodiversity hotspots and richness gradients. Coupled with geographic data and models of speciation and extinction, comprehensive phylogenies provide useful information about the pace of diversification (the balance of speciation and extinction) in space and time (*1*–*7*). Yet, large-scale phylogenetic trees are often constructed with few genetic markers, or even by including species without genetic data. Increasing the robustness of these trees is important to obtain robust diversification rate estimates. As Harvey *et al.* show, their analyses on a less robust suboscine tree (*2*) would not have detected the negative association between present-day speciation rates and species richness.

The relationship between existing species richness and speciation rates has often been discussed in the context of understanding why biodiversity peaks in the tropics. Different hypotheses for this latitudinal gradient in diversity lead to different predictions regarding diversification rates (*8*). According to the time hypothesis, the tropics are older and therefore have had more time to accumulate species, even in the absence of any difference between tropical and temperate diversification rates. Whereas the diversification rate hypothesis posits that the

IBENS, CNRS UMR 8197, PSL Research University46 rue d'Ulm, 75005 Paris, France. Email: helene.morlon@bio.ens.psl.eu

PHOTO: MURRAY COOPER/MINDEN PICTURES

Published by AAAS

The giant antpitta (*Grallaria gigantea*) in Ecuador is one species of thousands of suboscine passerines.

tropics accumulate species faster, either because they are "a cradle" of biodiversity with high speciation rates, or because they are "a museum" of biodiversity with low extinction rates. Until recently, phylogenetic and paleontological data generally supported the diversification rate hypothesis, with evidence for higher speciation rates in the tropics where diversity peaks (*8*).

The negative association between speciation rates and species richness found by Harvey *et al.* contravenes this evidence, echoing recent studies in birds at different elevations (*5*) and latitudinal studies in fish (*6*) and flowering plants (*7*). Harvey *et al.* suggest that species-rich areas are coldspots rather than hotspots of speciation and that the diversity of these areas is explained by their old age and/or low extinction rates. If this interpretation is correct and the results shared with other species groups, focus may need to change from understanding why speciation rates are higher in biodiverse regions such as the tropics to understanding why they are lower. But, should conclusions be drawn so hastily?

Recent studies have focused on speciation rates "at the tips" of phylogenies, i.e. present-day speciation rates (*9*). Earlier studies based on sister species comparisons and therefore also focused on the recent past had similarly found lower speciation rates where biodiversity peaks (*10*). By contrast, studies integrating information from the distant past have often found higher speciation rates in diversity hotspots (*3*, *4*, *8*). Therefore, one explanation for the apparent paradigm shift is that current hotspots of biodiversity were once hotspots of speciation, but speciation rates decreased drastically in these areas to reach present-day levels below those of species-poor areas. Such a decline in speciation rate is expected if species richness has an upper limit set by environmental conditions such as resource availability (*11*). Under this "ecological limits" hypothesis, regions that accumulated species rapidly in the past are closer to their biodiversity capacity than species-poor regions, and are therefore less prone to present-day speciation.

Harvey *et al.* found mixed support for this hypothesis. The best model to fit their phylogenetic tree is one that contains a causal link between climatic variables and species richness, and an inverse correlation between species richness and speciation rates. This suggests that the number of species in a given region modulates speciation rates. Conversely, Harvey *et al.* did not find evidence for the speciation slowdown expected under the ecological limits hypothesis. Given that the existence of ecological limits is highly debated (*12*), additional evidence is needed, or other plausible explanations should be investigated.

The analysis by Harvey *et al.* of recent speciation rates, along with that of others (*5*–*7*), forces reconsideration of traditional evolutionary hypotheses for explaining richness gradients. However, recent speciation rates are only one piece of the puzzle. A thorough understanding of species richness gradients requires mapping speciation and extinction rates to geography not only in the present, but also in the past, while allowing these rates to vary through time. This is a challenging task, but not one that cannot be tackled. For example, geographically dependent diversification models have allowed estimation of temporal variations in tropical and temperate speciation rates (*3*). An alternative approach for mapping past rates to geography is to couple recent approaches that allow estimation of lineage-specific diversification rates (*13*) to ancestral biogeographic reconstructions. Given that speciation rates estimated from only extant data have diminished reliability for historical rates, and that extinction rates, which may play an important role in modulating richness gradients, are notoriously difficult to estimate from reconstructed phylogenies (*14*), further efforts to incorporate direct information from the past (i.e., fossils) will be particularly useful (*15*).

The assembly and analysis of comprehensive phylogenomic data for the largest tropical bird radiation by Harvey *et al.* contribute to mounting evidence that new species may not always be preferentially generated in biodiversity hotspots. These findings suggest a new burning question: Have diversity hotspots always been coldspots of speciation, or did they turn from hotspots to coldspots of speciation as diversity accumulated? ■

REFERENCES AND NOTES

1. M. G. Harvey *et al.*, *Science* **370**, 1343 (2020).
2. W. Jetz, G. H. Thomas, J. B. Joy, K. Hartmann, A. O. Mooers, *Nature* **491**, 444 (2012).
3. J. Rolland, F. L. Condamine, F. Jiguet, H. Morlon, *PLOS Biol.* **12**, e1001775 (2014).
4. C. R. Hutter, S. M. Lambert, J. J. Wiens, *Am. Nat.* **190**, 828 (2017).
5. I. Quintero, W. Jetz, *Nature* **555**, 246 (2018).
6. D. L. Rabosky *et al.*, *Nature* **559**, 392 (2018).
7. J. Igea, A. J. Tanentzap, *Ecol. Lett.* **23**, 692 (2020).
8. G. G. Mittelbach *et al.*, *Ecol. Lett.* **10**, 315 (2007).
9. D. Schluter, M. W. Pennell, *Nature* **546**, 48 (2017).
10. J. T. Weir, D. Schluter, *Science* **315**, 1574 (2007).
11. D. L. Rabosky, A. H. Hurlbert, *Am. Nat.* **185**, 572 (2015).
12. L. J. Harmon, S. Harrison, *Am. Nat.* **185**, 584 (2015).
13. O. Maliet, F. Hartig, H. Morlon, *Nat. Ecol. Evol.* **3**, 1086 (2019).
14. D. L. Rabosky, *Evolution* **64**, 1816 (2010).
15. T. A. Heath, J. P. Huelsenbeck, T. Stadler, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2957 (2014).

ACKNOWLEDGMENTS

I thank my research group for comments and the CNRS for financial support.

10.1126/science.abf0830

STRUCTURAL BIOLOGY

Using genetics to reveal protein structure

Measurements of genetic interactions in vivo can be used to determine protein structure

By **Dong Wang**[1,2,3]

Structure explains function. Knowing the three-dimensional (3D) shape and the arrangement of atoms in a biomolecule can help us understand how it works. For example, the DNA duplex structure provides insights into how genetic information is copied (*1*). The current tools for structural determination are primarily biophysical methods, such as x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryo–electron microscopy. Now, structural biologists are about to embrace a new tool for structural determination in cells. On page 1294 of this issue, Braberg *et al.* (*2*) describe a genetic approach for structural determination using phenotypic readouts from genetic perturbations. This approach opens new doors for the structure determination of many biological complexes in their native environment in cells.

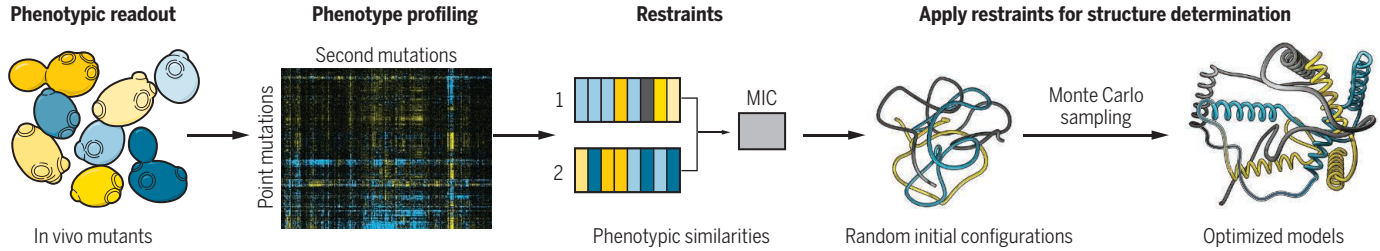
It has long been recognized that the inherent information that controls the folding of peptide chains into the specific 3D structures of proteins is stored in their linear amino acid sequence (*3*). However, how to decipher and use this for 3D structure determination remains a challenge. A breakthrough came when scientists compared the primary sequences of evolutionarily conserved protein homologs and found that the pairs of residues that are close in 3D space tend to have strong correlations in terms of sequence variations during evolution (covariation) (*4*). The information in these covariation patterns can be used for predicting protein folding with impressive success rates (*4*). Inspired by this success, scientists sought to test whether they

[1]Division of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego (UCSD), La Jolla, CA, USA. [2]Department of Cellular and Molecular Medicine, School of Medicine, UCSD, La Jolla, CA, USA. [3]Department of Chemistry and Biochemistry, UCSD, La Jolla, CA, USA. Email: dongwang@ucsd.edu

Published by AAAS

A genetic approach to determine protein structure

The genetic approach reported by Braberg *et al*. involves generation of a point mutant library from which in vivo phenotype profiles [such as point mutant epistatic miniarray profiles (pE-MAPs) or chemical genetics miniarray profiles (CGAPs)] are constructed. Phenotypic similarities are then quantified by assigning maximal information coefficient (MIC) values, and three-dimensional models are constructed using spatial restraints from pairwise genetic perturbations, structural refinement, and validation. Integrative structure determination can also include inputs from other methods.



could infer 3D structural information by analyzing the experimental covariation pattern of two mutations in the same protein.

Recently, several groups independently used deep mutational scanning (DMS), a technique that can generate and analyze functional consequences of a large set of point mutations in a protein, for structure determination (*5*–*7*). A dataset comprising the fitness effects of nearly all possible double mutants of a 56-amino-acid protein domain called IgG-binding domain of protein G (GB1) was analyzed and used to determine the structure of GB1 domains (*5*, *6*). The central idea of this genetic approach is that strong effects on protein function may occur when two genetic mutations affect amino acids that are proximal in 3D space. Thus, genetic interaction analysis can infer distance restraints that facilitate computing and refining protein structure models. In addition to GB1, structures of several other protein domains were also computed using the same approach (*5, 6*). DMS can also be used to determine biologically active conformations in cellular environments, even for highly dynamic multiconformational proteins such as α-synuclein (*7*). Moreover, an experimental method for protein structure determination, called 3D-seq, can compute 3D structures based on the analysis of sequence covariation patterns from experimental evolution of bacteria (*Escherichia coli*) that mimicked the natural evolution of two distinct antibiotic resistance proteins (*8*). However, it is not clear whether these approaches could be transferable to a much more complicated system, such as large protein complexes.

Braberg *et al*. describe an integrative structure determination approach for yeast (*Saccharomyces cerevisiae*) protein complexes. They constructed phenotypic profiles for point mutations of a target protein crossed against a library of genome-wide gene deletions or exposed to a collection of different environmental perturbations (see the figure). Phenotypic profiles for all mutations in the analysis were compiled, called a point mutant epistatic miniarray profile (pE-MAP). They defined the maximal information coefficient (MIC) as a quantitative measurement of similarities between pairs of phenotypic profiles in the pE-MAP. A pair of proximal point mutations is more likely to have a high MIC value than a pair of distal point mutations. Therefore, MIC values are informative about an upper distance bound between the residues. They then applied this distance restraint information for structural determination.

As a proof of principle, Braberg *et al*. determined the structure of the histone H3-H4 complex based on ~500,000 genetic interactions of 350 yeast mutants. The computed ensemble structural model is notably accurate and precise in comparison with the x-ray structure. They further estimated that 35 to 40 mutations per component are necessary to generate a reasonable model of a protein complex. Furthermore, the same pE-MAP could also be used to pair specific histone residues to their respective modifying enzymes, which are only transiently associated with the histones in vivo.

To demonstrate the generality of this approach, Braberg *et al*. also determined the structure of two large subunits of RNA polymerase II (Pol II) using a pE-MAP comprising ~60,000 genetic interactions (*9*). They found that the accuracy of the pE-MAP–derived structural model is comparable to those based on chemical cross-links or coevolution analysis. In addition, they also obtained a reasonable structural model of two large subunits of bacterial RNA polymerase (RNAP) based on a chemical genetics miniarray profile (CG-MAP) of 44 RNAP point mutations exposed to 83 different environmental stresses. Thus, this approach can be transferable to other types of phenotypic profiles. Integrative structure determination uses different types of data to improve the accuracy and precision of the model (*10*). Indeed, the authors showed that model accuracy and precision can be further improved if they use restraints from both genetic interactions and cross-links.

The genetic approach has several potential advantages over biophysical approaches. The structural data obtained from genetic mutations reflects the native state of proteins in vivo and does not require purification of samples. Therefore, it may help to solve the structures of complexes that are difficult to isolate and purify or those that are only transiently stable. However, there are some remaining challenges. The relationship between phenotypic pE-MAP measurements and structure can be very complicated. For example, some mutations in distant positions that are part of an allosteric network could give rise to similar profiles. Some mutations may perturb gene expression rather than protein function or cause major structural changes. Better analysis methods are needed to filter out these "abnormal mutants" that could compromise the accuracy of structure determination.

With recent advances in genome-editing, similar genetic approaches could be used for structure determination in human cells. Combinations of genetic approaches with other in situ approaches, such as superresolution imaging or cryo–electron tomography (*11*, *12*), and with deep-learning artificial intelligence approaches (*13*), could further improve the resolution of structures of important biological complexes in their native environment in cells. ■

REFERENCES AND NOTES

1. J. D. Watson, F. H. C. Crick, *Nature* **171**, 964 (1953).
2. H. Braberg *et al*., *Science* **370**, eaaz4910 (2020).
3. C. B. Anfinsen, *Science* **181**, 223 (1973).
4. D. S. Marks *et al*., *Nat. Biotechnol*. **30**, 1072 (2012).
5. N. J. Rollins *et al*., *Nat. Genet*. **51**, 1170 (2019).
6. J. M. Schmiedel, B. Lehner, *Nat. Genet*. **51**, 1177 (2019).
7. R. W. Newberry *et al*., *Nat. Chem. Biol*. **16**, 653 (2020).
8. M. A. Stiffler *et al*., *Cell Syst*. **10**, 15 (2020).
9. H. Braberg *et al*., *Cell* **154**, 775 (2013).
10. M. P. Rout, A. Sali, *Cell* **177**, 1384 (2019).
11. V. Marx, *Nat. Methods* **15**, 575 (2018).
12. L. Schermelleh *et al*., *Nat. Cell Biol*. **21**, 72 (2019).
13. R. F. Service *Science* 10.1126/science.abf9367 (2020).

ACKNOWLEDGMENTS

I thank J. Kadonaga, Y. Cheng, and N. Hao for helpful discussions.

10.1126/science.abf3863

GRAPHIC: A. KITTERMAN/*SCIENCE*

Published by AAAS

BIOPHYSICS

The glassiness of hardening protein droplets

Protein condensates can age to form glasses that increase in viscosity but retain elasticity

By **Huaiying Zhang**

In addition to dissolving in the watery cytosol, proteins can self-assemble into materials with different mechanical properties, such as solid filaments or soft gels, to facilitate various cellular functions. In the past decade, proteins have been shown to undergo liquid-liquid phase separation (LLPS) to form liquid droplets in which proteins are highly concentrated but are still dynamic and fluid (*1*, *2*). Many cellular compartments, such as nucleoli and stress granules, are phase-separated droplets. Although some droplets remain liquid for function, others harden into a less dynamic state over time (*3*, *4*). On page 1317 of this issue, Jawerth *et al.* (*5*) report that the hardening protein droplets are Maxwell glasses. These are Maxwell fluids that age like glasses in that viscosity increases with age, whereas the elasticity changes little over time (see the figure).

There are many distinctive features of LLPS that can be functionally relevant, one of which is the liquid property of the resulting droplets. Droplet fluidity is vital for ensuring proper chemical reactions occurring within the compartment. Disrupting the liquid properties of the nucleolus, for example, alters ribosomal RNA biogenesis (*6*). In addition, droplet fusion can be used for force generation to organize cellular space, such as clustering genomic loci (*7*, *8*) or bundling cytoskeleton filaments (*9*).

However, condensed phases formed with LLPS are not simple liquids but have diverse and changing material properties and are collectively called biomolecular condensates (*1*). The dependence of droplet fluidity on environmental factors (such as salt concentration) and condensate composition (such as RNA-to-protein ratio) has been revealed in vitro, with active rheology measured with optical tweezers and passive rheology measured with microbeads (*10*–*12*). Reconstituted droplets also undergo aging and maturation; that is, their material properties change with time.

Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
Email: huaiyinz@andrew.cmu.edu

Some condensates harden into a less dynamic state in which they do not fuse but stick together after collision, and the proteins rearrange less within them (*3*, *4*). Hardened condensates might be functional in that they can inhibit some chemical reactions or provide structural rigidity. The hardening process can also be vital for creating different material properties needed in a multistep cellular process. For example, in centrosome condensate formation with the protein spindle-defective protein 5 (SPD-5), the initial dynamic liquid may allow rapid protein incorporation early in the cell cycle, whereas the hardened condensate may provide the centrosome with the rigidity required to resist microtubule-pulling forces during mitosis (*4*).

In addition to hardening, some aged droplets even nucleate protein aggregates, including amyloids that are linked to various diseases (*10*, *13*, *14*). Thus, organizing biochemistry through LLPS may come at the cost of promoting pathological protein aggregation. However, the ability to nucleate solids is also exploited to seed cytoskeleton filaments, including actin and microtubules, from liquid condensates that concentrate monomers (*9*). Not only do condensate material properties change with environment, composition, and time, different material properties also coexist in subcompartments of a single condensate such as the nucleolus (*15*). These examples highlight the numerous ways that various condensate material properties can be combined to achieve complex functionality.

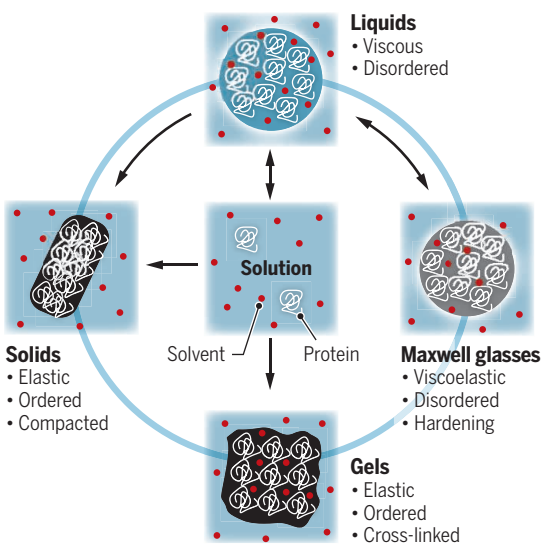
Jawerth *et al.* characterized the material properties of reconstituted protein condensates that harden over time by combining optical tweezer manipulation and microrheology. These hardening condensates behave like Maxwell fluids that have both viscous and elastic components present at all times. The elasticity of the Maxwell fluid changes little with age, indicating that hardening is not a gelation process in which molecules become cross-linked. However, the viscosity strongly increases with age, suggesting that the molecular dynamics are hindered by protein jamming within the liquid.

Because this material exhibits the behavior of Maxwell fluid but ages like glass, the authors call it Maxwell glass. Agreeing with the rheological results, no substantial structural changes within hardening condensates were observed with cryo–electron microscopy. Condensate-size shrinkage and increased protein density within the condensate were observed with fluorescent microscopy. The origin of protein jamming and how it links to protein chemistry await to be determined but will be needed to understand why some condensates harden over time and others do not.

Jawerth *et al.* followed the hardening of five different proteins [*Caenorhabditis elegans* protein PGL-3 (guanyl-specific ribonuclease pgl-3) and mammalian proteins FUS (RNA-binding protein FUS), EWSR1 (RNA-binding protein EWS), DAZAP1 (DAZ-associated protein 1), and TAF15 (TATA-binding protein–associated factor 2N)] with fluorescence recovery after photobleaching and studied the rheology of PGL-3 and FUS with optical tweezer and microrheology methods. Future work could expand the rheological studies to other condensates. For example, porous meshwork in hardened *Saccharomyces cerevisiae* and

Phases that proteins form

Proteins dissolved in solution (middle square) can self-assemble into liquids, solids, and gels. Jawerth *et al*. report that proteins can also harden into a new phase, a Maxwell glass. Protein liquids and Maxwell glasses are easier to reverse (double arrows) than protein gels and solids (single arrows).



GRAPHIC: C. BICKEL/*SCIENCE* FROM H. ZHANG

Published by AAAS

Schizosaccharomyces pombe Sup35 condensates was observed with cryo–electron microscopy, which suggests that they could be gels (*3*). Rheology studies on hardening Sup35 condensates would help to determine whether and how they age differently from condensates reported in this study. Nevertheless, it can no longer be assumed that all nondynamic condensates are gels because they may be Maxwell glasses.

Distinguishing glass-like and gel-like responses of hardened condensates is not only conceptually but functionally essential. Both gels and glasses can be structurally stable. Gel stiffness can be actively regulated by the degree of cross-linking and can be tailored to sustain different magnitudes of forces. A glass can act as a mechanical sensor, just like liquid droplets (*7*), because it can flow under stress. A jammed glass only allows small molecules to pass, whereas the larger pores of a gel permit diffusion of macromolecules such as proteins. However, glasses are more easily fluidized. A gel would be more desirable for a condensate where structure rigidity and chemical reactions are both needed, such as for centrosomes (*4*). A glass is suitable for slowing down all macromolecule movement through jamming and allowing small molecules to pass through and quickly fluidize the content when needed, such as for stress-sensing condensates (*3*).

It will be exciting to see which nondynamic condensates in cells are gels and which ones are glasses, and how cells exploit their material properties for functions. However, unlike the in vitro results reported by Jawerth *et al*., directly probing rheological properties of endogenous condensates remains technically challenging. Tools for forming de novo condensates in live cells in a controlled manner may be useful for engineering condensates suitable for optical tweezer manipulation or embedding microbeads to follow condensate rheology (*7*, *8*). ■

REFERENCES AND NOTES

1. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
2. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
3. T. M. Franzmann *et al*., *Science* **359**, eaao5654 (2018).
4. J. B. Woodruff *et al*., *Cell* **169**, 1066 (2017).
5. L. Jawerth *et al*., *Science* **370**, 1317 (2020).
6. L. Zhu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 17330 (2019).
7. Y. Shin *et al*., *Cell* **175**, 1481 (2018).
8. H. Zhang *et al*., *Mol. Biol. Cell* **31**, 2048 (2020).
9. T. Wiegand, A. A. Hyman, *Emerg. Top. Life Sci.* **0**, 1 (2020).
10. H. Zhang *et al*., *Mol. Cell* **60**, 220 (2015).
11. S. Elbaum-Garfinkle *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7189 (2015).
12. L. M. Jawerth *et al*., *Phys. Rev. Lett.* **121**, 258101 (2018).
13. Y. Lin, D. S. W. Protter, M. K. Rosen, R. Parker, *Mol. Cell* **60**, 208 (2015).
14. A. Patel *et al*., *Cell* **162**, 1066 (2015).
15. M. Feric *et al*., *Cell* **165**, 1686 (2016).

ACKNOWLEDGMENTS

This work was supported by the U.S. National Institutes of Health (1K22CA23763201).

10.1126/science.abe9745

CORONAVIRUS

Remembering seasonal coronaviruses

Antibodies against seasonal coronaviruses react with SARS-CoV-2

By **Jenna J. Guthmiller**[1] and **Patrick C. Wilson**[1,2]

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has differential effects according to age, with symptomatic and severe infections mostly occurring in older adults. One possible explanation for this variation is that children and younger adults have more preexisting immunity against seasonal human coronaviruses (HCoVs) that cross-react with SARS-CoV-2, providing protection from severe and even symptomatic SARS-CoV-2 infection. Consistently, SARS-CoV-2 cross-reactive memory $CD4^+$ and $CD8^+$ T cells against the spike protein, the major surface protein of coronaviruses, have been reported in unexposed individuals (*1*, *2*). Whether humoral immunity (antibodies and memory B cells) against SARS-CoV-2 cross-reacts with seasonal HCoVs is now emerging. On page 1339 of this issue, Ng *et al*. (*3*) and Shrock *et al*. (*4*) reveal that individuals exposed and unexposed to SARS-CoV-2 have cross-reactive serum antibodies against the spike protein of SARS-CoV-2 and seasonal HCoVs.

Four seasonal HCoV strains cause cold symptoms in humans: 229E, NL63, OC43, and HKU1. Despite using different host receptors for cellular entry, all HCoVs express the spike protein on their surface. The spike protein is composed of two subunits: S1 contains the receptor-binding domain (RBD), which is responsible for binding to host cell receptors, and S2 is critical for mediating viral and host cell membrane fusion and cell entry. The fusion peptide of the S2 subunit is highly conserved among seasonal HCoVs and zoonotic coronaviruses, including SARS-CoV-2 (*3*), whereas S1 is more variable.

Shrock *et al*. found that people unexposed to SARS-CoV-2 have cross-reactive antibody responses against an array of coronaviruses, including the recently emerged SARS-CoV-2. However, most of this preexisting immunity targets a few epitopes on the spike protein, nucleocapsid protein, and the nonstructural proteins that are encoded by open reading frame 1. Using a sensitive flow cytometric assay to detect cross-reactive serum antibodies, Ng *et al*. found that individuals unexposed to SARS-CoV-2 possessed neutralizing antibodies against the S2 protein. Cross-reactive antibodies were class-switched to the mature antibody isotypes immunoglobulin G (IgG) and IgA, suggesting that B cells producing these cross-reactive antibodies were induced by a previous immune response against infection with seasonal HCoVs.

Upon SARS-CoV-2 infection, individuals showed increased production of antibodies that cross-react with the spike proteins of SARS-CoV-2 and seasonal HCoVs, called back-boosting (see the figure). Back-boosting is a common phenomenon observed after influenza virus infection and vaccination that results in the recall of antibodies targeting conserved epitopes of past circulating influenza viruses (*5*). Back-boosting can be directed to nonprotective epitopes, such as those on the unexposed nucleocapsid protein, which is referred to as original antigenic sin. However, back-boosting can lead to the recall of broadly neutralizing antibodies, as observed with the 2009 pandemic H1N1 influenza virus (*6*). Shrock *et al*. and Ng *et al*. both identified that SARS-CoV-2 infection boosted antibody responses against several conserved epitopes, including the fusion peptide of the S2 subunit. Together, these studies solidify the existence of cross-reactive humoral immunity against HCoVs and SARS-CoV-2.

Many of the cross-reactive antibodies are specific for epitopes on the fusion peptide or nearby epitopes on the S2 subunit. These antibodies likely neutralize coronaviruses by blocking viral membrane fusion and host cell entry (*7*). Although Ng *et al*. showed that cross-reactive antibodies neutralized SARS-CoV-2 infection in vitro, it is unknown whether cross-reactive antibodies specific for the fusion peptide prevent SARS-CoV-2 infection or limit COVID-19 severity in vivo. Nonetheless, because the fusion peptide is highly conserved across coronaviruses (*8*), it is an attractive target for a universal coronavirus vaccine that could generate broadly neutralizing antibodies and thereby protect against seasonal HCoVs, SARS-CoV-2, and future zoonotic coronavirus spillovers.

[1]Section of Rheumatology, Department of Medicine, University of Chicago, Chicago, IL 60637, USA. [2]Committee on Immunology, University of Chicago, Chicago, IL 60637, USA. Email: wilsonp@uchicago.edu

Published by AAAS

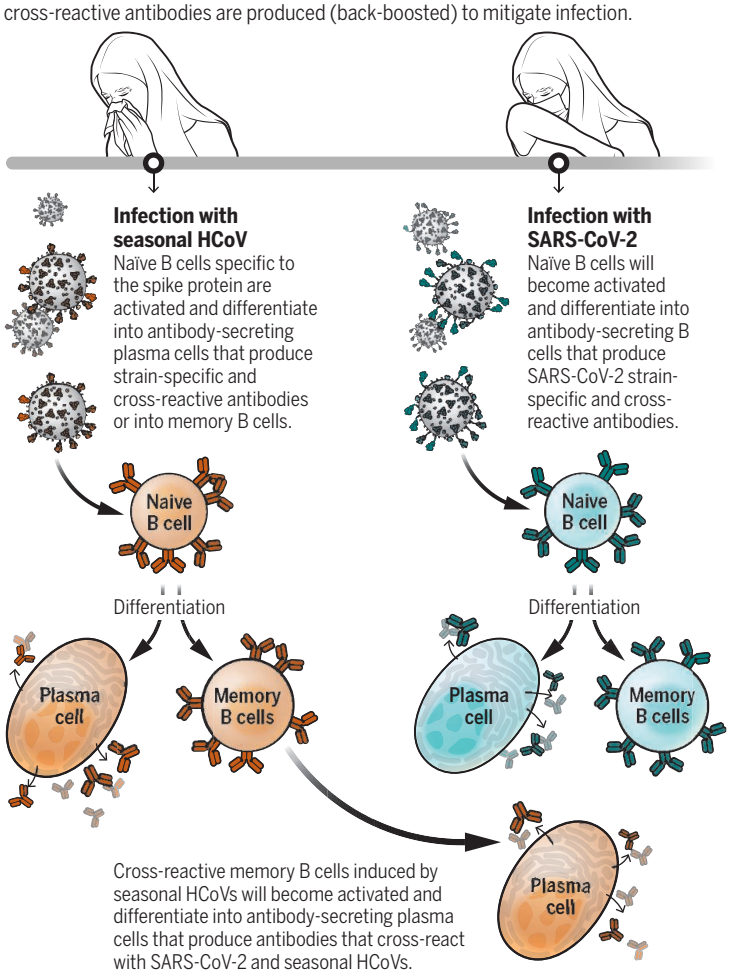
Most adults have been exposed to and have preexisting immunity against the seasonal HCoVs. Infection rates of seasonal HCoVs are higher in children and adolescents (*9*), potentially because children are in close contact in childcare settings and have little preexisting immunity against seasonal HCoVs. Ng *et al.* found that children and adolescents had the highest amounts of cross-reactive antibodies between seasonal HCoVs and SARS-CoV-2, suggesting that recurrent seasonal HCoV infections in young people may protect them from symptomatic and severe SARS-CoV-2 infections by limiting viral infection. Adults unexposed to SARS-CoV-2 have serum antibodies against seasonal HCoVs that do not cross-react with SARS-CoV-2, suggesting that recurrent HCoV infection may be necessary to induce and sustain protective cross-reactive antibody titers. A longitudinal study in Michigan found that HCoV-OC43 was the most common strain to infect children (*9*). Because the HCoV-OC43 and SARS-CoV-2 fusion peptides are highly conserved (*3*, *4*), antibodies produced from HCoV-OC43 infection may provide protection against SARS-CoV-2 infection in children. Furthermore, S2-specific memory B cells commonly cross-react between the spike protein of OC43 and SARS-CoV-2 (*10*). Thus, prior OC43 infection may train B cells toward conserved fusion peptide epitopes and subsequent exposure to OC43 or SARS-CoV-2 may boost cross-reactivity.

How preexisting humoral immunity against seasonal HCoVs affects the development of humoral immunity against SARS-CoV-2 remains unclear. Although humans have cross-reactive memory B cells that can be recalled by SARS-CoV-2 (*3*, *10*), cross-reactive antibodies and memory B cells against seasonal HCoVs may hinder the development and maturation of high-affinity B cells targeting conserved epitopes shared between seasonal HCoVs and SARS-CoV-2. Recent work on *Plasmodium falciparum*, the parasite that causes malaria, revealed that nonprotective antibody titers against the circumsporozoite protein limited the recruitment of high-affinity B cells against the same epitope (epitope masking) (*11*). In the context of coronaviruses, cross-reactive antibodies against the fusion peptide of seasonal HCoVs may mask this epitope and limit the recruitment and development of high-affinity B cells against the SARS-CoV-2 fusion peptide. Rather, the humoral immune response may be diverted toward SARS-CoV-2–specific epitopes, such as those of the RBD. Therefore, the generation of a universal coronavirus vaccine that boosts antibodies against the fusion peptide may be difficult to achieve. With the likely rollout of SARS-CoV-2 vaccines, it will be interesting to understand the relationship between cross-reactive preexisting antibodies before vaccination and the subsequent antibody specificity induced by vaccination, particularly in children and older people who experience different severity of disease.

Low antibody titers and low-affinity antibodies have the potential to enhance viral disease by facilitating viral uptake by host cells, a process called antibody-dependent enhancement (ADE). ADE is observed with atypical measles and serial infections with different dengue virus serotypes (*12*, *13*). Because preexisting antibody titers against SARS-CoV-2 are relatively low and potentially low affinity, concern remains that this could lead to ADE. However, there is no evidence for ADE in SARS-CoV-2 infection, particularly because thousands of acutely infected people have received convalescent plasma from recovered COVID-19 patients, with few adverse events (*14*). Additionally, ADE has been suggested as a potential mechanism leading to multisystem inflammatory syndrome in children (MIS-C), which is associated with SARS-CoV-2 infection. However, no distinct antibody response was predictive of MIS-C occurrence (*15*), although this study did not analyze the presence of cross-reactive antibodies against seasonal HCoVs. Therefore, antibodies that cross-react with seasonal HCoVs and SARS-CoV-2 are unlikely to worsen COVID-19 and instead are likely to provide some protection against SARS-CoV-2 infection and severe COVID-19. Together, the studies of Shrock *et al.* and Ng *et al.* highlight that further research is needed into how SARS-CoV-2 antibody responses are shaped by previous exposures to seasonal HCoVs and how this immunity can be harnessed to provide protection. ■

Humoral immunity against coronaviruses

Upon repeated exposure to seasonal human coronaviruses (HCoVs), strain-specific and cross-reactive antibodies are generated. Upon subsequent exposure to SARS-CoV-2, cross-reactive antibodies are produced (back-boosted) to mitigate infection.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. A. Grifoni *et al.*, *Cell* **181**, 1489 (2020).
2. J. Mateus *et al.*, *Science* **370**, 89 (2020).
3. K. W. Ng *et al.*, *Science* **370**, 1339 (2020).
4. E. Shrock *et al.*, *Science* **370**, eabd4250 (2020).
5. J. M. Fonville *et al.*, *Science* **346**, 996 (2014).
6. J. Wrammert *et al.*, *J. Exp. Med.* **208**, 181 (2011).
7. C. M. Poh *et al.*, *Nat. Commun.* **11**, 2806 (2020).
8. T. Tang *et al.*, *Antiviral Res.* **178**, 104792 (2020).
9. A. S. Monto *et al.*, *J. Infect. Dis.* **222**, 9 (2020).
10. P. Nguyen-Contant *et al.*, *mBio* **11**, e01991-20 (2020).
11. H. A. McNamara *et al.*, *Cell Host Microbe* **28**, 572 (2020).
12. F. P. Polack *et al.*, *Nat. Med.* **9**, 1209 (2003).
13. L. C. Katzelnick *et al.*, *Science* **358**, 929 (2017).
14. M. J. Joyner *et al.*, *Mayo Clin. Proc.* **95**, 1888 (2020).
15. S. P. Weisberg *et al.*, *Nat. Immunol.* 10.1038/s41590-020-00826-9 (2020).

ACKNOWLEDGMENTS

P.C.W. is funded by the National Institute of Allergy and Infectious Diseases (NIAID); National Institutes of Health (U19AI082724, U19AI109946, and U19AI057266). P.C.W. is also supported by the NIAID Collaborative Influenza Vaccine Innovation Centers (CIVIC; 75N93019C00051).

10.1126/science.abf4860

Published by AAAS

VIEWPOINT: COVID-19

Ensuring vaccine safety

Comprehensive safety testing is based on experience with prior vaccines

By David M. Knipe[1], Ofer Levy[2,3,4], Katherine A. Fitzgerald[5], Elke Mühlberger[6]

There is an urgent need for vaccines to protect against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection to reduce COVID-19 and stop the current pandemic. Although bureaucratic delays should be reduced to accelerate vaccine availability, there remains the need for extensive safety testing protocols developed by the U.S. Food and Drug Administration (FDA) and other regulatory agencies. COVID-19 vaccines will be safe if regulatory agencies maintain their well-documented safety testing protocols. Safety should be considered at every phase of vaccine discovery, development, and testing. History provides a strong scientific basis for safety evaluation of all vaccine candidates, which must be maintained to realize their enormous potential.

Vaccines are among the most successful medical and public health measures ever implemented (*1*). It is estimated that vaccines prevent ~6 million deaths globally per year (*2*). For example, vaccination eliminated variola virus (which causes smallpox) and nearly all wild poliovirus and has greatly reduced measles virus infection. Vaccines against hepatitis A and B viruses, rubella virus, mumps virus, influenza virus, human papilloma virus, varicella zoster virus (which causes chickenpox in children or shingles in adults), and yellow fever virus are broadly used and highly successful at reducing morbidity and mortality. The historical experience with vaccine development has paved the way for a well-developed path for preclinical and clinical testing of vaccines to ensure their safety and efficacy, leading to safe vaccines that have saved millions of lives.

Empirical experience, including evaluation of vaccine-associated adverse events, indicates the importance of thoroughly assessing safety of vaccine preparations before licensing and widespread use. Lessons learned include that all batches of vaccines must be tested for safety. For example, the need for systematic adherence to formulation standards and safety testing of each batch of formalin-inactivated Salk polio vaccine before deployment became apparent after a tragic incident in 1955. Although nearly all batches of the licensed vaccine were safe, two batches from Cutter Laboratories were contaminated with live poliovirus because of incomplete inactivation. This resulted in abortive polio (characterized by headache, stiff neck, fever, and muscle weakness) in 40,000 individuals, 51 cases of permanent paralysis, and five deaths and spread to family and community members (*3*). The Cutter incident led to strict new federal regulations, including current quality control measures to ensure strict adherence to vaccine formulation for inactivation of each vaccine batch. Inactivated polio vaccine preparations have been very safe since (*4*).

An additional level of vaccine safety is provided by understanding the mechanism of action and immune correlates of protection. For some vaccines, correlates of protection, such as the presence of specific neutralizing antibodies, are well established. A clear understanding of correlates of protection can ensure that vaccines induce the optimal immune response for protection while avoiding nonproductive or counterproductive immune responses or disease. Two examples illustrate this point. A formalin-inactivated respiratory syncytial virus (FI-RSV) vaccine developed in the 1960s stimulated moderate amounts of serum antibodies, measured by complement fixation, but failed to protect against RSV infection or disease (*5*). Moreover, most children immunized with this inactivated RSV vaccine and that subsequently became infected with RSV were hospitalized, with enhanced respiratory disease. The FI-RSV vaccine was later found not to induce neutralizing antibodies in vaccinees (*6*), and in a cotton rat infection model it induced a T helper cell 2 (T_H2)–biased CD4$^+$ T cell response (*7*), which can lead to lack of protective immunity and immune pathology. Lack of viral neutralization and altered T cell responses are believed to have contributed to vaccine-enhanced disease in these immunized individuals.

In another example, the first dengue vaccine, a tetravalent chimeric yellow fever–dengue virus vaccine, reduced incidence of severe dengue disease in older children and was licensed for use in children >9 years of age. However, although overall reduction of severe disease and hospitalization over a 5-year period was observed in the trials, protection was highest in individuals who had been exposed to dengue virus before immunization, whereas immunization increased severe disease in seronegative individuals (*8, 9*). Protection was lowest against one of the four dengue serotypes (DEN-2), suggesting that weak immune responses to DEN-2 or in general in seronegative individuals resulted in antibody-dependent enhancement (ADE) of disease. ADE is caused by non-neutralizing antibodies or antibodies at subneutralizing concentrations that promote infection by enhancing uptake of viral particles into host cells. ADE has been reported to increase infectivity of SARS-coronavirus in certain types of cells in culture (*10*). Given these experiences, current vaccine evaluation in preclinical and clinical studies (see the figure) includes scrutiny of vaccine immunogenicity in relation to correlates of protection to maximize efficacy and minimize potential detrimental effects.

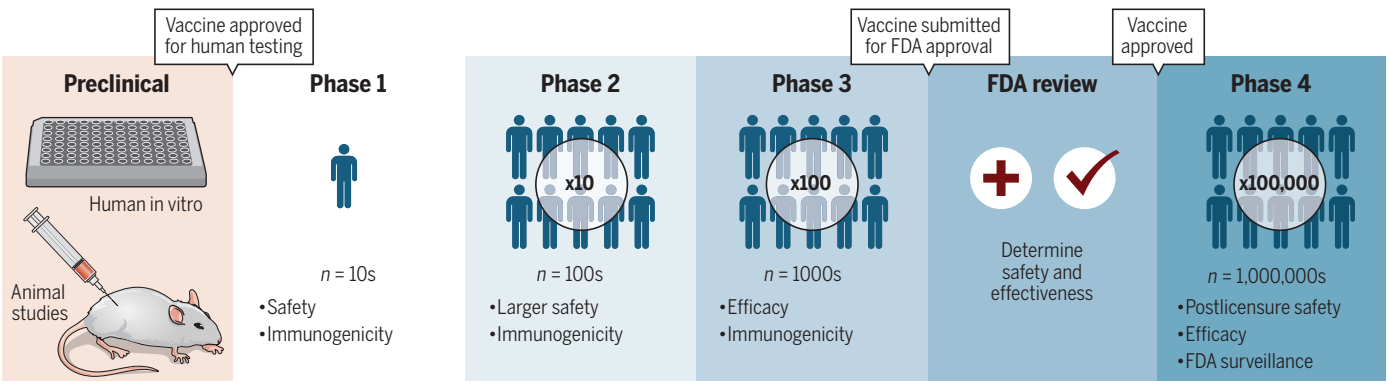
Another lesson learned is that if serious adverse events are detected in a clinical trial, then additional clinical testing is indicated. For example, the first rotavirus vaccine, a rhesus-human reassortant rotavirus tetravalent vaccine (RRV-TV) with genome segments of human and rhesus rotaviruses, was licensed in the United States in 1998. Rotaviruses cause severe and potentially fatal diarrhea in infants and children. Although the vaccine was effective in preventing gastroenteritis in infants, intussusception (a painful form of bowel obstruction due to bowel prolapse that can be fatal if left untreated) was reported in 5 of 10,054 vaccinees, compared with one case in 4633 placebo recipients, a difference that was not statistically significant. The perceived lack of serious adverse effects led to licensure of RRV-TV, with intussusception as a possible rare adverse reaction. However, after its approval and in the first year of vaccine use, 15 cases of intussusception were reported in vaccinees, in contrast to only 4 cases in the 7 years preceding vaccination, triggering suspension of the vaccine in 1999 (*11*). This example illustrates the importance of careful evaluation of any adverse reaction and postlicensure surveillance to ensure vaccine safety. Large prelicensure trials of two later rotavirus vaccines, a pentavalent human-bovine viral reassortant vaccine (RV5) and a monovalent single-strain human rotavirus vaccine (RV1) with compositions different from that of RRV-TV, demonstrated very low

[1]Department of Microbiology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA. [2]Precision Vaccines Program, Boston Children's Hospital, Boston, MA, USA. [3]Harvard Medical School, Boston, MA, USA. [4]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [5]Department of Medicine, Program in Innate Immunity, University of Massachusetts Medical School, Worcester, MA, USA. [6]Department of Microbiology and Immunology, National Emerging Infectious Diseases Laboratories, Boston University School of Medicine, Boston, MA, USA. Email: david_knipe@hms.harvard.edu

Published by AAAS

Vaccine safety evaluation

Safety is considered at every phase of vaccine discovery and development. Upon licensure, vaccines enter phase 4, whereby surveillance approaches by regulators, such as the U.S. Food and Drug Administration (FDA), monitor potential vaccine side effects.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

incidence of intussusception, and these vaccines are now widely approved for use.

In 1976, an outbreak of influenza that resulted in one death among Army recruits at Fort Dix, New Jersey, followed by spread through the base, was ascribed to influenza virus isolates similar to the 1918 pandemic H1N1 virus that killed millions of people. Fearful that a potentially pandemic influenza virus was emerging, an emergency vaccine initiative was approved by President Ford and funded by Congress. Meanwhile, scientific questions arose, including genetic evidence that the New Jersey isolates were similar to swine (H1N1) influenza viruses circulating in the United States at the time (*12*) and doubts about the virulence of the virus isolates, both suggesting that this was not an emerging pandemic virus. In addition, there was evidence that the vaccine being produced lacked an antigenic form of the viral neuraminidase protein, which could render it less efficacious. President Ford nevertheless implemented the vaccine initiative.

The influenza vaccine was tested in ~7000 individuals in the spring of 1976 and was deemed to be safe, with broad immunization starting on 1 October. About 25% of the U.S. population was immunized before ~450 cases of the paralytic Guillain-Barré syndrome (GBS) disease emerged, a statistically significant increase above the normal population incidence (*13*). This influenza immunization program was terminated in December 1976 with severe consequences, including diminished public confidence in vaccines and the public health care system. Later studies have shown minimal GBS associated with other influenza virus vaccines (*13*); thus, this particular vaccine formulation was likely problematic. Potential lack of neuraminidase activity, contamination with other microbes, and induction of autoimmune antibodies have all been suggested to account for these adverse effects.

As a result of the observations from the RRV-TV and 1976 swine flu vaccines, current regulatory practices require monitoring of rare adverse events pre- and postlicensure with detection of serious adverse events triggering a pause to trial or use. Such a safety pause enables study of adverse events to assess whether the vaccine trial or use can be resumed as deemed appropriate. Indeed, such safety pauses have occurred in current COVID-19 vaccine trials, underscoring the value of extensive regulatory safety protocols, which should not be rushed or undermined.

History has also taught us the importance of continued surveillance of potential vaccine-related adverse events and measurement of immunogenicity and outcomes even after licensure. Such surveillance—for example, through the U.S. Department of Health and Human Services Vaccine Adverse Event Reporting System (VAERS) and the FDA Biologics Effectiveness and Safety (BEST) electronic medical record–based platform—can confirm safety and increase understanding of the immune correlates of protection and mechanisms of immunogenicity to further enhance vaccine development.

Multiple COVID-19 vaccine trials are currently being conducted in parallel, with many additional candidates in preclinical development. A decision regarding one of these clinical trials that enables vaccine use before full regulatory approval—for example, through emergency use authorization (EUA)—would accelerate deployment at the potential risk of undermining the ability to recruit additional participants for that and other trials, thereby impairing collection of normally comprehensive data regarding safety, immunogenicity, efficacy, and durability of protection (*14*). To enhance public confidence in vaccines by providing transparency, the FDA Center for Biologics Evaluation and Research (CBER) has published master protocols for SARS-CoV-2 vaccine safety and effectiveness evaluation (*15*). Overall, the interests of vaccinees, vaccine developers, pharmaceutical companies, and regulatory agencies are aligned on the importance of ensuring vaccine safety.

There is an urgent need for COVID-19 vaccines and exciting progress to that end, but there remains a critical public health obligation to conduct rigorous evaluation to ensure safety as well as efficacy. Vaccines remain one of the most successful biomedical tools for prevention of disease. The urgent need for COVID-19 vaccines must be balanced with the imperative of ensuring safety and public confidence in vaccines by following the established clinical safety testing protocols throughout vaccine development, including both pre- and postdeployment. ■

REFERENCES AND NOTES

1. C. M. C. Rodrigues, S. A. Plotkin, *Front. Microbiol.* **11**, 1526 (2020).
2. J. Ehreth, *Vaccine* **21**, 4105 (2003).
3. N. Nathanson, A. D. Langmuir, *Am. J. Hyg.* **78**, 29 (1963).
4. P. A. Offit, *N. Engl. J. Med.* **352**, 1411 (2005).
5. H. W. Kim *et al.*, *Am. J. Epidemiol.* **89**, 422 (1969).
6. B. R. Murphy *et al.*, *J. Clin. Microbiol.* **24**, 197 (1986).
7. B. S. Graham *et al.*, *J. Immunol.* **151**, 2032 (1993).
8. S. Sridhar *et al.*, *N. Engl. J. Med.* **379**, 327 (2018).
9. S. B. Halstead, *Vaccine* **35**, 6355 (2017).
10. M. Jaume *et al.*, *J. Virol.* **85**, 10582 (2011).
11. Centers for Disease Control and Prevention (CDC), *MMWR Morb. Mortal. Wkly. Rep.* **48**, 577 (1999).
12. P. Palese, J. L. Schulman, *Nature* **263**, 528 (1976).
13. D. Evans, S. Cauchemez, F. G. Hayden, *J. Infect. Dis.* **200**, 321 (2009).
14. R. Gurahoy, E. P. Krenzelok, *Clin. Infect. Dis.* 10.1093/cid/ciaa1653 (2020).
15. U.S. Food and Drug Administration, Development and Licensure of Vaccines to Prevent COVID-19; www.fda.gov/media/139638/download.

ACKNOWLEDGMENTS

The authors are members of and funded by the Massachusetts Consortium on Pathogen Readiness. D.M.K. is a co-inventor, and O.L. is a named inventor on vaccine and vaccine adjuvant patents, respectively. D.M.K. and K.A.F. are supported by NIH grants. E.M. is supported by grants from NIH and Fast Grants. O.L. is supported by grants and contracts from NIH, the Bill & Melinda Gates Foundation, and the Mueller Health Foundation. We thank our colleagues for comments on the manuscript, P. T. Waters for manuscript preparation, and K. Johnson for help with the figure.

Published online 17 November 2020
10.1126/science.abf0357

Published by AAAS

RETROSPECTIVE

Angelika Amon (1967–2020)

Trailblazing cell cycle biologist

By **Jason M. Sheltzer**[1] *and* **Rosella Visintin**[2]

Angelika Amon, a pioneering geneticist and cell biologist, died on 29 October after a long battle with ovarian cancer. She was 53 years old. Angelika made important discoveries involving the control of chromosome segregation and the consequences of segregation errors. An innovative and outspoken scientist, she thrived on challenging dogmas and tackling unorthodox questions.

Born in Austria on 10 January 1967, Angelika first encountered the two loves of her life in high school. There, she met her future husband, Johannes Weis, who would become her constant companion for almost 40 years and the father of her beloved daughters, Theresa and Clara. In a science class, she also came across a grainy, black-and-white movie of chromosomes separating during mitosis. She would later identify that movie as the catalyst that spurred her to a lifetime of studying cell cycle genetics.

Angelika began her scientific career as a student in the laboratory of Kim Nasmyth at the Research Institute of Molecular Pathology in Vienna, where she made several fundamental discoveries illuminating how cyclins drive the yeast cell cycle. She showed that cyclin expression is carefully confined to discrete phases of the cell cycle by a combination of transcriptional and posttranslational regulation, which help to ensure proper progression through mitosis. After receiving her B.S. in 1989 and her Ph.D. in 1993, both from the University of Vienna, she joined Ruth Lehmann's lab at the Whitehead Institute in Cambridge, Massachusetts, for a brief postdoc on *Drosophila* embryogenesis. Angelika soon realized that flies grew too slowly for her liking, and she left the Lehmann lab to begin her independent career as a Whitehead Fellow in 1996.

In her own lab, Angelika decided to return to yeast as a model organism for cell cycle biology. She was fond of proclaiming, "When you work with yeast, the rate-limiting factor is your brain." She believed that the awesome power of yeast genetics gave scientists the ability to design clever experiments to uncover the secrets of the cell cycle and that these fundamental processes were likely to be conserved across the tree of life. One of us (R.V.) joined Angelika's lab as her first postdoc, and we embarked on an incredible scientific journey. Together, we discovered that exit from mitosis requires the phosphatase Cdc14, which is released from the nucleolus during anaphase. It was the first time that a phosphatase was shown to link mitosis and the subsequent G1 phase, and it established a new paradigm for how cell cycle–driven phosphorylation is regulated and reset after mitosis.

In 1999, Angelika became a faculty member at the Massachusetts Institute of Technology Center for Cancer Research (now the Koch Institute for Integrative Cancer Research), and she quickly secured an appointment as an investigator for the Howard Hughes Medical Institute. There, her lab continued to explore cell cycle control, and, among many seminal discoveries, she identified the genetic pathways that regulate chromosomal cohesion during both mitosis and meiosis. She also turned her attention to a new scientific problem: aneuploidy, or the consequences of chromosome segregation errors. For many years, it had been appreciated that the dysregulation of the cell cycle can lead to the generation of aneuploid progeny and that most human tumors are aneuploid. However, no scientist had ever attempted to systematically study aneuploid cells or understand how they are linked to tumorigenesis. As she did in so many cases, Angelika decided to push past what was known and directly interrogate the consequences of aneuploidy.

By using a series of clever genetic tricks, Angelika and the members of her lab created a set of yeast (and eventually, mouse) cells that had defined chromosomal aneuploidies. They discovered that aneuploidy of different chromosomes causes a stereotypical set of physiological changes, including altered proliferation, transcriptional deregulation, genomic instability, and metabolic abnormalities. One of us (J.M.S.) joined Angelika's lab as a graduate student to study the role of aneuploidy in tumor development. Unexpectedly, we found that aneuploid cells are resistant to oncogenic transformation, demonstrating that aneuploidy, rather than being just an early event in tumorigenesis, is an anticancer barrier that cells must overcome. Angelika's work on aneuploidy helped establish a new field of biological research, and her insights into the consequences of aneuploidy have spurred a widespread effort to identify aneuploidy-targeting compounds that could potentially be used to treat aneuploid cancers.

Angelika's scientific achievements were recognized by many awards and honors, including the National Science Foundation's Alan T. Waterman Award (the highest American honor for a scientist under 40), the Paul Marks Prize for Cancer Research, the Vanderbilt Prize in Biomedical Science, and the Breakthrough Prize in Life Sciences. She was elected to the National Academy of Sciences in 2010 and the American Academy of Arts and Sciences in 2017. Undoubtedly, she would have had many more discoveries ahead of her; on the day of her death, she published a preprint on bioRxiv that establishes a link between cytoplasmic volume and stem cell potential.

The life of every party, Angelika always had a brilliant story to tell, good advice to give, juicy gossip to loudly whisper, or a fascinating question to pose. In her lab, she brought together researchers from all over the globe and turned them into a surrogate family. She was a fierce defender of each of her students and was an especially strong advocate for the careers of the women whom she trained.

In her final weeks, Angelika remained a committed mentor and scientist. She continued to lead group meetings and advise her students from the confines of her hospital bed. As her current lab members wrote to her a few hours before her death, "Do not worry about us! You have equipped us with everything we need to succeed." While the scientific world will surely miss Angelika's joyful, boisterous presence, it is much richer for her remarkable discoveries and for the generations of scientists whose lives she inexorably changed. ■

10.1126/science.abf7124

PHOTO: SAMARA VISE

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [2]Department of Experimental Oncology, European Institute of Oncology, 20139 Milan, Italy. Email: sheltzer@cshl.edu

Published by AAAS

POLICY FORUM

ETHICS: COVID-19

COVID-19 vaccine trial ethics once we have efficacious vaccines

Some placebo-controlled trials can continue ethically after a candidate vaccine is found to be safe and efficacious

By **David Wendler,[1] Jorge Ochoa,[1] Joseph Millum,[1,2] Christine Grady,[1] Holly A. Taylor[1]**

The unprecedented effort to identify one or more safe and effective vaccines for COVID-19 includes more than 180 candidates in development (*1*), with at least 12 in phase 3 trials (*2*). The testing of so many vaccine candidates, in a pandemic of a disease for which there are to date limited treatment options, raises a critical challenge: What should researchers do if a vaccine candidate is judged to be safe and efficacious? Guidance from the U.S. Food and Drug Administration (FDA) states that in the event that a COVID-19 vaccine candidate is judged to be "safe and effective," discussion may be necessary "to address ethical arguments to break the blind and offer vaccine to placebo recipients" (*3*). We consider here two questions raised by this guidance: First, if a vaccine candidate is found to be safe and efficacious in a placebo-controlled trial, should the researchers continue that trial as designed? Second, should researchers continue to test other vaccine candidates using placebo-controlled trials? These two questions are especially timely given recent announcements by Pfizer and Moderna that their vaccine candidates have been found to be efficacious in preventing symptomatic COVID-19 (*4, 5*).

VACCINE SAFETY AND EFFICACY

A finding of efficacy occurs when a vaccine candidate being tested in a phase 3 trial meets its efficacy end point, either during an interim analysis or after study completion. The primary efficacy end point of current phase 3 trials is whether the vaccine candidate prevents symptomatic COVID-19 in individuals newly infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (*6*). A finding of efficacy in these trials thus implies that the vaccine candidate protects infected individuals from experiencing symptoms. A finding of safety occurs when there is sufficient data on the impact of the vaccine candidate to conclude that its potential benefits justify its risks in the target population.

A single finding of safety and efficacy may not be sufficient for a vaccine candidate to receive FDA approval. Depending on the strength of the data, it may be important to conduct additional research to learn more about the vaccine candidate and gain greater confidence in its safety and efficacy before offering it to millions of people. For example, the FDA frequently requires a finding of efficacy in two phase 3 trials before approving medical interventions for marketing to the public.

Still, a finding of safety and efficacy in a phase 3 trial is an ethically critical point in the vaccine development process. From that point, receiving the vaccine candidate is known to offer protection to many individuals. Hence, receiving a placebo instead is contrary to their interests, even though it may be necessary to collecting socially valuable data.

This tension between protecting research participants and gathering data that may benefit others is common to clinical research, and there is substantial literature on how to address it (*7*). Although this literature applies to the ethics of vaccine trials, the present situation is unparalleled.

There is a global crisis for which there are currently limited treatment options. Hence, billions of individuals could benefit from a safe and efficacious vaccine. And having so many vaccine candidates in development substantially increases the chances that one or more will be found to be safe and efficacious while others are still being tested. Limitations on current treatment options mean that it is in each individual's interests to receive the first vaccine found to be safe and efficacious, rather than participate in vaccine trials where they might receive placebo or an unproven vaccine candidate.

In addition, given the challenges of manufacturing sufficient supply for so many people, and given that different vaccine candidates may be efficacious in different populations, several vaccines may be needed to meet the global need. This highlights the potential social value of conducting additional trials after one or more vaccine candidates are found to be safe and efficacious.

Moreover, vaccine candidates typically have not been widely available until after they receive marketing approval by the FDA. However, the FDA has indicated that it may take the unprecedented step of making a vaccine candidate that is found to be safe and efficacious widely available through an emergency use authorization (EUA) before marketing approval. An EUA involves the FDA commissioner allowing unapproved medical products to be used in an emergency to diagnose, treat, or prevent serious or life-threatening diseases or conditions "when there are no adequate, approved, and available alternatives" (*8*).

PARTICIPANTS, NOT PATIENTS

Pfizer recently announced that at the first interim analysis, its vaccine candidate was found to be 95% effective at preventing symptomatic COVID-19 (*4, 5*). A week later, Moderna announced that its vaccine candidate was found to be 94.5% effective. At that point, the trials had enrolled more than 43,000 and 30,000 participants, respectively, of whom 94 and 95 developed symptomatic disease. If these findings are confirmed, they will represent an important public health advance, and provision of the vaccine candidates would be in the interests of many individuals. The FDA may thus grant an EUA, making either or both vaccines available before each study's completion and also before FDA marketing approval.

Some commentators argue that clinical trials are ethically acceptable only as long as there is insufficient evidence that the intervention offered in one arm is superior to what is offered in another arm, or to what is available outside the trial (*9*). This view implies that it would be unethical to continue Pfizer's and Moderna's placebo-controlled trials given the findings of efficacy. It also implies it would be unethical to test other unproven vaccine candidates against placebo.

Proponents defend this view on the grounds that it provides a way to protect participants from excessive risks. They argue further that researchers conducting clinical trials are obligated to treat participants consistent with their clinical interests and conclude that it is unethical to give participants a placebo once a safe and efficacious vaccine has been identified.

We disagree. This view fails to recognize that the obligations researchers have to their participants are distinct from the obli-

[1]Department of Bioethics, National Institutes of Health (NIH) Clinical Center, Bethesda, MD, USA. [2]Fogarty International Center, NIH, Bethesda, MD, USA. Email: dwendler@nih.gov

Published by AAAS

gations that clinicians have to their patients (*9*). It further ignores the existence of other mechanisms, including independent review and informed consent, to protect research participants from excessive risks. The differences between the ethics of clinical research and clinical care are reflected in the consensus that it can be ethically appropriate to invite research participants to accept some risks to collect socially valuable data. In particular, codes and guidelines around the world permit researchers to expose participants in clinical trials, including vaccine trials, to some risks to collect socially valuable data that cannot be obtained in a less risky way (*10*).

These guidelines have a critical implication for COVID-19 vaccine trial ethics: Researchers are not obligated to provide placebo recipients with a safe and efficacious vaccine once one has been identified. Instead, researchers are obligated to ensure that any plans to conduct placebo-controlled trials remain ethically appropriate given current evidence. Such plans must be reviewed and approved by an independent board, known as an institutional review board (IRB) in the United States, which must find that a number of conditions are satisfied (*11*). Two of these conditions are particularly important for determining the acceptability of conducting placebo-controlled trials after a vaccine candidate has been found to be safe and efficacious: Is the trial's risk-benefit profile still acceptable? Do participants consent?

ACCEPTABLE RISK-BENEFIT PROFILE

Continuing a blinded, placebo-controlled trial after the vaccine candidate being tested has been found to be safe and efficacious is justified only when the trial's risk-benefit profile remains acceptable. Making this determination requires answering three questions: (i) What is the social value of continuing the trial? (ii) What are the risks to participants of continuing the trial? (iii) Are the risks to participants sufficiently low and justified by the social value of continuing the trial?

Social value

Continuing a trial after the vaccine candidate has been found to be safe and efficacious can provide an opportunity to collect several types of socially valuable data. Of greatest importance, it can provide greater confidence in and a more precise point estimate of the vaccine's efficacy and offer an opportunity to collect additional safety data, including data on any uncommon or delayed side effects. It can also help to assess how long the vaccine's protective effect lasts; offer insight into the vaccine's impact in various subgroups, such as older individuals or those with comorbidities; and evaluate whether the vaccine candidate protects against infection itself.

Risks to participants

Once a vaccine candidate is found to be efficacious, participants in the placebo arm are known to be at higher risk of symptomatic disease than those in the active arm. The degree of risk depends on the chances that participants in the placebo arm will become infected, the risks they face if they are, and how much protection the vaccine candidate offers.

A volunteer receives an injection as part of a phase 3 trial of the vaccine candidate developed by Pfizer and BioNTech.

PHOTO: DOGUKAN KESKINKILIC/ANADOLU AGENCY VIA GETTY IMAGES

The chances that participants in the placebo arm will be infected depends on the local transmission rate, preventive measures they adopt, and the amount of time they remain on placebo. When participants are on placebo for a short time, the chances of infection are correspondingly low. For example, Moderna plans to continue its placebo-controlled trial until 151 cases of symptomatic disease are identified (*5*). Given current transmission rates, this may take as little as a few weeks.

Remaining on placebo for a few weeks, rather than accessing an efficacious vaccine, poses a low chance of substantial harm. Continuing on placebo for even longer periods also poses a low chance of substantial harm to individuals at low risk for severe disease.

Remaining on placebo for an extended period can pose considerable risks to individuals at high risk of severe disease. The extent of these risks depends critically on what options are available to them. In the setting of few effective treatments and potentially strained hospital systems, receiving placebo for an extended period rather than a safe and efficacious vaccine can pose substantial risks. However, if high-risk individuals would not have access to a safe and efficacious vaccine outside of research—for example, there is only enough supply for the trial or they are not part of a prioritized group that will receive the vaccine during the time of the trial—receiving placebo in a clinical trial poses few additional risks to them.

Are risks sufficiently low and justified?

There is no algorithm for determining how much social value a given clinical trial has and whether its social value justifies the risks participants face. As a result, IRBs tend to focus on ensuring that a trial has the potential to collect important data and that the risks of substantial harm are low.

Trials with the potential to collect data helpful for addressing a global pandemic have considerable social value. Inviting competent adults to participate in such trials can be ethical when doing so poses a small increase in their risk of experiencing substantial harm. This suggests that it can be ethically acceptable to continue a placebo-controlled trial for a short period after the vaccine candidate has been found to be safe and efficacious, even when par-

Published by AAAS

ticipants might be able to access the vaccine candidate outside the trial, for example, through an EUA (*12*). By contrast, if continuing the trial does not offer the opportunity to collect socially valuable data, or comparable data can be obtained in less risky ways, continuing the trial with a placebo arm for any length of time would be ethically problematic.

Inviting participants who are at low risk of severe disease to remain blinded and stay in the trial for a longer period can be acceptable when it offers the potential to collect data that might be helpful for addressing the pandemic. In most cases, continuing a blinded, placebo-controlled design with high-risk individuals for longer periods will not yield data of sufficient value to justify it. Exceptions might include when the individuals cannot access an efficacious vaccine outside the trial and their participation is needed to collect valuable data, or they are in a group for whom no efficacious vaccine candidate has been identified. Otherwise, individuals at high risk of severe disease should be unblinded and those on the placebo arm offered the vaccine within a redesigned study or given the opportunity to seek the vaccine outside the trial.

When the value of the data to be collected does not justify the risks of continuing the trial as designed, researchers have several options. They can unblind participants; offer placebo recipients the vaccine, possibly as part of an expanded access program; and follow them to collect additional data (*13*). Alternatively, researchers might redesign the trial, for example, to include a crossover in which the blind is maintained and those on the placebo arm receive the vaccine after they complete the placebo arm. Finally, in some cases, it may make sense to simply stop the trial and unblind participants, thus allowing those in the placebo arm to seek the vaccine elsewhere.

OTHER TRIALS

Even when a vaccine candidate is found to be safe and efficacious, there are likely to be good reasons to study others. Another vaccine candidate might be more effective, generate longer-lasting immunity, work better in certain subpopulations, provide greater protection against severe disease, or prevent infection better. Other candidates may also be superior with respect to cost or other practical considerations. For example, Pfizer's vaccine candidate must be stored at very low temperatures until 5 days before administration, a requirement which will reduce its availability in many parts of the world.

When there are safe and effective interventions available, the default is to use them rather than placebo as the comparator in clinical trials. Thus, if the Pfizer or Moderna vaccine candidates, or others, are confirmed to be safe and efficacious, researchers should consider whether their trials can be redesigned as a comparison between the vaccine candidate they are testing and a safe and efficacious vaccine. When making this decision, researchers and IRBs should take into account that an active comparison trial is likely to require larger sample sizes and extend the duration of the trial.

Redesigning ongoing trials to include an efficacious vaccine as an active comparator may not always be feasible or valuable. If continuing to evaluate the vaccine candidate nonetheless offers important social value, participants at low risk of severe disease might be invited to stay blinded and remain in the placebo-controlled trial. If continued participation appreciably increases participants' risk of severe disease, researchers should discuss participants' options for obtaining an efficacious vaccine outside research.

DISCLOSURE, UNBLINDING, CONSENT

Researchers are ethically obligated to inform participants of developments that might influence their willingness to remain in a clinical trial. Clearly, that a vaccine candidate has been found to be safe and efficacious meets this standard. Hence, investigators should inform participants in all trials of such a finding. This information should include the vaccine's safety record, the level of protection it provides, the populations for which it has been found to be safe and efficacious, and whether it might be available through an EUA or other means.

Researchers should then explain the plans for their trial going forward and solicit participants' consent. To minimize potential confusion and distrust (*14*), participants should understand the rationale behind the plan and that their prior consent does not obligate them to continue to participate. If researchers intend to continue with a placebo-controlled design, participants can be encouraged to remain in the trial. However, once a safe and efficacious vaccine candidate is available, knowing whether they received placebo or vaccine becomes relevant to a participant's decision whether to seek the vaccine outside research. Hence, participants who choose to leave the trial should be informed whether they received the vaccine candidate or placebo.

If one or more efficacious vaccines become widely available outside research, continuing a blinded, placebo-controlled trial might result in participants in both the active and placebo arms dropping out and seeking the vaccine elsewhere. Researchers should anticipate this possibility and consider whether they might unblind participants, offer an efficacious vaccine to the placebo recipients, and follow everyone. Alternatively, it might make sense to continue the blinded, placebo-controlled trial with a guarantee that individuals in the placebo arm will receive an efficacious vaccine once their participation in the study is completed. With careful planning and systematic assessment of the social value and risks, we believe it can be ethical to conduct some placebo-controlled trials of vaccine candidates for COVID-19 even after we have an efficacious vaccine. Doing so may be necessary to effectively address a pandemic that is causing so much harm around the world. ■

REFERENCES AND NOTES

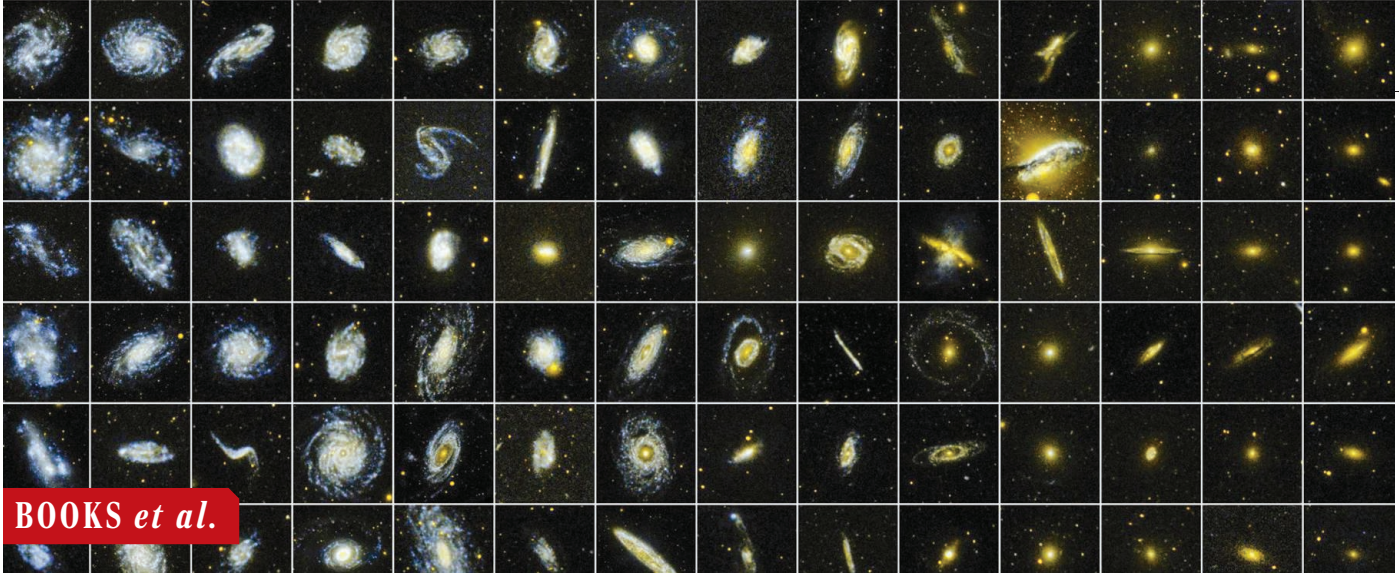
1. Bioworld, Biopharma products in development for COVID-19 (2020); www.bioworld.com/COVID19products.
2. J. Corum, S. L. Wee, C. Zimmer, Coronavirus vaccine tracker (2020); www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html.
3. FDA, "Development and licensure of vaccines to prevent COVID-19: Guidance for industry" (FDA-2020-D-1137, FDA, 2020); www.fda.gov/media/139638/download.
4. Pfizer, "Pfizer and BioNTech announce vaccine candidate against COVID-19 achieved success in first interim analysis from phase 3 study," 9 November 2020; www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-vaccine-candidate-against.
5. Moderna, "Moderna's COVID-19 vaccine candidate meets its primary efficacy endpoint in the first interim analysis of the phase 3 COVE study," 16 November 2020; https://investors.modernatx.com/news-releases/news-release-details/modernas-covid-19-vaccine-candidate-meets-its-primary-efficacy.
6. P. Doshi, *BMJ* **371**, m4037 (2020). doi:10.1136/bmj.m4037 Medline
7. J. Millum, C. Grady, *Contemp. Clin. Trials* **36**, 510 (2013). doi:10.1016/j.cct.2013.09.003 Medline
8. FDA, Emergency use authorization (2020); www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization.
9. S. P. Hey, A. J. London, C. Weijer, A. Rid, F. Miller, *BMJ* **359**, j5787 (2017).
10. Council for International Organizations of Medical Sciences (CIOMS), "International ethical guidelines for health-related research involving humans" (Report, CIOMS, 2016); https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf.
11. E. J. Emanuel, D. Wendler, C. Grady, *JAMA* **283**, 2701 (2000).
12. FDA, Emergency use authorization for vaccines to prevent COVID-19 guidance for industry (2020); www.fda.gov/media/142749/download.
13. M. W. Mccarthy, D. Oshinsky, A. Caplan, "Make pre-approval Covid-19 vaccines available through expanded access, not an EUA," *Stat*, 11 November 2020; www.statnews.com/2020/11/09/expanded-access-not-eua-for-distributing-preapproval-covid-19-vaccines/.
14. R. C. Warren, L. Forrow, D. A. Hodge Sr., R. D. Truog, *N. Engl. J. Med*. 10.1056/NEJMp2030033 (2020).

ACKNOWLEDGMENTS

Thanks to E. Erbelding for valuable discussion of this issue. This work was funded by the Intramural Research Program at the U.S. National Institutes of Health Clinical Center. The authors declare no competing interests. Disclaimer: The opinions expressed are the authors' own. They do not represent the position or policy of the National Institutes of Health, the U.S. Public Health Service, or the U.S. Department of Health and Human Services.

Published online 3 December 2020
10.1126/science.abf5084

Published by AAAS



Unlike this composite image of galaxies visualized by NASA's Galaxy Evolution Explorer (GALEX) spacecraft, the Universe is not boxed in, argues Barbour.

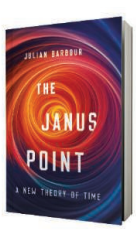
PHYSICS

An outside-the-box take on time

A physicist's provocative theory offers an optimistic view of our cosmic destiny

By **Matthew Johnson**

The history of the Universe thus far has certainly been eventful, marked by the primordial forging of the light elements, the birth of the first stars and their violent deaths, and the improbable origin of life on Earth. But will the excitement continue, or are we headed toward the ultimate mundanity of equilibrium in a so-called heat death? In *The Janus Point*, Julian Barbour takes on this and other fundamental questions, offering the reader a new perspective—illustrated with lucid examples and poetically constructed prose—on how the Universe started (or more precisely, how it did not start) and where it may be headed. This book is an engaging read, which both taught me something new about meat-and-potatoes physics and reminded me why asking fundamental questions can be so fun.



The Janus Point: A New Theory of Time
Julian Barbour
Basic Books, 2020. 398 pp.

Barbour argues that there is no beginning of time. The Big Bang, he maintains, was just a very special configuration of the Universe's fundamental building blocks, a shape he calls the Janus point. As we move away from this point, the shape changes, marking the passage of time. The "future," he argues, lies in both directions, hence the reference to Janus, the two-faced Roman god of beginnings and transitions.

Barbour illustrates his main points with a deceptively simple model known as the three-body problem, wherein three masses are subject to mutual gravitational attraction. In this context, the Janus point occurs when all three masses momentarily occupy the same point, in what is called a total collision. The special shape at the Janus point, explains Barbour, is an equilateral triangle, which is his model's version of the Big Bang. I found this imagery helpful when trying to understand the more abstract, and necessarily less technical, application of this concept to general relativity.

One of the main thrusts of Barbour's proposal is an attempt to reconcile cosmological evolution with the second law of thermodynamics and an arrow of time. As a reminder, the second law is the one that tells us that systems tend toward increasing entropy, sometimes colloquially referred to as disorder. In a closed system, if we begin with order (low entropy), then disorder will grow until the entropy reaches a maximum. The change in entropy can be used to define past (low entropy) and future (high entropy).

When applied to the Universe, the second law of thermodynamics evokes a number of questions. How did the Universe come to be in a state of low entropy, for example? Barbour's answer to this is the Janus point. In the three-body problem, the Janus point (i.e., a total collision) is an inevitability so long as the total angular momentum vanishes. Recall that angular momentum is among the sacred conserved quantities in classical mechanics, so vanishing at one time implies vanishing for all times. Barbour argues by analogy that the Universe, similarly restricted, must have a Janus point too. If true, the seemingly special configuration at the Janus point is not so special at all but in fact required by a conservation law.

So, what happens when entropy reaches a maximum? Barbour dodges this question by arguing that entropy is not the right factor to consider and that we should instead look to complexity. With time, he argues, the building blocks of the Universe can arrange themselves into increasingly complex structures.

Barbour does not buy the argument that the Universe is akin to a closed system whose entropy can reach a maximum and instead takes the optimistic outlook that complexity has no bound and our cosmic party can go on forever. But could this be so? The observed accelerated expansion of the Universe, the most likely explanation of which is a cosmological constant, suggests otherwise. A Universe with a cosmological constant is one that has event horizons—points of no return. And just as Stephen Hawking showed with black holes, the event horizon has a temperature and forms what can be thought of as a sort of box. Whether this means we should think of the Universe as a closed system remains unclear, but if that is the case, the entropy of the system will eventually be maximized. Change will cease, save for fleeting departures from equilibrium, and mundanity will prevail. Let us hope that Barbour is right. ■

10.1126/science.abf2440

The reviewer is at the Department of Physics and Astronomy, York University, Toronto, ON M3J 1P3, Canada, and the Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada. Email: mjohnson@perimeterinstitute.ca

IMAGE: NASA/JPL-CALTECH

Published by AAAS

SOCIAL SCIENCE

Social capital's value in the courtroom

Inequalities are reinforced during court proceedings and deliberations, finds a sociologist

By **Louise Seamster**

We are used to thinking about racial disparities in policing and arrests and disparities in sentencing and imprisonment, all of which produce pathways with multiple off-ramps for white and wealthy defendants. In his new book, *Privilege and Punishment*, sociologist Matthew Clair probes a less-examined period that exists between these two stages: the actual court proceedings and deliberations. Synthesizing knowledge gained from court watching, interviews with defendants and lawyers, and shadowing three public defenders in the Boston area, Clair zeroes in on this interval—when defendants' experiences should theoretically be relatively equal—to look at the subtle forms in which inequality can manifest.

Many treatments of racial inequality within the justice system assume that racial disparities must be produced through discrete acts of discrimination. But outright or obvious acts of racism are relatively rare in Clair's observations. His relational approach, which focuses on the relationship between lawyers and defendants, shows how inequality works more subtly, through ongoing interactions.

Researchers studying other interactional settings, such as schools and doctors' offices, have found that poor and working-class people are disadvantaged through patterns of deferential behavior. Meanwhile, middle-class actors are able to deploy assertive "styles of interaction that are either valued by gatekeepers…or that force these gatekeepers to provide them with more resources than their peers." But in interactions with one's lawyer, something like the opposite seems to happen: Poor and working-class defendants' attempts to wield knowledge, offer strategies, or otherwise advocate on their own behalf frequently backfire, finds Clair. Indeed, defendants' very familiarity with legal protocol, strategy, and baroque laws can serve as a strike against them, showing that they are not first-time offenders. Middle-class defendants, meanwhile, often take a more passive or deferential stance, conveying a trusting attitude toward their lawyer that apparently bodes well for the relationship.

Expertise is thus only valued in the "proper" context. Assertive defendants are perceived as potentially problematic and disrespectful of the process, privileging the authority and voice of lawyers and judges. Such defendants can gum up the works of an otherwise smoothly running system.

Inexpert attempts at self-advocacy could also serve a social bonding function for insiders. The development of professional in-groups can lead some clients to believe that their lawyer is working "hand-in-hand with the DA and the judge," as one defendant put it: "These people, after they leave [court], they go sit and do lunch."

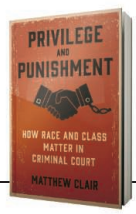
A defendant's lack of deference to the legal system can tip the scales of justice against him, finds Clair.

PHOTO: ISTOCK.COM/WESLEY VANDINTER

Clair's findings show the weak spot in exhortations that marginalized people should simply adopt the successful behaviors of the well-off in such situations. Just as "successful" parenting techniques of a given age tend to be whatever the upper class is doing at that moment, regardless of broader efficacy, there is no universally effective manual for navigating legal institutions.

The reviewer is at the Department of Sociology and Criminology and the African American Studies Program, Iowa University, Iowa City, IA 52242, USA. Email: louise-seamster@uiowa.edu

Privilege and Punishment: How Race and Class Matter in Criminal Court
Matthew Clair
Princeton University Press, 2020. 283 pp.



In the case of the courtroom, Clair shows that a defendant's social capital matters more than his or her technical knowledge. Several middle-class clients he observed were able to build a positive rapport with their lawyers through shared interests or culture. This is a complex performance of concurrent insider and outsider status, conveying ease and identification with the professional realm of the court and an outsider identity relative to the "criminal world." Clair suggests that these distinctions and cultural performances drive material outcomes, from verdicts to sentencing, as well as the degree to which a judgment affects one's life in the long term, thanks to expungement and record sealing. (Recall the judge sentencing Stanford rapist Brock Turner to just 6 months in jail, who predicted that prison "would have a severe effect" on Turner.)

Clair details some of the Kafkaesque bargains facing defendants. Accept the plea bargain or maintain your innocence and face greater retribution? Aim for probation or just take the jail sentence because probation is so hard to complete successfully without sufficient resources? In detailing these decisions, he successfully illuminates the distance between defendants' experiences and restorative justice.

The book is worth reading to the end for Clair's take on why this constitutes not just inequality but injustice—and what ought to be done about this injustice, in the context of contemporary abolition movements. Here, he suggests changes at multiple levels, ranging from shifting judge discretion and rules around lawyer selection to attenuating state involvement in punishment by removing laws and creating restorative justice processes. The biggest change he advocates involves shifting resources outside the criminal justice system altogether to meet people's other needs, rebuilding what sociologist Bruce Western, quoted here, calls "thick public safety." ■

10.1126/science.abe8289

Published by AAAS

Birds such as this red kite (*Milvus milvus*) are put at risk by the proliferation of Spanish wind farms.

PHOTO: © JOSÉ ANTONIO MARTÍNEZ

Edited by **Jennifer Sills**

Renewables in Spain threaten biodiversity

Reducing greenhouse gas emissions is a global priority. To meet this goal, the Spanish government is planning 89 GW of wind and solar photovoltaic energy in the draft of the National Integrated Energy and Climate Plan (PNIEC) for 2021–2030 (*1*). Despite the Spanish government's efforts to prevent a speculative bubble in the secondary market, there are already grid access permits for projects representing 121 GW that will be added to the 36 GW of renewables already installed (*2*), almost doubling the goals of the PNIEC. The new projects will affect hundreds of thousands of hectares, and there is no way to offset the huge quantity of valuable habitats that could be lost. Spain should adopt a more cautious approach to prevent a scenario in which energy goals are met at the expense of biodiversity.

Many renewables projects are planned in low-cost marginal soils of high ecological value, such as extensive cereal farmlands and wild mountain ranges harboring steppe birds and raptors that have in Spain their largest European or world populations. Photovoltaic energy needs huge amounts of land and will mostly affect declining species of steppe birds, which are poorly represented in the Spanish Natura 2000 network (*3*). Globally threatened large scavengers and other unique and scarce soaring birds are already paying a heavy toll due to the approximately 20,000 existing turbines, with demographic consequences for some threatened populations (*4*, *5*). Bat-killing figures are even higher, with a minimum of 200,000 deaths per year according to estimated mortality rates (*6*).

Studies forecasting mortality have shown scarce predictive power (*7*), and when mortality hotspots are detected, conflictive turbines are virtually never stopped to reduce bat and bird casualties as recommended (*8*). As a large-scale approach, the best way to reduce impacts is by choosing adequate locations (*9*). However, updated field information necessary to achieve this goal is often not available, and projects are authorized in areas with under-protected species because their status within regional and state listings is frequently outdated.

Studies designed to predict and monitor the incidence of renewable infrastructures are funded by energy companies (*10*), often with little supervision by governments, which precludes independence. The problem is exacerbated by the fragmentation of large projects (yielding reports of the smaller impacts associated with each part rather than a comprehensive assessment of the full project), the absence of an in-depth assessment of cumulative and synergistic environmental impacts, and decentralized administrative authority divided among the central state, regions, and municipalities.

We welcome renewables, but we urge Spanish authorities to correct these deficiencies and implement rigorous comprehensive planning based on the most updated ecological knowledge. We also call for a stronger commitment to more distributed and energy-saving policies that would reduce direct environmental impacts on biodiversity, such as energy efficiency, self-consumption, and improved energy performance in buildings.

David Serrano[1]*, Antoni Margalida[2], Juan M. Pérez-García[3], Javier Juste[4], Juan Traba[5], Francisco Valera[6], Martina Carrete[7], Joxerra Aihartza[8], Joan Real[9], Santi Mañosa[9], Carles Flaquer[10], Inazio Garin[8], Manuel B. Morales[5], J. Tomás Alcalde[11], Beatriz Arroyo[2], José A. Sánchez-Zapata[12], Guillermo Blanco[13], Juan J. Negro[4], José L. Tella[1], Carlos Ibañez[4], José L. Tellería[14], Fernando Hiraldo[1], José A. Donázar[1]

[1]Department of Conservation Biology, Estación Biológica de Donaña, Consejo Superior de Investigaciones Científicas (CSIC), E-41092 Sevilla, Spain. [2]Instituto de Investigación en Recursos Cinegéticos (CSIC–Universidad de Castilla-LaMancha), E-13005 Ciudad Real, Spain. [3]Department of Zoology, Universidad de Granada, E-18071 Granada, Spain. [4]Department of Evolutionary Biology, Estación Biológica de Donaña (CSIC), E-41092 Sevilla, Spain. [5]Department of Ecology, Centro de Investigación en Biodiversidad y Cambio Global, Universidad Autónoma de Madrid, E-28049 Madrid, Spain. [6]Estación Experimental de Zonas Áridas (CSIC), E-04120 Almería, Spain. [7]Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, E-41013 Sevilla, Spain. [8]Department of Zoology and Animal Cell Biology, Universidad del País Vasco/ Euskal Herriko Unibertsitatea, E-48940 Leioa, Spain. [9]Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals and Institut de la Recerca de la Biodiversitat, Universitat de Barcelona, E-08028 Barcelona, Spain. [10]Museu de Ciències Naturals de Granollers, E-08402 Granollers, Spain. [11]Spanish Association for Bat Research and Conservation, Universidad de Alcalá, 28805 Alcalá de Henares, Spain. [12]Department of Applied Biology, Miguel Hernández University, E-03202 Elche, Spain. [13]Department of Evolutionary Ecology, Museo Nacional de Ciencias Naturales (CSIC), E-28006 Madrid, Spain. [14]Department of Biodiversity, Ecology, and Evolution, Univesidad Complutente de Madrid, E-28040 Madrid, Spain.
*Corresponding author.
Email: serrano@ebd.csic.es

REFERENCES AND NOTES

1. Plan Nacional Integrado de Energía y Clima (PNIEC) 2021–2030 (2020); www.miteco.gob.es/es/prensa/pniec.aspx [in Spanish].

Published by AAAS

2. RED Eléctrica de España, "We help you connect your installation to the network" (2020); www.ree.es/es/actividades/acceso-conexion-y-puesta-en-servicio [in Spanish].
3. J. Traba, E. L. de la Morena, M. B. Morales, F. Suárez. *Biodivers. Conserv.* **16**, 3255 (2007).
4. Asociación Empresarial Eólica, "Wind Energy Yearbook 2020: All the information on the sector in 2019" (2020); www.aeeolica.org/comunicacion/publicaciones-aee/anuarios/4264-anuario-eolico-20-toda-la-informacion-del-sector-en-el-ano-2019 [in Spanish].
5. A. Sanz-Aguilar *et al.*, *Biol. Conserv.* **187**, 10 (2015).
6. C. C. Voigt, L. S. Lehnert, G. Petersons, F. Adorf, L. Bach, *Eur. J. Wildl. Res.* **61**, 213 (2015).
7. M. Ferrer *et al.*, *J. Appl. Ecol.* **49**, 38 (2012).
8. E. B. Arnett, M. M. P. Huso, M. R. Schirmacher, J. P. Hayes, *Front. Ecol. Environ.* **9**, 209 (2011).
9. C. Heuck *et al.*, *Biol. Conserv.* **236**, 44 (2019).
10. Law 21/2013, of December 9, on Environmental Evaluation (2013); www.boe.es/eli/es/l/2013/12/09/21/con.

10.1126/science.abf6509

Small Aral Sea brings hope for Lake Balkhash

In the recent Letter "Save Kazakhstan's shrinking Lake Balkhash" (16 October, p. 303), A. Ussenaliyeva calls for international attention to the ecological crisis now facing Lake Balkhash. Similar problems face other water bodies without effluents, such as the Great Salt Lake (*1*) and Lake Urmia in Iran (*2*). A parallel crisis severely damaged the Aral Sea (*2*, *3*), located between Kazakhstan and Uzbekistan. The action taken in response could provide a model for lakes suffering from desiccation.

Starting around 1960, diversion of inflowing water for irrigation purposes caused an accelerating regression and salinization of the Aral Sea (*3*). By 1990, the remaining waters sustained only organisms able to thrive in polyhaline conditions, and commercial fisheries, formerly of prime importance, had vanished (*4*). The desiccation also caused multiple severe health problems due to airborne dust (*5*). In 1992, a primitive dam was constructed across the Berg Strait to retain water in the northern Small Aral Sea, demonstrating a cost-effective way to maintain a low level of salinity. With financial help from the World Bank, the structure was later replaced by a stronger dam (*3*, *6*), which succeeded in rapidly restoring the Small Aral.

With increasing water volume and decreasing salinity, species that had become extinct during the regression crisis repopulated the sea by natural means from refugia in the affluent Syr Darya river system (*7*). Commercial fisheries that depend on the lake are thriving again (*7*, *8*). The Aral Sea crisis foreshadows what may face Lake Balkhash if no timely action is taken. It also shows that a relatively small financial effort can yield substantial results. The Aral Sea is surrounded by several countries that should, jointly with the international community, take renewed action to preserve even more of this system. This coalition should also lend its expertise to determine how attention and investment can best prevent a similar ecological crisis at Lake Balkhash and the Great Salt Lake.

Nikolai V. Aladin[1], Jens T. Høeg[2]*, Igor Plotnikov[3]
[1]Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia. [2]Department of Biology, University of Copenhagen, Copenhagen, Denmark. [3]Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia.
*Corresponding author. Email: jthoeg@bio.ku.dk

REFERENCES AND NOTES

1. M. L. Wine *et al.*, *Climate* **7**, 67 (2019).
2. S. Mischke, Ed., *Large Asian Lakes in a Changing World: Natural State and Human Impact* (Springer, Berlin, 2020).
3. P. Micklin, *Ann. Rev. Earth Planetary Sci.* **35**, 47 (2007).
4. I. S. Plotnikov, N. V. Aladin, A. A. Filippov, *Zool. Zhur.* **70**, 5 (1991) [in Russian].
5. I. Small *et al.*, *Environ. Health Perspect.* **109**, 547 (2001).
6. A. Ghany, A. S. Mohamed, "Project information document (concept stage)—Syr Darya control and Northern Aral Sea project, phase 2—P152001" (World Bank, 2014).
7. I. S. Plotnikov *et al.*, *Lakes Reserv.* **21**, 315 (2016).
8. D. H. Chen, "Once written off for dead, the Aral Sea is now full of life," *National Geographic* (2018).

10.1126/science.abf6682

Brazil's areas of not-so-permanent preservation

Recently, the National Council for the Environment in Brazil (CONAMA) repealed CONAMA Resolution 303/2002 (*1*), which has, for nearly two decades, set the rules for nation-wide Areas of Permanent Preservation (APPs). Unlike the Protected Areas officially set aside for conservation by the government or private property owners (*2*), APPs in Brazil require no governmental intervention to remain protected. Those who favor this repeal argue that CONAMA 303/2002 and its rules for APPs have been overwritten by the Forest Act approved in 2012 as Federal Law 12.651 (*3*, *4*). Although the Forest Act does cover the protection of some important ecosystems, such as mangroves (*3*), there are crucial environmental safeguards at risk should the repeal of CONAMA 303/2002 go into effect.

Under the Forest Act alone, coastal vegetation is an APP only if it acts to stabilize sand dunes, whereas CONAMA 303/2002 considers as APP all the coastal vegetation within 300 meters of the maximum height reached by a rising tide. Conversely, sand dunes will lose their protection as APPs unless they are stabilized by coastal vegetation. In addition, areas used for reproduction of migratory birds will no longer be classified as APPs. With the repeal of CONAMA 303/2002, habitat areas for threatened species will retain their categorization as APPs only if officially declared by the government, which rarely occurs.

After weeks under criticism (*5*), the repeal of CONAMA 303/2002 was temporarily suspended until the ultimate decision by the Supreme Federal Court in Brazil (*6*). We urge the Court to recognize the importance of these environmental safeguards remaining in place. In times of rapid climate change (*7*), sea level rise (*8*), and extinction crisis (*9*), the repeal of CONAMA 303/2002 will only jeopardize coastal management efforts and increase threats to species of conservation concern.

Milton A. U. de Andrade Junior[1] and Wagner Cleyton Fonseca[2,3]*
[1]State Department of Sustainable Economic Development, Florianópolis, Santa Catarina, Brazil. [2]Environmental Institute of Santa Catarina, Itajaí, Santa Catarina, Brazil. [3]Universidade do Vale do Itajaí, Itajaí, Santa Catarina, Brazil.
*Corresponding author.
Email: wagnerfonseca@univali.br

PHOTO: SANCHEZ/ALAMY STOCK PHOTO

The sand dunes of Joaquina Beach in southern Brazil could soon lose their protected status.

Published by AAAS

These views are the authors' and do not necessarily reflect those of the government of Santa Catarina.

REFERENCES AND NOTES

1. "CONAMA/Environmental Ministry Resolution 500/2020. 19 October 2020" (2020); www.in.gov.br/web/dou/-/resolucao-conama/mma-n-500-de-19-de-outubro-de-2020-284006009 [in Portuguese].
2. S. L. Maxwell *et al.*, *Nature* **586**, 217 (2020).
3. Federal Law 12.651, 25 May 2012 (2012); www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12651.htm [in Portuguese].
4. B. Soares-Filho *et al.*, *Science* **344**, 363 (2014).
5. C. Nugent, "Brazil Is burning—and President Bolsonaro's 'Terminator' environment minister is rolling back more protections," *Time* (2020).
6. Supreme Federal Court, "Motion of failure of fundamental precept 748" (2020); www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/adpf748.pdf [in Portuguese].
7. P. R. Elsen, W. B. Monahan, E. R. Dougherty, A. M. Merenlender, *Sci. Adv.* **6**, eaay0814 (2020).
8. C. Li, H. Held, S. Hokamp, J. Marotzke, *Sci. Adv.* **6**, eaaw9490 (2020).
9. G. Ceballos, P. R. Ehrlich, *Science* **360**, 1080 (2018).

10.1126/science.abf6705

TECHNICAL COMMENT ABSTRACTS

Comment on "Forest microclimate dynamics drive plant responses to warming"

Peter Schall and Steffi Heinrichs

Zellweger *et al.* (Reports, 15 May 2020, p. 772) claimed that a microclimatic debt, mainly controlled by canopy buffering, evolved in European forest understories. However, their analysis is based on circularity, as they explained the sum of three components by one of these components. The response of the understory to the thermal environment is generally weak.

Full text: dx.doi.org/10.1126/science.abd9920

Response to Comment on "Forest microclimate dynamics drive plant responses to warming"

Florian Zellweger, Pieter De Frenne, Jonathan Lenoir, Pieter Vangansbeke, Kris Verheyen, Markus Bernhardt-Römermann, Lander Baeten, Radim Hédl, Imre Berki, Jörg Brunet, Hans Van Calster, Markéta Chudomelová, Guillaume Decocq, Thomas Dirnböck, Tomasz Durak, Thilo Heinken, Bogdan Jaroszewicz, Martin Kopecký, František Máliš, Martin Macek, Marek Malicki, Tobias Naaf, Thomas A. Nagel, Adrienne Ortmann-Ajkai, Petr Petřík, Remigiusz Pielech, Kamila Reczyńska, Wolfgang Schmidt, Tibor Standovár, Krzysztof Świerkosz, Balázs Teleki, Ondřej Vild, Monika Wulf, David Coomes

Schall and Heinrichs question our interpretation that the climatic debt in understory plant communities is locally modulated by canopy buffering. However, our results clearly show that the discrepancy between microclimate warming rates and thermophilization rates is highest in forests where canopy cover was reduced, which suggests that the need for communities to respond to warming is highest in those forests.

Full text: dx.doi.org/10.1126/science.abf2939

Where Science Gets Social.

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY AAAS

Published by AAAS

Cite as: P. Schall, S. Heinrichs, *Science* 10.1126/science.abd9920 (2020).

Comment on "Forest microclimate dynamics drive plant responses to warming"

Peter Schall and Steffi Heinrichs*

Silviculture and Forest Ecology of the Temperate Zones, University of Göttingen, 37077 Göttingen, Germany.

*Corresponding author. Email: steffi.heinrichs@forst.uni-goettingen.de

Zellweger *et al.* (Reports, 15 May 2020, p. 772) claimed that a microclimatic debt, mainly controlled by canopy buffering, evolved in European forest understories. However, their analysis is based on circularity, as they explained the sum of three components by one of these components. The response of the understory to the thermal environment is generally weak.

In their recent paper, Zellweger *et al.* (*1*) quantified thermal macroclimate (MaC), thermal canopy buffering (CB), and community-based floristic temperature of understory plants (FT) for two points in time, a baseline survey and a recent resurvey, for plots located in 56 European regions. Decadal change rates of macroclimate (ΔMaC), canopy buffering (ΔCB), and floristic temperature (ΔFT, called thermophilization rate) were used to adjust for varying survey intervals. Microclimate change rate (ΔMiC) was defined as the sum of macroclimate change rate and canopy buffering change rate (ΔMiC = ΔMaC + ΔCB). The authors found a significant relationship between ΔMiC and ΔFT, whereas ΔMaC was not related to ΔFT. Reanalyzing this relationship using provided data (*2*) revealed a low effect size of 0.051° ± 0.012°C °C^{-1} (mean ± SEM) and a large scatter (R^2 = 1.1%; Fig. 1A). Thus, contrary to the authors' claim "that microclimate change ultimately drives organismal responses to climate change," microclimate change seems not to be the main driver of thermophilization in forest understories.

The authors continued by investigating how climate change affects a climatic debt. Climatic debt (ΔCD) was calculated by subtracting ΔFT from microclimate change rate (ΔCD_{MiC} = ΔMaC + ΔCB − ΔFT) and macroclimate change rate (ΔCD_{MaC} = ΔMaC − ΔFT) for each plot. Because ΔFT only weakly followed the climate change (Fig. 1A), a climatic debt automatically appeared. The authors argue that the understory lagged behind climate warming, leading to the accumulation of this climatic debt, but the validity of this assumption was not proven. Although a reanalysis of the baseline survey data (using floristic temperature data kindly provided by Zellweger) showed a stronger relationship between floristic temperature and climate than was found between rates, the slope was far from the assumed 1:1 relationship (Fig. 1B).

Thus, even for the baseline, floristic temperature was not strongly controlled by macroclimate and even less by microclimate; this does not correspond to the assumed direct linkage proposed by the climatic debt concept. These relationships have also not significantly changed for the resurvey, indicating that the understory community is stable in its thermal traits across space and time (Fig. 1, B and C). This also means that colder and warmer (closed and open) forests comprised nearly identical floristic temperatures. Rather than lagging behind temperature changes, forest plant communities are possibly well adapted to the periodic changes of canopy cover in forests (*3*) and often benefit from increased light availability after canopy opening (*4*). Therefore, climatic debt seems to be a construct not supported by the floristic data.

Zellweger *et al.* used this construct and related it to canopy buffering. The authors achieved an explained variance for microclimatic debt of 28.9%, suggesting that warming by reduced canopy buffering drives microclimatic debt and concluding that "increasing climatic debts in community responses to climate change mean that a growing number of species are occurring in suboptimal climate conditions, potentially accelerating the loss of biodiversity" and "that climate change impacts on forest plant communities have been reduced by higher standing stocks and associated cooling after increases in thermal buffering." These conclusions are, however, based on circular reasoning, as Zellweger *et al.* explained microclimatic and macroclimatic debt by a variant of the canopy buffering change rate, namely the canopy buffering absolute change (CB_{diff} = CB_R − CB_B), which is highly correlated with ΔCB (Pearson's r = 0.93), using linear mixed-effect models with region as a random effect. In doing so, Zellweger *et al.* explained a sum with a component of the sum, leading to dependency between explanatory and response variables. The authors basically analyzed the effect

of canopy buffering on climatic debt by modeling (ΔMaC + ΔCB – ΔFT) = f(ΔCB) for microclimatic debt and (ΔMaC – ΔFT) = f(ΔCB) for macroclimatic debt. Predictably, microclimatic debt but not macroclimatic debt was driven by canopy buffering, but this is only because canopy buffering explains itself in the microclimate case. We used hierarchical partitioning (*5*) to disentangle the contributions of the components of microclimatic debt (i.e., ΔMaC, ΔCB, ΔFT) on microclimatic debt (Table 1) and to explicitly show what the authors did. Microclimatic debt is consecutively driven by ΔMaC (45.9%), ΔCB (33.7%; ΔMiC = ΔMaC + ΔCB: 80.2%), and ΔFT (15.9%), and to 100% explained by all of its three components. Because of this circularity, the relationship between microclimatic debt and canopy buffering presented by Zellweger *et al.* does not provide any evidence for a response of the understory community to climate change. Thus, the conclusion of the authors that "a reduction of canopy cover impedes the ability of understory plant communities to respond to such high rates of warming" is not proven by the data.

We agree that forest understories experience climate change, but the floristic temperature seems to be relatively stable, at least until now (Fig. 1). There is no evidence that the relationship between the thermal affinity of understory plant communities and canopy closure or openness has changed markedly between the baseline survey and the resurvey. This is also supported by the weak relationship between thermophilization and climate change reported by Zellweger *et al.* Thus, there is no reason to suggest that canopy openings will create adverse thermal conditions for many species, in turn influencing forest management decisions.

Long-term data are highly valuable in detecting the effect of environmental change on ecosystems. Since the 2009 publication by Keith *et al.* (*6*), evidence for a homogenization and diversity loss of European forest understories has been increasing, with shade-tolerant, nutrient-demanding generalists expanding while light-demanding indicators of nutrient-poor sites disappear (*7*–*10*). Recently, Staude *et al.* (*11*) verified this using data from the same database as Zellweger *et al.* (*1*). Thus, when canopy openings (either natural or due to forest management) coexist with closed forests, they can rather be seen as a chance to maintain heterogeneity in forest understories (*12*) and to conserve plant species that benefited from traditional forest management practices centuries ago (*13*, *14*). By reducing the variation of canopy closure, we run the risk of accelerating the biodiversity loss in forest understories.

REFERENCES

1. F. Zellweger, P. De Frenne, J. Lenoir, P. Vangansbeke, K. Verheyen, M. Bernhardt-Römermann, L. Baeten, R. Hédl, I. Berki, J. Brunet, H. Van Calster, M. Chudomelová, G. Decocq, T. Dirnböck, T. Durak, T. Heinken, B. Jaroszewicz, M. Kopecký, F. Máliš, M. Macek, M. Malicki, T. Naaf, T. A. Nagel, A. Ortmann-Ajkai, P. Petřík, R. Pielech, K. Reczyńska, W. Schmidt, T. Standovár, K. Świerkosz, B. Teleki, O. Vild, M. Wulf, D. Coomes, Forest microclimate dynamics drive plant responses to warming. *Science* **368**, 772–775 (2020). doi:10.1126/science.aba6880 Medline
2. F. Zellweger, P. De Frenne, D. Coomes, Dataset for "Forest microclimate dynamics drive plant responses to warming". Dryad (2020); https://doi.org/10.5061/dryad.r7sqv9s83.
3. T. Hilmers, N. Friess, C. Bässler, M. Heurich, R. Brandl, H. Pretzsch, R. Seidl, J. Müller, Biodiversity along temperate forest succession. *J. Appl. Ecol.* **55**, 2756–2766 (2018). doi:10.1111/1365-2664.13238
4. L. Depauw, M. P. Perring, D. Landuyt, S. L. Maes, H. Blondeel, E. De Lombaerde, G. Brūmelis, J. Brunet, D. Closset-Kopp, J. Czerepko, G. Decocq, J. den Ouden, R. Gawryś, W. Härdtle, R. Hédl, T. Heinken, S. Heinrichs, B. Jaroszewicz, M. Kopecký, I. Liepiņa, M. Macek, F. Máliš, W. Schmidt, S. M. Smart, K. Ujházy, M. Wulf, K. Verheyen, Light availability and land-use history drive biodiversity and functional changes in forest herb layer communities. *J. Ecol.* **108**, 1411–1425 (2020). doi:10.1111/1365-2745.13339
5. A. Chevan, M. Sutherland, Hierachical partitioning. *Am. Stat.* **45**, 90–96 (1991).
6. S. A. Keith, A. C. Newton, M. D. Morecroft, C. E. Bealey, J. M. Bullock, Taxonomic homogenization of woodland plant communities over 70 years. *Proc. R. Soc. B* **276**, 3539–3544 (2009). doi:10.1098/rspb.2009.0938 Medline
7. K. Verheyen, L. Baeten, P. De Frenne, M. Bernhardt-Römermann, J. Brunet, J. Cornelis, G. Decocq, H. Dierschke, O. Eriksson, R. Hédl, T. Heinken, M. Hermy, P. Hommel, K. Kirby, T. Naaf, G. Peterken, P. Petřík, J. Pfadenhauer, H. Van Calster, G.-R. Walther, M. Wulf, G. Verstraeten, Driving factors behind the eutrophication signal in understorey plant communities of deciduous temperate forests. *J. Ecol.* **100**, 352–365 (2012). doi:10.1111/j.1365-2745.2011.01928.x
8. M. Kopecký, R. Hédl, P. Szabó, Non-random extinctions dominate plant community changes in abandoned coppices. *J. Appl. Ecol.* **50**, 79–87 (2013). doi:10.1111/1365-2664.12010 Medline
9. M. C. Jantsch, A. Fischer, H. S. Fischer, S. Winter, Shift in Plant Species Composition Reveals Environmental Changes During the Last Decades: A Long-Term Study in Beech (Fagus sylvatica) Forests in Bavaria, Germany. *Folia Geobot.* **48**, 467–491 (2013). doi:10.1007/s12224-012-9148-7
10. J. Reinecke, G. Klemm, T. Heinken, Vegetation change and homogenization of species composition in temperate nutrient deficient Scots pine forests after 45 yr. *J. Veg. Sci.* **25**, 113–121 (2014). doi:10.1111/jvs.12069
11. I. R. Staude, D. M. Waller, M. Bernhardt-Römermann, A. D. Bjorkman, J. Brunet, P. De Frenne, R. Hédl, U. Jandt, J. Lenoir, F. Máliš, K. Verheyen, M. Wulf, H. M. Pereira, P. Vangansbeke, A. Ortmann-Ajkai, R. Pielech, I. Berki, M. Chudomelová, G. Decocq, T. Dirnböck, T. Durak, T. Heinken, B. Jaroszewicz, M. Kopecký, M. Macek, M. Malicki, T. Naaf, T. A. Nagel, P. Petřík, K. Reczyńska, F. H. Schei, W. Schmidt, T. Standovár, K. Świerkosz, B. Teleki, H. Van Calster, O. Vild, L. Baeten, Replacements of small- by large-ranged species scale up to diversity loss in Europe's temperate forest biome. *Nat. Ecol. Evol.* **4**, 802–808 (2020). doi:10.1038/s41559-020-1176-8 Medline
12. P. Schall, M. M. Gossner, S. Heinrichs, M. Fischer, S. Boch, D. Prati, K. Jung, V. Baumgartner, S. Blaser, S. Böhm, F. Buscot, R. Daniel, K. Goldmann, K. Kaiser, T. Kahl, M. Lange, J. Müller, J. Overmann, S. C. Renner, E.-D. Schulze, J. Sikorski, M. Tschapka, M. Türke, W. W. Weisser, B. Wemheuer, T. Wubet, C. Ammer, The impact of even-aged and uneven-aged forest management on regional biodiversity of multiple taxa in European beech forests. *J. Appl. Ecol.* **55**, 267–278 (2018). doi:10.1111/1365-2664.12950
13. R. Hédl, M. Kopecký, J. Komárek, Half a century of succession in a temperate oakwood: From species-rich community to mesic forest. *Divers. Distrib.* **16**, 267–276 (2010). doi:10.1111/j.1472-4642.2010.00637.x
14. S. Heinrichs, W. Schmidt, Biotic homogenization of herb layer composition between two contrasting beech forest communities on limestone over 50 years. *Appl. Veg. Sci.* **20**, 271–281 (2017). doi:10.1111/avsc.12255

ACKNOWLEDGMENTS
We thank F. Zellweger for providing the floristic temperatures for the baseline survey and the resurvey, and for discussion. **Data and materials availability:** R script of the analyses conducted for the present comment is available at https://figshare.com/s/8b039b6782961eeaef92.

24 July 2020; accepted 9 November 2020
Published online 11 December 2020
10.1126/science.abd9920

Table 1. Independent and joint contribution of macroclimate change rate, canopy buffering change rate, and thermophilization rate to microclimatic debt from hierarchical partitioning. The total contribution reflects the explained variance R^2 of single linear models. A partitioning with microclimate change rate (ΔMiC) instead of its components (ΔMaC and ΔCB) contributes 80.2% to microclimatic debt. Considering region as a random effect in mixed models affects results only marginally ($R^2_{marginal}$: ΔMiC, 79.1%; ΔMaC, 45.7%; ΔCB, 33.6%; ΔFT, 13.5%).

	Independent	Joint contribution	Total
Macroclimate change rate (ΔMaC)	44.9%	1.0%	45.9%
Canopy buffering change rate (ΔCB)	37.3%	−3.6%	33.7%
Thermophilization rate (ΔFT)	17.8%	−1.9%	15.9%
Sum	100.0%	−4.5%	95.5%

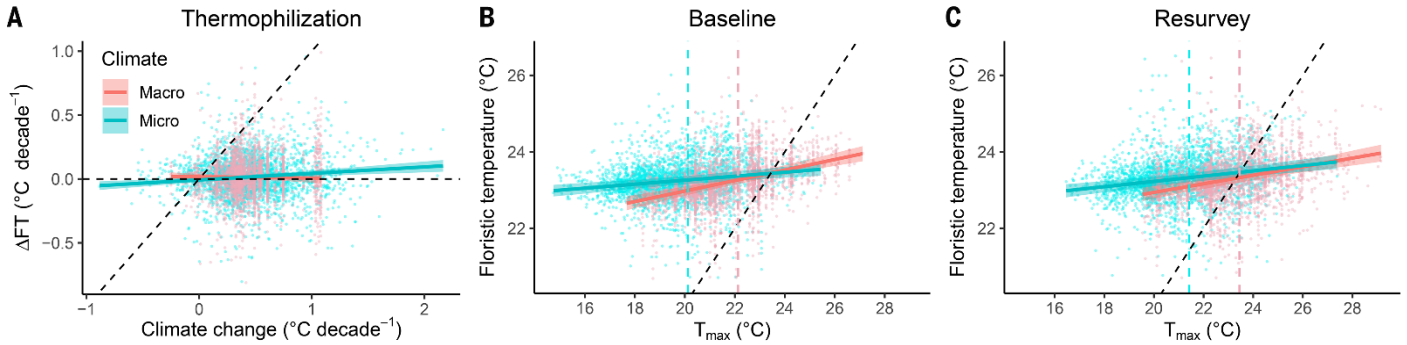


Fig. 1. Microclimate is only weakly related to floristic temperature and its change rate. (**A**) Mean effect size (±SEM) of microclimate change on thermophilization, based on linear mixed-effect models with region as a random effect, is 0.051° ± 0.012°C °C^{-1} decade^{-1}. It contributes only slightly to the explained variance of thermophilization as indicated by a marginal (conditional) R^2 value of 1.1% (11.8%). Mean effect size (±SEM) of macroclimate change on thermophilization is −0.010° ± 0.026°C °C^{-1} decade^{-1} [R^2: 0.0% (10.1%)]. The scatter shown is adjusted for the random effect. Black dashed lines indicate the 1:1 relationship and the zero line. (**B** and **C**) Mean effect size (±SEM) of maximum temperature on floristic temperature, based on linear mixed-effect models with region as a random effect, was about double for macroclimate relative to microclimate for the baseline survey (0.137° ± 0.017°C °C^{-1} versus 0.052° ± 0.013°C °C^{-1}; R^2: 11.1% versus 1.8%) (B) as well as for the resurvey (0.112° ± 0.018°C °C^{-1} versus 0.069° ± 0.013°C °C^{-1}; R^2: 6.6% versus 2.8%) (C). The scatter shown is adjusted for the random effect. Vertical dashed lines show the means of macro- and microclimate. The 1:1 relationship is indicated by the black dashed line. Slopes comparing the baseline and the resurvey are not significantly different.

Cite as: F. Zellweger *et al.*, *Science* 10.1126/science.abf2939 (2020).

Response to Comment on "Forest microclimate dynamics drive plant responses to warming"

Florian Zellweger[1,2]*, Pieter De Frenne[3], Jonathan Lenoir[4], Pieter Vangansbeke[3], Kris Verheyen[3], Markus Bernhardt-Römermann[5], Lander Baeten[3], Radim Hédl[6,7], Imre Berki[8], Jörg Brunet[9], Hans Van Calster[10], Markéta Chudomelová[11], Guillaume Decocq[4], Thomas Dirnböck[12], Tomasz Durak[13], Thilo Heinken[14], Bogdan Jaroszewicz[15], Martin Kopecký[16,17], František Máliš[18,19], Martin Macek[16], Marek Malicki[20], Tobias Naaf[21], Thomas A. Nagel[22], Adrienne Ortmann-Ajkai[23], Petr Petřík[16], Remigiusz Pielech[24], Kamila Reczyńska[25], Wolfgang Schmidt[26], Tibor Standovár[27], Krzysztof Świerkosz[28], Balázs Teleki[29,30], Ondřej Vild[11], Monika Wulf[21], David Coomes[1]

[1]Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK. [2]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland. [3]Forest and Nature Lab, Department of Environment, Faculty of Bioscience Engineering, Ghent University, Melle-Gontrode, Belgium. [4]UR "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN, UMR 7058 CNRS-UPJV), Université de Picardie Jules Verne, 80037 Amiens Cedex 1, France. [5]Institute of Ecology and Evolution, Friedrich Schiller University Jena, D-07743 Jena, Germany. [6]Institute of Botany of the Czech Academy of Sciences, CZ-602 00 Brno, Czech Republic. [7]Department of Botany, Faculty of Science, Palacký University in Olomouc, CZ-78371 Olomouc, Czech Republic. [8]Institute of Environmental and Earth Sciences, University of Sopron, H-9400 Sopron, Hungary. [9]Swedish University of Agricultural Sciences, Southern Swedish Forest Research Centre, Box 49, 230 53 Alnarp, Sweden. [10]Research Institute for Nature and Forest (INBO), B-1000 Brussels, Belgium. [11]Institute of Botany of the Czech Academy of Sciences, CZ-602 00 Brno, Czech Republic. [12]Environment Agency Austria, A-1090 Vienna, Austria. [13]Department of Plant Physiology and Ecology, University of Rzeszów, PL-35-959 Rzeszów, Poland. [14]General Botany, Institute of Biochemistry and Biology, University of Potsdam, 14469 Potsdam, Germany. [15]Białowieża Geobotanical Station, Faculty of Biology, University of Warsaw, 17-230 Białowieża, Poland. [16]Institute of Botany of the Czech Academy of Sciences, CZ-252 43 Průhonice, Czech Republic. [17]Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, CZ-165 21 Prague 6 - Suchdol, Czech Republic. [18]Faculty of Forestry, Technical University in Zvolen, SK-960 01 Zvolen, Slovakia. [19]National Forest Centre, SK-960 01 Zvolen, Slovakia. [20]Department of Botany, Faculty of Biological Sciences, University of Wrocław, Wrocław, Poland. [21]Leibniz Centre for Agricultural Landscape Research (ZALF), D-15374 Muencheberg, Germany. [22]Department of Forestry and Renewable Forest Resources, Biotechnical Faculty, University of Ljubljana, Ljubljana 1000, Slovenia. [23]Department of Hydrobiology, Institute of Biology, University of Pécs, H-7624 Pécs, Hungary. [24]Department of Forest Biodiversity, Faculty of Forestry, University of Agriculture, Kraków, Poland. [25]Department of Botany, Institute of Environmental Biology, University of Wrocław, PL-50-328 Wrocław, Poland. [26]Department of Silviculture and Forest Ecology of the Temperate Zones, University of Göttingen, Göttingen, Germany. [27]Department of Plant Systematics, Ecology and Theoretical Biology, Institute of Biology, L. Eötvös University, H-1117 Budapest, Hungary. [28]Museum of Natural History, University of Wrocław, PL-50-335 Wroclaw, Poland. [29]Institute for Regional Development, University of Pécs, H-7100 Szekszárd, Hungary. [30]Department of Ecology, University of Debrecen, H-4032 Debrecen, Hungary.

*Corresponding author. Email: florian.zellweger@wsl.ch

Schall and Heinrichs question our interpretation that the climatic debt in understory plant communities is locally modulated by canopy buffering. However, our results clearly show that the discrepancy between microclimate warming rates and thermophilization rates is highest in forests where canopy cover was reduced, which suggests that the need for communities to respond to warming is highest in those forests.

Biological communities accumulate a climatic (thermal) debt when their response to warming does not keep up with the warming rate itself. Forest understory plant communities appear to respond particularly slowly to warming, and thus climatic debts are commonly observed in forest understory plant communities (*1*, *2*). In line with conventional approaches used in the cited literature [e.g., (*1*, *2*)], we define the climatic debt as the difference between the thermophilization rate and the rate of climate warming, which are two independently calculated variables.

Schall and Heinrichs (*3*) question the validity of our climatic debt analysis because there is no 1:1 relationship between the floristic temperatures and the macro- and microclimate temperatures at one point in time. In this discussion, it is important to keep in mind that the commonly applied approach to inferring temperature conditions from plant species composition data (i.e., floristic temperatures) (*4*, *5*) is not designed to accurately reflect the actual temperatures at a particular time step across space, because these data are based on broad distribution ranges of individual species co-occurring in the community and are subject to considerable uncertainty, as outlined in detail by Rodríguez-Sańchez *et al.* (*6*). Thus, caution is needed when using such floristic temperatures to study spatial variation of community temperature preferences along a climatic gradient, as done by Schall and Heinrichs, especially when the studied macrocli-

matic gradient is relatively short, as is the case in our study (the interquartile ranges of spatial macroclimate temperatures in the baseline and resurveys were only 2.4° and 2.2°C, respectively). However, in our paper we studied temporal, not spatial, changes in floristic temperatures per unit time, and for this purpose floristic temperatures have been shown to be unbiased and robust, providing a solid database for our thermophilization and climatic debt analyses (*6*). Furthermore, our floristic temperatures per species are based on long-term macroclimate data during the vegetation period, whereas the macro- and microclimate change data were calculated independently on the basis of summer maximum temperatures prevailing in the 5 preceding years of the respective field surveys. Schall and Heinrichs' finding that the floristic temperatures at one time step more closely follow the macroclimate than the microclimate temperatures is also very plausible and expected, exactly because the floristic temperatures were calculated from macroclimate data, not microclimate data, which are not yet available at such scales. The patterns shown in Schall and Heinrichs' figure 1, B and C, can thus be explained very well and provide no reason to question the validity of our climate debt assessment. We also note that the mean thermophilization rates we found in our study compare well to the rates found in other lowland forests in Europe (*1*), as well as in forests in the Andes (*2*), and that the variation of our floristic temperatures for a given unit of spatial macroclimate change is similar to the respective variation found in these two studies. Schall and Heinrichs' claim that the climatic debt seems to be a construct not supported by the floristic data does not hold.

Schall and Heinrichs further question our interpretation of the effect of canopy buffering on climatic debt. As outlined above, and in agreement with Schall and Heinrichs' concerns, it is obvious that microclimate warming and microclimatic debt are not independent from each other. In fact, given the frequently observed slow response of forest plant communities (*1*, *2*), it can be expected that microclimate debts are more pronounced in areas with high rates of warming. We show that changes in temperature buffering due to dynamics in canopy cover are an important and integral component of forest microclimate warming. Regulating effects of canopy opening (e.g., due to clearcutting or tree mortality) on forest-floor temperature can act independently from macroclimate warming (*7*). It is thus reasonable to expect that changes in canopy-modulated temperature buffering are related to the microclimate debt, but not to macroclimate debt. What we show in figure 3A of (*7*) is the contribution of the change in temperature buffering to the microclimatic debt. We argue that this is an interesting contribution because (i) it can be substantial but has been largely ignored in the literature, and (ii) it shows how the need for communities to respond to warming is locally modulated by canopy cover dynamics. We agree with Schall and Heinrichs' claim that the relationship between microclimate debt and canopy buffering does not provide evidence of a response of the understory community to climate change. Evidence of such a response is provided in figure 2 of (*7*). However, the said relationship shows how the need for communities to respond to warming is locally modulated by canopy buffering, despite the non-independence between microclimate debt and microclimate warming. An ecologically realistic assessment of how much communities lag behind warming requires data about microclimate warming. In contrast to Schall and Heinrichs' critique and given our finding that thermophilization is more related to microclimate warming than to macroclimate warming [figure 2 of (*7*)], we indeed argued that microclimate warming, and not macroclimate warming, ultimately drives organismal responses to warming. These results also fully support our conclusion that a reduction in canopy buffering leads to higher rates of warming, thus seriously increasing the pressure for plant communities to respond to warming.

We agree with Schall and Heinrichs' conclusion that understory plant communities have responded only weakly to warming temperatures. That is actually the crucial point, because at the same time the temperatures themselves have warmed markedly, causing a climatic debt. It may indeed be that the microclimate temperature variations observed within our studied forests were within the thermal tolerances of most species, but given their slow response, the currently observed climate warming rates continue to erode their thermal safety margins, especially at their equatorward range boundaries. This has serious implications for future forest diversity and function (*8*) and will likely be felt first and most strongly in forests subject to relatively high local warming rates due to a reduction in canopy buffering.

Finally, we fully and evidently agree that deep shade is not beneficial for many species, and we have never claimed otherwise. Instead, we make the point that it is important to consider the effects of different forest management practices on local microclimates in any endeavor to safeguard forest biodiversity in a warming world. Forest ecologists are very much aware that many species benefit from canopy openings, and many of us have extensively worked on this topic (*9*–*11*). We also fully agree that at a landscape scale, biodiversity is enhanced by forest patches with deep shade and more open patches or edge habitat, as shown by many studies, including our own work (*12*, *13*). Environmental change affects forest understory plant communities in many ways, and we agree with the general remarks in Schall and Heinrichs' concluding paragraphs. However, in terms of recent warming effects on forest biodiversity, we argue that microclimate warming and its local drivers deserve increased attention.

REFERENCES

1. R. Bertrand, J. Lenoir, C. Piedallu, G. Riofrío-Dillon, P. de Ruffray, C. Vidal, J.-C. Pierrat, J.-C. Gégout, Changes in plant community composition lag behind climate warming in lowland forests. *Nature* **479**, 517–520 (2011). doi:10.1038/nature10548 Medline
2. B. Fadrique, S. Báez, Á. Duque, A. Malizia, C. Blundo, J. Carilla, O. Osinaga-Acosta, L. Malizia, M. Silman, W. Farfán-Ríos, Y. Malhi, K. R. Young, F. Cuesta C, J. Homeier, M. Peralvo, E. Pinto, O. Jadan, N. Aguirre, Z. Aguirre, K. J. Feeley, Widespread but heterogeneous responses of Andean forests to climate change. *Nature* **564**, 207–212 (2018). doi:10.1038/s41586-018-0715-9 Medline
3. P. Schall, S. Heinrichs, Comment on "Forest microclimate dynamics drive plant responses to warming". *Science* **370**, eabd9920 (2020).
4. P. De Frenne, F. Rodríguez-Sánchez, D. A. Coomes, L. Baeten, G. Verstraeten, M. Vellend, M. Bernhardt-Römermann, C. D. Brown, J. Brunet, J. Cornelis, G. M. Decocq, H. Dierschke, O. Eriksson, F. S. Gilliam, R. Hédl, T. Heinken, M. Hermy, P. Hommel, M. A. Jenkins, D. L. Kelly, K. J. Kirby, F. J. G. Mitchell, T. Naaf, M. Newman, G. Peterken, P. Petřík, J. Schultz, G. Sonnier, H. Van Calster, D. M. Waller, G.-R. Walther, P. S. White, K. D. Woods, M. Wulf, B. J. Graae, K. Verheyen, Microclimate moderates plant responses to macroclimate warming. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18561–18565 (2013). doi:10.1073/pnas.1311190110 Medline
5. V. Devictor, C. van Swaay, T. Brereton, L. Brotons, D. Chamberlain, J. Heliölä, S. Herrando, R. Julliard, M. Kuussaari, Å. Lindström, J. Reif, D. B. Roy, O. Schweiger, J. Settele, C. Stefanescu, A. Van Strien, C. Van Turnhout, Z. Vermouzek, M. WallisDeVries, I. Wynhoff, F. Jiguet, Differences in the climatic debts of birds and butterflies at a continental scale. *Nat. Clim. Change* **2**, 121–124 (2012). doi:10.1038/nclimate1347
6. F. Rodríguez-Sánchez, P. De Frenne, A. Hampe, Uncertainty in thermal tolerances and climatic debt. *Nat. Clim. Change* **2**, 636–637 (2012). doi:10.1038/nclimate1667
7. F. Zellweger, P. De Frenne, J. Lenoir, P. Vangansbeke, K. Verheyen, M. Bernhardt-Römermann, L. Baeten, R. Hédl, I. Berki, J. Brunet, H. Van Calster, M. Chudomelová, G. Decocq, T. Dirnböck, T. Durak, T. Heinken, B. Jaroszewicz, M. Kopecký, F. Máliš, M. Macek, M. Malicki, T. Naaf, T. A. Nagel, A. Ortmann-Ajkai, P. Petřík, R. Pielech, K. Reczyńska, W. Schmidt, T. Standovár, K. Świerkosz, B. Teleki, O. Vild, M. Wulf, D. Coomes, Forest microclimate dynamics drive plant responses to warming. *Science* **368**, 772–775 (2020). doi:10.1126/science.aba6880 Medline
8. D. Landuyt, E. De Lombaerde, M. P. Perring, L. R. Hertzog, E. Ampoorter, S. L. Maes, P. De Frenne, S. Ma, W. Proesmans, H. Blondeel, B. K. Sercu, B. Wang, S. Wasof, K. Verheyen, The functional role of temperate forest understorey vegetation in a changing world. *Glob. Change Biol.* **25**, 3625–3641 (2019). doi:10.1111/gcb.14756 Medline
9. O. Vild, J. Roleček, R. Hédl, M. Kopecký, D. Utinek, Experimental restoration of coppice-with-standards: Response of understorey vegetation from the conservation perspective. *For. Ecol. Manage.* **310**, 234–241 (2013). doi:10.1016/j.foreco.2013.07.056 Medline
10. H. Van Calster, L. Baeten, K. Verheyen, L. De Keersmaeker, S. Dekeyser, J. E. Rogister, M. Hermy, Diverging effects of overstorey conversion scenarios on the understorey vegetation in a former coppice-with-standards forest. *For. Ecol. Manage.* **256**, 519–528 (2008). doi:10.1016/j.foreco.2008.04.042
11. K. Verheyen, I. Fastenaekels, M. Vellend, L. De Keersmaeker, M. Hermy, Landscape factors and regional differences in recovery rates of herb layer richness in Flanders (Belgium). *Landsc. Ecol.* **21**, 1109–1118 (2006). doi:10.1007/s10980-006-7247-7
12. A. Valdés, J. Lenoir, P. De Frenne, E. Andrieu, J. Brunet, O. Chabrerie, S. A. O. Cousins, M. Deconchat, P. De Smedt, M. Diekmann, S. Ehrmann, E. Gallet-Moron, S. Gärtner, B. Giffard, K. Hansen, M. Hermy, A. Kolb, V. Le Roux, J. Liira, J. Lindgren, L. Martin, T. Naaf, T. Paal, W. Proesmans, M. Scherer-Lorenzen, M. Wulf, K. Verheyen, G. Decocq, High ecosystem service delivery potential of small woodlands in agricultural landscapes. *J. Appl. Ecol.* **57**, 4–16 (2020). doi:10.1111/1365-2664.13537
13. S. Govaert, C. Meeussen, T. Vanneste, K. Bollmann, J. Brunet, S. A. O. Cousins, M. Diekmann, B. J. Graae, P.-O. Hedwall, T. Heinken, G. Iacopetti, J. Lenoir, S. Lindmo, A. Orczewska, M. P. Perring, Q. Ponette, J. Plue, F. Selvi, F. Spicher, M. Tolosano, P. Vermeir, F. Zellweger, K. Verheyen, P. Vangansbeke, P. De Frenne, Edge influence on understorey plant communities depends on forest management. *J. Veg. Sci.* **31**, 281–292 (2020). doi:10.1111/jvs.12844

16 October 2020; accepted 9 November 2020
Published online 11 December 2020
10.1126/science.abf2939

RESEARCH

IN *SCIENCE* JOURNALS

Edited by Michael Funk

INDIAN MONSOON

Season of the drought

The Indian monsoon is a critical source of water for hundreds of millions of people, and when it fails to deliver its normal quantity of rain, enormous human, economic, and ecological costs can be incurred. Monsoon droughts are not always seasonal, however. Borah *et al*. found that nearly half of all monsoonal droughts were subseasonal and characterized by a steep decline in late-season rainfall. Moreover, this type of subseasonal drought appears to be related to a distinct cold anomaly in the North Atlantic Ocean, raising the possibility that monsoon droughts may be more predictable. —HJS *Science*, this issue p. 1335

A woman carries water across a dry lakebed in India, where a late-season decline in the monsoon can lead to drought.

SOLAR CELLS

Efficiency from hole-selective contacts

Perovskite/silicon tandem solar cells must stabilize a perovskite material with a wide bandgap and also maintain efficient charge carrier transport. Al-Ashouri *et al*. stabilized a perovskite with a 1.68–electron volt bandgap with a self-assembled monolayer that acted as an efficient hole-selective contact that minimizes nonradiative carrier recombination. In air without encapsulation, a tandem silicon cell retained 95% of its initial power conversion efficiency of 29% after 300 hours of operation. —PDS

Science, this issue p. 1300

CANCER IMMUNOTHERAPY

Stem-like T cells mediate response

Adoptive cell transfer (ACT) is a type of immunotherapy that uses a patient's own T lymphocytes to recognize and attack cancer. ACT has been effective in treating certain patients with metastatic melanoma and is being applied to treat some epithelial cancers. Krishna *et al*. investigated why some cancer patients respond to ACT, whereas others do not. They identified a population of $CD8^+$ T cells that had stem-like surface markers that were associated with effective tumor cell killing and favorable response of melanoma patients to ACT. Only a small subset of T cells specific against tumor mutations were found in this stem-like state, whereas most mutation-reactive T cells were terminally differentiated. These findings could be of value in improving cancer immunotherapy outcomes. —PNK

Science, this issue p. 1328

CELL DEATH

Corals catch fire

Pyroptosis is a form of inflammatory cell death that occurs in response to pathogen infection and results in the release of intracellular contents mediated by the pore-forming gasdermin family proteins. Jiang *et al*. identified a conserved gasdermin E homolog in corals that is cleaved by both coral and human caspase 3 to form two active N-terminal isoforms, each capable of inducing pyroptosis. After pathogen infection, caspase-dependent gasdermin E activation was associated with mitochondrial disruption and necrosis in the reef-building coral species *Pocillopora damicornis*. Gasdermin-mediated cell death is likely conserved in some invertebrates and may represent an immune defense activated in corals during bacterial infection resulting from environmental stress. —CO

Sci. Immunol. **5**, eabd2591 (2020).

CLIMATE CHANGE

A decline in the carbon fertilization effect

One source of uncertainty in climate science is how the carbon fertilization effect (CFE) will contribute to mitigation of anthropogenic climate

PHOTO: SAM PANTHAKY/STRINGER/GETTY IMAGES

Published by AAAS

change. Wang *et al*. explored the temporal dynamics of CFE on vegetation photosynthesis at the global scale. There has been a decline over recent decades in the contribution of CFE to vegetation photosynthesis, perhaps owing to the limiting effects of plant nutrients such as nitrogen and phosphorus. This declining trend has not been adequately accounted for in carbon cycle models. CFE thus has limitations for long-term mitigation of climate change, and future warming might currently be underestimated. —AMS

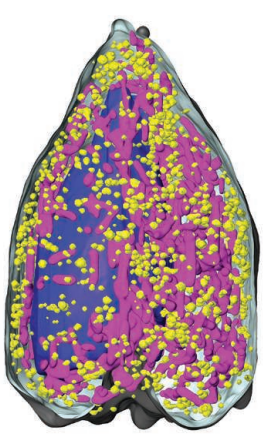
Science, this issue p. 1295

INTRACELLULAR IMAGING

Subcellular map of vesicle maturation

Insulin-containing vesicles in pancreatic β cells must migrate from the cell interior to the cell surface after stimulation by glucose. Two studies now use related whole-cell imaging techniques, soft x-ray tomography, and cryo–electron tomography to resolve the distribution, size, density, and location of insulin-containing vesicles as a function of time. White *et al*. visualized a mesoscale map of whole cells, and Zhang *et al*. provided higher-resolution structural information in specific "neighborhoods" of the cell. Understanding insulin regulation of circulating glucose by β cells will be advanced by this global picture of intracellular dynamics along the insulin secretion pathway. —PLY

Sci. Adv. 10.1126/sciadv.abc8262, 10.1126/sciadv.abc8258 (2020).



A colored soft x-ray tomogram shows insulin vesicles (yellow spheres) distributed within a pancreatic β cell

CREDITS (FROM LEFT TO RIGHT): WHITE *ET AL.*; XIAO XUE AT THE HONG KONG POLYTECHNIC UNIVERSITY

QUANTUM INFORMATION

Molecular qubits that respond to light

Spins in solid-state systems such as quantum dots and defect centers in diamond can easily be controlled by light for use in quantum information processing. More challenging is tuning their properties and making large arrays, something that can be done more easily with spins in molecules. Bayliss *et al*. combined the advantages of the two approaches by designing and characterizing three related molecular species that are optically addressable. The molecules consist of a central chromium ion surrounded by organic ligands, and their spin and optical properties can be tailored by simply changing the positions of methyl groups on the ligands. —JS

Science, this issue p. 1309

ULTRACOLD CHEMISTRY

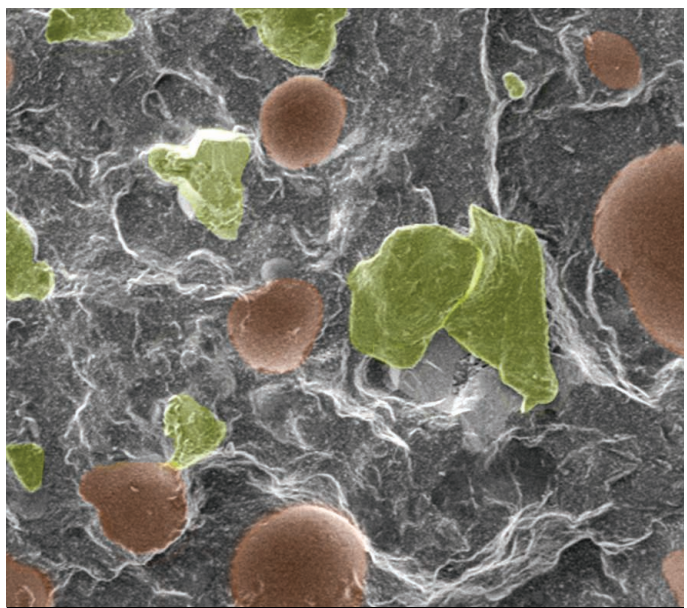
Electric field shielding of ultracold molecules

Because reactive collisions limit the lifetime of ultracold molecular ensembles, controlling chemical reactivity at ultralow temperatures has been a long-standing goal. Using large electric fields that trigger resonant dipolar interactions between potassium-rubidium molecules trapped in a quasi–two-dimensional geometry, Matsuda *et al*. report suppression of the reactive loss rate in the vicinity of the dipolar-mediated resonances by up to an order of magnitude below the background value. The proposed shielding mechanism is general and is expected to be effective in three-dimensional geometry. It could also be used for creating long-lived quantum molecular gases of other polar molecules under strong electric fields. —YS

Science, this issue p. 1324

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**



RADIATIVE COOLING

A smarter radiative cooler

Passive daytime radiative cooling materials allow heat to be transported into outer space through Earth's atmospheric infrared transparency window. Xue *et al*. focus on engineering inexpensive coating materials for buildings, which take advantage of daytime radiative cooling but suppress it in the night when cooling no longer is needed. This is accomplished by widening the emissivity spectrum over the entire mid-infrared range. The strategy also uses common materials for the coatings that are more likely to withstand long-term weathering. —BG

Adv. Mater. **32**, e1906751 (2020).

False-colored scanning electron microscope image of a coating for cooling buildings, containing hollow glass microspheres (red) and fluorescent pigment (green)

CANCER

Alleviating side effects

Platinum-based chemotherapy drugs, such as cisplatin and oxaliplatin, are commonly used to treat diverse cancer types. However, their use is limited by side effects, particularly vomiting, anorexia, muscle wasting, and weight loss. Breen *et al*. show that the amounts of the cytokine growth differentiation factor 15 (GDF15) increase in the circulation of patients with colorectal cancer, non–small cell lung cancer (NSCLC), and ovarian cancer who were treated with platinum chemotherapy. Moreover, the amount of circulating GDF15 was correlated with weight loss in metastatic colorectal cancer patients receiving oxaliplatin. Neutralization of GDF15 using monoclonal antibodies in nonhuman primates treated with cisplatin attenuated vomiting and anorexia. GDF15 also reversed weight loss in mice with NSCLC treated with cisplatin. This work shows that GDF15 has

Published by AAAS

DEFORESTATION

Health care and conservation

Logging in tropical forests affects the global climate, and there is an urgent need to reduce the rate of loss. Jones *et al*. show how access to health care for an indigenous community can contribute to this goal. Over a 10-year study in rural Borneo, they found that logging was reduced by 70% in forests adjacent to communities for whom a local health care clinic was established. The health outcomes were positive, too, with reductions in the incidence of tuberculosis, malaria, and other tropical diseases. This pattern of positive outcomes holds promise for human health and forest conservation in tropical forest areas that are home to impoverished communities. —AMS *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28515 (2020).

To reduce deforestation, rural communities in Borneo are exchanging chainsaws for affordable health care.

a causal role in these side effects and could be targeted to enable optimal chemotherapy treatment. —GKA

Cell Metab. **32**, 938 (2020).

HEMATOPOIESIS

Fetal versus adult hematopoiesis

Development of an organism over time requires a multitude of factors, cells, and movements. The fetal-to-adult transition is a time of major transformation. The hematopoietic system is one development process that sees important changes to enable differential cell composition and activity. Li *et al*. used single-cell RNA-seq, ATAC-seq, and ChIP-seq in mice and observed a gradual, rather than an abrupt, shift in gene expression and the epigenetic mechanism for hematopoietic stem cells and hematopoietic progenitor cells. Indeed, neonatal and adult expression can occur at the same time and the changes seem to occur in an uncoordinated manner rather than following a strict gene regulatory network. Just before birth, type I interferon is activated and plays a specific role at the perinatal to postnatal hematopoietic system switch. —BAP

Cell Stem Cell **27**, 732 (2020).

TOPOLOGICAL OPTICS

Confining light—bound and protected

The traditional method of confining light does so within the cavity formed by two mirrored surfaces. Cerjan *et al*. explored another form of light confinement using the higher-order topological insulating states of crystals. They show that these features provide protection against defects and chiral light propagation. Controlling the symmetry of artificial crystal structures to produce topologically protected bound states could prove a flexible route to fabricate devices from materials where the usual methods to confine and manipulate light fail. —ISO

Phys. Rev. Lett. **125**, 213901 (2020).

HIV

Wheeling out HIV elimination

Several sub-Saharan countries have been prioritized by the Joint United Nations Programme on HIV/AIDS (UNAIDS) for HIV elimination. This will require 90% treatment coverage by the end of 2030, but poor transport infrastructure means incomplete access to medication in some countries. Palk *et al*. developed a geospatial model for understanding the difficulties of reaching health care facilities in Malawi, a country severely affected by HIV. They mapped health care facilities with the density of HIV prevalence and quantified the difficulty of travel across Malawi's landscape by a friction surface raster map. If bicycles are used, then the catchment size for a health care facility is substantially larger than if people walk, and the required 90% treatment coverage for elimination becomes achievable. Bicycles are already used as "ambulances" in rural areas, but Malawi's bicycle fleet is small and in poor repair. One straightforward route for beating HIV (and many other health conditions) here and in similar countries could lie simply in boosting the supply of bicycles, an established and well-tested technology. —CA

Lancet Glob. Health **8**, E1555 (2020).

MEDICAL SCREENING

Who is getting mammograms?

Whether and when to recommended medical screening can be contentious. Einav *et al*. suggest that too little is known about how women who only seek mammograms in response to recommendations at age 40 ("compliers") might differ from women who don't wait for recommendations ("always-takers") or who never get mammograms ("never-takers"). They estimate that always-takers have higher rates of in situ and invasive cancer than the general population, whereas compliers are less likely to have invasive cancer and no more likely to have in situ cancer than never-takers. Screening debates must consider such selection effects because the typical woman who responds to the recommendation appears to be distinct from the broader population of women covered by the recommendation. —BW

Amer. Econ. Rev. **110**, 3836 (2020).

PHOTO: BAY ISMOYO/AFP VIA GETTY IMAGES

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

DISEASE DYNAMICS

Emergence to endemism

The emergence of a devastating transmissible facial cancer among Tasmanian devils over the past few decades has caused substantial concern for their future because these animals are already threatened by a regional distribution and other stressors. Little is known about the overall history and trajectory of this disease. Patton *et al*. used an epidemiological phylodynamic approach to reveal the pattern of disease emergence and spread. They found that low Tasmanian devil densities appear to be contributing to slower disease growth and spread, which is good news for Tasmanian devil persistence and suggests that care should be taken when considering options for increasing devil populations. —SNV

Science, this issue p. 1293

SYSTEMS BIOLOGY

From phenotype to structure

Much insight has come from structures of macromolecular complexes determined by methods such as crystallography or cryo–electron microscopy. However, looking at transient complexes remains challenging, as does determining structures in the context of the cellular environment. Braberg *et al*. used an integrative approach in which they mapped the phenotypic profiles of a comprehensive set of mutants in a protein complex in the context of gene deletions or environmental perturbations (see the Perspective by Wang). By associating the similarity between phenotypic profiles with the distance between residues, they determined structures for the yeast histone H3-H4 complex, subunits Rpb1-Rpb2 of yeast RNA polymerase II, and subunits RpoB-RpoC of bacterial RNA polymerase. Comparison with known structures shows that the accuracy is comparable to structures determined based on chemical cross-links. —VV

Science, this issue p. 1294; see also p. 1269

BIOPHYSICS

Rheology of aging protein condensates

Protein condensates that form by undergoing liquid-liquid phase separation will show changes in their rheological properties with time, a process known as aging. Jawerth *et al*. used laser tweezer–based active and microbead-based passive rheology to characterize the time-dependent material properties of protein condensates (see the Perspective by Zhang). They found that condensate aging is not gelation of the condensates, but rather a changing viscoelastic Maxwell liquid with a viscosity that strongly increases with age, whereas the elastic modulus stays the same. —MSL

Science, this issue p. 1317; see also p. 1271

BATTERIES

Cracking the problem of cracking cathodes

Polycrystalline cathode materials that contain a combination of nickel, manganese, and cobalt have been used for advanced lithium batteries. These materials fracture at high voltage, which increases surface area and leads to more side reactions and shorter cycle life. Using single-crystalline samples as model materials, Bi *et al*. observed changes in nickel-rich cathodes to study the fracture behavior under well-characterized conditions. As the material is charged and lithium is removed, specific planes glide over one another and microcracks are observed. However, this process is reversed on discharge, removing all traces of the microcracking. The authors developed a diffusion-induced stress model to understand the origin of the planar gliding and propose ways to stabilize these nickel-rich cathodes in working batteries. —MSL

Science, this issue p. 1313

PALEOCLIMATE

Controlling atmospheric carbon dioxide

The atmospheric concentration of carbon dioxide (CO_2) has varied substantially over the past million years in tandem with the glacial cycle. Although it is widely agreed that upwelling of Southern Ocean water is a key factor, the finer details about what caused these CO_2 variations are of great importance for understanding climate. Ai *et al*. identified three modes of change in Southern Ocean upwelling, adding a third to two previously recognized ones. This new mode can help explain better the relative timing of the glacial and CO_2 cycles. —HJS

Science, this issue p. 1348

EVOLUTION

Diversity does not drive speciation

The role of the environment in the origin of new species has long been debated. Harvey *et al*. examined the evolutionary history and species diversity of suboscine birds in the tropics (see the Perspective by Morlon). Contrary to expectations that the tropics have higher rates of speciation, the authors observed that higher and more constant speciation rates occur in harsh environments relative to the tropics. Thus, for this group of birds, diversification in temperate to Arctic regions followed by the movement and retention of species in the tropics results in their higher local levels of species diversity. —LMZ

Science, this issue p. 1343; see also p. 1268

CORONAVIRUS

Antibodies predating infection

Immunological memory after infection with seasonal human coronaviruses (hCoVs) may potentially contribute to cross-protection against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Ng *et al*. report that in a cohort of 350 SARS-CoV-2–uninfected individuals, a small proportion had circulating immunoglobulin G (IgG) antibodies that could cross-react with the S2 subunit of the SARS-CoV-2 spike protein (see the Perspective by Guthmiller and Wilson). By contrast, COVID-19 patients generated IgA, IgG, and IgM antibodies that recognized both the S1 and S2 subunits. The anti-S2 antibodies from SARS-CoV-2–uninfected patients showed specific neutralizing activity against both SARS-CoV-2 and SARS-CoV-2 S pseudotypes. A much higher percentage of SARS-CoV-2–uninfected children and adolescents were positive for these antibodies compared with adults. This pattern may be due to the fact that children and adolescents generally have higher hCoV infection rates and a more diverse antibody repertoire, which may explain the age distribution of COVID-19 susceptibility. —STS

Science, this issue p. 1339; see also p. 1272

IMMUNOLOGY

Small NOD to big changes

The innate pathogen-sensing protein NOD1 mediates proinflammatory responses to molecular signals from pathogens, but certain genetic variants in the *NOD1* gene are associated with an increased risk for developing gastric cancer. Rommereim *et al.* found that small increases in NOD1 abundance resulted in disproportionately large increases in the expression of inflammatory

Published by AAAS

genes and oncogenes. A 1.2- to 1.3-fold increase in NOD1 abundance reduced the amount of ligand required to activate NOD1 in a monocytic cell line. NOD1-mediated transcriptional responses became ligand independent upon a 1.5-fold increase in NOD1 abundance. —WW

Sci. Signal. **13**, eaba3244 (2020).

INFLUENZA

Influencing influenza immunity

Humans are exposed to influenza virus throughout their lifetimes through a combination of infections and vaccinations. It remains unclear whether different exposure routes induce distinct influenza-specific immunological memory. Dugan *et al.* found that infection-induced antibodies reacted to non-neutralizing epitopes of influenza virus, whereas vaccination-induced antibodies reacted to neutralizing epitopes. Infection-induced antibodies also preferentially responded to influenza strains present during an individual's childhood. Passive transfer of vaccination-induced antibodies, but not infection-induced antibodies, protected mice in a model of influenza infection. These findings demonstrate that existing influenza-specific memory and route of exposure influence influenza immunity. —CM

Sci. Transl. Med. **12**, eabd3601 (2020).

CORONAVIRUS

A history of vaccine safety

Vaccines are safe and effective and save millions of lives each year. Vaccine development has led to rigorous safety protocols to ensure that lessons from history are not repeated. In a Perspective, Knipe *et al.* discuss the history of vaccine safety issues that have led to current regulations and protocols, including pausing trials when any adverse event arises and ensuring safety monitoring after regulatory approval. The need for a COVID-19 vaccine and possible early regulatory approval before the completion of phase 3 trials could risk our ability to collect comprehensive safety data. It is important that safety, as well as efficacy, is evaluated before deployment. —GKA

Science, this issue p. 1274

ECONOMICS

Taxing mental health

Mental equilibrium is essential for an economically productive life in both industrialized and developing countries. Accumulating evidence shows that mental ill-health and poverty tend to be traveling partners, but which is the cause? Ridley *et al*. reviewed the literature on natural and controlled economic experiments involving individuals living in poverty. The authors sought to resolve the mechanisms whereby poverty triggers mental illness and how mental illness compounds poverty. Their results reveal the benefits of cash support and of low-cost therapeutic interventions for those suffering from mental illness under poverty. —CA

Science, this issue p. 1289

MOLECULAR BIOLOGY

CiBER-seq dissects genetic networks

Cells integrate environmental signals and internal states to dynamically control gene expression. Muller *et al*. developed a technique to dissect this cellular logic by linking targeted, genome-wide genetic perturbations with a deep-sequencing readout that quantitatively measured the expression phenotype induced by each perturbation. The method, dubbed CiBER-seq, was able to recapitulate known regulatory pathways linking protein synthesis with nutrient availability in budding yeast cells. Unexpectedly, the authors found that the cellular logic also appears to consider protein production machinery in this decision. By uncovering additional facets of this deeply conserved pathway, the findings demonstrate the utility of comprehensive and quantitative CiBER-seq profiling in mapping the gene networks underlying cellular decisions. —SMH

Science, this issue p. 1290

DEVELOPMENTAL BIOLOGY

Trunk formation in a dish

Building mammalian embryos from self-organizing stem cells in culture would accelerate the investigation of morphogenetic and differentiation processes that shape the body plan. Veenvliet *et al*. report a method for generating embryonic trunk-like structures (TLSs) with a neural tube, somites, and gut by embedding mouse embryonic stem cell aggregates in an extracellular matrix surrogate. Live imaging and comparative single-cell transcriptomics indicate that TLS formation is analogous to mouse development. TLSs therefore provide a scalable, tractable, and accessible high-throughput platform for decoding mammalian embryogenesis at a high level of resolution. —BAP

Science, this issue p. 1291

DEVELOPMENTAL BIOLOGY

Timing and trigger of cell polarization

During mammalian embryo development, the first cell fate decision separates the progenitors of the trophectoderm (destined to form the placenta) from the inner cell mass (which forms all tissues of the embryonic and yolk sac). A key event for this first lineage segregation is the establishment of apicobasal cell polarity. This event is set to occur at a fixed developmental stage. The factors that trigger the establishment of cell polarity as well as its temporal regulation have remained unknown. Zhu *et al*. show in mouse embryos that three molecular regulators—Tfap2c, Tead4, and RhoA—are sufficient to advance the timing of cell polarization with subsequent cell fate specification and morphogenesis. —BAP

Science, this issue p. 1292

Published by AAAS

REVIEW SUMMARY

ECONOMICS

Poverty, depression, and anxiety: Causal evidence and mechanisms

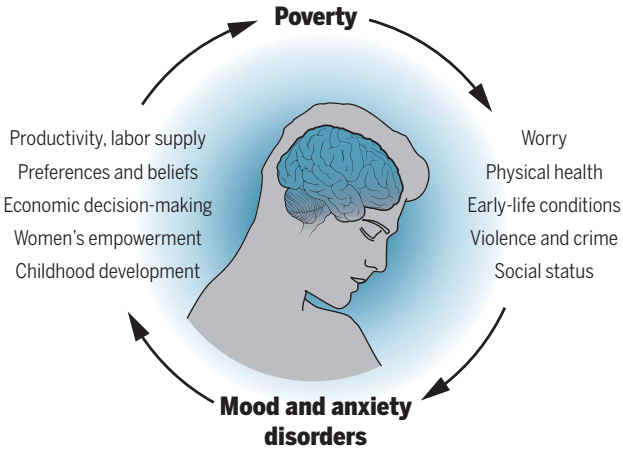
Matthew Ridley, Gautam Rao, Frank Schilbach*, Vikram Patel

BACKGROUND: Depression and anxiety disorders are together responsible for 8% of years lived with disability globally. Contrary to widely held preconceptions, these are not diseases of affluence. Within a given location, those with the lowest incomes are typically 1.5 to 3 times more likely than the rich to experience depression or anxiety.

ADVANCES: Recent research has established a bidirectional causal relationship between poverty and mental illness. Researchers have begun to isolate the underlying mechanisms, which can guide effective policies to protect the mental health of those living in poverty.

We now know that loss of income causes mental illness. Negative income shocks, such as bad harvests due to poor rainfall or job losses due to factory closures, worsen mental health. Conversely, cash transfers and broader antipoverty programs reduce depression and anxiety in randomized trials. Multiple mechanisms mediate this causal chain. Poverty is associated with volatile income and expenditures. The resulting worries and uncertainty can worsen mental health. Providing health, employment, or weather insurance, or other ways of smoothing shocks, may thus lower depression and anxiety. Living in inadequate housing in low-income neighborhoods, the poor are also more exposed to environmental stresses such as pollution, temperature extremes, and challenging sleep environments, which can cause mental illness. Early-life conditions—poverty experienced in childhood and in utero—increase the likelihood of poor nutrition and other stressors, resulting in impaired cognitive development and adult mental illness. This makes a strong case for providing financial support to pregnant women and caregivers of young children. Poverty is also associated with worse physical health; greater exposure to trauma, violence, and crime; and lower social status, each of which may affect mental health.

Mental illness in turn worsens economic outcomes for individuals. Studies show that randomized interventions to treat mental illnesses increase days worked. Depression and anxiety directly affect the way people think, by capturing their attention and distorting their memory. Such effects are likely to influence economic preferences and beliefs and thus distort important economic decisions made by individuals, such as how much to work, invest, and consume. Reduced concentration and greater fatigue reduce work productivity, and the social stigma of mental illness may further worsen labor-market outcomes. Mental illness appears to increase the likelihood of catastrophic health expenditures for individuals through its comorbidity with chronic illnesses such as diabetes and heart disease. Mental illness may also hinder education and skill acquisition among youth and exacerbate gender inequalities through its disproportionate prevalence among women. Parental mental illness can also influence children's cognitive development and educational attainment, transmitting mental illness and poverty across generations.



The causal relationship between poverty and common mental illnesses. This schematic shows the principal mechanisms we identify, on the basis of theory and empirical evidence, through which poverty and depressive and anxiety disorders interact.

OUTLOOK: The burden of mental illness is likely to increase in the coming decades. Although richer individuals within a given location are less likely to be mentally ill, richer countries do not have lower rates of mental illness. Thus, aggregate economic growth alone is unlikely to reduce mental illness. Climate change is likely to worsen mental health, both directly through the effect of higher temperatures on mood and through reductions in agricultural yields because of changes in rainfall and water supply, more frequent weather-related disasters, and an increased likelihood of violent conflict. Technological change and globalization create large overall economic gains but also concentrated groups of losers whose mental health may be compromised. The spread of social media and associated technologies may also be harming mental health, especially among adolescents.

Policy action on mental health is vital, as is interdisciplinary research on the mechanisms that link poverty and mental illness. Recently developed approaches to psychotherapy, delivered through nonspecialist providers, provide a scalable and effective approach to improving mental health in low-income countries. Given the associated economic benefits of improved mental health, such interventions should be a part of the antipoverty toolkit alongside more traditional economic interventions. Understanding the most effective combination of economic and psychological support in different populations is an important next step. A priority for research is testing for a mental health–based "poverty trap." If such poverty traps exist, then powerful one-time interventions will have large long-run effects as gains in mental health and economic outcomes reinforce one another. Evaluations of economic interventions should routinely measure mental health, and long-run evaluations of mental health interventions should measure potential impacts on poverty and other key economic outcomes. The causal relationship between poverty and mental health is even more pertinent given the ongoing pandemic, which has disproportionately affected the poor and may have lasting impacts on their economic and mental well-being. A massive investment in mental health was already long overdue. It has now become critically urgent. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: fschilb@mit.edu
Cite this article as M. Ridley *et al.*, *Science* **370**, eaay0214 (2020). DOI: 10.1126/science.aay0214

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aay0214

REVIEW

ECONOMICS

Poverty, depression, and anxiety: Causal evidence and mechanisms

Matthew Ridley[1], Gautam Rao[2], Frank Schilbach[1]*, Vikram Patel[3,4]

Why are people who live in poverty disproportionately affected by mental illness? We review the interdisciplinary evidence of the bidirectional causal relationship between poverty and common mental illnesses—depression and anxiety—and the underlying mechanisms. Research shows that mental illness reduces employment and therefore income, and that psychological interventions generate economic gains. Similarly, negative economic shocks cause mental illness, and antipoverty programs such as cash transfers improve mental health. A crucial step toward the design of effective policies is to better understand the mechanisms underlying these causal effects.

Depression and anxiety are the most common mental illnesses: 3 to 4% of the world's population suffers from each at any given time, and they are together responsible for 8% of years lived with disability globally (*1*). Contrary to widely held preconceptions from the 20th century, these are not "diseases of affluence" (*2*, *3*). Within a given location, those living in poverty are at least as likely to suffer as the rich. By some measures, the poor are substantially more likely than the affluent to experience mental ill-health. Rates of depression, anxiety, and suicide correlate negatively with income (*4*–*7*) and employment (*5*, *8*). Those with the lowest incomes in a community suffer 1.5 to 3 times more frequently from depression, anxiety, and other common mental illnesses than those with the highest incomes (*5*). For example, in India, 3.4% of those in the lowest income quintile experience depression at any given time, compared with 1.9% of those in the highest quintile (Fig. 1).

In this Review, we explore the evidence for the bidirectional causal relationship between poverty and mental health and its underlying mechanisms. Poverty is also correlated with poor physical health (*9*, *10*), but the relationship between mental illness and poverty has been overlooked and is worth emphasizing. Mental health has historically not been considered a priority by economists and policymakers, and until recently, mental health care had not been evaluated as an antipoverty tool. Mental health services are underresourced relative to physical health systems. On average, countries spend 1.7% of their health budgets on mental health, even though 14% of years lived with disability globally are known to be caused by depression, anxiety, and other mental illnesses (*1*). Low- and middle-income countries spend an even smaller share of already small health budgets on mental health (Fig. 2). Despite the existence of cost-effective treatments, such low investments in mental health have contributed to treatment gaps of more than 80% globally for common mental illnesses, which is much larger than for major physical health conditions (*11*–*14*). Mental and physical health are tightly connected: When mental health problems coexist with physical health problems, health outcomes, disability, and costs tend to be much worse (*15*–*17*). However, unlike most physical health conditions, mental disorders may directly distort economic decision-making in ways that perpetuate poverty, by directly affecting cognitive function, preferences, and beliefs.

What are the causal links between poverty and mental illness? Can economic policies improve psychological well-being? Can psychological interventions reduce poverty? Any attempt to understand this relationship must acknowledge the complexity and multidimensional nature of both mental health and poverty. Mental health in the broadest possible sense has been defined as "a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community" (*18*). This definition includes both happiness or life satisfaction, which also correlate positively with income (*19*), and symptoms associated with anxiety and mood disorders, such as depression. The two are clearly related; depression and anxiety are strong determinants of happiness (*20*), and ultimately, mental health and even mental illnesses such as depression and anxiety exist along a continuum.

We focus on the causal evidence that links poverty with depressive and anxiety disorders, the most common mental illnesses, which we refer to here using the more general terms "mental health" and "mental illness." Definitions of these illnesses and a brief primer on their measurement are provided in Box 1. Although other more serious mental illnesses, such as schizophrenia, are also correlated with poverty, and may have powerful effects on economic outcomes, we do not discuss them here (*21*).

Like mental health, poverty is multidimensional. We examine causal links between mental illness and important economic dimensions of poverty, particularly income and unemployment. We also touch on other dimensions

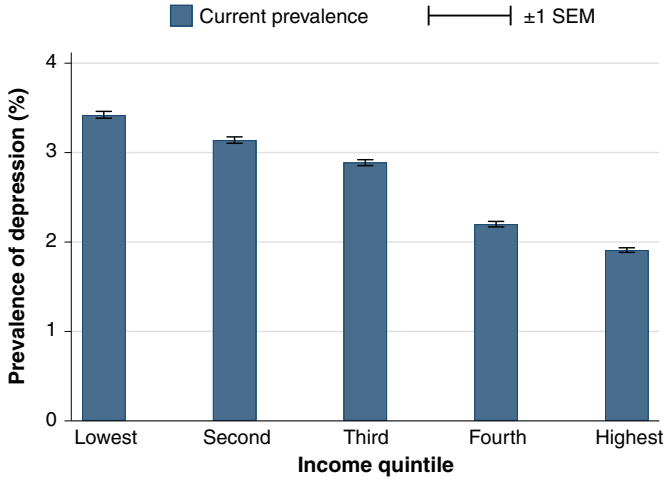


Fig. 1. Prevalence of depression by income quintile in India. The average percentage of people in each income quintile in India who have had depression within the past 2 weeks ("current" prevalence). Error bars show ±1 SEM. These numbers come from (*130*), an analysis of the Indian National Mental Health Survey, 2015–2016.

[1]Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Economics, Harvard University, Cambridge, MA 02138, USA. [3]Harvard Medical School, Harvard University, Boston, MA 02115, USA. [4]Harvard TH Chan School of Public Health, Harvard University, Boston, MA 02115, USA.
*Corresponding author. Email: fschilb@mit.edu

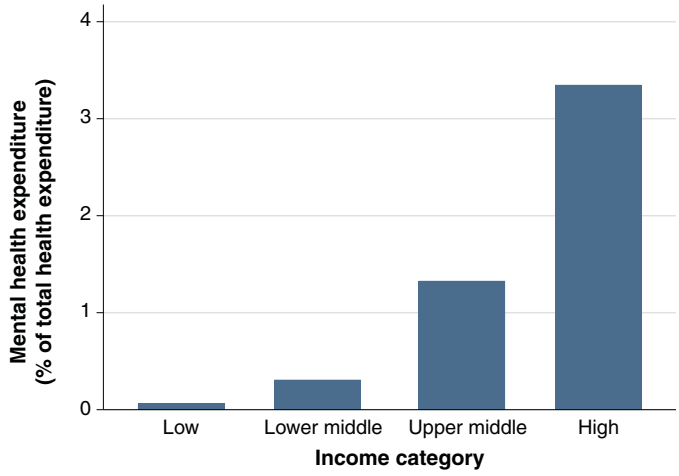


Fig. 2. Mental health expenditure by country income category. The average percentage of overall health budgets spent on mental health across countries in each of the four income categories used by the World Bank. Percent spent on mental health comes from the authors' own calculations, using data on overall mental health spending from the WHO Mental Health Atlas 2017 (www.who.int/mental_health/evidence/atlas), with data on total overall health spending from the WHO's Global Health Expenditure Database (https://apps.who.int/nha/database).

Box 1. Definition and measurement of depression and anxiety.

Depression, by which we refer here to major depressive disorder, is a constellation of symptoms that includes changes in psychomotor function, weight loss, oversleeping or undersleeping, decreased appetite, fatigue, difficulty concentrating, extreme feelings of guilt or worthlessness, and suicidal thoughts. According to the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-5), diagnosis of depression requires a set of these symptoms to be present over a 2-week period.

Anxiety, by which we refer here to generalized anxiety disorder, is characterized in the DSM-5 by long-lasting and excessive fear and worries over at least a 6-month period, with three or more of the following symptoms: restlessness, fatigue, concentration problems, irritability, muscle tension, and problems with sleep. Other definitions require the presence of at least one physical symptom such as heart palpitations, difficulty breathing, nausea or abdominal distress, dizziness, and numbness.

Measuring depression and anxiety in large population samples is feasible by using nonspecialist surveyors or even through self-administration of questionnaires. Reliable short-form diagnostic tools can predict professional diagnosis with rates of false positives and false negatives of 10 to 20% and have been validated in low-income countries (*151*–*154*). Widely used tools include the Generalized Anxiety Disorder 7-item (GAD-7) scale for anxiety, the Patient Health Questionnaire (PHQ-9) for depression, or the Self-Reporting Questionnaire 20-Item (SRQ-20) scale for any common mental illness. These scales typically ask respondents how much they experienced symptoms of depression or anxiety (such as sadness, lack of concentration, or poor sleep) in the past few weeks. The PHQ-9 and GAD-7 ask one question for each of the symptoms that are used to define major depressive disorder and generalized anxiety disorder, respectively. In practice, depression and anxiety are correlated, as evidenced by the fact that they share some symptoms.

The Center for Epidemiologic Studies Depression Scale (CES-D) is a popular measure among studies of the effect of economic interventions or shocks on mental health. Several studies also use custom indices of psychological well-being, typically an average of a life satisfaction scale, a "stress index," and some measure of worry or anxiety. In practice, such indices often measure several of the same symptoms as the PHQ-9 and GAD-7.

Some national surveys already include short-form screening tools, such as the UK Longitudinal Household Panel Survey and the South Africa National Income Dynamics Study.

of poverty, including a lack of capabilities resulting from low education and physical health, as well as relative poverty and associated low social status. Because of a relative scarcity of studies, we focus less on the relationship between mental health and the consumption of goods and services, which is a more direct economic measure of poverty. The existing evidence of this relationship is contentious, but the cross-sectional correlation of mental illness with consumption appears to be weaker than that between mental health and income (*22*–*25*). Income is more volatile than consumption in the short run. The stronger correlation of mental health with income suggests that mental health may be more affected by short-run changes to economic status than long-run or permanent changes.

We discuss evidence on poverty-alleviation programs and mental health treatments obtained from randomized controlled trials (RCTs). These generate variation in individuals' poverty and mental health status, respectively, that is entirely by chance and therefore uncorrelated with all other shared risk factors. Such studies allow us to isolate evidence of causal relationships. We also discuss studies of "natural experiments" in which naturally occurring variation in economic circumstances or mental health is argued to be "as good as random." Examples range from financial windfalls, such as lottery wins—for which lottery winners may be thought of as a treatment group and lottery losers as a control group—to weather shocks that affect some farmers' incomes more than others'.

The causal impact of poverty on mental ill-health

Job loss and income declines are drivers of poverty and often precede episodes of mental illness (*26*, *27*). Evidence from natural experiments confirms that this relationship is causal. For example, reduced agricultural output and income because of extreme rainfall caused increased rates of depression and suicide in rural parts of Indonesia (Fig. 3 and Box 2) (*28*). Similarly, job losses because of plant closures in Austria were associated with higher subsequent antidepressant use and mental health–related hospitalization (*29*). Areas in the United States more exposed to trade liberalization with China saw reduced income and employment for some groups of workers and increased mortality through drug overdoses among those same groups (*30*). Whether job loss worsens mental health beyond the impacts of the associated loss of income is unclear, but both mechanisms are argued to play a role in the phenomenon of "deaths of despair" (*31*).Conversely, income or wealth increases can improve mental health. For example, Native American tribes that opened casinos have seen substantial rises in income and reductions in anxiety relative to those that did not (*32*). Some studies have shown that lottery winners enjoy better mental health compared with those who win less or play but do not win (*33*). However, when fully controlling for the number and frequency of lottery tickets bought, winnings have small or no impacts on mental health (*34*, *35*).

The most compelling causal evidence that poverty causes mental illness comes from

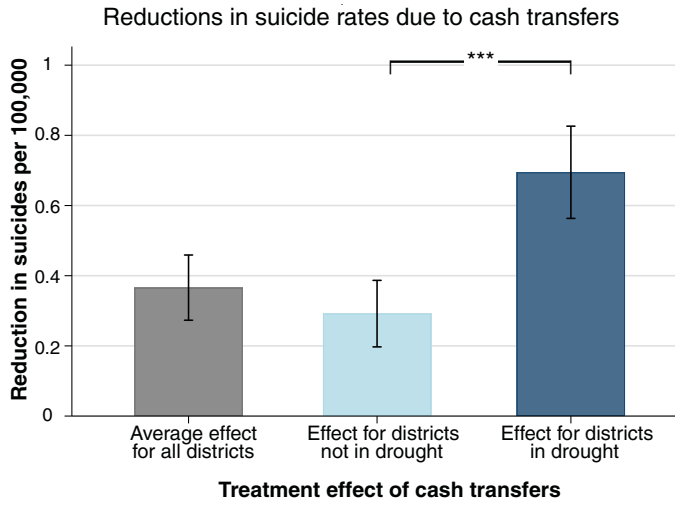


Fig. 3. Cash transfers, suicide rates, and droughts. The estimated effect of the cash transfer roll-out on district suicide rates, for all districts and separately by whether or not they were experiencing a drought (bottom 20% of the rainfall distribution) when the cash transfers reached them. Error bars show ±1 SEM. Asterisks denote a significant difference between effects: ***$P < 0.01$.

Box 2. Cash transfers, rainfall shocks, and suicides.

Christian, Hensel, and Roth examined how income shocks affect suicide rates and depression in Indonesia (*28*). They examined two natural experiments: the staggered roll-out across subdistricts of a conditional cash transfer program and annual and spatial variation in rainfall that affects farmers' incomes. They measured depression using a 10-question CES-D scale that is included in the Indonesian Family Life Survey. They also used the incidence of suicides, as measured by the reports of village leaders in census surveys.

Subdistricts that received the cash transfer program in the first wave of roll-out saw an 18% drop in suicides ($P < 0.01$) relative to those that received it later, even though both sets of districts had similar trends in suicide before the program's start. Meanwhile, rural subdistricts that experienced excess rainfall that increased crop yields between census years saw decreases in depression and suicides relative to subdistricts experiencing drought. The cash transfer had its largest effects on suicide in districts undergoing droughts, suggesting that policy can play a role in mitigating the mental health effects of economic shocks (Fig. 3).

Because suicide was only measured at the subdistrict level in this study, it is not possible to fully disentangle the direct effects of the cash transfers on recipients from spillover effects to others in the village. This highlights the need for better routine data collection on mental health outcomes alongside economic variables.

RCTs that evaluate antipoverty programs. Several studies that evaluated cash transfer and broader antipoverty programs have found substantial positive impacts on mental health, including over long time horizons, after the effects of any initial celebratory reactions among recipients have worn off. A meta-analysis of these studies is provided in Fig. 4. For example, cash transfers to Kenyan households worth $400 to $1500 at purchasing power parity (about 3 to 12 months of household income) increased consumption and happiness while reducing depression, stress, and worries (*36*, *37*). Scores on a depression scale were 0.12 standard deviations (SD; closely related to Cohen's *d*) lower 4 months after completion and 0.16 SD lower after almost 3 years, with larger transfers causing substantially larger effects.

Similarly, multifaceted antipoverty programs beyond cash transfers yield mental health benefits. A recent large-scale randomized evaluation of a "graduation program" in six countries that provided extremely poor participants with a mix of assets, intensive training, temporary cash support, savings incentives, and help to access health care found increases in consumption and assets 3 years later. The program also improved an index of psychological well-being by 0.1 SD, which was driven by an increase in happiness and a decrease in mental distress (*38*). Programs in other settings with similar approaches have found similar effects (*39*, *40*). Longer-run effects of such programs, when measured, appear to be even larger for both economic outcomes and mental health. In India, for example, an index of psychological well-being was 0.24 SD higher in the treatment group 7 years after the completion of a graduation program (*41*).

Hence, across a wide range of populations and study designs, positive economic shocks to individuals are shown to improve mental health, whereas negative economic shocks undermine mental health. This robust evidence, on the effects of changes in economic circumstances, indicates that poverty does cause mental illness. However, with two exceptions (*35*, *41*), the above studies consider the consequences of changes in economic status for a few years at most. An important question is whether these short-run effects can persist or grow over time. For example, some of the causal mechanisms we discuss below could take decades to manifest. However, there may be a hedonic adaptation effect in which mental health eventually adapts to the change in circumstances, so that even permanent increases in one's income level have a limited long-run effect. Of course, it could also be that the positive economic shocks are themselves undone by future negative shocks, causing mental health to revert to initial levels or even worsening. Ongoing long-run evaluations of cash-transfer programs are expected to provide evidence on this question (*42*).

Mechanisms for poverty causing mental ill-health

How does poverty cause mental illness? We discuss several plausible causal mechanisms and the limited evidence for each. The worries and uncertainty that come with living in poverty seem to be an important driver of mental illness, as do the effects of poverty on childhood development and one's living environment. We have more limited causal evidence for other plausible channels, including worsening of physical health caused by poverty, increased exposure to violence or crime, and the effects of low relative social status and social isolation. Understanding which of these mechanisms are important may have implications for policy choices. For example, if worries and uncertainty play a major role, then providing health and unemployment insurance may be crucial, whereas if early-life conditions are the key drivers, then cash transfers to parents of young children could be the most appropriate policy response.

Worries and uncertainty

The anticipation of economic shocks, not just their occurrence, may cause mental illness. People living in poverty face substantial

Study	Country	Outcome	Years elapsed since:		Intervention cost in:	
			Program Start	Program End	$ MER	$ PPP
Multifaceted antipoverty programs						
Blattman et al. (2019)	Ethiopia	PWB	5	4	450	1291
Edmonds and Theoharides (2020)	Philippines	PWB	3	0.75	184	518
Green et al. (2016)	Uganda	APAI–R	1.3	–	874	2150
Janzen et al. (2018)	Nepal	PWB	1.5	–	137	478
Banerjee et al. (2015)	Multiple	PWB	3	1	1467	3717
Bandiera et al. (2017)	Bangladesh	PWB	4	2.5	302	1120
Ismayilova et al. (2018)	Burkina Faso	CES–D	2	1	100	279
Glass et al. (2017)	DRC	HSCL	1.5	–	40	66
Banerjee et al. (2016)	India	PWB	7	5.5	357	1257
Bedoya et al. (2019)	Afghanistan	PWB	2	1	1688	6198
Cash transfers						
Heath et al. (2020)	Mali	PSS	3	–	649	1468
Fernald and Hidrobo (2011)	Ecuador	CES–D	1.4	–	255	676
Hjelm et al. (2017a)	Zambia	PSS	3	–	396	816
Macours and Vakis (2009)	Nicaragua	CES–D	0.75	–	260	758
Kilburn et al. (2019)	South Africa	CES–D	3	–	540	870
Baird et al. (2019)	Malawi	GHQ–12	4.3	2.3	180	440
Blattman et al. (2017)	Liberia	APAI–R	1	0.8	341	716
Haushofer et al. (2020)	Kenya	PWB	1	1	150	338
Blattman, Fiala and Martinez (2019)	Uganda	PWB	9	9	382	1175
Shapiro (2019)	Kenya	PWB	0.75	0.5	122	314
Hjelm et al. (2017b)	Zambia	PSS	3	–	432	891
Egger et al. (2019)	Kenya	PWB	1.5	1.5	1000	1871
Baird et al. (2013)	Malawi	GHQ–12	2.3	0.3	180	440
Kilburn et al. (2016)	Kenya	CES–D	4	–	960	2370
Aguilar (2012)	Mexico	SWB	5	2	900	1651
Haushofer and Shapiro (2018)	Kenya	PWB	3.4	3	521	709
Haushofer, Mudida and Shapiro (2020)	Kenya	PWB	1	1	534	1184
Angeles et al. (2019)	Malawi	CES–D	2	–	156	517
Multifaceted antipoverty programs effect (average: 0.138 SD)						
Cash transfers effect (average: 0.067 SD)						
Overall effect (average: 0.094 SD)						

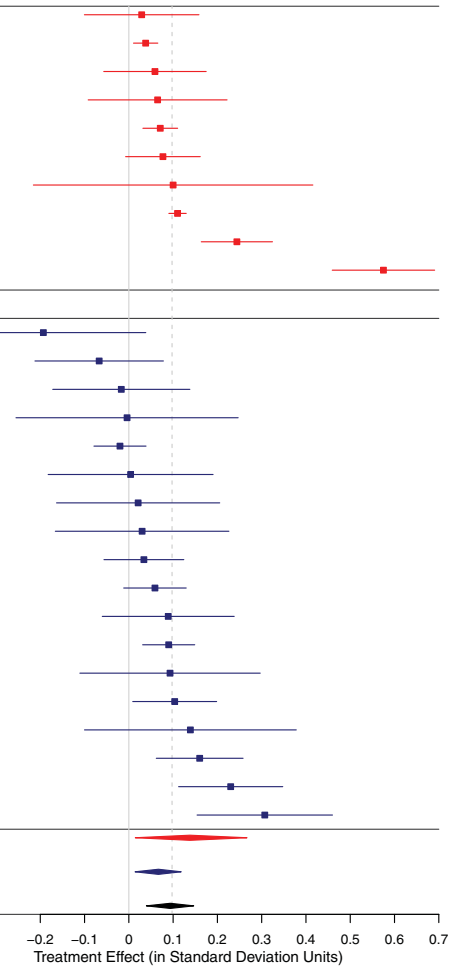


Fig. 4. The impacts of antipoverty programs on mental health. The estimated treatment effects of RCTs that evaluated antipoverty programs in low- or middle-income countries on indices of mental health (*37*–*41*, *56*, *119*, *131*–*149*). Positive treatment effects imply better mental health. "Cash Transfers" refers to studies of unconditional cash transfers to low-income households, with the exception of Baird *et al*. and Kilburn *et al*., who examined conditional cash transfer programs (*135*, *150*). "Multifaceted antipoverty programs" refers to interventions that aimed to lift people out of poverty by providing a range of elements, typically including asset transfers, skills training, cash support, and access to savings and health care opportunities. Treatment effects are expressed in SD units. If multiple follow-up measures were available, this figure shows the final measure. The outcomes vary across studies and include screening instruments to detect common mental illnesses [Short General Health Questionnaire (GHQ-12)] and symptoms of depression [Center for Epidemiologic Studies Depression Scale (CES-D) and Acholi Psychosocial Assessment (APAI-R)], indices of psychological well-being (PWB), and a perceived stress scale (PSS). "Intervention cost" refers to the total cost of implementing the intervention, when this is available. When implementation costs are unavailable (as with most of the cash transfer studies), it refers to the total undiscounted value of the transfer. MER, market exchange rates; PPP, purchasing power parity (which adjusts exchange rates to reflect the true cost of living). A missing value of years elapsed since program end means that the transfer was still ongoing when outcomes were measured. A complete description of the methodology of this analysis and details on each of the studies is provided in the supplementary materials.

uncertainty and income volatility and juggle what are, in effect, complex financial portfolios, often without access to formal insurance (*43*). Sustained long-run exposure to stress from managing this volatility may threaten mental health (*44*). Consistent with this hypothesis, a large-scale randomized experiment among low-income individuals in Oregon found that receiving largely free health insurance worth $550 to $750 per year reduced rates of depression by about a quarter within a few months (*45*). This effect did not appear to be explained by increased mental health care or changes in physical health. Although the increase in recipients' effective income may have played a role, it represented a much smaller relative increase than the cash transfer programs described above and yet generated a similar effect size on depression. Further suggestive evidence for uncertainty as a mechanism comes from the small or zero effect of wealth shocks on mental health in countries with generous and comprehensive systems of social insurance, such as Sweden (*34*, *35*).

Environmental factors

Those living in poverty are generally more exposed to environmental irritants posed by pollution, temperature extremes, and challenging sleep environments (*46*). Many of these factors have been linked directly to mental illness. Days with extreme heat see worse

self-reported mental health and increased rates of self-harm and suicide (*47*, *48*). Similarly, sleep deprivation is widespread among the urban poor in developing countries (*49*), and sleep is thought to be a mechanism affecting mental health (*50*, *51*). Some evidence exists that clinical interventions to improve sleep reduce depression (*49*, *52*). Exposure to air pollution is associated with living in poverty and may influence mental health through multiple channels, through restriction of physical activity or directly owing to neurotoxicity (*53*). Changes in air pollution in China have been associated with changes in mental health (*54*). In the United States, randomly selected low-income households that were paid to move to more affluent neighborhoods saw reductions in depression and anxiety despite little effect on income (*55*). However, it is not clear whether environmental factors or other features of high-income neighborhoods generated this effect.

Physical health

Lower income is robustly associated with worse physical health (*9*). Poverty increases exposure to the environmental factors described above and often also implies lower access to health care, which increases the burden of acute and chronic health conditions. Worse physical health may affect mental health through various channels. Chronic pain, worries about health and mortality, the financial costs of illness, and reduced physical activity may all worsen mental health. It is therefore unsurprising that physical ill-health often co-occurs with depressive and anxiety disorders (*15*). However, only limited causal evidence exists of poverty affecting mental health through changes in physical health. Many of the randomized interventions described in the previous section had no detectable effect on physical health even as they reduced mental illness (*36*, *38*, *56*). However, changes in physical well-being may manifest over a longer time frame, which may not be captured by these short-run studies.

Early-life conditions

Exposure to poverty early in life can threaten mental health in later years. Such effects can be generated in utero, by exposing pregnant women to malnutrition or stress. For example, the death of a mother's relative during pregnancy (compared with after childbirth) predicts depression and anxiety among her grown children later in life (*57*). Poverty may also disproportionately expose children to adverse shocks while their brains are highly plastic. Such shocks can profoundly affect brain development, cognitive ability, and mental health in adolescence and adulthood (*58*, *59*). Economic stresses around the time of birth also have long-term mental health costs; in Ghana, a decrease in crop prices by 1 SD at an individual's time of birth was found to increase incidence of anxiety or depression in adulthood by 50%, associating with maternal nutrition, breastfeeding duration, vaccination rates, and improved adult health (*60*). These results imply that programs that provide financial support for households with pregnant women or young children may have high long-run mental health and economic returns.

Trauma, violence, and crime

Living in poverty disproportionately exposes individuals to crime, including violent offenses (*61*). People living in poverty are also more likely to suffer traumatic events such as the early deaths of loved ones (*62*). Likewise, within the household, women and children in poor households are disproportionately affected by intimate partner violence (*63*). The relationship between poverty and experiencing violence itself may be causal: Cash transfers to households reduce intimate partner violence (*64*). In turn, both exposure to violence within the household and exposure to violent crime in general predict depression and other mental illnesses (*65*, *66*). Causal evidence on the effect of reductions in crime and violence on mental illness is needed to shed further light on this mechanism.

Social status, shame, and isolation

Relative poverty—consumption or income relative to others in one's society—may play a role in the relationship between poverty and mental illness through the resulting social status and interpersonal comparisons. In an interesting natural experiment, Norwegian tax records were posted online in 2001, making citizens' income easily searchable. Using survey data from 1985 to 2013, a study showed that the gap in happiness and life-satisfaction between the rich and poor within Norway increased sharply once relative income became easily visible (*67*). Although similar causal evidence is lacking for mental illness, it is plausible that diminished social status resulting from poverty causes or exacerbates depression and anxiety. Frequent marginalization of people living in poverty may also result in social isolation and loneliness (*68*), which in turn are correlated with depression (*69*).

The causal impact of mental ill-health on poverty

Mental illness predicts worse labor market outcomes later in life. After a diagnosis of depression or anxiety, employment rates and incomes have been estimated to fall by as much as half, relative to the nondepressed or nonanxious (*70*, *71*). Beyond such comparisons, which may be driven in part by unknown factors such as physical health, there is little evidence from natural experiments to link depression or anxiety to incomes. A study showed that the approval of lithium for treatment of bipolar disorder reduced the earnings penalty associated with bipolar illness by a third in Denmark, from 38 to 26%, with larger effects in the lower half of the earnings distribution (*72*). Studying similar natural experiments for depression and anxiety would be valuable.

There is, however, a substantial body of experiments that show a causal effect of treating mental illness on employment. A meta-analysis that aggregated results across 31 RCTs in developing countries showed a positive average effect of various interventions to treat mental illness on labor supply (*73*). Among these interventions, pharmacological and psychological treatments had similar positive effects on labor supply (0.1 to 0.15 SD), and combining both types of treatments had even larger effects (0.34 SD). For example, a cheap and scalable cognitive behavioral therapy (CBT) administered in India reduced depression by 25 percentage points compared with that in the control group and in turn increased reported days of work by 2.3 days per month (Fig. 5 and Box 3). Although these studies do not directly show that treatment of mental illness reduces poverty rates, higher labor supply and earnings naturally reduce the likelihood of living in poverty. Whether treatment of mental illness has larger long-run effects on consumption per unit cost than the cash transfers described above is unknown (*73*).

Mechanisms for mental ill-health causing poverty

Cognitive function

Like any illness, depression and anxiety may have economic effects because they directly reduce individuals' ability to work. Unlike most physical conditions, however, depression and anxiety also directly affect the way people think. Poverty itself can influence cognitive function by capturing attention and taxing mental bandwidth (*74*, *75*). Mental illness may have similar effects, by capturing attention, causing excessive rumination, and distorting people's memories and beliefs about their abilities (*76*). Such cognitive impacts could alter a range of economic decisions and outcomes, from finding jobs to saving to education and by exacerbating "behavioral biases" that economists increasingly recognize as important (*77*). For example, depressed individuals might avoid making active choices and may stick with "default options," may have decreased sensitivity to incentives because of anhedonia, or may have difficulty choosing among several options. Understanding the importance of this mechanism relative to more "direct" economic effects through disability or health expenditures

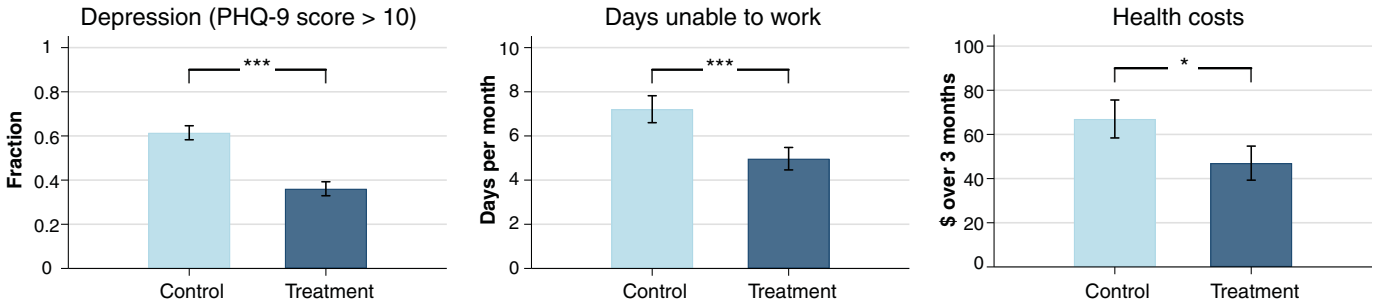


Fig. 5. Impacts of behavioral activation on depression, labor supply, and health costs. The mean values for the treatment and control groups of depression [measured with a Patient Health Questionnaire–9 (PHQ-9) score greater than 10], days unable to work, and health costs (excluding the intervention cost) at 3 months. Error bars denote ±1 SEM. Asterisks denote a significant difference between treatment and control after adjusting for covariates: *P < 0.1; ***P < 0.01.

Box 3. An example of a psychotherapy intervention with positive economic effects.

Patel *et al*. conducted an RCT of a brief behavioral activation (BA) therapy program, administered by nonspecialist counselors in a sample of 495 depressed adults in Goa, India (*155*). Compared with a control group that received enhanced usual care, treated patients were >60% more likely to be in remission 3 months later (64 versus 39%), as measured with a PHQ-9 score (Box 1) below 10, and maintained these gains after 12 months.

Those patients also reported being able to work 2.3 more days per month on average (P = 0.004) and reduced health costs, excluding intervention costs, of $20 per month ($P$ = 0.07) (Fig. 5). For comparison, a month's wages for a low-skilled worker in the study context was around $415. Given an average intervention cost of $66 per patient, in economic terms the intervention was highly cost-effective and may have paid for itself within a few months. After 12 months, the fall in treated patients' health costs alone had already significantly outpaced the cost of intervention, although the difference in days worked was no longer significant (*94*).

Other evaluations of inexpensive psychotherapies implemented by nonspecialist counselors in low-income settings have found similarly large effects on mental health (*122*). More evidence on the effect of such psychological interventions on broader economic outcomes would be highly valuable. For example, future trials could be linked to administrative or standardized survey data on wages, earnings, and consumption.

is crucial for correctly measuring the economic burden of mental illness and designing economic policy for those whose mental health is compromised.

Beliefs

Beliefs about one's own and others' abilities, circumstances, and actions are central to economic decision-making. Mental illness may distort such beliefs in various ways. Depression is associated with more negative beliefs about oneself and the external world (*78*, *79*). Depressed individuals are more likely to remember negative stimuli and have trouble disengaging from negative information once it grabs their attention (*76*). As such, although healthy individuals tend to protect overly optimistic beliefs about themselves by ignoring negative information (*80*), correlational evidence suggests that the depressed update their beliefs more pessimistically (*81*). Anxiety, meanwhile, is associated with greater selective attention toward threatening stimuli (*82*), which could lead to overestimation of risks and thus reduced risk-taking. Such evidence is consistent with mental illness causing pessimistic beliefs, pessimistic beliefs causing mental illness, or both. Causal evidence on how treatment of depression or anxiety affects beliefs would help disentangle these potential explanations.

Preferences

Mental illness may affect economic preferences, such as the extent to which people are willing to defer gratification (time preferences), tolerate risk for higher expected rewards (risk preferences), or split rewards between themselves and others (social preferences). For example, depression may diminish a person's patience and altruism. Similarly, anxiety disorders may reduce people's willingness to take on even modest levels of potentially profitable risk. Such impacts could in turn change a variety of economic behaviors, such as labor supply decisions, savings and investment choices, consumption behavior, and the take-up of social programs. The limited evidence on correlations between mental illness and economic preferences is mixed (*83*–*85*).

Labor supply and productivity

Depression and anxiety often affect individuals in the prime of their economic lives and can be highly recurrent (*86*). The depressed beliefs and distorted preferences described above may reduce motivation and labor supply. In addition, depression can have a direct effect on productivity, such as through increased fatigue and lack of concentration. Depressed individuals may therefore work fewer and shorter days and produce less per hour (*87*). Depressed workers might also be more easily discouraged during their job search or when facing setbacks at work. As described above, substantial causal evidence exists that treatment of mental illnesses increases employment (*73*). However, there is little evidence on whether this happens through higher at-work productivity, greater job search intensity, changed beliefs, or other mechanisms.

Stigma

Mentally ill individuals contend with substantial social stigma and negative stereotyping (*88*). This may result in discrimination in employment (*89*), which could lower wages and limit employment opportunities relative to equally productive mentally healthy workers. On top of this, those living with a mental illness are excluded from disability benefit schemes in many low-income countries (*90*). More generally, others' reluctance to interact socially with mentally ill people (*88*) may exclude them from social networks that provide economic opportunities. Stigma may also affect the formation and dissolution of households in ways that disadvantage the mentally ill (*91*). Depression and anxiety may come with a "discount" on the marriage market, causing mentally ill individuals to form households with less well-off partners, increasing the chances of living in poverty.

Health expenditures

Mental illnesses may deepen poverty through its impacts on health and health expenditures. In developing countries, people living in poverty usually pay most of their health costs out of pocket (*92*). Globally, 150 million people are estimated to have catastrophic health expenditures each year, which are defined as health care payments totaling more than 40% of a household's nonsubsistence expenditures (*93*). Costs associated with treatment of mental illness rarely account for large shares of individuals' budgets because most affected individuals remain untreated. However, depression and anxiety frequently co-occur with other health conditions (*15*), and such comorbidity with depression is associated with substantially higher health expenditures for a range of health conditions (*16*, *17*). Indeed, treatment of depression has been found to reduce overall health care costs (*94*).

Women's empowerment

The burden of mental illness falls disproportionately on women (*1*). A large-scale (n = 903 participants) RCT that evaluated CBT intervention for depressed pregnant women in Pakistan found a 17% reduction in depression rates compared with a control group 7 years after the intervention (*95*). Reduced depression among these women was accompanied by increased economic empowerment by 0.29 SD as measured in increased control over household and personal expenditures. Such impacts may have implications for women's consumption and relative poverty within the household.

Intergenerational effects

Improving a parent's mental health can benefit the next generation. In the above study in Pakistan, women who had received the intervention sent their children to better schools and had more learning materials in their homes (*95*). Similarly, other RCTs found that treatment of mothers' depression improves their interaction with their children and their children's mental health (*96*). Although little direct evidence shows that such interventions lead to improved educational outcomes or earnings, there is reason to believe they may. A substantial body of work from other contexts shows that early-childhood investments have large effects on children's income as adults (*97*).

Human capital accumulation

The onset of common mental illnesses often coincides with secondary and tertiary education and the early stages of an adult's work career (*86*). Mental illness may therefore cause long-run economic hardship by reducing school and college completion rates, worsening early-career job placements and hindering skill acquisition (*98*). This suggests the possibility of particularly high economic returns from improving mental health among adolescents and young adults. Although longitudinal studies show a substantial correlation between mental illness among students and subsequent educational outcomes, there is little experimental evidence to date that treatment of depression or anxiety among adolescents leads to improved educational outcomes (*99*).

Outlook

Having discussed some of the mechanisms that influence the relationship between poverty and mental health, we can speculate more broadly on how the relationship between poverty and mental illness may evolve, what this means for policy, and what directions may be fruitful for research.

Aggregate economic conditions

Economic growth and other ongoing global trends are unlikely to improve mental health by themselves. Higher income causes better mental health at the individual level, yet on average, the prevalence of mental illness is not lower in rich countries. To the contrary, existing evidence shows a higher prevalence of common mental illness in richer countries (Fig. 6) (*100*). This cross-country difference cannot be interpreted causally, and concerns remain about differences in methodology, diagnosis, or reporting across contexts (*101*). However, one way to reconcile the contrasting within-country and cross-country relationships is the possibility that relative, rather than absolute, poverty is the more relevant cause of mental illness. Risk factors, including inequality and relative poverty, or the stresses of urban environments, may deteriorate rather than improve as whole economies expand. Within-country inequality has increased in many countries in recent decades, despite substantial reductions in extreme poverty and global inequality (*102*). Complacency about mental health among the poor is therefore not warranted even in the presence of aggregate economic growth. For example, the burden of disease attributable to mental and neurological disorders in India and China increased by 61 and 28%, respectively, between 1990 and 2013, despite impressive economic growth (*103*).

Climate change

The more frequent occurrence of extreme heat because of climate change is anticipated to exacerbate mental illness directly (*104*, *105*). The increased frequency of weather-related disasters, such as floods and hurricanes, poses a threat to mental health through greater exposure to trauma (*106*). Climate change also threatens mental health through its negative economic consequences, which are likely to be more pronounced in low-income countries (*104*). Extreme temperatures during the agricultural growing season that damage crops and thus economic well-being have been reported to increase suicides in agricultural regions in India (*107*). Predicted increases in water scarcity and droughts are also likely to worsen economic and in turn psychological well-being. Climate change is also expected to lead to increased violence and political conflict over the next century through increased pressure on resources, such as productive land and, possibly, psychological effects of heat on aggression (*108*). This combination of economic and political consequences of climate change may increase the flow of refugees and economic migrants, with concomitant challenges to mental health (*109*).

Technological change and globalization

For many of those living in poverty across the world, technological change and globalization offer enormous economic opportunities; however, both phenomena produce winners and losers. The costs to losers, especially low-wage workers in high- and middle-income countries who lose jobs as a result of changes in patterns of trade or automation, can be long lasting and substantial (*110*), resulting in worse mental health (*29*, *30*). Although most economic research on these topics focuses on rich countries, there is an urgent need to understand the mental-health effects of these economic changes in poorer countries.

Offering social insurance and welfare, skills training, and job transition programs, including psychotherapies for workers exposed to the harmful effects of technological change and globalization, will be important to protect mental health. One example is the ongoing Building Bridges and Bonds evaluation, which provides a tailored CBT intervention to unemployed (formerly incarcerated) men in the United States, in conjunction with traditional job-seeking services, in an attempt to increase employment and wages.

Pandemics

Public-health crises such as the ongoing COVID-19 pandemic tend to disproportionately affect those living in poverty (*111*). They may worsen mental health on average and particularly among the poor. Income and employment losses as a result of morbidity can be large, which in turn can reduce mental health through the mechanisms described above. In addition, the exposure to trauma, increased worries and uncertainty, and worse physical health will tend to impair mental health, in turn reducing income and employment. However, ongoing progress in reducing the burden of other diseases that disproportionately affect the poor—such as cholera and diarrhoeal infections, tuberculosis, malaria, and other

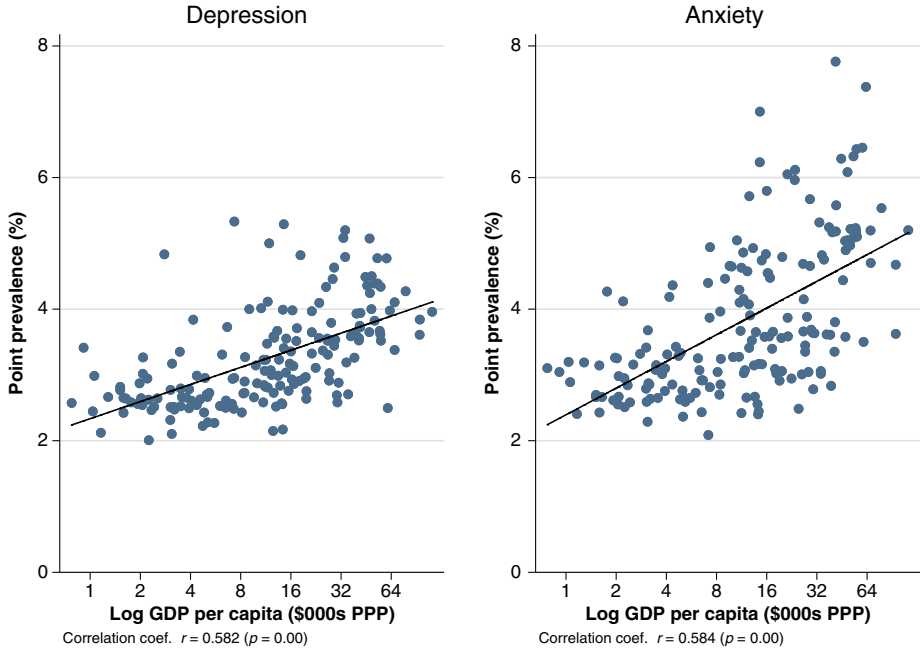


Fig. 6. Prevalence of common mental illnesses by country. For each country, the percentage of the population estimated to have a (left) depressive disorder or (right) anxiety disorder is shown at a given point in time against that country's log gross domestic product (GDP) per capita. Each scatter point represents one country. The line shown is an ordinary least squares regression line of country prevalence rates on a constant and log GDP per capita. Prevalence rate data come from the Global Burden of Disease Study, 2017 (http://ghdx.healthdata.org/gbd-results-tool). GDP per capita data are for 2017, measured in constant 2017 international $, and come from the World Bank's World Development Indicators dataset (https://databank.worldbank.org/source/world-development-indicators).

insect-borne diseases—will provide a countervailing force likely to improve mental health among the poor. The COVID-19 pandemic has the potential to set back progress in reducing the enormous burden of these diseases.

Social media

The spread of mobile phones and the internet opens up opportunities for poverty alleviation (*112*) and new ways to deliver mental health care. However, some of these technologies may pose new threats to mental health. Although more causal evidence is needed, some studies have found that depression is correlated with internet addiction and with the intensity of use of social media among young adults and adolescents (*113*, *114*). Recent randomized interventions show that deactivating social media accounts for 4 weeks led to 0.1 SD reductions in depression and anxiety scores (*115*, *116*). Increased access to the internet and social media among the global poor, although presenting enormously valuable benefits, may therefore also pose some threat to mental health among the poor (*117*).

Implications for research and policy

Because mental health and poverty are intimately linked, interdisciplinary collaborations between mental-health researchers and social scientists studying poverty are essential (Box 4). Evaluations of economic interventions should carefully measure impacts on mental health, and vice versa. An example of such work is Bhat *et al.*, in which a team of psychiatrists and economists followed up on psychotherapy clinical trials in India and deployed the tools of behavioral economics and psychiatry to study long-run effects of psychotherapy on mental health, economic well-being, and decision-making (*118*).

Policy tools

There is a strong economic case for investing in the mental health of people in poverty. A recent meta-analysis showed that mental health interventions in low- and middle-income countries, including psychotherapy and pharmacotherapy, had similar or larger effects on employment than economic interventions such as cash transfers, job training programs, and multifaceted antipoverty programs (*73*). Yet, they were an order of magnitude less expensive to deliver. Such treatments could therefore be the most cost-effective antipoverty intervention, at least among the subset of people who are mentally ill. However, we know little about how to optimally combine, dose, sequence, and target economic and mental health interventions. For example, combining psychological and economic interventions may lead to treatment effects that are greater than the sum of their parts. In particular, improved mental health could increase the economic returns of receiving cash or asset transfers by improving decision-making and productivity.

Recently, innovative studies have compared the effects of providing psychotherapy, cash support, or both among low-income populations. An RCT measured the effect of 8 weeks of CBT and/or $200 in cash support to 999 criminally engaged men in Liberia (*119*). Although the psychotherapy targeted antisocial behavior rather than mental illness per se, the study found that the combination of cash transfer and psychotherapy improved an index of self-regard and mental health by 0.2 SD a year later (P = 0.024), accompanied by a modest reduction in depression and psychological distress (–0.11 SD, P = 0.24). The combined treatment not only reduced antisocial behavior but also increased patience and self-control, whereas neither cash nor therapy alone had detectable effects. However, none of the treatments apparently influenced consumption or income a year later. More studies along these lines would be valuable.

Treatment gaps

In poor countries, the fraction of individuals diagnosed with depression and anxiety who do not receive treatment often exceeds 90% (*11*–*14*). Such treatment gaps likely result from a combination of poor supply and low demand for mental health services.

Increasing supply

Resources for mental health care are extremely limited in low-income countries (Fig. 7), and

Box 4. Priorities for future research on poverty and common mental illnesses.

1. Measurement of mental health in economic surveys to estimate

The comparative impacts of diverse economic interventions, such as insurance, social safety, and employment support, relative to cash transfers
The longer-run effects of antipoverty programs
The effects of absolute versus relative poverty
The effect of technological change and globalization on high and low-wage workers
The impact of the looming global economic recession resulting from COVID-19

2. Measurement of economic outcomes in intervention studies for depression and anxiety, which includes

Income, labor supply, productivity, and profits from self-employment
Economic preferences and beliefs; investment and savings behaviors
Household expenditures and consumption, including within-household allocation of resources

3. Evaluations of interventions to reduce stigma and to boost demand for mental health care, which includes

Diverse mental health literacy approaches, from mass-media campaigns to grass-root awareness building
Subsidies and incentives for seeking and engaging with mental health care
The effects of interventions on marginalized and underserved communities

4. Evaluation of technologies to increase the coverage of effective psychotherapies, which includes

Text, phone, or video delivery
Digital approaches to training, supervision, and quality assurance for frontline providers
Artifical intelligence bot-based and other guided self-help approaches, adapted to different languages and cultural contexts

5. Evaluation of interventions to interrupt the intergenerational transmission of poverty and mental illness, such as through

School mental health–promotion programs and measurement of effects also on educational attainment, labor supply, productivity, and earnings
Treatment of parental mental illness and measurement of effects on children's cognitive and educational outcomes

people living in poverty often lack access to basic mental health care (*120*). For example, some estimates suggest that India has only 3900 psychiatrists for more than 1.3 billion people, and 13 psychiatrists serve Zimbabwe's 14 million people (*121*). However, cost-effective and scalable strategies for treating mental illness in low-resource settings do exist. Evidence from multiple countries shows that "psychosocial" treatments such as manualized talk therapies can be highly effective at low cost, even when delivered by nonspecialist community health workers (*122*, *123*). An even more pared-down—but still effective—approach is the "friendship benches" of Zimbabwe, in which nonspecialist health workers (popularly known as "community grandmothers") provide problem-solving therapy with components of behavior activation to patients (*124*).

Reliance on in-person training and supervision of the community health workers by experts is a major barrier to scaling these approaches. The Empower initiative seeks to address these structural barriers through the deployment of digital platforms to enable frontline workers to learn, deliver, and master evidence-based brief psychological treatments for mental health problems, using innovative practices such as coach-supported learning and peer-supervised quality assurance, all delivered through digital tools (*125*). Empower will roll out the brief behavioral activation treatment developed and evaluated in India and Nepal (Box 3) to community health workers in India and the United States from 2021 onward.

Stimulating demand

Even in settings with affordable and effective mental health services, many people do not seek or adhere to treatment (*126*). People often lack mental health literacy and may not possess basic information about mental health conditions and their risk factors, symptoms, and potential treatment options (*127*). Stigma and shame can further depress demand for mental health services. But there are examples of successful community-based programs that increase mental-health literacy and boost the share of mentally ill individuals who seek treatment; for example, the VISHRAM (Vidarbha Stress and Health Programme) initiative, rolled out in 30 villages in rural India, deployed "change agents"—typically, persons already playing leadership roles in the villages—to engage the community in conversations around mental health and illness, using a range of contextually appropriate strategies such as group discussions; the program led to a sixfold increase in help-seeking by persons with depression in just 18 months (*128*). A priority for future work should be the evaluation of such programs at scale, as well as testing approaches that bundle mental health treatments with other unstigmatized services, subsidizing or even rewarding the take-up of treatment, or using remote technologies such as app-based therapy that are less likely to expose individuals to stigma.

Poverty traps

A classic idea in development economics is the existence of a "poverty trap" (*129*). This is the idea that below a certain threshold level of wealth, people are not able to earn and save their way out of poverty. They remain trapped in poverty precisely because of the deprivations associated with poverty and not because of any intrinsic lack of ability. Such poverty traps could exist for many reasons. For example, the very poor may not be able to afford enough food to be productive at work. Recent evaluations of multifaceted antipoverty programs have provided some evidence for the existence of such poverty traps (*39*, *41*). However, the underlying mechanisms are unclear. The evidence of a bidirectional causal relationship between poverty and mental health presented in this Review suggests that mental health could be a key mechanism: There could be psychological poverty traps. Some of those living in poverty may be ensnared in a vicious cycle of poverty and mental illness (*5*). If this is the case, a one-time economic or psychological intervention of sufficient magnitude may "push" people into a state of sustainably higher income and better mental health.

Conclusion

The causal relationship between poverty and mental health we have described could not be more pertinent than in the ongoing pandemic, which has already adversely affected both of these outcomes. Given the surge of deaths

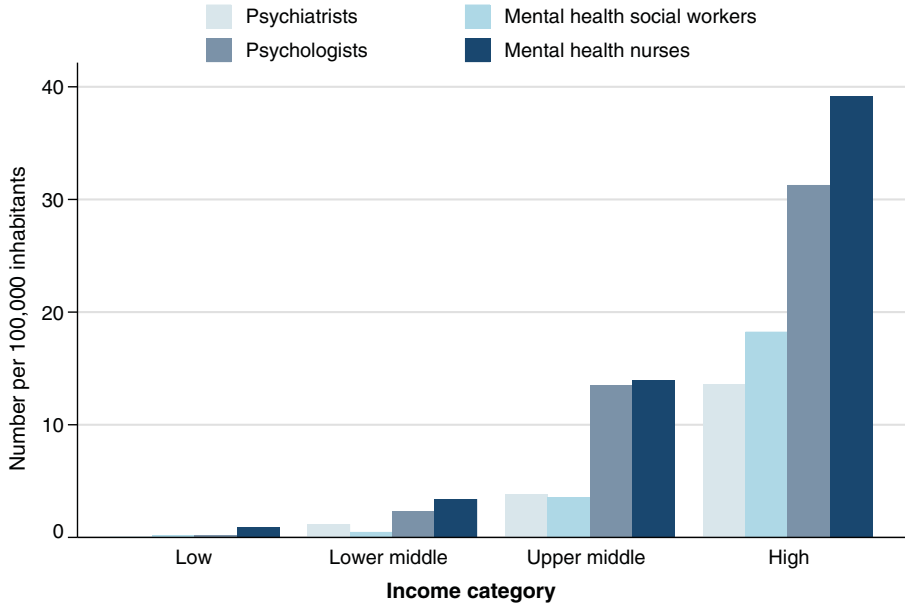


Fig. 7. Availability of mental health workers across countries. The mean numbers of psychiatrists, social workers, psychologists, and nurses working in the mental health sector (per 100,000 people) are shown for countries in each of the four income categories used by the World Bank. Data on mental health workers come from the WHO's Global Health Observatory (http://apps.who.int/gho/data/node.main.MHHR?lang=en) and are for the most recent year available (which ranges from 2013 to 2017).

of despair in the United States in the wake of the Great Recession of 2008 (*31*), we have grave concerns about the mental-health implications of the economic recession confronting the world. The pandemic has disproportionately affected the poor and may have lasting adverse impacts on their economic and mental well-being. A massive investment in mental health was long overdue even before the pandemic and has become critically urgent now. Beyond more money, this is also an important opportunity to invest wisely in lower-cost innovations that provide quality care to low-income and disadvantaged communities and to integrate economic interventions with mental health care to reduce historic disparities in both wealth and mental health.

REFERENCES AND NOTES

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018). doi: 10.1016/S0140-6736(18)32279-7; pmid: 30496104
2. R. Desjarlais, *World Mental Health: Problems and Priorities in Low-income Countries* (Oxford Univ. Press, 1995).
3. C. J. L. Murray, A. D. Lopez, Others, "The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary" [World Health Organization (WHO), 1996]; https://apps.who.int/iris/handle/10665/41864.
4. J. Sareen, T. O. Afifi, K. A. McMillan, G. J. G. Asmundson, Relationship between household income and mental disorders: Findings from a population-based longitudinal study. *Arch. Gen. Psychiatry* **68**, 419–427 (2011). doi: 10.1001/archgenpsychiatry.2011.15; pmid: 21464366
5. C. Lund *et al*., Poverty and common mental disorders in low and middle income countries: A systematic review. *Soc. Sci. Med.* **71**, 517–528 (2010). doi: 10.1016/j.socscimed.2010.04.027; pmid: 20621748
6. V. Iemmi *et al*., Suicide and poverty in low-income and middle-income countries: A systematic review. *Lancet Psychiatry* **3**, 774–783 (2016). doi: 10.1016/S2215-0366(16)30066-9; pmid: 27475770
7. L. M. Banks, H. Kuper, S. Polack, Poverty and disability in low- and middle-income countries: A systematic review. *PLOS ONE* **12**, e0189996 (2017). doi: 10.1371/journal.pone.0189996; pmid: 29267388
8. D. Frasquilho *et al*., Mental health outcomes in times of economic recession: A systematic literature review. *BMC Public Health* **16**, 115 (2016). doi: 10.1186/s12889-016-2720-y; pmid: 26847554
9. D. M. Cutler, A. Lleras-Muney, T. Vogl, in *The Oxford Handbook of Health Economics*, S. Glied, P. Smith, Eds. (Oxford Univ. Press, 2011), pp. 124–163.
10. H. Bleakley, Health, human capital, and development. *Annu. Rev. Econom.* **2**, 283–310 (2010). doi: 10.1146/annurev.economics.102308.124436; pmid: 24147187
11. D. Chisholm *et al*., Scaling-up treatment of depression and anxiety: A global return on investment analysis. *Lancet Psychiatry* **3**, 415–424 (2016). doi: 10.1016/S2215-0366(16)30024-4; pmid: 27083119
12. V. Patel, G. Simon, N. Chowdhary, S. Kaaya, R. Araya, Packages of care for depression in low- and middle-income countries. *PLOS Med.* **6**, e1000159 (2009). doi: 10.1371/journal.pmed.1000159; pmid: 19806179
13. P. S. Wang *et al*., Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet* **370**, 841–850 (2007). doi: 10.1016/S0140-6736(07)61414-7; pmid: 17826169
14. R. Kohn, S. Saxena, I. Levav, B. Saraceno, The treatment gap in mental health care. *Bull. World Health Organ.* **82**, 858–866 (2004). pmid: 15640922
15. K. M. Scott *et al*., Association of mental disorders with subsequent chronic physical conditions: World mental health surveys from 17 countries. *JAMA Psychiatry* **73**, 150–158 (2016). doi: 10.1001/jamapsychiatry.2015.2688; pmid: 26719969
16. E. P. Vamos, I. Mucsi, A. Keszei, M. S. Kopp, M. Novak, Comorbid depression is associated with increased healthcare utilization and lost productivity in persons with diabetes: A large nationally representative Hungarian population survey. *Psychosom. Med.* **71**, 501–507 (2009). doi: 10.1097/PSY.0b013e3181a5a7ad; pmid: 19528291
17. C. Shen, U. Sambamoorthi, G. Rust, Co-occurring mental illness and health care utilization and expenditures in adults with obesity and chronic physical illness. *Dis. Manag.* **11**, 153–160 (2008). doi: 10.1089/dis.2007.0012; pmid: 18564027
18. WHO, "Promoting mental health: concepts, emerging evidence, practice: a report of the World Health Organization, Department of Mental Health and Substance Abuse" (WHO/Victorian Health Promotion Foundation/Univ. Melbourne, 2005); https://apps.who.int/iris/bitstream/handle/10665/43286/9241562943-tha.pdf.
19. J. Haushofer, E. Fehr, On the psychology of poverty. *Science* **344**, 862–867 (2014). doi: 10.1126/science.1232491; pmid: 24855262
20. R. Layard, J. F. Helliwell, J. D. Sachs, "World Happiness Report 2012" (Earth Institute, Columbia University, 2012).
21. C. Lund, S. Stansfeld, M. De Silva, Social determinants of mental health, in *Global Mental Health: Principles and Practice* (Oxford Univ. Press, 2014), pp. 116–136.
22. G. Tampubolon, W. Hanandita, Poverty and mental health in Indonesia. *Soc. Sci. Med.* **106**, 20–27 (2014). doi: 10.1016/j.socscimed.2014.01.012; pmid: 24524962
23. J. Das, Q.-T. Do, J. Friedman, D. McKenzie, Mental health patterns and consequences: Results from survey data in five developing countries. *World Bank Econ. Rev.* **23**, 31–55 (2009). doi: 10.1093/wber/lhn010
24. J. Corrigall, C. Lund, V. Patel, S. Plagerson, M. K. Funk, Poverty and mental illness: fact or fiction? A commentary on Das, Do, Friedman, McKenzie & Scott (65:3, 2007, 467–480). *Soc. Sci. Med.* **66**, 2061–2063, discussion 2064–2066 (2008). doi: 10.1016/j.socscimed.2008.01.005; pmid: 18325649
25. J. Das, Q.-T. Do, J. Friedman, D. McKenzie, K. Scott, Revisiting the relationship between mental health and poverty in developing countries: A response to Corrigall. *Soc. Sci. Med.* **66**, 2064–2066 (2008). doi: 10.1016/j.socscimed.2008.01.004
26. S. C. Olesen, P. Butterworth, L. S. Leach, M. Kelaher, J. Pirkis, Mental health affects future employment as job loss affects mental health: Findings from a longitudinal population study. *BMC Psychiatry* **13**, 144 (2013). doi: 10.1186/1471-244X-13-144; pmid: 23705753
27. M. Alloush, Income, psychological well-being, and the dynamics of poverty: Evidence from South Africa (2018); https://ageconsearch.umn.edu/record/274223/files/Abstracts_18_05_23_17_38_40_27__23_123_4_199_0.pdf.
28. C. Christian, L. Hensel, C. Roth, Income shocks and suicides: Causal evidence from Indonesia. *Rev. Econ. Stat.* **101**, 1–16 (2019).
29. A. Kuhn, R. Lalive, J. Zweimüller, The public health costs of job loss. *J. Health Econ.* **28**, 1099–1115 (2009). doi: 10.1016/j.jhealeco.2009.09.004; pmid: 19833399
30. J. R. Pierce, P. K. Schott, Trade liberalization and mortality: Evidence from U.S. counties. *Am. Econ. Rev. Insights* **2**, 47–63 (2020). doi: 10.1257/aeri.20180396
31. A. Case, A. Deaton, *Deaths of Despair and the Future of Capitalism* (Princeton Univ. Press, 2020).
32. B. Wolfe, J. Jakubowski, R. Haveman, M. Courey, The income and health effects of tribal casino gaming on American Indians. *Demography* **49**, 499–524 (2012). doi: 10.1007/s13524-012-0098-8; pmid: 22427279
33. B. Apouey, A. E. Clark, Winning big but feeling no better? The effect of lottery prizes on physical and mental health. *Health Econ.* **24**, 516–538 (2015). doi: 10.1002/hec.3035; pmid: 24677260
34. D. Cesarini, E. Lindqvist, R. Östling, B. Wallace, Wealth, health, and child development: Evidence from administrative data on Swedish lottery players. *Q. J. Econ.* **131**, 687–738 (2016). doi: 10.1093/qje/qjw001
35. E. Lindqvist, R. Östling, D. Cesarini, Long-run effects of lottery wealth on psychological well-being. *Rev. Econ. Stud.* **87**, 2703–2726 (2020). doi: 10.1093/restud/rdaa006
36. J. Haushofer, J. Shapiro, The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *Q. J. Econ.* **131**, 1973–2042 (2016). doi: 10.1093/qje/qjw025; pmid: 33087990
37. J. Haushofer, J. Shapiro, "The long-term impact of unconditional cash transfers: experimental evidence from Kenya" (Busara Center for Behavioral Economics, 2018); https://pdfs.semanticscholar.org/2454/746163854dd855ac0d93f13f39556232bdf7.pdf.

38. A. Banerjee *et al*., Development economics. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* **348**, 1260799 (2015). doi: 10.1126/science.1260799; pmid: 25977558
39. O. Bandiera *et al*., Labor markets and poverty in village economies. *Q. J. Econ.* **132**, 811–870 (2017). doi: 10.1093/qje/qjx003
40. G. Bedoya, A. Coville, J. Haushofer, M. R. Isaqzadeh, J. Shapiro, *No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation*, policy research working papers (The World Bank, 2019).
41. A. Banerjee, E. Duflo, R. Chattopadhyay, J. Shapiro, The long term impacts of a "Graduation" program: Evidence from West Bengal. Unpublished paper, Massachusetts Institute of Technology, Cambridge, MA (2016); https://trickleup.org/wp-content/uploads/2016/12/2016_Banerjee_Long_Term_Impacts_Bandhan_Full.pdf.
42. A. Banerjee, P. Niehaus, T. Suri, Universal basic income in the developing world. *Annu. Rev. Econ.* **11**, 959–983 (2019). doi: 10.1146/annurev-economics-080218-030229
43. D. Collins, J. Morduch, S. Rutherford, O. Ruthven, *Portfolios of the Poor: How the World's Poor Live on $2 a Day* (Princeton Univ. Press, 2009).
44. S. M. Staufenbiel, B. W. J. H. Penninx, A. T. Spijker, B. M. Elzinga, E. F. C. van Rossum, Hair cortisol, stress exposure, and mental health in humans: A systematic review. *Psychoneuroendocrinology* **38**, 1220–1235 (2013). doi: 10.1016/j.psyneuen.2012.11.015; pmid: 23253896
45. A. Finkelstein *et al*., The Oregon health insurance experiment: Evidence from the first year. *Q. J. Econ.* **127**, 1057–1106 (2012). doi: 10.1093/qje/qjs020; pmid: 23293397
46. E. B. Dean, F. Schilbach, H. Schofield, in *The Economics of Poverty Traps*, C. B. Barrett, M. R. Carter, J.-P. Chavas, Eds. (Univ. Chicago Press, 2018), pp. 57–118.
47. N. Ding, H. Berry, L. O'Brien, The effect of extreme heat on mental health—Evidence from Australia. *Int. J. Epidemiol.* **44** (suppl. 1), i64–i64 (2015). doi: 10.1093/ije/dyv097.238
48. M. N. Williams, S. R. Hill, J. Spicer, Do hotter temperatures increase the incidence of self-harm hospitalisations? *Psychol. Health Med.* **21**, 226–235 (2016). doi: 10.1080/13548506.2015.1028945; pmid: 25849087
49. P. Bessone, G. Rao, F. Schilbach, H. Schofield, M. Toma, The Economic Consequences of Increasing Sleep Among the Urban Poor, National Bureau of Economic Research (NBER) working paper no. 26746 (2020); www.nber.org/papers/w26746.
50. A. G. Harvey, Sleep and circadian functioning: Critical mechanisms in the mood disorders? *Annu. Rev. Clin. Psychol.* **7**, 297–319 (2011). doi: 10.1146/annurev-clinpsy-032210-104550; pmid: 21166537
51. S. Grubb, M. Lauritzen, Deep sleep drives brain fluid oscillations. *Science* **366**, 572–573 (2019). doi: 10.1126/science.aaz5191; pmid: 31672882
52. R. Manber *et al*., Cognitive behavioral therapy for insomnia enhances depression outcome in patients with comorbid major depressive disorder and insomnia. *Sleep* **31**, 489–495 (2008). doi: 10.1093/sleep/31.4.489; pmid: 18457236
53. Z. Jia *et al*., Exposure to ambient air particles increases the risk of mental disorder: Findings from a natural experiment in Beijing. *Int. J. Environ. Res. Public Health* **15**, 160 (2018). doi: 10.3390/ijerph15010160; pmid: 29351245
54. T. Xue, T. Zhu, Y. Zheng, Q. Zhang, Author correction: Declines in mental health associated with air pollution and temperature variability in China. *Nat. Commun.* **10**, 3609 (2019). doi: 10.1038/s41467-019-11660-5; pmid: 31383856
55. J. Ludwig *et al*., Neighborhood effects on the long-term well-being of low-income adults. *Science* **337**, 1505–1510 (2012). doi: 10.1126/science.1224648; pmid: 22997331
56. D. Egger, J. Haushofer, E. Miguel, P. Niehaus, M. W. Walker, General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya. NBER working paper no. 26600 (2019); www.nber.org/papers/w26600.
57. P. Persson, M. Rossin-Slater, Family ruptures, stress, and the mental health of the next generation. *Am. Econ. Rev.* **108**, 1214–1252 (2018). doi: 10.1257/aer.20141406; pmid: 30091569
58. K. G. Noble *et al*., Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.* **18**, 773–778 (2015). doi: 10.1038/nn.3983; pmid: 25821911
59. C. Blair, C. C. Raver, Poverty, stress, and brain development: New directions for prevention and intervention. *Acad. Pediatr.* **16** (Suppl), S30–S36 (2016). doi: 10.1016/j.acap.2016.01.010; pmid: 27044699
60. A. Adhvaryu, J. Fenske, A. Nyshadham, Early life circumstance and adult mental health. *J. Polit. Econ.* **127**, 1516–1549 (2019). doi: 10.1086/701606
61. P. Sharkey, M. Besbris, M. Friedson, in *The Oxford Handbook of the Social Science of Poverty*, D. Brady, L. M. Burton, Eds. (Oxford Univ. Press, 2016).
62. M. Marmot, Social determinants of health inequalities. *Lancet* **365**, 1099–1104 (2005). doi: 10.1016/S0140-6736(05)71146-6; pmid: 15781105
63. C. B. Cunradi, R. Caetano, C. Clark, J. Schafer, Neighborhood poverty as a predictor of intimate partner violence among White, Black, and Hispanic couples in the United States: A multilevel analysis. *Ann. Epidemiol.* **10**, 297–308 (2000). doi: 10.1016/S1047-2797(00)00052-1; pmid: 10942878
64. J. Haushofer, C. Ringdal, J. P. Shapiro, X. Y. Wang, Income changes and intimate partner violence: Evidence from unconditional cash transfers in Kenya. NBER working paper no. 25627 (2019); www.nber.org/papers/w25627.
65. L. A. Goodman, K. F. Smyth, A. M. Borges, R. Singer, When crises collide: How intimate partner violence and poverty intersect to shape women's mental health and coping? *Trauma Violence Abuse* **10**, 306–329 (2009). doi: 10.1177/1524838009339754; pmid: 19776085
66. F. Cornaglia, N. E. Feldman, A. Leigh, Crime and mental well-being. *J. Hum. Resour.* **49**, 110–140 (2014). doi: 10.3368/jhr.49.1.110
67. R. Perez-Truglia, The effects of income transparency on well-being: Evidence from a natural experiment. *Am. Econ. Rev.* **110**, 1019–1054 (2020). doi: 10.1257/aer.20160256
68. R. Walker, G. Bantebya-Kyomuhendo, *The Shame of Poverty* (Oxford Univ. Press, 2014).
69. J. T. Cacioppo, M. E. Hughes, L. J. Waite, L. C. Hawkley, R. A. Thisted, Loneliness as a specific risk factor for depressive symptoms: Cross-sectional and longitudinal analyses. *Psychol. Aging* **21**, 140–151 (2006). doi: 10.1037/0882-7974.21.1.140; pmid: 16594799
70. C. Hakulinen *et al*., Mental disorders and long-term labour market outcomes: Nationwide cohort study of 2 055 720 individuals. *Acta Psychiatr. Scand.* **140**, 371–381 (2019). doi: 10.1111/acps.13067; pmid: 31254386
71. R. Mojtabai *et al*., Long-term effects of mental disorders on employment in the National Comorbidity Survey ten-year follow-up. *Soc. Psychiatry Psychiatr. Epidemiol.* **50**, 1657–1668 (2015). doi: 10.1007/s00127-015-1097-z; pmid: 26211661
72. B. Biasi, M. S. Dahl, P. Moser, Career effects of mental health. Working paper (2018); https://papers.ssrn.com/abstract=2544251.
73. C. Lund, K. Orkin, M. Witte, T. Davies, J. Haushofer, J. Bass, P. Bolton, S. Murray, L. Murray, W. Tol, G. Thornicroft, V. Patel, The economic effects of mental health interventions in low and middle-income countries. Working paper (2019); https://custom.cvent.com/4E741122FD8B4A1B97E483EC8BB51CC4/files/Event/159bd4dc083941a79dd0211437d5d7dc/d79476623ef7451884092864b1d7116b.pdf.
74. A. Mani, S. Mullainathan, E. Shafir, J. Zhao, Poverty impedes cognitive function. *Science* **341**, 976–980 (2013). doi: 10.1126/science.1238041; pmid: 23990553
75. S. Kaur, S. Mullainathan, F. Schilbach, Does financial strain lower productivity? Working paper (2019); https://economics.mit.edu/files/16997.
76. I. H. Gotlib, J. Joormann, Cognition and depression: Current status and future directions. *Annu. Rev. Clin. Psychol.* **6**, 285–312 (2010). doi: 10.1146/annurev.clinpsy.121208.131305; pmid: 20192795
77. M. Kremer, G. Rao, F. Schilbach, in *Handbook of Behavioral Economics, Foundations and Applications 2*, B. D. Bernheim, S. DellaVigna, D. Laibson, Eds. (North-Holland, 2019), vol. 2 of *Handbooks in Economics*, pp. 345–458.
78. A. T. Beck, *Depression: Clinical, Experimental, and Theoretical Aspects* (Univ. Pennsylvania Press, 1967).
79. J. de Quidt, J. Haushofer, Depression for economists. NBER working paper no. 22973 (2016), doi: 10.3386/w22973
80. D. Eil, J. M. Rao, The good news-bad news effect: Asymmetric processing of objective information about yourself. *Am. Econ. J. Microecon.* **3**, 114–138 (2011). doi: 10.1257/mic.3.2.114
81. C. W. Korn, T. Sharot, H. Walter, H. R. Heekeren, R. J. Dolan, Depression is related to an absence of optimistically biased belief updating about future life events. *Psychol. Med.* **44**, 579–592 (2014). doi: 10.1017/S0033291713001074; pmid: 23672737
82. T. Armstrong, B. O. Olatunji, Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clin. Psychol. Rev.* **32**, 704–723 (2012). doi: 10.1016/j.cpr.2012.09.004; pmid: 23059623
83. Y. M. Bayer, Z. Shtudiner, O. Suhorukov, N. Grisaru, Time and risk preferences, and consumption decisions of patients with clinical depression. *J. Behav. Exp. Econ.* **78**, 138–145 (2019). doi: 10.1016/j.socec.2019.01.003
84. D. A. Cobb-Clark, S. Dahmann, N. Kettlewell, "Depression, risk preferences and risk-taking behavior," research paper 12285 (Institute for the Study of Labor, 2019); https://papers.ssrn.com/abstract=3390275.
85. J. S. Rounds, J. G. Beck, D. M. Grant, Is the delay discounting paradigm useful in understanding social anxiety? *Behav. Res. Ther.* **45**, 729–735 (2007). doi: 10.1016/j.brat.2006.06.007; pmid: 16890909
86. R. C. Kessler *et al*., Age of onset of mental disorders: A review of recent literature. *Curr. Opin. Psychiatry* **20**, 359–364 (2007). doi: 10.1097/YCO.0b013e32816ebc8c; pmid: 17551351
87. S. Mall *et al*., Days out of role due to mental and physical illness in the South African stress and health study. *Soc. Psychiatry Psychiatr. Epidemiol.* **50**, 461–468 (2015). doi: 10.1007/s00127-014-0941-x; pmid: 25096982
88. B. A. Pescosolido, T. R. Medina, J. K. Martin, J. S. Long, The "backbone" of stigma: Identifying the global core of public prejudice associated with mental illness. *Am. J. Public Health* **103**, 853–860 (2013). doi: 10.2105/AJPH.2012.301147; pmid: 23488508
89. J. Sharac, P. McCrone, S. Clement, G. Thornicroft, The economic impact of mental health stigma and discrimination: A systematic review. *Epidemiol. Psichiatr. Soc.* **19**, 223–232 (2010). doi: 10.1017/S1121189X00001159; pmid: 21261218
90. S. Saxena, P. Sharan, M. Garrido, B. Saraceno, World Health Organization's Mental Health Atlas 2005: Implications for policy development. *World Psychiatry* **5**, 179–184 (2006). pmid: 17139355
91. C. Lauber, W. Rössler, Stigma towards people with mental illness in developing countries in Asia. *Int. Rev. Psychiatry* **19**, 157–178 (2007). doi: 10.1080/09540260701278903; pmid: 17464793
92. A. V. Banerjee, E. Duflo, The economic lives of the poor. *J. Econ. Perspect.* **21**, 141–167 (2007). doi: 10.1257/jep.21.1.141; pmid: 19212450
93. K. Xu *et al*., Protecting households from catastrophic health spending. *Health Aff. (Millwood)* **26**, 972–983 (2007). doi: 10.1377/hlthaff.26.4.972; pmid: 17630440
94. B. Weobong *et al*., Sustained effectiveness and cost-effectiveness of the Healthy Activity Programme, a brief psychological treatment for depression delivered by lay counsellors in primary care: 12-month follow-up of a randomised controlled trial. *PLOS Med.* **14**, e1002385 (2017). doi: 10.1371/journal.pmed.1002385; pmid: 28898283
95. V. Baranov, S. R. Bhalotra, P. Biroli, J. Maselko, Maternal depression, women's empowerment, and parental investment: Evidence from a large randomized control trial (2017); https://papers.ssrn.com/abstract=3081421.
96. P. Cuijpers, E. Weitz, E. Karyotaki, J. Garber, G. Andersson, The effects of psychological treatment of maternal depression on children and parental functioning: A meta-analysis. *Eur. Child Adolesc. Psychiatry* **24**, 237–245 (2015). doi: 10.1007/s00787-014-0660-6; pmid: 25522839
97. J. J. Heckman, S. Mosso, The economics of human development and social mobility. *Annu. Rev. Econom.* **6**, 689–733 (2014). doi: 10.1146/annurev-economics-080213-040753; pmid: 25346785
98. G. C. Patton *et al*., Our future: A Lancet commission on adolescent health and wellbeing. *Lancet* **387**, 2423–2478 (2016). doi: 10.1016/S0140-6736(16)00579-1; pmid: 27174304
99. D. Prinz, M. Chernew, D. Cutler, A. Frakt, "Health and economic activity over the lifecycle: Literature review," NBER working paper series 24865 (NBER, 2018), doi: 10.3386/w24865
100. M. L. A. Dückers, L. Reifels, D. P. De Beurs, C. R. Brewin, The vulnerability paradox in global mental health and its applicability to suicide. *Br. J. Psychiatry* **215**, 588–593 (2019). doi: 10.1192/bjp.2019.41
101. P. Brhlikova, A. M. Pollock, R. Manners, Global burden of disease estimates of depression—How reliable is the epidemiological evidence? *J. R. Soc. Med.* **104**, 25–34 (2011). doi: 10.1258/jrsm.2010.100080; pmid: 21205775
102. B. Milanovic, *Global Inequality: A New Approach for the Age of Globalization* (Harvard Univ. Press, 2016).

103. F. J. Charlson, A. J. Baxter, H. G. Cheng, R. Shidhaye, H. A. Whiteford, The burden of mental, neurological, and substance use disorders in China and India: A systematic analysis of community representative epidemiological studies. *Lancet* **388**, 376–389 (2016). doi: 10.1016/S0140-6736(16)30590-6; pmid: 27209143
104. V. Masson-Delmotte, *Global Warming of 1.5 OC: An IPCC Special Report on the Impacts of Global Warming of 1.5° C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (World Meteorological Organization, 2018).
105. P. M. Trang, J. Rocklöv, K. B. Giang, G. Kullgren, M. Nilsson, Heatwaves and hospital admissions for mental disorders in northern Vietnam. *PLOS ONE* **11**, e0155609 (2016). doi: 10.1371/journal.pone.0155609; pmid: 27195473
106. H. L. Berry, T. D. Waite, K. B. G. Dear, A. G. Capon, V. Murray, The case for systems thinking about climate change and mental health. *Nat. Clim. Chang.* **8**, 282–290 (2018). doi: 10.1038/s41558-018-0102-4
107. T. A. Carleton, Crop-damaging temperatures increase suicide rates in India. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8746–8751 (2017). doi: 10.1073/pnas.1701354114; pmid: 28760983
108. M. Burke, S. M. Hsiang, E. Miguel, Climate and conflict. *Annu. Rev. Econ.* **7**, 577–617 (2015). doi: 10.1146/annurev-economics-080614-115430
109. C. McMichael, J. Barnett, A. J. McMichael, An ill wind? Climate change, migration, and health. *Environ. Health Perspect.* **120**, 646–654 (2012). doi: 10.1289/ehp.1104375; pmid: 22266739
110. D. H. Autor, D. Dorn, G. H. Hanson, The China shock: Learning from labor-market adjustment to large changes in trade. *Annu. Rev. Econ.* **8**, 205–240 (2016). doi: 10.1146/annurev-economics-080315-015041
111. A. Adams-Prassl, T. Boneva, M. Golin, C. Rauh, "Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys," Cambridge-INET working paper no. 2020/18 (2020); https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3594297.
112. T. Suri, W. Jack, The long-run poverty and gender impacts of mobile money. *Science* **354**, 1288–1292 (2016). doi: 10.1126/science.aah5309; pmid: 27940873
113. J. H. Ha *et al*., Depression and Internet addiction in adolescents. *Psychopathology* **40**, 424–430 (2007). doi: 10.1159/000107426; pmid: 17709972
114. L. Y. Lin *et al*., Association between social media use and depression among U.S. young adults. *Depress. Anxiety* **33**, 323–331 (2016). doi: 10.1002/da.22466; pmid: 26783723
115. H. Allcott, L. Braghieri, S. Eichmeyer, M. Gentzkow, The welfare effects of social media. *Am. Econ. Rev.* **110**, 629–676 (2020). doi: 10.1257/aer.20190658
116. R. Mosquera, M. Odunowo, T. McNamara, X. Guo, R. Petrie, The economic effects of Facebook. *Exp. Econ.* 10.2139/ssrn.3312462 (2019).
117. World Bank Group, *World Development Report 2016: Digital Dividends* (World Bank Publications, 2016).
118. B. Bhat, J. Haushofer, J. de Quidt, V. Patel, F. Schilbach, G. Rao, P.-L. Vautrey, The economics consequences of depression: Evidence from Goa. AEA RCT Registry (2019), doi: 10.1257/rct.3823-1.0
119. C. Blattman, J. C. Jamison, M. Sheridan, Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia. *Am. Econ. Rev.* **107**, 1165–1206 (2017). doi: 10.1257/aer.20150503; pmid: 29553237
120. M. Thirunavukarasu, P. Thirunavukarasu, Training and national deficit of psychiatrists in India—A critical analysis. *Indian J. Psychiatry* **52** (suppl. 1), S83–S88 (2010). doi: 10.4103/0019-5545.69218; pmid: 21836723
121. WHO, "Mental health atlas 2017" (WHO, 2017);available at www.who.int/mental_health/evidence/atlas/mental_health_atlas_2017/en.
122. D. R. Singla *et al*., Psychological treatments for the world: Lessons from low- and middle-income countries. *Annu. Rev. Clin. Psychol.* **13**, 149–181 (2017). doi: 10.1146/annurev-clinpsy-032816-045217; pmid: 28482687
123. V. Patel, N. Chowdhary, A. Rahman, H. Verdeli, Improving access to psychological treatments: Lessons from developing countries. *Behav. Res. Ther.* **49**, 523–528 (2011). doi: 10.1016/j.brat.2011.06.012; pmid: 21788012
124. D. Chibanda *et al*., Effect of a primary care–based psychological intervention on symptoms of common mental disorders in Zimbabwe: A randomized clinical trial. *JAMA* **316**, 2618–2626 (2016). doi: 10.1001/jama.2016.19102; pmid: 28027368
125. R. Waters, Community workers lend human connection to COVID-19 response. *Health Aff. (Millwood)* **39**, 1112–1117 (2020). doi: 10.1377/hlthaff.2020.00836; pmid: 32634347
126. V. Patel *et al*., The magnitude of and health system responses to the mental health treatment gap in adults in India and China. *Lancet* **388**, 3074–3084 (2016). doi: 10.1016/S0140-6736(16)00160-4; pmid: 27209149
127. A. F. Jorm, Mental health literacy. Public knowledge and beliefs about mental disorders. *Br. J. Psychiatry* **177**, 396–401 (2000). doi: 10.1192/bjp.177.5.396; pmid: 11059991
128. R. Shidhaye *et al*., The effect of VISHRAM, a grass-roots community-based mental health programme, on the treatment gap for depression in rural communities in India: A population-based study. *Lancet Psychiatry* **4**, 128–135 (2017). doi: 10.1016/S2215-0366(16)30424-2; pmid: 28063879
129. P. Dasgupta, D. Ray, Inequality as a determinant of malnutrition and unemployment: Theory. *Econ. J. (Lond.)* **96**, 1011 (1986). doi: 10.2307/2233171
130. B. A. Arvind *et al*., T. M. Shibukumar, NMHS Collaborators Group, Prevalence and socioeconomic impact of depressive disorders in India: Multisite population-based cross-sectional study. *BMJ Open* **9**, e027250 (2019). doi: 10.1136/bmjopen-2018-027250; pmid: 31253618
131. E. P. Green, C. Blattman, J. Jamison, J. Annan, Does poverty alleviation decrease depression symptoms in post-conflict settings? A cluster-randomized trial of microenterprise assistance in Northern Uganda. *Glob. Ment. Health (Camb.)* **3**, e7 (2016). doi: 10.1017/gmh.2015.28; pmid: 28596876
132. C. Blattman, S. Dercon, S. Franklin, Impacts of industrial and entrepreneurial jobs on youth: 5-year experimental evidence on factory job offers and cash grants in "Ethiopia," NBER working paper no. 25788 (2019); www.nber.org/papers/w25788.
133. C. Blattman, N. Fiala, S. Martinez, "The Long Term Impacts of Grants on Poverty: 9-year Evidence From Uganda's Youth Opportunities Program," NBER working paper no. 24999 (2018); www.nber.org/papers/w24999.
134. L. Hjelm, S. Handa, J. de Hoop, T. Palermo; Zambia CGP and MCP Evaluation Teams, Poverty and perceived stress: Evidence from two unconditional cash transfer programs in Zambia. *Soc. Sci. Med.* **177**, 110–117 (2017). doi: 10.1016/j.socscimed.2017.01.023; pmid: 28167339
135. S. Baird, B. de Hoop, B. Özler, Income shocks and adolescent mental health. *J. Hum. Resour.* **48**, 370–403 (2013). doi: 10.3368/jhr.48.2.370
136. K. Kilburn, H. Thirumurthy, C. T. Halpern, A. Pettifor, S. Handa, Effects of a large-scale unconditional cash transfer program on mental health outcomes of young people in Kenya. *J. Adolesc. Health* **58**, 223–229 (2016). doi: 10.1016/j.jadohealth.2015.09.023; pmid: 26576822
137. G. Angeles *et al*., Government of Malawi's unconditional cash transfer improves youth mental health. *Soc. Sci. Med.* **225**, 108–119 (2019). doi: 10.1016/j.socscimed.2019.01.037; pmid: 30826585
138. J. Haushofer, R. Mudida, J. Shapiro, "The comparative impact of cash transfers and psychotherapy on psychological and economic well-being," working paper (2020); www.researchgate.net/profile/Robert_Mudida/publication/340435415_The_Comparative_Impact_of_Cash_Transfers_and_Psychotherapy_on_Psychological_and_Economic_Well-being.
139. J. Haushofer, M. Chemin, C. Jang, J. Abraham, Economic and psychological effects of health insurance and cash transfers: Evidence from a randomized experiment in Kenya. *J. Dev. Econ.* **144**, 102416 (2020). doi: 10.1016/j.jdeveco.2019.102416
140. E. Edmonds, C. Theoharides, The short term impact of a productive asset transfer in families with child labor: Experimental evidence from the Philippines. *J. Dev. Econ.* **146**, 102486 (2020). doi: 10.1016/j.jdeveco.2020.102486
141. S. Janzen, N. Magnan, S. Sharma, W. Thompson, Short-term impacts of a pay-it-forward livestock transfer and training program in Nepal. *AEA Pap. Proc.* **108**, 422–425 (2018). doi: 10.1257/pandp.20181120
142. N. Glass, N. A. Perrin, A. Kohli, J. Campbell, M. M. Remy, Randomised controlled trial of a livestock productive asset transfer programme to improve economic and health outcomes and reduce intimate partner violence in a postconflict setting. *BMJ Glob. Health* **2**, e000165 (2017). doi: 10.1136/bmjgh-2016-000165; pmid: 28589002
143. S. Baird, C. McIntosh, B. Özler, When the money runs out: Do cash transfers have sustained effects on human capital accumulation? 30 March 2019 (socialprotection.org, 2019); https://socialprotection.org/sites/default/files/publications_files/Baird%20McIntosh%20Ozler_When%20the%20Money%20Runs%20Out_COMBINED_Revision_FINAL.pdf.
144. J. Shapiro, The impact of recipient choice on aid effectiveness. *World Dev.* **116**, 137–149 (2019). doi: 10.1016/j.worlddev.2018.10.010; pmid: 30944504
145. A. Aguilar, Cash and conditionalities' effects on early childhood: medium-term effects on physical and cognitive development (2012); www.aguilaresteva.com/wp-content/uploads/2012/09/Progresa_ECD_paper.pdf.
146. K. Macours, R. Vakis, *Changing Households' Investments and Aspirations Through Social Interactions: Evidence From a Randomized Transfer Program* (The World Bank, 2009).
147. L. C. H. Fernald, M. Hidrobo, Effect of Ecuador's cash transfer program (Bono de Desarrollo Humano) on child development in infants and toddlers: A randomized effectiveness trial. *Soc. Sci. Med.* **72**, 1437–1446 (2011). doi: 10.1016/j.socscimed.2011.03.005; pmid: 21531060
148. L. Ismayilova *et al*., Improving mental health among ultra-poor children: Two-year outcomes of a cluster-randomized trial in Burkina Faso. *Soc. Sci. Med.* **208**, 180–189 (2018). doi: 10.1016/j.socscimed.2018.04.022; pmid: 29743136
149. R. Heath, M. Hidrobo, S. Roy, Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali. *J. Dev. Econ.* **143**, 102410 (2020). doi: 10.1016/j.jdeveco.2019.102410
150. K. Kilburn *et al*., Cash transfers, young women's economic well-being, and HIV risk: Evidence from HPTN 068. *AIDS Behav.* **23**, 1178–1194 (2019). doi: 10.1007/s10461-018-2329-5; pmid: 30415429
151. V. Patel *et al*., Detecting common mental disorders in primary care in India: A comparison of five screening questionnaires. *Psychol. Med.* **38**, 221–228 (2008). doi: 10.1017/S0033291707002334; pmid: 18047768
152. K. Kroenke, R. L. Spitzer, The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr. Ann.* **32**, 509–515 (2002). doi: 10.3928/0048-5713-20020901-06
153. R. L. Spitzer, K. Kroenke, J. B. W. Williams, B. Löwe, A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch. Intern. Med.* **166**, 1092–1097 (2006). doi: 10.1001/archinte.166.10.1092; pmid: 16717171
154. G.-C. Ali, G. Ryan, M. J. De Silva, Validated screening tools for common mental disorders in low and middle income countries: A systematic review. *PLOS ONE* **11**, e0156939 (2016). doi: 10.1371/journal.pone.0156939; pmid: 27310297
155. V. Patel *et al*., The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: A randomised controlled trial. *Lancet* **389**, 176–185 (2017). doi: 10.1016/S0140-6736(16)31589-6; pmid: 27988143

ACKNOWLEDGMENTS

We thank V. Baranov, T. Cruz Vital, J. Das, E. Gallagher, J. Haushofer, A. Karing, J. Li, C. Lund, M. Mani, K. Orkin, J. Pan, and K. Rao, for helpful comments and suggestions. We thank six anonymous referees for detailed comments and helpful suggestions. We thank C. Roth and L. Hensel for kindly providing us the data needed to construct Fig. 3. **Funding:** None declared. **Author contributions:** All authors contributed equally to this work. **Competing interests:** The authors have no competing interests or financial conflicts of interest.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eaay0214/suppl/DC1
Supplementary Text
Figs. S1 to S3
References (*156*–*164*)

10.1126/science.aay0214

RESEARCH ARTICLE SUMMARY

MOLECULAR BIOLOGY

CiBER-seq dissects genetic networks by quantitative CRISPRi profiling of expression phenotypes

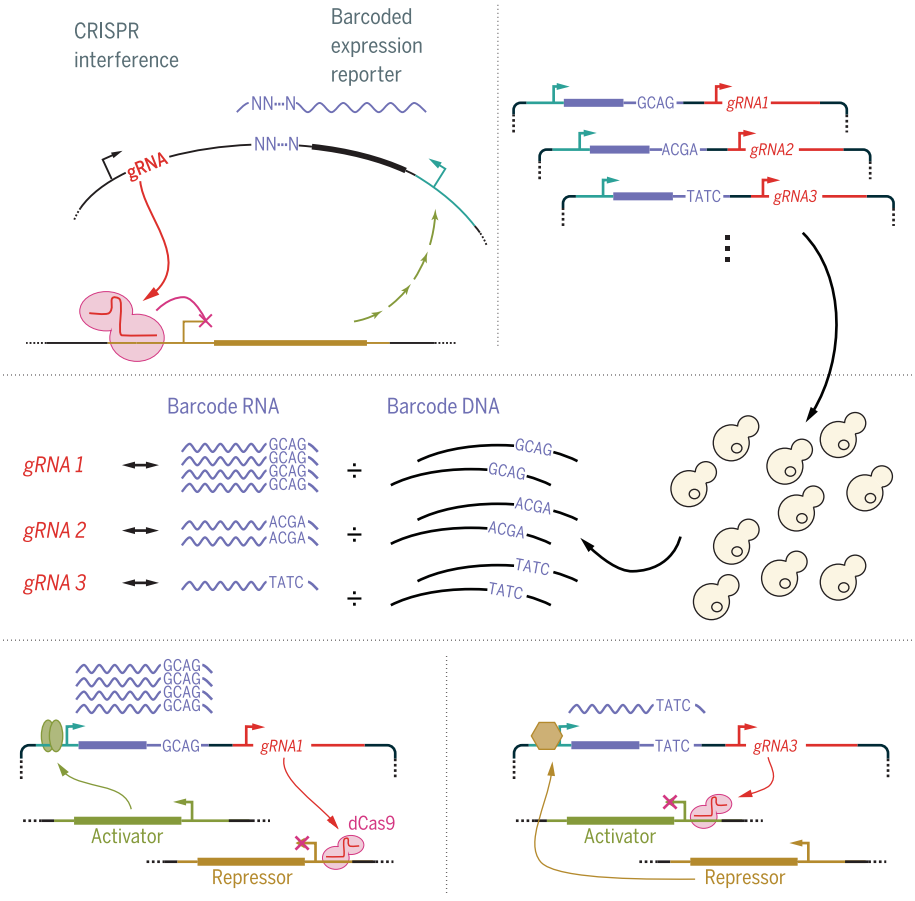
Ryan Muller, Zuriah A. Meacham, Lucas Ferguson, Nicholas T. Ingolia*

INTRODUCTION: Systematically profiling the effects of genetic perturbations is a powerful approach that has revealed the molecular basis for a wide range of biological phenomena. The simple, programmable DNA recognition of CRISPR-Cas9 enables genome-wide genetic analysis in human cells and many other systems. Cas9 is guided by a short RNA to a complementary sequence in the genome, where it can introduce mutations or alter gene expression. Pooled libraries of guide RNAs (gRNAs) that individually target each gene in the genome allow us to introduce genetic perturbations systematically into a population of cells. A key challenge is measuring the phenotypic effects caused by individual guides in these pooled libraries and linking these phenotypes back to the associated gRNA, thereby finding the gene that is responsible.

RATIONALE: Molecular phenotypes such as gene expression changes provide a clear and sensitive measure for many cellular processes. We sought a general approach to profile how the expression of a particular gene of interest changed when other genes were perturbed. We began with a library of gRNAs, each disrupting one gene, and linked these guides with an expression reporter containing a guide-specific nucleotide barcode. gRNAs that alter reporter expression will change the abundance of the expressed RNA barcode specifically associated with that guide. Deep sequencing of these expressed barcodes quantifies each of these guide-specific reporter expression effects individually within a pooled, complex population. We have implemented this strategy by combining CRISPR interference (CRISPRi) with barcoded expression reporter sequencing (CiBER-seq).

RESULTS: We used CiBER-seq to profile the responses of several yeast promoters tied to a range of biological functions. Each promoter yielded a distinct pattern of responses that could be understood in terms of its known function and regulation. For example, we rediscover the control of *MET6* expression by regulatory ubiquitylation and connect the bud scar protein Cwp1 to other genes required for budding and cytokinesis. Our analysis of the *HIS4* promoter, a well-characterized target of the integrated stress response, yielded a range of genetic perturbations that activate this pathway by causing the accumulation of uncharged transfer RNAs (tRNAs). We also uncovered a notable role for tRNA depletion in this response, as impaired tRNA biogenesis activated *HIS4* expression through a distinct pathway. In order to understand this regulation, we carried out genetic interaction analysis and looked for quantitative deviations in CiBER-seq profiles caused by the introduction of a second genetic perturbation. We also developed an indirect CiBER-seq approach to measure translational and posttranslational regulation, which both play roles in the signaling pathways upstream of *HIS4*.

CONCLUSION: CiBER-seq produces comprehensive phenotypic profiles that offer insights into gene function and regulation. These high-throughput and quantitative phenotypic measurements are also well suited for the systematic measurement of genetic interactions, which contain rich information about the operation of biological processes. This approach can be applied to study a wide range of transcriptional, translational, and posttranslational regulatory responses, and it has the potential to shed light on many areas of biology. ■



CRISPRi with barcoded expression reporters. CRISPR-Cas9 gRNA cassettes are linked with transcriptional reporters containing specific barcodes. The RNA-to-DNA ratio for each barcode, measured by deep sequencing, reveals the reporter expression phenotype induced by each gRNA.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: ingolia@berkeley.edu
Cite this article as R. Muller *et al.*, *Science* **370**, eabb9662 (2020). DOI: 10.1126/science.abb9662

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb9662

RESEARCH ARTICLE

MOLECULAR BIOLOGY

CiBER-seq dissects genetic networks by quantitative CRISPRi profiling of expression phenotypes

Ryan Muller[1], Zuriah A. Meacham[1], Lucas Ferguson[1], Nicholas T. Ingolia[1,2]*

To realize the promise of CRISPR-Cas9–based genetics, approaches are needed to quantify a specific, molecular phenotype across genome-wide libraries of genetic perturbations. We addressed this challenge by profiling transcriptional, translational, and posttranslational reporters using CRISPR interference (CRISPRi) with barcoded expression reporter sequencing (CiBER-seq). Our barcoding approach allowed us to connect an entire library of guides to their individual phenotypic consequences using pooled sequencing. CiBER-seq profiling fully recapitulated the integrated stress response (ISR) pathway in yeast. Genetic perturbations causing uncharged transfer RNA (tRNA) accumulation activated ISR reporter transcription. Notably, tRNA insufficiency also activated the reporter, independent of the uncharged tRNA sensor. By uncovering alternate triggers for ISR activation, we illustrate how precise, comprehensive CiBER-seq profiling provides a powerful and broadly applicable tool for dissecting genetic networks.

CRISPR-Cas9 has emerged as a powerful and versatile tool for creating precise, programmable genetic perturbations (*1*). CRISPR-based knockout (*2*–*4*) and knockdown (*5*) approaches enable systematic, genome-wide genetic analysis in a wide range of cells and organisms. The Cas9 protein binds short RNAs that guide this protein-RNA complex to complementary sites in the genome, where it can induce mutations or silence promoters (*1*). Libraries of guide RNAs (gRNAs), each targeting one individual gene, can be used to create a population of cells that each express one distinct guide (*6*–*8*). The phenotype of each cell then reflects the impact of the single guide that it expresses. It is straightforward to identify guides that affect cell survival or proliferation, but growth is a crude phenotype that is poorly suited to address many important biological questions (*9*). The scope of CRISPR-based genetics would be expanded by improved techniques to measure more-specific and relevant phenotypes across this diverse population and link these measurements back to individual guides.

Molecular phenotypes, such as the expression level of a critical gene or the stability of a key protein, provide a focused and sensitive gauge for many aspects of cell physiology. We devised an approach for profiling a transcriptional, translational, or posttranslational regulatory response comprehensively across CRISPR-based perturbations genome-wide. We adapted barcoded expression reporters (*10*) to produce quantitative phenotypic profiles from bulk sequencing of highly diverse populations. These profiles also enabled high-precision genetic interaction analyses, which use double-mutant phenotypes to map the structure of regulatory networks (*11*). This direct sequencing approach offers substantial advantages over fluorescent reporters for CRISPR-based genetics. Fluorescence phenotypes are typically analyzed by cell sorting (*9*), which imposes bottlenecks on the cell population size and discretizes quantitative fluorescence measurements into a few broad gates. Our approach circumvented both of these limitations. It also complemented broader expression profiles from single-cell approaches such as Perturb-seq (*12*, *13*), CRISP-seq (*14*), and CROP-seq (*15*), which cannot currently capture enough cells to approach genome-scale coverage. Thus, better ways to profile molecular phenotypes across genome-scale guide libraries stand to benefit many areas of biology. Here, we combined CRISPR interference (CRISPRi) with barcoded expression reporter sequencing (CiBER-seq) to measure cellular responses provoked by gRNA-mediated knockdown.

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [2]California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, CA 94720, USA.
*Corresponding author. Email: ingolia@berkeley.edu

Results

Barcoded expression reporters linked transcriptional responses with gRNA-mediated perturbations in massively parallel screens

The development of CiBER-seq relied on massively parallel measurements of reporter expression in a diverse population by the deep sequencing of short-sequence barcodes embedded in the reporter transcript (*10*). Each barcode was linked to one gRNA, and the RNA abundance of each barcode reflected reporter expression levels in the cells containing that barcode and expressing the associated guide (fig. S1). Barcode RNA levels are also affected by the variable abundance of cells containing the barcode and its associated guide and by nonspecific disruptions of cell viability or transcription. To correct for these confounding effects, we paired each barcoded reporter with a barcoded control transcript driven from a housekeeping promoter, which will experience the same overall cellular environment. We implemented CiBER-seq in budding yeast and delivered guide-reporter pairs on autonomous low-copy plasmids. We tagged each guide sequence in a comprehensive yeast gRNA library with pairs of random nucleotide barcodes and determined the linkage between barcode pairs and guides by long-read next-generation sequencing (Fig. 1A). Our library contained 10 guides per gene (~60,000 in total) (*16*) and ~240,000 distinct barcode pairs (approximately four per guide on average). By linking multiple barcodes with each guide, we were able to obtain independent measurements of guide effects in a single experiment.

We drove gRNA expression from a tetracycline-inducible promoter (*17*), which allowed us to specifically measure the difference in barcode expression before and after guide induction. This experimental design allowed us to exclude technical effects resulting from the sequence of a barcode. Inducible CRISPRi also facilitated the measurements of guides with strong fitness effects by allowing us to propagate cells without guide induction and thereby avoid the premature loss of guides with growth defects. We analyzed the multifactorial barcode abundance data in a generalized linear model framework, implemented by a massively parallel reporter assay linear model (mpralm) (*18*), which estimates the change in reporter expression caused by CRISPRi guide induction while controlling for baseline reporter expression and changes in the paired housekeeping barcode. This framework allowed us to incorporate replicate measurements, and we carried out all CiBER-seq experiments in biological duplicate. It also allowed us to identify guides causing a statistically significant change in normalized reporter expression, using statistical approaches developed for gene expression measurements in limma (linear models for microarray data) (*19*).

To explore whether CiBER-seq could dissect a range of regulatory circuits, we profiled the responses of four promoters—P(*MET6*), P(*CWP1*), P(*PHO5*), and P(*HIS4*)—whose gene products encompass a wide range of cellular roles. The protein product of *MET6* catalyzes the conversion of homocysteine to L-methionine and plays a central role in sulfur and one-carbon metabolism. The activity of P(*MET6*) is controlled by the transcription factor Met4, which is ubiquitinated and inactivated by the SCF^{MET30} complex (*20*) (Fig. 1B). In agreement with the inhibitory effect of SCF^{MET30}, CiBER-seq analysis of P(*MET6*) regulation identified several guides

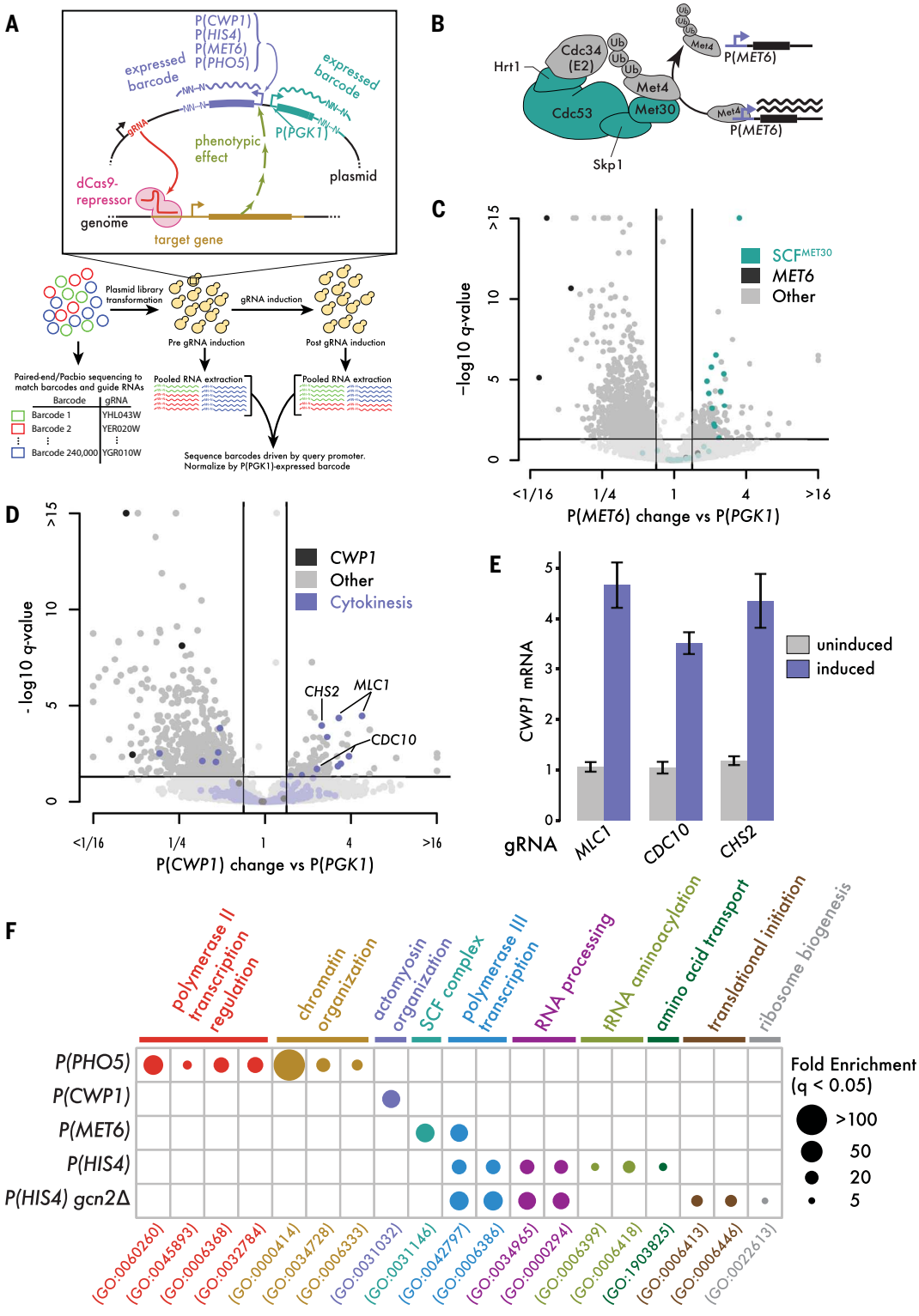


Fig. 1. Barcoded expression reporters linked transcriptional responses with gRNA-mediated perturbations in massively parallel screens. (**A**) Schematic of CiBER-seq profiling experiment. Barcode expression is driven by a query promoter or matched P(*PGK1*) normalizer. Changes in relative barcode expression after gRNA induction link individual gRNAs with their corresponding phenotypic effect. (**B**) Model diagram of the SCF^{MET30} complex. Subunits with at least one significant guide [$q < 0.05$ and >1.5-fold P(*MET6*) increase] are colored green. (**C**) Genome-wide CiBER-seq profile of P(*MET6*) transcription changes upon gRNA induction, relative to P(*PGK1*). Each point represents one gRNA, with guides against SCF^{MET30} complex or *MET6* itself colored as indicated. Lines indicate cutoffs for significant ($q < 0.05$) and substantial (>1.5-fold change) effects. (**D**) Genome-wide CiBER-seq profile of P(*CWP1*) transcription, as in (C), with guides against cytokinesis GO term genes and *CWP1* itself colored as indicated. Significant guides against cytokinesis genes *MLC1*, *CDC10*, and *CHS2* are labeled; the guide with the strongest *q* value was selected for validation. (**E**) Endogenous *CWP1* mRNA measurements before and after *MLC1*, *CDC10*, and *CHS2* knockdown. Error bars are the standard deviation across $N = 3$ biological replicates. (**F**) GO analysis of each CiBER-seq query promoter profile. Guides were filtered by $q < 0.05$ and >2-fold increase, and resulting gene lists were analyzed for GO category overrepresentation using Fisher's exact test with false discovery rate (FDR)–adjusted $P < 0.05$. The most statistically significant entry was chosen from chains of hierarchically nested categories, and all chains with significant categories for any promoter are represented in the plot.

targeting this complex that increased P(*MET6*)-driven barcode expression relative to P(*PGK1*) controls (Fig. 1, B and C, fig. S2A, and data S1). We found a 32-fold enrichment ($q < 1.3 \times 10^{-5}$) for this functional category among guides activating P(*MET6*), spanning nearly every subunit of this complex.

CiBER-seq analysis indicated that P(*CWP1*) activity was increased by genetic perturbation of cytokinesis. The Cwp1 protein localizes to bud scars, the cell wall structures formed at the site of cytokinesis in budding yeast (*21*). Appropriately regulated transcription of *CWP1* is important for this localization, although the mode of its regulation is unknown. We found that gRNAs targeting genes involved in budding and cytokinesis—including septins, myosin, and chitin synthase (*22*)—all activated P(*CWP1*) (Fig. 1D and data S1). We validated that knockdown of three representative genes, *MLC1*, *CDC10*, and *CHS2*, all induced the endogenous *CWP1* gene

(Fig. 1E). Thus, P(*CWP1*) is activated by a coherent set of gRNAs that is consistent with the link between Cwp1 and budding.

Each of the four promoters displayed a distinctive and largely nonoverlapping phenotypic profile consistent with their known regulation and function (Fig. 1F and tables S1 to S4). CiBER-seq profiling of P(*PHO5*), a model promoter for early studies of chromatin-based transcription regulation (*23*), identified a collection of genes involved in nucleosome remodeling and RNA polymerase II initiation (fig. S2B and data S1). Similarly, P(*HIS4*) CiBER-seq yielded a complex profile of responses, including activation of P(*HIS4*) by knockdown of aminoacyl-tRNA synthetases (fig. S2C and data S1). The *HIS4* gene, which encodes an amino acid biosynthetic enzyme, is a well-characterized transcriptional target of the yeast integrated stress response (ISR), also known as the general amino acid control (GAAC) response. This deeply conserved pathway up-regulates biosynthetic genes in response to elevated levels of uncharged tRNAs that arise during amino acid starvation (*24*). It is therefore easy to understand how directly impairing tRNA charging by knockdown of the synthetase enzyme that carries out this reaction would trigger the ISR and thereby activate P(*HIS4*) as well as P(*MET6*), which is also an ISR transcriptional target.

We also observed clear P(*HIS4*) and P(*MET6*) activation from guides targeting RNA polymerase III, a scenario that should result in the absence of tRNA rather than the accumulation of uncharged tRNA. Because these uncharged tRNAs are sensed by the Gcn2 kinase, we investigated how the CiBER-seq profile of P(*HIS4*) would differ in a gcn2Δ knockout strain. Deletion of this sensor kinase eliminated the response of P(*HIS4*) to guides against tRNA synthetases, whereas the effects of RNA polymerase III knockdown remained at least as strong (Fig. 1F, fig. S2D, and data S1). Because we did not observe any effect of these guides on P(*PHO5*) or P(*CWP1*), it seemed unlikely that our results reflected a change in the activity of the P(*PGK1*) promoter that we used as a common point of reference in these experiments. Nonetheless, we set out to profile P(*HIS4*) activity more directly, normalizing the RNA expression level of the barcode against the DNA abundance rather than a control promoter.

CiBER-seq recapitulated known genetic regulators of ISR and identified regulators related to tRNA insufficiency

CiBER-seq profiling of P(*HIS4*) and P(*PGK1*) individually, with normalization against barcode DNA abundance (Fig. 2A and fig. S3), allowed us to unambiguously attribute observed transcriptional responses to one promoter. Our unbiased gene ontology (GO) analysis captured specific and distinct categories of P(*PGK1*) and P(*HIS4*) activators (Fig. 2B and tables S6 and S7), in agreement with a direct comparison of P(*HIS4*) and P(*PGK1*) effects during CRISPRi perturbation (Fig. 2A and fig. S3, A and B). Although guides that activated P(*HIS4*) did not induce P(*PGK1*), knockdown of general RNA polymerase II transcription reduced expression in both reporters (fig. S3, A to E). Additionally, analysis of P(*PGK1*) in isolation demonstrated a clear induction in response

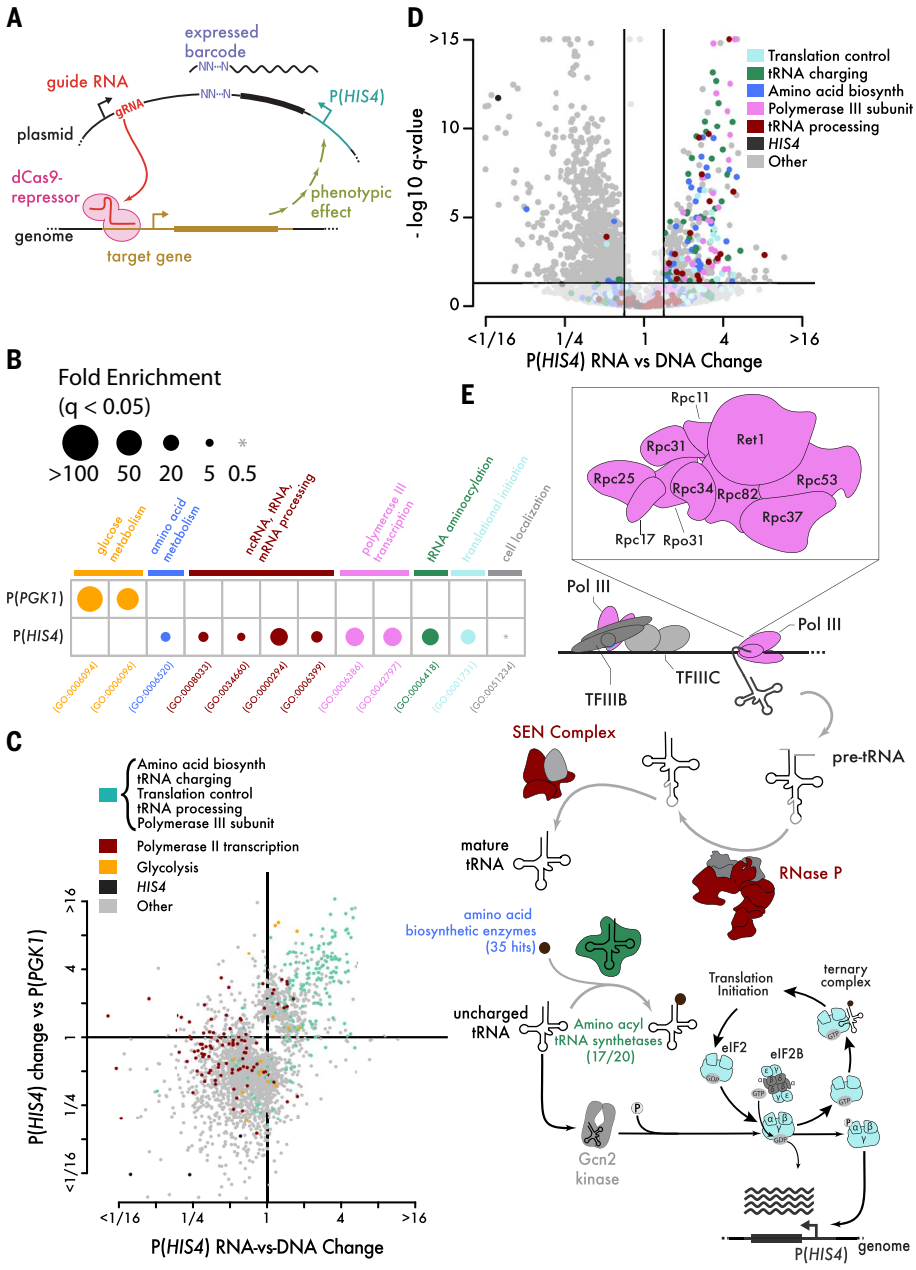


Fig. 2. CiBER-seq recapitulated known genetic regulators of ISR and identified regulators related to tRNA insufficiency. (**A**) Schematic of CiBER-seq profiling experiment, with modifications to isolate the regulatory effects of guide-mediated knockdown on a single promoter. (**B**) GO analysis of P(*PGK1*) and P(*HIS4*) with DNA normalization, as in Fig. 1F. (**C**) Comparison of CiBER-seq profiles for P(*PGK1*)–normalized and DNA-normalized P(*HIS4*) CiBER-seq analysis. Guides with significant effects in either profile are shown, and all GO terms highlighted in (B) are condensed to one group. (**D**) Genome-wide CiBER-seq profile of P(*HIS4*) transcription relative to DNA barcode abundance. Each point represents a different gRNA, analyzed to determine the change in P(*HIS4*) barcode RNA levels upon gRNA induction, normalized against the change in barcode DNA. The lines indicate cutoffs for significant ($q < 0.05$) and substantial (>1.5-fold change) effects. Guides are color coded by relevant statistically overrepresented GO terms with strong effects on P(*HIS4*). (**E**) Schematic of biological complexes with significant guides [as in (D)], with colors corresponding to GO terms. Subunits without a significant guide are displayed in gray. Pol III, polymerase III; TFIII, transcription factor for Pol III.

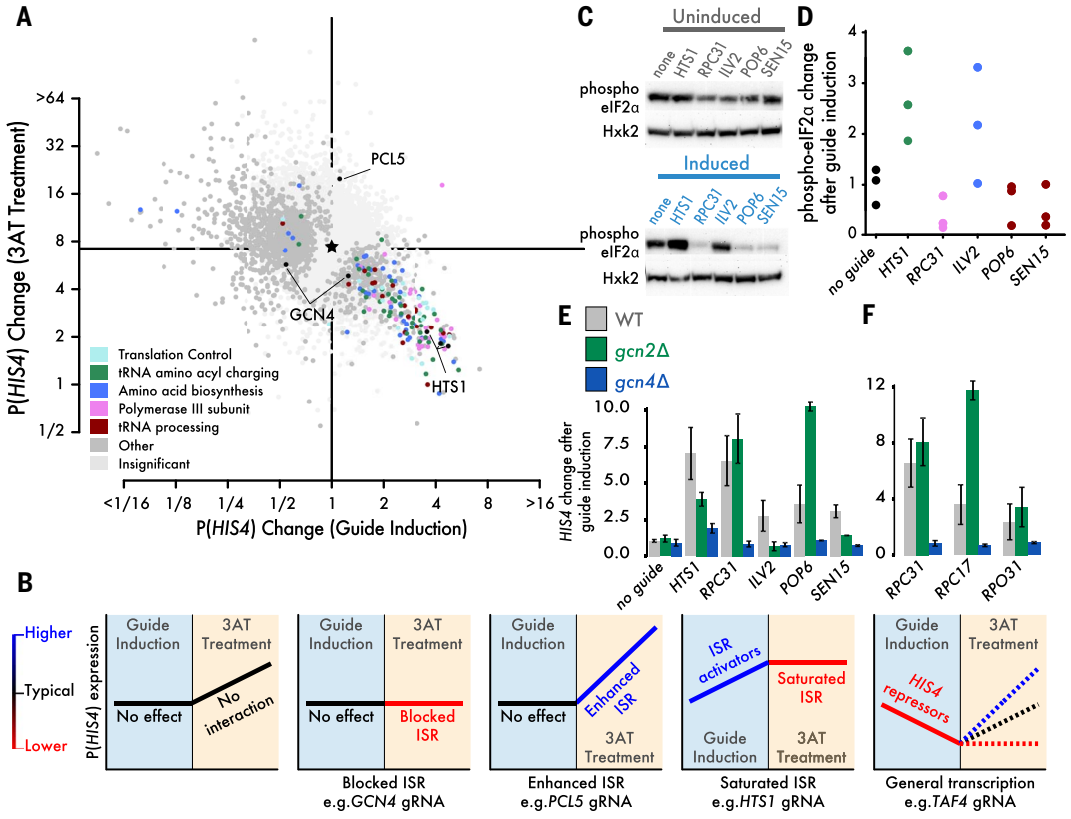


Fig. 3. tRNA insufficiency triggered HIS4 transcription independently of eIF2α phosphorylation and Gcn2 kinase. (**A**) Comparison of P(*HIS4*) CiBER-seq profiles before and after 3AT treatment, analyzed and colored as in Fig. 2D. (**B**) Schematic outlining expected ISR responses for CRISPRi knockdown of genes in various functional categories, before and after 3AT treatment. (**C**) Western blot for eIF2α phosphorylation relative to hexokinase (Hxk2) loading control. Knockdown of amino acid biosynthesis (*ILV2*) and aminoacyl tRNA charging (*HTS1*) increases eIF2α phosphorylation, whereas knockdown of RNA polymerase III (*RPC31*) or tRNA processing (*POP6* and *SEN15*) does not. (**D**) Quantification of eIF2α phosphorylation relative to hexokinase loading control across *N* = 3 biological replicates. (**E**) Change in endogenous *HIS4* mRNA levels after guide induction measured by qPCR. Endogenous *HIS4* activation by *ILV2*, *HTS1*, *RPC31*, *POP6*, and *SEN15* knockdown is completely *GCN4* dependent. *HIS4* activation by *ILV2* knockdown is also completely *GCN2* dependent, whereas effects of *RPC31* and *POP6* knockdown are entirely *GCN2* independent, and *HTS1* and *SEN15* show intermediate dependency. Error bars represent standard deviation for *N* = 3 biological replicates. (**F**) As in (E) for guides targeting three distinct RNA polymerase III subunits. Error bars represent standard deviation for *N* = 3 biological replicates.

to knockdown of glycolytic enzymes (fig. S3, A to C), which suggests a homeostatic transcriptional activation of *PGK1* expression in response to impaired glycolysis. Although the P(*PGK1*) promoter is often used with the intent to produce constitutive expression, it is subject to regulation (*25*), and Pgk1 activity increases when glycolysis is inhibited (*26*).

Our P(*HIS4*) CiBER-seq profile with DNA normalization confirmed observations from a previous profile based on P(*PGK1*) normalization—namely, that well-characterized defects in tRNA charging and newly identified deficiencies in tRNA synthesis affected P(*HIS4*) specifically (Fig. 2C). We thus surveyed the molecular complexes whose knockdown activated P(*HIS4*) (Fig. 2D). We identified 35 different gRNAs targeting amino acid biosynthesis pathways, along with guides against 17 of the 20 of the aminoacyl-tRNA synthetases (Fig. 2, D and E), which are all expected to interfere with tRNA charging. We also observed P(*HIS4*) activation in response to knockdown of each individual component of the eIF2 translation initiation complex (Fig. 2, D and E); depletion of these proteins directly increases translation of the Gcn4 transcription factor, leading to P(*HIS4*) induction (*24*, *27*).

We also found P(*HIS4*) activation in response to guides targeting many steps of tRNA biogenesis, which suggests that overall tRNA depletion triggered the ISR transcriptional program. Guides targeting subunits of RNA polymerase III, which transcribes tRNAs, as well as the tRNA processing complex ribonuclease (RNAse) P and the tRNA splicing endonuclease (SEN) complex (*28*) all activated P(*HIS4*) transcription (Fig. 2, A and C). In contrast to known ISR triggers, these genetic perturbations should not lead to the accumulation of uncharged tRNAs but should reduce overall tRNA levels. Although depletion of initiator methionyl-tRNA can induce P(*HIS4*) (*29*) and RNA polymerase III defects can lead to initiator methionyl-tRNA depletion (*30*), this effect could not explain our observations here. Initiator tRNA does not contain an intron, and so the loss of SEN should not reduce initiator tRNA levels. To further exclude initiator tRNA depletion as an explanation for the effects of RNA polymerase III knockdown, we overexpressed initiator tRNA during CRISPRi-mediated ISR activation (fig. S3, J and K). Impaired tRNA transcription showed no particular susceptibility to suppression of the ISR response by high-copy overexpression of initiator tRNA, relative to disruption of amino acid biosynthesis or tRNA charging (fig. S3K). Thus, elongator tRNA depletion can directly activate ISR transcription, perhaps through its effects on translation elongation.

tRNA insufficiency triggered HIS4 transcription independently of eIF2α phosphorylation and Gcn2 kinase

Because the ISR is not activated during log-phase growth, our initial CiBER-seq data did

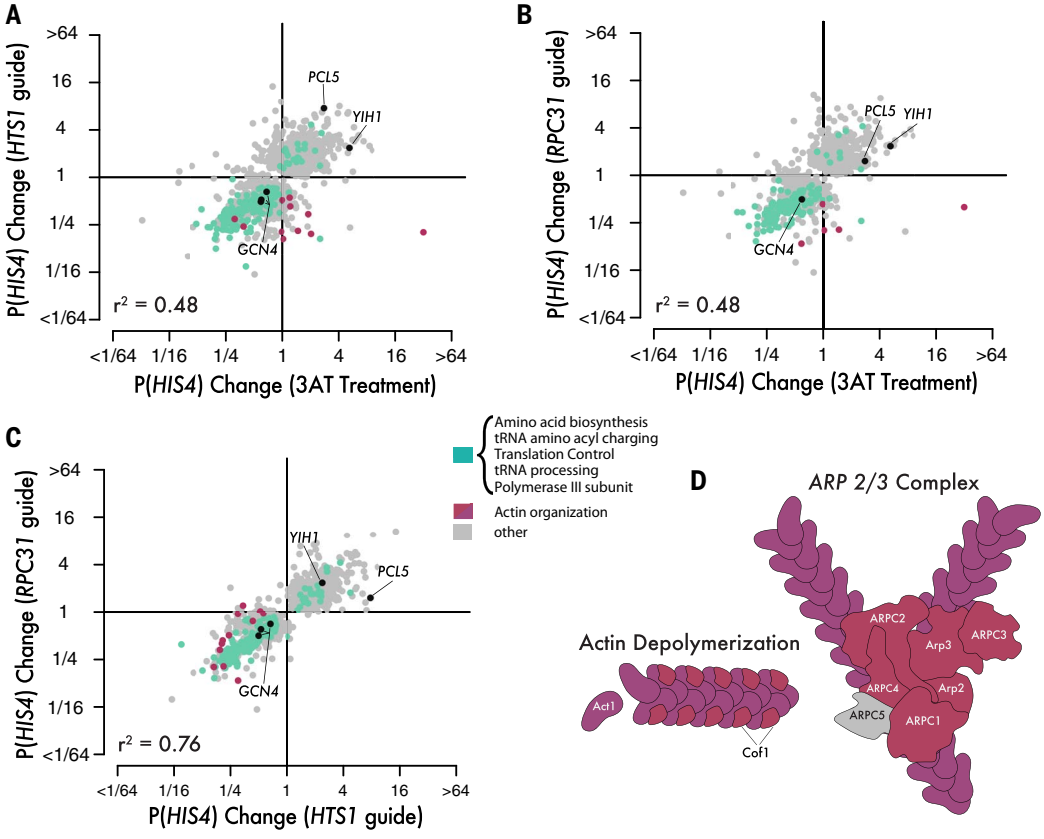


Fig. 4. Perturbations of the ARP2/3 complex prevented ISR activation by HTS1 or RPC31 knockdown. (**A** to **C**) Pair-wise comparison of P(*HIS4*) CiBER-seq profiles between ISR activation by 3AT treatment, *HTS1* knockdown, and *RPC31* knockdown. CiBER-seq profiles represent changes in the guide effect on P(*HIS4*) expression in the context of an ISR activator (as in Fig. 3A) relative to the change caused by the guide in isolation (as in Fig. 2D). Plotted guides were significant ($q < 0.05$) in at least one of the three epistatic profiles. All guides, regardless of significance, were used to calculate the pairwise Pearson correlations. (**D**) Coverage of actin cytoskeleton and ARP2/3 complex by guides that block ISR induction during *HTS1* or *RPC31* knockdown but not 3AT treatment.

not identify expression changes caused by loss of canonical, positive transducers of ISR signaling, such as *GCN2* and *GCN4*. To uncover the phenotypes of these and other ISR pathway genes, we next looked for regulators whose depletion would block P(*HIS4*) activation triggered by the toxic histidine analog 3-amino-1,2,4-triazole (3AT), which blocks histidine biosynthesis (Fig. 3, A and B). Guides that activated P(*HIS4*) under replete nutrient conditions did not further elevate P(*HIS4*) expression upon 3AT treatment (Fig. 3A). We inferred that Gcn4-mediated activation is saturated under these conditions; however, knockdown of Gcn4 degradation factors, such as *PCL5* (*31*), can enhance 3AT-mediated P(*HIS4*) induction. Meanwhile, guides that reduced P(*HIS4*) transcription, including guides targeting the core RNA polymerase II transcription machinery (fig. S2, A to C), did not interfere with P(*HIS4*) activation upon 3AT treatment. We therefore reasoned that guides affecting amino acid biosynthesis, as well as tRNA charging, transcription, and processing, each saturated P(*HIS4*) transcription through a shared, *GCN4*-dependent mechanism.

Because a hallmark of ISR activation is the phosphorylation of eIF2α, we investigated whether tRNA insufficiency provoked this response. Although CRISPRi knockdown of the amino acid biosynthetic enzyme *ILV2* or the histidyl-tRNA synthetase *HTS1* both induced eIF2α phosphorylation, knockdown of the RNA polymerase III subunit *RPC31* or tRNA processing and maturation factors *POP6* and *SEN15* did not (Fig. 3, C and D, and fig. S4A), consistent with previous observations (*30*). Deletion of *GCN2*, the only known eIF2α kinase in yeast (*32*), completely blocked eIF2α phosphorylation in all conditions tested (fig. S4B). Activation of P(*HIS4*) in the absence of eIF2α phosphorylation further distinguished the response to tRNA depletion upon *RPC31* knockdown and the classical ISR pathway.

We thus tested the genetic requirements for activation of endogenous *HIS4* transcription in response to these CRISPRi-mediated perturbations. Deletion of *GCN4* blocked *HIS4* activation by *RPC31*, *POP6*, and *SEN15* knockdown as well as *ILV2* and *HTS1* knockdown, and so the effects of tRNA insufficiency and defects in tRNA maturation reflected a *GCN4*-mediated ISR. Furthermore, *GCN2* deletion produced distinct effects across these five CRISPRi knockdowns (Fig. 3E). Deletion of *GCN2* completely blocked *HIS4* induction in response to *ILV2* depletion, consistent with canonical models of ISR signaling (*33*). Although *HTS1* knockdown triggered strong eIF2α phosphorylation, *GCN2* deletion only partially abrogated its transcriptional effects; *SEN15*, which did not induce eIF2α phosphorylation, nonetheless showed a weakened response in the deletion as well. Finally, *RPC31* and *POP6* knockdowns induced *HIS4* transcription independent of *GCN2*, as we expected on the basis of the lack of eIF2α phosphorylation and P(*HIS4*) CiBER-seq analysis in a *gcn2* null background (figs. S2D and S4C). Epistatic characterization of additional RNA polymerase III subunits *RPC17* and *RPO31* confirmed that tRNA insufficiency broadly activated P(*HIS4*) by means of a *GCN2*-independent mechanism and that this was not a *RPC31*-specific effect (Fig. 3F). Thus, translation elongation defects arising from impaired tRNA recruitment directly triggered *GCN4*-mediated transcriptional responses. Knockdown of tRNA synthetases activated the ISR by this

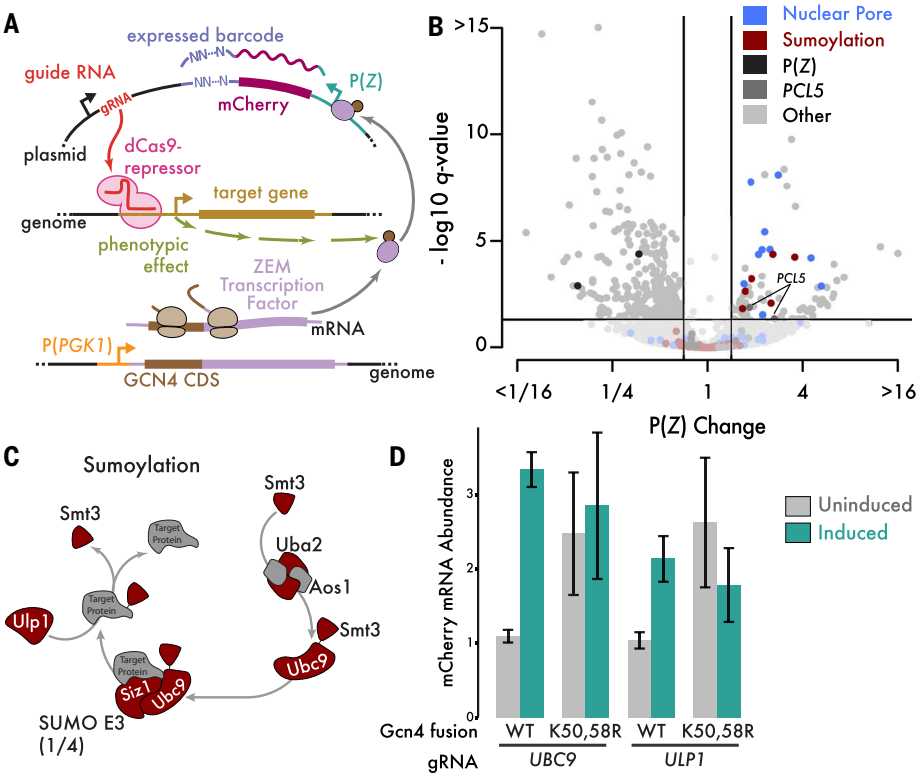


Fig. 5. Disrupting sumoylation enhanced the activity of Gcn4. (**A**) Schematic of indirect CiBER-seq experiment to identify posttranslational regulators of Gcn4. (**B**) CiBER-seq response of P(*Z*) driven by Gcn4-ZEM fusion, as in Fig. 2D. (**C**) Schematic of SUMOylation cycle. Proteins targeted by significant (q < 0.05 and >1.5-fold change) guides from (A) are colored, whereas proteins without a significant guide are gray. (**D**) Measurement of *mCherry* reporter mRNA abundance by qPCR in cells expressing the indicated Gcn4 fusion with the ZEM transcription factor and a gRNA targeting either *UBC9* or *ULP1*.

pathway in parallel with the *GCN2*-dependent response to uncharged tRNAs.

Perturbations of the ARP2/3 complex prevented ISR activation by HTS1 or RPC31 knockdown

Given the fact that both Gcn2 activity and eIF2α phosphorylation were dispensable for ISR activation in response to tRNA depletion, we next sought to systematically identify genes required for this response. We looked for genetic perturbations that modified P(*HIS4*) responses to different ISR triggers by performing dual-guide CiBER-seq—combining one guide that individually activated the ISR with a second guide from our genome-wide CRISPRi library to obtain quantitative, genome-wide genetic interaction profiles. We compared the interaction profiles of the histidyl-tRNA synthetase *HTS1*, which is required for tRNA charging, with those of the RNA polymerase III subunit *RPC31*, which is required for tRNA transcription (Fig. 4, A to C, and tables S8 and S9). We observed saturated P(*HIS4*) induction after knockdown of *HTS1* or *RPC31*, similar to the P(*HIS4*) saturation that we observed upon 3AT treatment (Fig. 3, A and B). Guides that activated the ISR on their own did not further enhance transcription in these dual-guide epistasis experiments. We also observed guides that had no effect on their own but either suppressed ISR activation, such as guides against *GCN4* itself, or enhanced the strength of the response, including guides against the degradation factor *PCL5* or the ISR inhibitor *YIH1* (*34*, *35*).

We compared the genetic interaction profiles across three different ISR stimuli to test whether they were acting through similar or distinct pathways. Epistatic profiles of *HTS1* and *RPC31* knockdown resembled each other (R^2 = 0.76, where R^2 is the coefficient of determination) more closely than either of them resembled the profile during 3AT treatment (R^2 = 0.48), although all three profiles did overlap substantially. This pattern aligned with our observation that P(*HIS4*) activation by either tRNA charging defects or tRNA depletion were not completely *GCN2* dependent, whereas the amino acid starvation response was fully *GCN2* dependent (Fig. 3E). Thus, we looked for epistatic modifiers of P(*HIS4*) activation in response to *HTS1* and *RPC31* knockdown. Using unbiased GO enrichment analysis on the subset of observed guides that blocked ISR-mediated transcriptional activation (tables S8 and S9), we identified an unexpected requirement for the actin cytoskeletal components in this response (Fig. 4). Although knockdown of actin itself (*ACT1*) or genes encoding members of the Arp2/3 complex (fold enrichment >14, q < 0.002 in both dual-guide datasets) did not affect P(*HIS4*) transcription in normal growth or 3AT treatment (fig. S5A), it blocked P(*HIS4*) activation by tRNA charging defects or tRNA depletion (fig. S5, B and C). This could reflect interactions between the ISR inhibitor *YIH1* and free actin monomers (*35*), or it could arise because of nuclear actin's role in transcription (*36*).

Disrupting sumoylation enhanced the activity of Gcn4

CiBER-seq analysis of P(*HIS4*) regulation provided a comprehensive view of yeast ISR signaling, which revealed distinct triggers for *GCN4*-dependent transcriptional responses. However, this approach did not allow us to distinguish the different layers of Gcn4 regulation. Gcn4 abundance is controlled by regulated protein degradation (*37*) in addition to its well-characterized translational induction upon eIF2α phosphorylation (*24*). We observed epistatic enhancement of ISR activation from guides depleting the decay factor *PCL5*. Thus, we wanted to specifically analyze the regulators of Gcn4 stability and translation in isolation and separate them from other effects on P(*HIS4*) activity. To couple these protein-level phenotypes with a transcriptional readout suitable for CiBER-seq, we returned to the synthetic, ZEM chimeric transcription factor used to initially validate barcode sequencing. Because barcode transcription was directly linked to the abundance of this synthetic transcription factor (fig. S1, E and I), we reasoned that it could couple protein-level regulation with expressed RNA barcode abundance (Fig. 5A).

We first profiled posttranslational control of Gcn4 by carrying out CiBER-seq in yeast that constitutively expressed a protein fusion between Gcn4 and the chimeric ZEM transcription factor (Fig. 5A and tables S10 and S11). Levels of this transcription factor should reflect only posttranslational regulation of Gcn4—and not its translational regulation—because it is expressed using the *PGK1* promoter and 5′ untranslated region (5′UTR). Knockdown of *PCL5*, which regulates Gcn4 degradation, affected the activity of the Gcn4-ZEM posttranslational reporter. However, neither tRNA charging nor tRNA biogenesis had an effect. Instead, disruption of sumoylation increased expression from P(*Z*) (Fig. 5, B and C). We identified individually significant gRNAs targeting nearly every step of the sumoylation pathway, including SUMO itself, SUMO-activating and SUMO-conjugating enzymes, and the Siz1 SUMO E3 ligase (*38*). Likewise, we identified guides against most components of the Nup84 subcomplex of the nuclear pore (*39*) that increased Gcn4-ZEM activity (Fig. 5B).

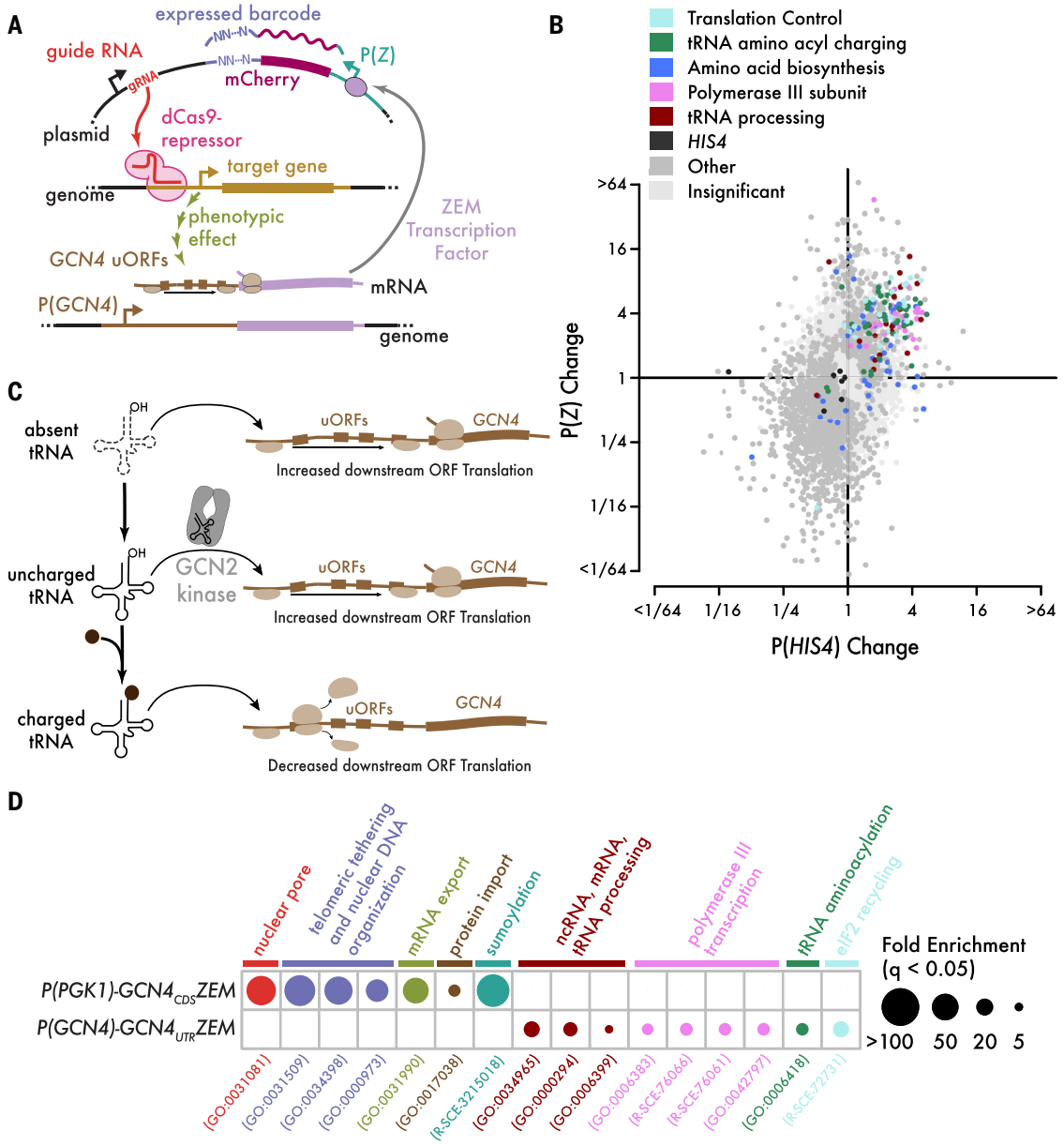


Fig. 6. The GCN4 5′ leader sequence is an intrinsic biosensor of translation stress. (**A**) Schematic of indirect CiBER-seq experiment to identify mediators of *GCN4* 5′ leader translation regulation. (**B**) Comparison of P(*HIS4*) DNA-normalized CiBER-seq profile (Fig. 2D) versus indirect P(*Z*) profile generated with P(*GCN4*)–UTR_{GCN4}–ZEM, with ISR activators colored according to their function, as in Fig. 2D. (**C**) Model of *GCN4* 5′ leader sequence as an intrinsic biosensor of translation stress. Either tRNA insufficiency or uncharged tRNAs increase *GCN4* translation. (**D**) GO analysis of indirect CiBER-seq experiments, as in Fig. 1F. Significant annotations were collected from both the GO biological process complete and reactome pathway annotation datasets. ORF, open reading frame.

Sumoylation of Gcn4 promotes its eviction from chromatin—even when it is recruited through heterologous DNA binding domains—and causes its subsequent degradation. Removal of Gcn4-ZEM from target promoters and degradation of the protein are both expected to limit transactivation of barcoded reporters (*40*, *41*), and knockdown of the SUMO conjugation machinery would relieve these limiting effects and enhance barcode expression. Elevated barcode expression during knockdown of the Ulp1 deconjugating protease could result from its role in SUMO maturation (*42*) or reflect a more-complex requirement for a sumoylation-desumoylation cycle in Gcn4 regulation. Our Gcn4-ZEM fusion did retain the Gcn4 DNA binding domain, in addition to the heterologous Zif268 DNA binding domain, which could also contribute to the overall profile of responses that we observed.

The sites of SUMO modification on Gcn4 are known, and unSUMOylatable mutants have been reported. To directly test the role of SUMOylation in regulation of Gcn4p-mediated transcription, we generated a transcription factor fusion with the unSUMOylatable Gcn4 mutation (*40*) and measured its activity (Fig. 5D). The unSUMOylatable mutant caused higher baseline P(*Z*) expression that was not further increased upon SUMO cascade knockdown (Fig. 5D). Thus, Gcn4-specific SUMOylation is reducing the activity of the Gcn4-ZEM fusion, and this reduction is relieved when SUMOylation is impaired globally or blocked by Gcn4-specific mutations. The role of the nuclear pore, although less clear, may be linked to sumoylation. In yeast, Ulp1 binds physically to the nuclear pore (*43*, *44*), mutations in the Nup84 complex can mimic some *ulp1* phenotypes (*45*, *46*), and Ulp1 at the

nuclear pore has been implicated directly in transcriptional activation (*47*). Alternately, this effect may be specific to the Gcn4-ZEM fusion, although it did not arise in other CiBER-seq profiles based on the ZEM transcription factor, as described below.

The GCN4 5′ leader sequence is an intrinsic biosensor of translation stress

We observed clear and coherent patterns of genetic perturbation affecting Gcn4 protein activity, but we saw no evidence that tRNA insufficiency affected its posttranslational regulation. We next explored the translational control of *GCN4*, which results from regulatory upstream open reading frames in the *GCN4* 5′ leader sequence (*24*). To capture perturbations that affect this translation regulation, we drove expression of the ZEM synthetic transcription factor from the *GCN4* promoter and 5′ leader sequence and assessed guide effects on barcode expression (Fig. 6A).

The same guides that activated P(*HIS4*), including those that block tRNA biogenesis as well as tRNA charging, also increased P(*Z*) in our CiBER-seq analysis of translational control through the *GCN4* 5′ leader (>32-fold, hypergeometric $P < 2.04 \times 10^{-156}$) (Fig. 6B, fig. S6A, and tables S12 and S13). By assessing this translational response in isolation, we observed that effects occur in the same direction but with larger magnitude than P(*HIS4*). Thus, translational regulation is central to ISR activation, even when it proceeds independently of *GCN2* and eIF2α phosphorylation (Fig. 6C). To exclude transcriptional effects mediated by the *GCN4* promoter (*48*), we repeated this CiBER-seq profiling experiment with a version of the ZEM transcription factor driven from the P(*PGK1*) promoter, but still containing the *GCN4* 5′ leader (fig. S6B). This profile enriched most of the same functional categories as the profile using the *HIS4* promoter, including tRNA transcription and processing (fig. S6C). We observed an overall enrichment of shared activators between the two profiles (>11-fold, hypergeometric $P < 1.62 \times 10^{-61}$) and specific enrichment in tRNA biogenesis factors (>100-fold, hypergeometric $P < 8.88 \times 10^{-59}$).

Discussion

By linking CRISPRi guides with barcoded expression reporters, we generated quantitative, genome-wide phenotypic profiles for specific molecular events in the yeast cell. Our CiBER-seq approach allowed us to address transcriptional, translational, and posttranslational regulation, enabling the systematic genetic analysis of diverse biological processes. We characterized distinctive CiBER-seq profiles for five different promoters, each of which could be understood in light of the function of the associated gene. We leveraged these quantitative phenotypic profiles to gain insights into the ISR, identifying tRNA depletion as a previously unappreciated trigger for this well-characterized pathway.

Because translation is a resource-intensive biosynthetic process, most organisms sense translational stresses and respond with physiological changes that maintain homeostasis. In eukaryotes, amino acid starvation, which directly affects tRNA charging and translation elongation, triggers a global decrease in translation initiation while also increasing the transcription of amino acid biosynthesis genes (*24*). In many cases, reduced initiation is sufficient to restore normal elongation profiles (*49*). In fact, when aminoacyl-tRNA synthetases are depleted, yeast reduces translation initiation until tRNA charging and utilization are balanced. Furthermore, in normal circumstances, these synthetases appear to buffer tRNA levels by sequestering uncharged tRNAs (*50*). Here, we found that the opposite situation—tRNA depletion in the presence of adequate synthetases and amino acids—triggered a *GCN2*-independent ISR. Our comprehensive genetic data, implicating many stages of tRNA biogenesis, suggest that *GCN4* translation responds directly to elongator tRNA insufficiency, rather than to depletion of initiator tRNA (*30*) or the accumulation of unprocessed nuclear tRNA precursors (*51*). Because we have not identified other trans-acting regulatory pathways under these circumstances, we propose that the *GCN4* 5′ leader is an intrinsic biosensor for translation elongation stress (Fig. 6C). Artificial stimuli, including tRNA depletion, can activate *GCN4* translation independent of eIF2α phosphorylation. In the natural history of budding yeast, the ISR likely evolved to sense nitrogen starvation and elicit general inhibition of translation through *GCN2*, along with *GCN4*-dependent activation of homeostatic transcriptional programs (*24*). Our work expands the range of stresses known to activate the ISR and points to a general mechanism for sensing translational perturbations.

Although we showed that the ISR is activated by artificial, genetic depletion of tRNAs, similar situations may arise in a natural context (*52*). Starvation and other stresses can induce tRNA cleavage (*53*, *54*), and although much attention has focused on the positive roles of the resulting fragments (*55*), cleavage can also deplete the tRNA substrate and alter translation (*56*). Effective tRNA depletion may also arise when tRNAs are sequestered in the nucleus and thus unavailable for cytosolic translation (*28*). Finally, an array of human disease mutations that affect tRNA biogenesis factors lead to neurodegenerative disorders, although the molecular basis for this effect is not clear (*52*, *57*). Mutations in tRNA genes themselves can lead to tRNA insufficiency, ISR activation, and neurodegeneration, although in mammals this effect is *GCN2* dependent (*58*, *59*).

Our results demonstrate the power of CiBER-seq in elucidating the genetic architecture of regulation in the cell. We observed distinct patterns of response for different promoters and comprehensive coverage of pathways and molecular complexes by guides targeting each individual component, which indicate that our approach is both specific and sensitive. We identified ISR activation upon knockdown of each subunit of RNA polymerase III along with most proteins involved in tRNA processing, which reveals tRNA depletion as the underlying trigger for the ISR. We identified other modes of regulation as well, such as sumoylation of Gcn4, and even detected the compensatory P(*PGK1*) activation in response to impaired glycolysis. Quantitative CiBER-seq profiles enable genetic interaction analysis that provides further insight into regulatory networks. Correlated patterns of epistasis are a powerful tool for identifying genes that function together in pathways and complexes, and likewise chemogenomic comparison between chemical and genetic interaction profiles can reveal functional drug targets (*11*). Here, we leveraged this advantage of CiBER-seq to identify a role for actin cytoskeletal components in ISR signaling. More broadly, our results highlight the power of CiBER-seq to combine specific and quantitative molecular phenotypes with targeted genetic perturbations and thereby precisely dissect regulatory pathways. A similar Reporter-seq technique has been developed simultaneously with this work, and it has been applied to elucidate known and unknown pathways linking diverse stressors to the yeast heat shock response (*60*). Because the key components of CiBER-seq translate into nearly any organism, we anticipate many biological insights will arise from broad application of our approach.

Materials and methods

Plasmid materials

pRS416-dCas9-Mxi1 + TetR + pRPR1(TetO)-NotI-gRNA was a gift from R. Davis (Addgene plasmid #73796; http://addgene.org/73796; RRID: Addgene_73796). pKT0139 was a gift from K. Thorn (Addgene plasmid #8731; http://addgene.org/8731; RRID: Addgene_8731). pHES836 was a gift from H. El-Samad (Addgene plasmid #89195; http://addgene.org/89195; RRID: Addgene_89195). pHES795 was a gift from H. El-Samad (Addgene plasmid #87943; http://addgene.org/87943; RRID: Addgene_87943). pCfB2337 was a gift from I. Borodina (Addgene plasmid #67555; http://addgene.org/67555; RRID: Addgene_67555). pCfB2226 was a gift from I. Borodina (Addgene plasmid #67533; http://addgene.org/67533; RRID: Addgene_67533). pCfB2189 was a gift from I. Borodina (Addgene plasmid #67532; http://addgene.org/67532; RRID: Addgene_67532). Plasmids used in this study are listed in table S1.

Plasmid construction

Oligonucleotides used in this study are listed in table S2. Plasmid assembly was carried out using standard molecular biology techniques as described below, and verified by Sanger sequencing. All polymerase chain reaction (PCR) reactions were performed using Q5 polymerase (NEB M0491S) according to manufacturer protocols. Restriction enzymes were obtained from NEB and high-fidelity (HF) variants were used when available. DNA fragments were size-selected by gel electrophoresis in 1% agarose. Gibson assembly reactions were carried out using NEBuilder HiFi DNA Assembly Master Mix (NEB E2621L) with DNA fragments containing homology arms between 15 and 20 base pairs (bp) long. DNA was purified and concentrated as necessary with DNA Clean & Concentrator (Zymo D4013). Transformations were performed in Stbl3 chemically competent cells provided by the QB3 Berkeley Macrolab facility and plated on appropriate antibiotic plates for colony selection. Plasmid DNA was purified from liquid cultures using DNA miniprep kit (NEB #T1010).

pNTI660 was constructed in several steps. First, pNTI601 (pRS416-dCas9-Mxi1 + TetR + pRPR1(TetO)-NotI-gRNA, AddGene #73796) (*17*) was amplified with NM717 and NM724 (*16*) to isolate the gRNA expression cassette and the resulting PCR product was recircularized by Gibson assembly to create pNTI646. pNTI646 was then digested with MfeI and combined with oligonucleotide NI1025 to create pNTI660.

pNTI725 was created in several steps. First, the promoter P(*PGK1*) was amplified from *Saccharomyces cerevisiae* BY4741 genomic DNA using NI553 and NI554, and the resulting product was subcloned and used as a template for P(*PGK1*) amplification. Yeast-optimized yECitrine coding sequence was amplified from pNTI189 (pKT0139 AddGene #8731) (*61*) using primers RM151 and RM155 and cloned downstream of P(*PGK1*). Next, the P(*PGK1*)-yECitrine expression fragment was isolated by digestion with SacI and BsrgI restriction sites, and pNTI660 was linearized by digestion with MfeI. Double-stranded DNA splints were created by annealing oligonucleotides RM317 and RM318 and oligonucleotides RM319 and RM320. Digested P(*PGK1*)-yECitrine and pNTI660 were joined together using these splints in a four-piece Gibson assembly reaction. The NotI site for gRNA insertion was replaced with an NruI site by digesting the backbone with NotI and then introducing single-stranded oligonucleotide RM321 by Gibson assembly. Finally, the Illumina TruSeq Read1 site was added to the beginning of the terminator T(*ADH1*) by annealing single-stranded oligonucleotides RM323 and RM324, extending them using Q5 polymerase, digesting the backbone with AscI and introducing the double-stranded product into the digested plasmid by Gibson assembly.

pNTI726 was constructed by replacing P (*PGK1*) in pNTI725 with estradiol-responsive P(*GAL1*). P(*GAL1*) with an embedded Zif268 binding site was PCR amplified from pHES836 using primers RM348 and RM349. pNTI725 was digested with SacI and EcoRI and amplified P(*GAL1*) was then Gibson assembled into the backbone.

pNTI727 was generated by inserting the gRNA sequence targeting P(*Z*) into pNTI726. Plasmid pNTI726 was digested with NruI, single-stranded oligonucleotides RM389 and RM390 were annealed, extended using Q5 DNA polymerase, and the double-stranded product was introduced into the linearized plasmid by Gibson assembly.

pNTI728 was generated by inserting the gRNA sequence targeting P(*ADH1*) into pNTI726, using the same methods as pNTI727, using oligonucleotides RM383 and RM384.

pNTI729 was constructed by first amplifying the ZEM transcription factor from pNTI638 (pHES795 Addgene #87943) (*62*) in two pieces, using oligonucleotide primers RM352 with RM353 and RM354 with RM355. The easy clone vector pCfB2337 (Addgene #67555) (*63*) was then digested with HindIII and the three fragments were combined by Gibson assembly.

pNTI730 was constructed by replacing the promoter P(*ADH1*) in pNTI729 with the P (*PGK1*) promoter. pNTI729 was digested with PvuII and NheI, P(*PGK1*) was amplified from pNTI725 using primers RM417 and RM418, and the two fragments were combined by Gibson assembly.

pNTI731 through pNTI737, expressing gRNAs against *ILV2*, *HTS1*, *RPC31*, *NUP133*, *NUP145*, *ULP1*, and *UBC9*, respectively, were subcloned into the integrating plasmid pCfB2189. First, DNA fragments containing the gRNA expression cassette were amplified from the pool of gRNA plasmids using forward primers RM639, RM517, RM518, and RM641 through 644, respectively, that anneal to the specific nucleotide barcode sequence, and a reverse primer RM519 that binds downstream of the gRNA scaffold. Fragments were amplified with Gibson homology arms on each end. The integration vector pCfB2189 (Addgene #67532) (*63*) was digested with HindIII and gRNA-expressing fragments were transferred into this backbone by Gibson assembly.

pNTI738 for initiator methionyl-tRNA overexpression was constructed by amplifying the initiator methionyl-tRNA locus *IMT4* from yeast genomic DNA using primers RM636 and RM637, digesting the high-copy yeast plasmid pRS426 (ATCC #77107) (*64*) with EcoRI, and combining them by Gibson assembly.

pNTI739 was constructed from pNTI730. The *GCN4* coding sequence (CDS) was amplified from yeast genomic DNA using primers RM515 and RM516 and inserted upstream of, and in frame with, the synthetic transcription factor by digesting pNTI730 with NheI and introducing *GCN4* by Gibson assembly.

pNTI740 was constructed by digesting NTI730 with PvuII and NheI, amplifying the *GCN4* promoter and 5′UTR from yeast genomic DNA using primers RM513 and RM514, and combining the two fragments by Gibson assembly.

pNTI741, the P(*HIS4*)-yECitrine PCR template used for downstream CiBER-seq plasmid library preparation, was generated by replacing the P(*PGK1*) in pNTI725 with P(*HIS4*). pNTI725 was digested with SacI and NheI, P(*HIS4*) was amplified from yeast genomic DNA using primers RM499 and RM500, and the two fragments were combined by Gibson assembly.

pNTI742 was made in several steps. First, the NotI restriction site of pNTI660 was replaced with an AvrII restriction site by digesting with NotI and introducing a replacement cassette, composed of annealed oligonucleotides RM459 and RM460, by Gibson assembly. Next, the vector was digested with SacI and MfeI. The T (*ADH1*) terminator and TruSeq Read1 priming site were amplified from pNTI726 using primers RM489, RM490, and RM491, in a nested PCR approach where RM490 was used at 0.1 times the normal concentration. The resulting PCR product was joined with the digested vector by Gibson assembly.

pNTI743 was made in several steps. First, the yeast *ACT1* 3′UTR and transcription termination sequence was PCR amplified from yeast genomic DNA using RM633 and RM634. pNTI742 was linearized with BamHI and the amplified PCR product was inserted by Gibson assembly. This intermediate vector was then digested with HindIII and AflII, and oligo RM 719 was inserted by Gibson assembly.

pNTI744 through pNTI750, expressing gRNAs against *POP6*, *SEN15*, *RPC17*, *RPO31*, *MLC1*, *CDC10*, and *CHS2*, respectively, were subcloned into the integrating plasmid pCfB2189. First, DNA fragments containing the gRNA expression cassette were amplified from the pool of gRNA plasmids using forward primers RM722-725, RM813, RM814, and RM817, respectively, that anneal to the specific nucleotide barcode sequence, and a reverse primer RM519 that binds downstream of the gRNA scaffold. Fragments were amplified with Gibson homology arms on each end. The integration vector pCfB2189 (Addgene #67532) (*63*) was digested with HindIII and gRNA-expressing fragments were transferred into this backbone by Gibson assembly.

pNTI751 was constructed by three-piece Gibson of the following DNA fragments. First, P(*PGK1*) was PCR amplified from pNTI725 using primers RM726 and RM727. Second, the GCN4 5′UTR was amplified from pNTI740 using primers RM728 and RM729. Third, pNTI740 was digested with PvuII and HindIII and the 8600-bp fragment was gel extracted.

pNTI752 was constructed by first digesting pNTI739 with NheI and SapI, and gel extracting the 10-kb fragment. The K50,58R mutation was introduced into the *GCN4* CDS by PCR amplifying the CDS from pNTI739 using primers RM730, RM731, and RM732 (731 was used at 0.1 times the concentration of the other two). pNTI752 was then assembled by three-piece Gibson of the linearized backbone, PCR fragment, and RM733 splint oligo.

pNTI753 was constructed in a several steps. First, P(*PGK1*) was subcloned into pNTI742. This plasmid was then linearized with SpeI and gel extracted. P(*HIS4*) was PCR amplified from yeast genomic DNA with primers RM739 and RM740 and mCherry was PCR amplified from a subcloned plasmid template using RM741 and RM742. The three resulting fragments were Gibson assembled.

pNTI754 through pNTI756 were cloned using a similar method, by first digesting pNTI753 with SpeI and BamHI and gel extracting the 7-kb backbone fragment. The following promoters were then amplified from yeast genomic DNA as follows: (i) P(*MET6*) with RM745 and RM746; (ii) P(*CWP1*) with RM747 and RM748; and (iii) P(*PHO5*) with RM749 and RM750. Each PCR-amplified promoter was then assembled into the linearized backbone by Gibson assembly.

pNTI757 through pNTI760 were all cloned using a similar method, by first PCR amplifying respective divergent promoters using the following combinations of PCR template primers: (i) pNTI753 with RM789 and RM 790; (ii) pNTI754 with RM793 and RM794; (iii) pNTI755 with RM795 and RM796; and (iv) pNTI756 with RM797 and RM798. Respective PCR products were then TOPO cloned using the TOPO Blunt system (Thermo Fisher Scientific #451245) according to manufacturer protocol.

Plasmid library construction

All pooled plasmid libraries were constructed by Gibson-style assembly using HiFi DNA assembly master mix (NEB E2621L) and transformed using either ElectroMAX DH10B (ThermoFisher #18290015) or NEB 10-beta Competent *Escherichia coli* (NEB #C3019H) according to manufacturer protocol. All column purifications were performed using Zymo DNA Clean and Concentrator-5 (Zymo #D4013). Dilutions of reach transformation were plated to estimate library size and ensure sufficiency library diversity. Plasmid libraries were harvested from liquid culture using the Monarch Plasmid DNA Miniprep Kit (NEB #T1010). Miniprep kit reagents were scaled to accommodate one spin column for every 5 ml of culture.

To generate estradiol-inducible barcode-expressed plasmid libraries (fig. S1), pNTI726 and pNTI728 were first digested with BamHI. Five barcodes were generated by annealing primer RM396 with one of five barcode primers, RM391, RM392, RM393, RM394, or RM399, and extending the annealed duplex using Q5 polymerase. Four of these barcodes were inserted by Gibson assembly into BamHI-linearized pNTI726 and the last barcode DNA fragment inserted into BamHI-linearized pNTI728. Gibson reactions were purified and concentrated, transformed into ElectroMAX DH10B cells, and inoculated directly into 200 ml of lysogeny broth (LB)–carbenicillin media. Pooled libraries were harvested when batch liquid culture reached an optical density (OD) of 4.

To generate gRNA plasmid libraries, 100 pg of CRISPRi gRNA oligonucleotide library was amplified by PCR using Q5 polymerase using primers NM636 and NM637 (*16*). pNTI742 was digested with AvrII and the amplified guide fragments were introduced into the digested vector by Gibson assembly. The assembly product was purified and transformed into ElectroMAX DH10B cells by electroporation. Electroporated cells were inoculated directly into 500 ml of LB-carbenicillin media and harvested at an OD of 2. Serial dilutions of the initial transformation were plated to ensure sufficient library diversity (>50× coverage of 60,000 guides). Plasmid library was purified using NEB miniprep kit and pooled library was analyzed by Sanger sequencing.

Barcodes were then added to the gRNA plasmid library, targeting an average of four barcodes per gRNA (240,000 barcodes in total). First, barcode DNA fragments were generated by annealing oligos RM504 and RM505 and extending the duplex using Q5 polymerase. The following thermocycler conditions were used to avoid amplification of any specific barcode sequence. Initial denaturation at 98°C for 45 s, six cycles of annealing and extension at 68°C for 15 s and 72°C for 5 s, respectively, and finally a 12°C hold. The barcode fragments were digested with BamHI and MfeI to eliminate barcode fragments that contain these restriction sites and then purified using a DNA Clean and Concentrator-5. The library of gRNA-expressing plasmids was then digested with AflII and the barcode DNA fragments were introduced by Gibson assembly. The assembly reaction was purified using a DNA Clean and Concentrator-5 then transformed into NEB 10-beta Competent *E. coli* (High Efficiency) according to manufacturer protocol. Transformation dilutions were plated to estimate total number of transformants, and in parallel inoculated at several dilutions directly into LB-carbenicillin media. Liquid culture corresponding to roughly 240,000 transformants was harvested at an OD of 2 and plasmid library was extracted using NEB miniprep kit. Reagents were scaled to accommodate one spin column for every 5 ml of culture. This barcoded gRNA library was paired-end sequenced (see Barcode to gRNA assignment section) to assign barcodes with gRNAs.

P(*HIS4*)-yECitrine was amplified from pNTI741 using primers RM501 and RM502, P(*PGK1*)-yECitrine was amplified from pNTI725 using primers RM501 and RM503, and two versions of estradiol-inducible P(*Z*)-mCherry with specific nucleotide identifiers were amplified with either RM522 and RM524 or RM523 and RM524. The barcoded gRNA plasmid library was digested with BamHI and the four PCR products were introduced into the digested library by Gibson assembly to create the four CiBER-seq libraries. These libraries were then transformed into ElectroMAX DH10B cells by electroporation. Electroporated cells were inoculated directly into 500 ml of LB-carbenicillin media and harvested at an OD of 2. Serial dilutions of the initial transformation were plated to ensure sufficient library diversity (>50× coverage of 60,000 guides). Plasmid library was purified using NEB miniprep kit and pooled library was analyzed by Sanger sequencing.

CiBER-seq plasmid libraries for studying P(*MET6*), P(*CWP1*), P(*PHO5*), and P(*HIS4*) regulation (Fig. 1) were constructed in a similar fashion, but with key differences. pNTI743 containing divergent terminating sequences was used as the parent vector into which the gRNA library was inserted at the AvrII restriction site. Dual barcodes were next inserted into this gRNA plasmid library by digesting the plasmid library with AscI, annealing and extending RM720 and RM721, and Gibson assembling. Finally, the divergent promoters containing P(*PGK1*)-citrine normalizer and P(*query*)-mCherry expression cassette were inserted in between the two barcodes by: (i) digesting pNTI757 through pNTI760 with BciVI and gel extracting divergent promoter template; (ii) digesting dual barcoded gRNA plasmid library with AscI which cuts in between the two barcodes; and (iii) Gibson assembling each divergent promoter into the barcoded gRNA plasmid library and separately transforming and preparing these plasmid libraries as outlined above.

Construction of yeast strains

All yeast strains were constructed by transforming PCR products or plasmids linearized by restriction enzyme digestion into yeast by standard lithium acetate transformation (*65*). Selected, clonal transformants were confirmed by two genotyping PCRs, using primer pairs that amplify across the junction on either side of the integration. Genotyping reactions were performed with OneTaq Hotstart Polymerase (NEB #M0481S).

Yeast strains used in this study are listed in table S3. Strains with *gcn2*Δ or *gcn4*Δ genotype were constructed by first amplifying the hygromycin resistance cassette from pNTI729 by PCR, with flanking sequence homologous to the 5′ and 3′ ends of the *GCN2* or *GCN4* coding sequence, using primers RM579 through

RM582 for *gcn2*Δ and primers RM583 through RM586 for *gcn4*Δ. Primers RM589 and RM590 were used to genotype the upstream junction for *gcn2*Δ, whereas RM591 and RM592 were used for the downstream junction. Primers RM593 and RM590 were used to genotype the upstream junction for *gcn4*Δ, whereas RM591 and RM594 were used for the downstream junction.

gRNA expression cassettes were integrated using vectors from the EasyClone 2.0 toolkit for yeast genomic integration (*63*). Plasmids using backbone pCfB2189 were integrated by digesting with NotI, transforming this digestion, and selecting transformants on synthetic complete drop-out (SCD) – Leu plates. Plasmids using backbone pCfB2337 were integrated by digesting with NotI, transforming this digestion, plating transformations on nonselective yeast extract peptone dextrose (YEPD) media overnight, and then replica plating on hygromycin plates. Histidine prototrophy was restored by digesting pCfB2226 with NotI, transforming into the appropriate strain, and selecting transformants on SCD – His plates. Plasmid pNTI647, expressing dCas9 and tetR, was integrated as described in (*16*). Integration of gRNA expression cassettes was validated by amplifying across the upstream junction with RM527 and RM373 and across the downstream junction with RM374 and RM528. Integration of synthetic transcription factor variants were genotyped by amplifying across the upstream junction with RM372 and RM373 and across the downstream junction with RM374 and RM375.

Media

LB media was prepared by dissolving LB medium capsules (MP Biomedicals #3002-31) in ultrapure water and sterilized by autoclaving, according to manufacturer instructions. LB media was supplemented with 50 μg/ml carbenicillin (Sigma #C1389) for antibiotic selection. YEPD was prepared by dissolving yeast extract (RPI # Y20026) and peptone (RPI #20241) in ultrapure water, sterilizing by autoclaving, then supplementing with 2% final concentration of sterilized dextrose (Fisher Chemical #D16-500). SCD media was prepared by dissolving yeast nitrogen base (BD Difco #291940), dextrose to 2% final concentration, and the appropriate drop-out mix in ultrapure water, then sterilizing by filtration. SCD – Leu drop-out media was prepared using drop-out mix synthetic minus leucine (US Biological #D9525), SCD – Ura drop-out media was prepared using drop-out mix synthetic minus uracil (US Biological #D9535).

Quantitative reverse transcription PCR (qRT-PCR)

RNA was isolated from yeast by acid phenol extraction (*66*). Reverse transcription was carried out using Protoscript II reverse transcriptase (NEB #M0368L) with oligo-dT priming according to manufacturer instructions. qRT-PCR was carried out using DyNAmo qPCR mastermix (ThermoFisher #F410L) according to manufacturer protocols, using a CFX96 Touch Real-Time PCR Detection System. Primers, provided in table S4, were designed using Primer Blast and validated by using a cDNA dilution standard curve. Oligonucleotide sequence for RT-PCR are listed in table S5.

Fluorescence measurements

Expression of yECitrine, under the control of the ZEM transcription factor, (Fig. 1) was monitored using a 96-well plate reader (Tecan SPARK Multimode Microplate Reader). Overnight cultures were back-diluted in 96-well round bottom plates to OD 0.05 in SCD – Ura selective media containing beta-estradiol and/or anhydrotetracycline as indicated. Fluorescence (excitation at 516 nm and emission at 540 nm) and OD_{600} were measured in triplicate every 15 min.

Turbidostat continuous culture

Yeast populations transformed with plasmid libraries were inoculated into a custom turbidostat (*67*) and maintained in SCD – Ura media at a density corresponding to an optical density 600 nm (OD_{600}) of roughly 0.8. When growth rate reached a steady state (corresponding to a doubling time of roughly 90 min), a 50-ml preinduction sample was collected. gRNA expression was then induced by injecting concentrated anyhydrotetracycline into both the growth chamber and turbidostat media reservoir to obtain a final concentration of 250 ng/ml of anyhydrotetracycline. Six doublings later (9 hours) a 50-ml postinduction sample was collected. For 3AT treatment, samples were induced by injecting concentrated 3AT into both the growth chamber and turbidostat media reservoir to obtain a final concentration of 90 mM 3AT. Two hours later, a postinduction sample was collected. For CiBER-seq experiments involving the ZEM transcription factor, media was supplemented with 8 nM beta-estradiol throughout the course of the experiment. Collected samples were pelleted by centrifugation at 4000 × *g* for 5 min, the media was aspirated, and the pellets were stored at −80°C.

Barcode sequencing

All PCR reactions were performed using Q5 polymerase (NEB M0491S) according to manufacturer protocols. PCR cycle numbers were adjusted as needed (between 6 and 12 cycles) to obtain adequate concentration of product for sequencing while avoiding overamplification. DNA was purified using DNA Clean & Concentrator (Zymo #D4013) according to the manufacturer instructions. When necessary, AMPure XP beads (Beckman Coulter #A63881) were used to purify full-length DNA product. Size distributions and concentrations of the sequencing libraries were measured before sequencing using an Agilent TapeStation 2200 and High Sensitivity D1000 ScreenTape. To ensure proper clustering on the Illumina HiSeq4000, CiBER-seq libraries were either pooled with RNA sequencing libraries and standard 2% PhiX, or individually with 10% PhiX. PCR products were pooled and sequenced on HISeq4000 SR50 (Vincent J. Coates sequencing facility). Adapter sequences for each library are listed in table S6, and sequencing data accession numbers are available in table S7.

DNA library preparation

Plasmid DNA was extracted from yeast pellets using Zymo yeast miniprep kit (Zymo #D2004) according to manufacturer protocol, except as described below. Reagent volumes were scaled to accommodate roughly 250 million cells, as yeast pellets included 25 OD_{600} ml of yeast. Zymolyase concentration was increased to 1 μl for every 10 million cells and the digestion time was doubled to ensure complete cell wall digestion.

Extracted barcode expression plasmid libraries derived from pNTI726 and pNTI728 backbones (Fig. 1) were prepared by PCR amplification of barcode sequences with primers that incorporated flanking TruSeq adapters. Barcode sequences were first amplified by PCR with RM411 and an i5 primer from NEBNext Multiplex Oligos (NEB #E7600S) for 12 cycles. PCR products were purified and concentrated using a DNA Clean and Concentrator-5, and then used as the template for an additional six cycles of PCR amplification with NEBNext i5 and i7 primers (NEB #E7600S).

To achieve a more linear amplification, all other CiBER-seq DNA libraries were prepared using an initial linear amplification by in vitro transcription. Extracted plasmid was linearized by restriction digestion with MfeI for P(*PGK1*) and P(*HIS4*) plasmid libraries (Figs. 2 to 4) or PvuII for P(*Z*) libraries (Figs. 5 and 6). The digestion product was purified using a DNA Clean and Concentrator-5, and used as the template for an IVT reaction using T7 HiScribe (NEB #E2040S) according to the manufacturer protocol for short transcripts, with an overnight incubation at 37°C. The reaction was then treated with DNase I to remove template DNA and RNA product was purified using an RNA Clean and Concentrator-5 (Zymo #R1016). Purified RNA was used as the template for reverse transcription using ProtoScript II according to manufacturer protocol, using sequence-specific primer RM511 for *P*(*PGK1*) and *P*(*HIS4*) plasmid libraries (Figs. 2 to 4) or RM546 for *P*(*Z*) libraries (Figs. 5 and 6). Reverse transcription cDNA product was treated for 30 min with 0.5 μl RNAse A (ThermoFisher #EN0531) and 0.5 μl RNAse H (NEB #M0297S) to remove RNA, and then DNA was purified using a DNA Clean and Concentrator-5 column. The sequence-specific

reverse transcription primers RM511 and RM546 incorporate the i7 priming site, and the i5 priming site is included in the plasmid. The final library was generated using eight cycles of PCR amplification using NEBNext i5 and i7 dual index primers, and purified cDNA product as template. The resulting product was purified using a DNA Clean and Concentrator-5 and submitted for Illumina high-throughput sequencing.

RNA library preparation

Total RNA was harvested from yeast pellets as described above for quantitative PCR (qPCR) analysis. Extracted RNA from libraries using the pNTI726 or pNTI728 backbones (Fig. 1) was used as the template for reverse transcription by ProtoScript II using sequence-specific primer RM411, according to the manufacturer instructions. Reverse transcription product was purified using a DNA Clean and Concentrator-5 and used as the template for PCR amplification using i5 and i7 dual indexing primers (NEB #E7600S).

In CiBER-seq libraries, the barcode is present in the opposite orientation. Reverse transcription was carried out using ProtoScript II according to the manufacturer protocol for oligo-dT priming. Reverse transcription product was treated with 0.5 μl of RNAse A and 0.5 μl of RNAse H for 30 min to remove RNA, and the DNA product was purified using a DNA Clean and Concentrator-5. The resulting cDNA product was first amplified by PCR for six cycles using primers RM511 and RM512 for isolated P(*HIS4*) and P(*PGK1*) libraries, or primers RM512 and RM546 for P(*Z*) libraries, primers RM511 and RM810 for P(*PGK1*) in dual barcoded CiBER-seq, and primers RM810 and RM812 for P(*query*) in dual barcoded CiBER-seq. The PCR products were then purified and subsequently amplified in eight cycles of PCR using NEBNext i5 and i7 dual indexing primers. PCR products were purified using a DNA Clean and Concentrator-5 and submitted for Illumina high-throughput sequencing.

Barcode to gRNA assignment

The sequencing library for paired-end barcode-to-gRNA assignment sequencing was constructed by PCR amplification from the barcoded gRNA library using primers RM506 and RM509 for six cycles using Q5 polymerase according to the manufacturer instructions. The resulting PCR product was purified using a DNA Clean and Concentrator-5 and then used as a template for a second, 15-cycle PCR using NEBNext i5 and i7 dual indexing primers. The PCR product was purified using a DNA Clean and Concentrator-5 and analyzed using an Agilent TapeStation 2200 and High Sensitivity D1000 ScreenTape. The PCR product library was then sequenced using 150 + 150 base paired–end sequencing on an Illumina MiSeq. Barcode to gRNA assignment sequencing data are available under accession SRR10327353.

The PacBio sequencing library for dual barcodes to gRNA assignment was prepared without PCR amplification to avoid sequencing errors associated with the rolling circle polymerase and to maintain sequence diversity. First, 5 μg of dual barcoded gRNA plasmid library was digested with HindIII and PmeI and another 5 μg digested with BglII and XhoI. The ~1200-bp fragment for each digestion, containing dual barcode and gRNA were gel extracted and prepared for sequencing according to SMRTbell Express Template Prep Kit 2.0 (#100-938-900) according to manufacturer instructions.

Single gRNA induction

Yeast strains expressing individual gRNAs were grown overnight in YEPD, then diluted to OD_{600} of 0.01 in – Leu media containing either 250 ng/ml anhydrotetracycline, or no inducer. Yeast were harvested 12 hours later (after roughly 6 doublings), pelleted at 4000 × *g* for 5 min, and pellets were frozen and stored at −80°C for subsequent RNA extraction or protein isolation.

Immunoblotting

Total protein was isolated from frozen yeast pellets by resuspending in 1.5 ml of 5% trichloroacetic acid and incubated at 4°C for <10 min. Protein was pelleted by centrifugation at 16,000 × *g* for 2 min and washed in 0.5 ml of acetone, vortexed briefly, and collected by centrifugation for 5 min at 16,000 × *g*. Protein pellets were washed once more in 1 ml of acetone, vortexed, and collected by centrifugation for 5 min at 16,000 × *g*. Pellets were dried overnight and then resuspended in 100 μl of freshly prepared protein breakage buffer [50 mM Tris•HCl pH 7.5, 1 mM EDTA, and 3 mM dithiothreitol (DTT)] containing 100 μl of glass beads (BioSpec 11079105). Samples were vortexed in five cycles of 1-min vortexing followed by 1 min on ice. Lysates were transferred to new tubes and collected by centrifugation at 16,000 × *g*, then resuspended in 150 μl resuspension buffer (100 mM Tris•HCl pH 11, 3% SDS). Samples were boiled for 5 min then allowed to cool and pelleted at 16,000 × *g* for 30 s. Then, 120 μl of lysate was transferred to a new tube, and a bicinchoninic acid (BCA) assay was performed to determine protein abundance. Equal amounts of total protein were loaded on a 4 to 12% polyacrylamide Bis-Tris gel (Thermo Scientific #NW04120BOX) and separated by electrophoresis in MOPS buffer at 200 V in a Bolt gel tank (Thermo Scientific #A25977) according to manufacturer instructions. Protein was then transferred to a nitrocellulose membrane (Thermo Scientific #88018) according to manufacturer guidelines. Membranes were blocked for 1 hour in tris-buffered saline with Tween 20 (TBST) with 5% milk, washed with TBST three times for 10 min of shaking, and incubated with primary antibodies in TBST plus 0.5% milk overnight. Hexokinase loading control was probed with rabbit anti-Hxk2 (Rockland #100-4159) and phosphorylated eIF2α was detected using Phospho-eIF2a (Ser52) Polyclonal Antibody (Invitrogen #44-728G). Membranes were then washed and probed with secondary antibody anti-rabbit immunoglobulin G (IgG), HRP-linked antibody (Cell Signaling Technology #7074S). Membrane was imaged by Pierce enhanced chemiluminesence (ECL) Western blotting substrate (ThermoScientific #32209) and the chemiluminescence measured on a FluorChem R (ProteinSimple).

Sequencing data analysis

Analysis software is archived (*68*). Sequencing data were processed using Cutadapt to remove constant adapter sequences and demultiplex libraries based on embedded nucleotide identifiers. The adapter sequences for each experimental dataset are provided in table S6. The underlined sequence represents the library-specific nucleotide identifier.

For dual-guide experiments (Fig. 4), demultiplexing was not required as each pool derived from a single P(*HIS4*) library. For this reason, adapter trimming was performed by instead taking the first 25 nucleotides using fastx_trimmer.

Trimmed barcodes were then counted using the custom "bc-seqs" program, which collapses barcode variants separated by single-nucleotide mismatches (*16*). Using the "bc-tabulate" script, these barcode counts were then collected into a matrix, tabulating samples within an experiment. The P(*HIS4*) and P(*PGK1*) libraries were sequenced in two Illumina HiSeq4000 runs and the barcode counts for the runs were summed.

Barcodes were first filtered to remove those that lacked at least 32 counts for at least one replicate in the preinduction DNA sample, or the preinduction P(*PGK1*)-normalizing sample for dual-barcoded experiments. The remaining barcodes were evaluated by differential activity analysis using mpralm (*18*). Barcode expression was compared between samples before and after gRNA induction or before and after 3AT treatment. Analysis was performed using the aggregate = "sum" parameter to sum barcodes that corresponded to the same gRNA and model_type = "indep_groups," which treats the replicates as independent experiments. Output tables were merged with gene information from the *Saccharomyces* Genome Database. Scatterplots comparing log fold change values of individual gRNAs showing were first thresholded for guides in which $q <$ 0.05 in at least one of the two expression analyses. Barcodes corresponding to no-guide

plasmids were used as negative controls to determine the distribution of *P* values. Genes involved in amino acid biosynthesis, aminoacyl tRNA charging, ISR-controlled translation initiation, RNA polymerase III transcription, tRNA processing, or actin cytoskeletal arrangement were hand-curated and used to annotate volcano and scatter plots.

REFERENCES AND NOTES

1. G. J. Knott, J. A. Doudna, CRISPR-Cas guides the future of genetic engineering. *Science* **361**, 866–869 (2018). doi: 10.1126/science.aat5011; pmid: 30166482
2. L. Cong *et al*., Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi: 10.1126/science.1231143; pmid: 23287718
3. M. Jinek *et al*., RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013). doi: 10.7554/eLife.00471; pmid: 23386978
4. P. Mali *et al*., RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013). doi: 10.1126/science.1232033; pmid: 23287722
5. L. A. Gilbert *et al*., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013). doi: 10.1016/j.cell.2013.06.044; pmid: 23849981
6. H. Koike-Yusa, Y. Li, E.-P. Tan, M. D. C. Velasco-Herrera, K. Yusa, Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32**, 267–273 (2014). doi: 10.1038/nbt.2800; pmid: 24535568
7. O. Shalem *et al*., Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014). doi: 10.1126/science.1247005; pmid: 24336571
8. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014). doi: 10.1126/science.1246981; pmid: 24336569
9. M. Kampmann, CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* **13**, 406–416 (2018). doi: 10.1021/acschembio.7b00657; pmid: 29035510
10. R. P. Patwardhan *et al*., High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* **27**, 1173–1175 (2009). doi: 10.1038/nbt.1589; pmid: 19915551
11. M. Breker, M. Schuldiner, Explorations in topology-delving underneath the surface of genetic interaction maps. *Mol. Biosyst.* **5**, 1473–1481 (2009). doi: 10.1039/b907076c; pmid: 19763324
12. B. Adamson *et al*., A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867–1882.e21 (2016). doi: 10.1016/j.cell.2016.11.048; pmid: 27984733
13. A. Dixit *et al*., Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853–1866.e17 (2016). doi: 10.1016/j.cell.2016.11.038; pmid: 27984732
14. D. A. Jaitin *et al*., Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **167**, 1883–1896.e15 (2016). doi: 10.1016/j.cell.2016.11.039; pmid: 27984734
15. P. Datlinger *et al*., Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017). doi: 10.1038/nmeth.4177; pmid: 28099430
16. N. J. McGlincy *et al., A genome-scale CRISPR interference guide library enables comprehensive phenotypic profiling in yeast. bioRxiv 2020.03.11.988105 [Preprint]. 12 March 2020. doi: 10.1101/2020.03.11.988105*
17. J. D. Smith *et al*., Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol.* **17**, 45 (2016). doi: 10.1186/s13059-016-0900-9; pmid: 26956608
18. L. Myint, D. G. Avramopoulos, L. A. Goff, K. D. Hansen, Linear models enable powerful differential activity analysis in massively parallel reporter assays. *BMC Genomics* **20**, 209 (2019). doi: 10.1186/s12864-019-5556-x; pmid: 30866806
19. M. E. Ritchie *et al*., limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015). doi: 10.1093/nar/gkv007; pmid: 25605792
20. P. Kaiser, K. Flick, C. Wittenberg, S. I. Reed, Regulation of transcription by ubiquitination without proteolysis: Cdc34/SCF^{Met30}-mediated inactivation of the transcription factor Met4. *Cell* **102**, 303–314 (2000). doi: 10.1016/S0092-8674(00)00036-2; pmid: 10975521
21. G. J. Smits, L. R. Schenkman, S. Brul, J. R. Pringle, F. M. Klis, Role of cell cycle-regulated expression in the localized incorporation of cell wall proteins in yeast. *Mol. Biol. Cell* **17**, 3267–3280 (2006). doi: 10.1091/mbc.e05-08-0738; pmid: 16672383
22. Y. P. Bhavsar-Jog, E. Bi, Mechanics and regulation of cytokinesis in budding yeast. *Semin. Cell Dev. Biol.* **66**, 107–118 (2017). doi: 10.1016/j.semcdb.2016.12.010; pmid: 28034796
23. P. Korber, S. Barbaric, The yeast PHO5 promoter: From single locus to systems biology of a paradigm for gene regulation through chromatin. *Nucleic Acids Res.* **42**, 10888–10902 (2014). doi: 10.1093/nar/gku784; pmid: 25190457
24. A. G. Hinnebusch, Translational regulation of GCN4 and the general amino acid control of yeast. *Annu. Rev. Microbiol.* **59**, 407–450 (2005). doi: 10.1146/annurev.micro.59.031805.133833; pmid: 16153175
25. A. Chambers, E. A. Packham, I. R. Graham, Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*). *Curr. Genet.* **29**, 1–9 (1995). doi: 10.1007/BF00313187; pmid: 8595651
26. P. K. Maitra, Z. Lobo, Control of glycolytic enzyme synthesis in yeast by products of the hexokinase reaction. *J. Biol. Chem.* **246**, 489–499 (1971). pmid: 4250748
27. N. P. Williams, A. G. Hinnebusch, T. F. Donahue, Mutations in the structural genes for eukaryotic initiation factors 2 alpha and 2 beta of *Saccharomyces cerevisiae* disrupt translational control of GCN4 mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7515–7519 (1989). doi: 10.1073/pnas.86.19.7515; pmid: 2678106
28. A. K. Hopper, Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. *Genetics* **194**, 43–67 (2013). doi: 10.1534/genetics.112.147470; pmid: 23633143
29. T. E. Dever, W. Yang, S. Aström, A. S. Byström, A. G. Hinnebusch, Modulation of tRNA(iMet), eIF-2, and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of eIF-2.GTP.Met-tRNA(iMet) ternary complexes. *Mol. Cell. Biol.* **15**, 6351–6363 (1995). doi: 10.1128/MCB.15.11.6351; pmid: 7565788
30. C. Conesa *et al*., Modulation of yeast genome expression in response to defective RNA polymerase III-dependent transcription. *Mol. Cell. Biol.* **25**, 8631–8642 (2005). doi: 10.1128/MCB.25.19.8631-8642.2005; pmid: 16166643
31. R. Shemer, A. Meimoun, T. Holtzman, D. Kornitzer, Regulation of the transcription factor Gcn4 by Pho85 cyclin PCL5. *Mol. Cell. Biol.* **22**, 5395–5404 (2002). doi: 10.1128/MCB.22.15.5395-5404.2002; pmid: 12101234
32. R. C. Wek, H. Y. Jiang, T. G. Anthony, Coping with stress: eIF2 kinases and translational control. *Biochem. Soc. Trans.* **34**, 7–11 (2006). doi: 10.1042/BST0340007; pmid: 16246168
33. N. S. Bae, A. P. Seberg, L. P. Carroll, M. J. Swanson, Identification of Genes in *Saccharomyces cerevisiae* that Are Haploinsufficient for Overcoming Amino Acid Starvation. *G3* **7**, 1061–1084 (2017). doi: 10.1534/g3.116.037416; pmid: 28209762
34. H. Kubota, Y. Sakaki, T. Ito, GI domain-mediated association of the eukaryotic initiation factor 2α kinase GCN2 with its activator GCN1 is required for general amino acid control in budding yeast. *J. Biol. Chem.* **275**, 20243–20246 (2000). doi: 10.1074/jbc.C000262200; pmid: 10801780
35. E. Sattlegger *et al*., YIH1 is an actin-binding protein that inhibits protein kinase GCN2 and impairs general amino acid control when overexpressed. *J. Biol. Chem.* **279**, 29952–29962 (2004). doi: 10.1074/jbc.M404009200; pmid: 15126500
36. I. A. Olave, S. L. Reck-Peterson, G. R. Crabtree, Nuclear actin and actin-related proteins in chromatin remodeling. *Annu. Rev. Biochem.* **71**, 755–781 (2002). doi: 10.1146/annurev.biochem.71.110601.135507; pmid: 12045110
37. D. Kornitzer, B. Raboy, R. G. Kulka, G. R. Fink, Regulated degradation of the transcription factor Gcn4. *EMBO J.* **13**, 6021–6030 (1994). doi: 10.1002/j.1460-2075.1994.tb06948.x; pmid: 7813440
38. E. S. Johnson, Protein modification by SUMO. *Annu. Rev. Biochem.* **73**, 355–382 (2004). doi: 10.1146/annurev.biochem.73.011303.074118; pmid: 15189146
39. A. Hoelz, E. W. Debler, G. Blobel, The structure of the nuclear pore complex. *Annu. Rev. Biochem.* **80**, 613–643 (2011). doi: 10.1146/annurev-biochem-060109-151030; pmid: 21495847
40. E. Rosonina, S. M. Duncan, J. L. Manley, Sumoylation of transcription factor Gcn4 facilitates its Srb10-mediated clearance from promoters in yeast. *Genes Dev.* **26**, 350–355 (2012). doi: 10.1101/gad.184689.111; pmid: 22345516
41. A. Akhter, E. Rosonina, Chromatin Association of Gcn4 Is Limited by Post-translational Modifications Triggered by its DNA-Binding in *Saccharomyces cerevisiae*. *Genetics* **204**, 1433–1445 (2016). doi: 10.1534/genetics.116.194134; pmid: 27770033
42. S. J. Li, M. Hochstrasser, A new protease required for cell-cycle progression in yeast. *Nature* **398**, 246–251 (1999). doi: 10.1038/18457; pmid: 10094048
43. S. J. Li, M. Hochstrasser, The Ulp1 SUMO isopeptidase: Distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J. Cell Biol.* **160**, 1069–1082 (2003). doi: 10.1083/jcb.200212052; pmid: 12654900
44. V. G. Panse, B. Küster, T. Gerstberger, E. Hurt, Unconventional tethering of Ulp1 to the transport channel of the nuclear pore complex by karyopherins. *Nat. Cell Biol.* **5**, 21–27 (2003). doi: 10.1038/ncb893; pmid: 12471376
45. B. Palancade *et al*., Nucleoporins prevent DNA damage accumulation by modulating Ulp1-dependent sumoylation processes. *Mol. Biol. Cell* **18**, 2912–2923 (2007). doi: 10.1091/mbc.e07-02-0123; pmid: 17538013
46. J. O. Rouvière *et al*., A SUMO-dependent feedback loop senses and controls the biogenesis of nuclear pore subunits. *Nat. Commun.* **9**, 1665 (2018). doi: 10.1038/s41467-018-03673-3; pmid: 29695777
47. L. Texari *et al*., The nuclear pore regulates GAL1 gene transcription by controlling the localization of the SUMO protease Ulp1. *Mol. Cell* **51**, 807–818 (2013). doi: 10.1016/j.molcel.2013.08.047; pmid: 24074957
48. G. Albrecht, H. U. Mösch, B. Hoffmann, U. Reusser, G. H. Braus, Monitoring the Gcn4 protein-mediated response in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 12696–12702 (1998). doi: 10.1074/jbc.273.21.12696; pmid: 9582292
49. A. M. Darnell, A. R. Subramaniam, E. K. O'Shea, Translational Control through Differential Ribosome Pausing during Amino Acid Limitation in Mammalian Cells. *Mol. Cell* **71**, 229–243.e11 (2018). doi: 10.1016/j.molcel.2018.06.041; pmid: 30029003
50. M. R. McFarland *et al*., The molecular aetiology of tRNA synthetase depletion: Induction of a GCN4 amino acid starvation response despite homeostatic maintenance of charged tRNA levels. *Nucleic Acids Res.* **48**, 3071–3088 (2020). doi: 10.1093/nar/gkaa055; pmid: 32016368
51. H. Qiu *et al*., Defects in tRNA processing and nuclear export induce GCN4 translation independently of phosphorylation of the alpha subunit of eukaryotic translation initiation factor 2. *Mol. Cell. Biol.* **20**, 2505–2516 (2000). doi: 10.1128/MCB.20.7.2505-2516.2000; pmid: 10713174
52. S. Kirchner, Z. Ignatova, Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat. Rev. Genet.* **16**, 98–112 (2015). doi: 10.1038/nrg3861; pmid: 25534324
53. S. R. Lee, K. Collins, Starvation-induced cleavage of the tRNA anticodon loop in Tetrahymena thermophila. *J. Biol. Chem.* **280**, 42744–42749 (2005). doi: 10.1074/jbc.M510356200; pmid: 16272149
54. D. M. Thompson, C. Lu, P. J. Green, R. Parker, tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* **14**, 2095–2103 (2008). doi: 10.1261/rna.1232808; pmid: 18719243
55. S. Yamasaki, P. Ivanov, G. F. Hu, P. Anderson, Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J. Cell Biol.* **185**, 35–42 (2009). doi: 10.1083/jcb.200811106; pmid: 19332886
56. D. Huh *et al*., An adaptive stress-induced tRNA depletion response mediates codon-based gene repression and growth suppression. bioRxiv 416727 [Preprint]. 13 September 2018. doi: 10.1101/416727
57. M. Kapur, C. E. Monaghan, S. L. Ackerman, Regulation of mRNA Translation in Neurons-A Matter of Life and Death. *Neuron* **96**, 616–637 (2017). doi: 10.1016/j.neuron.2017.09.057; pmid: 29096076
58. R. Ishimura *et al*., Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science* **345**, 455–459 (2014). doi: 10.1126/science.1249749; pmid: 25061210
59. R. Ishimura, G. Nagy, I. Dotu, J. H. Chuang, S. L. Ackerman, Activation of GCN2 kinase by ribosome stalling links translation elongation with translation initiation. *eLife* **5**, e14295 (2016). doi: 10.7554/eLife.14295; pmid: 27085088
60. B. D. Alford, G. Valiant, O. Brandman, Genome-wide, time-sensitive interrogation of the heat shock response under

diverse stressors via ReporterSeq. bioRxiv 2020.03.29.014845 [Preprint]. 31 March 2020. .doi: 10.1101/2020.03.29.014845
61. M. A. Sheff, K. S. Thorn, Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast* **21**, 661–670 (2004). doi: 10.1002/yea.1130; pmid: 15197731
62. A. Aranda-Díaz, K. Mace, I. Zuleta, P. Harrigan, H. El-Samad, Robust Synthetic Circuits for Two-Dimensional Control of Gene Expression in Yeast. *ACS Synth. Biol.* **6**, 545–554 (2017). doi: 10.1021/acssynbio.6b00251; pmid: 27930885
63. V. Stovicek, G. M. Borja, J. Forster, I. Borodina, EasyClone 2.0: Expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. *J. Ind. Microbiol. Biotechnol.* **42**, 1519–1531 (2015). doi: 10.1007/s10295-015-1684-8; pmid: 26376869
64. T. W. Christianson, R. S. Sikorski, M. Dante, J. H. Shero, P. Hieter, Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**, 119–122 (1992). doi: 10.1016/0378-1119(92)90454-W; pmid: 1544568
65. R. D. Gietz, R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007). doi: 10.1038/nprot.2007.13; pmid: 17401334
66. M. Ares, Isolation of total RNA from yeast cell cultures. *Cold Spring Harb. Protoc.* **2012**, 1082–1086 (2012). doi: 10.1101/pdb.prot071456; pmid: 23028070
67. A. M. McGeachy, Z. A. Meacham, N. T. Ingolia, An Accessible Continuous-Culture Turbidostat for Pooled Analysis of Complex Libraries. *ACS Synth. Biol.* **8**, 844–856 (2019). doi: 10.1021/acssynbio.8b00529; pmid: 30908907
68. R. Muller, Z. A. Meacham, L. Ferguson, N. T. Ingolia, CiBER-seq dissects genetic networks by quantitative CRISPRi profiling of expression phenotypes, version 1, Zenodo (2020); doi: 10.5281/zenodo.4035711

ACKNOWLEDGMENTS

We thank G. Brar, L. Lareau, E. Ünal, and members of the Ingolia and Lareau laboratories for discussion and advice. **Funding:** This work was funded by NIH grants DP2 CA195768, R01 GM130996, and R01 GM135233 (to N.T.I.). This work used the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley, supported by NIH S10 OD018174 instrumentation grant. **Author contributions:** R.M. and N.T.I. designed the experiments. R.M., Z.A.M., and L.F. carried out the experiments. R.M., L.F., and N.T.I. analyzed high-throughput sequencing data. R.M. and N.T.I. conceived the study and wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All high-throughput sequencing data are available from the National Center for Biotechnology Information, Sequence Read Archive under BioProject PRJNA578818. All source code is available at https://github.com/ingolia-lab/CiBER_seq and archived at Zenodo (*68*). Plasmids and libraries will be made available through AddGene.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eabb9662/suppl/DC1
Figs. S1 to S6
Tables S1 to S7
MDAR Reproducibility Checklist
Data S1

View/request a protocol for this paper from *Bio-protocol*.

30 March 2020; accepted 22 October 2020
10.1126/science.abb9662

RESEARCH ARTICLE SUMMARY

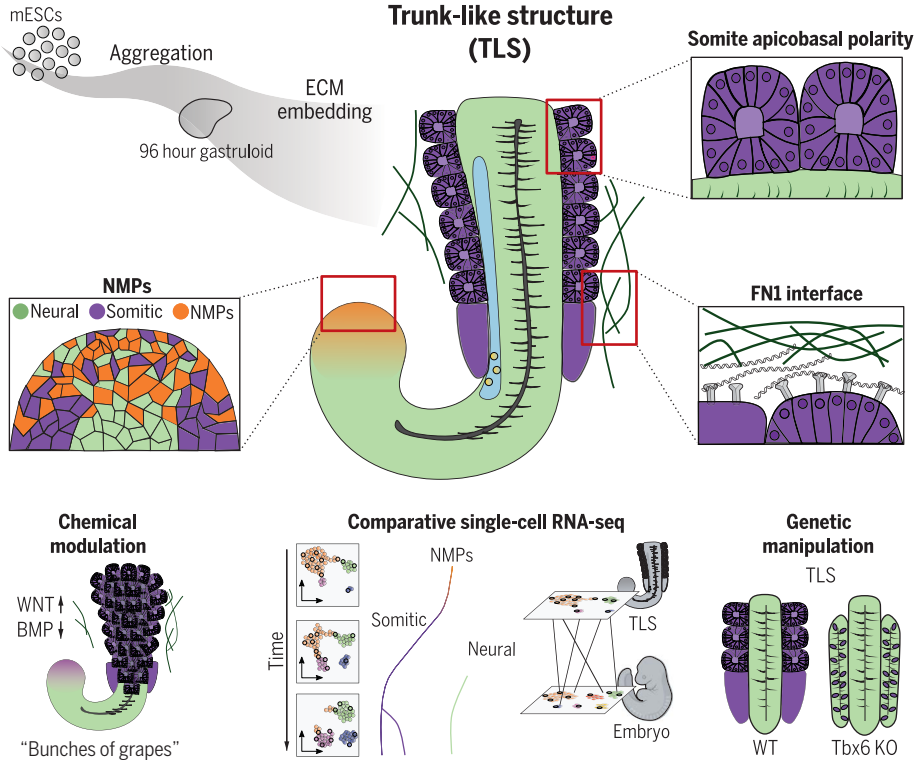
DEVELOPMENTAL BIOLOGY

Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites

Jesse V. Veenvliet*†, Adriano Bolondi†, Helene Kretzmer‡, Leah Haut‡, Manuela Scholze-Wittler, Dennis Schifferl, Frederic Koch, Léo Guignard, Abhishek Sampath Kumar, Milena Pustet, Simon Heimann, René Buschow, Lars Wittler, Bernd Timmermann, Alexander Meissner*, Bernhard G. Herrmann*

INTRODUCTION: Vertebrate development comprises multiple complex morphogenetic processes that shape the embryonic body plan through self-organization of pluripotent stem cells and their descendants. Because mammalian embryogenesis proceeds in utero, it is difficult to study the dynamics of these processes, including much-needed analysis at the cellular and molecular level. Various three-dimensional stem cell systems ("embryoids") have been developed to circumvent this impediment. The most advanced models of post-implantation development achieved so far are gastruloids, mouse embryonic stem cell (mESC)–derived aggregates with organized gene expression domains but lacking proper morphogenesis.

RATIONALE: To advance the current models, we explored the usage of Matrigel, an extracellular matrix (ECM) surrogate. During embryonic development, the ECM provides essential chemical and mechanical cues. In vitro, lower percentages of Matrigel can drive complex tissue morphogenesis in organoids, which led us to use Matrigel embedding in various media conditions to achieve higher-order embryo-like architecture in mESC-derived aggregates.



Engineering the embryonic trunk in a dish. Embedding of mouse embryonic stem cell (mESC) aggregates in an extracellular matrix (ECM) enables generation of trunk-like structures (TLSs) with an in vivo–like architecture including gut, and neuromesodermal progenitor (NMP)–derived neural tube and somites. Comparative single-cell RNA sequencing revealed that TLS cell states and differentiation dynamics match those of the embryo. Chemical modulation and genetic manipulation highlight the utility of the TLS as a scalable, tractable, and accessible model for investigating mid-gestational embryogenesis. FN1, fibronectin.

RESULTS: We found that embedding of 96-hour gastruloids in 5% Matrigel is sufficient to induce the formation of highly organized "trunk-like structures" (TLSs), comprising the neural tube and bilateral somites with embryo-like polarity. This high level of self-organization was accompanied by accumulation of the matrix protein fibronectin at the Matrigel-TLS interface and the transcriptional up-regulation of fibronectin-binding integrins and other cell adhesion molecules. Chemical modulation of signaling pathways active in the developing mouse embryo [WNT and bone morphogenetic protein (BMP)] resulted in an excess of somites arranged like a "bunch of grapes." Comparative time-resolved single-cell RNA sequencing of TLSs and embryos revealed that TLSs follow the same stepwise gene regulatory programs as the mouse embryo, comprising expression of critical developmental regulators at the right place and time. In particular, trunk precursors known as neuromesodermal progenitors displayed the highest differentiation potential and continuously contributed to neural and mesodermal tissue during TLS formation. In addition, live imaging demonstrated that the segmentation clock, required for rhythmic deposition of somites in vivo, ticks at an embryo-like pace in TLSs. Finally, a proof-of-principle experiment showed that Tbx6-knockout TLSs generate ectopic neural tubes at the expense of somite formation, mirroring the embryonic phenotype.

CONCLUSION: We showed that embedding of embryonic stem cell–derived aggregates in an ECM surrogate generates more advanced in vitro models that are formed in a process highly analogous to embryonic development. Trunk-like structures represent a powerful tool that is easily amenable to genetic, mechanical, chemical, or other modulations. As such, we expect them to facilitate in-depth analysis of molecular mechanisms and signaling networks that orchestrate embryonic development as well as studies of the ontogeny of mutant phenotypes in the culture dish. The scalable, tractable, and highly accessible nature of the TLS makes it a complementary in vitro platform for deciphering the dynamics of the molecular, cellular, and morphogenetic processes that shape the post-implantation embryo, at an unprecedented spatiotemporal resolution. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: veenvlie@molgen.mpg.de (J.V.V.); meissner@molgen.mpg.de (A.M.); herrmann@molgen.mpg.de (B.G.H.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
Cite this article as J. V. Veenvliet *et al.*, *Science* **370**, eaba4937 (2020). DOI: 10.1126/science.aba4937

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba4937

RESEARCH ARTICLE

DEVELOPMENTAL BIOLOGY

Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites

Jesse V. Veenvliet[1]*†, Adriano Bolondi[2]†, Helene Kretzmer[2]‡, Leah Haut[1,2]‡, Manuela Scholze-Wittler[1], Dennis Schifferl[1], Frederic Koch[1], Léo Guignard[3], Abhishek Sampath Kumar[2], Milena Pustet[1], Simon Heimann[1], René Buschow[4], Lars Wittler[1], Bernd Timmermann[5], Alexander Meissner[2,6,7,8]*, Bernhard G. Herrmann[1,9]*

Post-implantation embryogenesis is a highly dynamic process comprising multiple lineage decisions and morphogenetic changes that are inaccessible to deep analysis in vivo. We found that pluripotent mouse embryonic stem cells (mESCs) form aggregates that upon embedding in an extracellular matrix compound induce the formation of highly organized "trunk-like structures" (TLSs) comprising the neural tube and somites. Comparative single-cell RNA sequencing analysis confirmed that this process is highly analogous to mouse development and follows the same stepwise gene-regulatory program. *Tbx6* knockout TLSs developed additional neural tubes mirroring the embryonic mutant phenotype, and chemical modulation could induce excess somite formation. TLSs thus reveal an advanced level of self-organization and provide a powerful platform for investigating post-implantation embryogenesis in a dish.

Vertebrate post-implantation development comprises a multitude of complex morphogenetic processes resulting from self-organization of stem cells and their descendants shaping the embryonic body plan (*1*). Recently developed stem cell models represent powerful platforms for deconstructing the dynamics of these processes in vitro (*1*, *2*). The most advanced models in terms of developmental stage are gastruloids, self-organizing mESC aggregates that form elongating structures that comprise postoccipital embryo derivatives of all three germ layers but lack proper morphogenesis (*1*–*3*). In vivo, the extracellular matrix (ECM) provides chemical signals and exerts mechanical constraints via the basement membrane, which has a critical role in tissue morphogenesis (*4*, *5*). In vitro, Matrigel can serve as an ECM surrogate, and embedding of gastruloids in 10% Matrigel allowed the formation of a string of somite-like structures with anterior-posterior polarity (*6*). Lower percentages of Matrigel facilitated complex morphogenesis in organoids (*7*).

[1]Department of Developmental Genetics, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [2]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [3]Max Delbrück Center for Molecular Medicine and Berlin Institute of Health, 10115 Berlin, Germany. [4]Microscopy and Cryo-Electron Microscopy, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [5]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [6]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA. [7]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [8]Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany. [9]Institute for Medical Genetics, Charité–University Medicine Berlin, Campus Benjamin Franklin, 12203 Berlin, Germany. ***Corresponding author. Email: veenvlie@molgen.mpg.de (J.V.V.); meissner@molgen.mpg.de (A.M.); herrmann@molgen.mpg.de (B.G.H.)** †These authors contributed equally to this work. ‡These authors contributed equally to this work.

Matrigel embedding drives trunk-like morphogenesis

To achieve more advanced embryo-like morphogenetic features in gastruloids, we used Matrigel with various media conditions (Fig. 1A). To facilitate high-throughput characterization and quantification, we generated mESCs with *T::H2B-mCherry* (hereafter T^{mCH}) and *Sox2::H2B-Venus* (hereafter $Sox2^{VE}$) reporters, marking the mesodermal (ME) or neural (NE) lineage, respectively (fig. S1A). Embedding of 96-hour aggregates in 5% Matrigel resulted in segmentation in the T^{mCH+} ME domain and formation of a $Sox2^{VE+}$ neural tube–like structure (Fig. 1, A to C, fig. S1, B and C, and movies S1 and S2). The vast majority of structures [hereafter referred to as trunk-like structures (TLSs)] elongated and formed a T^{mCH+} pole at the posterior end, with segmentation occurring in about half of the TLSs (Fig. 1D). Whole-mount in situ hybridization for Tcf15 and Uncx confirmed somite identity and revealed embryo-like anterior-posterior polarity (Fig. 1E) (*8*). In 61% of the segmented TLSs, bilateral somites were observed (fig. S1D). Additional WNT activation using CHIR99021 (hereafter TLS^{C}) or combined with bone morphogenetic protein (BMP) inhibition by LDN193189 (hereafter TLS^{CL}) improved the physical separation of neighboring segments without affecting T^{mCH+} pole formation or elongation (Fig. 1, A to D, and fig. S1, E to H) and resulted in an excess of segments at the anterior end, arranged like a "bunch of grapes" (Fig. 1, B, C, and F) (*9*). Moreover, the ME domain expanded at the expense of the NE compartment, with apparent disorganization of the posterior end and neural tissue (Fig. 1B and fig. S1, G and H). This phenotype has not been observed in vivo and may be explained by shifting the lineage choice of neuromesodermal progenitors (NMPs)—bipotent cells giving rise to both postoccipital NE and ME (*10*)—toward ME as a result of dominant WNT signaling. In addition, $T^{mCH+}/Sox2^{VE+}$ putative NMPs were reduced, as confirmed by flow cytometry (fig. S1I).

Phalloidin and N-cadherin staining demonstrated that the cells of the neural tube and somites of TLS, TLS^{C}, and TLS^{CL} show proper apicobasal polarity, a characteristic of epithelial tissues, with F-actin and N-cadherin accumulating at the apical side (Fig. 1, G and H, and fig. S2, A and B). In gastruloids, by contrast, F-actin distribution appeared random and organized epithelial structures were not detected (fig. S2C).

Whole-mount immunofluorescence analysis of FOXA2 and SOX17, transcription factors characteristic for endoderm, revealed gut formation in a subset of TLSs (fig. S3, A to D). Cells at the posterior base were SOX17-negative but coexpressed FOXA2 with high levels of T^{mCH} (fig. S3B). Our data show that embedding in Matrigel is both necessary and sufficient to drive complex, embryo-like tissue morphogenesis of the three embryonic germ layers.

We next performed a detailed morphometric analysis of TLSs and their substructures. The data demonstrate reproducibility of the three protocols with respect to size and shape of the whole structure, somites, and neural tube, whereas the gut-like structure shows more variation (fig. S3E and fig. S4, A to D). Time-resolved whole-structure morphometry showed similar morphogenetic changes over time (fig. S4B). TLSs in general were larger than gastruloids (fig. S4A). Comparisons of the different TLS protocols revealed that, relative to TLS, TLS^{C} and TLS^{CL} were slightly larger and formed more somites, whereas their neural domains were reduced in length and narrowed toward the anterior end (fig. S4, A to D). In all protocols, somites were similar in shape but smaller than their embryonic counterparts (fig. S4D).

To assess whether the segmentation clock, an oscillator involved in somitogenesis, is active in TLSs, we performed live imaging (*11*). In line with recent observations, we found that segmentation occurs in a rhythmic fashion at an embryo-like pace in all three TLS conditions (fig. S5, A to C, and movies S3 to S5) (*6*). TLS^{C} and TLS^{CL} showed consecutive formation of multiple somites (movies S6 to S9).

Transcriptional characterization highlights selective responses to chemical modulation

To characterize the structures in detail at the molecular level, we performed RNA sequencing

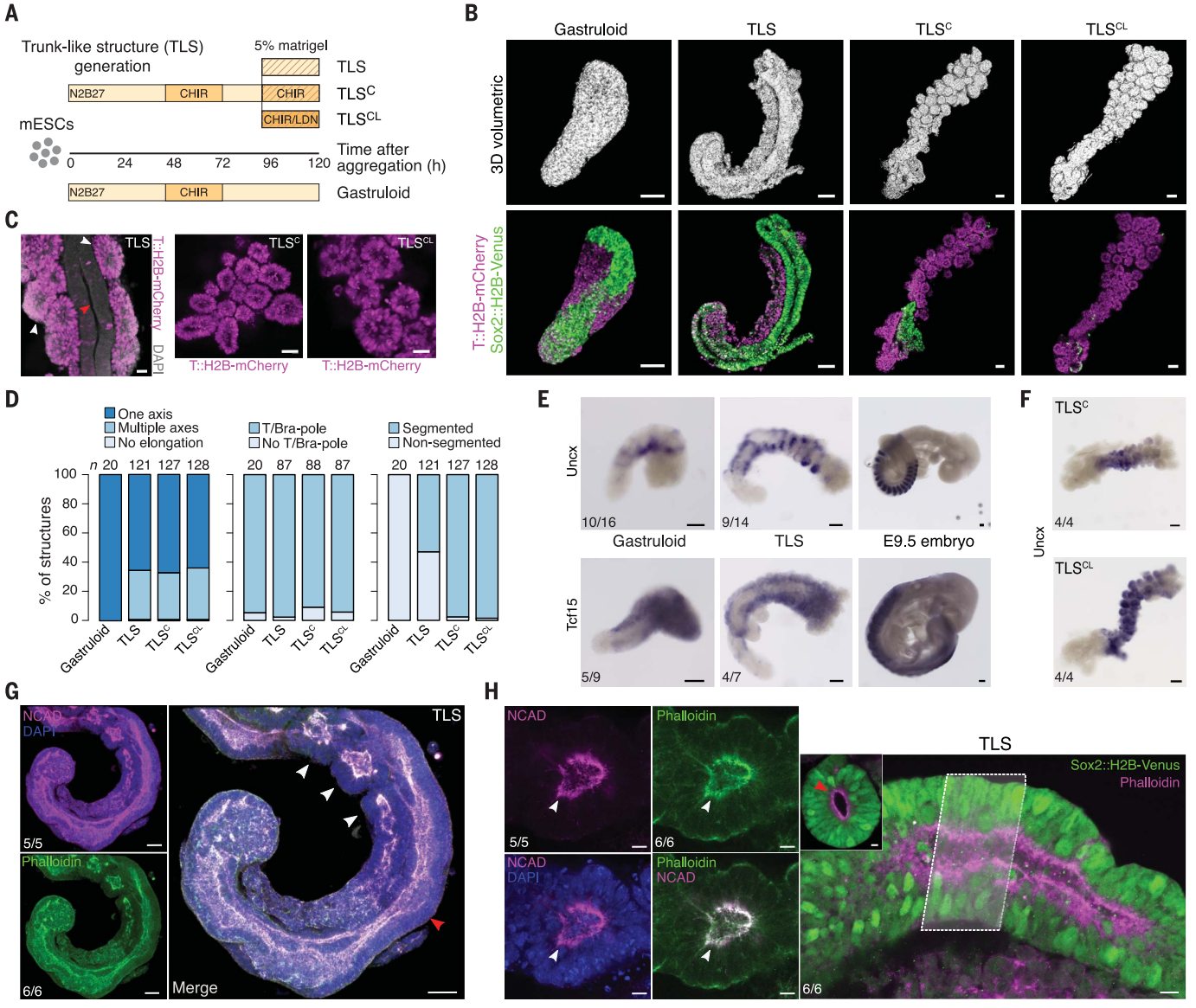


Fig. 1. Generation of trunk-like structures (TLSs) with somites and a neural tube. (**A**) Schematic overview: 200 to 250 mESCs were aggregated in ultralow-attachment plates; *Wnt* agonist CHIR99021 (CHIR) was added between 48 and 72 hours (*3*). At 96 hours, aggregates were (i) cultured as gastruloids (*3*), (ii) embedded in 5% Matrigel (TLS), and (iii) treated with WNT signaling activator CHIR99021 (CHIR) alone (TLS^C), or (iv) treated with CHIR and the BMP signaling inhibitor LDN193189 (TLS^{CL}). The two compounds have been reported to induce a (pre-)somitic mesoderm fate in 2D and 3D differentiation protocols (*8*). (**B**) 3D volumetric renderings (top) and confocal sections (bottom) of a gastruloid, TLS, TLS^C, and TLS^{CL}. Scale bars, 100 μm. Each image is representative of at least 10 biological replicates with similar morphology and expression patterns. (**C**) Segments in TLS are T^{mCH+} and are positioned adjacent to the neural tube. In TLS^C and TLS^{CL} the segments are arranged in "bunches of grapes." Scale bars, 25 μm. Red arrowhead, neural tube; white arrowheads, somites. (**D**) Quantification of morphogenetic features in gastruloids, TLSs, TLS^C, and TLS^{CL} (see supplementary materials for scoring criteria). (**E**) Segments express somitic markers Uncx and Tcf15 as shown by whole-mount in situ hybridization. Note the characteristic stripe-like expression pattern of Uncx in TLS due to posterior restriction, whereas Tcf15 is expressed throughout the segments (as in the embryo). Scale bars, 100 μm. (**F**) In TLS^C and TLS^{CL}, Uncx is detected throughout the segments, indicating loss of anterior-posterior polarity. Scale bars, 100 μm. (**G** and **H**) Confocal sections showing that cells of somites and neural tube display apical-basal polarity with NCAD and F-actin (phalloidin) accumulating at the apical surface. Inset represents an optical section of the neural tube shown in the main panel. Scale bars, 50 μm (G), 10 μm (H). White arrowheads, somites; red arrowheads, neural tubes.

(RNA-seq) analysis (fig. S6A) and found that the TLS models the postoccipital embryo, similar to gastruloids (Fig. 2A) (*3*). Relative to TLS, both TLS^C and TLS^{CL} showed significant up-regulation of genes involved in (pre)somitic development [e.g., *Tbx6*, *Msgn1*, *Hes7* (*8*, *10*, *12*)] at the expense of NE marker genes [e.g., *Sox1*, *Pax6*, *Irx3* (*13*)], corroborating the flow cytometry and imaging results (Fig. 2A and fig. S6, B and C). The analysis of marker gene sets for NMPs, for their direct descendants undergoing lineage choice (NMP ME and NMP NE), and for committed NE and ME cells substantiated this finding. Moreover, TLS^C and TLS^{CL} displayed reduced expression of markers in all clusters, including ME, relative to TLS (Fig. 2B). In contrast, on average (pre)somitic mesoderm [(P)SM]–specific markers were up-regulated, whereas intermediate ME (IM) and lateral plate ME (LPM) markers were down-regulated in TLS^C and further reduced

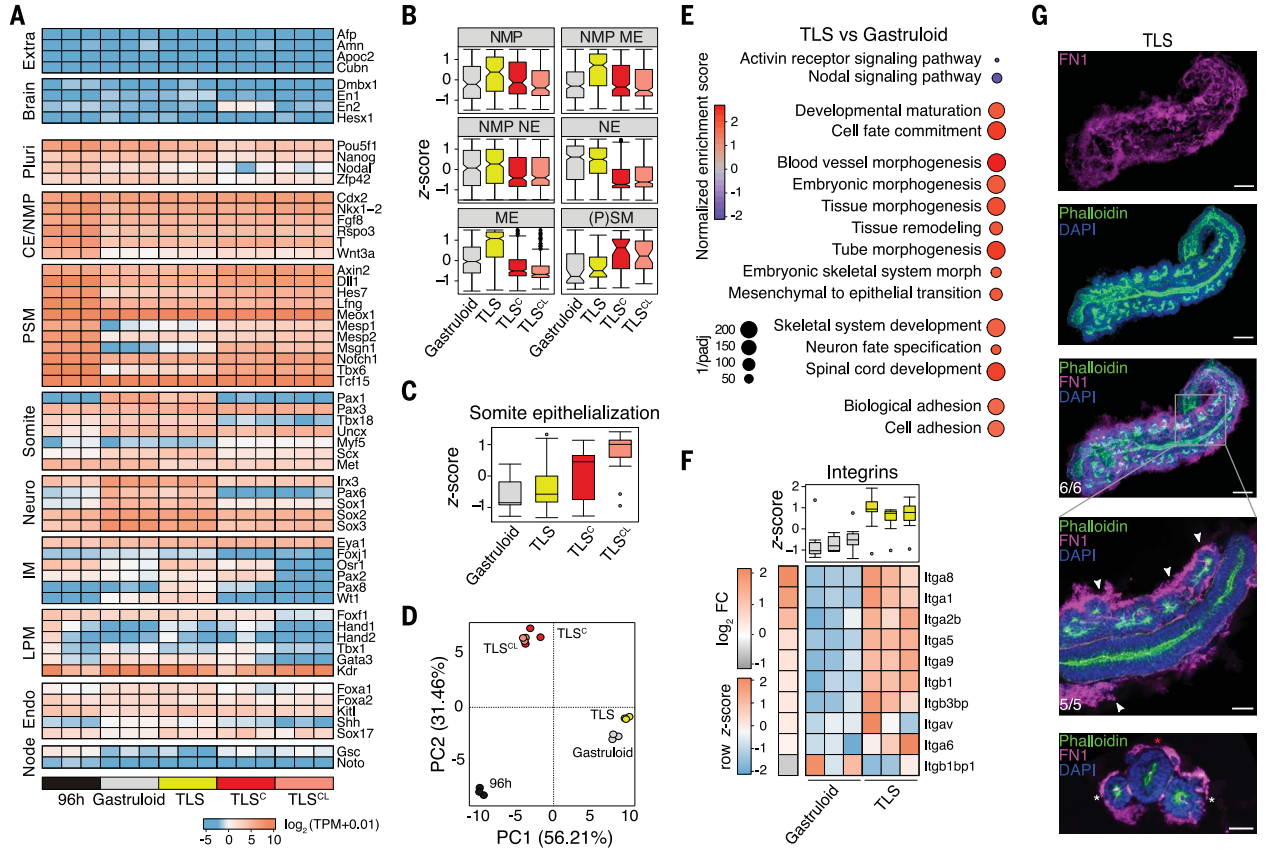


Fig. 2. Transcriptional characterization of TLSs. (**A**) Heatmap of $\log_2$(TPM + 0.01) expression (TPM, transcripts per million) of selected genes associated with development of indicated embryonic structures in 96-hour aggregates and 120-hour gastruloids, 120-hour TLSs, 120-hour TLS^{C}, and 120-hour TLS^{CL}, as measured by RNA-seq. Replicates were derived from pools of independent biological samples (see fig. S6A for experimental setup). CE, caudal end; NMP, neuromesodermal progenitors; PSM, presomitic mesoderm; LPM, lateral plate mesoderm; IM, intermediate mesoderm. (**B**) Box plots showing distribution of marker genes for indicated cell types (average *z*-score). Boxes indicate interquartile range (IQR); whiskers extend to 1.5 × IQR from the hinge. Dots indicate outliers. Notches are centered on the median. See data S1 for lists of genes used for each category and statistical analysis. (**C**) Box plot representing average *z*-score per column (pool of three replicates) for somite epithelialization factors (see fig. S6E for individual genes). (**D**) PCA analysis of samples from (A) with color coding of individual samples (dots) as in (A). PC1 and PC2 represent the two components with highest percentage of explained variance. (**E**) Selected significant terms of gene set enrichment analysis (GSEA) enriched in TLSs as compared to gastruloids at 120 hours. See data S2 for full list of significant terms [false discovery rate (FDR) < 0.05]. (**F**) Heatmap of scaled expression (row *z*-score) and $\log_2$ FC of integrins with significantly different expression ($P_{adj(FDR)}$ < 0.05) in 120-hour TLSs versus gastruloids. Box plot represents *z*-score per column (sample), with boxes indicating interquartile range, whiskers extending to 1.5 × IQR from the hinge, dots showing outliers, and central line representing median. Each column represents one of three biological replicates. See data S1 for statistical analysis. (**G**) 3D maximum-intensity projection (top three images) and confocal section showing FN1 accumulation around TLS somites and neural tube (zoomed-in image, white arrowheads). Phalloidin staining shows apical-basal polarity. Scale bars, 100 μm, 50 μm for magnification. Bottom: Light-sheet optical transversal section showing FN1 accumulation around the somites (white asterisk) and neural tube (red asterisk) in TLSs. Scale bar, 50 μm.

in TLS^{CL} (Fig. 2B and fig. S6D) (*3*, *12*), in line with the known role of WNT versus BMP signaling in PSM versus IM and LPM specification (Fig. 2A and fig. S6D) (*8*).

We next searched for gene expression differences possibly underlying improved physical separation of somites observed in TLS^{C} and TLS^{CL}. Among the most strongly up-regulated genes relative to TLSs was *Wnt6*, which acts as a somite epithelialization factor in vivo (fig. S6E) (*14*). In addition, multiple ephrins and their receptors, and other factors involved in somite epithelialization, were up-regulated (Fig. 2C and fig. S6E) (*15*, *16*). We observed expression changes of selected somite polarity markers and their inducers, in line with the role played by WNTs, SHH, BMPs, and their antagonists in somite compartmentalization in vivo (fig. S6, F and G) (*8*, *17*). Our data show that exposure to CHIR or CHIR/LDN improved segment boundary formation but affected somite cellular composition.

Tissue morphogenesis and remodeling genes are up-regulated in TLSs

Principal components analysis (PCA) indicated a high transcriptional similarity between gastruloids and TLSs despite profound morphological differences (Fig. 2D). The latter are better highlighted by gene set enrichment analysis (GSEA), which showed that Matrigel embedding promotes tissue morphogenesis and remodeling (Fig. 2E and fig. S6H). Zooming in on embryonic and tissue morphogenesis gene sets revealed markers of blood vessel development among up-regulated genes, which suggests the induction of capillary morphogenesis in TLSs (fig. S6I). GSEA also showed an enrichment of cell adhesion terms and overall a significant up-regulation of corresponding marker genes in TLSs (Fig. 2E and fig. S7, A and B). The most pronounced increase was observed for integrins, transmembrane receptors mediating cell adhesion to the ECM (important for, e.g., neural tube formation, blood vessel development, and segmentation) (Fig. 2F and fig. S7, A and B) (*18*–*21*). Because binding of integrin to the glycoprotein fibronectin (FN1) and matrix

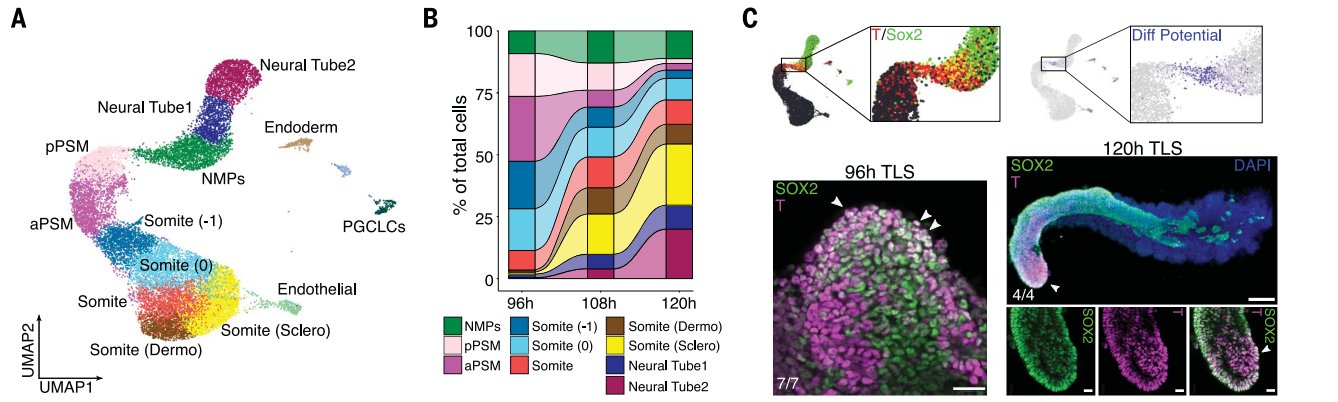


Fig. 3. Time-resolved single-cell RNA-seq of TLSs. In total, 20,294 cells were sampled from TLSs at 96, 108, and 120 hours (see fig. S9A for experimental setup). (**A**) UMAP (uniform manifold approximation and projection) colored by the 14 clusters identified. (**B**) Alluvial plot of percentage of neuromesodermal progenitors (NMPs), posterior presomitic mesoderm (pPSM), anterior PSM, somitic, and neural tube cells over time. (**C**) NMPs coexpress Sox2 and T (top left; blending with normalized expression threshold = 0.25) and are characterized by the highest differentiation potential (top right; see supplementary materials for differentiation potential calculation). NMPs coexpress T and SOX2 at 96 and 120 hours and reside at the posterior end of the TLS (confocal sections, bottom left and magnifications, 3D maximum-intensity projection, whole structure); white arrowheads, NMPs. Scale bars, 50 μm for 96-hour TLS, 100 μm for 3D maximum-intensity projection, 20 μm for magnifications.

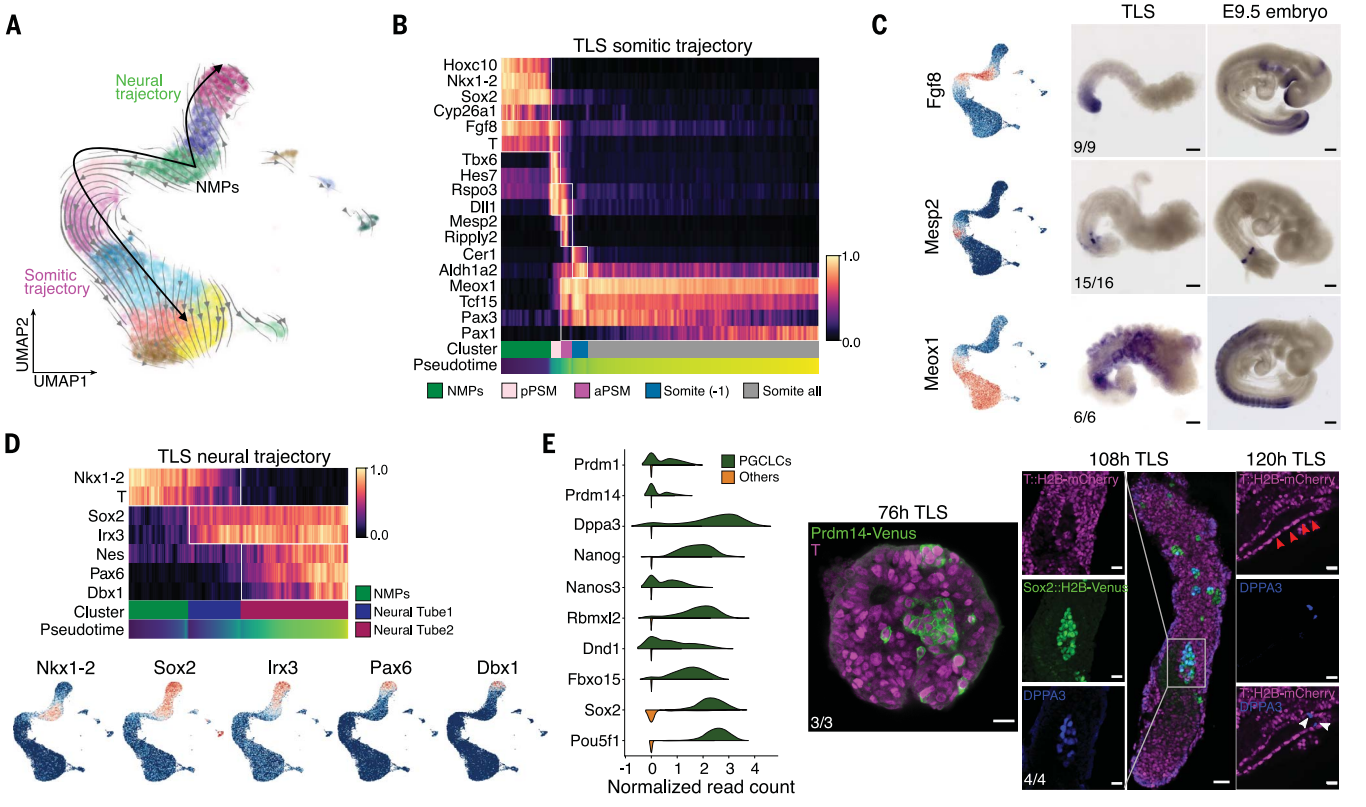


Fig. 4. TLS differentiation trajectories. (**A**) UMAP colored by identified clusters with trajectories inferred from RNA velocity. Gray arrow flows represent calculated velocity trajectories. (**B**) Heatmap with scaled expression of genes involved in somitogenesis measured in 5966 cells from 120-hour TLSs rooted in NMPs and ordered by pseudotime. (**C**) UMAP colored by expression of indicated genes (left) and whole-mount in situ hybridization for the same genes in TLS and E9.5 embryos (right). Numbers indicate the fraction of TLSs with embryo-like expression. Scale bars, TLS, 100 μm; embryo, 200 μm. (**D**) Heatmap with scaled expression of genes involved in neural development measured in 3462 cells from 120-hour TLSs rooted in NMPs and ordered by pseudotime (top) and UMAP colored by expression of indicated genes (bottom). (**E**) Split violin plots showing expression of marker genes for primordial germ cell–like cells (PGCLCs, left) and confocal sections of TLS showing PGCLC specification dynamics: T/PRDM14^{VE} double positive at 76 hours, SOX2$^{VE\text{-}high}$/DPPA3^{+} PGCLCs at 108 hours, and DPPA3^{+} cells in close contact with the T^{mCH+} gut-like structure at 120 hours. At 120 hours, Sox2$^{VE\text{-}high}$ cells in contact with FOXA2^{+} gut-like domain or DPPA3^{+} cells in contact with T^{mCH+} gut-like domain were observed in seven of nine TLSs (see also fig. S14D). Scale bars, 20 μm for 76-hour TLS, 50 μm for 108-hour TLS (20 μm for magnifications), 25 μm for 120-hour TLS. Red arrowheads, gut-like structure; white arrowheads, DPPA3^{+} PGCLCs.

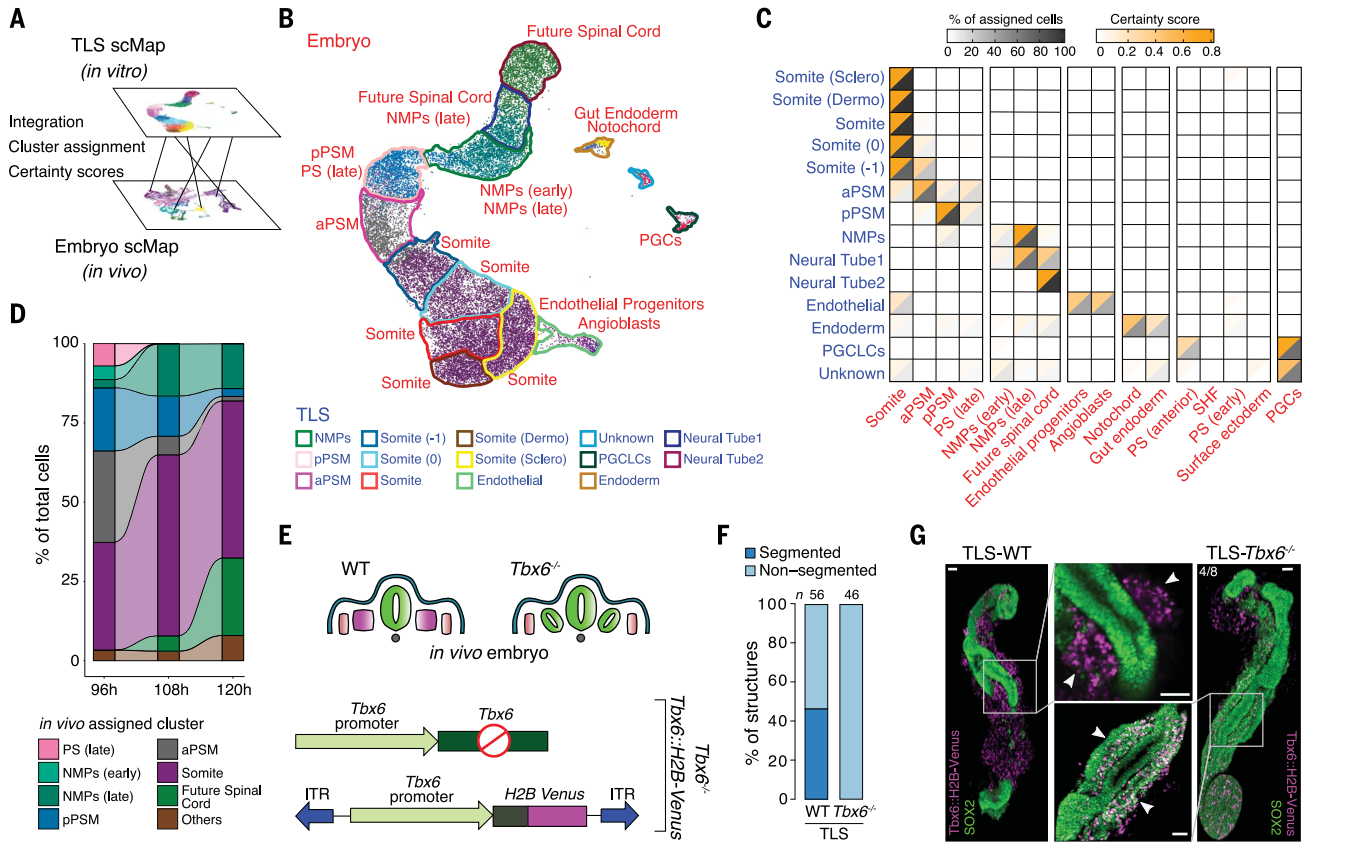


Fig. 5. TLS cell states are embryo-like and *Tbx6*$^{-/-}$ TLSs recapitulate the embryonic knockout phenotype. (**A**) Schematic of our comparative transcriptome analysis of TLSs with postoccipital E7.5 and E8.5 embryos at the single-cell level. (**B**) TLS UMAP colored by assigned embryonic cell states. TLS clusters are projected as corresponding colored contours. Blue font, TLS clusters; red font, embryo clusters. (**C**) Split heatmap with percentage of assigned cells (dark gray) and certainty score (orange) for TLS cells from the indicated cluster upon unbiased mapping to the in vivo counterpart. Font colors as in (B). (**D**) Alluvial plot of percentage of cells assigned to the indicated in vivo clusters in 96-, 108-, and 120-hour TLSs. (**E**) Simplified schematics of *Tbx6*$^{-/-}$ in vivo phenotype and knockout/reporter constructs. (**F**) Quantification of segmentation phenotype in TLS-*Tbx6*$^{-/-}$. Data represent three different experiments performed with two independent mESC lines of each genotype. (**G**) Formation of ectopic neural tubes in TLS-*Tbx6*$^{-/-}$. Ectopic neural tubes are identified as SOX2^{+}/Tbx6^{VE+} tubular structures flanking the main SOX2^{+}-only neural tube. Green, SOX2; magenta, Tbx6^{VE}. White arrowheads indicate Tbx6^{VE+} somites in the wild type (WT) and Tbx6^{VE+}/SOX2^{+} ectopic neural tubes in *Tbx6*$^{-/-}$. Scale bars, 50 μm.

assembly play an important role in somite and neural tube formation (*21*, *22*), we asked whether FN1, which is highly expressed in TLSs, was retained by Matrigel, as recently shown in 2D culture (fig. S7C) (*23*). FN1 protein clearly accumulated at the TLS-Matrigel interface; this was not observed in gastruloids, where FN1 likely diffuses into the medium (Fig. 2G, fig. S7D, and movie S10). Taken together, our data suggest that the activation of morphogenetic programs by Matrigel embedding is driven by ECM assembly involving integrins and fibronectin.

Single-cell RNA-seq demonstrates embryo-like dynamics of cell differentiation

We next focused on the TLS condition because it produced the most in vivo–like structures. After confirming reproducibility at the molecular level (fig. S8, A to C), we performed a time-resolved single-cell RNA-seq (scRNA-seq) analysis on a total of 20,294 postprocessed cells sampled from TLSs at 96, 108, and 120 hours (fig. S9A). Clustering analysis identified 14 different cell states. The larger clusters corresponded to derivatives of the PSM and NE that flank putative NMPs, whereas smaller clusters comprised endoderm, endothelial cells, and primordial germ cell–like cells (PGCLCs) (fig. S9B). The main clusters organized into a continuum of states recapitulating spatiotemporal features of the developing postoccipital embryo (Fig. 3A). Across the three time points sampled, progenitor cells gradually decreased in favor of more mature neural and somitic cells as development progressed (Fig. 3B and fig. S9C). Putative NMPs coexpressing T and SOX2, or T^{mCH}, Sox2^{VE}, and CDX2, were located at the posterior end at 96 and 120 hours (Fig. 3C and fig. S10, A to C) (*10*, *24*). TLS-NMPs thus display an in vivo–like NMP signature and have the highest differentiation potential, as they give rise to differentiating cells of both neural and mesodermal lineages (Fig. 3C and fig. S10, D to F).

RNA velocity analysis revealed neural and somitic trajectories rooted in the NMPs, further suggesting that the TLS recapitulates the developmental dynamics observed in the mid-gestational embryo (Fig. 4A and fig. S11A) (*25*). In vivo, NMPs and their descendants are arranged in order of progressive maturity along the posterior-to-anterior axis (*8*). Accordingly, ordering of TLS-derived cells along a pseudotemporal trajectory showed that the somitic trajectory reflects the genetic cascade observed in the embryo (Fig. 4B and fig. S12A). For example, the trajectory from Fgf8^{+} NMPs and PSM, through the determination front marked by Mesp2, to Meox1^{+} somites, was faithfully recapitulated, and the embryo-like spatial arrangement was confirmed by whole-mount in situ hybridization (Fig. 4C) (*8*). Likewise, the genetic cascade from NMPs to neural progenitors reflected the in vivo differentiation path in space and time (Fig. 4D). Subclustering of the neural cells demonstrated that TLSs generate

both dorsal and ventral neural subtypes, with dorsal subtypes being more prevalent (fig. S12B) (*13*). The analysis of *Hox* gene expression at consecutive time points showed in vivo–like collinear activation, as described for gastruloids (fig. S12C) (*3*). To test whether TLS somites establish cell states segregated along the dorsal-ventral (D-V) and anterior-posterior (A-P) somite axes in vivo, we reclustered all somitic cells. At 96 hours, we detected two main groups corresponding to the Uncx^{+} posterior and Tbx18^{+} anterior somite domains, in line with the A-P polarity established during segmentation (fig. S13, A and B) (*6*, *8*). At 120 hours, we found distinct clusters of Pax3^{+} (putative dorsal dermomyotomal) and Pax1^{+} (putative ventral sclerotomal) cells, as well as a small cluster of Lbx1^{+}/Met^{+} putative migratory limb muscle precursors (fig. S13, C to F) (*8*, *26*). In addition, Scx^{+} syndetome cells were detected (fig. S13G), and Uncx and Tbx18 expression were anticorrelated (fig. S13H).

Primordial germ cell (PGC) specification in the embryo occurs between embryonic day (E) 6.0 and E6.5 via T-mediated activation of *Blimp1* and *Prdm14*, and at E7.5, nascent PGCs can be identified as a group of DPPA3^{+} cells in the posterior primitive streak, which later migrate along the hindgut to the gonads (*27*, *28*). We assigned PGCLC identity using marker genes characteristic for PGCs and identified their location in the TLS (Fig. 4E and fig. S14). At 76 hours, roughly corresponding to stage E6.5, we detected T/Prdm14^{VE}–coexpressing cell clusters (Fig. 4E and fig. S14, A and B). At 108 hours, we found a group of Sox2$^{VE\text{-}high}$ cells that coexpressed DPPA3 (Fig. 4E and fig. S14C). At 120 hours, Sox2$^{VE\text{-}high}$ cells were detected in contact with FOXA2^{+} cells, and DPPA3^{+} cells in contact with a T^{mCH+} gut-like epithelial structure (Fig. 4E and fig. S14D). These data show that TLSs contain cells displaying characteristics typical for PGCs.

TLSs display a high complexity of cell states that match their in vivo counterparts

Single-cell comparison of 120-hour gastruloids with 120-hour TLSs identified different proportions of the major cell states (fig. S15, A to C). A more refined analysis revealed a higher complexity of cell states in TLSs (fig. S16, A to E, fig. S17, A to F, and fig. S18, A to D), and expression of later (more posterior) *Hox* genes suggests development into more advanced trunk stages (fig. S18, E and F). The comparison of TLSs with TLS^{CL} showed that in the latter, (i) sclerotomal and more mature neural cells are virtually absent, and (ii) somitic as well as endothelial cell identities are altered (fig. S15, B and C, fig. S16, A and B, fig. S17, A to G, and fig. S18, A and B). Application of RNA velocity confirmed that in TLS^{CL}, NMPs are highly biased toward the mesodermal lineage, whereas contribution to the neural lineage is diminished relative to TLS-NMPs; this is further corroborated by up-regulation of posterior PSM and down-regulation of neural marker genes (fig. S15, D and E) (*8*, *13*).

To investigate how close the cellular states identified in TLSs resemble those in embryos, we mapped TLS single-cell transcriptomes to a scRNA-seq compendium of postoccipital embryonic cellular subtypes (Fig. 5A) (*29*). The data revealed globally high accordance of TLS and embryonic cell states including characteristic marker genes, and pairwise comparison of mapped clusters identified only a small fraction of differentially expressed genes (Fig. 5, B and C, and fig. S19, A to D). Of note, PS- and early NMP–like cells were exclusively present at 96 hours and were replaced by late NMP–like cells at 108 and 120 hours (Fig. 5D). Taken together, our scRNA-seq analyses demonstrate that the TLS executes gene regulatory programs in a spatiotemporal order resembling that of the embryo.

Knockout TLSs display the embryonic mutant phenotype

Finally, to explore the utility of TLSs further, we conducted a proof-of-concept experiment to test whether gene ablation would reproduce the embryonic mutant phenotype. In vivo, loss of *Tbx6* results in transdifferentiation of prospective PSM and subsequent formation of ectopic neural tubes at the expense of PSM and somites (Fig. 5E) (*30*, *31*). We deleted *Tbx6* from *Tbx6::H2B-Venus* (*Tbx6*^{Ve}) mESCs and generated TLSs, which clearly failed to form somites even upon treatment with CHIR or CHIR/LDN (Fig. 5, E and F, and fig. S20, A and B). Quantitative polymerase chain reaction analysis on fluorescence-activated cell sorter–purified Tbx6^{VE+} cells revealed up-regulation of neural markers at the expense of (P)SM markers in *Tbx6*$^{-/-}$ cells, thus recapitulating the in vivo phenotype at the molecular level (fig. S20C). Finally, whole-mount immunofluorescence analysis for SOX2 showed that *Tbx6*$^{-/-}$ TLSs generated ectopic Tbx6^{VE+} neural tubes, whereas gastruloids, TLS^{C}, and TLS^{CL} formed an excess of morphologically indistinct SOX2^{+} tissue (Fig. 5G and fig. S20, D to F).

Discussion

The TLS model provides a scalable, tractable, readily accessible platform for investigating the lineage decisions and morphogenetic processes that shape the mid-gestational embryo with high spatiotemporal resolution. Our results show that the TLS faithfully reproduces key features of postoccipital embryogenesis, including axial elongation with coordinated neural tube, gut, and somite formation as well as PGCLCs. Accordingly, genetic manipulation of the TLS faithfully reproduced the morphogenetic and molecular changes observed in vivo. Thus, the TLSs will enable deeper analysis of the ontogeny of mutant phenotypes and provide an additional tool for investigating morphogenetic mechanisms unavailable in vivo. We also envision that the TLS^{C} and TLS^{CL} models may become important for testing current concepts of somitogenesis—for instance, the hypothesis that somite size and shape are controlled by local cell-cell interactions (*9*).

Mechanistically, our data highlight a crucial role for the ECM surrogate in unlocking the potential of in vitro derived mESC aggregates, although future efforts will have to address the exact functional contribution of individual components and biophysical properties (fig. S21), possibly using modular synthetic 3D matrices (*7*, *32*). Alternatively, the single-cell expression catalog of TLSs and gastruloids can provide initial guidance for further exploration of cell-cell and cell-matrix interactions and their control of embryonic architecture (fig. S7, A and B, and figs. S22 and S23).

REFERENCES AND NOTES

1. M. N. Shahbazi, E. D. Siggia, M. Zernicka-Goetz, Self-organization of stem cells into embryos: A window on early mammalian development. *Science* **364**, 948–951 (2019). doi: 10.1126/science.aax0164; pmid: 31171690
2. M. N. Shahbazi, M. Zernicka-Goetz, Deconstructing and reconstructing the mouse and human early embryo. *Nat. Cell Biol.* **20**, 878–887 (2018). doi: 10.1038/s41556-018-0144-x; pmid: 30038253
3. L. Beccari *et al*., Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276 (2018). doi: 10.1038/s41586-018-0578-0; pmid: 30283134
4. I. Matsuo, R. Hiramatsu, Mechanical perspectives on the anterior-posterior axis polarization of mouse implanted embryos. *Mech. Dev.* **144A**, 62–70 (2017). doi: 10.1016/j.mod.2016.09.002; pmid: 27697519
5. H. K. Kleinman, D. Philp, M. P. Hoffman, Role of the extracellular matrix in morphogenesis. *Curr. Opin. Biotechnol.* **14**, 526–532 (2003). doi: 10.1016/j.copbio.2003.08.002; pmid: 14580584
6. S. C. van den Brink *et al*., Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409 (2020). doi: 10.1038/s41586-020-2024-3; pmid: 32076263
7. J. A. Brassard, M. P. Lutolf, Engineering Stem Cell Self-organization to Build Better Organoids. *Cell Stem Cell* **24**, 860–876 (2019). doi: 10.1016/j.stem.2019.05.005; pmid: 31173716
8. J. Chal, O. Pourquié, Making muscle: Skeletal myogenesis *in vivo* and *in vitro*. *Development* **144**, 2104–2122 (2017). doi: 10.1242/dev.151035; pmid: 28634270
9. A. S. Dias, I. de Almeida, J. M. Belmonte, J. A. Glazier, C. D. Stern, Somites without a clock. *Science* **343**, 791–795 (2014). doi: 10.1126/science.1247575; pmid: 24407478
10. F. Koch *et al*., Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev. Cell* **42**, 514–526.e7 (2017). doi: 10.1016/j.devcel.2017.07.021; pmid: 28826820
11. O. Pourquié, The segmentation clock: Converting embryonic time into spatial pattern. *Science* **301**, 328–330 (2003). doi: 10.1126/science.1085887; pmid: 12869750
12. J. Chal *et al*., Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969 (2015). doi: 10.1038/nbt.3297; pmid: 26237517
13. A. Sagner, J. Briscoe, Establishing neuronal diversity in the spinal cord: A time and a place. *Development* **146**, dev182154 (2019). doi: 10.1242/dev.182154; pmid: 31767567
14. J. von Maltzahn, N. C. Chang, C. F. Bentzinger, M. A. Rudnicki, Wnt signaling in myogenesis. *Trends Cell Biol.* **22**, 602–609 (2012). doi: 10.1016/j.tcb.2012.07.008; pmid: 22944199
15. M. Rowton *et al*., Regulation of mesenchymal-to-epithelial transition by PARAXIS during somitogenesis. *Dev. Dyn.* **242**, 1332–1344 (2013). doi: 10.1002/dvdy.24033; pmid: 24038871

16. L. Durbin *et al*., Eph signaling is required for segmentation and differentiation of the somites. *Genes Dev.* **12**, 3096–3109 (1998). doi: 10.1101/gad.12.19.3096; pmid: 9765210
17. C. Marcelle, M. R. Stark, M. Bronner-Fraser, Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development* **124**, 3955–3963 (1997). pmid: 9374393
18. D. Jülich *et al*., Cross-Scale Integrin Regulation Organizes ECM and Tissue Topology. *Dev. Cell* **34**, 33–44 (2015). doi: 10.1016/j.devcel.2015.05.005; pmid: 26096733
19. J. H. Shawky, L. A. Davidson, Tissue mechanics and adhesion during embryo development. *Dev. Biol.* **401**, 152–164 (2015). doi: 10.1016/j.ydbio.2014.12.005; pmid: 25512299
20. S. Koshida *et al*., Integrinalpha5-dependent fibronectin accumulation for maintenance of somite boundaries in zebrafish embryos. *Dev. Cell* **8**, 587–598 (2005). doi: 10.1016/j.devcel.2005.03.006; pmid: 15809040
21. E. Nikolopoulou, G. L. Galea, A. Rolo, N. D. E. Greene, A. J. Copp, Neural tube closure: Cellular, molecular and biomechanical mechanisms. *Development* **144**, 552–566 (2017). doi: 10.1242/dev.145904; pmid: 28196803
22. G. G. Martins *et al*., Dynamic 3D cell rearrangements guided by a fibronectin matrix underlie somitogenesis. *PLOS ONE* **4**, e7429 (2009). doi: 10.1371/journal.pone.0007429; pmid: 19829711
23. J. Lu *et al*., Basement membrane regulates fibronectin organization using sliding focal adhesions driven by a contractile winch. *Dev. Cell* **52**, 631–646.e4 (2020). doi: 10.1016/j.devcel.2020.01.007; pmid: 32004443
24. S. Amin *et al*., Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Rep.* **17**, 3165–3177 (2016). doi: 10.1016/j.celrep.2016.11.069; pmid: 28009287
25. G. La Manno *et al*., RNA velocity of single cells. *Nature* **560**, 494–498 (2018). doi: 10.1038/s41586-018-0414-6; pmid: 30089906
26. M. Buckingham *et al*., The formation of skeletal muscle: From somite to limb. *J. Anat.* **202**, 59–68 (2003). doi: 10.1046/j.1469-7580.2003.00139.x; pmid: 12587921
27. S. Aramaki *et al*., A mesodermal factor, T, specifies mouse germ cell fate by directly activating germline determinants. *Dev. Cell* **27**, 516–529 (2013). doi: 10.1016/j.devcel.2013.11.001; pmid: 24331926
28. M. Saitou, M. Yamaji, Primordial germ cells in mice. *Cold Spring Harb. Perspect. Biol.* **4**, a008375 (2012). doi: 10.1101/cshperspect.a008375; pmid: 23125014
29. S. Grosswendt *et al*., Epigenetic regulator function through mouse gastrulation. *Nature* **584**, 102–108 (2020). doi: 10.1038/s41586-020-2552-x; pmid: 32728215
30. D. L. Chapman, V. E. Papaioannou, Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature* **391**, 695–697 (1998). doi: 10.1038/35624; pmid: 9490412
31. D. Concepcion *et al*., Cell lineage of timed cohorts of *Tbx6*-expressing cells in wild-type and *Tbx6* mutant embryos. *Biol. Open* **6**, 1065–1073 (2017). doi: 10.1242/bio.026203; pmid: 28606934
32. A. Ranga *et al*., Neural tube morphogenesis in synthetic 3D microenvironments. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6831–E6839 (2016). doi: 10.1073/pnas.1603529113; pmid: 27742791

ACKNOWLEDGMENTS

We are grateful for the support and feedback received from members of the Herrmann and Meissner laboratories, in particular S. Grosswendt, Z. Smith, and A. Taguchi. We thank D. Micic and J Fiedler for animal care; N. Mages for assistance with (sc)RNA-seq; C. Giesecke-Thiel and U. Marchfelder for assistance with FACS; T. Mielke and B. Fauler for help with microscopy; S. Währisch, F. Tobor, I. Gassaloglu, and P. Burton for technical assistance; and N. Rajewsky (MDC/BIMSB) for providing access to the NanoString. **Funding:** Supported by an Alexander von Humboldt Fellowship (J.V.V.), NIH grant HG006193 (A.M.), and the Max Planck Society. **Author contributions:** B.G.H. initiated the study; J.V.V. and B.G.H. conceived the project; J.V.V., A.M., and B.G.H. supervised the project; J.V.V., A.B., and L.H. designed, performed, and quantified most experiments; H.K. performed bulk and scRNA-seq computational analysis with help from J.V.V. and A.B.; M.S.-W., D.S., F.K., M.P., and A.S.K. generated and derived the mESC reporter lines; L.W. performed tetraploid complementation; S.H. performed pilot experiments to optimize culture media; R.B. helped with image acquisition and performed image analysis with L.G., A.B., and J.V.V.; B.T. supervised next-generation sequencing; J.V.V. drafted the first version of the manuscript; and the final manuscript was written by J.V.V., A.B., A.M., and B.G.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Bulk and scRNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE141175. Computational code is available at https://github.com/HeleneKretzmer/TLS and https://github.com/leoguignard/TLS-morpho.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eaba4937/suppl/DC1
Materials and Methods
Figs. S1 to S24
Movies S1 to S10
Data S1 to S10
References (*33*–*71*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

9 December 2019; resubmitted 13 May 2020
Accepted 24 September 2020
10.1126/science.aba4937

RESEARCH ARTICLE SUMMARY

DEVELOPMENTAL BIOLOGY

Developmental clock and mechanism of de novo polarization of the mouse embryo

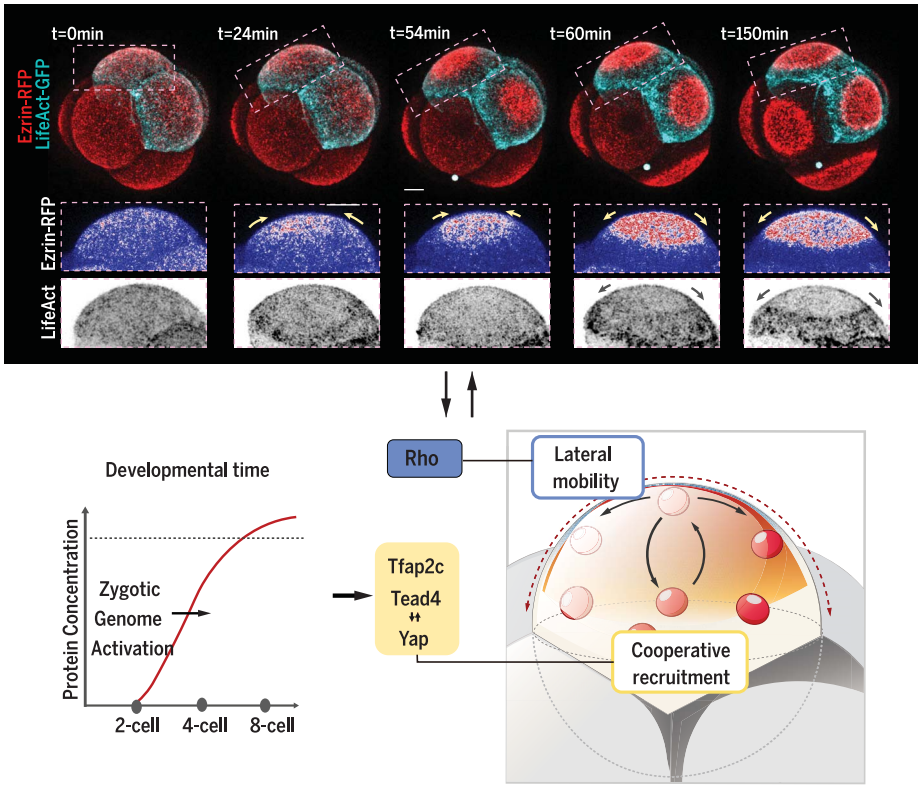
Meng Zhu, Jake Cornwall-Scoones, Peizhe Wang, Charlotte E. Handford, Jie Na, Matt Thomson, Magdalena Zernicka-Goetz*

INTRODUCTION: During preimplantation development, the establishment of apicobasal cell polarity is key for the transition from totipotency to pluripotency, which induces cell differentiation toward trophectoderm (TE). In the mouse embryo, this event is programmed to occur at the eight-cell stage, and this timing follows an intrinsic developmental clock that is independent of embryo size or cell cycle progression. Despite the importance of apical domain formation, the molecular mechanisms to establish cell polarization and the temporal regulation of this event in mouse and human embryos have remained largely elusive.

RATIONALE: In different mammalian species, zygotic genome activation (ZGA) is evolutionarily conserved to occur before the establishment of cell polarization. We therefore hypothesized that zygotic transcription regulates the timing of polarization. To test this, we deployed assays to alter the cellular concentration of zygotic transcripts and to assess the consequences of these changes for the timing of embryo polarization. We also performed an RNA interference (RNAi) screen on 124 zygotically expressed transcripts to determine the molecular identity of zygotic transcripts crucial to cell polarization. Finally, we combined cutting-edge imaging methods with biophysical modeling to account for how the factors we identified regulate the de novo establishment of cell polarization.

RESULTS: Cell polarity in the mouse embryo is marked by the appearance of a cap-shaped apical domain. Consistent with our hypothesis, an increase or decrease of zygotic transcripts respectively accelerated or inhibited apical domain formation. Our RNAi screen identified two transcription factors—transcription factor AP-2 gamma (Tfap2c) and TEA domain transcription factor 4 (Tead4), which play a redundant role in regulating cell polarization timing. Both Tfap2c and Tead4 proteins accumulate after ZGA, and elevation of their expression allows polarity proteins to anchor to the apical surface prematurely at the four-cell stage. However, these apical proteins failed to organize into an expanded apical domain, instead becoming hypercentralized to form membrane protrusions. This indicates that an additional condition is required for the apical domain formation. We have previously characterized that Rho guanosine triphosphatase (GTPase) signaling, which regulates the actomyosin apical localization and becomes activated around the eight-cell stage, is important for cell polarity. In this study, we found that premature activation of Rho GTPase with expression of Tfap2c and Tead4 allows a complete, precocious induction of the apical domain, leading to the premature expression of TE transcription factors and to morphogenesis events downstream of cell polarization. By combining quantitative imaging measurements and mathematical modeling, we show that apical domain formation is driven by the dynamic interplay between two key processes: (i) the cooperative recruitment of ezrin via the actin network and (ii) the lateral mobility of ezrin on the membrane. On the basis of the experimental evidence and biophysical simulations of these interactions, we show that Tfap2c and Tead4 control the cooperative recruitment of ezrin, whereas RhoA promotes membrane mobility.

CONCLUSION: The timing and mechanisms for cell polarization have remained largely unknown. We now identify molecules that are necessary and sufficient for the de novo establishment of cell polarization in the mouse embryo. Our results indicate a direct role of ZGA in regulating the timing of cell polarization. Beyond identifying the key molecules sufficient to establish cell polarization, we also provide biophysical understanding of the mechanism by which these molecules act to build cell polarization in the mammalian embryo. ■



Molecular mechanism and temporal regulation for the apical domain formation. Zygotic genome activation enables the expression of Tfap2c and Tead4, which regulates the cooperative membrane recruitment of apical proteins. Cooperative recruitment interacts with Rho GTPase–regulated apical protein lateral mobility to establish the apical domain.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: mz205@cam.ac.uk
Cite this article as M. Zhu *et al.*, *Science* **370**, eabd2703 (2020). DOI: 10.1126/science.abd2703

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abd2703

RESEARCH ARTICLE

DEVELOPMENTAL BIOLOGY

Developmental clock and mechanism of de novo polarization of the mouse embryo

Meng Zhu[1], Jake Cornwall-Scoones[2], Peizhe Wang[3], Charlotte E. Handford[1,4], Jie Na[3], Matt Thomson[2], Magdalena Zernicka-Goetz[1,2,4]*

Embryo polarization is critical for mouse development; however, neither the regulatory clock nor the molecular trigger that it activates is known. Here, we show that the embryo polarization clock reflects the onset of zygotic genome activation, and we identify three factors required to trigger polarization. Advancing the timing of transcription factor AP-2 gamma (Tfap2c) and TEA domain transcription factor 4 (Tead4) expression in the presence of activated Ras homolog family member A (RhoA) induces precocious polarization as well as subsequent cell fate specification and morphogenesis. Tfap2c and Tead4 induce expression of actin regulators that control the recruitment of apical proteins on the membrane, whereas RhoA regulates their lateral mobility, allowing the emergence of the apical domain. Thus, Tfap2c, Tead4, and RhoA are regulators for the onset of polarization and cell fate segregation in the mouse.

The totipotent mammalian zygote can produce any embryonic or extraembryonic tissue, but this ability becomes restricted in the first cell fate decision that generates distinct inner cell mass (ICM) and outer extraembryonic trophectoderm (TE). The ICM will form epiblast (EPI) and extraembryonic primitive endoderm (PE), generating the future fetus and the yolk sac, respectively. The TE will form the placenta. Formation of these three lineages by implantation is a prerequisite for successful pregnancy.

Embryo polarization is key to the segregation of the ICM and TE lineages (*1*, *2*). In the mouse, this process happens at the eight-cell stage (*2*–*4*), when each blastomere acquires an apical domain, comprising the Par protein complex and ERM proteins (ezrin, radixin, and moesin) enclosed by an actomyosin ring (*5*, *6*). The apical domain enables expression of transcription factors such as Cdx2 and Gata3, which drive differentiation into TE, whereas apolar cells maintain pluripotency to become ICM (*7*, *8*).

Mammalian embryo development is regulative, yet the timing of embryo polarization remains unchanged even if embryos are split into individual blastomeres, when cells are aggregated together, or when cell divisions are prevented. Thus, the polarization seems set to a strict developmental clock that is independent of cell number (*9*, *10*). Here, we show that this clock reflects activation of the zygotic genome, and we identify three factors whose convergent activity triggers self-organization of the apical domain.

[1]Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK. [2]Division of Biology and Biological Engineering, California Institute of Technology (Caltech), Pasadena, CA 91125, USA. [3]Centre for Stem Cell Biology and Regenerative Medicine, School of Medicine, Tsinghua University, Beijing 100084, China. [4]Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK.
*Corresponding author. Email: mz205@cam.ac.uk

Critical threshold of transcripts required for embryo polarization

Polarization timing varies between species, reflecting the onset of zygotic genome activation (ZGA) (*11*, *12*). In mouse, the major wave of ZGA occurs at the two-cell stage, but an additional transcriptional wave also occurs at the early eight-cell stage, just before polarization (*13*). To determine whether the latter transcriptional wave is associated with polarization, we first treated embryos from the early eight-cell stage with two transcription inhibitors: 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) and triptolide. Each drug prevented apical localization of the polarity marker Pard6 (Fig. 1, A to D, and fig. S1, A to G). Polarization was restored by washing out the reversible transcription inhibitor DRB (fig. S1, H to J). Thus, transcription at the early eight-cell stage appears to be required for embryo polarization. As a second test, we reduced the cytoplasmic volume (Fig. 1E), shown by others to increase the concentration of newly synthesized mRNA (*14*, *15*). We injected zygotes with an apical marker [ezrin–red fluorescent protein (RFP) mRNA] and resected 30 to 40% of the cytoplasm from a two- or four-cell-stage blastomere using a technique that does not compromise development (*16*) (Fig. 1F). Single-molecule fluorescence in situ hybridization confirmed that this reduction increased the concentration of newly transcribed mRNAs (*17*) (fig. S2, A to E). Such blastomere resection advanced embryo polarization by 2.1 hours when performed at the two-cell stage (Fig. 1, F to I; N = 62 pairs; movie S1) and by 3.3 hours at the four-cell stage (fig. S2, F to H, and Fig. 1J; N = 76 pairs; movie S2). Both experimental and control embryos established all three lineages as blastocysts (fig. S2, I to K). Inhibiting transcription with a 3-hour pulse of DRB led both resected and control blastomeres to polarize simultaneously (Fig. 1K and fig. S2, L to N). These results indicate that de novo synthesis of transcripts and their accumulation to a critical threshold is required for embryo polarization.

Tfap2c and Tead4 are required for embryo polarization

We hypothesized that the requirement for eight-cell-stage transcription could result from direct expression of cytoskeletal regulators of cell polarization or their indirect expression through a transcriptional hierarchy. We therefore interrogated published single-cell RNA sequencing (RNA-seq) data (*18*) and selected genes for 118 cytoskeletal polarity regulators and six transcription factors that show increased expression by the eight-cell stage and hence are likely to be active according to assay for transposase-accessible chromatin sequencing (ATAC-seq) (*19*) (fig. S3 and tables S1 and S2). We down-regulated each of these 124 genes by RNA interference (RNAi) (figs. S3 and S4, A and B) and scored the timing of embryo polarization from time-lapse imaging of the distribution of ezrin-RFP (Fig. 2, A to D). Only depletion of the transcription factors Tfap2c (transcription factor AP-2 gamma) or Tead4 (TEA domain transcription factor 4) prevented embryo polarization (Fig. 2, A to C, and E, and fig. S4, A to E). Individual depletion of Tfap2c and Tead4 delayed polarization from the 8- to 16-cell stage (Fig. 2, A to C, E, and F), whereas polarization was entirely abolished by their co-depletion (Fig. 2, D to H). Depletion of Tfap2c and Tead4 also prevented precocious polarization resulting from blastomere resection (fig. S4, F and G).

To confirm the requirement for Tfap2c and Tead4 in polarization, we deleted both genes by CRISPR-Cas9 mutagenesis. We designed three single-guide RNAs (sgRNAs) to target a single protein-coding exon of each gene (Fig. 2I) and injected them into the zygote together with Cas9 mRNA and ezrin-RFP mRNA as an apical marker. We categorized blastomeres on the basis of whether they had undetectable, moderate, or wild-type levels of Tfap2c or Tead4 proteins at the 8- to 16-cell stage (fig. S5, A and B) and confirmed by DNA sequencing that blastomeres with undetectable Tfap2c or Tead4 were homozygous mutants (Fig. 2J and fig. S5, C and D). Simultaneous deletion of *Tfap2c* and *Tead4* completely abolished embryo polarization, in contrast to their individual deletions, which were less severe (Fig. 2, J to L, and fig. S5, E and F). Thus, zygotic expression

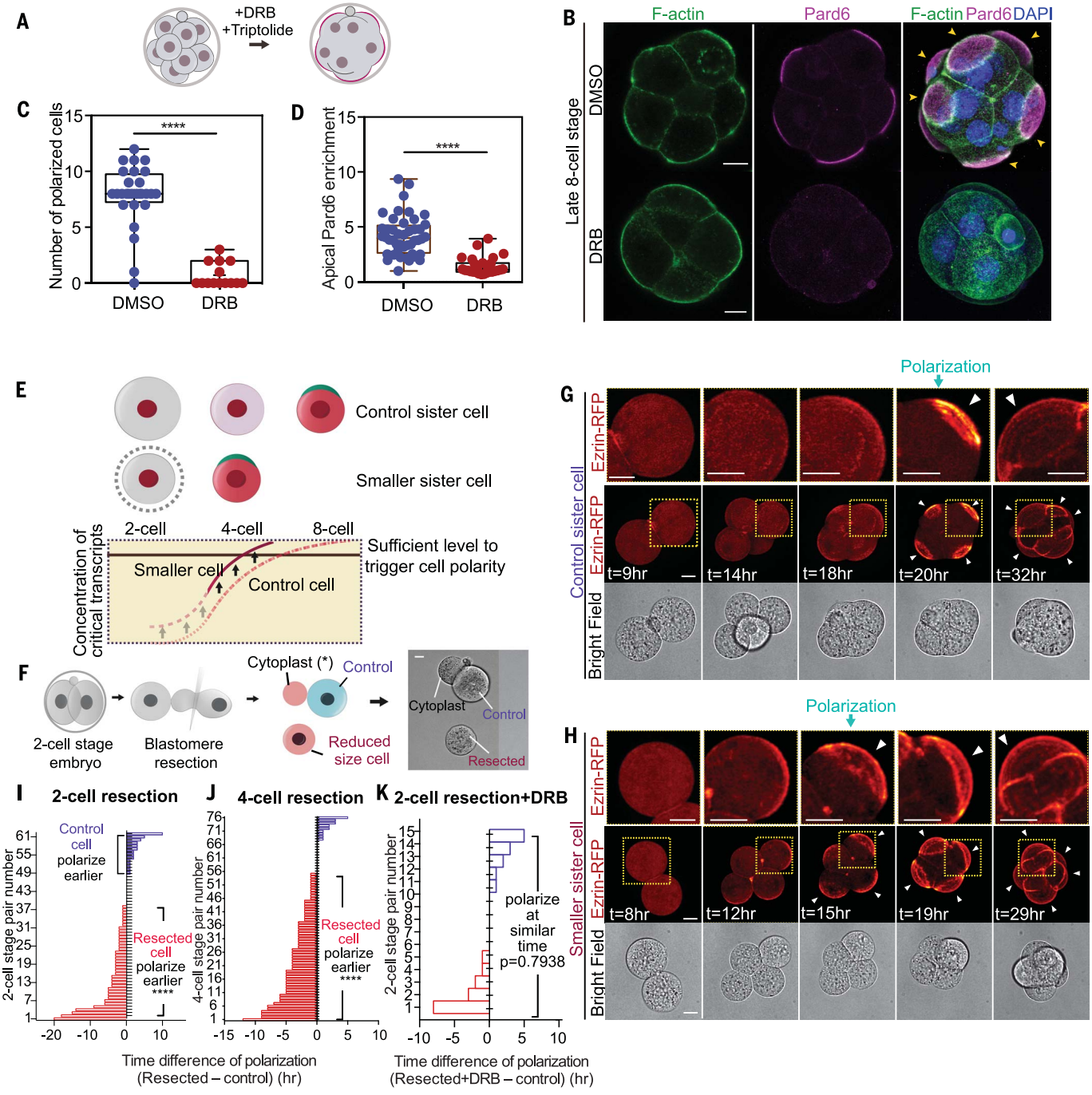


Fig. 1. The dependency of polarization on nascent transcripts. (**A**) Scheme indicating inhibitor treatments. (**B**) Dimethyl sulfoxide (DMSO)–treated (control) or DRB-treated 8- to 16-cell embryos were analyzed for the localization of F-actin, Pard6, and DNA. Arrowheads indicate the apical domain. (**C**) Polarized cell number in DMSO (control) or DRB-treated embryos. ****$P < 0.0001$, Mann-Whitney U test. $N = 2$ experiments. (**D**) Apical enrichment of Pard6 (see methods) in 8- to 16-cell-stage cells treated with DMSO (control) or DRB. ****$P < 0.0001$. Mann-Whitney U test. (**E**) Scheme of the hypothesis: Newly synthesized factors important for polarization accumulate up to a point at which polarization is induced at the eight-cell stage. Decreasing the cell size elevates the concentration of such factors, leading to an advance in polarization timing. (**F**) Scheme showing the blastomere resection procedure. (**G** and **H**) Time-lapse of control or smaller sister blastomeres from the experiment described in (F). Arrowheads indicate the apical domain. Dotted yellow squares indicate the magnified regions (top row). (**I** and **J**) Polarization time difference (see methods) between smaller and control sister blastomeres from (F) or fig. S2F; each bar represents one comparison. Smaller cells polarize earlier in the significant majority of cases. $N = 13$ experiments for (I), $N = 6$ experiments for (J). ****$P < 0.0001$, one-sample t test, hypothetical mean = 0. (**K**) Polarization time difference between control and smaller DRB-treated sister cells, from experiments in fig. S2L. Each bar represents one comparison. Pulsed transcription inhibition prevents the early polarization of smaller cells. $N = 3$ experiments. One-sample t test, hypothetical mean = 0. Arrowheads indicate the apical domain. Scale bars, 15 μm.

of *Tfap2c* and *Tead4* is required for embryo polarization at the eight-cell stage.

Tead4 had previously been shown to function only downstream of polarization, after nuclear relocalization of its transcriptional coactivator Yap, to induce TE transcription factor expression (7). To gain further insight into the earlier role of Tead4, we examined the localization of Yap. We found Yap localization in the nucleus before polarization at the eight-cell stage (20) (fig. S6, A to C) that was diminished by down-regulation (fig. S6, D and E) and enhanced by up-regulation of Tead4 (fig. S6, F and G). Thus, Tead4 affects the localization of Yap before polarization,

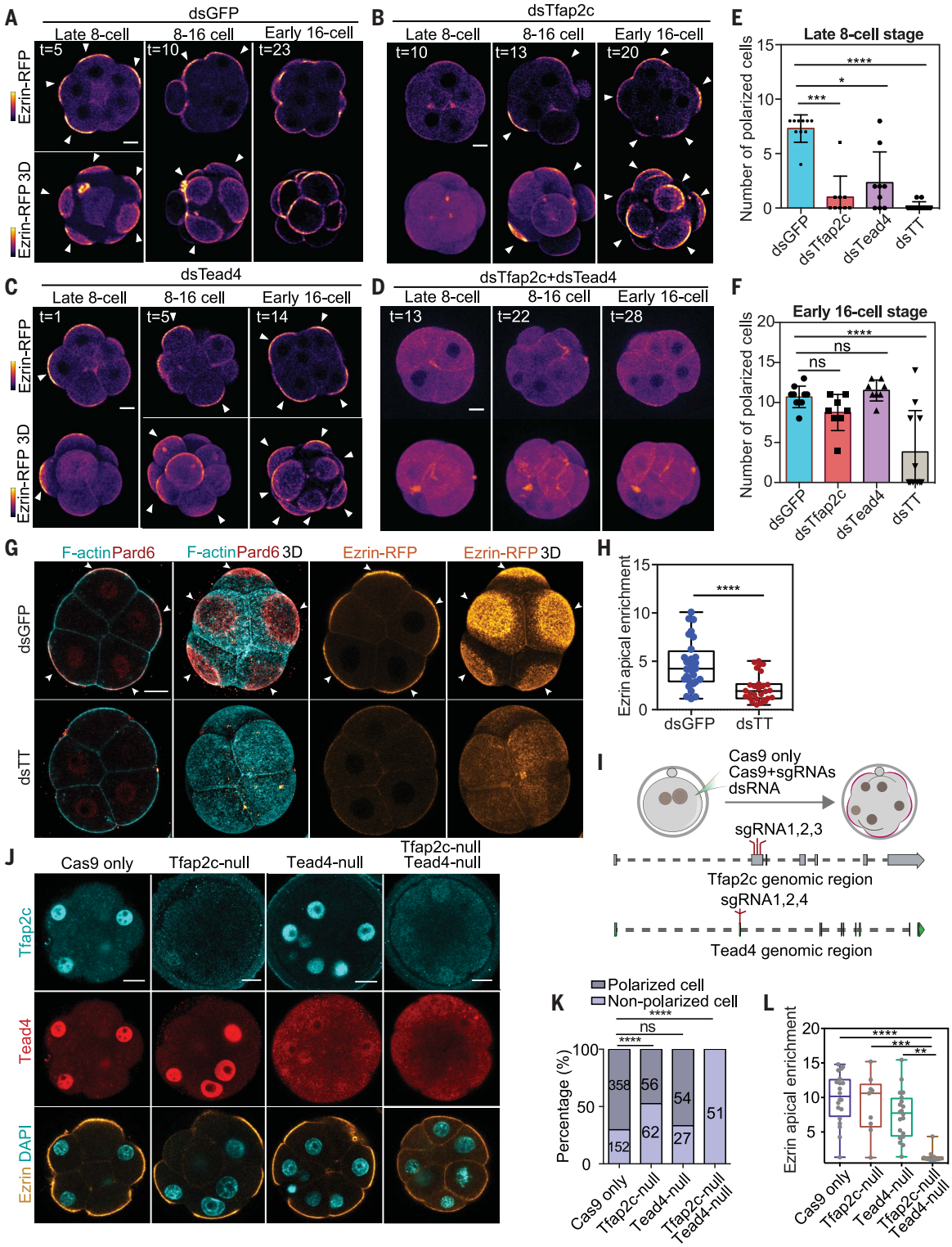


Fig. 2. Zygotic expression of Tfap2c and Tead4 is essential for polarization. (**A** to **D**) Time lapse of ezrin-RFP localization in embryos with or without Tfap2c and/or Tead4. Tfap2c and Tead4 co-depletion causes polarization failure until the 16-cell stage. Arrowheads indicate the apical domain. (**E** and **F**) Polarized cell number in different conditions and stages. Each dot, square, or triangle represents an embryo. dsTT, dsTfap2c+dsTead4. ns, not significant; **P* = 0.0306; ****P* = 0.0006; *****P* < 0.0001. Kruskal-Wallis test for (E), one-way analysis of variance (ANOVA) test for (F). (**G**) F-actin, Pard6, and ezrin localization in late eight-cell-stage embryos injected with double-strand RNA targeting GFP (dsGFP; control) or dsTT. (**H**) Quantification of ezrin apical enrichment. *****P* < 0.0001, Student's *t* test. (**I**) CRISPR-Cas9 strategy to deplete Tfap2c and Tead4. (**J**) Tfap2c and Tead4 protein levels in wild-type (Cas9 mRNA, control), Tfap2c-depleted (with Tfap2c sgRNAs), Tead4-depleted (with Tead4 sgRNAs), Tfap2c and Tead4–co-depleted (with sgRNAs targeting both Tfap2c and Tead4), and ezrin-RFP–expressing embryos. (**K**) Proportions of polarized cells in different genotypes presented in (J). The number of cells analyzed is shown within each bar. *****P* < 0.0001, Fisher's exact test. *N* = 2 experiments. Tfap2c and Tead4 co-depletion represses apical domain formation. (**L**) Quantifications of ezrin apical enrichment. Each dot represents a cell. ***P* = 0.0012, ****P* = 0.0007, *****P* < 0.0001, Kruskal-Wallis test. *N* = 4 experiments. Scale bars, 15 μm.

indicating a previously undescribed, polarity-independent Tead4 function.

Advancing expression of Tfap2c, Tead4, and Rho GTPase advances polarization timing

Having found that Tfap2c and Tead4 are required for embryo polarization, we sought to determine whether advancing their expression would advance the timing of polarization. We therefore injected Tfap2c and Tead4 mRNAs into one blastomere at the late two-cell stage to elevate their expression at the four-cell stage (fig. S7, A to C). Advancing expression of Tead4 did not induce premature polarization (fig. S7, D, E, I, and J). Advancing expression of Tfap2c led to formation of cell protrusions enriched in apical polarity proteins, including Pard6 and ezrin at the four-cell stage (fig. S7, F to I). Advancing the expression of Tfap2c and Tead4 together also induced premature formation of protrusions (fig. S7, G to I; Fig. 3, B and C; and movies S3 and S4). In all cases, the induced protrusions were smaller than the natural apical domain at the eight-cell stage (Fig. 3B; movies S3 and S4; and fig. S7H). These results suggested that Tfap2c and Tead4 might be sufficient to trigger polarization of apical proteins but that other factors are required for proper apical domain formation.

We previously found that actomyosin activation by protein kinase C–Rho guanosine triphosphatase (GTPase) signaling was necessary but not sufficient to trigger apical domain formation (Fig. 3, B and C, and movie S5) (*21*). We therefore hypothesized that Rho GTPase activation might be required in addition to Tfap2c and Tead4 to achieve complete polarization. To test this, we injected mRNAs for Tfap2c and Tead4 at the two-cell stage (with ezrin-RFP mRNA as apical marker) and mRNA for constitutively active Ras homolog family member A (RhoA)–Q63L (Gln^{63}→Leu) at the four-cell stage [with LifeAct–green fluorescent protein (GFP) as injection marker] (Fig. 3A and fig. S7K). Expression of all three factors established complete apical domains at the four-cell stage (Fig. 3, B and C, and movie S6). These induced apical domains were enriched with ezrin and Pard6, strongly resembling apical domains that normally form only at the eight-cell stage (Fig. 3B and fig. S8, B and C).

To confirm these results, we overexpressed these factors in just half the embryo and found that the targeted blastomeres polarized earlier than controls in the other embryo half (Fig. 3, D to F, and movies S7 and S8). We observed no differences in division timing between blastomeres, suggesting that induced four-cell-stage polarization is not caused by a cytokinesis delay (fig. S8A). Our results indicate that a transcriptional program triggered by Tfap2c and Tead4 alongside activation of actomyosin downstream of Rho GTPase signaling triggers polarization at a specific stage of embryogenesis.

Advancing Tfap2c, Tead4, and Rho GTPase expression advances morphogenesis and differentiation

Embryo polarization at the eight-cell stage is followed by the zippering of adjacent apical domains, which expand and seal their boundaries at the late 16-cell stage to enable blastocyst formation (*22*). To determine whether premature polarization could also advance the zippering process, we induced expression of Tfap2c, Tead4, and RhoA-Q63L either in the whole embryo or in half of the embryo to trigger four-cell-stage polarization and followed subsequent development by time-lapse microscopy. The induced premature polarization resulted in zippered domains associated with the tight junction protein ZO-1, not at the 16-cell stage but at the eight-cell stage (fig. S8, B to F). Thus, embryo polarization is sufficient to advance the subsequent step of embryogenesis leading to blastocyst formation.

As polarization in the mouse embryo is followed by cell fate specification, we determined whether overexpressing Tfap2c, Tead4, and RhoA-Q63L to induce polarization at the four-cell stage would advance differentiation of cells inheriting an apical domain into TE. Premature polarization induced premature expression of the TE transcription factors Cdx2 and Gata3 in a cell-autonomous manner (fig. S8, G to K, and movies S9 and S10). Thus, the combined activities of Tfap2c, Tead4, and RhoA-Q63L are sufficient to advance the timing not only of polarization but also the differentiation program.

Tfap2c and Tead4 are required for apical protein centralization

To define the relative roles of RhoA, Tfap2c, and Tead4 in driving polarization, we visualized events leading to apical domain formation in living embryos expressing LifeAct-GFP and ezrin-RFP from the mid to late 8-cell stages. During apical protein polarization, ezrin-RFP first became concentrated at the center of the cell contact–free surface to form an apical patch concomitant with a local reduction of actin. We refer to this stage as centralization (Fig. 4, A and B, and movie S11). This apical patch of ezrin-GFP expanded, and actin became concentrated in a ring around it. We refer to this phase as expansion (Fig. 4, A and C, and movie S11). Downregulation of Tfap2c and Tead4 diminished the initial centralization of ezrin-RFP (Figs. 2G and 4D), implying that centralization is required for apical domain formation.

Tfap2c and Tead4 control polarized growth of apical protein clusters

Imaging of apical domain centralization in mid to late eight-cell-stage embryos with higher temporal-spatial resolution (movie S12) revealed that ezrin formed clusters that colocalized with actin clusters when the embryo had just compacted but was not yet polarized (Fig. 4, E and J). As polarization progressed, the ezrin clusters grew; the more-distant clusters grew faster than those near cell-cell contacts, resulting in ezrin enrichment toward the middle of the cell-contact free surface (Fig. 4, E and F). The amount of membrane-associated ezrin increased as the ezrin clusters grew (Fig. 4G), suggesting that cluster growth is driven by ezrin's recruitment to the membrane.

Overexpression of Tfap2c and Tead4 led to an increase in membrane enrichment of ezrin and precocious growth of both ezrin and actin clusters at the late-four-cell stage (Fig. 4, H, I, and L to N). By contrast, Tfap2c and Tead4 depletion decreased ezrin's membrane enrichment and prevented growth of ezrin and actin clusters at the mid-eight-cell stage (Fig. 4, J to N). Together, this suggests that Tfap2c and Tead4 are required for the growth of actin and apical protein clusters, recruitment of apical protein to the membrane, and centralization of apical protein.

Tfap2c and Tead4 regulate actin dynamics to promote apical protein cluster growth

As our results suggested that Tfap2c and Tead4 could regulate actin dynamics to direct the growth of apical protein clusters, we next examined the cortical movements generated by the contractile actomyosin network during apical domain formation (*23*) (movie S13). We first tested whether such cortical movements can drive asymmetric growth of apical protein clusters by tracking LifeAct and ezrin clusters using particle image velocimetry (PIV) (movie S14; and see methods section in the supplementary materials). In contrast to actin flows of postmitotic cells (*22*) (fig. S9C), PIV did not detect any obvious movement toward the center of the cell-contact free surface (fig. S9, A and B). Accordingly, we found that inhibiting actin flows with blebbistatin failed to prevent asymmetric ezrin cluster growth (fig. S9, D to G, and movie S15). Thus, asymmetric clustering of ezrin is not driven by the cortical flow mediated by actomyosin contractility.

Time-lapse observations revealed that ezrin cluster growth occurred during the merging and splitting of actin clusters and was unimpeded by blebbistatin (Fig. 5A), suggesting that cortical actin remodeling may allow ezrin's recruitment to growing clusters. Consistently, perturbing actin depolymerization in eight-cell embryos with Jasplakinolide (JASP) prevented apical domain formation (fig. S9, H and K). Moreover, when we inhibited the Arp2/3 complex with CK666 to prevent actin nucleation, ezrin cluster growth (Fig. 5, B and C) and apical domain formation (fig. S9, I and K) were also blocked. By contrast, treatment of

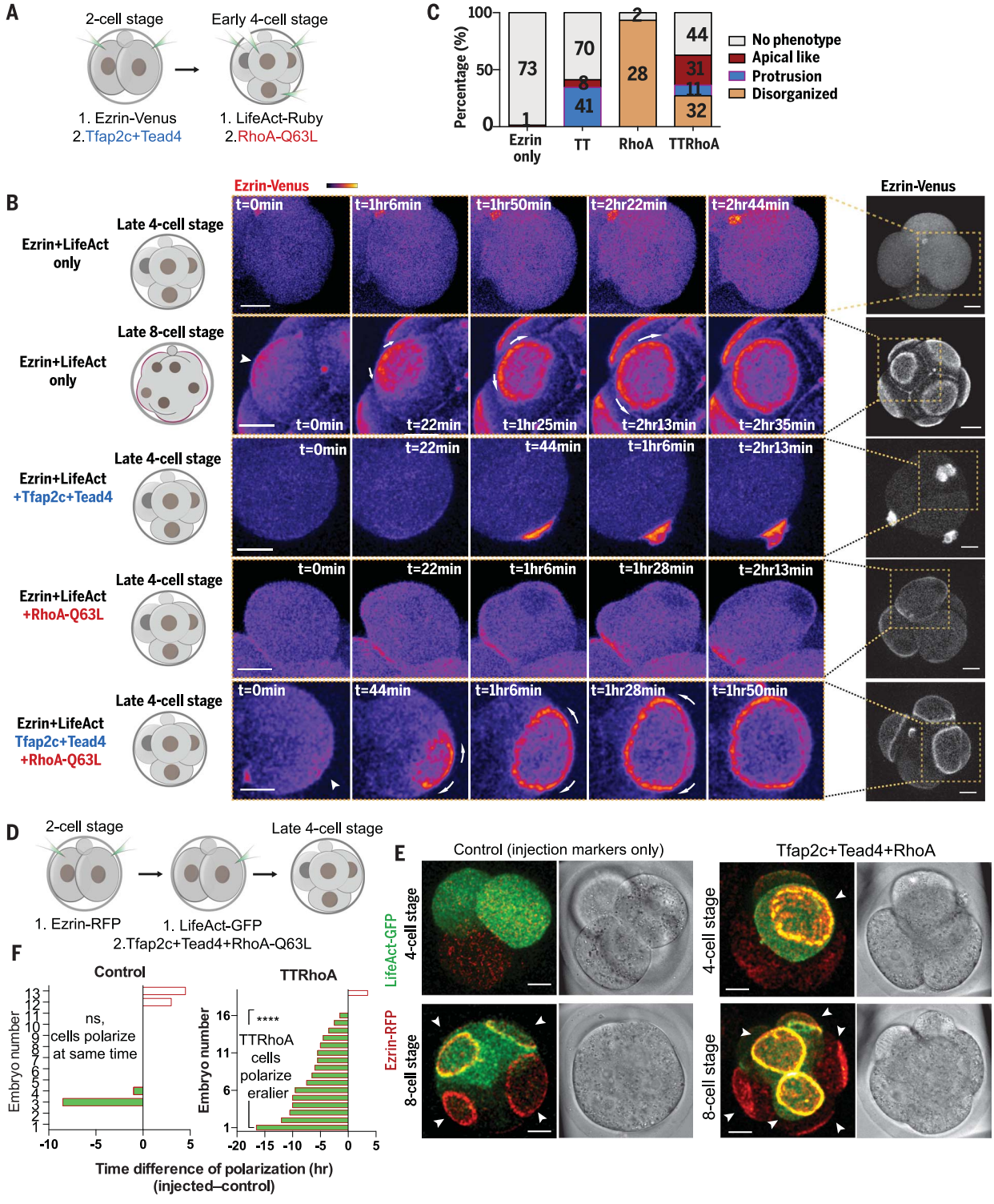


Fig. 3. Premature expression of Tfap2c, Tead4, and activated RhoA is sufficient to advance polarization timing. (**A**) Scheme of Tfap2c, Tead4, and RhoA-Q63L overexpression. (**B**) Ezrin-Venus dynamics in late four-cell or eight-cell stage control cells (ezrin-Venus only), or in late four-cell stage cells overexpressing (i) Tfap2c+Tead4 (TT), (ii) RhoA-Q63L (RhoA), or (iii) Tfap2c+Tead4+RhoA-Q63L (TTRhoA). TTRhoA overexpression induces a premature, full apical domain. In all conditions, cell divisions were not affected. Arrowheads indicate ezrin-Venus enrichment at the cell-contact free surface. Arrows indicate apical domain expansion. Dashed squares indicate the magnified regions. (**C**) Quantification of morphologies induced by the conditions in (B). The number of cells analyzed is shown within each bar. (**D**) Scheme of Tfap2c, Tead4, and RhoA-Q63L overexpression. (**E**) Representative images of embryos overexpressed with ezrin-RFP, LifeAct-GFP mRNA only (control), or with Tfap2c, Tead4, and RhoA-Q63L mRNA at the four- or eight-cell stage. The cells overexpressed with Tfap2c, Tead4, and RhoA-Q63L polarize significantly earlier than control cells in the same embryos, or cells in the control embryos. Arrowheads indicate the apical domain. (**F**) Polarization time difference between cells with or without LifeAct-GFP or TTRhoA overexpression in the same embryo, or in control embryos. ****$P < 0.0001$, one-sample t test, hypothetical mean = 0. Control group, $N = 13$ embryos ; TTRhoA group, $N = 19$ embryos. $N = 2$ experiments. Scale bars, 15 μm.

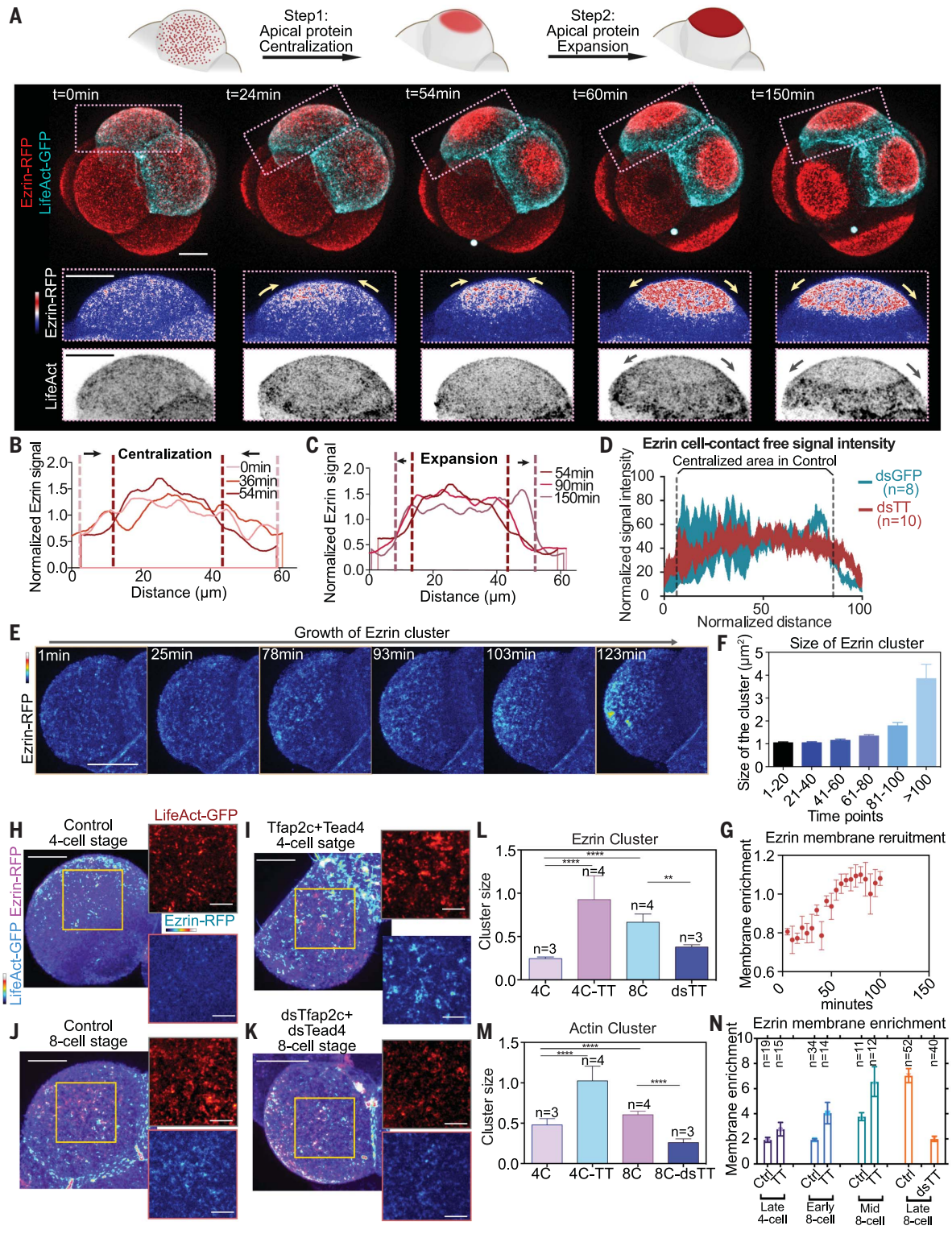


Fig. 4. Tfap2c and Tead4 regulate apical domain centralization by regulating apical protein clustering. (**A**) LifeAct-GFP and ezrin-RFP dynamics during polarization. Squares denote magnified regions; yellow and gray arrows indicate apical protein or actin ring movements. The apical domain forms after centralization and expansion steps. *N* = 7 embryos, *N* = 4 experiments. (**B** and **C**) Ezrin-RFP signal at the cell-contact free surface during centralization or expansion steps in (A). Ezrin signal is normalized against average membrane signal intensity. (**D**) Ezrin-RFP distribution on cell-contact free surface at the late eight-cell stage in dsGFP (control) or dsTfap2c+dsTead4 injected cells. Numbers indicate examined cells. *N* = 2 experiments. (**E**) Ezrin-RFP distribution during apical centralization. (**F**) Ezrin cluster size during polarization. More than 1500 clusters were analyzed for each time point. Data shown as mean ± SD. *N* = 2 experiments. (**G**) Ezrin membrane enrichment during centralization. *N* = 4 cells from four embryos. Data shown as mean ± SEM. (**H** to **K**) Localization of LifeAct-GFP and ezrin-RFP in embryos injected with or without Tfap2c+Tead4 mRNA at late four-cell stage, or embryos injected with or without dsTfap2c+dsTead4 at eight-cell stage. Yellow squares indicate the magnified regions (right). (**L** and **M**) Size of actin (L) or ezrin (M) clusters in embryos shown in (H) to (K). Data shown as mean ± SEM. $^{**}P < 0.01$, $^{****}P < 0.0001$, one-way ANOVA test. Numbers (n) indicate examined cells. More than 500 clusters were measured in each condition. (**N**) Ezrin membrane enrichment in different conditions and stages. TT, Tfap2c+Tead4 overexpression; dsTT, dsTfap2c+dsTead4 knockdown. Numbers indicate examined cells. Data presented as mean ± SEM. Scale bars for magnified images in (H) to (K), 5 μm. All other scale bars, 15 μm.

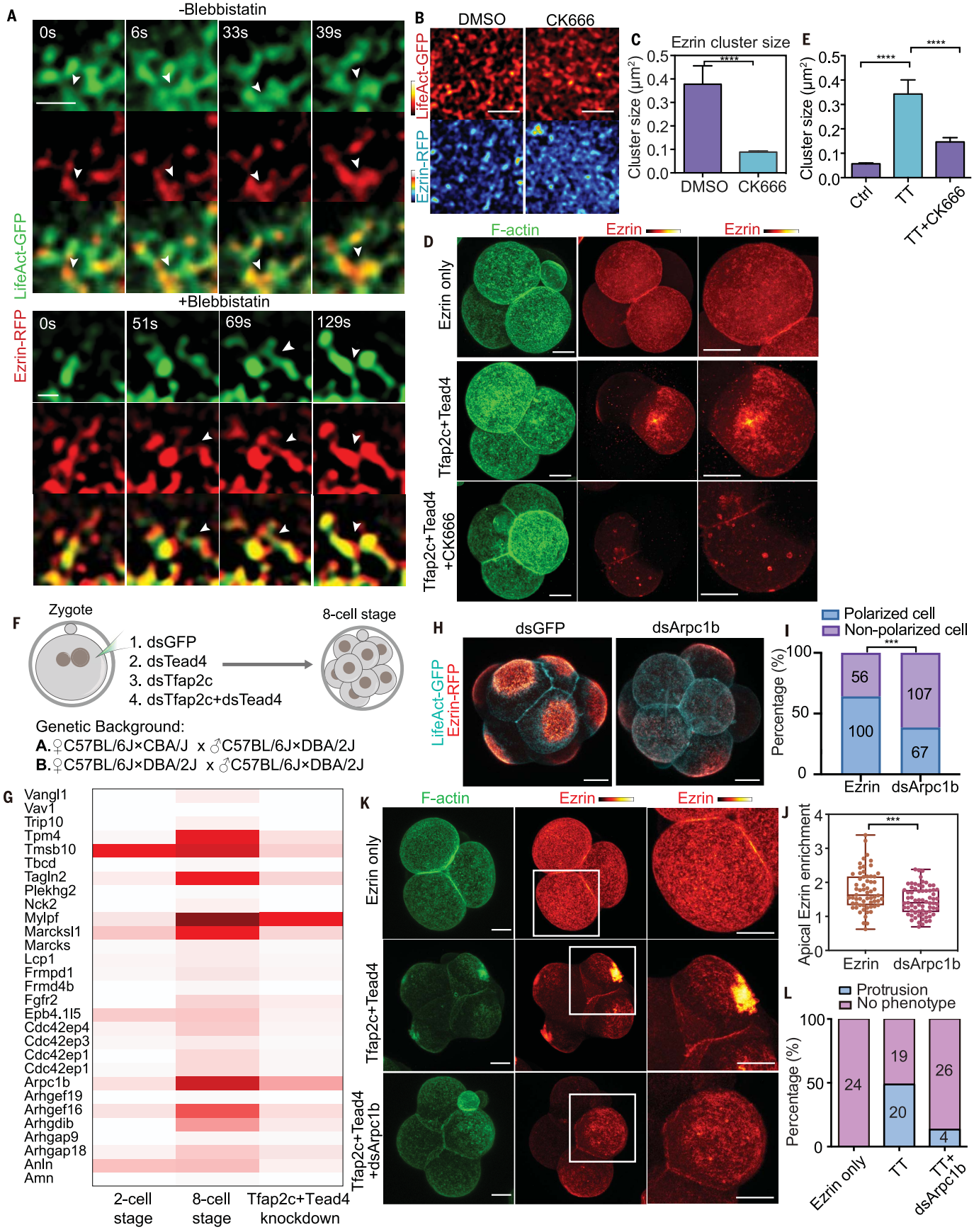

A
-Blebbistatin
0s
6s
33s
39s
LifeAct-GFP
+Blebbistatin
Ezrin-RFP
0s
51s
69s
129s
B
DMSO
CK666
LifeAct-GFP
Ezrin-RFP
C
Ezrin cluster size
Cluster size (μm²)
DMSO
CK666
E
Cluster size (μm²)
Ctrl
TT
TT+CK666
D
F-actin
Ezrin
Ezrin only
Tfap2c+Tead4
Tfap2c+Tead4 +CK666
F
Zygote
8-cell stage
1. dsGFP
2. dsTead4
3. dsTfap2c
4. dsTfap2c+dsTead4
Genetic Background:
A. ♀C57BL/6J×CBA/J x ♂C57BL/6J×DBA/2J
B. ♀C57BL/6J×DBA/2J x ♂C57BL/6J×DBA/2J
G
Vangl1
Vav1
Trip10
Tpm4
Tmsb10
Tbcd
Tagln2
Plekhg2
Nck2
Mylpf
Marcksl1
Marcks
Lcp1
Frmpd1
Frmd4b
Fgfr2
Epb4.1l5
Cdc42ep4
Cdc42ep3
Cdc42ep1
Cdc42ep1
Arpc1b
Arhgef19
Arhgef16
Arhgdib
Arhgap9
Arhgap18
Anln
Amn
2-cell stage
8-cell stage
Tfap2c+Tead4 knockdown
H
dsGFP
dsArpc1b
LifeAct-GFP
Ezrin-RFP
I
Polarized cell
Non-polarized cell
Percentage (%)
Ezrin
dsArpc1b
K
F-actin
Ezrin
Ezrin
Ezrin only
Tfap2c+Tead4
Tfap2c+Tead4 +dsArpc1b
J
Apical Ezrin enrichment
Ezrin
dsArpc1b
L
Protrusion
No phenotype
Percentage (%)
Ezrin only
TT
TT+ dsArpc1b

Fig. 5. Clustering of apical proteins is regulated by local actin dynamics. (**A**) LifeAct-GFP and ezrin-RFP dynamics with or without blebbistatin treatment during polarization. Arrowheads indicate the merging of ezrin clusters during actin polymerization. Scale bar, 1 μm. N = 5 cells for each condition. N = 3 experiments. Blebbistatin treatment did not prevent the clustering of actin or ezrin proteins. (**B**) LifeAct-GFP and ezrin-RFP localization in mid eight-cell stage embryos treated with DMSO (control) and CK666. Scale bars, 5 μm. (**C**) Ezrin cluster size in cells treated with DMSO or CK666 in (B). ****P < 0.0001, Mann-Whitney U test. More than 1500 clusters were measured in each condition, N = 2 experiments. (**D**) F-actin and ezrin-RFP localization in late four-cell-stage embryos expressing ezrin-RFP only, or with Tfap2c+Tead4 treated with or without CK666. (**E**) Ezrin cluster size in (D). ****P < 0.0001, Kruskal-Wallis test. More than 1100 clusters were calculated in each condition, N = 2 experiments. (**F**) Experimental strategy for RNA-seq. (**G**) Heatmap showing the expression of selective cytoskeleton regulators downstream of Tfap2c and Tead4. (**H**) LifeAct-GFP and ezrin-RFP localization in dsGFP or dsArpc1b injected embryos. (**I**) Polarized cell numbers in dsGFP and dsArpc1b groups. Numbers within bars represent cell number. ***P < 0.001, Fisher's exact test. (**J**) Ezrin-RFP apical enrichment in dsGFP or dsArpc1b cells. Each dot represents a cell. ***P < 0.001, Student's t test. (**K**) LifeAct-GFP and ezrin-RFP localization in different conditions. Squares indicate magnified regions (right). (**L**) Number of cells showing apical protrusions in (K). Scale bars, 15 μm.

embryos with a formin inhibitor (SMiFH2) did not affect apical domain formation (fig. S9, J and L). These observations indicate that actin remodeling is required for ezrin clustering and apical domain formation. We also found that CK666 abolished apical protein polarization induced by Tfap2c and Tead4 overexpression in four-cell embryos (Fig. 5, D and E). Thus, Tfap2c- and Tead4-dependent regulation of actin dynamics is required for the growth of ezrin clusters and apical protein formation.

Tfap2c and Tead4 control expression of actin regulators

To view events downstream of Tfap2c and Tead4, we carried out RNA-seq of eight-cell-stage embryos depleted of Tfap2c and/or Tead4. For each group of embryos (control *GFP* RNAi, *Tfap2c* RNAi, *Tead4* RNAi, and *Tfap2c*/*Tead4* co-RNAi), two biological replicates were collected with 10 embryos per sample, and experiments were performed on two strains to eliminate genetic background effects (Fig. 5F and methods). The effect of Tfap2c and Tead4 depletion was highly reproducible between biological replicates and between genetic backgrounds (fig. S10A). Depletion of Tfap2c led to the down-regulation of 749 or 929 genes (with a twofold differential cutoff), depending on the strain, whereas Tead4 depletion led to down-regulation of 242 or 314 genes (fig. S10, B and C). Their co-depletion led to an additional 135 or 95 genes being down-regulated compared with single knockdown embryos, depending on the strain (fig. S10, B and C).

A significant proportion of down-regulated genes in double-knockdown embryos had actin polymerization functions (Fig. 5G and table S3). These included known actin regulators for apical domain formation, such as Cdc42 effector protein family members (Borg) (*24*), and other actin regulators, including the Arp2/3 complex component Arpc1b, the tropomyosin protein Tpm4, Marcks and Marcksl1 proteins, and the FREM family member Ebp4.1l5, whose functions have not been explored in the mouse embryo. Expression of these actin regulators becomes up-regulated between the two- and eight-cell stages and correlates with the size increase of actin clusters during polarization (Fig. 5G). Depletion of Tfap2c and Tead4 eliminated expression of actin regulators and accordingly led to a decreased actin cluster size (Figs. 4, H and J to M, and 5G).

To test whether these actin regulators participate in apical domain formation, we depleted *Arpc1b*, *Tpm4*, *Marcksl1*, or *Ebp4.1l5* individually from two-cell embryos and determined the apical domain formation efficiency at the late eight-cell stage (Fig. 5, H to J, and fig. S11, A to D). Consistent with the effects of CK666, depletion of Arpc1b led to defective apical domain formation (Fig. 5, H to J) in natural eight-cell embryos and prevented Tfap2c- and Tead4-induced apical protein polarization at the four-cell stage (Fig. 5, K and L). Together, these results suggest that Tfap2c and Tead4 control embryo polarization by activating expression of actin regulatory proteins.

RhoA signaling reorganizes the actin network during polarization

Knowing that not only Tfap2c and Tead4 but also RhoA-Q63L were required for apical protein clustering and apical domain formation in four-cell embryos, we aimed to determine how excess RhoA activity (by overexpressing RhoA-Q63L) or reduced RhoA activity (by treatment with RhoA inhibitor C3-transferase) at the mid eight-cell stage would affect ezrin's membrane distribution (fig. S12A). Overexpression of RhoA-Q63L eliminated actin and ezrin cluster formation, resulting in the homogeneous distribution of actin and ezrin on the membrane (fig. S12, B and C). By contrast, C3-transferase treatment resulted in the ectopic clustering of actin and ezrin on the cell membrane (fig. S12, D and E), reminiscent of four-cell embryos overexpressing Tfap2c and Tead4 but lacking RhoA activity (Fig. 3B and fig. S7, G and H). Thus, RhoA signaling is required to reorganize the actin network in the embryo and thereby the clustering of apical proteins induced by Tfap2c and Tead4.

Positive feedback and lateral mobility govern apical domain formation

To gain biophysical understanding of how Tfap2c, Tead4, and RhoA regulate the timing and pattern of apical domain formation, we measured the growth of ezrin cluster size during apical protein centralization and found that it increases exponentially (Fig. 4F), suggesting involvement of a positive feedback mechanism. To gain understanding of this mechanism, we tagged ezrin with green to red photoactivatable Dendra2 fluorescent protein to track Dendra2-ezrin movement after blue light conversion. RFP signal dynamically dissipated within seconds when Dendra2 was photoconverted at the mid eight-cell stage either within or outside the nascent apical domain, suggesting rapid ezrin membrane turnover (Fig. 6, A to D). Irrespective of the site of photoconversion, ezrin-Dendra2 relocated to the nascent domain in proportion to the concentration of ezrin in this area, indicating positive feedback of ezrin on its own accumulation. This correlates with previous measurements of cooperative recruitment of ezrin to phosphatidylinositol 4,5-bisphosphate (PIP2) membranes (*25*, *26*).

To determine whether ezrin dynamics can account for apical protein centralization, we constructed a model (see methods) based on four empirically grounded assumptions (Fig. 6E): (i) a cooperative increase in ezrin binding rate (k_{on}) with increasing ezrin concentration saturating above a critical concentration (E_{crit}) suggested by the positive feedback above (Fig. 6, B and D) and known cooperative binding of ezrin to membranes (*25*, *26*); (ii) a limit to ezrin membrane loading by the finite pool of PIP2 (P_{tot}) based on colocalization of ezrin with PIP2 in the apical domain (fig. S12F) and prevention of apical domain formation by reduced PIP2 (*21*); (iii) lateral motility of ezrin along the membrane, as observed (Fig. 6, A to D, and fig. S12G), which can be modeled as effective diffusion with diffusivity (D_E); and (iv) dissociation of ezrin at a uniform rate from the membrane (k_{off}) (Fig. 6E).

We simulated ezrin dynamics in one dimension and estimated parameter values for simulations best fitting our experimental measurements (see methods and "Supplementary Modeling" section in the supplementary materials). Our simulations reproduced the dynamic changes of ezrin distribution during centralization in vivo (Fig. 6F). The model also recapitulated the ability of single cells to form a centralized ezrin domain in the absence of

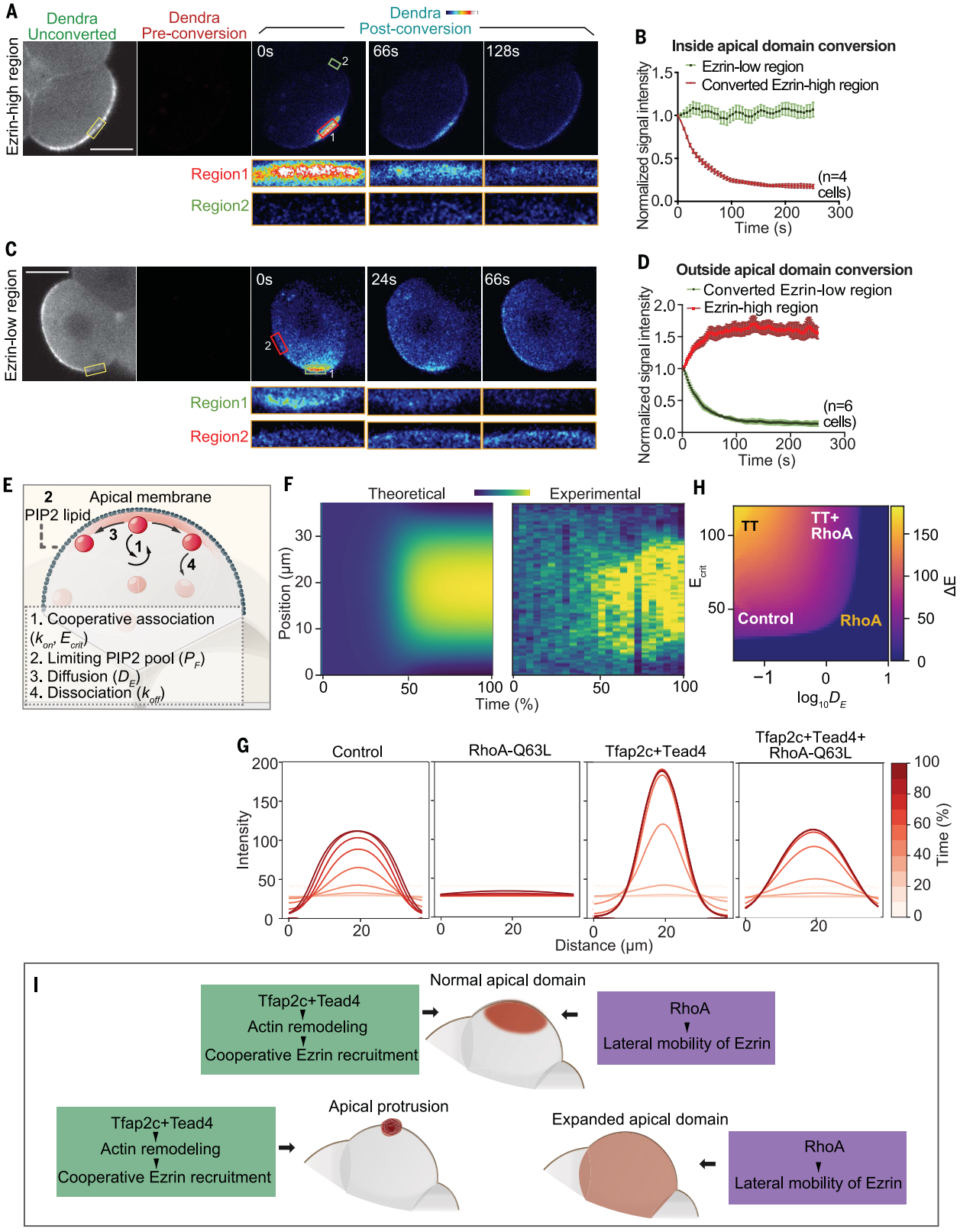


Fig. 6. Tfap2c, Tead4, and RhoA regulate apical domain formation through a positive feedback and mobility system. (**A** and **C**) Ezrin-Dendra2 localization during and after photoconversion. Yellow squares indicate converted region, red and green squares indicate ezrin-high or ezrin-low regions, respectively, measured in (B) and (D). (**B** and **D**) Signal intensity of converted ezrin-Dendra2 [experimental setting shown in (A) and (C)] in ezrin-low and ezrin-high regions within 6 min after conversion. Scale bars, 15 μm. (**E**) Structure of the biophysical model. (**F**) Kymographs comparing in silico and in vivo polarization dynamics. Same color corresponds to the same ezrin intensity. (**G**) Time courses of simulated ezrin apical distribution at different regions of parameter space. Colors corresponding to the elapsed simulation time. (**H**) Phase space of polarization shape within the monopolar regime. (**I**) Summary of the results.

cell-cell contacts (fig. S12H and methods) and the ability of cell-cell contacts to constrain apical ezrin to the center of cell-contact free surface (fig. S12I). Thus, experimental observations and computational simulation together reveal that positive feedback, lateral mobility, and competition for limiting PIP2 are sufficient to explain the apical protein centralization step.

Tfap2c, Tead4, and RhoA control cooperative recruitment and lateral mobility of apical proteins

Our model predicts that elevating the saturation threshold (E_{crit}) for cooperative recruitment of ezrin would change the steady-state distribution of membrane-bound ezrin, resulting in a narrowed peak resembling the Tfap2c and Tead4 overexpression phenotype (Fig. 6G and "Supplementary Modeling" section of the supplementary materials). This suggests that Tfap2c and Tead4 regulate the kinetics of the cooperative ezrin recruitment. To test this, we computed the rate of change in membrane-bound ezrin concentration ($\Delta E/\Delta T$) as a function of the ezrin concentration during apical centralization, for both control and Tfpa2c and Tead4 overexpressing embryos (fig. S12J). This rate of change revealed that when the ezrin concentration falls below a threshold value, local ezrin subsequently decreases, but when the local concentration exceeds this value, more ezrin gets recruited to the membrane (fig. S12J).

Myosin motor activity within the actin cortex can affect the lateral mobility of membrane proteins that can negatively regulate the formation of membrane clusters (*27*), suggesting that RhoA might regulate the lateral mobility of ezrin. To test this, we quantified the "spread" of photoconverted signal (see methods) and compared the difference between control and RhoA-Q63L–overexpressing cells (fig. S12K). Elevating RhoA activity increased the spreading effect of photoconverted ezrin on the membrane, suggesting that RhoA positively regulates the lateral mobility of ezrin (fig. S12K). In line with this, our simulations predict that increasing lateral mobility leads to homogeneously distributed ezrin (Fig. 6G) and that insufficient lateral mobility results in multiple peaks of apical protein (fig. S12L). These predictions are concordant with the phenotypes resulting from overexpression of RhoA-Q63L and depletion of RhoA in the eight-cell embryo (fig. S12, B to E).

Our model suggests that lateral mobility and the cooperative recruitment threshold have opposing effects on the shape of the apical patch (Fig. 6H and fig. S12M). Specifically, a normal apical domain will form only when both processes are activated at an appropriate level. This prediction provides a qualitative explanation for the concurrent requirements for Tfap2c, Tead4, and RhoA signaling in regulating apical domain formation (Figs. 3B and 6I).

Discussion

The importance of the first appearance of apical-basal cell polarity in mammalian development is evident from its requirement for triggering the first cell fate diversification event. Here, we show that zygotic expression of Tfap2c and Tead4 is a prerequisite for such polarization. We have been able to induce precocious embryo polarization and thereby advance subsequent embryogenesis by driving the ectopic expression of Tfap2c, Tead4, and activated Rho GTPase. Our findings help account for the temporal relationship between zygotic genome activation and the establishment of embryo polarization across multiple mammalian species (*11*, *12*).

Embryo polarization at the eight-cell stage has been viewed as a model for epithelial polarization. However, formation of the apical domain is distinct from that of many other cell types, as it can occur in the absence of external cues, such as the extracellular matrix or cell adhesion. The mechanisms behind such distinctive spontaneous symmetry-breaking properties have remained elusive. Here, we show that the initial step for symmetry breaking is the centralization of apical proteins through their two types of behavior on the membrane: actin-mediated cooperative recruitment and lateral mobility. These two processes act as opposing forces to regulate the shape of the apical domain; a cooperative recruitment mechanism enables symmetry breaking and concentration of the apical proteins, whereas lateral mobility allows apical proteins to diffuse, thereby establishing a crescent-shaped patch (Fig. 6, G and H, and fig. S12M). The balanced activity between the two processes ensures the proper shape of the apical domain because an excessive cooperative recruitment force would lead to small and often multiple domains, whereas excessive lateral mobility would lead to the uniform distribution of apical proteins and thereby inhibit symmetry breaking (Fig. 6G and fig. S12L).

Our results suggest that cooperative recruitment is regulated by actin remodeling controlled by Tfap2c and Tead4. Although the detailed mechanism is beyond the scope of this work, it is possible that ezrin is preferentially recruited to the actin structure promoted by the Arp2/3 complex in a process similar to protein condensation (*28*). In such a case, the density of the branched actin network would positively affect the saturation level of ezrin in the cooperative recruitment process. It has been observed in vitro that the actin clusters formed by Arp2/3 activity are degraded by high levels of cortical myosin (*29*), which could explain the opposing effects between the two transcription factors and RhoA in regulating the apical protein clustering and accordingly the apical domain shape.

The regulatory regime we describe is based on the behavior of ezrin, and it is likely to apply to other apical proteins, such as the Par complex, whose polarization dynamics are highly similar and also require the actin network and membrane binding, the key conditions of the process we describe (*21*). Our work illustrates how the embryo establishes cell polarization at a specific developmental stage under the regulation of stage-dependent pathways. Our results also provide a biophysical explanation for how polarization is established, indicating that positive feedback combined with lateral mobility are sufficient to drive this self-organization process. These results therefore provide insight into both the timing and mechanism of the establishment of de novo polarization in the mouse embryo, the critical event for the first cell fate specification.

Methods summary

This work followed regulations of the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 reviewed by the University of Cambridge Animal Welfare and Ethical Review. Embryos were collected from superovulated F1 females (C57BI6xCBA) crossed with F1 males. For embryo culture and inhibitor treatment, embryos were recovered at the zygote or two-cell stage in M2 medium and transferred to KSOM medium for long-term culture. The inhibitors—or the same amount of vehicles (for control conditions)—were applied to the culture. The microinjection procedure, immunostaining, static imaging and image processing, and real-time quantitative polymerase chain reaction were carried out as previously described (*21*). For the photoconversion experiment, the region of interest (ROI) covered a rectangular area of ~5 μm length and ~2 μm width on the membrane of cells expressing ezrin-Dendra2, using the midplane of the blastomeres as a reference. The ROI was illuminated at 405 nm for 5 s, after which converted proteins were imaged with a 568-nm laser at an emission wavelength between 580 and 620 nm every 2 s per frame for 5 min. The converted scanning speed is 200 Hz, and the normal scanning speed is 700 Hz. For all imaging settings, the images have been recorded using the 1024 pixel by 1024 pixel format. PIV analysis was performed using the PIVlab MATLAB algorithm (https://pivlab.blogspot.com/). For statistics, the sample distribution as well as statistical tests were performed using Prism software (www.graphpad.com/). Details of the materials and methods, as well as details for RNA-seq and modeling methods, can be found in the supplementary materials.

REFERENCES AND NOTES

1. E. Korotkevich *et al*., The apical domain is required and sufficient for the first lineage segregation in the mouse embryo. *Dev. Cell* **40**, 235–247.e7 (2017). doi: 10.1016/j.devcel.2017.01.006; pmid: 28171747
2. M. H. Johnson, C. A. Ziomek, The foundation of two distinct cell lineages within the mouse morula. *Cell* **24**, 71–80 (1981). doi: 10.1016/0092-8674(81)90502-X; pmid: 7237545
3. M. H. Johnson, C. A. Ziomek, Induction of polarity in mouse 8-cell blastomeres: Specificity, geometry, and stability. *J. Cell Biol.* **91**, 303–308 (1981). doi: 10.1083/jcb.91.1.303; pmid: 7298724
4. T. P. Fleming, P. M. Cannon, S. J. Pickering, The cytoskeleton, endocytosis and cell polarity in the mouse preimplantation embryo. *Dev. Biol.* **113**, 406–419 (1986). doi: 10.1016/0012-1606(86)90175-2; pmid: 3512332
5. S. Louvet, J. Aghion, A. Santa-Maria, P. Mangeat, B. Maro, Ezrin becomes restricted to outer cells following asymmetrical division in the preimplantation mouse embryo. *Dev. Biol.* **177**, 568–579 (1996). doi: 10.1006/dbio.1996.0186; pmid: 8806832
6. B. Plusa *et al*., Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. *J. Cell Sci.* **118**, 505–515 (2005). doi: 10.1242/jcs.01666; pmid: 15657073
7. N. Nishioka *et al*., The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev. Cell* **16**, 398–410 (2009). doi: 10.1016/j.devcel.2009.02.003; pmid: 19289085
8. A. Ralston *et al*., Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**, 395–403 (2010). doi: 10.1242/dev.038828; pmid: 20081188
9. M. H. Johnson, J. McConnell, J. Van Blerkom, Programmed development in the mouse embryo. *J. Embryol. Exp. Morphol.* **83** (suppl.), 197–231 (1984). pmid: 6085344
10. S. A. Morris, Y. Guo, M. Zernicka-Goetz, Developmental plasticity is bound by pluripotency and the Fgf and Wnt signaling pathways. *Cell Rep.* **2**, 756–765 (2012). doi: 10.1016/j.celrep.2012.08.029; pmid: 23041313
11. H. Koyama, H. Suzuki, X. Yang, S. Jiang, R. H. Foote, Analysis of polarity of bovine and rabbit embryos by scanning electron microscopy. *Biol. Reprod.* **50**, 163–170 (1994). doi: 10.1095/biolreprod50.1.163; pmid: 8312441
12. G. Nikas, A. Ao, R. M. Winston, A. H. Handyside, Compaction and surface polarity in the human embryo in vitro. *Biol. Reprod.* **55**, 32–37 (1996). doi: 10.1095/biolreprod55.1.32; pmid: 8793055
13. T. Hamatani, M. G. Carter, A. A. Sharov, M. S. Ko, Dynamics of global gene expression changes during mouse preimplantation development. *Dev. Cell* **6**, 117–131 (2004). doi: 10.1016/S1534-5807(03)00373-3; pmid: 14723852
14. O. Padovan-Merhar *et al*., Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol. Cell* **58**, 339–352 (2015). doi: 10.1016/j.molcel.2015.03.005; pmid: 25866248
15. M. Bao, J. Xie, A. Piruska, W. T. S. Huck, 3D microniches reveal the importance of cell size and shape. *Nat. Commun.* **8**, 1962 (2017). doi: 10.1038/s41467-017-02163-2; pmid: 29213086
16. M. Zernicka-Goetz, Fertile offspring derived from mammalian eggs lacking either animal or vegetal poles. *Development* **125**, 4803–4808 (1998). pmid: 9806928
17. F. Wang *et al*., RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012). doi: 10.1016/j.jmoldx.2011.08.002; pmid: 22166544
18. Q. Deng, D. Ramsköld, B. Reinius, R. Sandberg, Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193–196 (2014). doi: 10.1126/science.1245316; pmid: 24408435
19. J. Wu *et al*., The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016). doi: 10.1038/nature18606; pmid: 27309802
20. Y. Hirate *et al*., Par-aPKC-dependent and -independent mechanisms cooperatively control cell polarity, Hippo signaling, and cell positioning in 16-cell stage mouse embryos. *Dev. Growth Differ.* **57**, 544–556 (2015). doi: 10.1111/dgd.12235; pmid: 26450797
21. M. Zhu, C. Y. Leung, M. N. Shahbazi, M. Zernicka-Goetz, Actomyosin polarisation through PLC-PKC triggers symmetry breaking of the mouse embryo. *Nat. Commun.* **8**, 921 (2017). doi: 10.1038/s41467-017-00977-8; pmid: 29030553
22. J. Zenker *et al*., Expanding actin rings zipper the mouse embryo for blastocyst formation. *Cell* **173**, 776–791.e17 (2018). doi: 10.1016/j.cell.2018.02.035; pmid: 29576449
23. J. L. Maître, R. Niwayama, H. Turlier, F. Nédélec, T. Hiiragi, Pulsatile cell-autonomous contractility drives compaction in the mouse embryo. *Nat. Cell Biol.* **17**, 849–855 (2015). doi: 10.1038/ncb3185; pmid: 26075357
24. Q. P. Vong *et al*., A role for Borg5 during trophectoderm differentiation. *Stem Cells* **28**, 1030–1038 (2010). doi: 10.1002/stem.428; pmid: 20506138
25. A. Herrig *et al*., Cooperative adsorption of ezrin on PIP2-containing membranes. *Biochemistry* **45**, 13025–13034 (2006). doi: 10.1021/bi061064a; pmid: 17059219
26. J. J. Jayasundar *et al*., Open conformation of ezrin bound to phosphatidylinositol 4,5-bisphosphate and to F-actin revealed by neutron scattering. *J. Biol. Chem.* **287**, 37119–37133 (2012). doi: 10.1074/jbc.M112.380972; pmid: 22927432
27. B. D. Slaughter *et al*., Non-uniform membrane diffusion enables steady-state cell polarization via vesicular trafficking. *Nat. Commun.* **4**, 1380 (2013). doi: 10.1038/ncomms2370; pmid: 23340420
28. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017). doi: 10.1038/nrm.2017.7; pmid: 28225081
29. K. A. Ganzinger, S. K. Vogel, J. Mücksch, P. Blumhardt, P. Schwille, Myosin-II activity generates a dynamic steady state with continuous actin turnover in a minimal actin cortex. *J. Cell Sci.* **132**, jcs219899 (2019). doi: 10.1242/jcs.219899

ACKNOWLEDGMENTS

We thank E. Munro, D. Glover, A. Andersen, and M. Shahbazi for helpful discussion; S. Shadkhoo for comments on the model; and S. Malas for the Gata3-GFP transgenic line. Some of the computations were conducted on the Caltech High Performance Cluster, supported by a Gordon and Betty Moore Foundation grant. **Funding:** This work was supported by grants from the Wellcome Trust (098287/Z/12/Z), ERC (669198), Leverhulme Trust (RPG-2018-085), Open Philanthropy/Silicon Valley, Weston Havens Foundations and NIH R01 HD100456-01A1 to M.Z.-G; Packard Foundation, Heritage Medical Research Institute, NIH U01CA244109 to M.T.; and the National Key R&D Program of China grants 2017YFA0102802 and 2019YFA0110001 to J.N. **Author contributions:** Conceptualization: M.Z. and M.Z.-G. Investigation: M.Z., J.C.-S., P.W., and C.E.H. Writing: M.Z. and M.Z.-G. Supervision: M.Z.-G., M.T., and J.N. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All raw data are available upon request from the corresponding author. The RNA-seq data have been deposited in the Gene Expression Omnibus database (accession number GSE124755). The code for computation simulation has been deposited at https://jakesorel.github.io/Apical_Domain_2020/.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eabd2703/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S4
References (*30*–*69*)
MDAR Reproducibility Checklist
Movies S1 to S15

View/request a protocol for this paper from *Bio-protocol*.

10 June 2020; accepted 14 October 2020
10.1126/science.abd2703

RESEARCH ARTICLE SUMMARY

DISEASE DYNAMICS

A transmissible cancer shifts from emergence to endemism in Tasmanian devils

Austin H. Patton, Matthew F. Lawrance, Mark J. Margres, Christopher P. Kozakiewicz, Rodrigo Hamede, Manuel Ruiz-Aravena, David G. Hamilton, Sebastien Comte, Lauren E. Ricci, Robyn L. Taylor, Tanja Stadler, Adam Leaché, Hamish McCallum, Menna E. Jones, Paul A. Hohenlohe, Andrew Storfer*

INTRODUCTION: Emerging infectious diseases pose one of the greatest threats to human health and biodiversity. Phylodynamics is an effective tool for inferring epidemiological parameters to guide intervention strategies, particularly for human viruses such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). However, phylodynamic analysis has historically been limited to the study of rapidly evolving viruses and, in rare cases, bacteria. Nonetheless, application of phylodynamics to nonviral pathogens has immense potential, such as for predicting disease spread and informing the management of wildlife diseases.

We conducted a phylodynamics analysis of devil facial tumor disease (DFTD), a transmissible cancer that has spread across nearly the entire geographic range of Tasmanian devils and threatens the species with extinction. DFTD is transmitted as an allograft through biting during common social interactions, susceptibility is nearly universal, and case fatality rates approach 100%. The goals of our study were to (i) characterize the geographic spread of DFTD, (ii) identify whether there are different circulating tumor lineages, and (iii) quantify rates of transmission among lineages.

PHOTOS: DAVID G. HAMILTON (TOP), ALEXANDRA K. FRAIK (BOTTOM).

Tasmanian devils and their transmissible cancer. Healthy (top) and DFTD-infected (bottom) Tasmanian devils.

RATIONALE: In principle, phylodynamics should be readily extended to the study of slowly evolving pathogens with large genomes through careful interrogation of genes to identify those that are measurably evolving. By testing individual genes for a clocklike signal, these genes may then be used for phylodynamic analysis. We demonstrate this proof of concept in DFTD.

RESULTS: We screened >11,000 genes across the DFTD genome, identifying 28 that exhibited a strong, clocklike signal, and performed the first phylodynamic analysis of a genome larger than a bacterium. We demonstrate here, contrary to field observations, that DFTD spread omnidirectionally throughout the epizootic, leaving little signal of geographic structuring of tumor lineages across Tasmania. Despite predictions of devil extinction, we found that the effective reproduction number (R_E), a summary of the rate at which disease spreads, has declined precipitously after the initial epidemic spread of DFTD. Specifically, R_E peaked at a high of ~3.5 shortly after the discovery of DFTD in 1996 and is now ~1 in both extant tumor lineages. This is consistent with a shift from emergence to endemism. Except for a single gene, we found little evidence for convergent molecular evolution among tumor lineages.

CONCLUSION: We have demonstrated that phylodynamics can be applied to virtually any pathogen. In doing so, we show that through careful interrogation of the pathogen genome, a measurably evolving set of genes can be identified to characterize epidemiological dynamics of nonviral pathogens with large genomes. By applying this approach to DFTD, we have shown that the disease appears to be transitioning from emergence to endemism. Consistent with recent models, our inference that R_E ~1 predicts that coexistence between devils and DFTD is a more likely outcome than devil extinction. Therefore, our findings present cautious optimism for the continued survival of the iconic Tasmanian devil but emphasize the need for evolutionarily informed conservation management to ensure their persistence. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: astorfer@wsu.edu
Cite this article as: A. H. Patton *et al.*, *Science* **370**, eabb9772 (2020). DOI: 10.1126/science.abb9772

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb9772

RESEARCH ARTICLE

DISEASE DYNAMICS

A transmissible cancer shifts from emergence to endemism in Tasmanian devils

Austin H. Patton[1,2], Matthew F. Lawrance[1], Mark J. Margres[3], Christopher P. Kozakiewicz[1], Rodrigo Hamede[4,5], Manuel Ruiz-Aravena[4,6], David G. Hamilton[4], Sebastien Comte[4,7], Lauren E. Ricci[1,8], Robyn L. Taylor[4], Tanja Stadler[9,10], Adam Leaché[11], Hamish McCallum[7,12], Menna E. Jones[4], Paul A. Hohenlohe[13], Andrew Storfer[1]*

Emerging infectious diseases pose one of the greatest threats to human health and biodiversity. Phylodynamics is often used to infer epidemiological parameters essential for guiding intervention strategies for human viruses such as severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2). Here, we applied phylodynamics to elucidate the epidemiological dynamics of Tasmanian devil facial tumor disease (DFTD), a fatal, transmissible cancer with a genome thousands of times larger than that of any virus. Despite prior predictions of devil extinction, transmission rates have declined precipitously from ~3.5 secondary infections per infected individual to ~1 at present. Thus, DFTD appears to be transitioning from emergence to endemism, lending hope for the continued survival of the endangered Tasmanian devil. More generally, our study demonstrates a new phylodynamic analytical framework that can be applied to virtually any pathogen.

Emerging infectious diseases (EIDs) threaten the health of wildlife, livestock, domestic animals, and humans (*1*). One of the primary contributors to species endangerment (*2*), EIDs have also led or contributed to notable extinctions, including dozens of amphibian species by chytridiomycosis (*3*), the Polynesian tree snail (*Partula turgida*) by a microsporidian infection (*4*), and 16 species of Hawaiian honeycreepers (Drepanidini, Fringillidae) by avian malaria (*Plasmodium relictum*) and avian pox [*Poxvirus avium* (*2*)]. EIDs also can have profound impacts on society; indeed, the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) has led to socioeconomic consequences that will surely last for years to come (*5*).

[1]School of Biological Sciences, Washington State University, Pullman, WA 99164, USA. [2]Department of Integrative Biology and Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720, USA. [3]Department of Integrative Biology, University of South Florida, Tampa, FL 33620, USA. [4]School of Biological Sciences, University of Tasmania, Hobart, Tasmania 7001, Australia. [5]CANECEV, Centre de Recherches Ecologiques et Evolutives sur le Cancer (CREEC), Montpellier 34090, France. [6]Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA. [7]Vertebrate Pest Research Unit, Invasive Species and Biosecurity, NSW Department of Primary Industries, Orange, New South Wales 2800, Australia. [8]Department of Wildland Resources, Utah State University, Logan, UT 84322, USA. [9]Department for Biosystems Science and Engineering, ETH Zürich, Basel 4058, Switzerland. [10]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [11]Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA. [12]Environmental Futures Research Institute, Griffith University, Brisbane, Queensland 4111, Australia. [13]Department of Biological Science, University of Idaho, Moscow, ID 83844, USA.
*Corresponding author. Email: astorfer@wsu.edu

Recently, phylodynamics has emerged as an invaluable tool for the characterization of the epidemiological dynamics of such rapidly evolving pathogens (*6*, *7*). By reconstructing pathogen phylogenies, phylodynamic analyses elucidate critically relevant epidemiological parameters such as the effective reproduction number (R_E) (*8*) and the effective number of infections (N_E) (*9*). R_E is a generalization of the basic reproduction number (R_0) in epidemiological models, or the expected number of secondary infections from a single infected individual entering a wholly susceptible population. Similarly, R_E quantifies transmission in populations already infected, predicts when pathogen prevalence will increase ($R_E > 1$) or decrease ($R_E < 1$), and helps to estimate the vaccination fraction necessary to achieve herd immunity.

Although phylodynamics has contributed substantially to the management of human diseases, including the responses to SARS-Cov-2 (*10*), its application to nonviral pathogens such as bacteria has been limited by their slower rate of molecular evolution and larger genome sizes (*11*). Past efforts to apply phylodynamic approaches to pathogens other than viruses have been impeded by the challenges associated with large genome size and the identification of a measurably evolving portion of the pathogen genome with which to reconstruct phylogenies. However, the extension of phylodynamics to nonviral pathogens including many wildlife EIDs would prove invaluable for management and intervention.

A marquee example is Tasmanian devils (*Sarcophilus harrisii*; Fig. 1A), which are endangered by an unusual class of emerging infectious disease: a transmissible cancer (*12*). Devil facial tumor disease (DFTD; Fig. 1B) has spread across 95% of the devil's geographic range since its discovery in 1996, causing localized population declines exceeding 90% (*13*) and a species-wide decline of 80% (*12*, *14*). DFTD replicates clonally and is transmitted as an allograft through biting during social interactions (*15*). Case fatality rates are nearly 100%, and devil susceptibility appears to be largely universal (*12*) due in part to limited genetic variation caused by historical population bottlenecks (*16*).

Although efforts have been made to describe DFTD transmission dynamics early in the epizootic [e.g., (*17*)], little is known how its epidemiology has changed since emergence. Despite initial model predictions of devil extinction resulting from frequency-dependent transmission (*17*), populations persist even in long-diseased areas, and some populations may even be recovering (*14*, *18*). Clearly, more analyses are needed to reconcile the discrepancy between model predictions and empirical observations.

Herein, we characterized the epidemiological history of DFTD using whole-genome sequencing of 51 tumor samples selected to maximize spatiotemporal variation across Tasmania between 2003 and 2018 (fig. S1). By screening >11,000 genes distributed across the tumor genome, we identified a subset that is measurably evolving, thus demonstrating an approach that enables the application of a suite of phylodynamic methods to virtually any pathogen. We also dated the approximate emergence time of DFTD, determined the number of lineages present, tracked the rate and directionality of lineage spread, and estimated N_E and R_E.

Results

Identification of clocklike genes

We screened 11,359 total genes across the DFTD genome and identified 28 genes (totaling 431,608 bp) that were sufficiently variable (>50 parsimony-informative sites) and evolved at an appropriate clocklike manner for use in tip-dating our phylogeny (table S1). These 28 genes were distributed widely across the genome without any discernible pattern with respect to gene function. Once concatenated and aligned for the 51 final samples, the aligned matrix was composed of 431,608 total columns, 2520 parsimony-informative sites, 1893 singletons, 802 doubletons, and 2711 variants found in three or more individuals (fig. S2).

DFTD phylogeography

Using these 28 genes, we estimated the time of DFTD origination to be between 1977 and 1987 (mean = 1983.93; Fig. 1C and figs. S3, S7, and S8), which is compatible with its discovery in 1996 (*12*). Field studies indicate an origin of DFTD in northeastern Tasmania, with subsequent southern and westward spread across the

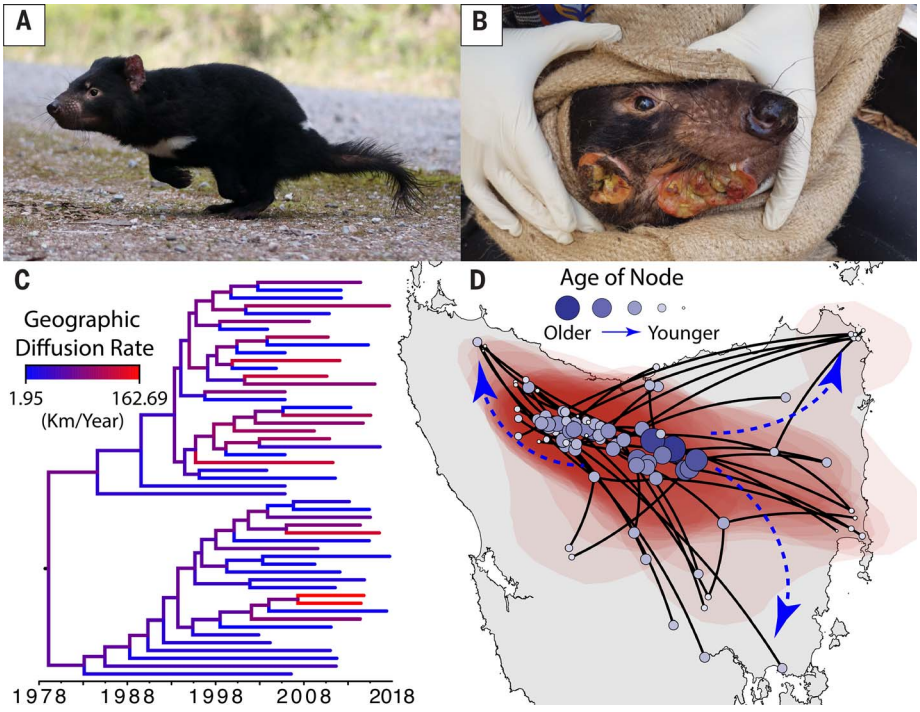


Fig. 1. Spread of DFTD lineages across Tasmania. (**A**) Healthy devil (Photo: David G. Hamilton). (**B**) Tasmanian devil infected with DFTD (Photo: Alexandra K. Fraik). (**C**) Phylogeny of 51 tumor samples; branch colors indicate the estimated geographic diffusion rate. Shown are branch lengths in units of time; the *x*-axis corresponds to year. (**D**) The same phylogeny mapped in geographic space. Blue circles correspond to both terminal and internal nodes, with size and color corresponding to age. Inferred locations of internal nodes correspond to our sampled lineages; no samples exist at the time and location of disease origination. Red polygons indicate posterior probability locations of historical infections. Dashed blue arrows indicate omnidirectional DFTD spread.

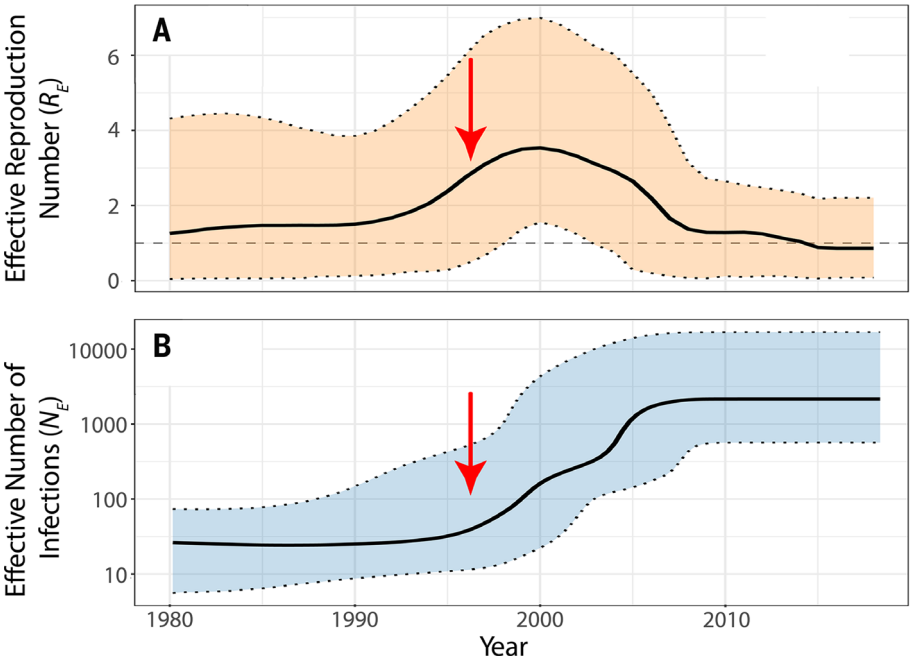


Fig. 2. Epidemiological dynamics of DFTD. Red arrows indicate the year of DFTD discovery (1996). Shading indicates the 95% credible intervals for parameter estimates; thick black lines are medians of the posterior distribution. (**A**) R_E through time under the birth-death skyline model. Dashed gray line indicates R_E = 1. Above this line, disease spreads; below it, the number of infections decreases. (**B**) N_E under the coalescent Bayesian skyline.

island (*13*). However, our phylogeographic reconstruction showed a more complex pattern of DFTD spread (Fig. 1D and movie S1). We found evidence for two contemporary, monophyletic DFTD lineages that emerged early in the epizootic and completely overlapped in their geographic distributions (Fig. 1, C and D), consistent with a recent landscape genetics study (*19*). After reaching central Tasmania, the cancer subsequently recolonized previously infected eastern and southern populations after their well-documented and extensive declines [e.g., (*13*, *14*)]. Our results suggest that the spread of DFTD continued omnidirectionally toward the present.

DFTD phylodynamics

Estimates of R_E through time are consistent with the discovery of DFTD in 1996; a low R_E at the time of disease origination would translate to low prevalence before discovery (Fig. 2A). The birth-death skyline, which assembles piecewise constant estimates of R_E, inferred a sudden, increased rate of transmission in the late 1990s (Fig. 2A). Specifically, R_E increased from ~1 around 1980 to a maximum of ~3.5. Currently R_E is <1, suggesting that DFTD will decrease in prevalence. The N_E (Fig. 2B) supports the same dynamics, showing a rapid increase in number of infections when R_E was >1, followed by a stabilization of the infected population size coinciding with an R_E of ~1.

The methods used above rely on the assumption that transmission dynamics do not vary among tumor lineages. Therefore, we relaxed these assumptions using a multistate birth-death model to determine whether tumor lineages differ in their epidemiological dynamics. We found that, indeed, transmission rate dynamics were not uniform across tumor lineages (Fig. 3A and fig. S9), with two detectable shifts in transmission rate. R_E (as estimated under the birth-death skyline) declined to just above or below 1 toward the present in each of the two, reciprocally monophyletic contemporary transmission clusters, respectively (Fig. 3B). Neither transmission cluster was geographically discrete; each was distributed island wide.

Genomic differentiation among transmission clusters

Our results show that accelerated transmission rates of DFTD began in the early to mid-1990s (Figs. 2A and 3B). With our data, we were unable to unequivocally determine what specifically led to this change. However, we screened the 51 tumor genomes and identified 791 unique variants that differentiated the three identified transmission clusters found in the top 0.1% of pF_{ST} (a measure of genomic differentiation) values. Of these variants, 687 were intergenic and the remaining 104 fell within a total of 68 unique genes and 70 unique transcripts (fig. S10).

These genes are associated with regulation of (i) STAT3 (NFATC3 and PRKG1), a key DFTD

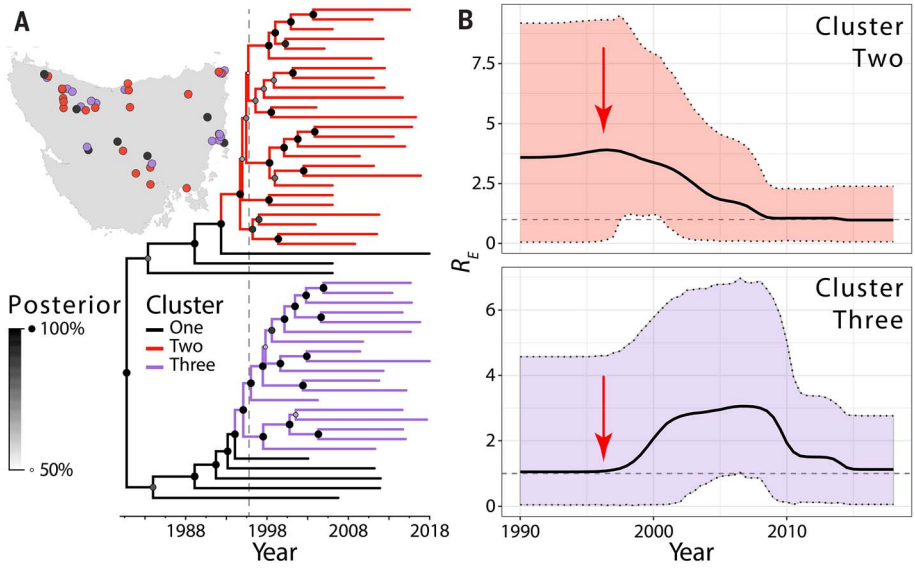


Fig. 3. Heterogeneous transmission rate dynamics in DFTD. (**A**) Phylogeny inferred under the birth-death skyline model. Branch colors show three distinct transmission rate regimes identified under the multistate birth-death model. Circles show posterior probability of interior nodes. Vertical dashed gray line indicates year of disease detection (1996). (**B**) Transmission rates through time as inferred under the birth-death skyline model for clusters 2 (red) and 3 (green). Median estimates are bold black lines; the 95% credible interval is shaded. Red arrows indicate the year of DFTD discovery (1996). Inset map shows the distribution of samples belonging to each transmission cluster, with points colored the same as in (A).

immune evasion protein (*20*); (ii) Schwann cell differentiation, the cell type of DFTD origin (*21*); and (iii) Wnt signaling (ARMC8, CDH17, DMXL1, and PKP2), a key process in cancer progression (*22*) and identified across previous DFTD genomic studies (*23*). The above genes were found to be highly differentiated between clusters 1 (black; Fig. 3A) and 3 (green), as well as between clusters 2 (red) and 3, but not between clusters 1 and 2.

Discussion

Herein, we have demonstrated that the phylodynamic analytical framework, although previously applied exclusively to viruses and a few bacteria, can be used to assess more slowly evolving pathogens with larger genomes. Through careful assessment of clocklike evolution across the genome, we were able to extract a measurably evolving signal across 28 genes totaling 431.6 kb. Note that this amount of sequence is far greater than that used in phylodynamic studies of viruses, including Ebola [~19.9 kb (*24*)], influenza A [13.6 kb (*25*)], and SARS-Cov-2 [~29.8 kb (*26*)]. Therefore, our approach can be readily applied to the study of other nonviral pathogens that previously fell outside of the scope of phylodynamic study. When applied to devil facial tumor disease, our analyses show (i) that DFTD originated well before its discovery in 1996; (ii) no geographic substructuring among two extant tumor lineages; (iii) omnidirectional spread of tumors; and (iv) a precipitous decline in transmission to replacement at present, indicating a shift to endemism.

Our study indicates that DFTD may have originated nearly a decade or more before its discovery in 1996. These findings are compatible with field observations; upon initial discovery in northeastern Tasmania, tumors were large and widespread in the population (*12*), so the disease had likely been circulating for some time. Nonetheless, it is possible that some host variants persisted in our multiple sequence alignments (MSAs) despite our best efforts to remove potential contamination (see the materials and methods). Because devil genomic variants will coalesce earlier than DFTD variants, a consequence of such contamination is that our inferred dates of disease origination are likely slightly earlier than the true origination date.

One surprising finding of our phylogeographic analysis was the apparent lack of geographical structuring of DFTD lineages. Rather, our results indicate that the disease spread omnidirectionally throughout the epizootic, repeatedly recolonizing previously infected populations that experienced substantial population declines [e.g., (*13*, *14*)]. That is, as devil populations reached low densities, they likely received infected migrant devils from neighboring areas, which is compatible with the observed disease-induced metapopulation dynamics (*27*) and local patterns of DFTD lineage replacement (*28*). Thus, our phylogeographic analysis challenges the conventional narrative of an east-to-west disease wave emanating from northeastern Tasmania (*12*) and instead suggests continuous spread in all cardinal directions (Fig. 1D).

These results differ slightly from those of Murchison *et al.* (*21*), who recovered evidence of fine-scale geographic structuring with the Forestier peninsula of southeastern Tasmania. Whereas we sampled 51 tumors across Tasmania between 2003 and 2018, 36 of the 68 tumor samples in Murchison *et al.* (*21*) came from the Forestier peninsula over a 4-year period. Additionally, the tumor phylogeny inferred by Murchison *et al.* (*21*) was based on 16 nuclear and 21 mitochondrial variants that were presumably somatic. By contrast, our phylogenies are based on 5406 total variants distributed across 28 measurably evolving genes.

Terminal branches of our inferred phylogenies are longer than internal branches, differing slightly from typical viral phylogenies. High within-host tumor diversity is one potential explanation, but evidence of this is limited and preliminary. The only study that compared within-host tumor variation found that only six tumors from 20 individuals could be distinguished. Furthermore, three of these six differed by only a single variant, thus demonstrating limited within-host variation (*21*) and the potential for superinfection by more than one tumor lineage.

An alternative explanation for long terminal branches comes from the widespread geographic distribution and complete geographic overlap of tumor lineages, which were also seen in (*19*). That is, because of the low-density sampling scheme of our study with respect to geographic and temporal distribution of samples, the probability of sampling two closely related samples at any time or place is low.

Although we infer a geographical origin in north-central Tasmania, this is likely due to the earliest DFTD samples being collected from this location. Note that this is not to be interpreted as the location at which DFTD originated because it is much more probable that the disease originated on the eastern coast, near the site of discovery (*12*). Our inference of the root location in central Tasmania is likely a consequence of the absence of samples collected at the general location and time of disease origin in northeastern Tasmania. Tissue samples of tumors were not collected until the early 2000s, so no samples exist from the time and location of disease origin. Further, the earliest available samples are from central Tasmania, not the site of disease origination. In turn, we view the reconstructed geographic root state as being the location of the most recent common ancestor of our samples rather than the location of disease origin.

We found three transmission clusters that differ in their epidemiological dynamics and have identified a number of genes that may contribute to the observed transmission rate

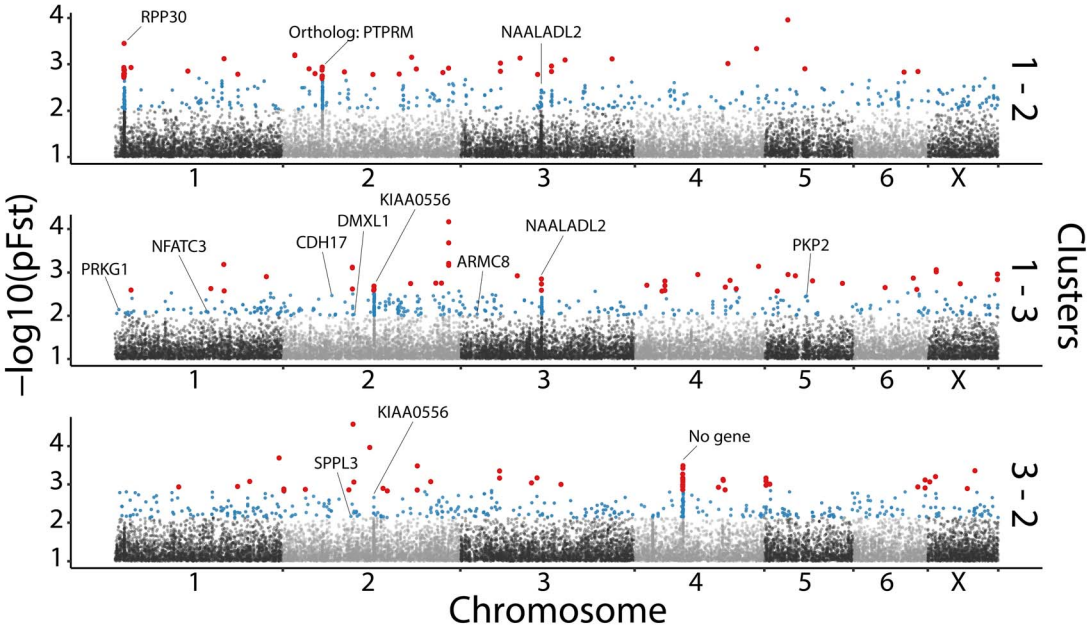


Fig. 4. Manhattan plot illustrating differentiation of SNPs in pairwise comparisons of transmission clusters using pF_{ST}. The top 0.1% most differentiated sites are shown in blue, and the top 0.01% most differentiated are shown in red. Select highly differentiated variants including those of functional importance are annotated. DFTD-associated genes discussed in the text are annotated.

variation. The absence of highly differentiated sites among clusters 1 and 2 could be caused by the minimal change in transmission rate at the time of origination of cluster 2 (Fig. 3B). These genes may thus contribute to the observed transmission rate variation and are candidates in need of further functional validation in vitro (Fig. 4 and tables S4 and S5).

Molecular evolution among DFTD transmission clusters could occur in two ways. In the first, convergent molecular evolution would be evidenced by the same genes becoming differentiated in each derived cluster (Fig. 3A). Alternatively, evolution could result in the differentiation of unique genes in each cluster. In general, each derived transmission cluster harbors a largely unique set of candidate SNPs, implicating unique mechanisms by which transmission rates vary. We only found a single gene (*NAALADL2*) consistent with the former (convergent) pattern of molecular evolution between clusters 2 and 3 (fig. S10 and table S4). This gene represents the only evidence recovered for potential convergent molecular evolution and may be important with respect to the oncogenicity of DFTD. *NAALADL2* has been demonstrated to be overexpressed in colon and prostate tumors compared with benign tumors in humans (*29*), and it promotes a tumor phenotype that exhibits greater capacity to migrate and metastasize. Future study of this gene in the context of DFTD evolution is thus warranted. Nonetheless, the observed transmission rate declines could be due to the accumulation of mutation load, which has been found in the canine transmissible cancer (*30*). Additionally, study of patterns of molecular evolution in the 28 clocklike genes identified in our study presents an exciting avenue of research.

Contrary to expectations of Tasmanian devil extinction (*13*, *17*), our results suggest that DFTD is transitioning from emergence to endemism. The large decline of R_E in the two pandemic tumor lineages suggests either coexistence or DFTD extinction, a result consistently supported by recent models and field data (*27*, *31*). This decrease in R_E may be a consequence of devil population declines, leading to fewer transmission opportunities. Current low R_E values are consistent with a reduced force of infection estimated using mark-recapture field data (*32*) and possible demographic recovery of some populations (*33*). Contemporary transmission dynamics may result in either the long-term coexistence of DFTD and devils or DFTD extinction, outcomes predicted by individual-based models (*32*). However, the potential for coexistence does not imply recovery of devils to pre-DFTD population sizes; current models predict persistent, substantially reduced devil densities (*31*).

Our results support the growing body of evidence that if DFTD continues to progress naturally, then devil extinction is unlikely. Therefore, we urge caution in the consideration of introductions of captive-bred devils to infected populations, a practice that has already been set into action (*34*). If reduced devil densities are contributing to reduced transmission rates (*27*), then an artificial increase of population densities through such introductions may be unwise. Further, introduction of disease-naïve devils may swamp adaptation to this new selective pressure (*35*). Instead, the preferred management option may be to allow natural evolution to occur and only introduce captive devils as a last resort if demographic rescue is urgently needed.

The existence of two widespread transmission clusters with unique epidemiological dynamics also suggests that mitigation strategies such as an oral bait vaccine (*36*) should account for this diversity. Unfortunately, a second, independently evolving transmissible cancer, DFT2, has been discovered in southeastern Tasmania (*37*) and may become a growing threat. Transmissible tumors may thus be commonplace throughout the evolutionary history of devils, but initial studies do not support this conclusion (*38*). Comparative phylodynamics of DFT2 and DFTD will be essential to understand their relative transmission dynamics. Our results suggest cautious optimism for the continued survival of the iconic Tasmanian devil but emphasize the need for evolutionarily informed management practices to ensure its persistence (*18*).

Materials and Methods

Sample collection

Tumor biopsies were not collected from Tasmanian devils until the early 2000s, corresponding to the time of our earliest samples. Samples from the University of Tasmania collection were obtained under University of Tasmania ethics approval A13326 and Washington State University institutional animal

care and use committee approval ASAF 6796. Tumor tissue biopsies were obtained using a 3-mm biopsy punch from wild devils and detailed field trapping protocols that have been previously described (*12*, *39*).

Early sampling efforts across Tasmania have historically been heterogeneous, with much of the effort being focused on collection at disease fronts. Therefore, we made a concerted effort to incorporate samples that maximized the available temporal range (2003 to 2018) and geographic distribution of samples available in the two largest sample repositories at the University of Tasmania and the Tasmanian Department of Primary Industries, Parks, Water and Entertainment, Tasmania, Australia. Where possible, samples were obtained such that the dataset represented the geographical distribution of DFTD in Tasmania through time, with the caveat that samples from eastern Tasmania are unavailable earliest in the collections, and collection has ceased at a number of sites in central Tasmania. Further, for devils that had tumor biopsies taken multiple times, we selected tumor biopsies that were associated with the first tumor biopsy for that devil.

Thus, in 2018 we obtained 50 tumor samples spread both geographically and temporally (2003 to 2018; fig. S1) across Tasmania. Whole-genomic libraries were prepared using NEB-next (New England Biolabs, Ipswitch, MA, USA) kits and sequenced on 26 lanes of a 10× Illumina Platform at Northwestern Genomics (Seattle, WA). An additional six whole-genome tumor samples originally sequenced for another study (*23*) were also included (BioProject PRJNA472767; BioSamples SAMN09242213, SAMN09242220, SAMN09242222, SAMN09242223, SAMN09242224, and SAMN09242226).

Sequencing read assembly

Raw reads were merged using flash2 (*40*) and adaptors trimmed with sickle (*41*). Trimmed reads were then aligned to the reference genome of *S. harrisii* (*21*) using bwa mem (*42*). Bam files produced from this step were subsequently sorted using Samtools (v1.9) sort (*43*). Samtools merge was then used to combine the sorted paired-end and single-end bam files for each sample; these combined bams were subsequently merged across lanes. Polymerase chain reaction duplicates were removed using Picard Markduplicates (*44*), and final bam files were indexed using the Samtools index.

Phylogenetic sequence preparation

To date, most phylodynamic studies have been conducted on extremely rapidly evolving viruses. Given that DFTD evolves much more slowly, we needed to ensure that the loci used to generate our phylogeny were measurably evolving. To accomplish this, we used an iterative approach that first identified a candidate set of genes evolving in a clocklike manner in both devils and tumors.

We first generated consensus sequences for the 22,391 annotated genes for the Tasmanian devil genome (Devil_ref v7.0, INSDC Assembly GCA_000189315.1). We generated consensus sequences for 48 of the 56 tumor isolates and 14 devil individuals using Samtools faidx. We then generated individual MSAs for 11,359 of the 22,391 genes using Clustal Omega (*45*). We only generated ~50% of the alignments because of computational constraints; these alignments took ~3 months to finish. However, the aligned genes were distributed across chromosomes 1 through 6, as well as the X chromosome, so we are not concerned that we are including a biased sample of the genome. We then summarized each alignment using AMAS (*46*), which quantifies the number of variable and parsimony-informative sites per gene.

Using these aligned genes, we subsequently inferred individual gene trees using IQ-TREE (*47*). Specifically, we fitted models of nucleotide substitution using the -m TEST flag, and the best-fit model was subsequently used for tree inference. We then tested, using each of the 11,359 gene trees, for a clocklike substitution process using the scripts provided by Murray *et al.* (*48*). These scripts regress phylogenetic root-to-tip distance against sampling date in a manner analogous to TempEst (*49*) to find the root branch that maximizes the fit of the regression. Significance was assessed by randomly permuting sampling dates across the tips 500 times, with the correlation coefficient as the test statistic.

We sought to identify a candidate set of genes that both exhibited a significant and realistic clocklike substitution process and harbored a sufficient number of phylogenetically informative sites to resolve the phylogeny. Therefore, we filtered these 11,359 genes such that only 364 were found to be evolving in a significant, clocklike manner ($P < 0.05$), harboring ≥50 parsimony sites, and with time to most recent common ancestor (MRCA) being inferred as <200 years before present (ybp). From this point forward, an additional eight tumor samples were included into our analyses, bringing the total number of sampled tumors to 56. These eight samples were previously excluded because of spontaneous regression [see (*23*)] but are included herein because we are specifically interested in broadscale patterns of transmission across Tasmania. The phylogenetic positions of these samples are indicated in fig. S3.

Next, we sought to remove single-nucleotide polymorphisms (SNPs) present in both devils and tumors, which would be indicative of potential host contamination in the tumor sample. To accomplish this, we called SNPs and indels for all 56 tumor samples using bcftools mpileup and bcftools call -mv. Indels were then normalized using bcftools norm -m. We then compared these SNPs and indels with those identified in the available 12 Tasmanian devil high-coverage (~30×) whole-genome sequences from Margres *et al.* (*35*) using bcftools isec to identify which SNPs and indels were unique to the tumor samples. We filtered out all variants shared between devils and tumors, leaving putatively somatic tumor variants that are not found as standing genetic variation in devil populations. This approach has been successfully applied to the study of DFTD recently (*50*) and is analogous to the approach used to study somatic variants in canine transmissible venereal tumor (*51*). Per-sample consensus sequences were then called using bcftools consensus in the manner described above.

We then inferred a final set of MSAs of all 56 tumor samples using MAFFT (*52*). Specifically, we ran MAFFT using the following settings: –local-pair –max-iterate 1000 and the accurate L-INS-I alignment method. We then repeated the workflow outlined above to infer gene trees using IQ-TREE, to test for clocklike signal, and to summarize alignments. As above, we subsequently filtered resultant MSAs such that only 28 genes were retained that were evolving in a significant ($P < 0.05$) clocklike manner, harboring >50 parsimony-informative sites, and exhibiting an MRCA < 200 ybp.

Phylodynamic inference

To expedite tree inference in our Bayesian phylodynamic analyses, we inferred a starting tree using concatenated sequences after automated model selection with IQ-TREE. We leveraged three complementary Bayesian phylodynamic methods in Beast v2.5.1 (*53*): Bayesian continuous phylogeography [visualized in SpreadD3 (*54*, *55*)], coalescent Bayesian skyline (*9*), and birth-death skyline (*8*) to: (i) characterize the phylogeographic history of DFTD as it spread across Tasmania, (ii) infer the demographic history of the disease, and (iii) quantify the rates of transmission throughout the epizootic. To test for among-lineage variation in transmission rates, we applied the multistate birth-death model (*56*), as implemented in R v3.6.1 (*57*), to the birth-death skyline maximum clade consensus tree. To quantify transmission rate variation in each cluster, we used the birth-death skyline model for the samples within each cluster. To prevent overfitting, we reduced the number of time series estimated R_E per cluster to five. We additionally used TreeTime (*58*) to conduct a root-to-tip regression of molecular divergence in the ML tree against time of sampling to obtain an estimate of the time of disease origination. Details specific to each analysis are described in the extended materials and methods in the supplementary materials.

Phylogenetic inference: Starting tree

As before, we implemented model selection in IQ-TREE on the concatenated sequence before topology inference. To quantify topological uncertainty, we conducted 100 rounds of bootstrapping. The ML tree was then read into TempEst to identify problematic sequences. We identified and removed five tumors that had outlying residuals or appeared to have mislabeled tip-dates. These five samples were excluded from all downstream analyses. In turn, we generated our final set of 51 tumor samples presented in the main text. After removal of these tumors, the inferred topology was subsequently visualized using FigTree (*59*).

Bayesian phylodynamic inference

For each of the Bayesian phylodynamic analyses described above, the best-fit model of nucleotide evolution as determined by IQ-TREE for the concatenated sequence was set in BEAST. Nondefault priors are shown in table S2. Four independent Markov chain Monte Carlo (MCMC) algorithms were run for 50 million generations, sampling every 1000 generations. Individual MCMC traces were compared to assess convergence using Tracer. Convergent chains were combined using LogCombiner, discarding the first 20% of samples for each chain as burn-in. Last, a single chain of 50 million generations was run, sampling from the prior only; these MCMC traces were used to assess deviation of the posterior from the prior for each estimated parameter. Maximum clade consensus trees from combined posterior distributions for each model were obtained using TreeAnnotater and visualized in FigTree.

Genomic differentiation among transmission clusters

To characterize genetic differences among transmission clusters identified under the multistate birth-death model, we identified SNPs in the genome exhibiting particularly strong differentiation among clusters, which are interpreted as candidates for the explanation of transmission rate variation using the pF_{ST} association test implemented in the GPAT++ software package (*60*, *61*). Specifically, we used this approach to identify candidate genomic variants that could explain the observed transmission rate variation. The pF_{ST} association test conducts a likelihood ratio test of allele frequency differences among populations, correcting for sequencing error using genotype likelihood scores in the calculation of model parameters. This approach has previously been successfully applied to the study of DFTD (*50*). To ensure that only confidently identified SNPs were retained in this analysis, we filtered the dataset using VCFtools v0.1.16 (*62*) using the following flags: –mac 4, –max-alleles 2, –min-alleles 2, –minDP 10, –max-missing 0.8, and –remove-indels. Specifically, we made all pairwise comparisons among identified transmission clusters and/or regimes. We subsequently characterized and extracted annotations from Ensembl for differentiated SNPs using the variant effect predictor (*63*).

REFERENCES AND NOTES

1. P. Daszak, A. A. Cunningham, A. D. Hyatt, Emerging infectious diseases of wildlife—Threats to biodiversity and human health. *Science* **287**, 443–449 (2000). doi: 10.1126/science.287.5452.443; pmid: 10642539
2. K. F. Smith, D. F. Sax, K. D. Lafferty, Evidence for the role of infectious disease in species extinction and endangerment. *Conserv. Biol.* **20**, 1349–1357 (2006). doi: 10.1111/j.1523-1739.2006.00524.x; pmid: 17002752
3. V. T. Vredenburg, R. A. Knapp, T. S. Tunstall, C. J. Briggs, Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9689–9694 (2010). doi: 10.1073/pnas.0914111107; pmid: 20457913
4. A. A. Cunningham, P. Daszak, Extinction of a species of land snail due to infection with a microsporidian parasite. *Conserv. Biol.* **12**, 1139–1141 (1998). doi: 10.1046/j.1523-1739.1998.97485.x
5. I. Chakraborty, P. Maity, COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Sci. Total Environ.* **728**, 138882 (2020). doi: 10.1016/j.scitotenv.2020.138882; pmid: 32335410
6. B. T. Grenfell *et al*., Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332 (2004). doi: 10.1126/science.1090727; pmid: 14726583
7. E. M. Volz, K. Koelle, T. Bedford, Viral phylodynamics. *PLOS Comput. Biol.* **9**, e1002947 (2013). doi: 10.1371/journal.pcbi.1002947; pmid: 23555203
8. T. Stadler, D. Kühnert, S. Bonhoeffer, A. J. Drummond, Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl. Acad. Sci. U.S.A.* **110**, 228–233 (2013). doi: 10.1073/pnas.1207965110; pmid: 23248286
9. A. J. Drummond, A. Rambaut, B. Shapiro, O. G. Pybus, Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005). doi: 10.1093/molbev/msi103; pmid: 15703244
10. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020). doi: 10.1126/science.abb5793; pmid: 32291278
11. R. Biek, O. G. Pybus, J. O. Lloyd-Smith, X. Didelot, Measurably evolving pathogens in the genomic era. *Trends Ecol. Evol.* **30**, 306–313 (2015). doi: 10.1016/j.tree.2015.03.009; pmid: 25887947
12. C. E. Hawkins *et al*., others, Emerging disease and population decline of an island endemic, the Tasmanian devil *Sarcophilus harrisii*. *Biol. Conserv.* **131**, 307–324 (2006). doi: 10.1016/j.biocon.2006.04.010
13. H. McCallum *et al*., Distribution and impacts of Tasmanian devil facial tumor disease. *EcoHealth* **4**, 318–325 (2007). doi: 10.1007/s10393-007-0118-0
14. B. T. Lazenby *et al*., Density trends and demographic signals uncover the long-term impact of transmissible cancer in Tasmanian devils. *J. Appl. Ecol.* **55**, 1368–1379 (2018). doi: 10.1111/1365-2664.13088; pmid: 30089931
15. A. M. Pearse, K. Swift, Allograft theory: Transmission of devil facial-tumour disease. *Nature* **439**, 549 (2006). doi: 10.1038/439549a; pmid: 16452970
16. A. H. Patton *et al*., Contemporary demographic reconstruction methods are robust to genome assembly quality: A case study in Tasmanian devils. *Mol. Biol. Evol.* **36**, 2906–2921 (2019). doi: 10.1093/molbev/msz191; pmid: 31424552
17. H. McCallum *et al*., Transmission dynamics of Tasmanian devil facial tumor disease may lead to disease-induced extinction. *Ecology* **90**, 3379–3392 (2009). doi: 10.1890/08-1763.1; pmid: 20120807
18. P. A. Hohenlohe *et al*., Conserving adaptive potential: Lessons from Tasmanian devils and their transmissible cancer. *Conserv. Genet.* **20**, 81–87 (2019). doi: 10.1007/s10592-019-01157-5; pmid: 31551664
19. C. P. Kozakiewicz *et al*., Comparative landscape genetics reveals differential effects of environment on host and pathogen genetic structure in Tasmanian devils (*Sarcophilus harrisii*) and their transmissible tumour. *Mol. Ecol.* **29**, 3217–3233 (2020). doi: 10.1111/mec.15558; pmid: 32682353
20. L. Kosack *et al*., The ERBB-STAT3 Axis Drives Tasmanian Devil Facial Tumor Disease. *Cancer Cell* **35**, 125–139.e9 (2019). doi: 10.1016/j.ccell.2018.11.018; pmid: 30645971
21. E. P. Murchison *et al*., Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. *Cell* **148**, 780–791 (2012). doi: 10.1016/j.cell.2011.11.065; pmid: 22341448
22. G. Davidson, C. Niehrs, Emerging links between CDK cell cycle regulators and Wnt signaling. *Trends Cell Biol.* **20**, 453–460 (2010). doi: 10.1016/j.tcb.2010.05.002; pmid: 20627573
23. M. J. Margres *et al*., The genomic basis of tumor regression in Tasmanian devils (*Sarcophilus harrisii*). *Genome Biol. Evol.* **10**, 3012–3025 (2018). pmid: 30321343
24. T. Wawina-Bokalanga *et al*., Complete genome sequence of a new Ebola virus strain isolated during the 2017 Likati outbreak in the Democratic Republic of the Congo. *Microbiol. Resour. Announc.* **8**, e00360-19 (2019). doi: 10.1128/MRA.00360-19; pmid: 31097506
25. G. Winter, S. Fields, Nucleotide sequence of human influenza A/PR/8/34 segment 2. *Nucleic Acids Res.* **10**, 2135–2143 (1982). doi: 10.1093/nar/10.6.2135; pmid: 6281731
26. A. Wu *et al*., Genome composition and divergence of the novel Coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **27**, 325–328 (2020). doi: 10.1016/j.chom.2020.02.001; pmid: 32035028
27. V. Siska, A. Eriksson, B. Mehlig, A. Manica, A metapopulation model of the spread of the Devil Facial Tumour Disease predicts the long term collapse of its host but not its extinction. bioRxiv 347062 [Preprint]. 14 June 2018. https://doi.org/10.1101/347062
28. R. K. Hamede *et al*., Transmissible cancer in Tasmanian devils: Localized lineage replacement and host population response. *Proc. Biol. Sci.* **282**, 20151468 (2015). doi: 10.1098/rspb.2015.1468; pmid: 26336167
29. H. C. Whitaker *et al*., N-acetyl-L-aspartyl-L-glutamate peptidase-like 2 is overexpressed in cancer and promotes a pro-migratory and pro-metastatic phenotype. *Oncogene* **33**, 5274–5287 (2014). doi: 10.1038/onc.2013.464; pmid: 24240687
30. P. W. Ewald, H. A. Swain Ewald, Infection and cancer in multicellular organisms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20140224 (2015). doi: 10.1098/rstb.2014.0224; pmid: 26056368
31. K. Wells *et al*., Individual and temporal variation in pathogen load predicts long-term impacts of an emerging infectious disease. *Ecology* **100**, e02613 (2019). doi: 10.1002/ecy.2613; pmid: 30636287
32. K. Wells *et al*., Infection of the fittest: Devil facial tumour disease has greatest effect on individuals with highest reproductive output. *Ecol. Lett.* **20**, 770–778 (2017). doi: 10.1111/ele.12776; pmid: 28489304
33. S. Comte, S. Carver, R. Hamede, M. Jones, Changes in spatial organization following an acute epizootic: Tasmanian devils and their transmissible cancer. *Glob. Ecol. Conserv.* **22**, e00993 (2020). doi: 10.1016/j.gecco.2020.e00993
34. S. Fox, P. J. Seddon, "Wild devil recovery: Managing devils in the presence of disease," in *Saving the Tasmanian Devil: Recovery Through Science-based Management*, C. Hogg, S. Fox, D. Pemberton, K. Belov, Eds. (CSIRO, 2019), pp. 157–164.
35. M. J. Margres *et al*., Large-effect loci affect survival in Tasmanian devils (*Sarcophilus harrisii*) infected with a transmissible cancer. *Mol. Ecol.* **27**, 4189–4199 (2018). doi: 10.1111/mec.14853; pmid: 30171778
36. A. S. Flies *et al*., An oral bait vaccination approach for the Tasmanian devil facial tumor diseases. *Expert Rev. Vaccines* **19**, 1–10 (2020). doi: 10.1080/14760584.2020.1711058; pmid: 31971036
37. R. J. Pye *et al*., A second transmissible cancer in Tasmanian devils. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 374–379 (2016). doi: 10.1073/pnas.1519691113; pmid: 26711993
38. A. R. Stahlke, B. Epstein, S. Barbosa, A. Patton, S. A. Hendricks, A. Veillet, A. K. Fraik, B. Schönfeld, H. I. McCallum, R. Hamede, M. E. Jones, A. Storfer, P. A. Hohenlohe, Historical and contemporary signatures of selection in response to transmissible cancer in the Tasmanian devil (*Sarcophilus harrisii*). bioRxiv 241885 [Preprint]. 7 August 2020. https://doi.org/10.1101/2020.08.07.241885
39. R. K. Hamede, H. McCallum, M. Jones, Biting injuries and transmission of Tasmanian devil facial tumour disease. *J. Anim. Ecol.* **82**, 182–190 (2013). doi: 10.1111/j.1365-2656.2012.02025.x; pmid: 22943286
40. T. Magoč, S. L. Salzberg, FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**,

2957–2963 (2011). doi: 10.1093/bioinformatics/btr507; pmid: 21903629
41. N. A. Joshi, J. N. Fass, Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (2011); https://github.com/najoshi/sickle.
42. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN] (26 May 2013).
43. H. Li *et al*., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi: 10.1093/bioinformatics/btp352; pmid: 19505943
44. Broad Institute, "Picard tools" (Broad Institute, 2018); http://broadinstitute.github.io/picard/.
45. F. Sievers, D. G. Higgins, Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105–116 (2014). doi: 10.1007/978-1-62703-646-7_6; pmid: 24170397
46. M. L. Borowiec, AMAS: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ* **4**, e1660 (2016). doi: 10.7717/peerj.1660; pmid: 26835189
47. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015). doi: 10.1093/molbev/msu300; pmid: 25371430
48. G. G. R. Murray *et al*., The effect of genetic structure on molecular dating and tests for temporal signal. *Methods Ecol. Evol.* **7**, 80–89 (2016). doi: 10.1111/2041-210X.12466; pmid: 27110344
49. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016). doi: 10.1093/ve/vew007; pmid: 27774300
50. M. J. Margres *et al*., Spontaneous tumor regression in Tasmanian devils associated with RASL11A activation. *Genetics* **215**, 1143–1152 (2020). doi: 10.1534/genetics.120.303428; pmid: 32554701
51. B. Decker *et al*., Comparison against 186 canid whole-genome sequences reveals survival strategies of an ancient clonally transmissible canine tumor. *Genome Res.* **25**, 1646–1655 (2015). doi: 10.1101/gr.190314.115; pmid: 26232412
52. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi: 10.1093/molbev/mst010; pmid: 23329690
53. R. Bouckaert *et al*., BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* **10**, e1003537 (2014). doi: 10.1371/journal.pcbi.1003537; pmid: 24722319
54. P. Lemey, A. Rambaut, J. J. Welch, M. A. Suchard, Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010). doi: 10.1093/molbev/msq067; pmid: 20203288
55. F. Bielejec *et al*., SpreaD3: Interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol. Biol. Evol.* **33**, 2167–2169 (2016). doi: 10.1093/molbev/msw082; pmid: 27189542
56. J. Barido-Sottani, T. G. Vaughan, T. Stadler, Detection of HIV transmission clusters from phylogenetic trees using a multi-state birth-death model. *J. R. Soc. Interface* **15**, 20180512 (2018). doi: 10.1098/rsif.2018.0512; pmid: 30185544
57. R Core Team, "R: A language and environment for statistical computing" (R Foundation for Statistical Computing, 2019); www.R-project.org.
58. P. Sagulenko, V. Puller, R. A. Neher, TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018). doi: 10.1093/ve/vex042; pmid: 29340210
59. A. Rambaut, FigTree (2018), (available at http://tree.bio.ed.ac.uk/software/figtree/).
60. E. Garrison, "vcflib: C++ library and cmdline tools for parsing and manipulating VCF files" (2012); https://github.com/vcflib/vcflib.
61. Z. Kronenberg, M. D. Shapiro, "Association testing with GPAT++" (2014); https://github.com/zeeev/vcflib/wiki/Association-testing-with-GPAT.
62. P. Danecek *et al*., The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi: 10.1093/bioinformatics/btr330; pmid: 21653522
63. W. McLaren *et al*., The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016). doi: 10.1186/s13059-016-0974-4; pmid: 27268795
64. M. A. Suchard *et al*., Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018). doi: 10.1093/ve/vey016; pmid: 29942656
65. G. Baele, P. Lemey, Bayesian evolutionary model testing in the phylogenomics era: Matching model complexity with computational efficiency. *Bioinformatics* **29**, 1970–1979 (2013). doi: 10.1093/bioinformatics/btt340; pmid: 23766415
66. V. N. Minin, E. W. Bloomquist, M. A. Suchard, Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **25**, 1459–1471 (2008). doi: 10.1093/molbev/msn090; pmid: 18408232
67. O. G. Pybus, A. Rambaut, P. H. Harvey, An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* **155**, 1429–1437 (2000). pmid: 10880500
68. K. Strimmer, O. G. Pybus, Exploring the demographic history of DNA sequences using the generalized skyline plot. *Mol. Biol. Evol.* **18**, 2298–2305 (2001). doi: 10.1093/oxfordjournals.molbev.a003776; pmid: 11719579

ACKNOWLEDGMENTS

We thank L. Harmon, R. Gomulkiewicz, J. Kelley, O. Cornejo, A. Stalhke, and S. Barbosa for insightful discussions and recommendations. **Funding:** This work was funded by NIH grant no. R01-GM126563 as part of the NIH-NSF-USDA Ecology and Evolution of Infectious Diseases Program (to A.S., M.E.J., H.M., and P.A.H.) and by ARC DECRA grant no. DE170101116 (to R.H.). The research presented in this study was performed under University of Tasmania ethics approval A13326 and WSU IACUC approval ASAF 6796. **Author contributions:** A.H.P. conducted all analyses after the assembly of raw reads. M.L. extracted genomic DNA, assembled reads, and assisted in initial multiple sequence alignments. R.H., M.R.-A., D.G.H., S.C., and M.J. collected tumor biopsies. A.H.P. and A.S. wrote the manuscript with contributions from all other authors. H.M., M.J., P.A.H., and A.S. initially conceived of the project and all other authors contributed to its development. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Newly generated assembled sequence data are available from the NCBI Sequence Read Archive (SRA) under BioProject PRJNA613730, BioSample Accessions SAMN14418888-906, SAMN14418908-910, SAMN14418912, SAMN14418914-931, and SAMN15869204-7. All other supporting data and code are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eabb9772/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 to S5
References (*64*–*68*)
Movie S1
MDAR Reproducibility Checklist

30 March 2020; accepted 21 October 2020
10.1126/science.abb9772

RESEARCH ARTICLE SUMMARY

SYSTEMS BIOLOGY

Genetic interaction mapping informs integrative structure determination of protein complexes

Hannes Braberg*, Ignacia Echeverria*, Stefan Bohn*, Peter Cimermancic*, Anthony Shiver, Richard Alexander, Jiewei Xu, Michael Shales, Raghuvar Dronamraju, Shuangying Jiang, Gajendradhar Dwivedi, Derek Bogdanoff, Kaitlin K. Chaung, Ruth Hüttenhain, Shuyi Wang, David Mavor, Riccardo Pellarin, Dina Schneidman, Joel S. Bader, James S. Fraser, John Morris, James E. Haber, Brian D. Strahl, Carol A. Gross, Junbiao Dai, Jef D. Boeke†, Andrej Sali†, Nevan J. Krogan†

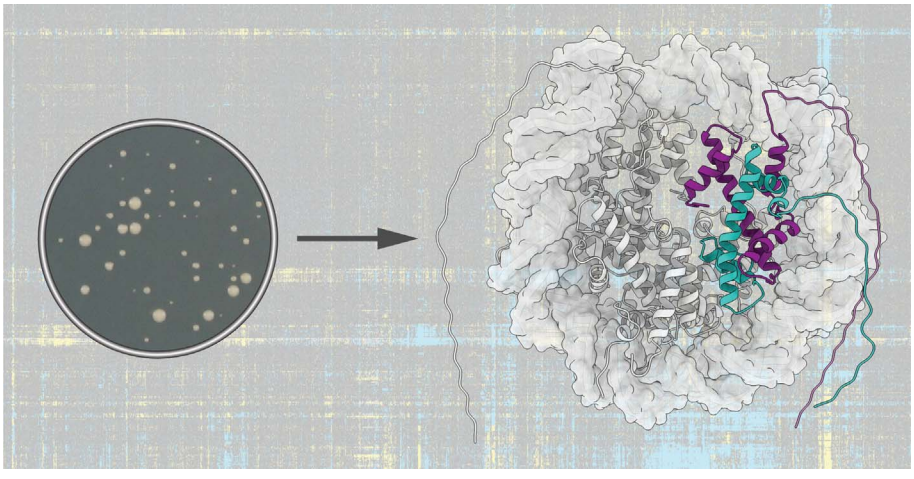
INTRODUCTION: Determining the structures of protein complexes is crucial for understanding cellular functions. Here, we describe an integrative structure determination approach that relies on in vivo quantitative measurements of genetic interactions. Genetic interactions report on how the effect of one mutation is altered by the presence of a second mutation and have proven effective for identifying groups of genes or residues that function in the same pathway. The point mutant epistatic miniarray profile (pE-MAP) platform allows for rapid measurement of genetic interactions between sets of point mutations and deletion libraries. A pE-MAP is made up of phenotypic profiles, each of which contains all genetic interactions between a single point mutant and the entire deletion library.

RATIONALE: We observe a statistical association between the distance spanned by two mutated residues in a protein complex and the similarity of their phenotypic profiles (phenotypic similarity) in a pE-MAP. This observation is in agreement with the expectation that mutations within the same functional region (e.g., active, allosteric, and binding sites) are likely to share more similar phenotypes than those that are distant in space. Here, we explore how to use these associations for determining in vivo structures of protein complexes using integrative modeling.

RESULTS: We generated a large pE-MAP by crossing 350 mutations in yeast histones H3 and H4 against 1370 gene deletions (or hypomorphic alleles of essential genes). The phenotypic similarities were then used to generate spatial restraints for integrative modeling of the H3-H4 complex structure. The resulting ensemble of H3-H4 configurations is accurate and precise, as evidenced by its close similarity to the crystal structure. This finding indicates the utility of the pE-MAP data for integrative structure determination. Furthermore, we show that the pE-MAP provides a wealth of biological insight into the function of the nucleosome and can connect individual histone residues and regions to associated complexes and processes. For example, we observe very high phenotypic similarities between modifiable histone residues and their cognate enzymes, such as H3K4 and COMPASS, or H3K36 and members of the Set2 pathway. Furthermore, the pE-MAP reveals several residues involved in DNA repair and others that function in cryptic transcription.

We demonstrate that the approach is transferable to other complexes and other types of phenotypic profiles by determining the structures of two complexes of known structure: (i) subunits Rpb1 and Rpb2 of yeast RNA polymerase II, using a pE-MAP of 53 point mutants crossed against 1200 deletions and hypomorphic alleles; and (ii) subunits RpoB and RpoC of bacterial RNA polymerase, using a chemical genetics map of 44 point mutants subjected to 83 environmental stresses. The accuracy and precision of the models are comparable to those based on chemical cross-linking, which is commonly used to determine protein complex structures. Moreover, the accuracy and precision improve when using pE-MAP and cross-linking data together, indicating complementarity between these methods and demonstrating a premise of integrative structure determination.

CONCLUSION: We show that the architectures of protein complexes can be determined using quantitative genetic interaction maps. Because pE-MAPs contain purely phenotypic measurements, collected in living cells, they generate spatial restraints that are orthogonal to other commonly used data for integrative modeling. The pE-MAP data may also enable the characterization of complexes that are difficult to isolate and purify, or those that are only transiently stable. Recent advances in CRISPR-Cas9 genome editing provide a means for extending our platform to human cells, allowing for identification and characterization of functionally relevant structural changes that take place in disease alleles. Expanding this analysis to look at structural changes in host-pathogen complexes and how they affect infection will also be feasible by introducing specific mutations into the pathogenic genome and studying the phenotypic consequences using genetic interaction profiling of relevant host genes. ■



In vivo structure determination using genetic interactions. pE-MAPs are generated by measuring the growth of yeast colonies (left) and visualized as a heatmap (background). We present an application of pE-MAPs to determine protein complex structures, using integrative modeling, and apply it to histones H3 and H4 (right) and other complexes. H3 (purple) and H4 (teal) are highlighted in the context of the nucleosome [gray, modified Protein Data Bank (PDB) 1ID3].

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: nevan.krogan@ucsf.edu (N.J.K.); sali@salilab.org (A.Sa.); jef.boeke@nyulangone.org (J.D.B.)
Cite this article as H. Braberg *et al.*, *Science* **370**, eaaz4910 (2020). DOI: 10.1126/science.aaz4910

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aaz4910

RESEARCH ARTICLE

SYSTEMS BIOLOGY

Genetic interaction mapping informs integrative structure determination of protein complexes

Hannes Braberg[1,2]*, Ignacia Echeverria[1,2,3]*, Stefan Bohn[1,2,4]*†‡, Peter Cimermancic[3]*§, Anthony Shiver[5]¶, Richard Alexander[1], Jiewei Xu[1,2,4], Michael Shales[1,2], Raghuvar Dronamraju[6], Shuangying Jiang[7], Gajendradhar Dwivedi[8]#, Derek Bogdanoff[9], Kaitlin K. Chaung[9], Ruth Hüttenhain[1,2,4], Shuyi Wang[1], David Mavor[2,3], Riccardo Pellarin[3]**, Dina Schneidman[3], Joel S. Bader[10], James S. Fraser[2,3], John Morris[11], James E. Haber[8], Brian D. Strahl[6], Carol A. Gross[12], Junbiao Dai[7], Jef D. Boeke[13,14,15,16]††, Andrej Sali[2,3,11]††, Nevan J. Krogan[1,2,4,17]††

Determining structures of protein complexes is crucial for understanding cellular functions. Here, we describe an integrative structure determination approach that relies on in vivo measurements of genetic interactions. We construct phenotypic profiles for point mutations crossed against gene deletions or exposed to environmental perturbations, followed by converting similarities between two profiles into an upper bound on the distance between the mutated residues. We determine the structure of the yeast histone H3-H4 complex based on ~500,000 genetic interactions of 350 mutants. We then apply the method to subunits Rpb1-Rpb2 of yeast RNA polymerase II and subunits RpoB-RpoC of bacterial RNA polymerase. The accuracy is comparable to that based on chemical cross-links; using restraints from both genetic interactions and cross-links further improves model accuracy and precision. The approach provides an efficient means to augment integrative structure determination with in vivo observations.

A mechanistic understanding of cellular functions requires structural characterization of the corresponding macromolecular assemblies (*1*). Traditional structural biology methods—such as x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and electron microscopy (EM)—rely on purified samples and are generally not applicable to heterogeneous samples, such as those of large, membrane-bound, or transient assemblies (*2*). Moreover, these methods do not determine the structures in their native environments, therefore increasing the risk of producing structures in nonfunctional states or missing relevant functional states.

Integrative structure determination has emerged as a powerful approach for determining the structures of biological assemblies (*3*). The motivation is that any system can be described most accurately, precisely, completely, and efficiently by using all available information about it, including varied experimental data (e.g., chemical cross-links, protein interaction data, small-angle x-ray scattering profiles) and prior models (e.g., atomic structures of the subunits). Integrative methods can often tackle protein assemblies that are difficult to characterize using traditional structural biology methods alone (*1*, *4*–*10*). Spatial data generated by in vivo methods are especially useful for integrative structure determination (*11*). Therefore, high-throughput in vivo methods are needed to supplement low-throughput in vivo methods, such as single-molecule Förster resonance energy transfer spectroscopy (*12*).

Here, we describe how integrative structure modeling can benefit from spatial restraints derived from in vivo quantitative measurements of genetic interactions. A genetic interaction between two mutations occurs when the effect of one mutation is altered by the presence of the second mutation (Fig. 1A) (*13*). Positive genetic interactions (epistasis or suppression) arise when the double mutant is healthier than expected, whereas negative interactions (synthetic sickness) arise in relationships where the double mutant is sicker than expected. Single genetic interactions can often be difficult to interpret in isolation. A phenotypic profile, defined as a set of genetic interactions between a given mutation (e.g., a point mutation) and a library of secondary mutations (e.g., gene deletions), can be more informative (Fig. 1B) (*14*). A point mutant epistatic miniarray profile (pE-MAP) is composed of such phenotypic profiles for all mutations in the analysis (Fig. 1C) (*15*). We have previously found a statistical association between the distance between two mutated residues in the wild-type (WT) structure and the similarity between their phenotypic profiles (i.e., phenotypic similarity) (*15*, *16*) (Fig. 1D). This observation is in agreement with the expectation that mutations within the same functional region (e.g., active, allosteric, and binding sites) are likely to share more similar phenotypes than those that are distant in space (*17*–*19*). Here, we explore how to use these associations for determining in vivo structures of macromolecular assemblies using integrative modeling (Fig. 1E). To enable this analysis, we generated a large pE-MAP, by designing a comprehensive set of 350 mutations in histones H3 and H4 and crossing these against 1370 gene deletions (or hypomorphic alleles for essential genes). We describe this pE-MAP and illustrate integrative structure determination by its application to three complexes of known structure: (i) the yeast histones H3 and H4; (ii) subunits Rpb1 and Rpb2 of yeast RNA polymerase II (RNAPII), using a pE-MAP dataset of 53 point mutants crossed against a library of 1200 deletions and hypomorphic alleles (*15*); and (iii) subunits RpoB and RpoC of bacterial RNA polymerase (RNAP), using a chemical genetics miniarray profile (CG-MAP), where 44 point mutants were subjected to 83 different environmental stresses (e.g., treatments with chemicals and temperature shocks) (*20*).

A comprehensive pE-MAP of histones H3 and H4

Histones are central to chromatin structure and dynamics because they make up the core of the nucleosome, the fundamental repeating unit of chromatin. The state of the nucleosome

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA. [2]Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA 94158, USA. [3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA. [4]Gladstone Institutes, San Francisco, CA 94158, USA. [5]Graduate Group in Biophysics, University of California San Francisco, San Francisco, CA 94158, USA. [6]Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. [7]CAS Key Laboratory of Quantitative Engineering Biology, Guangdong Provincial Key Laboratory of Synthetic Genomics and Shenzhen Key Laboratory of Synthetic Genomics, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. [8]Department of Biology and Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454, USA. [9]Center for Advanced Technology, Department of Biophysics and Biochemistry, University of California, San Francisco, San Francisco, CA 94158, USA. [10]Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [11]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA. [12]Department of Microbiology and Immunology and Department of Cell and Tissue Biology, University of California, San Francisco, San Francisco, CA 94158, USA. [13]NYU Langone Health, New York, NY 10016, USA. [14]High Throughput Biology Center and Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [15]Institute for Systems Genetics and Department of Biochemistry and Molecular Pharmacology, NYU Langone Health, New York, NY 10016, USA. [16]Department of Biomedical Engineering, NYU Tandon School of Engineering, Brooklyn, NY 11201, USA. [17]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
*These authors contributed equally to this work. †Present address: Department of Molecular Machines and Signaling, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. ‡Present address: Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. §Present address: Verily Life Sciences, 269 E. Grand Ave., South San Francisco, CA 94080, USA. ¶Present address: Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. #Deceased. **Present address: Institut Pasteur, Structural Bioinformatics Unit, Department of Structural Biology and Chemistry, CNRS UMR 3528, C3BI USR 3756 CNRS & IP, Paris, France.
††Corresponding author. Email: nevan.krogan@ucsf.edu (N.J.K.); sali@salilab.org (A.Sa.); jef.boeke@nyulangone.org (J.D.B.)

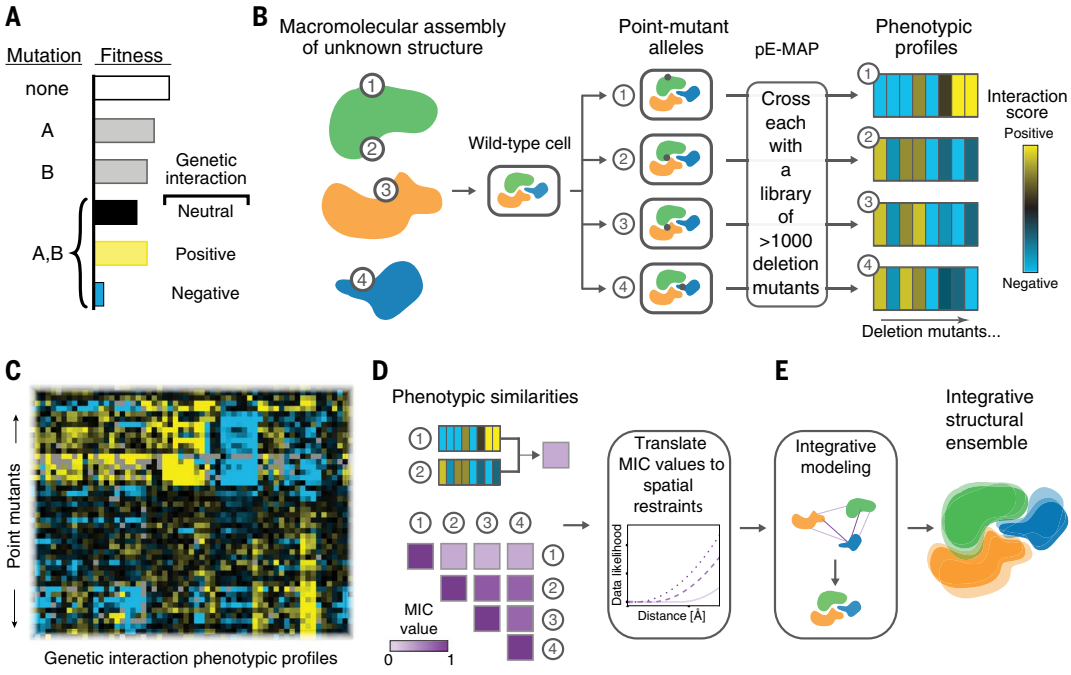


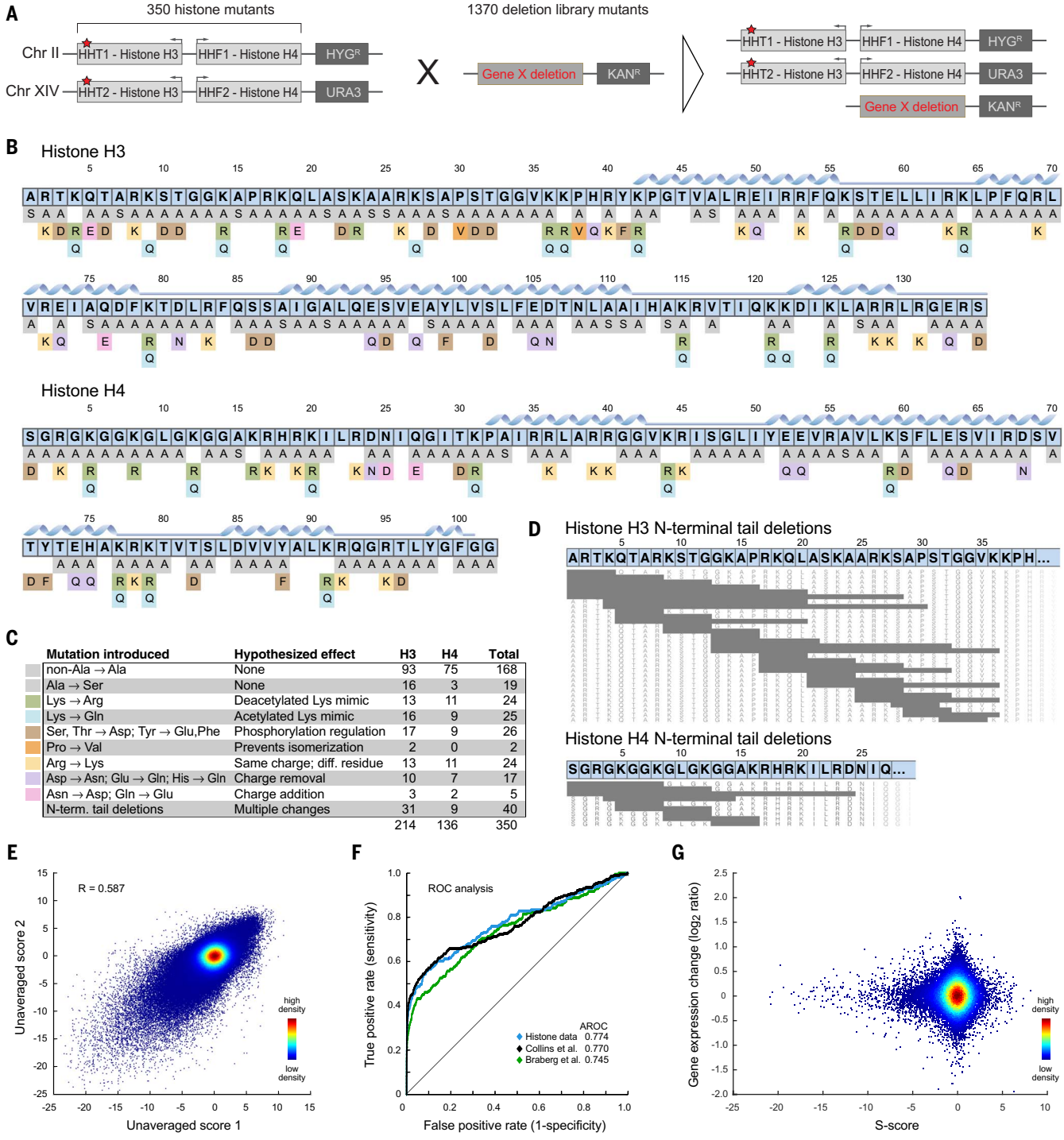
Fig. 1. Building spatial restraints from pairwise genetic perturbations. (**A**) Genetic interactions arise when the combined fitness defect of a double mutant deviates from the expected multiplicative growth defect of the two single mutants. (**B**) The generation of a pE-MAP relies on a collection of point mutations, which is constructed by systematic mutagenesis of genes that encode the subunits of a macromolecular assembly (mutations labeled 1 to 4). The point-mutant strains are then crossed against a library of gene deletions, followed by fitness measurement and subsequent calculation of genetic interaction scores to obtain the phenotypic profiles. (**C**) An example subset of a pE-MAP of point mutants crossed against a library of gene deletions. (**D**) Each pairwise combination of phenotypic profiles is transformed into a single MIC value that reflects the similarity between the two profiles. The MIC values are translated into spatial restrains for integrative modeling. (**E**) The MIC values and other input information are used for integrative structure modeling. An ensemble of structures that satisfy the input information is obtained.

is controlled by histone posttranslational modifications (PTMs) (*21*)—including acetylation, methylation, phosphorylation, and ubiquitination—that help maintain and regulate chromatin structure and transcription. Our library of point mutations in the core histones H3 and H4 was designed to comprise a comprehensive alanine scan, as well as context-specific mutations of modifiable residues (e.g., lysine and arginine), such as charge removal or reversal and substitutions mimicking PTMs (*22*, *23*). Partial deletions of the N-terminal tails of H3 and H4 were also included, because these regions play important and sometimes redundant roles in chromatin biology (*24*, *25*). In budding yeast, histones H3 and H4 are expressed from two loci each, *HHT1/HHT2* and *HHF1/HHF2*, respectively. To ensure preservation of the native expression levels, we engineered each strain to include identical point mutations in both relevant loci with separate selection markers (HYG^R and *URA3*) (Fig. 2A). In total, we designed 479 histone mutants, of which 350 were amenable to pE-MAP analysis (Fig. 2, B to D, and table S1); the remaining 129 mutants either were lethal or exhibited very poor growth, rendering them inaccessible to genetic analysis (fig. S1). The histone mutants were crossed against a library of 1370 gene deletions and hypomorphic alleles (table S1) using our triple-mutant selection strategy (*26*, *27*) involving three different selectable markers (HYG^R and *URA3* to select for both copies of the histone alleles and KAN^R for the knockout library strains) (Fig. 2A and Methods) (*26*). Genetic interactions were quantified using the S-score (*28*), which measures the deviation of the double mutant fitness from the expected combined effect of the individual mutations (Methods). The pE-MAP screen was carried out in three biological replicates (Methods), which exhibit a high reproducibility (Fig. 2E), and the final S-scores (as depicted in Fig. 1C) are the averages of these replicates.

It has been shown that a pE-MAP can be used to predict protein-protein interactions (PPIs) by comparing the genetic interaction patterns between pairs of deletion mutants across all the point mutants (*15*, *29*). On a global level, this is only possible if the point mutant set affects a broad group of processes and exhibits genetic interactions with the many different deletion mutants that encode the PPI proteins. Because the histone mutant collection perturbs only two proteins (H3 and H4), we set out to investigate whether the resulting phenotypic profiles are sufficient to predict PPIs among the 1370 deletion mutants. Using a receiver operating characteristic (ROC) curve, we find that the histone pE-MAP predicts PPIs similarly to previous E-MAPs that affect more genes (*15*, *29*) (Fig. 2F and Methods). This finding indicates that the combined set of histone point mutants affects a broad set of cellular processes, reflecting the multifunctional nature of histones H3 and H4 and their central role in controlling the global genetic environment of cells.

To gain insight into the regulatory hierarchy that drives the widespread functional effects of histone perturbations, we set out to examine the relationship between genetic interactions and gene expression changes. To this end, we determined the genome-wide gene expression levels for 29 representative histone mutants using RNA sequencing (RNA-seq) and found no correlation between the expression change of a gene resulting from a given histone mutation and the corresponding S-score (Fig. 2G and table S2). This indicates that observed genetic interactions between histone mutations and deletion mutants are due to complex regulatory patterns, rather than the histone mutation directly modulating the expression of the interacting gene.

The pE-MAP was clustered hierarchically along both dimensions (Fig. 3A and data S1) and effectively recapitulates known protein complex and pathway memberships. For example, the pE-MAP identified COMPASS (*30*, *31*),



Mutation introduced	Hypothesized effect	H3	H4	Total
non-Ala → Ala	None	93	75	168
Ala → Ser	None	16	3	19
Lys → Arg	Deacetylated Lys mimic	13	11	24
Lys → Gln	Acetylated Lys mimic	16	9	25
Ser, Thr → Asp; Tyr → Glu,Phe	Phosphorylation regulation	17	9	26
Pro → Val	Prevents isomerization	2	0	2
Arg → Lys	Same charge; diff. residue	13	11	24
Asp → Asn; Glu → Gln; His → Gln	Charge removal	10	7	17
Asn → Asp; Gln → Glu	Charge addition	3	2	5
N-term. tail deletions	Multiple changes	31	9	40
		214	136	350

Fig. 2. Genetic interrogation of histones H3 and H4 at a residue-level resolution. (**A**) Each histone mutant strain was modified at both native loci (*HHT1* and *HHT2* for H3 or *HHF1* and *HHF2* for H4, red stars) and crossed against a library of 1370 different deletion mutants (or hypomorphic alleles for essential genes). Chr, chromosome. (**B**) Schematic of the histone point mutants analyzed in this study (table S1). Secondary structure elements are indicated as ribbons above the amino acid sequence. The mutations are color coded according to the mutation introduced (C). Mutations resulting in inviable strains or strains too sick for genetic analysis are shown in fig. S1. (**C**) Table of histone mutant categories and their hypothesized effects [color coding as in (B)]. (**D**) Overview of viable H3 and H4 tail deletion mutants amenable to pE-MAP analysis. The amino acid sequences of the WT alleles are shown on top (residues 1 to 39 of histone H3 and 1 to 27 of histone H4). Gray bars represent the deleted residues in H3 and H4. (**E**) Reproducibility of histone pE-MAP S-scores between biological replicates. Plotted are all S-score pairs among the biological replicas, which include triplicate measurements for 346 histone alleles and duplicates of four alleles (H4E73Q, H4H18A, H4121A, and H4K44Q). (**F**) ROC curves showing the power to predict physical interactions between pairs of proteins from this pE-MAP (blue) as well as a previously published pE-MAP [green, (*15*)] and E-MAP [black, (*29*)] data. (**G**) Relationship between gene expression (log_2 fold change over WT) and S-scores of 29 H3 and H4 alleles (table S2). Data from all 1256 deletion library mutants that were measured in both RNA-seq expression and pE-MAP analysis are plotted. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

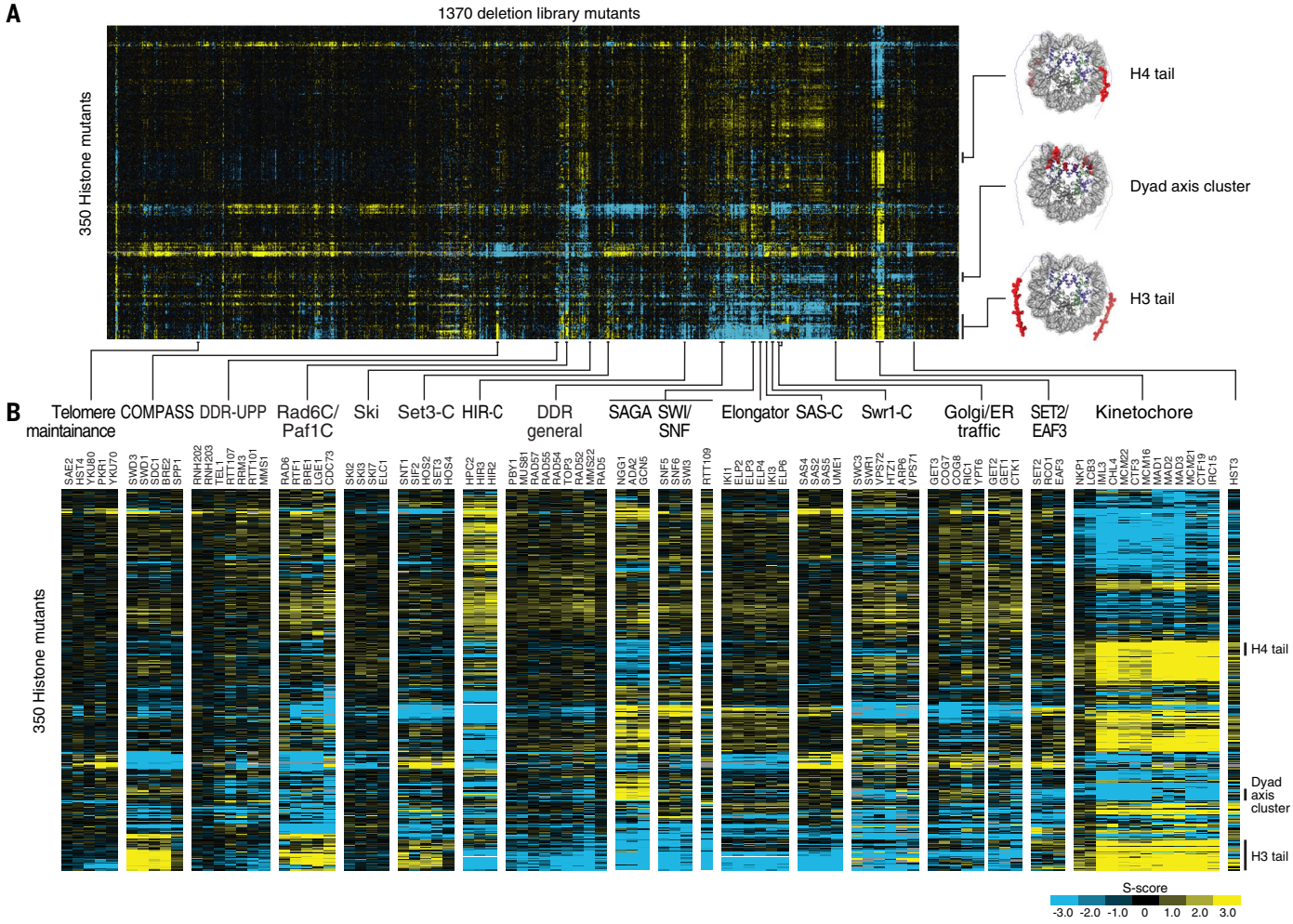


Fig. 3. The genetic interaction landscape of histones H3 and H4. (**A**) Hierarchically clustered pE-MAP of 350 histone H3 and H4 alleles screened against a library of 1370 deletion mutants or hypomorphic alleles. The pE-MAP consists of more than 479,000 genetic interactions. Positive (suppressive or epistatic) and negative (synthetic sick) genetic interactions are colored in yellow or blue, respectively. Examples of histone alleles with similar genetic interaction profiles are highlighted on the right side in the context of the nucleosome structure. The nucleosome structure is modified from PDB 1ID3 (data S2), with H3 in purple, H4 in green, and mutated or deleted residues highlighted in red. N-terminal tail residues of H3 and H4 not included in PDB 1ID3 are visualized as strings on the periphery. (**B**) Examples of genetic interaction profiles of gene clusters belonging to known protein complexes or biological pathways are highlighted and their genetic interaction profiles enlarged from (A). DDR, DNA damage or repair; UPP, ubiquitin proteasome pathway; SAGA, Spt-Ada-Gcn5 acetyltransferase; SWI/SNF, SWItch/sucrose non-fermentable.

Swr1-C (*32*), and the Set2-Eaf3 pathway (*33*–*35*), as well as clusters of genes linked to telomere maintenance and Golgi-ER traffic (Fig. 3B and data S1). Furthermore, mutations of histone residues in close proximity to each other (e.g., mutants of the H3 or H4 N-terminal tails) tend to show similar phenotypic profiles (Fig. 3 and fig. S2A). Overall, we find that histone tail deletion mutants give rise to stronger phenotypic profiles than the point mutants (fig. S2B), reflecting the multiple residue perturbations and the importance of functional histone tails for cell homeostasis.

Phenotypic profile similarities are correlated with structural proximity

Similarities between pairs of phenotypic profiles in the histone H3-H4 pE-MAP were quantified by the maximal information coefficient (MIC) (*36*, *37*) (Fig. 1D, fig. S3, and Methods). The MIC values between pairs of phenotypic profiles do not linearly correlate with the distances between the mutated residues in the WT structure (Pearson correlation coefficient of −0.07, Cα-Cα distances) but are informative about an upper distance bound between the residues (Fig. 4A and fig. S3C). The upper distance bound was obtained by binning the MIC values into 20 intervals and selecting the maximum distance spanned by any pair of residues in each bin, followed by fitting a logarithmic decay function to these maximum distances (Fig. 4A, fig. S3C, and Methods). The data show that a pair of proximal point mutations are more likely to have a high MIC value than a pair of distal point mutations. However, not all proximal mutations have a high MIC value: Most pairs of phenotypic profiles, even those for residues that are less than 16 Å apart, are highly dissimilar (94% of all pairs exhibit a MIC value <0.3). These observations justify converting the pE-MAP data into a Bayesian data likelihood that provides an upper bound on the distance spanned by the mutated residues (Fig. 4B and Methods). This Bayesian term objectively interprets the noise in the experimental data and allows us to quantify the uncertainty of the resulting structural models. The complete scoring function for evaluating any structural model also includes simple terms accounting for excluded volume and sequence connectivity, in addition to the Bayesian terms for all pairs of profiles in the pE-MAP with a MIC value above 0.3 (Fig. 5).

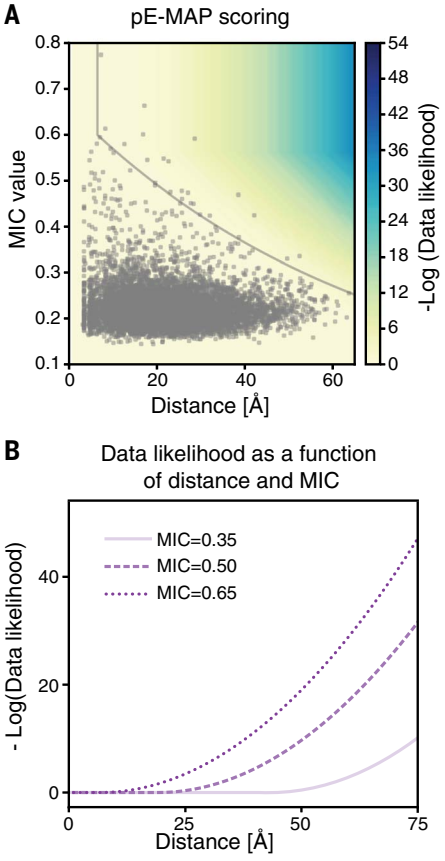


Fig. 4. Generation of the scoring function. (**A**) Relationship between pairwise distances and MIC values. The solid gray line represents the logarithmic decay fit to the upper distance bounds (Methods, Eq.1). The background color gradient reflects how the data likelihood depends on MIC value and distance. (**B**) –Log of the data likelihood as a function of distance for different MIC values (Methods).

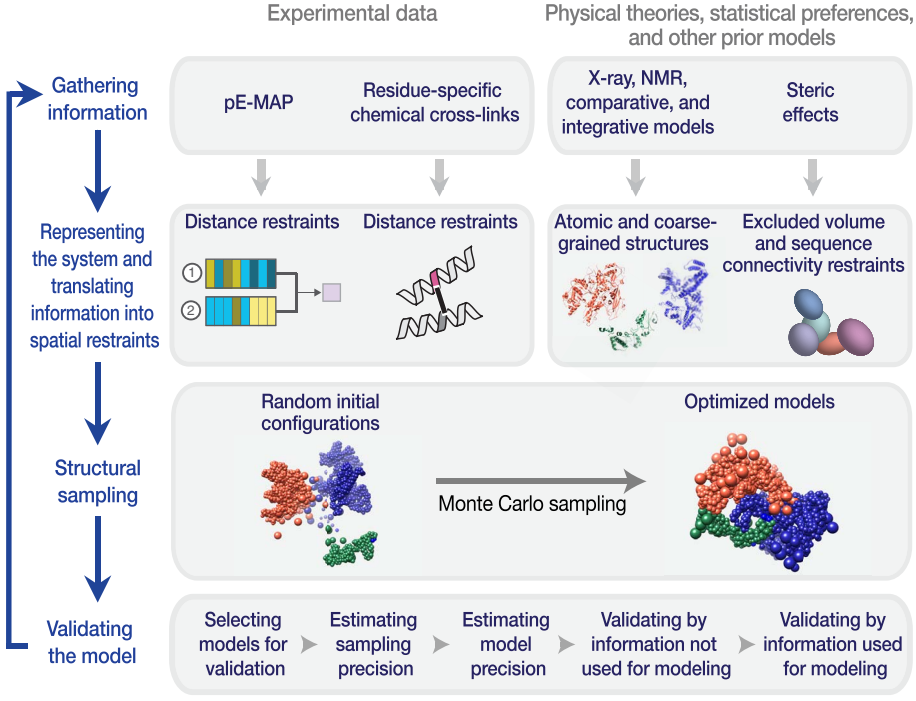


Fig. 5. Description of the integrative modeling workflow. The four stages include (i) gathering all available experimental data and prior information, (ii) translating all information into a representation of the assembly components and a scoring function for ranking alternative assembly structures, (iii) sampling structural models, and (iv) validating the model. In this example, the representation of the components of a complex is based on comparative models of its components. The scoring function consists of spatial restraints that are obtained from pE-MAP and/or cross-linking experiments (evolutionary coupling analysis is not indicated in this scheme) as well as excluded volume and sequence connectivity restraints. The sampling explores the configurations of rigid components, searching for those assembly structures that satisfy the spatial restraints as well as possible. The goal is to obtain an ensemble of structures that satisfy the input data within the uncertainty of the data used to compute them. The sampling precision is estimated and models are clustered and evaluated by the degree to which they satisfy the input information used to construct them as well as omitted information. The protocol can iterate through the four stages until the models are judged to be satisfactory, most often based on their precision and the degree to which they satisfy the data.

Spatial restraints derived from pE-MAP data can be used for integrative structure determination

An ensemble of the H3-H4 dimer configurations that satisfy the input information (i.e., the model) was found by exhaustive Monte Carlo sampling guided by the scoring function, starting with random initial configurations of the rigid comparative models of the H3 and H4 subunits (Fig. 5 and Methods). The resulting ensemble is accurate and precise, as demonstrated by the similarity between the x-ray structure [Protein Data Bank (PDB) 1ID3, (*38*)] and model contact maps (Fig. 6, A and B). Specifically, the mean accuracy is 3.8 Å (Fig. 6C); the accuracy is defined as the average Cα root-mean-square deviation (RMSD) between the x-ray structure and each of the structures in the ensemble. The precision is 1.0 Å (Fig. 6C), which is defined as the average RMSD between all solutions in the ensemble. As a control, we also computed a model from randomly shuffled MIC values. The resulting model (Fig. 6D) is incorrect (mean accuracy of 15.8 Å and incorrect contact map; Fig. 6, C and D) and imprecise (7.6 Å; Fig. 6C). As another control, we computed a model by a state-of-the-art protein-protein docking method (*39*), resulting in a model with an inferior accuracy of 6.9 Å (fig. S4). Finally, we also mapped the accuracy and precision of the model as a function of the fraction of the pE-MAP data used (Methods). As expected, the more pE-MAP data that are used, the more accurate and precise is the model (Fig. 6C).

To compute the structure of a protein complex for which the structures of the components are known, we estimate that 35 to 40 mutations per component are necessary to generate a complex model with precision sufficient to map the positions and relative orientations of the components (fig. S5 and Methods). What is a useful model precision depends on the questions asked (*40*). Fortunately, many questions can often be answered by models as precise as those obtained based on pE-MAP data (RMSD range of 1 to 15 Å). Some examples include describing the architecture and evolution of protein assemblies (*8*, *41*), designing interface mutations (*42*), characterizing structural heterogeneity of protein complexes (*42*, *43*), and mapping binding-induced structural changes (*44*). Importantly, the estimate of 35 to 40 mutations is an upper bound, and, in many cases, the number might be reduced by specifically exploiting point mutations that target surface residues and/or residues known to be functionally important and by choosing substitutions likely to give rise to functional perturbations. The outcome of these calculations indicates the utility of the pE-MAP data for integrative structure determination.

The pE-MAP connects individual histone residues and regions to other associated complexes and processes

To examine whether the pE-MAP can identify interactions with complexes that are not stably associated with histones, we investigated the relationships between modifiable histone residues and their cognate enzymes (modifier pairs). Interestingly, we observed a dramatic

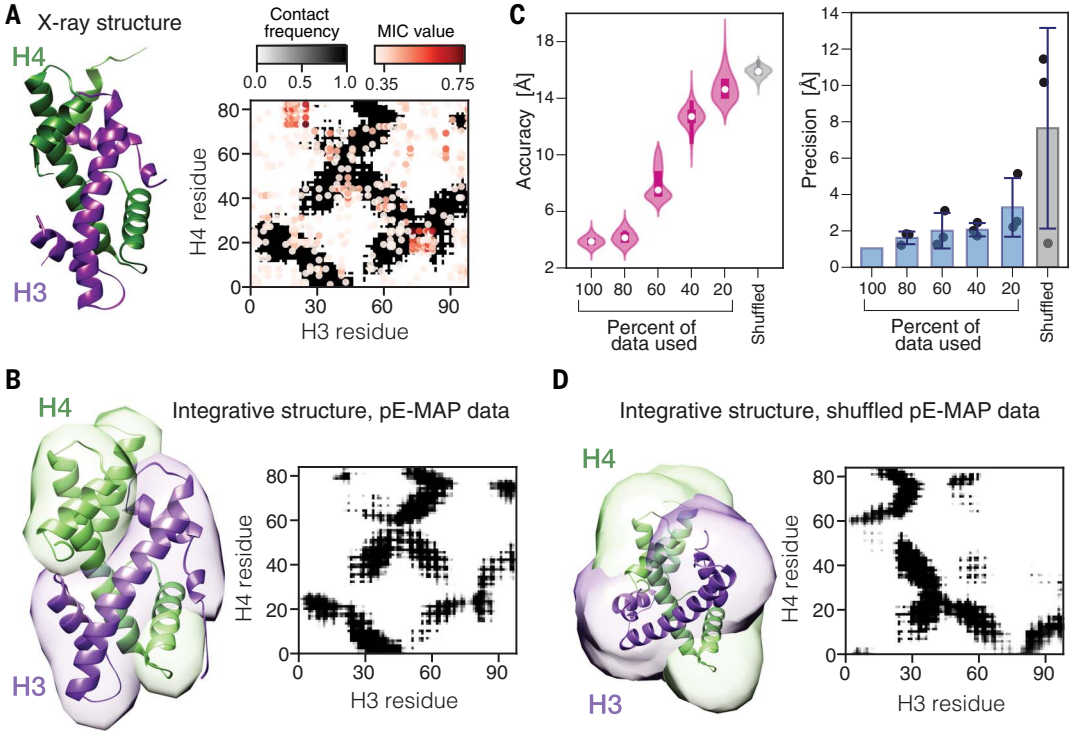


Fig. 6. Integrative structure determination of histones H3 and H4. (**A**) The native structure of the histone H3-H4 dimer (PDB 1ID3, left) and its contact map (right). In contact maps, the intensity of gray is proportional to the relative frequency of residue-residue contacts in the models (cutoff distance of 12 Å). For x-ray structures, the contact frequency is either 0 (white) or 1 (black). The circles correspond to the pairs of restrained residues, with the intensity of red proportional to the MIC value (MIC > 0.3), showing that the pairs of residues with high MIC values are distributed throughout the proteins. (**B**) The localization probability density of the ensemble of structures is shown with a representative (centroid) structure from the computed ensemble embedded within it (left) and the corresponding contact map (right). The localization probability density map represents the probability of any volume element being occupied by a given protein. (**C**) Distributions of accuracy (left) of structures in the ensembles and model precisions (right) based on the full pE-MAP dataset, resampled datasets that consider fractions of the data, and using shuffled MIC values. On the left, the white dots represent median accuracies, and on the right, the error bars represent the standard deviations of model precision over three independent realizations (shown as dots). (**D**) Localization probability density and centroid structure (left) and contact map (right) computed with shuffled MIC values (Methods).

increase in S-scores within specific modifier pairs, as compared with the overall genetic interaction distribution (Fig. 7A and table S4). The positive S-scores reflect that a modifier and its target residue often are epistatic or suppressive because they function in the same pathway. To test if this pattern extends to phenotypic profile similarities, we integrated the histone pE-MAP into a merged map of previously collected genetic interaction data for gene deletions and hypomorphic alleles (*45*). We computed Pearson correlation coefficients for each histone mutant phenotypic profile across the merged map, generating a correlation map of 350 histone mutants against 4414 whole-gene perturbations (data S3 and Methods). In agreement with the individual S-scores, the specific modifier pairs exhibit significantly higher phenotypic profile correlations than the overall map (Fig. 7A and table S4). These findings show that the pE-MAP can be used to pair specific residues to their respective modifiers, even though these are not stably associated with the histones. For example, when components of COMPASS, which methylates histone H3K4 (*30*, *31*, *46*), are deleted (*swd1*Δ, *swd3*Δ, *sdc1*Δ, *bre2*Δ), we observe strong positive S-scores with both H3K4 mutants [K4R (Lys^4→Arg) and K4Q (Lys^4→Gln)] as well as high correlations of the phenotypic profiles between the H3K4 mutants and COMPASS deletions (Fig. 7B and data S1 and S2).

To explore these relationships in a structural context, we developed a Cytoscape (*47*) app named stE-MAP (structure E-MAP) that interactively maps the genetic interactions of pE-MAP gene clusters onto the point-mutated protein structure. stE-MAP connects Cytoscape to ChimeraX (*48*) and displays connections between a predefined set of genes and all mutated residues for which the underlying interactions pass user-defined criteria (fig. S6A). We mapped the genetic connections between COMPASS and all histone residues with which it exhibits >0.2 median correlation (Methods). Only five residues pass this threshold, and the strongest connection is displayed by H3K4. The other four residues (H3K1 to H3K3 and H3K5) are proximal and are thus likely to interfere with the interaction between COMPASS and H3K4 (Fig. 7C). This finding is particularly notable because these residues reside in the most distal region of the unstructured H3 N-terminal tail. Given that we do not have COMPASS point mutations in our dataset, we did not attempt to model this interaction. However, analysis of the MIC values associated with the H3 and H4 tails, and their relationship with the core domains, indicates that distance restraints for the histone tails could be derived from the pE-MAP data. Specifically, the MIC value distributions for the tail-core and tail-tail pairs of mutations are similar to that of the core-core mutations (fig. S5C). This similarity indicates that we can derive distance restraints for the histone tails, thus, in principle, supporting the feasibility of integrative structure modeling of disordered regions. In such modeling, to avoid overinterpretation of the data and to account for a possibility that the pE-MAP data of the histone tails reflect interactions with neighboring nucleosomes, we would have to include multiple nucleosome copies in the model and allow for assignment ambiguity, just as we do for distances inferred from chemical cross-links and protein proximity inferred from affinity copurification (*49*).

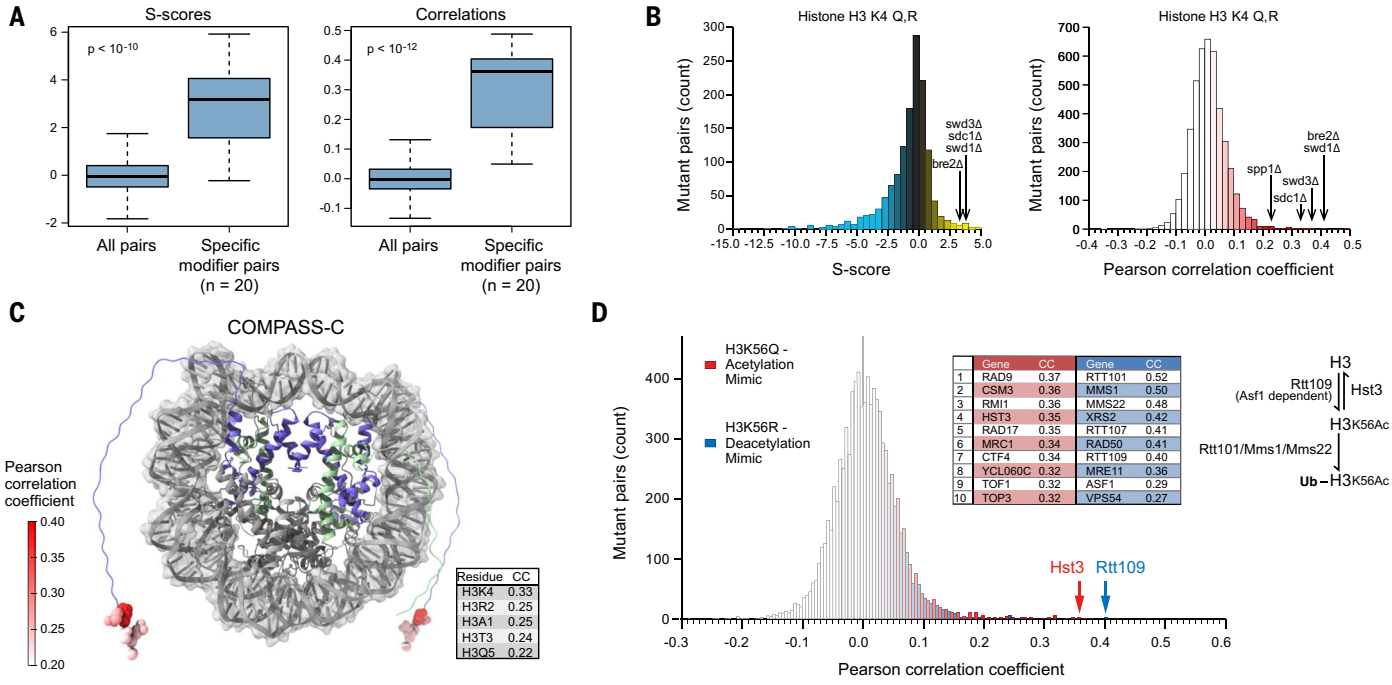


Fig. 7. Connecting individual histone residues and regions to other associated complexes. (**A**) Comparison of S-scores and Pearson correlation coefficients of phenotypic profiles of modifier-residue pairs to the overall data. Only residues with a single known modifier and modifiers with a single known target residue were included (table S4). *p* values were calculated using two-sided Wilcoxon rank sum tests. The whiskers of the boxplots extend to a maximum of 1.5 × IQR (interquartile range), and outliers are not plotted. (**B**) Average distributions of S-scores (left) and phenotypic profile correlations (right) of H3K4 mutants (mean of H3K4Q and H3K4R). Members of the COMPASS complex that exhibit a mean S-score >2.5 or a mean genetic interaction profile correlation >0.2 with H3K4 mutants are highlighted. The COMPASS complex is responsible for H3K4 methylation. (**C**) Mapping of genetic interaction profile correlations to COMPASS complex members on the structure of the nucleosome (modified PDB 1ID3, data S2). N-terminal tail residues of H3 and H4 not included in 1ID3 are visualized as strings on the periphery. Only residues that exhibit a median genetic profile correlation >0.2 with the COMPASS subunits are highlighted (Methods). H3 is depicted in purple, H4 in light green, and H2A/H2B and DNA in gray. The red color gradient reflects the strength of the correlation between each residue and the COMPASS members, calculated as the median correlation between the residue's tested mutations and the COMPASS members. (**D**) Distributions of genetic interaction profile correlations of H3K56Q (acetylation mimic) and H3K56R (deacetylation mimic). Correlations of key H3K56ac-level regulators, Rtt109 (acetylating) and Hst3 (deacetylating), are highlighted. The cartoon outlines the H3K56 acetylation pathway and its role in H3 ubiquitylation. Rtt109 acetylates H3K56 through an Asf1-dependent mechanism, which promotes ubiquitylation of H3 by Rtt101-Mms1 and Mms22. These five gene deletions are all found among the top 10 most similar to the deacetylation mimic H3K56R, whereas deletion of the H3K56 deacetylase Hst3 instead gives rise to a profile similar to the acetylation mimic H3K56Q (table inset). CC, Pearson correlation coefficient.

We observe similar trends for other histone modifiers. For example, members of the Set2 pathway (Set2, Eaf3, Rco1, Ctk1) (*33*–*35*) rank highly in the distributions of S-scores or correlations for mutations of their target residue, H3K36 (fig. S6, B and C). Interestingly, we also found instances where different mutations of a single residue identify connections to different modifiers. For example, the phenotypic profile of the deacetylation mimic H3K56R is similar to that of deletion of *RTT109*, which encodes the H3K56 acetylase (Pearson correlation coefficient of 0.4) (*29*), whereas the acetylation mimic H3K56Q instead correlates with the profile generated from the deletion of the corresponding deacetylase, *HST3* (Pearson correlation coefficient of 0.35) (*50*) (Fig. 7D). H3K56R further correlates with *asf1*Δ, *rtt101*Δ, *mms1*Δ, and *mms22*Δ, whose corresponding proteins play key roles in the H3K56 acetylation pathway and downstream H3 ubiquitylation (*51*, *52*) (Fig. 7D). Accordingly, the stE-MAP app identified strong links between Hst3-Hst4 and H3K56Q, as well as Rtt109-Asf1 and H3K56R (fig. S6D). Although we find that it is often informative to group different mutations of the same residue together, these examples highlight the potential of these maps for deeper mechanistic insights where required.

Expanding on these findings, we built a gene set enrichment map connecting the modifiable histone residues to nuclear processes (fig. S7A, table S5, and Methods). We observe both known and previously unknown connections. For example, "DNA recombination and repair" is connected to four residues, and two of these, H3K56 and H3K79, have been shown to play key roles in yeast's DNA repair (*53*–*57*). Interestingly, we find that mutations of the other two residues (H3R63K and H4R36K) result in increased spontaneous mutation frequencies at the *URA3* locus, indicating that these residues also function in DNA repair (fig. S7, B and C; table S6; and Methods).

The gene set enrichment analysis also identified 13 residues connected to cryptic transcription (fig. S7A). The pE-MAP includes 24 different mutations of these residues, and we tested their involvement in cryptic transcription by quantifying the abundance of transcripts at the 5′ and 3′ end of the *STE11* gene, using quantitative polymerase chain reaction (qPCR) (fig. S7D and Methods). In total, 16 mutations, distributed among 10 residues, increase 3′ transcript abundance by >50% compared with WT (table S7), and nine mutants among five residues increase 3′ transcription more than twofold, without major changes in 5′ transcription (fig. S7E). As expected, H3K36A, H3K36R, H3K36Q, and *set2*Δ increase 3′ transcript abundance strongly, as do mutations of H4K44, which is a residue known to affect cryptic transcription (*58*). Interestingly, H3K122A increases 3′-transcript abundance >15-fold and, using ATAC-seq, we find that the mutation gives rise to nucleosome-free regions in *STE11* and other genes known to produce cryptic transcripts (fig. S7, F to H). H3K122A exhibits positive genetic interactions with deletion of the

histone chaperone *SPT2* (S-score = 2.4) and the nucleosome remodeling factor *CHD1* (S-score = 4.9), which are both involved in cryptic transcription (*59*, *60*). Accordingly, we find that deletion of either *SPT2* or *CHD1* suppresses the cryptic transcription phenotype observed in H3K122A to WT levels, even though *spt2*Δ or *chd1*Δ alone has no effect (fig. S7, I to K).

The integrative structure determination approach is transferable to other complexes

To test whether genetic interaction mapping can be used to determine the structure of other complexes, we examined a pE-MAP of RNAPII in budding yeast (*15*). This pE-MAP consists of 53 point mutants crossed against a library of 1200 gene deletions and hypomorphic alleles. Interestingly, the association between MIC values and the upper distance bound is also apparent in this dataset (fig. S8, A and B), even though the protein sizes and mutational coverage of the polymerase system (up to ~1700 residues and 1 to 2%, respectively) are vastly different from those of the histones (<140 residues and 85 to 90%). These observations suggest that our parameterization of the pE-MAP spatial restraint based on the histone data may be generally applicable. To evaluate this expectation directly, we next modeled subunits Rpb1 and Rpb2 of RNAPII using the Bayesian likelihood parametrization based on the histone pE-MAP. To illustrate the modeling of higher-order complexes, we divided Rpb1 into two domains, thereby representing the system with three rigid bodies (Methods and table S8). We obtained a model with a mean accuracy of 16.8 Å and precision of 9.8 Å (Fig. 8, A to D, and table S8). This positive result illustrates the generality of the pE-MAP–based spatial restraints.

To further assess the utility of pE-MAP data for structure determination, we compared the RNAPII model obtained using the pE-MAP to a model using 22 previously published chemical cross-links (*61*). Cross-linking is widely used for integrative structure determination of macromolecular assemblies (*2*, *8*). Interestingly, a model of yeast RNAPII based on the pE-MAP data is as accurate as that based on the cross-links (16.8 and 16.7 Å, respectively; Fig. 8D). Moreover, the accuracy and precision of the model improves if both datasets are used simultaneously (10.2 and 3.7 Å, respectively; Fig. 8D), indicating complementarity between the two types of data and demonstrating a premise of integrative structure determination (Fig. 5). Although a cross-link between two residues may provide more direct structural information than the corresponding pE-MAP pair, the number of possible cross-links is limited by the number of proximal reactive residue pairs, whereas the number of pE-MAP pairs grows quadratically with every additional point mutation introduced. Therefore, the larger number of less precise pE-MAP restraints can lead to a more accurate model than a smaller number of more precise cross-links.

The integrative structure determination approach is transferable to other types of phenotypic profiles

To examine the applicability of our approach to other types of phenotypic profiles, we turned to a CG-MAP of 44 bacterial RNAP point mutations exposed to 83 different environmental stresses (e.g., chemical perturbations, temperature stress, and pH change) (*20*). We observe an association between MIC values and the upper distance bound, similar to that of the pE-MAP datasets (fig. S8 and Methods). We modeled the structure of subunits RpoB and RpoC of the bacterial RNAP with a mean accuracy of 15.0 Å and precision of 6.6 Å (Fig. 8, E to H, and table S9). This result suggests that maps with relatively small numbers of orthogonal phenotypes per point mutation can be used to accurately predict the architecture of macromolecular assemblies. Considering that constructing large gene deletion libraries and crossing them against point mutations can be laborious, environmental phenotypic profiles may be a more efficient alternative for generating spatial restraints for integrative structure determination than genetic interaction phenotypic profiles.

Spatial restraints derived from pE-MAP data are comparable to other commonly used data types

Coevolution information can also be used to predict the structure of protein assemblies (*18*, *62*, *63*). However, the success of such modeling is heavily dependent on the number of sequences in the input sequence alignments and the ability to discriminate interacting from noninteracting homologs in genomes with multiple paralogs (*64*). Using the RaptorX protein complex contact prediction server (*65*, *66*), we predicted the interfacial contacts between RpoB and RpoC of the bacterial RNAP; the numbers of homologous sequences were insufficient for the yeast histones and RNAPII (Methods). Importantly, RaptorX is based on a combination of coevolution analysis and a deep-learning algorithm that reduces the requirement for sequence homologs and improves accuracy (*67*). Other commonly used coevolution methods (*18*, *62*) did not identify any interfacial contacts. Similar to the pE-MAP and CG-MAP datasets, we observe a negative statistical association between the residue pair coupling strengths and the upper distance bounds (fig. S9). To mimic the pE-MAP restraint, we converted the top coupling strengths into upper distance bound restraints (Methods). The model ensemble computed from coevolution-derived restraints includes two different sets of configurations (mean accuracy of 22.6 Å; model precision of 9.0 Å; Fig. 8H). Only a fraction of the bacterial RNAP structures computed using coevolution-derived restraints are as accurate as those computed using the CG-MAP restraints. The model precision and accuracy of the model improve slightly if both types of restraints are combined (mean accuracy of 14.5 Å; model precision of 6.5 Å; Fig. 8H).

Discussion

We show that the architectures of macromolecular assemblies can be determined using quantitative genetic interaction data collected in vivo. The accuracy and precision of such models are comparable to those of models based on chemical cross-linking or coevolution analysis. A key premise of integrative modeling is that using several different types of data improves the accuracy and precision of the model. Because the pE-MAPs and CG-MAPs contain purely phenotypic measurements, collected in living cells, these datasets generate spatial restraints that are orthogonal to other commonly used data for integrative modeling. Because these data reflect in vivo structures, and are thus unlikely to share artifacts of biophysical methods, they could be of particularly high value in the integrative modeling process. The genetic interaction data may also allow for the characterization of complexes that are difficult to isolate and purify or those that are only transiently stable. Importantly, the equipment required for generating these data is basic, and, in particular, the CG-MAPs can be generated efficiently. Recent developments in CRISPR-Cas9 based approaches have paved the way for multiplexed precision genome editing in yeast (*68*), allowing for rapid generation of CG-MAPs. Together, these methods make feasible the proteome-wide modeling of protein complex structures, guided by global protein-protein interaction maps (*69*). In addition to proteins, the approach is also applicable to assemblies containing nucleic acids, thus further expanding the scope of integrative structural biology. pE-MAPs and CG-MAPs are complementary to other high-throughput functional assays. For example, two recent studies have shown that deep mutational scans can be used for determining the fold of a small protein domain (*70*, *71*). If these methods prove useful for multidomain proteins or disordered regions (*72*), then the measured phenotype changes could be used to derive additional restraints for the integrative structure determination.

The relationship between phenotypic pE-MAP measurements and structure can be uncertain. The reasons for this include mutations in distant positions that are part of an allosteric network and could give rise to similar profiles, mutations that are functionally irrelevant, and mutations that perturb gene expression, mRNA stability, or translation. Additionally, the approach relies on the introduction of point mutations

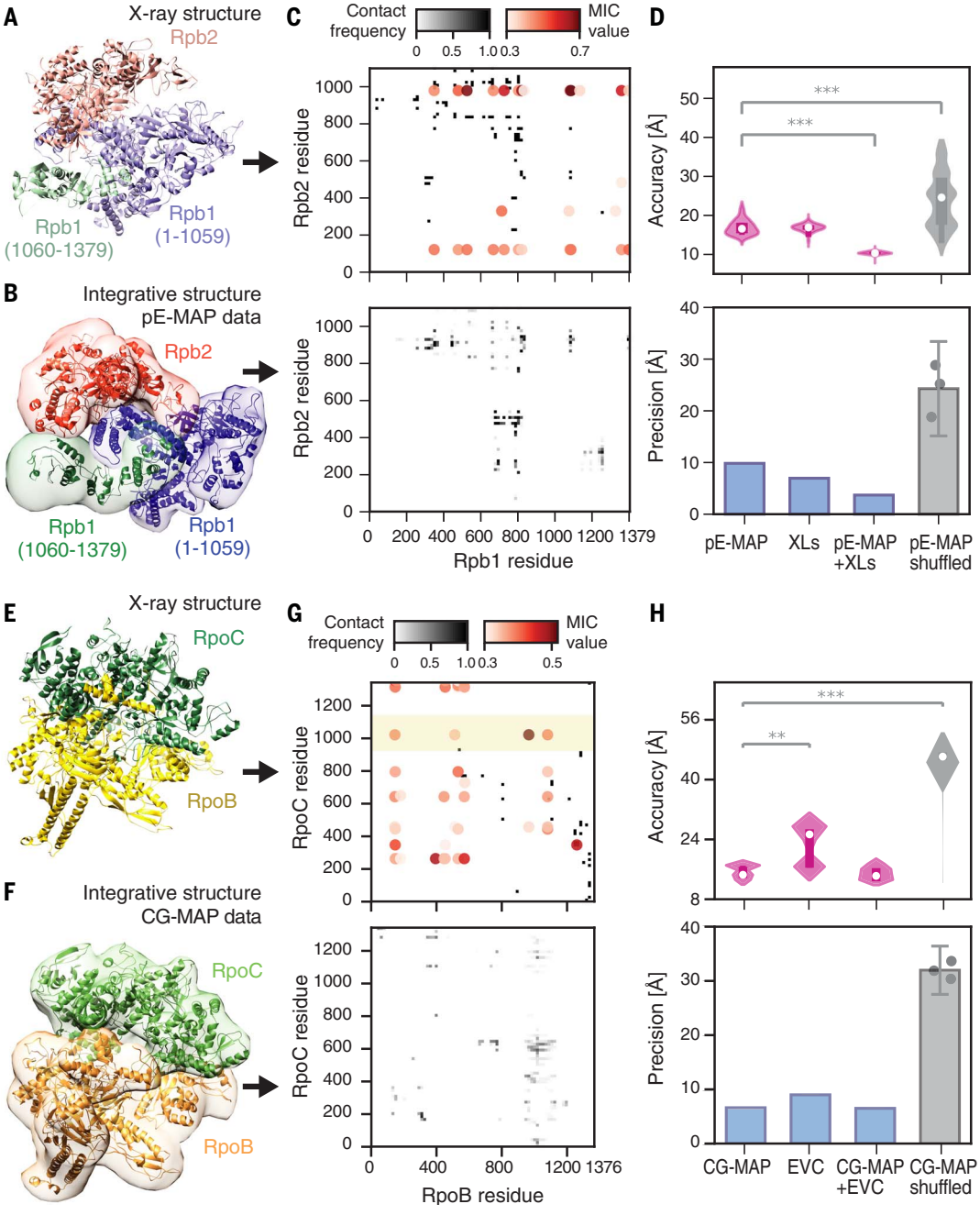


Fig. 8. Integrative structure determination of yeast RNAPII and bacterial RNAP. (**A**) The native structure of Rpb1-Rpb2 (PDB 2E2H) showing its three rigid-body components. Rpb1 was split into two domains, as shown. (**B**) The localization probability density of the ensemble of the three rigid-body structures is shown with a representative (centroid) structure from the computed ensemble embedded within it. (**C**) Contact maps computed for the x-ray structure (top) and model using the pE-MAP dataset (bottom). The circles correspond to the pairs of restrained residues, with the intensity of red proportional to the MIC value (MIC > 0.3). (**D**) Distributions of accuracy (top) for all structures in the ensemble and model precisions (bottom) for the computed ensembles based on pE-MAP and cross-link data. The white dots represent median accuracies. Error bars represent the standard deviations of model precisions over three independent realizations (shown as dots); $***p < 10^{-12}$. (**E**) Structure of subunits RpoB and RpoC from bacterial RNAP (PDB 4YG2). (**F**) The localization probability density of the ensemble of the RpoB-RpoC structures with a representative (centroid) structure from the computed ensemble embedded within it. (**G**) Contact maps computed for the x-ray structure (top) and model using the CG-MAP dataset (bottom). The shaded yellow band represents a region missing in the x-ray structure. (**H**) Distributions of accuracy (top) for all structures in the ensemble and model precisions (bottom) for the ensembles based on CG-MAP and evolutionary coupling (EVC) data. The white dots represent median accuracies. The error bars represent the standard deviations of model precision over three independent realizations (shown as dots). $**p < 10^{-6}$; $***p < 10^{-12}$.

into the proteins of interest, which may result in structural changes. However, proteins often adapt to mutations by small local changes in their structure, maintaining their overall fold and function (*73*). Mutations that cause major misfolding of essential proteins and/or assemblies are uncommon in pE-MAPs because the resulting fitness defects typically prevent successful screening. The method could be improved by specifically designing point mutants that do not alter the structure and/or lead to aggregation, by selecting commonly allowed mutations as determined by divergent protein sequence alignments (*74*).

The aim of integrative structure determination is to model the structures of macromolecular assemblies. This often requires the structures of the individual components [from x-ray crystallography, NMR, cryo-EM, comparative modeling, or, increasingly, ab initio structure prediction (*65*, *67*, *75*, *76*)]. The quality of the structures of the individual components and input data are crucial for integrative (or indeed any other) structural approaches, and one cannot achieve a precise structure from low-quality starting structures or data. Even so, there are numerous examples of utility of structural models at lower resolution (*3*). For example, these models can be used to explain the architectural principles of large assemblies (*8*, *41*, *77*, *78*), describe the structural dynamics of protein complexes (*42*, *43*), or rationalize the impact of many mutations (*41*). A lower-resolution structure is also often a useful starting point for higher-resolution structure characterization.

CRISPR-Cas9 genome editing (*79*) has proven highly effective for high-throughput genetic interaction mapping in mammalian cells (*80*, *81*). To date, these efforts have relied on whole-gene perturbations, but methods for systematic generation of point mutants using CRISPR-Cas9 have recently been developed (*82*, *83*), paving the way for mammalian pE-MAP screening. This advance provides a means for integrative structure determination of assemblies in human cells and also allows for identification and characterization of functionally relevant structural changes that take place in disease alleles. Expanding this analysis to host-pathogen complexes (*84*–*86*) will be feasible by introducing specific mutations into the pathogenic genome and studying the phenotypic consequences using genetic interaction profiling of relevant host genes (*87*). Furthermore, several efforts are under way to generate multiscale models of entire cells (*88*–*92*). In such instances, high-throughput genetic interaction mapping could provide global insights into cellular organization and dynamics of different components while also informing on the structures of individual assemblies.

Methods

Histone mutant strain construction

The histone H3 and H4 mutant strain library was constructed essentially as described (*22*, *23*). Briefly, the mutants (tail deletions, complete alanine-scan, and context specific point mutations) were generated in the YMS196 background (MATα *his3*Δ *leu2*Δ *ura3*Δ *can1*::*STE2pr-spHIS5 lyp1*::*STE3pr-LEU2*) (table S1). First, the base strains were created by replacing the *HHT2-HHF2* locus with a *URA3*-containing cassette carrying a mutated *HHT2-HHF2* locus with their endogenous promoters. We randomly picked a few base strains, and the mutated *HHT2-HHF2* loci were PCR amplified and validated by sequencing. Then, the *HHT1-HHF1* locus was replaced with a HYG^R-containing cassette carrying a mutated version of the *HHT1-HHF1* locus, resulting in pE-MAP-amenable strains (Matα *his3*Δ *leu2*Δ *ura3*Δ *hht1-hhf1*::HYG^R *hht2-hhf2*::*URA3 can1*::*STE2pr-spHIS5 lyp1*::*STE3pr-LEU2*).

Histone mutant library validation

Libraries were validated in three steps: (i) Each mutant was constructed, transformed into bacteria, and sequenced; 100% sequence identity was required to pass quality control. (ii) After the correct integration of histone mutants in the *HHT2-HHF2* locus, 5 to 10 yeast base strains from each 96-well plate were randomly selected, and corresponding histone fragments were amplified and sequenced to ensure the identity of each mutant in the well and no cross-contamination during plasmid preparation and yeast transformation; 100% of these were correct. (iii) After obtaining the yeast library with the second (HYG^R) cassette integrated, 5 to 10 yeast strains from each 96-well plate were also randomly selected. Both copies of histone mutants in each strain were amplified and sequenced to confirm the identity of mutations. All of them were correct.

pE-MAP analysis

Each of the histone H3 and H4 mutant strains was crossed with 1370 MATa KAN^R marked deletion (nonessential genes) or DAmP (decreased abundance by mRNA perturbation; essential genes) strains by pinning on solid media as described (*15*). Sporulation was induced, and MATa haploid spores were selected by replica plating onto media containing canavanine (selecting *can1Δ* haploids) and S-AEC (selecting *lyp1Δ* haploids) and lacking histidine (selecting MATa spores). Triple-mutant haploids were isolated on media containing hygromycin (selecting *hht1-hhf1 mutant cassette*) and G418 (selecting KAN^R marked deletion or DAmP), and lacking uracil (selecting *hht2-hhf2 mutant cassette*). Finally, triple-mutant colony sizes were extracted using imaging software. The screen was carried out in three biological replicates with three technical replicates in each. Four mutants (H4E73Q, H4H18A, H4121A, and H4K44Q) failed screening in one biological replicate, and the results for these are based on the two successful replicates. Detailed E-MAP experimental procedures are described in (*26*, *29*, *93*). Genetic interactions were quantified using S-scores (*28*), which are closely related to *t* values. The S-score quantifies the deviation of the double (or triple) mutants from the expected combined fitness effects of the individual mutants and incorporates the reproducibility between technical replicates. The published S-scores represent the average S-scores across biological replicates.

Design of the pE-MAP spatial restraints

The distance restraint based on pE-MAP data was designed using the 308 single-point mutants from the histone pE-MAP and the structure from the PDB 1ID3, as follows: (i) postprocessing of the genetic interaction phenotypic profiles, (ii) devising a phenotypic similarity metric between the phenotypic profiles, and (iii) designing spatial restraints for integrative structure modeling using the phenotypic similarity values and the known nucleosome x-ray structure. Next, we describe each of these three steps in turn.

(i) Postprocessing of the genetic interaction phenotypic profiles: All missing values in the pE-MAP were imputed as the mean of the S-scores between the corresponding deletion mutant and all histone point mutants. To increase the signal-to-noise ratio of the pE-MAP, gene deletion mutants that mostly exhibited weak genetic interactions with the histone mutants were filtered out. To this end, the gene deletion profiles were ranked in descending order based on the counts of their S-scores that fell in either the top 2.5% of positive S-scores or the bottom 5% of negative S-scores, from the complete point mutant pE-MAP (cutoffs calculated after imputation). The more stringent cutoff for positive S-scores was chosen to reflect the smaller dynamic range for positive genetic interactions compared to negative genetic interactions. Gene deletions with the same count were then ranked in descending order by the mean of the absolute values of their highest and lowest score (fig. S3A). The top fraction of the deletions, determined in step (iii) below, were retained for computing the histone point mutant phenotypic profile similarities (below, fig. S3).

(ii) Devising a phenotypic similarity metric between the phenotypic profiles: We computed the similarity between all pairs of histone phenotypic profiles using the maximal information coefficient (MIC; fig. S3B), with the MIC parameters *alpha* and *c* set to 0.6 and 15, respectively, as suggested (*36*, *37*). Many positions in the histones were mutated to several different residue types, giving rise to several phenotypic profiles for each of these positions. As a result, more than one MIC value would often be computed for a single residue pair. In such cases, only the highest MIC value was retained.

(iii) Designing spatial restraints for integrative structure modeling: Using the histone x-ray structure [PDB 1ID3; (*38*)], we measured the Cα-Cα distance between all pairs of residues for which we computed a MIC phenotypic similarity score. The percentage of the top scoring phenotypic profiles [ranked by the genetic interaction scores; step (i)] retained for further analysis was determined as follows. We compared the statistical association of the distances between two mutated residues with their phenotypic similarity by selecting the top 10, 25, 50, and 100% of the ranked deletions (fig. S3C). Although MIC values between phenotypic profiles do not linearly correlate with the distances spanned by the mutated residues in the WT structure (Pearson correlation coefficient of −0.07 when using the top 25% or top 50% of deletions), the MIC values provide an upper distance bound between the residues. The upper distance bound was obtained by binning the MIC values into 20 intervals and selecting the maximum distance spanned by any pair of residues in each bin, followed by fitting a logarithmic decay function (d_U) to the upper distance bounds:

$$d_U(\mathrm{MIC}) = \begin{cases} \dfrac{\log(\mathrm{MIC}) - n}{k} & \text{if MIC} \le 0.6 \\ 6.84 & \text{if MIC} > 0.6 \end{cases} \quad (1)$$

where k and n are −0.0147 and −0.41, respectively (fig. S3C). We find that selecting the top 25 or 50% of the deletions had a comparable

association between the upper distance bounds and the computed MIC values. The association was determined by computing the R and p values of the Pearson correlation coefficient and association significance, respectively, for the log-transformed MIC values. In this work, we retained the top 25% of the ranked phenotypic profiles for computing the phenotypic profile similarities.

To effectively handle the uncertain relationship between the data and model, we use Bayesian inference for scoring alternative models by formulating spatial restraints as Bayesian data likelihoods (*94*). Formally, the posterior probability of model M given data D and prior information I is $p(M|D,I) \propto p(D|M,I) \cdot p(M|I)$. The model, M, consists of a structure X and unknown parameters Y, such as noise in the data. The prior $p(M|I)$ is the probability density of model M given I. The prior can in general reflect information such as statistical potentials or a molecular mechanics force field; here, we only used excluded volume and sequence connectivity. The likelihood function $p(D|M,I)$ is the probability density of observing data D given M and I. The pE-MAP data was used to compute phenotypic similarities (i.e., MIC values) that inform distances between mutated residues pairs i,j. The likelihood of the entire pE-MAP dataset is the product over the individual observations between residue pairs i,j: $p(D|M,I) = \prod_{i,j} N[d_{i,j}|\, f_{i,j}(X), \sigma_{i,j}]$, where $f_{i,j}(X)$ is a forward model that predicts the data point $d_{i,j}$ in D that would have been observed for structure X in an experiment without noise; $N[d_{i,j}|\, f_{i,j}(X), \sigma_{i,j}]$ is a noise model that quantifies the deviation between the predicted and observed data points.

We defined the forward model by inverting the relation between the upper distance bound and observed MIC values [$d_{\mathrm{U}}(\mathrm{MIC})$, Eq. 1]:

$$f_{i,j}(X) = \mathrm{MIC}(d_{i,j}) = \begin{cases} \exp(k \cdot d_{i,j} + n) & \text{if } d_{i,j} \le d_0 \\ 0.6 & \text{if } d_{i,j} > d_0 \end{cases} \quad (2)$$

where $d_0 = d_{\mathrm{U}}(0.6)$. Our choice of a noise model is a lognormal distribution with a flat plateau for MIC values below the upper bound on the experimentally observed MIC values ($\mathrm{MIC}^{\mathrm{obs}}$):

$$P(\mathrm{MIC}_{i,j}^{\mathrm{obs}}|\mathrm{MIC}_{i,j}, X, \sigma_{i,j}) = \begin{cases} \frac{1}{N} & \text{if } \mathrm{MIC}_{i,j}^{\mathrm{obs}} \ge \mathrm{MIC}_{i,j} \\ \frac{1}{M} \frac{1}{\sqrt{2\pi\sigma_{i,j}^2}\, \mathrm{MIC}_{i,j}^{\mathrm{obs}}} \exp\left[-\frac{1}{2\sigma_{i,j}^2} \log^2 \frac{\mathrm{MIC}_{i,j}^{\mathrm{obs}}}{\mathrm{MIC}_{i,j}}\right] & \text{if } \mathrm{MIC}_{i,j}^{\mathrm{obs}} < \mathrm{MIC}_{i,j} \end{cases} \quad (3)$$

Here, $\sigma_{i,j}$ are the noise parameters that can optionally be determined as part of the model, and N and M are normalization factors necessary to make the likelihood continuous. Lognormal noise models have previously been used to describe errors of inherently positive quantities (*95*). For computational efficiency, we used a single σ value for all residue pairs. An uninformative Jeffrey's prior is applied to σ to represent a lack of information on the bounds and distribution of this parameter (*96*).

Finally, a Bayesian term in the scoring function is defined as the negative logarithm of the posterior probability density: $S(M) = -\log p(M|D,I)$. In the Bayesian view, the output model is in fact best equated to the posterior model density that specifies a distribution of alternative single models M with varying probability density, not a single model, although single representative or average models can always be proposed based on the posterior model density.

Calculation of similarity metrics for yeast RNAPII and bacterial RNAP datasets

Steps (i) and (ii) from "Design of the pE-MAP spatial restraints" were repeated for the yeast RNAPII and bacterial RNAP datasets to generate the similarity metrics (MIC values) for these two systems, with the following modifications:

For yeast RNAPII, before imputing missing values in the pE-MAP, any deletion mutants that exhibit missing values with more than 15% of the point mutants were filtered out. This step is part of our pipeline but had no effect on the histone pE-MAP (because this pE-MAP does not contain any deletion mutant with more than 15% values missing). The number of ranked deletion mutants retained at the end of pE-MAP postprocessing was chosen to be 25% of the number of deletions in the original unfiltered pE-MAP (in accordance with the histone pE-MAP processing).

For bacterial RNAP, owing to the very small number of perturbations in this dataset, all the perturbations (instead of the top 25%) were retained for computing point mutant phenotypic profile similarities. In addition, owing to differences in the experimental design for generating the yeast pE-MAPs and the bacterial RNAP CG-MAP, the bacterial RNAP MIC distribution had a ~2-fold higher median and greater spread than the other datasets. Correspondingly, the bacterial RNAP MIC distribution was normalized using linear scaling, decreasing its median to match that of the histone MIC distribution. Importantly, this step was based solely on the MIC distributions, without reliance on any structural information.

Integrative structure determination

Integrative structure determination for each system proceeded through the standard four stages (*3*, *4*, *5*, *8*, *41*, *97*) (Fig. 5 and tables S3, S8, and S9): (i) gathering data, (ii) representing subunits and translating data into spatial restraints, (iii) configurational sampling to produce an ensemble of structures that satisfies the restraints, and (iv) analyzing and validating the ensemble structures and data. The integrative structure modeling protocol [i.e., stages (ii), (iii), and (iv) was scripted using the *Python Modeling Interface* (PMI) package, a library for modeling macromolecular complexes based on our open-source *Integrative Modeling Platform* (IMP) package (*5*), version 2.8 (https://integrativemodeling.org). Files containing the input data, scripts, and output results are available at https://integrativemodeling.org/systems/pemap and the nascent integrative methods benchmarking section of the worldwide Protein Data Bank (wwPDB) PDB-Dev repository for integrative structures and corresponding data (http://pdb-dev.wwpdb.org) (*98*).

(i) Gathering data

To mimic realistic integrative structure determination, we did not rely on the known atomic structures of the subunits in the actual modeled complex [correct docking of exact bound structures based on geometric complementarity is easy, (*99*)]. Instead, we computed comparative models of histones H3 and H4 based on their alignments with structures of the 1TZY (*100*) (89 and 92% sequence identity, respectively), using MODELLER, version 9.21 (*101*). The Cα-atom RMSDs between the crystal structures and comparative models is 2.8 and 5.5 Å for H3 and H4, respectively, corresponding to medium- and low-accuracy comparative models. The second major input information source was a pE-MAP dataset of 308 point mutations in histones H3 and H4 crossed against an array of ~1370 gene deletion alleles, resulting in 946 MIC values above 0.3. Of these, 170 MIC values were converted into distance restraints between H3 and H4 residues (fig. S8 and table S3).

Comparative models of subunits Rpb1 and Rpb2 of yeast RNAPII were computed based on template structures 6GMH (*102*) (54% sequence identity) and 4AYB (*103*) (43% sequence identity), respectively. The Cα RMSD between the crystal structures of subunit Rpb1 and Rpb2 [2E2H (*104*)] and their comparative models are 7.3 and 5.2 Å, respectively. A pE-MAP dataset of 53 single point mutants in yeast RNAPII (44 of which reside in subunits Rpb1 and Rpb2) and a library of ~1200 gene-deletions resulted in 195 MIC values above 0.3. Of these, 123 MIC values were converted into distance restraints (fig. S8 and table S8). In addition, we compared the RNAPII model based on the pE-MAP to a model based on 22 previously published chemical cross-links (*61*).

The structures of subunits RpoB and RpoC of bacterial RNAP were obtained from the x-ray structure of the entire complex (4YG2) (*105*). A CG-MAP of 44 single point mutants of the two subunits and a library of 83 conditions (e.g., treatments with chemicals and temperature shocks) resulted in 109 MIC values above 0.3. Of these, 63 MIC values were converted into distance restraints between the subunits

(fig. S8 and table S9). In addition, we compared the bacterial RNAP model based on the CG-MAP to a model computed based on distance restraints derived from the interfacial contacts predicted using the RaptorX protein complex contact prediction server (*65*, *66*)

(ii) Representing subunits and translating data into spatial restraints

To maximize computational efficiency while avoiding using too coarse a representation, we represented each complex in a multiscale fashion. In particular, the subunits and domains of each complex were coarse-grained using beads of varying sizes representing either a rigid body or a flexible string, based on the available comparative models, as follows (tables S3, S8, and S9). The comparative models were coarse-grained into two representations at different resolutions. First, we identified loop regions of at least eight residues using DSSP (*106*, *107*) and represented them by flexible strings of beads of up to 10 residues each. Second, for the remaining residues each bead corresponded to an individual residue, centered at the position of its Cα atom. With this representation in hand, we next translated the input information into spatial restraints as follows.

The defining and most important restraint for our method is extracted from the pE-MAP and CG-MAP data. The collected pE-MAP and CG-MAP MIC values were used to construct the Bayesian term in the scoring function that restrained the distances spanned by the mutated residues as described above. The pE-MAP restraint was applied to the one residue-per-bead representation for the comparative models as well as to the flexible beads. To improve computational efficiency, we only considered point mutation pairs with MIC values greater than 0.3. This restraint was applied to all three complexes (tables S3, S8, and S9). In addition to the pE-MAP data, integrative modeling can benefit from many other types of input information. Here, we have supplemented the pE-MAP and CG-MAP data by additional simple terms accounting for excluded volume and sequence connectivity. First, the excluded volume restraints were applied to each bead in the one-residue (or the closest) bead representations, using the statistical relationship between the volume and the number of residues that it covered (*4*, *108*). Second, we applied the sequence connectivity restraint, using a harmonic upper bound on the distance between consecutive beads in a subunit, with a threshold distance equal to four times the sum of the radii of the two connected beads. The bead radius was calculated from the excluded volume of the corresponding bead, assuming standard protein density (*4*, *108*). Moreover, we evaluated the utility of the pE-MAP and CG-MAP data by considering two additional types of restraints. First, the 22 previously determined BS3 RNAPII cross-links (*61*) were used to construct a Bayesian term that restrained the distances spanned by the cross-linked residues (30 Å) (*109*, *110*). The cross-link restraints were applied to the one residue-per-bead representation for the comparative models as well as flexible beads, only for RNAPII (table S8). Second, we applied the evolutionary coupling restraints to determine the structures of the RpoB and RpoC subunits of bacterial RNA polymerase. Coupling strengths between residue pairs were obtained using the RaptorX ComplexContact server (http://raptorx.uchicago.edu/ComplexContact/) (*65*, *66*) with default parameters. The top L/50 coupling strengths (fig. S9) with sequence separation of three or greater were converted into distance restraints using a harmonic upper bound on the distances between the residues. The threshold distance was set to 12 Å. This restraint was applied only to a subset of bacterial RNAP modeling instances (table S9).

Configurational sampling to produce an ensemble of structures that satisfy the restraints

The initial positions and orientations of rigid bodies and flexible beads were randomized. The generation of structural models was performed using Replica Exchange Gibbs sampling, based on the Metropolis Monte Carlo (MC) algorithm (*110*, *111*). Each MC step consisted of a series of random transformations (i.e., rotation and translation) of the positions of the flexible beads and rigid bodies. Details about the MC runs for each system are in tables S3, S8, and S9.

Analyzing and validating the ensemble structures and data

Model validation follows four major steps (*3*, *112*): (i) selection of the models for validation, (ii) estimation of sampling precision, (iii) estimation of model precision, and (iv) quantification of the degree to which a model satisfies the information used to compute it. These validations are based on the nascent wwPDB effort on archival, validation, and dissemination of integrative structures (*98*, *113*). We now discuss each one of these validations in turn.

(i) Selection of models for validation: The first step is to objectively define the ensemble of models that will be further analyzed. For each trajectory, we automatically determined the MC step at which all data likelihoods and priors have equilibrated (run equilibration step), and all prior frames are discarded (*114*). Discarding the initial, nonequilibrated steps of each run is helpful because nontypical early configurations (e.g., a random configuration of beads, an extended configuration of beads, and beads far apart from each other) are removed from the statistical sample used for posterior model estimates.

With this ensemble of sampled structures and their corresponding scores in hand, we analyze the data likelihoods and priors. We used HDBSCAN clustering, a hierarchical density-based clustering algorithm, to identify all high-density regions in the likelihoods and priors (*115*). If a single cluster was identified, we consider all the models after discarding the initial steps; otherwise, we consider all models in the clusters that satisfy the input information, within the uncertainty of the data, for further analysis (below).

(ii) Estimation of sampling precision: Next, we estimate the precision at which sampling sampled the selected structures (sampling precision) (*112*); the sampling precision must be at least as high as the precision of the structure ensemble consistent with the input data (model precision). As a proxy for testing the thoroughness of sampling, we performed four sampling convergence tests: (i) verify that the scores of refined structures do not continue to improve as more structures are computed, (ii) confirm that the selected structures in independent sets of sampling runs (sample A and sample B) satisfy the data equally well, (iii) cluster the structural models and determine the sampling precision at which the structural features can be interpreted (fig. S10), and (iv) compare the localization probability density maps for each protein obtained from independent sets of runs. Details about all the tests are described in (*112*). For each modeling instance, the results from the convergence tests are summarized in tables S3, S8, and S9.

(iii) Estimation of model precision: In the third step, the model uncertainty (precision) is estimated. The most explicit description of model uncertainty is provided by the set of all models that are sufficiently consistent with the input information (i.e., the ensemble). Model precision can be quantified by the variability among the models in the ensemble; in the end, the ensemble can be described by one or more representative models and their uncertainties. For example, if the structures in the ensemble are clustered into a single cluster, the model precision is defined as the RMSD between models in the cluster. Importantly, the uncertainty may not be distributed evenly across the ensemble, such that some regions are determined at a higher precision than others.

(iv) Quantification of the degree to which a model satisfies the data used to compute it: An accurate structure needs to satisfy the input information used to compute it; all structures at computed precision that are consistent with the data are provided in the ensemble. A pE-MAP derived restraint is satisfied by a cluster of structures if the corresponding Cα-Cα distance in any of the structures in the cluster is lower than the distance predicted by the MIC value (Eq. 1). A BS3 cross-link restraint is satisfied

by a cluster of structures if the corresponding Cα-Cα distance in any of the structures in the cluster is less than 30 Å (*116*). The remainder of the restraints are harmonic, with a specified standard deviation. Therefore, a restraint is satisfied by a cluster of structures if the restrained distance in any structure in the cluster is violated by less than three standard deviations, specified for the restraint. Tables S3, S8, and S9 show that all models satisfy the input information within its uncertainty.

Benchmark

To benchmark the four-stage protocol described above, we computed the distribution of the accuracy for each structure in the ensemble of solutions obtained by integrative modeling. The accuracy is defined as the mean of Cα RMSD between the x-ray structure and each of the structures in the ensemble. The PDB accession code and accuracies for each modeling instance are summarized in tables S3, S8, and S9.

To assess the information content of the histone pE-MAP, we computed the models of the H3-H4 complex based on random subsets of the data. To this end, from the dataset of computed MIC values for pairs of mutated residues, we performed three independent random selections of 80, 60, 40, and 20% of the data each. As expected, the more pE-MAP data that are used, the more accurate and precise are the models (Fig. 6C).

Finally, as another test, we computed the model based on datasets with randomly shuffled MIC values for the same pE-MAP or CG-MAP residue pairs, for each of the complexes.

Estimation of the number of mutations per protein

To estimate the suggested number of mutated positions per protein for integrative structure determination, we computed the number of mutations that would result in four or more MIC values above a 0.4, 0.45, 0.5, or 0.55 threshold. Based on our scoring function, MIC values above these thresholds will result in distance restraints with an upper distance bound in the 12- to 34-Å range. These distances are comparable to the upper distance bounds used for chemical cross-links (e.g., DSS, DSSO, EDC). A previous systematic study established that at least four chemical cross-links are needed to determine the binding mode of protein dimers if the subunit structures are known (e.g., from x-ray, NMR, or comparative models) (*110*). In general, adding more chemical cross-links does not further improve the accuracy, although it increases the precision of the resulting ensemble. By analogy, we estimate that, for systems in which the structures of the components are known, a good number of mutations per protein is 35 to 40 (fig. S5A). This data can be used as a guideline to decide on the number of mutations to use for generating a pE-MAP or CG-MAP. Importantly, this estimate is an upper bound on the number of mutations, and in many cases, the number might be smaller for the following two reasons. First, this estimation was done assuming protein-wide mutations of residues, often to alanine. In practice, the number of necessary mutations can be reduced by specifically designing point mutations that target surface residues and/or residues known to be functionally important, and by choosing substitutions likely to give rise to functional perturbations. In general, we did not find a correlation between the secondary structure of the residue pairs and their associated MIC value (fig. S5B). Second, this estimation only relies on the residue pairs with high MIC values. In contrast to chemical cross-links, the upper distance bound of pE-MAP derived restraints are obtained from the statistical association between the MIC values and distance between residues. Consequently, residue pairs with low MIC values still carry structural information, even if at low resolution. Consistent with these considerations, the RNAPII dataset contains only 31 and 9 mutated residues for Rpb1 and Rpb2, respectively. Similarly, the bacterial RNAP dataset contains 23 and 15 mutated residues for Rpob and Rpoc, respectively.

Docking

To assess the relative value of pE-MAP restraints for structure determination, we computed the structures of the H3-H4 and RNAPII complexes by molecular docking. Specifically, we followed an integrative docking protocol (*117*) using the rigid-body docking program PatchDock (*39*). In each case, we used the same comparative models and rigid-body definitions used for integrative modeling (figs. S4 and S11) and default parameter values (figs. S4 and S11).

Visualizations

The pE-MAP was hierarchically clustered in both histone mutant and gene deletion dimensions using Cluster 3.0 (*118*) and displayed using Java Treeview (*119*) (Fig. 3). Images highlighting histone residues in context of the nucleosome structure (PDB 1ID3 or its modified version in data S2) were created using ChimeraX (Figs. 3A and 7C; and figs. S1D and S6) (*48*).

Distance of clustered pE-MAP profiles

First, all histone alleles affecting residues not included in the structural reference (PDB 1ID3, H3A1-H3K37 and H4S1-H4R17) were removed and the remaining data (n = 222) clustered hierarchically using Cluster 3.0 (*118*). For each node of the clustergram, the mean distance among member residues was calculated and plotted versus the normalized branch length (where the first node is set to branch length = 0 and the last node to branch length = 1) of the respective node (fig. S2A, red dots, and random distribution plotted in black).

Generation of the correlation map

Pearson correlation coefficients were computed for each of the 350 H3 and H4 mutants against all genes and alleles (rows) in a merged map of previously published genetic interaction data [dataset S4 from (*45*)]. If the overlap between a histone mutant and a S-score vector from the merged map was <150 scores, the resulting correlation was not considered (i.e., replaced by "NaN") (data S3). Pearson correlation coefficients were chosen over MIC for this analysis because we found Pearson correlation more robust than MIC when many missing values were present.

Structural mapping of genetic interactions—stEMAP app

The hierarchically clustered pE-MAP data was imported into Cytoscape (*47*), creating an initial network, and then linked to a modified version of the nucleosome structure 1ID3 (data S2) using the stEMAP app, developed to facilitate interactive exploration of the pE-MAP. The original nucleosome structure was modified by adding the N-terminal disordered regions of histone H3 and H4 and manually positioning them for clarity. The linking proceeds as follows: First, the structure is opened in ChimeraX (*48*) by structureVizX (*120*) and positioned in response to commands from the stEMAP app. Then, a residue interaction network (RIN) is created by the structureVizX app where nodes are positioned to reflect the nucleosome structure through the help of the RINalyzer app (*121*). Finally, the RIN network and the network created by the original cluster files are merged, and edges are drawn between genes and residues with interactions passing a user-defined threshold (fig. S6A). All of the preceding steps happen automatically through the stEMAP app interface, which takes as input any given PDB file and a short user-defined JSON configuration file defining interaction thresholds (here: correlation > 0.2), colors of edges (here: color-gradient from white to red for positive correlations), and display style of the structure in ChimeraX.

Selection of individual genes triggers the interacting residues to be selected, and, in the ChimeraX window, those residues are shown as space-filling atoms, which are colored according to the edge colors. When multiple genes are selected (e.g., genes belonging to the same complex), there might be multiple edges connecting an individual residue. In this case, the color reflects the significance and consistency of the interactions (see below). To assist in interpretation and interactive exploration of complex data sets (i) colors are quantized into 10 bins, five positive and five negative; (ii) a heatmap is presented that shows only the

values for the selected genes and their interacting residues; (iii) sets of genes belonging to a complex can be selected using the setsApp (*122*); and (iv) a slider provides a filter to restrict the selection to only those mutations with a minimum number of interactions.

To determine if a gene set is connected to a given residue, the stEMAP app calculates the median Pearson correlation coefficient across all genes of the gene set against all different mutations at that residue. If this median correlation is above the threshold of 0.2 (defined in the JSON file), the respective residue is colored according to the median. To instead determine if a gene set is connected to an individual mutation, the same method is used, except the median correlation coefficient is now calculated across all genes of the gene set against the single given mutation (instead of all mutations of the residue).

ROC curves

Only library deletion mutants that exist in both this study and the two previously published E-MAP datasets [Braberg *et al.* (*15*) and Collins *et al.* (*29*)] were included (n = 389) in this analysis. Based on their pE-MAP profiles, Pearson correlation coefficients were calculated for all pairwise combinations of these 389 deletion mutants. To determine the power of these correlations to predict physical interactions between encoded proteins, an ROC curve was computed, where a physical interaction between proteins was defined if their PE score is greater than 2 (*69*). From the Collins *et al.* E-MAP, query strain profiles with more missing data than the sparsest histone mutant were removed, as were query mutants that also existed in the library mutant set. Because the Braberg *et al.* pE-MAP only includes 53 query mutants (rows), we used subsets of 53 query mutants each for the histone and Collins *et al.* E-MAPs when generating their ROC curves, to make all three systems comparable. To this end, for the Collins *et al.* E-MAP and histone pE-MAP, 53 query mutant profiles were randomly selected 1000 times, and a ROC curve was generated for each run. The median areas under the ROC curves (AROCs) and corresponding ROC curves are reported together with the ROC curve of the pE-MAP from Braberg *et al.* in Fig. 2F.

RNA-seq expression analysis

Ten ml of overnight cultures of 29 histone mutant strains (table S2) were harvested in mid-log phase [optical density at 600 nm (OD_{600}) ≈ 1.0] and washed with DEPC-ddH_2O. RNA was extracted with hot acidic phenol as described previously (*123*). RNA-seq libraries were generated using the QuantSeq 3′ mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen). Single-end, 50-base reads were sequenced using an Illumina HiSeq 4000 sequencer. Reads were filtered for quality and aligned to the yeast genome using tophat (*124*). Nonunique reads and reads mapping to ribosomal RNA were removed before analysis. Transcript counts were extracted using htseq-count (*125*), and differential expression was measured using the Dseq2 package in R (*126*).

Identification of functional links between H3 and H4 mutants and biological processes

The correlation map (data S3) was used as the basis for this analysis. First, a curated annotation of all genes in the correlation map relevant to nuclear function was devised. Biological process definitions for genes in nuclear processes were assigned manually based on literature and annotations from previous genetic interaction maps (*29*, *127*, *128*) (table S5). To identify links between H3 and H4 residues and nuclear processes that were highly correlated, we used a one-sided Mann-Whitney U test to compare the correlation distribution between the mutants of each H3 and H4 residue and the members of each process to (i) the correlations between the same H3 and H4 mutants and all genes not in that process and to (ii) the correlations between the same process and all other H3 and H4 mutants. The highest p value of the comparison to (i) or (ii) was recorded. False discovery rates (FDRs) for the links were then computed using the method of Benjamini and Hochberg (*129*) and are reported in table S5. The most significant links with FDR $<10^{-6}$ were used for follow-ups.

Spontaneous mutation frequency

Cells were grown to saturation and then plated on yeast extract peptone dextrose (YEPD)– and 5-fluoroorotic acid (5-FOA)–supplemented media. Mutants growing on 5-FOA were counted only after confirming that colonies growing on YEPD for all the strains were of equal size. The assay was repeated three times independently (table S6).

MS quantification of H3K56ac levels

Sample preparation

Histone mutant cultures (WT, H3R63K, and H4R36K) were harvested in mid-log phase (OD_{600} ≈ 1.0) using a 250-mm ceramic filter funnel and 30-μm nitrocellulose membranes connected to high continuous wall suction. Yeast were removed from the nitrocellulose membrane and flash frozen for storage or used immediately for protein extraction. Per gram of yeast pellet, 3 ml of Yeast-Protein Extract Reagent (Y-PER; ThermoFisher Scientific) with added protease inhibitors (cOmplete Sigma-Aldrich, one tablet per 50 ml), phosphatase inhibitors (PhosSTOP Sigma-Aldrich; one tablet per 50 ml), histone deacetylase inhibitors (sodium butyrate 100 mM and nicotinamide 100 mM), and β-mercaptoethanol (15 mM) were added. The suspension was mixed on a gyrator at 4°C for 30 min and centrifuged. Pellets were resuspended in fresh Y-PER medium, and extraction was repeated two additional times for a total of three extractions. Pellets were sequentially washed twice with 3 ml ddH_2O per gram of yeast. Histone extraction was performed in the presence of 2.5 ml of 8M urea/0.4N sulfuric acid per gram of yeast protein pellets, incubated for 1 hour, centrifuged, and supernatants collected. Proteins were precipitated using a methanol-chloroform precipitation as previously described (*130*). Extracted proteins were trypsin digested; desalted and acetylated peptides were enriched as previously described (*131*).

Generation of selected reaction monitoring (SRM) assays for acetylation sites

Peptide mixtures (obtained from ThermoFisher) were analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) on a Thermo Scientific Orbitrap Fusion mass spectrometry system equipped with a Proxeon Easy nLC 1200 ultra-high-pressure LC and autosampler system. Samples were injected onto a C18 column (25 cm × 75 μm I.D. packed with ReproSil Pur C18 AQ 1.9-μm particles) in 0.1% formic acid and then separated with a 60-min gradient from 5 to 40% buffer B (90% ACN/10% water/0.1% formic acid) at a flow rate of 300 nl/min. The mass spectrometer collected data in a data-dependent fashion, collecting one full scan in the Orbitrap followed by collision-induced dissociation MS/MS scans in the dual linear ion trap for the 20 most intense peaks from the full scan. Dynamic exclusion was enabled for 30 s with a repeat count of 1. Charge state screening was used to reject analysis of singly charged species or species for which a charge could not be assigned. The raw data was matched to protein sequences using the MaxQuant algorithm (version 1.5.2.8) (*132*). Data were searched against a database containing SwissProt Human protein sequences concatenated to a decoy database where each protein sequence was randomized in order to estimate the FDR. Variable modifications were allowed for methionine oxidation and protein N-terminal acetylation and lysine acetylation. A fixed modification was indicated for cysteine carbamidomethylation. Full trypsin specificity was required. The first search was performed with a mass accuracy of ±20 parts per million and the main search was performed with a mass accuracy of ±4.5 parts per million. A maximum of five modifications were allowed per peptide. A maximum of two missed cleavages were allowed. The maximum charge allowed was 7+. Individual peptide mass tolerances were allowed. For MS/MS matching, a mass tolerance of 0.8 Da was allowed, and the top eight peaks per 100 Da were analyzed. MS/MS matching was allowed for higher charge states and water and ammonia

loss events. The data were filtered to obtain a peptide, protein, and site-level FDR of 0.01. The minimum peptide length was seven amino acids. Selected reaction monitoring (SRM) assays were generated for selected acetylation sites. SRM assay generation was performed using Skyline (*133*). For all targeted proteins, proteotypic peptides and optimal transitions for identification and quantification were selected based on a spectral library generated from the shotgun MS experiments. The Skyline spectral library was used to extract optimal coordinates for the SRM assays, for example, peptide fragments and peptide retention times. For each peptide, the five best SRM transitions were selected based on intensity and peak shape.

Acquisition and quantification of acetylation sites by SRM

Digested peptide mixtures were analyzed by LC-SRM on a Thermo Scientific TSQ Quantiva MS system equipped with a Proxeon Easy nLC 1200 ultra-high-pressure LC and autosampler system. Samples were injected onto a C18 column (25 cm × 75 μm I.D. packed with ReproSil Pur C18 AQ 1.9-μm particles) in 0.1% formic acid and then separated with a 60-min gradient from 5 to 40% buffer B (90% ACN/10% water/0.1% formic acid) at a flow rate of 300 nl/min. SRM acquisition was performed operating Q1 and Q3 at 0.7-unit-mass resolution. For each peptide, the best five transitions were monitored in a scheduled fashion with a retention time window of 5 min and a cycle time fixed to 2 s. Argon was used as the collision gas at a nominal pressure of 1.5 mTorr. Collision energies were calculated by, $CE = 0.0348 \times (m/z) + 0.4551$ and $CE = 0.0271 \times (m/z) + 1.5910$ (*CE*, collision energy and m/z, mass to charge ratio) for doubly and triply charged precursor ions, respectively. SRM data was processed using Skyline (*133*). Protein significance analysis was performed using MSstats (*134*). Normalization of the intensities across samples was performed using the acetylated peptides H3K9_H3K14 (peptide containing both acetylation sites), H3K23 and H3K14 as global standards, which did not show any change across the mutants. Log_2 fold changes were calculated from three independent runs and plotted (fig. S7C and table S6).

Cryptic transcription—qPCR

Total RNA was extracted from 10 OD_{600} units of mid-log phase cells (WT and respective mutant strains) using hot acid phenol-chloroform extraction method as described. Ten μg of total RNA was DNAse I treated (Promega) followed by purification using an RNeasy minikit (Qiagen). One μg of DNAse I–treated total RNA was used to synthesize cDNA using SuperScript III first strand synthesis system (Life Technologies) and random hexamer primers. cDNA was diluted 1:25 before amplification by PCR using primers designed for the 5′ and the 3′ ends of the *STE11* gene. qPCR was performed using SYBR green (Biorad) as described previously (*135*). Relative change in the transcript levels were estimated using the $\Delta\Delta C_t$ method described in (*136*) and were normalized to *ACT1* transcript (table S7). Primer sequences are available upon request.

Western blotting

Whole-yeast-cell lysates were prepared using trichloroacetic acid (TCA) lysis as described previously (*137*). Lysates were subjected to immunoblotting according to standard procedures, and proteins were detected using ECL Prime (Amersham Biosciences). Membranes were probed with αH3K36me3 antibody purchased from (Abcam, catalog no. 9050). Glyceraldehyde-phosphate dehydrogenase (GAPDH) was used as a loading control and detected using an antibody purchased from Sigma (catalog no. A9521).

ATAC-seq

Yeast cells (2.5×10^6) were grown to mid-log phase, pelleted, washed with SB-buffer (1.4 M Sorbitol, 40 mM HEPES-KOH pH 7.5, 0.5 mM MgCl2), resuspended in 200 μl SB buffer + 10 mM DTT with 10 μl of 10 mg/ml 100T zymolyase (MP Biomedicals) solution and incubated for 5 min at 30°C. Spheroblasted cells were washed with SB-buffer and incubated for 15 min at 37°C in 25 μl of transposase solution (12.5 μl 2x TD buffer, 1.25 μl Nextera enzyme, 11.25 μl water). DNA was purified (Qiagen MinElute DNA Purification Kit), amplified, and barcoded by PCR. Purified PCR products were sequenced using an Illumina HiSeq 4000 sequencer. Sequence reads were trimmed and aligned to the genome of *Saccharomyces cerevisiae* (version SacCer3 from hgdownload.cse.ucsc.edu/downloads.html), and reads with a length <100 basepairs (bp) were removed. Replicates belonging to an allele (WT, H3K36A, H3K122A, *set2*Δ) were merged and normalized to the smallest read number. For visualization of *STE11* read coverage using the IGV genome browser (*138*) (fig. S7G), each track was scaled linearly so that the largest peak in the displayed window is the same height for all tracks. Count files were generated with "featureCounts v1.5.3" (*139*).

Gene body plots (fig. S7H) were generated as follows: First, counts from genes reported to be targets of cryptic transcription ($n = 11$; *FLO8*, *AVO1*, *LCB5*, *SMC3*, *SPB4*, *APM2*, *DDC1*, *SYF1*, *OMS1*, *PUS4*, *STE11*), as well as counts 400 bp up- and downstream of the respective gene bodies, were extracted. Second, up- and downstream regions were split into 50 bins of equal size (8 bp), whereas the gene body was split into 300 equal bins, resulting in 400 bins for each gene in each tested strain (WT, *set2*Δ, H3K36A and H3K122A). Third, for each of the 400 bins, the average for the 11 target genes was calculated. Fourth, each mutant allele was then scaled linearly so that the first bin (i.e., 400 bp upstream of the gene body start) was equal to that of WT. Finally, the WT counts were subtracted from the mutant counts for each bin.

REFERENCES AND NOTES

1. F. Alber, F. Förster, D. Korkin, M. Topf, A. Sali, Integrating diverse data for structure determination of macromolecular assemblies. *Annu. Rev. Biochem.* **77**, 443–477 (2008). doi: 10.1146/annurev.biochem.77.060407.135530; pmid: 18318657
2. F. Herzog *et al*., Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **337**, 1348–1352 (2012). doi: 10.1126/science.1221483; pmid: 22984071
3. M. P. Rout, A. Sali, Principles for integrative structural biology studies. *Cell* **177**, 1384–1403 (2019). doi: 10.1016/j.cell.2019.05.016; pmid: 31150619
4. F. Alber *et al*., Determining the architectures of macromolecular assemblies. *Nature* **450**, 683–694 (2007). doi: 10.1038/nature06404; pmid: 18046405
5. D. Russel *et al*., Putting the pieces together: Integrative modeling platform software for structure determination of macromolecular assemblies. *PLOS Biol.* **10**, e1001244 (2012). doi: 10.1371/journal.pbio.1001244; pmid: 22272186
6. A. B. Ward, A. Sali, I. A. Wilson, Integrative structural biology. *Science* **339**, 913–915 (2013). doi: 10.1126/science.1228565; pmid: 23430643
7. F. Alber *et al*., The molecular architecture of the nuclear pore complex. *Nature* **450**, 695–701 (2007). doi: 10.1038/nature06405; pmid: 18046406
8. K. Lasker *et al*., Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1380–1387 (2012). doi: 10.1073/pnas.1120559109; pmid: 22307589
9. A. Loquet *et al*., Atomic model of the type III secretion system needle. *Nature* **486**, 276–279 (2012). doi: 10.1038/nature11079; pmid: 22699623
10. Z. Duan *et al*., A three-dimensional model of the yeast genome. *Nature* **465**, 363–367 (2010). doi: 10.1038/nature08973; pmid: 20436457
11. J. M. Plitzko, B. Schuler, P. Selenko, Structural biology outside the box—inside the cell. *Curr. Opin. Struct. Biol.* **46**, 110–121 (2017). doi: 10.1016/j.sbi.2017.06.007; pmid: 28735108
12. M. Dimura *et al*., Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **40**, 163–185 (2016). doi: 10.1016/j.sbi.2016.11.012; pmid: 27939973
13. S. R. Collins, A. Roguev, N. J. Krogan, Quantitative genetic interaction mapping using the E-MAP approach. *Methods Enzymol.* **470**, 205–231 (2010). doi: 10.1016/S0076-6879(10)70009-4; pmid: 20946812
14. P. Beltrao, G. Cagney, N. J. Krogan, Quantitative genetic interactions reveal biological modularity. *Cell* **141**, 739–745 (2010). doi: 10.1016/j.cell.2010.05.019; pmid: 20510918
15. H. Braberg *et al*., From structure to systems: High-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* **154**, 775–788 (2013). doi: 10.1016/j.cell.2013.07.033; pmid: 23932120
16. H. Braberg, E. A. Moehle, M. Shales, C. Guthrie, N. J. Krogan, Genetic interaction analysis of point mutations enables interrogation of gene function at a residue-level resolution: Exploring the applications of high-resolution genetic interaction mapping of point mutations. *BioEssays* **36**, 706–713 (2014). doi: 10.1002/bies.201400044; pmid: 24842270
17. N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, Protein sectors: Evolutionary units of three-dimensional structure. *Cell* **138**, 774–786 (2009). doi: 10.1016/j.cell.2009.07.038; pmid: 19703402
18. D. S. Marks *et al*., Protein 3D structure computed from evolutionary sequence variation. *PLOS ONE* **6**, e28766 (2011). doi: 10.1371/journal.pone.0028766; pmid: 22163331
19. G. Diss, B. Lehner, The genetic landscape of a physical interaction. *eLife* **7**, e32472 (2018). doi: 10.7554/eLife.32472; pmid: 29638215
20. A. L. Shiver, H. Osadnik, J. M. Peters, R. A. Mooney, P. I. Wu, J. C. Hu, R. Landick, K. C. Huang, C. A. Gross,

Chemical-genetic interrogation of RNA polymerase mutants reveals structure-function relationships and physiological tradeoffs. bioRxiv 2020.2006.2016.155770 [Preprint]. 17 June 2020; .doi: 10.1101/2020.06.16.155770
21. H. Huang, S. Lin, B. A. Garcia, Y. Zhao, Quantitative proteomic analysis of histone modifications. *Chem. Rev.* **115**, 2376–2418 (2015). doi: 10.1021/cr500491u; pmid: 25688442
22. J. Dai *et al*., Probing nucleosome function: A highly versatile library of synthetic histone H3 and H4 mutants. *Cell* **134**, 1066–1078 (2008). doi: 10.1016/j.cell.2008.07.019; pmid: 18805098
23. S. Jiang *et al*., Construction of comprehensive dosage-matching core histone mutant libraries for *Saccharomyces cerevisiae*. *Genetics* **207**, 1263–1273 (2017). pmid: 29084817
24. J. E. Brownell *et al*., Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–851 (1996). doi: 10.1016/S0092-8674(00)81063-6; pmid: 8601308
25. L. K. Durrin, R. K. Mann, P. S. Kayne, M. Grunstein, Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65**, 1023–1031 (1991). doi: 10.1016/0092-8674(91)90554-C; pmid: 2044150
26. H. Braberg *et al*., Quantitative analysis of triple-mutant genetic interactions. *Nat. Protoc.* **9**, 1867–1881 (2014). doi: 10.1038/nprot.2014.127; pmid: 25010907
27. J. E. Haber *et al*., Systematic triple-mutant analysis uncovers functional connectivity between pathways involved in chromosome regulation. *Cell Rep.* **3**, 2168–2178 (2013). doi: 10.1016/j.celrep.2013.05.007; pmid: 23746449
28. S. R. Collins, M. Schuldiner, N. J. Krogan, J. S. Weissman, A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol.* **7**, R63 (2006). doi: 10.1186/gb-2006-7-7-r63; pmid: 16859555
29. S. R. Collins *et al*., Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**, 806–810 (2007). doi: 10.1038/nature05649; pmid: 17314980
30. T. Miller *et al*., COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12902–12907 (2001). doi: 10.1073/pnas.231473398; pmid: 11687631
31. A. Roguev *et al*., The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* **20**, 7137–7148 (2001). doi: 10.1093/emboj/20.24.7137; pmid: 11742990
32. G. Mizuguchi *et al*., ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343–348 (2004). doi: 10.1126/science.1090701; pmid: 14645854
33. M. J. Carrozza *et al*., Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005). doi: 10.1016/j.cell.2005.10.023; pmid: 16286007
34. S. Venkatesh *et al*., Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* **489**, 452–455 (2012). doi: 10.1038/nature11326; pmid: 22914091
35. M. C. Keogh *et al*., Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593–605 (2005). doi: 10.1016/j.cell.2005.10.025; pmid: 16286008
36. D. N. Reshef *et al*., Detecting novel associations in large data sets. *Science* **334**, 1518–1524 (2011). doi: 10.1126/science.1205438; pmid: 22174245
37. D. Albanese *et al*., Minerva and minepy: A C engine for the MINE suite and its R, Python and MATLAB wrappers. *Bioinformatics* **29**, 407–408 (2013). doi: 10.1093/bioinformatics/bts707; pmid: 23242262
38. C. L. White, R. K. Suto, K. Luger, Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207–5218 (2001). doi: 10.1093/emboj/20.18.5207; pmid: 11566884
39. D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson, PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005). doi: 10.1093/nar/gki481; pmid: 15980490
40. D. Baker, A. Sali, Protein structure prediction and structural genomics. *Science* **294**, 93–96 (2001). doi: 10.1126/science.1065659; pmid: 11588250
41. S. J. Kim *et al*., Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**, 475–482 (2018). doi: 10.1038/nature26003; pmid: 29539637
42. C. Gutierrez *et al*., Structural dynamics of the human COP9 signalosome revealed by cross-linking mass spectrometry and integrative modeling. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4088–4098 (2020). doi: 10.1073/pnas.1915542117; pmid: 32034103
43. K. S. Molnar *et al*., Cys-scanning disulfide crosslinking and bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **22**, 1239–1251 (2014). doi: 10.1016/j.str.2014.04.019; pmid: 25087511
44. Y. Kwon *et al*., Structural basis of CD4 downregulation by HIV-1 Nef. *Nat. Struct. Mol. Biol.* **27**, 822–828 (2020). doi: 10.1038/s41594-020-0463-z; pmid: 32719457
45. C. J. Ryan *et al*., Hierarchical modularity and the evolution of genetic interactomes across species. *Mol. Cell* **46**, 691–704 (2012). doi: 10.1016/j.molcel.2012.05.028; pmid: 22681890
46. J. Wysocka *et al*., A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 (2006). doi: 10.1038/nature04815; pmid: 16728976
47. P. Shannon *et al*., Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003). doi: 10.1101/gr.1239303; pmid: 14597658
48. T. D. Goddard *et al*., UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi: 10.1002/pro.3235; pmid: 28710774
49. D. Schneidman-Duhovny, R. Pellarin, A. Sali, Uncertainty in integrative structural modeling. *Curr. Opin. Struct. Biol.* **28**, 96–104 (2014). doi: 10.1016/j.sbi.2014.08.001; pmid: 25173450
50. N. L. Maas, K. M. Miller, L. G. DeFazio, D. P. Toczyski, Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4. *Mol. Cell* **23**, 109–119 (2006). doi: 10.1016/j.molcel.2006.06.006; pmid: 16818235
51. J. Han *et al*., A Cul4 E3 ubiquitin ligase regulates histone hand-off during nucleosome assembly. *Cell* **155**, 817–829 (2013). doi: 10.1016/j.cell.2013.10.014; pmid: 24209620
52. T. Tsubota *et al*., Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol. Cell* **25**, 703–712 (2007). doi: 10.1016/j.molcel.2007.02.006; pmid: 17320445
53. E. M. Hyland *et al*., Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **25**, 10060–10070 (2005). doi: 10.1128/MCB.25.22.10060-10070.2005; pmid: 16260619
54. H. Masumoto, D. Hawke, R. Kobayashi, A. Verreault, A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**, 294–298 (2005). doi: 10.1038/nature03714; pmid: 16015338
55. M. Giannattasio, F. Lazzaro, P. Plevani, M. Muzi-Falconi, The DNA damage checkpoint response requires histone H2B ubiquitination by Rad6-Bre1 and H3 methylation by Dot1. *J. Biol. Chem.* **280**, 9879–9886 (2005). doi: 10.1074/jbc.M414453200; pmid: 15632126
56. I. Celic *et al*., The sirtuins Hst3 and Hst4p preserve genome integrity by controlling histone H3 lysine 56 deacetylation. *Curr. Biol.* **16**, 1280–1289 (2006). doi: 10.1016/j.cub.2006.06.023; pmid: 16815704
57. I. Celic, A. Verreault, J. D. Boeke, Histone H3 K56 hyperacetylation perturbs replisomes and causes DNA damage. *Genetics* **179**, 1769–1784 (2008). doi: 10.1534/genetics.108.088914; pmid: 18579506
58. H. N. Du, S. D. Briggs, A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *J. Biol. Chem.* **285**, 11704–11713 (2010). doi: 10.1074/jbc.M109.085043; pmid: 20139424
59. S. Chen *et al*., Structure-function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. *Genes Dev.* **29**, 1326–1340 (2015). doi: 10.1101/gad.261115.115; pmid: 26109053
60. H. G. Tran, D. J. Steger, V. R. Iyer, A. D. Johnson, The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO J.* **19**, 2323–2331 (2000). doi: 10.1093/emboj/19.10.2323; pmid: 10811623
61. Z. A. Chen *et al*., Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717–726 (2010). doi: 10.1038/emboj.2009.401; pmid: 20094031
62. S. Ovchinnikov, H. Kamisetty, D. Baker, Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* **3**, e02030 (2014). doi: 10.7554/eLife.02030; pmid: 24842992
63. Q. Cong, I. Anishchenko, S. Ovchinnikov, D. Baker, Protein interaction networks revealed by proteome coevolution. *Science* **365**, 185–189 (2019). pmid: 31296772
64. T. Gueudré, C. Baldassi, M. Zamparo, M. Weigt, A. Pagnani, Simultaneous identification of specifically interacting paralogs and interprotein contacts by direct coupling analysis. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12186–12191 (2016). doi: 10.1073/pnas.1607570113; pmid: 27729520
65. S. Wang, S. Sun, Z. Li, R. Zhang, J. Xu, Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLOS Comput. Biol.* **13**, e1005324 (2017). doi: 10.1371/journal.pcbi.1005324; pmid: 28056090
66. H. Zeng *et al*., ComplexContact: A web server for inter-protein contact prediction using deep learning. *Nucleic Acids Res.* **46**, W432–W437 (2018). doi: 10.1093/nar/gky420; pmid: 29790960
67. S. Wang, S. Sun, J. Xu, Analysis of deep learning methods for blind protein contact prediction in CASP12. *Proteins* **86** (suppl. 1), 67–77 (2018). doi: 10.1002/prot.25377; pmid: 28845538
68. K. R. Roy *et al*., Multiplexed precision genome editing with trackable genomic barcodes in yeast. *Nat. Biotechnol.* **36**, 512–520 (2018). doi: 10.1038/nbt.4137; pmid: 29734294
69. S. R. Collins *et al*., Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **6**, 439–450 (2007). doi: 10.1074/mcp.M600381-MCP200; pmid: 17200106
70. J. M. Schmiedel, B. Lehner, Determining protein structures using deep mutagenesis. *Nat. Genet.* **51**, 1177–1186 (2019). doi: 10.1038/s41588-019-0431-x; pmid: 31209395
71. N. J. Rollins *et al*., Inferring protein 3D structure from deep mutation scans. *Nat. Genet.* **51**, 1170–1176 (2019). doi: 10.1038/s41588-019-0432-9; pmid: 31209393
72. R. W. Newberry, J. T. Leong, E. D. Chow, M. Kampmann, W. F. DeGrado, Deep mutational scanning reveals the structural basis for α-synuclein activity. *Nat. Chem. Biol.* **16**, 653–659 (2020). doi: 10.1038/s41589-020-0480-6; pmid: 32152544
73. E. Eyal, R. Najmanovich, M. Edelman, V. Sobolev, Protein side-chain rearrangement in regions of point mutations. *Proteins* **50**, 272–282 (2003). doi: 10.1002/prot.10276; pmid: 12486721
74. R. Sasidharan, C. Chothia, The selection of acceptable protein mutations. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10080–10085 (2007). doi: 10.1073/pnas.0703737104; pmid: 17540730
75. J. Schaarschmidt, B. Monastyrskyy, A. Kryshtafovych, A. M. J. J. Bonvin, Assessment of contact predictions in CASP12: Co-evolution and deep learning coming of age. *Proteins* **86** (suppl. 1), 51–66 (2018). doi: 10.1002/prot.25407; pmid: 29071738
76. S. Ovchinnikov, H. Park, D. E. Kim, F. DiMaio, D. Baker, Protein structure prediction using Rosetta in CASP12. *Proteins* **86** (suppl. 1), 113–121 (2018). doi: 10.1002/prot.25390; pmid: 28940798
77. P. J. Robinson *et al*., Molecular architecture of the yeast Mediator complex. *eLife* **4**, e08719 (2015). doi: 10.7554/eLife.08719; pmid: 26402457
78. J. Luo *et al*., Architecture of the human and yeast general transcription and DNA repair factor TFIIH. *Mol. Cell* **59**, 794–806 (2015). doi: 10.1016/j.molcel.2015.07.016; pmid: 26340423
79. M. Jinek *et al*., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi: 10.1126/science.1225829; pmid: 22745249
80. J. P. Shen *et al*., Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. *Nat. Methods* **14**, 573–576 (2017). doi: 10.1038/nmeth.4225; pmid: 28319113
81. D. Du *et al*., Genetic interaction mapping in mammalian cells using CRISPR interference. *Nat. Methods* **14**, 577–580 (2017). doi: 10.1038/nmeth.4286; pmid: 28481362
82. L. Ma *et al*., CRISPR-Cas9-mediated saturated mutagenesis screen predicts clinical drug resistance with improved accuracy. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11751–11756 (2017). doi: 10.1073/pnas.1708268114; pmid: 29078326
83. A. V. Anzalone *et al*., Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019). doi: 10.1038/s41586-019-1711-4; pmid: 31634902

84. D. E. Gordon *et al*., A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020). doi: 10.1038/s41586-020-2286-9; pmid: 32353859
85. M. Eckhardt, J. F. Hultquist, R. M. Kaake, R. Hüttenhain, N. J. Krogan, A systems approach to infectious disease. *Nat. Rev. Genet.* **21**, 339–354 (2020). doi: 10.1038/s41576-020-0212-5; pmid: 32060427
86. D. E. Gordon *et al*., Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, eabe9403 (2020). doi: 10.1126/science.abe9403; pmid: 33060197
87. D. E. Gordon *et al*., A quantitative genetic interaction map of HIV infection. *Mol. Cell* **78**, 197–209.e7 (2020). doi: 10.1016/j.molcel.2020.02.004; pmid: 32084337
88. S. R. McGuffee, A. H. Elcock, Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLOS Comput. Biol.* **6**, e1000694 (2010). doi: 10.1371/journal.pcbi.1000694; pmid: 20221255
89. S. Takamori *et al*., Molecular anatomy of a trafficking organelle. *Cell* **127**, 831–846 (2006). doi: 10.1016/j.cell.2006.10.030; pmid: 17110340
90. B. G. Wilhelm *et al*., Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* **344**, 1023–1028 (2014). doi: 10.1126/science.1252884; pmid: 24876496
91. J. Singla *et al*., Opportunities and challenges in building a spatiotemporal multi-scale model of the human pancreatic β cell. *Cell* **173**, 11–19 (2018). doi: 10.1016/j.cell.2018.03.014; pmid: 29570991
92. P. J. Thul *et al*., A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017). doi: 10.1126/science.aal3321; pmid: 28495876
93. M. Schuldiner, S. R. Collins, J. S. Weissman, N. J. Krogan, Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. *Methods* **40**, 344–352 (2006). doi: 10.1016/j.ymeth.2006.07.034; pmid: 17101447
94. W. Rieping, M. Habeck, M. Nilges, Inferential structure determination. *Science* **309**, 303–306 (2005). doi: 10.1126/science.1110428; pmid: 16002620
95. W. Rieping, M. Habeck, M. Nilges, Modeling errors in NOE data with a log-normal distribution improves the quality of NMR structures. *J. Am. Chem. Soc.* **127**, 16026–16027 (2005). doi: 10.1021/ja055092c; pmid: 16287280
96. H. Jeffreys, An invariant form for the prior probability in estimation problems. *Proc. R. Soc. London Ser. A* **186**, 453–461 (1946). doi: 10.1098/rspa.1946.0056; pmid: 20998741
97. A. Sali *et al*., Outcome of the first wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **23**, 1156–1167 (2015). doi: 10.1016/j.str.2015.05.013; pmid: 26095030
98. S. K. Burley *et al*., PDB-Dev: A prototype system for depositing integrative/hybrid structural models. *Structure* **25**, 1317–1318 (2017). doi: 10.1016/j.str.2017.08.001; pmid: 28877501
99. R. Chen, J. Mintseris, J. Janin, Z. Weng, A protein-protein docking benchmark. *Proteins* **52**, 88–91 (2003). doi: 10.1002/prot.10390; pmid: 12784372
100. C. M. Wood *et al*., High-resolution structure of the native histone octamer. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **61**, 541–545 (2005). doi: 10.1107/S1744309105013813; pmid: 16511091
101. A. Šali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993). doi: 10.1006/jmbi.1993.1626; pmid: 8254673
102. S. M. Vos *et al*., Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560**, 607–612 (2018). doi: 10.1038/s41586-018-0440-4; pmid: 30135578
103. M. N. Wojtas, M. Mogni, O. Millet, S. D. Bell, N. G. A. Abrescia, Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. *Nucleic Acids Res.* **40**, 9941–9952 (2012). doi: 10.1093/nar/gks692; pmid: 22848102
104. D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, R. D. Kornberg, Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006). doi: 10.1016/j.cell.2006.11.023; pmid: 17129781
105. K. S. Murakami, X-ray crystal structure of *Escherichia coli* RNA polymerase σ70 holoenzyme. *J. Biol. Chem.* **288**, 9126–9134 (2013). doi: 10.1074/jbc.M112.430900; pmid: 23389035
106. R. P. Joosten *et al*., A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411–D419 (2011). doi: 10.1093/nar/gkq1105; pmid: 21071423
107. W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983). doi: 10.1002/bip.360221211; pmid: 6667333
108. M.-Y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006). doi: 10.1110/ps.062416606; pmid: 17075131
109. J. P. Erzberger *et al*., Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* **158**, 1123–1135 (2014). doi: 10.1016/j.cell.2014.07.044; pmid: 25171412
110. Y. Shi *et al*., Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell. Proteomics* **13**, 2927–2943 (2014). doi: 10.1074/mcp.M114.041673; pmid: 25161197
111. R. H. Swendsen, J. S. Wang, Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.* **57**, 2607–2609 (1986). doi: 10.1103/PhysRevLett.57.2607; pmid: 10033814
112. S. Viswanath, I. E. Chemmama, P. Cimermancic, A. Sali, Assessing exhaustiveness of stochastic sampling for integrative modeling of macromolecular structures. *Biophys. J.* **113**, 2344–2353 (2017). doi: 10.1016/j.bpj.2017.10.005; pmid: 29211988
113. B. Vallat, B. Webb, J. D. Westbrook, A. Sali, H. M. Berman, Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules. *Structure* **26**, 894–904.e2 (2018). doi: 10.1016/j.str.2018.03.011; pmid: 29657133
114. J. D. Chodera, A simple method for automated equilibration detection in molecular simulations. *J. Chem. Theory Comput.* **12**, 1799–1805 (2016). doi: 10.1021/acs.jctc.5b00784; pmid: 26771390
115. L. McInnes, J. Healy, S. Astels, hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2**, 205 (2017). doi: 10.21105/joss.00205
116. E. D. Merkley *et al*., Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23**, 747–759 (2014). doi: 10.1002/pro.2458; pmid: 24639379
117. D. Schneidman-Duhovny *et al*., A method for integrative structure determination of protein-protein complexes. *Bioinformatics* **28**, 3282–3289 (2012). doi: 10.1093/bioinformatics/bts628; pmid: 23093611
118. M. J. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004). doi: 10.1093/bioinformatics/bth078; pmid: 14871861
119. A. J. Saldanha, Java Treeview—Extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004). doi: 10.1093/bioinformatics/bth349; pmid: 15180930
120. J. H. Morris, C. C. Huang, P. C. Babbitt, T. E. Ferrin, structureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics* **23**, 2345–2347 (2007). doi: 10.1093/bioinformatics/btm329; pmid: 17623706
121. N. T. Doncheva, K. Klein, F. S. Domingues, M. Albrecht, Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.* **36**, 179–182 (2011). doi: 10.1016/j.tibs.2011.01.002; pmid: 21345680
122. J. H. Morris *et al*., setsApp: Set operations for Cytoscape Nodes and Edges. *F1000 Res.* **3**, 149 (2014). doi: 10.12688/f1000research.4392.1; pmid: 25352980
123. M. A. Collart, S. Oliviero, Preparation of yeast RNA. *Curr. Protoc. Mol. Biol.* **23**, 13.12.1–13.12.5 (2001). doi: 10.12688/f1000research.4392.1; pmid: 25352980
124. D. Kim *et al*., TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013). doi: 10.1186/gb-2013-14-4-r36; pmid: 23618408
125. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015). doi: 10.1093/bioinformatics/btu638; pmid: 25260700
126. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
127. M. Costanzo *et al*., The genetic landscape of a cell. *Science* **327**, 425–431 (2010). doi: 10.1126/science.1180823; pmid: 20093466
128. G. M. Wilmes *et al*., A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Mol. Cell* **32**, 735–746 (2008). doi: 10.1016/j.molcel.2008.11.012; pmid: 19061648
129. Y. H. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995). doi: 10.1111/j.2517-6161.1995.tb02031.x
130. D. Wessel, U. I. Flügge, A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138**, 141–143 (1984). doi: 10.1016/0003-2697(84)90782-6; pmid: 6731838
131. M. Downey *et al*., Acetylome profiling reveals overlap in the regulation of diverse processes by sirtuins, gcn5, and esa1. *Mol. Cell. Proteomics* **14**, 162–176 (2015). doi: 10.1074/mcp.M114.043141; pmid: 25381059
132. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008). doi: 10.1038/nbt.1511; pmid: 19029910
133. B. MacLean *et al*., Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010). doi: 10.1093/bioinformatics/btq054; pmid: 20147306
134. M. Choi *et al*., MSstats: An R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014). doi: 10.1093/bioinformatics/btu305; pmid: 24794931
135. R. Dronamraju *et al*., Spt6 association with RNA polymerase II directs mRNA turnover during transcription. *Mol. Cell* **70**, 1054–1066.e4 (2018). doi: 10.1016/j.molcel.2018.05.020; pmid: 29932900
136. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods* **25**, 402–408 (2001). doi: 10.1006/meth.2001.1262; pmid: 11846609
137. R. Dronamraju, B. D. Strahl, A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res.* **42**, 870–881 (2014). doi: 10.1093/nar/gkt1003; pmid: 24163256
138. J. T. Robinson *et al*., Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011). doi: 10.1038/nbt.1754; pmid: 21221095
139. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014). doi: 10.1093/bioinformatics/btt656; pmid: 24227677
140. R. Edgar, M. Domrachev, A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002). doi: 10.1093/nar/30.1.207; pmid: 11752295

ACKNOWLEDGMENTS

We thank M. Eckhardt, I. Chemmama, A. Verreault, D. Truong, P. Beltrao, and A. Roguev for helpful discussions. **Funding:** We acknowledge funding through NIH grants R01 GM084448, R01 GM084279, P50 GM081879, P50 AI150476, P01 HL089707, U54 NS100717, and R01 GM098101 to N.J.K.; R01 GM083960, P50 GM082250, U19 AI135990, S10 OD021596, and P41 GM109824 to A.Sa.; U54 RR020839 and U54 GM103520 to J.D.B.; R35 GM118061 to C.G.; GM126900 to B.D.S.; P01GM105473 to J.E.H.; U54 RR020839 and R01 GM084448 to J.S.B.; OD009180 and GM123159 to J.S.F.; and GM008284 to A.Sh. We also acknowledge funding through the National Key Research and Development Program of China (2017YFA0505103), Shenzhen Science and Technology Program (KQTD20180413181837372), Guangdong Provincial Key Laboratory of Synthetic Genomics (2019B030301006), and Bureau of International Cooperation, Chinese Academy of Sciences (172644KYSB20180022) to J.D. and the National Natural Science Foundation of China (31800069) to S.J. **Author contributions:** H.B., I.E., P.C., J.D.B., A.Sa., and N.J.K. designed the research; H.B., S.B., A.Sh., R.A., J.X., R.D., S.J., G.D., D.B., K.K.C., R.H., and S.W. performed experimental assays; I.E., P.C., R.P., and D.S. developed and performed integrative modeling; H.B., S.B., M.S., and D.M. analyzed experimental data; J.M. developed software; J.S.B., J.S.F., J.E.H., B.D.S., C.A.G., J.D., J.D.B., A.Sa., and

N.J.K. supervised the research; H.B., I.E., S.B., A.Sa., and N.J.K. wrote the paper with input from all authors. **Competing interests:** J.D.B. is a founder and director of CDI Labs, Inc.; a founder of Neochromosome, Inc.; and a founder and SAB member of ReOpen Diagnostics. J.D.B. also serves or served on the scientific advisory board of Sangamo, Inc.; Modern Meadow, Inc.; Sample6, Inc.; and the Wyss Institute. J.S.B. is a founder and director of Neochromosome, Inc. **Data and materials availability:** IMP modeling scripts, input data, and output results are available at https://integrativemodeling.org/systems/pemap. IMP is an open-source program, available at http://integrativemodeling.org. The clustered pE-MAP and correlation maps are available in the supplementary data and can be visualized using the Java Treeview app (*119*), available at http://jtreeview.sourceforge.net/. The stEMAP app source is publicly available at https://github.com/RBVI/StEMAPApp, and stEMAP will be available through the Cytoscape App store. stEMAP requires installation of Cytoscape and ChimeraX. Cytoscape is available at https://cytoscape.org/download.html and ChimeraX at www.rbvi.ucsf.edu/chimera/download.html. stEMAP also requires installation of the Sets, RINalyzer, StructureVizX, clusterMaker2, and ctdReader apps, all available through the Cytoscape apps menu. The modified version of PDB 1ID3 (including the unstructured H3 and H4 tails for illustrative purposes) is available in the supplementary data. The yeast RNAPII pE-MAP is available in the supplementary materials of (*15*). The bacterial RNAP point mutation data are available in the supplementary materials of (*20*). The library of histone mutant strains is available from J. Dai under a material transfer agreement with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. RNA-seq and ATAC-seq data have been deposited in NCBI's Gene Expression Omnibus (*140*) and are accessible through GEO Series accession number GSE160350 (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE160350).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eaaz4910/suppl/DC1
Figs. S1 to S11
Tables S1 to S9
References (*141*–*144*)
MDAR Reproducibility Checklist
Data S1 to S3

View/request a protocol for this paper from *Bio-protocol*.

14 October 2019; resubmitted 23 July 2020
Accepted 23 October 2020
10.1126/science.aaz4910

RESEARCH ARTICLE

CLIMATE CHANGE

Recent global decline of CO_2 fertilization effects on vegetation photosynthesis

Songhan Wang[1,2], Yongguang Zhang[1,2,3]*, Weimin Ju[1,2], Jing M. Chen[1,4], Philippe Ciais[5], Alessandro Cescatti[6], Jordi Sardans[7,8], Ivan A. Janssens[9], Mousong Wu[1,2], Joseph A. Berry[10], Elliott Campbell[11], Marcos Fernández-Martínez[9], Ramdane Alkama[6], Stephen Sitch[12], Pierre Friedlingstein[13], William K. Smith[14], Wenping Yuan[15], Wei He[1,2], Danica Lombardozzi[16], Markus Kautz[17], Dan Zhu[5], Sebastian Lienert[18], Etsushi Kato[19], Benjamin Poulter[20], Tanja G. M. Sanders[21], Inken Krüger[21], Rong Wang[22], Ning Zeng[23,24], Hanqin Tian[25], Nicolas Vuichard[5], Atul K. Jain[26], Andy Wiltshire[12], Vanessa Haverd[27], Daniel S. Goll[5,28], Josep Peñuelas[7,8]

The enhanced vegetation productivity driven by increased concentrations of carbon dioxide (CO_2) [i.e., the CO_2 fertilization effect (CFE)] sustains an important negative feedback on climate warming, but the temporal dynamics of CFE remain unclear. Using multiple long-term satellite- and ground-based datasets, we showed that global CFE has declined across most terrestrial regions of the globe from 1982 to 2015, correlating well with changing nutrient concentrations and availability of soil water. Current carbon cycle models also demonstrate a declining CFE trend, albeit one substantially weaker than that from the global observations. This declining trend in the forcing of terrestrial carbon sinks by increasing amounts of atmospheric CO_2 implies a weakening negative feedback on the climatic system and increased societal dependence on future strategies to mitigate climate warming.

Terrestrial ecosystems have accounted for more than half of the global carbon sink during the last six decades and have thus substantially mitigated climate warming (*1*). Global process-based models attribute part of the increasing land carbon sink (*2*) to the increase in vegetation productivity driven by the fertilization effect of increasing atmospheric CO_2 concentration (*3*), i.e., the CO_2 fertilization effect (CFE), a process that acts as a negative feedback in the climate system (*4*). First introduced by Keeling (*5*), the β factor generally is used to characterize the plant response to increasing CO_2 concentration. To compare the results between global analysis and experimental measurements, we used an approximation form of the original β ($\beta = \frac{\partial GPP}{\partial C_a}$) (*6*), which is defined as the relative increase in gross primary production (GPP) in response to a 100-ppm increase in atmospheric CO_2 concentration (C_a).

Increases of GPP originate from the direct acceleration of photosynthesis in response to the increased supply of CO_2 (*7*), but they are also modified by a suite of indirect responses, including water saving because of the reduced stomatal conductance under increasing atmospheric CO_2 concentrations (*8*, *9*) and nutrient limitation (*10*). Enhanced GPP and net primary production (NPP) are commonly observed in field experiments, such as free-air CO_2 enrichment (FACE) experiments (*7*) and open-top chamber experiments (*11*), where ecosystems are exposed to elevated CO_2 in the range of two times ambient values, albeit with different response magnitudes across sites (*7*).

The fingerprint of CFE on the carbon and water cycles on global scales (*3*, *8*) is, however, more elusive, given the covariation of atmospheric CO_2 with other environmental drivers of vegetation productivity. Elucidating this fingerprint of CFE is a scientific problem involving detection and attribution methods, requiring either statistical methods and long-term data such that effects of non-CFE drivers can be removed empirically (*12*, *13*), or process-based models with which CFE can be isolated by deliberate factorial simulations (*14*).

Process-based models of the terrestrial carbon cycle have indicated that CFE accounts for ~70% of the increasing global trend in foliar area, i.e., global greening (*15*), and up to 60% of the current terrestrial carbon sink (*3*). These models project that CFE will induce increased land carbon storage by the end of this century, despite opposing effects from climate change (*3*, *16*). A cascade of uncertainties in projecting future land carbon storage arises from the responses of photosynthesis, NPP, and ecosystem carbon turnover times to increasing CO_2, especially in the presence of increasing limitations from nutrient and water availability. Therefore, accurately quantifying the temporal dynamics of CFE on GPP is essential for reducing the uncertainties of future land carbon storage and climate projections derived on the basis of Earth system models (*4*).

A study using the long-term CO_2 concentration measurements at the Barrow station (Alaska, 71°N) found that the sensitivity of the seasonal amplitude of atmospheric CO_2 to the increase in CO_2 concentration has decreased, which suggests a declining response of GPP to CO_2 at northern high latitudes (*17*). If such declining trends prevail across the globe, the terrestrial vegetation carbon sink response to increasing CO_2 will decrease, with important consequences for the global carbon budget and for the effort required from mitigation policies to meet future climate targets. For this reason, a comprehensive assessment of the temporal dynamics of CFE on global GPP is timely as a first step to understand the impacts of CFE on the trends of the land carbon sinks. We used three long-term satellite datasets collected during 1982 to 2015: a recently developed vegetation index (NIR_V) (*18*) from the Advanced Very High Resolution Radiometer (AVHRR NIR_V), the fusion of NIR_V from AVHRR and the Moderate Resolution Imaging Spectroradiometer (AVHRR+MODIS NIR_V), and the fusion of NIR_V and sun-induced chlorophyll fluorescence (SIF) (*19*) (NIR_V+SIF), which are proxies for GPP (hereafter, satellite GPP proxies) (supplementary text S1 and figs. S1 to S5) (*20*). We also corroborated our findings using long-term GPP time series from

[1]International Institute for Earth System Science, Nanjing University, Nanjing, Jiangsu 210023, China. [2]Jiangsu Provincial Key Laboratory of Geographic Information Technology, Key Laboratory for Land Satellite Remote Sensing Applications of Ministry of Natural Resources, School of Geography and Ocean Science, Nanjing University, Nanjing, Jiangsu 210023, China. [3]Huangshan Park Ecosystem Observation and Research Station, Ministry of Education, Huangshan, China. [4]Department of Geography and Planning, University of Toronto, Toronto, Ontario, Canada. [5]Laboratoire des Sciences du Climat et de l'Environnement, CEA CNRS UVSQ, Gif-sur-Yvette, France. [6]European Commission, Joint Research Centre (JRC), Ispra, Italy. [7]CSIC, Global ecology Unit CREAF-CSIC-UAB, Bellaterra 08193, Catalonia, Spain. [8]CREAF, Cerdanyola del Vallès 08193, Catalonia, Spain. [9]Department of Biology, Centre of Excellence PLECO (Plant and Vegetation Ecology), University of Antwerp, Wilrijk, Belgium. [10]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305, USA. [11]Sierra Nevada Research Institute, University of California, Merced, CA 95343, USA. [12]College of Life and Environmental Sciences, University of Exeter, Exeter, UK. [13]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK. [14]School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA. [15]School of Atmospheric Sciences, Center for Monsoon and Environment Research, Sun Yat-Sen University, Guangzhou, China. [16]Terrestrial Sciences Section, National Center for Atmospheric Research, Boulder, CO, USA. [17]Forest Research Institute Baden-Württemberg, Freiburg, Germany. [18]Climate and Environmental Physics, Physics Institute, and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland. [19]Institute of Applied Energy, Tokyo, Japan. [20]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [21]Thünen Institute of Forest Ecosystems, Alfred-Möller-Str. 1, 16225 Eberswalde, Germany. [22]Department of Environmental Science and Engineering, Fudan University, Shanghai, 200433, China. [23]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742, USA. [24]LASG, Institute of Atmospheric Physics, Chinese Academy of Science, Beijing 100029, China. [25]International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA. [26]Department of Atmospheric Sciences, University of Illinois, 105 South Gregory Street, Urbana, IL 61801-3070, USA. [27]CSIRO Oceans and Atmosphere, Canberra, ACT 2601, Australia. [28]Institute of Geography, University of Augsburg, Augsburg, Germany.
*Corresponding author. Email: yongguang_zhang@nju.edu.cn

eddy covariance (EC) flux towers, from a light use efficiency (LUE) model (*21*), and from an ensemble of terrestrial carbon cycle models (TRENDY v6, trends and drivers of the regional scale sources and sinks of carbon dioxide) (*14*).

Global temporal trend of β from observations and models

Linear and nonlinear multiple regression approaches were used to estimate β from satellite GPP proxies at each pixel across the globe (*12*, *13*, *20*). Pixels with large land-cover changes were excluded from the analyses (supplementary text S2 and fig. S6). The global median β during 1982 to 2015 was 16.1 ± 11.5% 100 ppm^{-1} (fig. S7), consistent with the FACE experiments (*7*), which on average suggested a 15.5% increase in light-saturated photosynthesis per 100 ppm CO_2. Results showed that the estimated β values for different vegetation physiology and biome types were also aligned with experimental results (*7*), e.g., the β values of C_4 plants were much smaller than those of C_3 plants (supplementary text S3 and fig. S8). We next calculated the time series of β with 15-year moving windows during 1982 to 2015 and found that β significantly decreased at a rate of −0.92 ± 0.12% 100 ppm^{-1} $year^{-1}$ ($p < 0.01$) (Fig. 1A). This decrease was also evident after excluding the crop areas (fig. S9). Compared with the global median of the declining trend in β, grasslands and plants in cold climate zones exhibited a larger declining trend of β, whereas the β trends of shrubs and plants in tropical areas were slightly lower; nevertheless, the decreases in β across all the climate zones and vegetation types were notable (fig. S10). Across the global terrestrial areas, β

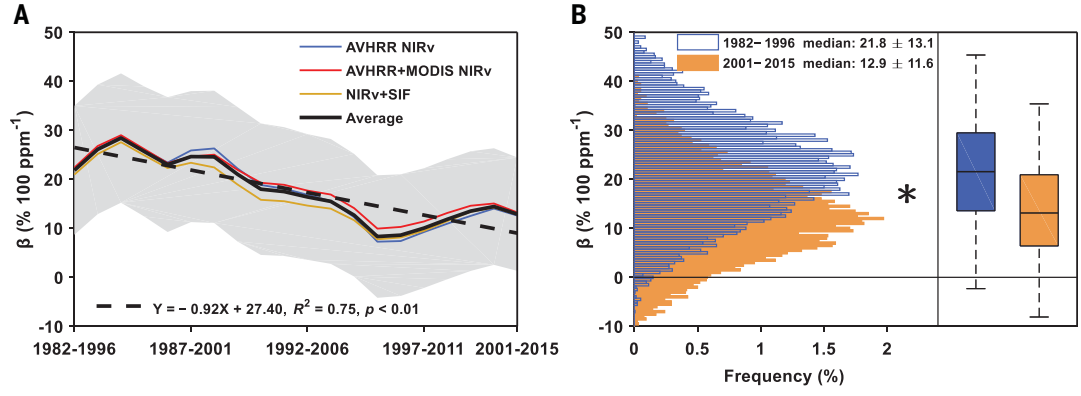


Fig. 1. Declining trend of global β. (**A**) Temporal dynamics of β for three satellite GPP proxies with 15-year moving windows during 1982 to 2015. The gray area indicates 1 SD on either side of the mean (n = 12,850). The trend and statistical significance (p value) of the β time series were estimated using the Mann-Kendall test. (**B**) (Left) The histogram distribution of β across all pixels in two 15-year periods. β was the average of these three satellite GPP proxies. (Right) Boxes represent the interquartile ranges of the β values (solid lines represent medians), and whiskers extend to one times the interquartile range. Median β values for these two periods and their SDs are shown at the top of the graphs. The asterisk indicates a significantly different β between these two periods, on the basis of a two-sample Kolmogorov-Smirnov test at $p < 0.01$.

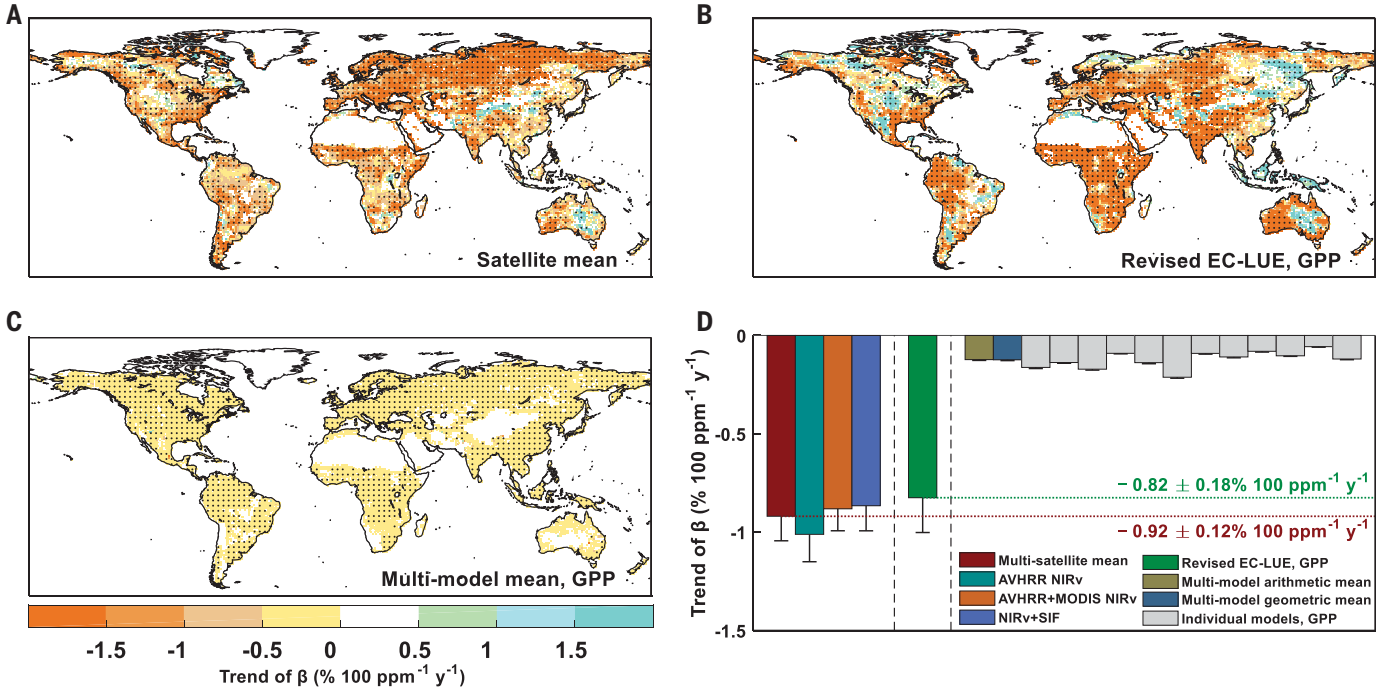


Fig. 2. Global declining trend of β. Spatial patterns of trends in β derived from three datasets using 15-year moving windows, reported as the percentage per 100 ppm per year. (**A**) Means from three satellite GPP proxies. (**B**) The revised EC-LUE GPP. (**C**) Multimodel mean GPP determined from the TRENDY v6 ensemble. All datasets are from 1982 to 2015. The Mann-Kendall test was used to estimate the trend of β pixel by pixel, and the regions with black dots indicate significant trends ($p < 0.05$). The pixel size is 1°. (**D**) Mean declining trend of β from different datasets. The error bars represent SEs.

significantly declined from 21.8 ± 13.1% 100 ppm^{-1} during 1982 to 1996 to 12.9 ± 11.6% 100 ppm^{-1} during 2001 to 2015 (Fig. 1B).

We also found a robust decrease in β when we used another satellite NIRv time series (fig. S11), various land cover change masks (fig. S12), different window lengths (supplementary text S4 and fig. S13), various combinations of explanatory variables (supplementary text S5 and figs. S14 and S15), and different definitions of growing seasons (fig. S16) and when we included the consideration of seasonal precipitation (supplementary text S6 and fig. S17). After considering the uncertainties from original satellite data and from the methods, the global β estimated from satellite GPP proxies during 2001 to 2015 was still significantly lower than that during 1982 to 1996 (supplementary text S7 and figs. S18 and S19). To verify whether the trend of β was an artefact of our regression method, we also estimated the temporal dynamics of β using the optimal fingerprint attribution method (supplementary text S8) and again found that the declining trend of β was significant (fig. S20). Altogether, these results based on satellite observations of GPP proxies suggest a significant declining trend in the response of GPP to increasing atmospheric CO_2.

To further assess the decreases in β inferred from satellite GPP proxies, we used other independent datasets: (i) the GPP time series derived from the revised EC-LUE model (*21*) (which also accounted for the direct CO_2 effects on LUE); (ii) satellite-based leaf area index (LAI) time series; (iii) multiyear GPP estimations from 22 EC flux sites (table S4); and (iv) a global long-term GPP dataset from the Carbon Cycle Data Assimilation System (CCDAS) (supplementary text S9 to S12). We also found a significant declining trend of β on the basis of the revised EC-LUE GPP, with a rate of −0.82 ± 0.18% 100 ppm^{-1} year^{-1} (fig. S21), similar to the estimates from satellite GPP proxies. However, the result from satellite-based LAI showed a smaller declining rate of β (−0.59 ± 0.11% 100 ppm^{-1} year^{-1}) (fig. S22). This divergence suggested that the decrease in β was likely due to two factors: the direct effect on foliar physiology and the indirect effect on LAI. The former refers to the CO_2-induced stimulation of carboxylation per unit leaf area, whereas the latter depicts the increased carbon sequestration used for leaf area expansion. The smaller declining β trend from satellite LAI suggested that these two effects were both relevant for the full description of the global decreases in the CO_2 effect on GPP. Ground GPP estimates from EC flux sites also confirmed our findings, by indicating an average declining rate of β at −0.70% 100 ppm^{-1} year^{-1}, which was comparable to the results from satellite GPP proxies around these sites (fig. S23). On the basis of the CCDAS GPP, a data–model fusion product constrained by atmospheric CO_2 observations and which is independent from satellite data, we also found a declining β trend at a rate of −0.62% 100 ppm^{-1} year^{-1} (fig. S24). The overall consistency between various remote-sensing data, ground measurements, and a carbon assimilation system suggests that the global decline of β is robust and coherent across multiple, independent observations.

Spatial pattern of the temporal trend of β across the globe

The geographic distribution of the temporal trends of β from satellite GPP proxies (Fig. 2A) revealed that β decreased in ~86% of global lands (fig. S25, A to C). Areas with declining β spanned over most of the globe. By contrast, increasing β was found in some limited regions of Southeast Asia, east Australia, and North America (Fig. 2, A and B). These increasing β trends are likely driven by an intensification of management in croplands (i.e., irrigation and fertilization) or related to increasing atmospheric nutrients deposition in recent years (fig. S33). We also observed similar spatial patterns for β trends calculated using both 10-year and 17-year moving windows (fig. S26). Overall, the decreases of β in tropical areas were smaller, whereas cold regions had a slightly stronger declining trend (figs. S10A and S25, A to C). The latter may be because the declining signals of β in northern high latitudes include both the indirect effects on LAI and the direct effect on foliar physiology, whereas those in warmer climates include only the physiological effect, given that LAI in some tropical humid regions may already be close to saturation. Results of the satellite LAI confirmed this interpretation by suggesting a larger reduction of β in colder northern regions than in warmer climates (fig. S25, E and F). Similar declines in β were obtained from the revised EC-LUE GPP, which covered ~74% of the global terrestrial area (Fig. 2B). In general, all datasets indicated that β has been declining in most global lands and across various vegetation types and climate zones.

We also investigated whether an ensemble of state-of-the-art carbon cycle models (*14*) was able to reproduce these observed global declines of β on GPP. We used results from 12 models that contributed to the TRENDY v6 ensemble to calculate β (*20*) (supplementary text S13). For each model, we extracted the differences in simulated GPP between S1 (time-varying CO_2 only) and S0 (constant CO_2) scenarios, which represented the CO_2 effect on vegetation productivity. These models predicted a negative trend in β, both for the multimodel mean and for the individual one, but the declining rate was clearly lower than that derived from satellite GPP proxies and with no evident spatial variations (Fig. 2C and fig. S27). Grouping the estimates of CFE trends, we found that the global β declined at a rate of −0.92 ± 0.12% 100 ppm^{-1} year^{-1} for satellite GPP proxies and only −0.12 ± 0.01% 100 ppm^{-1}

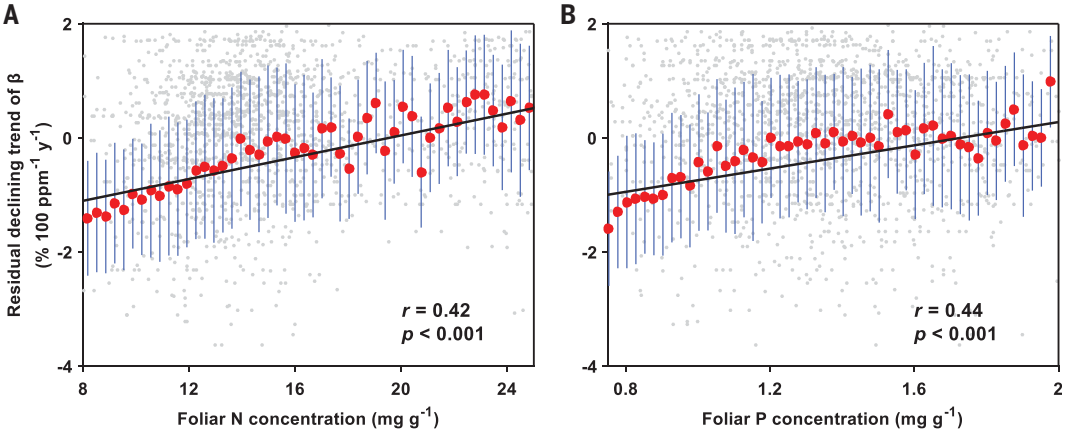


Fig. 3. Relationship between the residual trend of β and foliar nutrient concentrations. Results for foliar N (**A**) and foliar P (**B**) concentrations after accounting for site mean annual air temperature, and mean annual precipitation from 3846 samples on the basis of a spatial mixed-effects model. Model performance and results are presented in table S5. Data are classified into 50 bins for clear visualization on the basis of foliar N or P concentrations. The red dots represent the means for each bin, and the blue lines represent the SDs of the means. The gray dots represent the raw individual samples, and black lines represent the linear regressions of these gray dots. The correlation coefficients (*r*) and *p* values were calculated on the basis of the raw data (*n* = 3846).

year^{-1} for TRENDY GPP (Fig. 2D). β trends of individual models were variable, from −0.06 to −0.21% 100 ppm^{-1} year^{-1} (fig. S28), but all of them were lower than that from satellite GPP proxies. Given that we defined β as the percent increase in GPP per 100-ppm increase in CO_2 and that the S1 simulations solely considered the CO_2 effect, the smaller decreases in β from TRENDY GPP were likely caused by the saturating physiological response of GPP to CO_2 (*22*), without adequately capturing the concurrent emergence of other limiting factors driven by the changing environmental conditions. We also used GPP from the TRENDY S2 and S3 scenarios to estimate the trends in β by using the regression approaches (*20*). The temporal trends of β remained unaltered, with a value of approximately −0.12 ± 0.12% 100 ppm^{-1} year^{-1} for both TRENDY S2 and S3 simulations (fig. S29). These results highlight that the ongoing strong decrease in β inferred from satellite datasets is probably underestimated by TRENDY models.

Possible mechanisms accounting for the declining β

Two possible non–mutually exclusive hypotheses were proposed to account for the declining β and to explain why TRENDY models failed to adequately replicate the magnitude of this decline: (i) the increasing constraints on vegetation productivity from emerging nutrient limitations and imbalances that were not adequately represented in models and (ii) current models underestimated the sensitivity of terrestrial GPP to changes in water availability.

The first hypothesis relates to the possible effect of the growing limitation of key nutrients, including nitrogen (N) and phosphorus (P). Using 410 groups of ground-based foliar N and P observations (supplementary text S14 and fig. S30), we found a general decrease in foliar N and P concentrations, with mean values of −0.24 ± 0.06% year^{-1} and −0.55 ± 0.06% year^{-1}, respectively (fig. S31, A and B). Our findings are consistent with a recent study suggesting a general global pattern of decreasing foliar N concentration (*23*) and with many examples of local to regional decreases in foliar N and P concentrations (*24*, *25*). Enhanced GPP from the increasing atmospheric CO_2 concentrations, leading to larger NPP and higher nutrient demands by plants, may partly explain the observed declines of foliar N and P concentrations. Concurrently, these decreases in key foliar nutrients may impose limitations on GPP and thereby limit β. To test this hypothesis, we applied a linear spatial mixed-effects model (table S5) to investigate the relationship between the trend of β and foliar N or P concentrations (supplementary text S14 and fig. S32). After accounting for the trends in β explained statistically by mean annual air temperature and mean annual precipitation, we found clear positive correlations between the residual declining trends of β with both foliar N and P concentrations across European forests (Fig. 3). These results suggested that vegetation with lower foliar nutrient concentrations generally showed larger declines in β, therefore supporting our hypothesis of a role for nutrient limitation in the temporal dynamics of β.

This phenomenon was also verified from global trends in atmospheric N and P depositions, which suggested that areas with decreasing atmospheric nutrient supplies probably had larger decreases in β (such as Europe and Siberia) and vice versa (such as East Asia) (fig. S33). This may be because vegetation in areas with declining nutrient supplies from atmospheric depositions tend to have larger N and P limitations on GPP. Moreover, the increases of N and P depositions in some regions of East Asia may explain the increasing trend of β in these areas (Figs. 2A and fig. S33). The ongoing decreases in foliar nutrients might constrain the plant photosynthetic capacity and result in the decline of β, which might not have been adequately represented in current models (supplementary text S14). Regarding this aspect, the models that included C-N cycle interactions to emulate nutrient constraints exhibited a larger declining rate of β (fig. S34), partly confirming this interpretation. The role of N limitation on β has been widely suggested by experimental evidence (*10*, *26*, *27*), model analyses (*28*), and synthesis reviews (*29*, *30*), all of which have been consistent with our analyses. Our finding regarding P limitation was also consistent with FACE experiments, which demonstrated that insufficient P availability generally had negative impacts on β (*31*). Furthermore, as foliar P:N ratios positively correlated with plant net photosynthesis and growth (*32*), the decreasing foliar P:N ratios in European forests (fig. S31C) suggested a worsening nutrient imbalance that may partly account for the observed decline in β.

The second hypothesis to explain the divergence of the β trend between the TRENDY models and the satellite-derived estimates stated that these models underestimated the sensitivity of GPP to water availability, because the coupling between water and carbon in models is underestimated (*33*). To test this hypothesis, we analyzed the sensitivity of satellite GPP

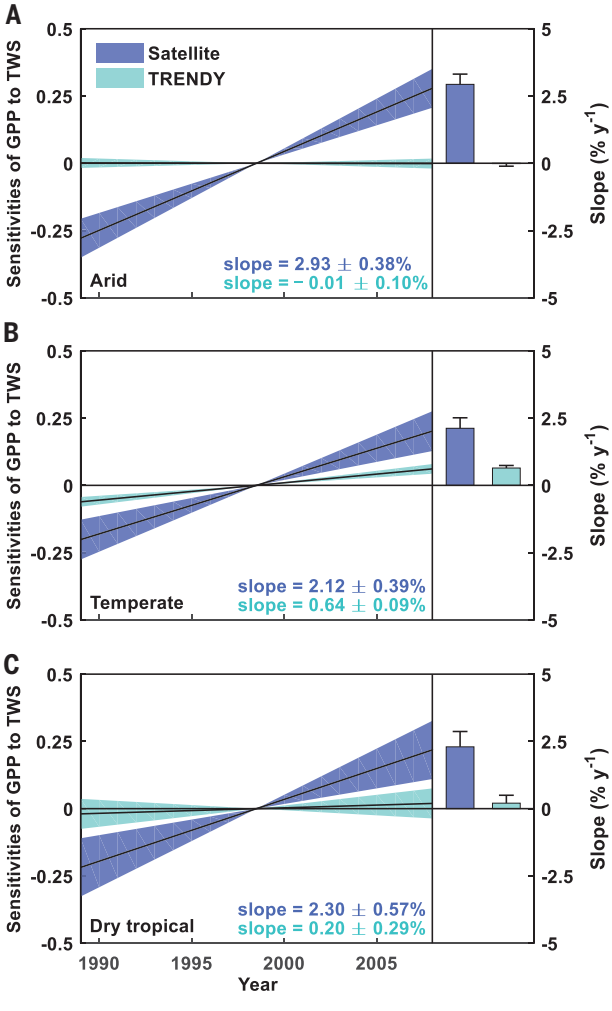


Fig. 4. Sensitivities of GPP to TWS. The changes in satellite GPP proxies or TRENDY GPP per unit change in TWS were estimated for arid (**A**), temperate (**B**), and dry tropical (**C**) climate zones using a moving window of 15 years. The sensitivity time series have been standardized. The solid lines in the left panels represent the linear regressions. The shaded areas represent the SEs. The bars and error bars in the right panels represent the slopes and their SEs, respectively.

proxies and TRENDY GPP to water availability using a moving window of 15 years (supplementary text S15). We used the terrestrial water storage (TWS) data to represent the availability of water to plants and selected arid, temperate, and dry tropical climate zones as the research areas (fig. S35), as the GPP in these zones was found to be highly sensitive to TWS (fig. S36). We observed that the sensitivity of GPP to TWS was relatively constant across these three zones for the multimodel mean of the TRENDY GPP, whereas the TWS sensitivities derived from satellite GPP proxies exhibited significantly increasing trends for arid (2.93 ± 0.38% $year^{-1}$), temperate (2.12 ± 0.39% $year^{-1}$), and dry tropical (2.30 ± 0.57% $year^{-1}$) ecosystems (Fig. 4). In temperate and dry tropical areas, the majority of models exhibited divergent results compared to satellite GPP proxies, and almost all of them largely underestimated the GPP sensitivities to TWS in arid areas (fig. S37). The decreases in β were larger in the regions where the increases of TWS sensitivities were higher (fig. S38). These strongly divergent sensitivities between satellite GPP proxies and TRENDY GPP were also confirmed when using various TWS products (*20*) (supplementary text S15 and fig. S39) and when using a shorter moving window (fig. S40).

This finding implies that GPP and consequently land carbon uptake are more sensitive to the variations in water availability than assumed by the TRENDY models, as previously suggested by several recent studies (*34*, *35*). Moreover, a recent study showed that the CFE of grasslands could be reduced under drier conditions (*36*). The relationship between drought stress and CFE is complex and ecosystem specific and may be affected by the total annual rainfall as well as rainfall seasonality. It was shown, for instance, in 19 grasslands experiments that the annual CFE was negated when spring precipitation became too low (*37*). Field experiments also support our results by suggesting that the CFE on vegetation productivity is at least partly limited by water availability (*38*). The significant increases in the GPP sensitivities to TWS may thus partly explain the decreases of β in arid, temperate, and dry tropical climate zones.

From a theoretical perspective, the global declines of CFE may result from several factors. First, given that the CO_2-induced photosynthesis stimulation at the leaf level is scaled up to the canopy level through LAI (*39*, *40*), the declining β on the basis of satellite LAI (fig. S22) could partly explain the global decreases of the GPP response to CO_2. Consistent with our findings, a recent study using the FACE experiments on mature forests found a relatively low CO_2 effect on GPP (*41*), possibly because the LAI of these forests did not change much (*39*). Second, the CO_2 effect on photosynthesis at the leaf level involves both the stimulation of carboxylation and increases in water use efficiency (WUE). The former may possibly be regulated by the foliar key nutrients (N and P), and the latter is related to the water availability. According to the progressive N limitation theory, the soil N availability for plant growth may be expected to diminish over time (*42*), possibly leading to the observed global decreases in β. A recent study using carbon isotope measurements revealed diminishing CO_2-induced WUE gains across global forests (*43*), which supported our findings of the effects of water supply limitation on CFE and could partly explain the global declines in β. Moreover, complex interactive effects between nutrients and water supply may also have impacts on CFE. For example, low soil water supply could possibly strengthen the nutrient constraints on CFE through limiting nutrient decomposition and diffusion in soils (*38*). Our further analysis using model simulations showed that the GPP trends have clear reductions in arid areas, once the interactions between N limitation and climate constraints are considered, supporting this hypothesis (fig. S41B and supplementary text S16). Nevertheless, an excess of N fertilizer may result in a reduction in the soil water and therefore possibly lead to drought stress (*44*). Regarding this aspect, our analysis also highlights the need for future efforts to better understand the complex interactions between nutrients, water, and CO_2 effects on vegetation from a climate change perspective. Additional mechanisms (e.g., plant acclimation or changes in plant species over time) may also explain the observed global decrease in CFE (supplementary text S17).

Conclusion

Our analyses showed a significant and spatially extensive decline in β, which implies a substantial reduction of the positive effects of increasing atmospheric CO_2 on terrestrial carbon uptake. A recent study suggested that the CO_2 effect on the carbon cycle in tropical regions (*3*) might be counteracted by impacts from climate-driven changes (*45*), in agreement with our findings. Although still under debate, the possible increasing trend of the airborne fraction of anthropogenic CO_2 may imply a saturation of the CO_2 sinks from land and oceans (*46*–*49*), which may be partly caused by the global decline in CFE. Current carbon cycle models also exhibit such global decreases in β but fail to adequately detect the sharp declining trend that we identified from satellite data. This divergence between observations and process-based models possibly originates from the models' limitations in adequately representing the emerging decline in key foliar nutrient concentrations and the increasing constraints of water limitations on vegetation productivity. Ultimately, these results indicate that terrestrial photosynthesis may not increase as much as models project, possibly reducing the potential of land-based climate mitigation, further accelerating global warming and exacerbating the efforts required for meeting climate targets. Our findings also highlight the need for better characterizations of the biogeochemical and hydrological effects on vegetation in current carbon cycle models to produce more robust projections of the terrestrial carbon budget for the next decades.

REFERENCES AND NOTES

1. P. Friedlingstein *et al*., *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
2. A. P. Ballantyne, C. B. Alden, J. B. Miller, P. P. Tans, J. W. White, *Nature* **488**, 70–72 (2012).
3. D. Schimel, B. B. Stephens, J. B. Fisher, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 436–441 (2015).
4. P. Friedlingstein *et al*., *J. Clim.* **27**, 511–526 (2014).
5. C. D. Keeling, "The carbon dioxide cycle. Reservoir models to depict the exchange of atmospheric carbon dioxide with the oceans and land plants" in *Chemistry of the Lower Atmosphere*, S. I. Rasool, Ed. (Plenum, 1973), pp. 251–329.
6. R. Bacastow, C. D. Keeling, "Atmospheric carbon dioxide and radio-carbon in the natural carbon cycle. II: Changes from A.D. 1700 to 2070 as deduced from a geochemical model" in *Carbon in the Biosphere*, G. M. Woodwell, E. V. Pecan, Eds. (Atomic Energy Commission, 1973), pp. 86–136.
7. E. A. Ainsworth, A. Rogers, *Plant Cell Environ.* **30**, 258–270 (2007).
8. T. F. Keenan *et al*., *Nature* **499**, 324–327 (2013).
9. J. A. Morgan *et al*., *Nature* **476**, 202–205 (2011).
10. P. B. Reich *et al*., *Nature* **440**, 922–925 (2006).
11. W. I. Dieleman *et al*., *Glob. Change Biol.* **18**, 2681–2693 (2012).
12. S. Piao *et al*., *Glob. Change Biol.* **19**, 2117–2132 (2013).
13. W. Kolby Smith *et al*., *Nat. Clim. Chang.* **6**, 306–310 (2016).
14. S. Sitch *et al*., *Biogeosciences* **12**, 653–679 (2015).
15. Z. Zhu *et al*., *Nat. Clim. Chang.* **6**, 791–795 (2016).
16. P. M. Cox *et al*., *Nature* **494**, 341–344 (2013).
17. J. Peñuelas *et al*., *Nat. Ecol. Evol.* **1**, 1438–1445 (2017).
18. G. Badgley, C. B. Field, J. A. Berry, *Sci. Adv.* **3**, e1602244 (2017).
19. L. Guanter *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1327–E1333 (2014).
20. Materials and methods are available as supplementary materials.
21. W. Yuan *et al*., *Sci. Adv.* **5**, eaax1396 (2019).
22. M. G. De Kauwe, T. F. Keenan, B. E. Medlyn, I. C. Prentice, C. Terrer, *Nat. Clim. Chang.* **6**, 892–893 (2016).
23. J. M. Craine *et al*., *Nat. Ecol. Evol.* **2**, 1735–1744 (2018).
24. M. Jonard *et al*., *Glob. Change Biol.* **21**, 418–430 (2015).
25. K. K. McLauchlan *et al*., *Sci. Rep.* **7**, 7856 (2017).
26. R. E. McMurtrie *et al*., *Funct. Plant Biol.* **35**, 521–534 (2008).
27. R. J. Norby, J. M. Warren, C. M. Iversen, B. E. Medlyn, R. E. McMurtrie, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19368–19373 (2010).
28. *Global Biogeochem. Cycles* **9**, 407–437 (1995).
29. H. Mooney, B. G. Drake, R. Luxmoore, W. Oechel, L. Pitelka, *Bioscience* **41**, 96–104 (1991).
30. A. D. McGuire, J. M. Melillo, L. A. Joyce, *Annu. Rev. Ecol. Syst.* **26**, 473–503 (1995).
31. C. Terrer *et al*., *Nat. Clim. Chang.* **9**, 684–689 (2019).
32. J. Peñuelas *et al*., *Nat. Commun.* **4**, 2934 (2013).
33. V. Humphrey *et al*., *Nature* **560**, 628–631 (2018).
34. P. B. Reich *et al*., *Nature* **562**, 263–267 (2018).
35. J. K. Green *et al*., *Nature* **565**, 476–479 (2019).
36. W. Obermeier *et al*., *Nat. Clim. Chang.* **7**, 137–141 (2017).
37. M. J. Hovenden *et al*., *Nat. Plants* **5**, 167–173 (2019).
38. P. B. Reich, S. E. Hobbie, T. D. Lee, *Nat. Geosci.* **7**, 920–924 (2014).
39. Y. Luo, S. Niu, *Nature* **580**, 191–192 (2020).
40. Q. Li *et al*., *Biogeosciences* **15**, 6909–6925 (2018).
41. M. Jiang *et al*., *Nature* **580**, 227–231 (2020).
42. Y. Luo *et al*., *Bioscience* **54**, 731–739 (2004).
43. M. A. Adams, T. N. Buckley, T. L. Turnbull, *Nat. Clim. Chang.* **10**, 466–471 (2020).
44. E. C. da Silva, R. Nogueira, M. A. da Silva, M. B. de Albuquerque, *Plant Stress* **5**, 32–41 (2011).
45. J. Liu *et al*., *Science* **358**, eaam5690 (2017).
46. J. G. Canadell *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18866–18870 (2007).

47. M. Raupach, J. Canadell, C. Le Quéré, *Biogeosciences* **5**, 1601–1613 (2008).
48. W. Knorr, *Geophys. Res. Lett.* **36**, L21710 (2009).
49. M. R. Raupach *et al*., *Biogeosciences* **11**, 3453–3475 (2014).

ACKNOWLEDGMENTS

We acknowledge constructive and insightful comments from S. Piao at Peking University. We thank H. Müller Schmied for providing the WaterGAP TWS data. We are deeply indebted to the data providers and the managers of the ICP Forests data. We also give our sincere thanks to all data providers listed in table S1 for their continuous efforts and for sharing their data. We also thank the three anonymous reviewers for their constructive and insightful suggestions and comments, which have greatly improved our manuscript. **Funding:** This research was supported by the National Key R&D Program of China (2016YFA0600202), Jiangsu Provincial NSF for Distinguished Young Scholars (BK20170018), the General Program of Natural Science Foundation of China (42071388). J.P., P.C., I.A.J., J.S., and D.S.G. acknowledge financial support from the European Research Council Synergy grant ERC-SyG-2013-610028 IMBALANCE-P. S.L. and P.F. have received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 821003 (project CCiCC/4C, Climate-Carbon Interactions in the Coming Century), and S.L. received funding from SNSF (grant no. 20020_172476). H.T. is supported by U.S. National Science Foundation grants (1903722 and 1243232). M.F.-M. is a postdoctoral fellow of the Research Foundation-Flanders (FWO). W.K.S. acknowledges support from NASA Terrestrial Ecosystems Grant 80NSSC19M0103. N.Z. is supported by U.S. NOAA (NA18OAR4310266). A.W. is supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. Y.Z. acknowledges support from Introducing Talents of Discipline to Universities (Resource and Environmental Sciences) and CAS Interdisciplinary Innovation Team. **Author contributions:** Y.Z., S.W., and J.P. designed the research; S.W. performed the analysis; S.W., Y.Z., J. P., and W.J. drafted the paper; P.C., A.C., J.M.C., and J.S. contributed to the interpretation of the results and to the writing; I.A.J., M.W., J.A.B., E.C., M.F.-M., R.A., S.S., P.F., W.K.S., W.Y., W.H., D.L., M.K., D.Z., S.L., E.K., B.P., T.G.M.S., I.K., R.W., N.Z., H.T., N.V., A.K.J., A.W., V.H., and D.S.G. provided the data and contributed to the writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability**: All data acquired or used in this analysis are available from the links in table S1. The code is available at https://figshare.com/s/518a4bdcb14ae95dbcc7. The analysis of foliar nutrients was based on data that was collected by partners of the official UNECE ICP Forests Network (http://icp-forests.net/contributors) (data accessed on 30 October 2018). Part of the data was cofinanced by the European Commission.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1295/suppl/DC1
Materials and Methods
Supplementary Text S1 to S18
Figs. S1 to S47
Tables S1 to S5
References (*50–138*)
MDAR Reproducibility Checklist

31 March 2020; accepted 23 October 2020
10.1126/science.abb7772

SOLAR CELLS

Monolithic perovskite/silicon tandem solar cell with >29% efficiency by enhanced hole extraction

Amran Al-Ashouri[1]*, Eike Köhnen[1]*, Bor Li[1], Artiom Magomedov[2], Hannes Hempel[3], Pietro Caprioglio[1,4], José A. Márquez[3], Anna Belen Morales Vilches[5], Ernestas Kasparavicius[2], Joel A. Smith[6,7], Nga Phung[6], Dorothee Menzel[1], Max Grischek[1,4], Lukas Kegelmann[1], Dieter Skroblin[8], Christian Gollwitzer[8], Tadas Malinauskas[2], Marko Jošt[1,9], Gašper Matič[9], Bernd Rech[10,11], Rutger Schlatmann[5,12], Marko Topič[9], Lars Korte[1], Antonio Abate[6], Bernd Stannowski[5,13], Dieter Neher[4], Martin Stolterfoht[4], Thomas Unold[3], Vytautas Getautis[2], Steve Albrecht[1,11]†

Tandem solar cells that pair silicon with a metal halide perovskite are a promising option for surpassing the single-cell efficiency limit. We report a monolithic perovskite/silicon tandem with a certified power conversion efficiency of 29.15%. The perovskite absorber, with a bandgap of 1.68 electron volts, remained phase-stable under illumination through a combination of fast hole extraction and minimized nonradiative recombination at the hole-selective interface. These features were made possible by a self-assembled, methyl-substituted carbazole monolayer as the hole-selective layer in the perovskite cell. The accelerated hole extraction was linked to a low ideality factor of 1.26 and single-junction fill factors of up to 84%, while enabling a tandem open-circuit voltage of as high as 1.92 volts. In air, without encapsulation, a tandem retained 95% of its initial efficiency after 300 hours of operation.

A tandem solar cell, consisting of a silicon cell overlaid by a perovskite solar cell (PSC) (*1*), could increase efficiencies of commercial mass-produced photovoltaics beyond the single-junction cell limit (*1*, *2*) without adding substantial cost (*3*, *4*). The certified power conversion efficiency (PCE) of PSCs has reached up to 25.5% for single-junction solar cells (usual active area of ~0.1 cm^2) (*5*), 24.2% for perovskite/CIGSe (copper-indium-gallium-selenide) tandem cells (~1 cm^2) (*5*–*7*), 24.8% for all-perovskite tandem cells (0.05 cm^2) (*8*, *9*), and 26.2% for the highest openly published perovskite/silicon tandem efficiency (~1 cm^2) (*10*). Perovskite/silicon tandem cells have additionally undergone technological advances in both stability and compatibility with textured silicon substrates (*11*–*13*). However, these perovskite-based tandem solar cells still have room for improvement, as practical limits for all these tandem technologies are well above 30% (*14*, *15*).

The increase in PSC efficiency has been driven in part by advances in physical and chemical understanding of the defect and recombination mechanisms. Some reports presented near-perfect passivation of surfaces and grain boundaries, with photoluminescence quantum yields (PLQYs) approaching theoretical limits (*16*–*18*). Consequently, PSCs were reported with open-circuit voltage (V_{OC}) values of only a few tens of meV below their radiative limit (*19*–*23*). These values surpass those reached with crystalline silicon absorbers and are comparable with solar cells based on epitaxially grown GaAs (*23*, *24*). However, perovskite compositions with a wider bandgap that are needed for high-efficiency tandem solar cells still show considerable V_{OC} losses (*14*, *25*). The main reasons include comparably low PLQYs of the absorber material itself, an unsuitable choice of selective contacts, and phase instabilities. Even state-of-the-art perovskite/silicon tandem cells still have V_{OC} values well below 1.9 V.

We present a strategy to overcome these issues simultaneously, demonstrated with a triple-cation perovskite composition with a bandgap of 1.68 eV, which enables photostable tandem devices with a V_{OC} of 1.92 V. We note that the charge extraction efficiency, and hence the fill factor (FF), of PSCs is still poorly understood. Although reported PSCs usually feature a small active area (~0.1 cm^2) with small absolute photocurrents (a few milliamperes), and thus small series resistance losses at the contacts, typical FFs of high-efficiency devices generally range from 79 to 82%. However, on the basis of the detailed balance limit, PSCs should be able to deliver a FF of 90.6% at a bandgap of 1.6 eV. Wider-bandgap perovskite compositions near 1.7 eV seem especially prone to low FFs, resulting in tandem cell FF values commonly below 77%, near current-matching conditions (*11*, *12*, *26*). In optimized perovskite single-junction devices, the FFs only recently exceeded 80%, with a maximum value of 84.8% (*27*).

One reason for the low FF might be that there are only a few techniques for quantifying and analyzing the FF losses in PSCs. We show that intensity-dependent transient photoluminescence in combination with absolute photoluminescence is a viable technique for doing so. A main FF limitation of high-efficiency PSCs is the ideality factor n_{ID}, with typical values of 1.4 to 1.8 for high-V_{OC} devices (*28*), whereas established solar cell technologies reach values of 1 to 1.3 (*29*). Thus, an important goal for perovskite photovoltaics is to lower the ideality factor while minimizing nonradiative interface recombination to achieve a high V_{OC} (*28*). We designed a self-assembled monolayer (SAM) with methyl group substitution as a hole-selective layer, named Me-4PACz ([4-(3,6-dimethyl-9H-carbazol-9-yl)butyl]phosphonic acid) and show that a fast hole extraction went along with a lower ideality factor. Thus, FFs of up to 84% in p-i-n single-junction PSCs and >80% in tandem devices were achieved.

The SAM provided both fast extraction and efficient passivation at the hole-selective interface. This combination slowed light-induced halide segregation of a tandem-relevant perovskite composition with 1.68-eV bandgap, allowed a PLQY as high as on quartz glass, and led to high single-junction device V_{OC} values of >1.23 V. The single-junction improvements transferred into tandem devices, which allowed us to fabricate perovskite/silicon tandem solar cells with a certified PCE of 29.15%. This value surpasses the best silicon single-junction cell (26.7%) and is comparable to the best GaAs solar cell (*27*) at the same area of 1 cm^2. Under maximum power point (MPP) tracking in ambient air without encapsulation, a Me-4PACz tandem cell retained 95% of its initial efficiency after 300 hours. We used injection-dependent absolute electroluminescence (EL) spectroscopy to reconstruct the individual subcell current-voltage curves without the influence of series resistance (pseudo–*J*-*V* curves), which showed that the tandem device design that features only a standard perovskite film without additional bulk passivation could in principle realize PCE values up to 32.4%.

Stabilization of wide-bandgap perovskite with the hole-selective layer

The ideal top cell bandgap for perovskite absorbers in conjunction with CIGSe and Si bottom cells is ~1.68 eV (*30*–*32*). These wider-bandgap compositions often feature a Br/I ratio of >20%, which can lead to phase instabilities caused by light-induced halide segregation, most strikingly evident from photoluminescence (PL)

[1]Young Investigator Group Perovskite Tandem Solar Cells, Helmholtz-Zentrum Berlin, 12489 Berlin, Germany. [2]Department of Organic Chemistry, Kaunas University of Technology, Kaunas LT-50254, Lithuania. [3]Department of Structure and Dynamics of Energy Materials, Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, 14109 Berlin, Germany. [4]Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany. [5]PVcomB, Helmholtz-Zentrum Berlin, 12489 Berlin, Germany. [6]Young Investigator Group Active Materials and Interfaces for Stable Perovskite Solar Cells, Helmholtz-Zentrum Berlin, 12489 Berlin, Germany. [7]Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, UK. [8]Physikalisch-Technische Bundesanstalt, 10587 Berlin, Germany. [9]Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia. [10]Scientific Management, Helmholtz-Zentrum Berlin, 12489 Berlin, Germany. [11]Faculty of Electrical Engineering and Computer Science, Technical University Berlin, 10587 Berlin, Germany. [12]HTW Berlin–University of Applied Sciences, 12459 Berlin, Germany. [13]Beuth University of Applied Sciences Berlin, 13353 Berlin, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: steve.albrecht@helmholtz-berlin.de

spectra that show a double-peak formation under continuous illumination (*33*, *34*). Upon generation of charge carriers in the perovskite film, iodide-rich clusters can form that are highly luminescent because they serve as charge carrier sinks, given their lower bandgap relative to the surrounding material (*35*). As quantified by Mahesh *et al*., although some portion of the V_{OC} loss is related to halide segregation, the dominant source of V_{OC} loss is likely the generally low optoelectronic quality of the Br-rich mixed-halide perovskite absorbers, or high nonradiative recombination rates at their interfaces (*35*). Hence, to unambiguously determine the limitations and potentials of wide-bandgap compositions, it is necessary to find suitable charge-selective contacts that do not introduce further losses or instabilities.

We show that fast charge extraction paired with surface passivation can effectively suppress the formation of a double-peak emission in the PL, indicative of phase stabilization, and simultaneously enable a high quasi–Fermi level splitting (QFLS) and device performance. Rather than optimizing the perovskite composition or passivating the film, we chose a variant of the widely used Cs-, FA-, and MA-containing "triple-cation" perovskite (*36*) that is highly reproducible (FA, formamidinium; MA, methylammonium) and focused on preparing an optimal charge-selective contact on which the perovskite film was deposited. We enlarged the bandgap by increasing the Br/I ratio to obtain a 1.68-eV (23% Br) absorber instead of the commonly used 1.60 to 1.63 eV (~17% Br), yielding a nominal precursor composition of $Cs_{0.05}(FA_{0.77}MA_{0.23})_{0.95}Pb(I_{0.77}Br_{0.23})_3$.

A schematic of the device stack and the hole-selective layers (commonly abbreviated as HTLs, "hole-transporting layers") used for PL measurements is shown in Fig. 1. We first compared the QFLS measured by absolute PL and then the PL stability of this perovskite composition prepared on indium tin oxide (ITO) substrates covered by the HTLs. In recently published high-PCE p-i-n ("inverted") single-junction and tandem PSCs, the polymer poly[bis(4-phenyl)(2,4,6-trimethylphenyl)amine] (PTAA) or the comparable poly[*N*,*N*′-bis(4-butylphenyl)-*N*,*N*′-bis(phenyl)-benzidine] (polyTPD) is typically used (*10*, *11*, *37*, *38*). Alternatively, SAMs based on carbazole, such as MeO-2PACz and 2PACz, can form passivated interfaces while allowing for low transport losses because they are ultrathin (<1 nm) (*7*). The introduction of a methyl-group substitution to the "lossless" hole-selective interface created by 2PACz (*7*) led to a more optimized alignment with the perovskite valence band edge (see energetic band edge diagram in fig. S1) with a similar dipole moment (~1.7 D) and resulted in faster charge extraction. The supplementary materials contain the synthesis scheme of the SAMs we used. In the literature concerning the n-i-p configuration of PSCs, methoxy substituents are prevalent in HTLs, with some reports of a possible passivation function at the perovskite interface (*39*–*42*). For the p-i-n configuration, however, the standard high-performance HTLs PTAA and polyTPD contain alkyl substituents. In the present study, we directly compared methoxy and methyl substituents in p-i-n cells with MeO-2PACz and Me-4PACz, with the results showing advantages for the methyl substitution with respect to both passivation and hole extraction. We tested the influence of the aliphatic chain length (n) in carbazole-based SAMs without (nPACz) and with methyl substitution (Me-nPACz) on PSC performance and found an optimum FF at n = 2 for nPACz and n = 4 for Me-nPACz (see fig. S23). For n = 6, both SAMs led to current-voltage hysteresis.

The QFLS values of bare perovskite films (Fig. 1B) deposited on 2PACz and Me-4PACz were similar to that on quartz glass, commonly regarded as a perfectly passivated substrate (*16*). Perovskite compositions with high Br content typically segregate into I-rich phases indicated by increased PL intensity at lower photon energies, here at a wavelength of 780 nm (*33*). Pristine regions of the nonsegregated perovskite film emitted photons at a peak wavelength of ~740 nm for perovskite deposited on glass (Fig. 1C) or ITO/PTAA (Fig. 1D), and a similar response was seen for the SAM MeO-2PACz (fig. S3) on ITO. However, the perovskite emission was more stable over time on ITO/2PACz and ITO/Me-4PACz substrates (Fig. 1E and fig. S3). The raw spectra are shown in fig. S4.

Among the studied HTLs, phase segregation was inhibited only if the perovskite was grown on a substrate that fulfilled the requirements of both fast charge extraction and good passivation; Fig. 1F shows that passivation alone was insufficient. The black curve shows a PL spectrum of the perovskite film on an insulating glass substrate that was covered by Me-4PACz after 10 min of continuous spot illumination with 1-sun equivalent photon flux. The illuminated film showed signs of I-rich phases emitting at a center wavelength of ~780 nm. The glass substrate ensured that no hole transfer out of the perovskite bulk occurred. In contrast, a conductive ITO substrate that allowed hole transmission in combination with Me-4PACz increased the PL stability, as evidenced by the sharp peak with emission centered at ~740 nm even after 10 min of spot illumination.

A bare ITO substrate seemed to prevent charge accumulation as well, allowing a stable PL peak position at 1-sun intensity (spot size 0.12 cm^2; see fig. S6). The connection between charge accumulation in the perovskite and phase instability was reported in previous studies in which a reduced density of carriers increased the activation energy of mobile ion species and allowed the film to remain in its initial form (*43*, *44*). Spot illumination (0.12 cm^2 with 1-sun photon flux) represented increased stress testing on phase stability compared to full illumination because it created an outward driving force for ions from the illuminated area (*45*). Consequently, a smaller illumination spot (i.e., larger edge-to-area ratio) at the same illumination intensity showed a faster PL redshift (see figs. S5 and S6). To compare the degree of PL redshift and double-peak formation, we evaluated the ratio of the two emission center intensities at 740 and 780 nm for two different excitation fluences equivalent to 1-sun and 30-sun illumination (Fig. 1, G and H). At 1-sun–equivalent intensity, only 2PACz and Me-4PACz on ITO had a stable ratio. However, upon increasing the intensity and thus the charge carrier generation rate by a factor of 30, a Me-4PACz–covered ITO substrate differed from the 2PACz–covered substrate by still displaying a similarly stable PL intensity ratio.

We used transient photoluminescence (TrPL) to analyze charge carrier transfer into adjacent charge-selective layers (*46*). The full decay is governed by nonradiative, trap-assisted surface/bulk recombination (mostly monoexponential decay), radiative recombination ("bimolecular," second-order decay), and charge transfer effects, which can be disentangled if these time constants differ sufficiently from each other (*18*). Figure 2A presents PL transients of 1.68 eV–bandgap perovskite films on ITO/HTL substrates. With MeO-2PACz and PTAA, it was not possible to clearly differentiate between charge extraction and trap-assisted recombination because the nonradiative recombination was high (as evidenced by lower QFLS values relative to quartz glass; Fig. 1B) and because the transients did not saturate toward one process. In contrast, the PL transients for 2PACz and Me-4PACz showed a clear monoexponential decay at later times, indicating Shockley-Read-Hall recombination (*47*). Fits to the TrPL transients (fig. S8) were used to compute the differential lifetime $\tau = -\{d\ \ln[\phi(t)]/dt\}^{-1}$ (Fig. 2B), where $\phi(t)$ is the time-dependent PL photon flux. In this representation, the processes that reduce the PL counts over time are separable, and the transient decay time (or "lifetime") is directly readable at each time point (*46*).

The asymptotically reached high TrPL lifetimes of >5 μs for both 2PACz and Me-4PACz suggests that there were minimal nonradiative recombination losses at the SAM interfaces. The charge transfer process at early times (until ~1 μs) led to a sharp rise of τ, resembling simulated curves by Krogmeier *et al*. (*46*). The transition from increasing lifetime to the plateau marks the end of charge transfer, and nonradiative first-order recombination becomes dominant. Because PLQY measurements of films on 2PACz and Me-4PACz indicated a similar level of interface recombination under

the same charge generation conditions (see also fig. S9), the steepness of this rise was influenced by the charge transfer speed. The observed gradient for Me-4PACz implied a faster hole transfer to the underlying ITO relative to 2PACz, with the saturation starting after ~300 ns rather than ~1 µs.

In the charge carrier generation regime of this experiment (~1 sun, ~3 × 10^{15} cm^{-3}), trap-assisted recombination dominated, with the PL flux scaling proportionally to the density of photogenerated carriers n, as evidenced by intensity-dependent TrPL shown in fig. S9. Figure S9 further demonstrates that at higher generation conditions, the PL flux scaled proportionally to n^2, where transients usually show a multiexponential signature, as seen with 2PACz and quartz (fig. S10). Nonetheless, in this regime the Me-4PACz transients remained monoexponential until generation densities exceeded ~35 suns equivalent. We interpret this as a consequence of a large hole-extraction flux, which causes first-order recombination to dominate even in this injection regime.

We quantify this phenomenon of persisting domination of first-order recombination in Fig. 2C by displaying the ratio of higher-order

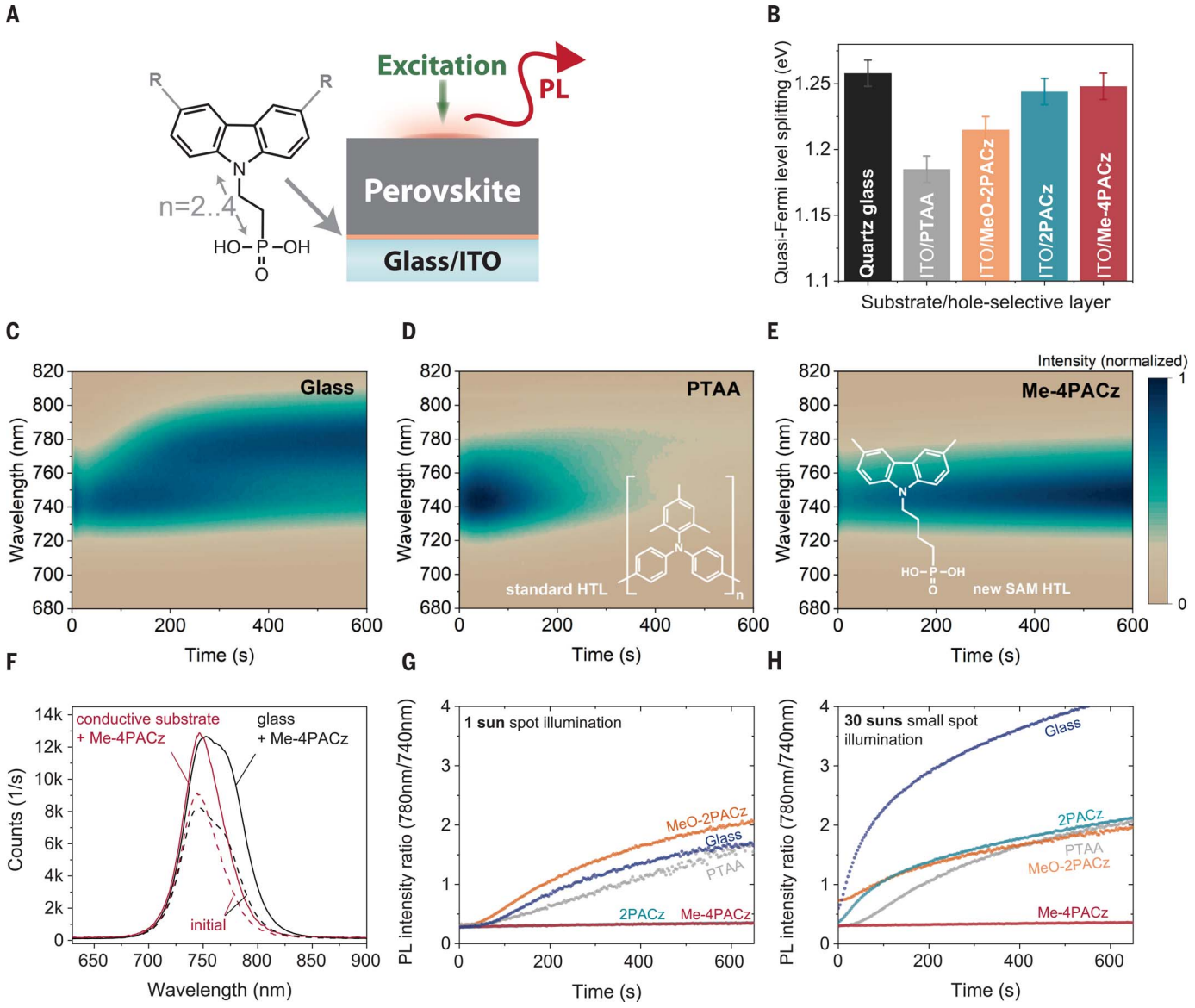


Fig. 1. Photoluminescence properties and stability assessment of perovskite films on different substrates. (**A**) Schematic description of the photoluminescence (PL) experiment and chemical structure of a general carbazole-based SAM, with R denoting a substitution, which in this work is either nothing (2PACz), a methoxy group (MeO-2PACz), or a methyl group (Me-4PACz). The number 2 or 4 denotes the number of the linear C atoms between the phosphonic acid anchor group and the conjugated carbazole main fragment. (**B**) Quasi–Fermi level splitting (QFLS) values of nonsegregated 1.68-eV bandgap perovskite films on a bare glass substrate and different hole-selective layers on the transparent and conductive indium tin oxide (ITO). Error bars denote the global error of the evaluation method (~20 meV). (**C** to **E**) Time-dependent photoluminescence spectra analyzing phase stability of perovskite absorbers with 1.68-eV bandgap. The perovskite films were deposited either on glass (C) or on ITO substrates with different hole-selective layers [(D) and (E)]. The color scale is at the far right. (**F**) PL spectra before (dashed lines) and after 600 s of light-soaking (solid lines) under 1-sun equivalent illumination in air, comparing the perovskite grown on Me-4PACz that had been deposited on a glass substrate and a conductive ITO substrate. (**G**) Ratio of PL intensities at 780 nm (I-rich domains) and 740 nm (neat perovskite) from the PL evolutions in (C), (D), (E), and two other hole-selective layers (see fig. S4; illumination spot size ~0.12 cm^2), shown as a figure of merit for phase stability. (**H**) Ratio of PL intensities as in (G), but at higher illumination intensity through decrease of the excitation spot size to 0.4 mm^2.

to first-order recombination for the different generation conditions (see supplementary text for the evaluation method). Comparison of Me-4PACz to 2PACz indicates that the hole-extraction flux of Me-4PACz was larger by a factor of >10, because the curvature of the TrPL transient only begins to resemble that of 2PACz at a factor of >10 higher generation density (indicated by the blue dashed line in Fig. 2C).

The carrier mobilities determined by optical pump terahertz probe measurements (fig. S12) were similar between perovskite films grown on the different HTLs. To also exclude differences in perovskite composition and crystal orientation due to possible growth differences, we probed the effect of the HTL on these properties by grazing-incidence wide-angle x-ray scattering at the four-crystal monochromator beamline of the Physikalisch-Technische Bundesanstalt (*48*). Azimuthally integrated diffraction patterns collected on a movable PILATUS detector module (*49*) showed comparable composition in each case (fig. S13), with marginally increased PbI_2 scattering intensity on PTAA as we observed in our previous work (*7*). Comparing azimuthal intensity profiles for perovskite scattering features (fig. S14), we found a negligible difference in crystallographic orientation between the samples.

Our complete solar cells were capped by C_{60} as the electron-selective contact. The electron extraction speed did not limit the cell operation, as demonstrated by a time-resolved terahertz photoconductivity measurement combined with TrPL on a quartz/perovskite/C_{60} sample (fig. S11). We compared the decays of free charge carriers after interface-near carrier generation on both sample sides and found an electron transfer time constant of ~1 ns, substantially faster than hole transfer at the hole-selective layer (in the range of ~100 ns). The extraction velocity into the C_{60} in our model was 1.6×10^4 cm/s (see fig. S11 for details), a value similar to earlier reported velocities (*46*).

Performance of perovskite single-junction solar cells

For analysis at the solar cell level, we focused on the simple single-junction device stack glass/ITO/HTL/perovskite/C_{60}/SnO_2/Ag, with the SnO_2 serving as a buffer layer for indium zinc oxide (IZO) sputtering in the fabrication of tandem solar cells (*50*). We found that the combination of fast charge extraction and passivated interface not only mitigated phase instability (see Fig. 1) but was also linked to an increased FF of solar cell devices, mainly by a decreased diode ideality factor of the PSCs. The FF is the major remaining parameter for which PSCs have not yet come close to the values of established solar cell technologies (*24*, *51*) (see fig. S16 for FF comparisons), with the ideality factor being one of the main properties that limit high-efficiency PSCs (*29*). MeO-2PACz and 2PACz led to FFs of up to 82% (Fig. 3A), whereas with Me-4PACz the values were as high as 84%, representing ~93% of the radiative limit.

Figure 3B shows *J-V* curves recorded at simulated AM1.5G illumination conditions, comparing champion PTAA and Me-4PACz cells of the same batch and showing the superior performance of the SAM. The ideality factors n_{ID} for PSCs with different HTLs (Fig. 3C and Table 1) were ~1.26 for Me-4PACz, ~1.42 for 2PACz, 1.51 for MeO-2PACz, and ~1.55 for PTAA cells. Figure S20 compares the V_{OC} values achieved with the different HTLs. Despite the large differences in passivation at the hole-selective interface, the differences in V_{OC} were not as large (average difference of 30 mV between PTAA and Me-4PACz) because of the limiting nonradiative recombination at the C_{60} interface. However, as reasoned above, the C_{60} layer did not limit charge extraction, hence the different extraction speeds invoked by the HTLs directly influenced the FF values. The high FF with Me-4PACz was accompanied by high V_{OC} values of up to 1.16 V; when a LiF interlayer was placed between the perovskite and C_{60}, we achieved a maximum voltage of 1.234 V (*52*, *53*) (Fig. 3D and fig. S20). The combination of a high V_{OC} with low n_{ID} was previously considered as challenging for PSCs (*28*), and it allowed us to fabricate a perovskite single junction with a PCE of 20.8% with Me-4PACz (fig. S18) and a perovskite bandgap of 1.68 eV.

To investigate the FF values without the influence of series resistance losses, we measured intensity-dependent absolute PL spectra and computed the QFLS values [or implied V_{OC} (iV_{OC})] as a function of the illumination intensity. The derived data pairs of generation currents and iV_{OC} values allowed the reconstruction of hypothetical, so-called pseudo–*J-V* curves, as recently shown in (*54*) (Fig. 3D). The extracted FF and pseudo-FF values (FF in absence of transport losses) of bare perovskite films grown on different HTLs are summarized in Table 1, row 1. Both 2PACz and Me-4PACz enabled high "pseudo-FF" (pFF) values of ~88%, which is 96.8% of the detailed balance limit and similar to the value achieved on a bare quartz substrate. PTAA allowed for a pFF of only 85.6%.

This analysis highlights how the SAMs formed a practically lossless interface between ITO and perovskite. Interestingly, when including a C_{60} layer on top of the perovskite film, no

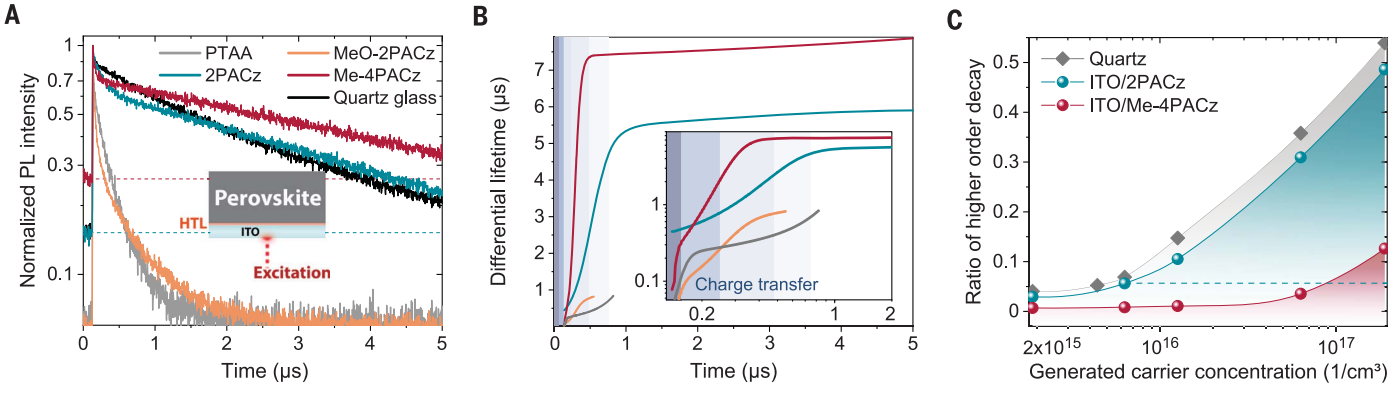


Fig. 2. Role of charge transfer in transient photoluminescence (TrPL). (**A**) PL transients of perovskite on ITO/hole-selective layer substrates. The dashed lines indicate the background levels. (**B**) Computed differential lifetimes from fits to the transients in (A), showing the single-exponential decay time at each time of the transient, with early times shown in the inset. The inset highlights the region of the Me-4PACz and 2PACz transients that is governed by hole transfer into the ITO. Excitation density is similar to 1-sun conditions (fluence of ~30 nJ/cm^2, 2×10^{15} to 3×10^{15} cm^{-3}). Colors are as in (A). The shaded areas are a guide marking the approximate time domain in which the Me-4PACz transient is governed by charge transfer. (**C**) Ratio of higher-order processes to monoexponential decay in the TrPL transients, revealing that Me-4PACz not only extracts holes faster [inset in (B)] but does so at ~10 times the efficiency of 2PACz, because the Me-4PACz transient shows the same magnitude of radiative recombination only with charge carrier generation that is higher by a factor of ~10 (comparison along the dashed line; see fig. S10 for details).

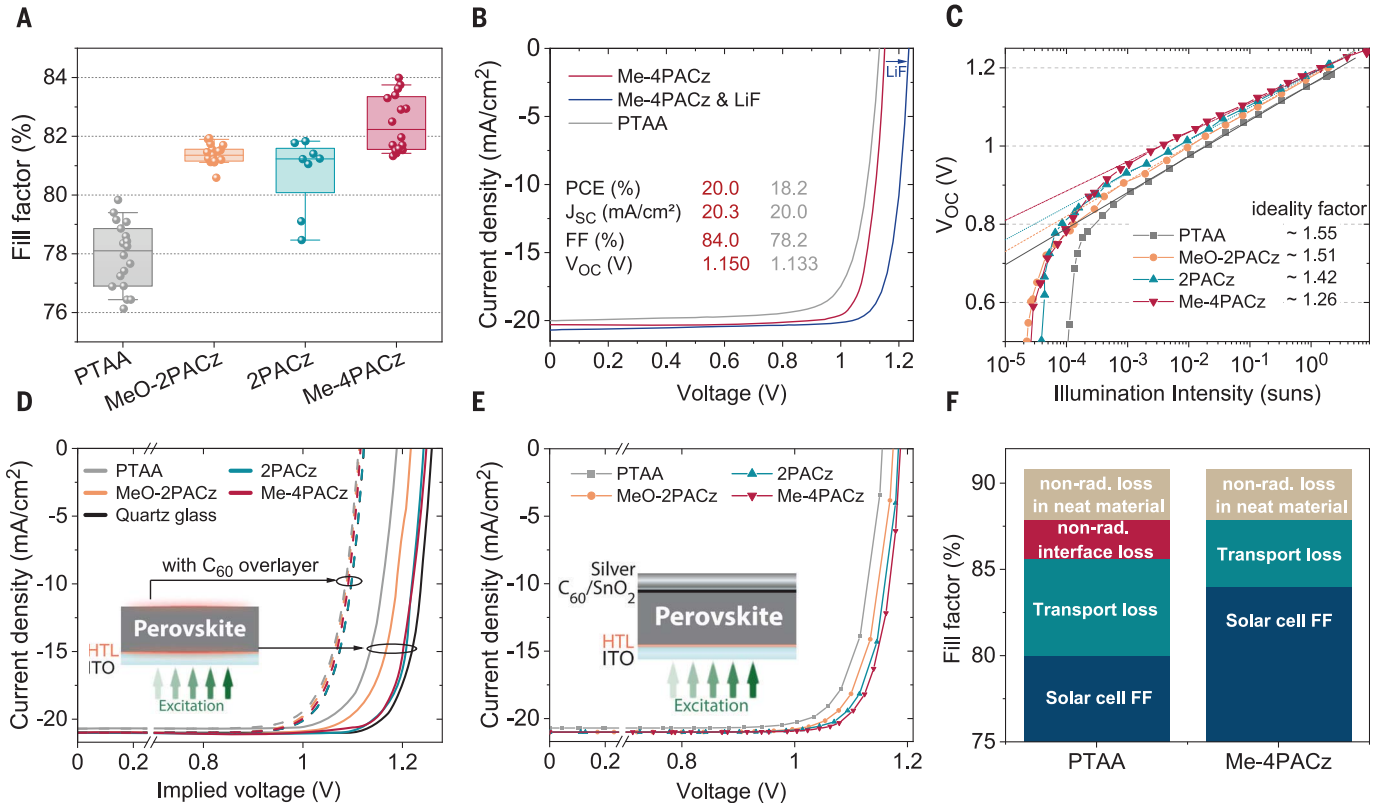


Fig. 3. Performance and fill factor loss analysis of p-i-n solar cells with different hole-selective layers. (**A**) Comparison of fill factor values of PSCs with the stack glass/ITO/HTL/perovskite/C_{60}/SnO_2/Ag, triple-cation perovskite absorber with 1.68-eV bandgap. All data are from cells made from the same perovskite precursor and contact processing batch. The boxes indicate the 25/75 percentiles; the whiskers indicate the 10/90 percentiles. (**B**) *J-V* curves of the best cells of the batch in (A) and a *J-V* curve of a Me-4PACz cell from another batch with LiF interlayer between C_{60} and perovskite, reaching a V_{OC} of 1.234 V. (**C**) Intensity-dependent open-circuit voltage V_{OC} with linear fits (dashed lines). (**D**) Pseudo–*J-V* curves reconstructed from intensity-dependent absolute PL measurements on the illustrated sample stack. The 2PACz and Me-4PACz curves almost coincide; the dashed lines represent pseudo–*J-V* curves from the sample variations including the electron-selective C_{60} layer, with which all curves are comparable because of the limiting nonradiative recombination at the C_{60} interface. (**E**) Pseudo–*J-V* curves reconstructed from the measurements in (C). Table 1 summarizes the FF values extracted from the pseudo–*J-V* curves. (**F**) Repartition of loss mechanisms lowering the cell's FF below the detailed balance limit, comparing PTAA and Me-4PACz cells: nonradiative loss in neat material (= radiative FF limit minus pFF of neat film), nonradiative interface loss (= pFF of neat film minus pFF of full cell), and transport loss (= pFF of full cell minus FF of measured solar cell).

Table 1. Comparison of "pseudo" fill factors (pFF) and implied open-circuit voltages (iV_{OC}). The values were derived from suns-PL and suns-V_{OC} measurements for our perovskite film on all studied hole-selective layers and on quartz glass. The table also shows the maximum FF attained in *J-V* measurements (max FF) (see also Fig. 3). "Half cell" refers to substrate/HTL/absorber, whereas "full cell" denotes the complete solar cell including C_{60}, SnO_2 and Ag metal electrode.

	Quartz glass	PTAA	MeO-2PACz	2PACz	Me-4PACz
pFF (%), half cell (suns-PL)	87.9	85.6	85.5	88.3	87.5
pFF (%), half cell + C_{60} (suns-PL)	85.3	85.3	85.3	85.3	85.3
pFF (%), full cell (suns-V_{OC})		85.8	85.9	86.9	87.9
max FF (%), full cell (*J-V*)		79.8	81.9	81.8	84.0
iV_{OC} (V), half cell (absolute PL)	1.258	1.185	1.215	1.244	1.248
n_{ID}, full cell (suns-V_{OC})		1.55	1.51	1.42	1.26

differences between the studied HTLs for the iV_{OC} and pFF were apparent (Fig. 3D, dashed lines; Table 1, row 2), as the C_{60} layer sets an iV_{OC} limitation through high nonradiative recombination rates (*53*). This limitation was only overcome with a counter electrode on the C_{60} (Fig. 3E and full devices), which underscores the role of the dipoles that Me-4PACz and 2PACz created at the ITO surface. The calculated molecular dipole value of the hole-transporting fragment is ~0.2 D for MeO-2PACz, ~1.7 D for Me-4PACz, and ~2 D for 2PACz. The positive dipoles shifted the work function of the ITO toward higher absolute numbers (fig. S2A), which presumably resulted in a higher built-in potential throughout the device (*55*, *56*). A well-defined built-in potential can exist with the presence of a second electrode countering the ITO—in this case, Ag or Cu. Thus, when reconstructing the *J-V*s from the suns-V_{OC} measurement on full devices in Fig. 3C to extract the pFF (Table 1, row 3), both 2PACz and Me-4PACz overcame the pFF and iV_{OC} limitations imposed by the C_{60} layer (Fig. 3E).

The differences between the electrical *J-V* curves (max. FF 84%) in Fig. 3B and pseudo–*J-V* curves (max. FF ~88%) arose from transport losses caused by the finite mobility of the C_{60}, non-optimized sample design, and ITO sheet resistance, as well as the measurement setup. Figure 3F summarizes a comparison of the different contributions to FF losses for PTAA

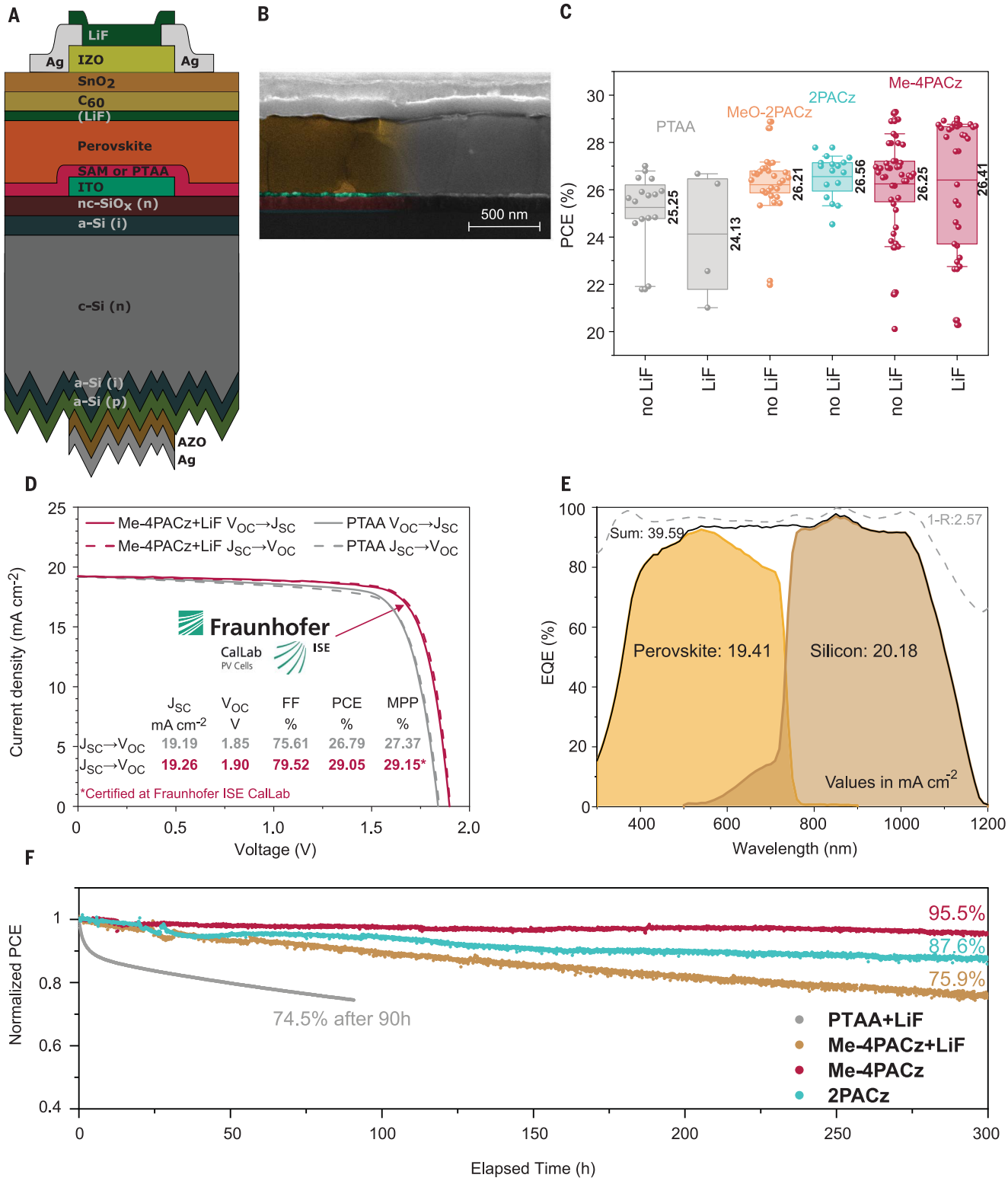


Fig. 4. Characteristics of monolithic perovskite/silicon tandem solar cells using various HTLs. (**A**) Schematic stack of the monolithic perovskite/silicon tandem solar cell. (**B**) SEM image of a tandem cross section with Me-4PACz as HTL. (**C**) Statistics of the PCE of PTAA, MeO-2PACz, 2PACz, and Me-4PACz tandem solar cells from *J-V* scans. (**D**) Certified *J-V* curve measured at Fraunhofer ISE, including the MPP value and the device parameters (red), in comparison to a tandem cell with PTAA (gray) as HTL measured in-house. (**E**) External quantum efficiency (EQE) and reflection (denoted as 1-R) of the certified tandem cell measured in-house. The AM1.5G-equivalent current densities are given. (**F**) Long-term MPP track using a dichromatic LED illumination of nonencapsulated solar cells in air at a controlled temperature of 25°C and relative humidity of 30 to 40%. The data are normalized to the MPP average of the first 60 min of each individual track to account for measurement noise. Because of the fast degradation, the MPP track of the PTAA + LiF cell is normalized to the first recorded value. The legend specifies each HTL and notes whether a LiF interlayer was used.

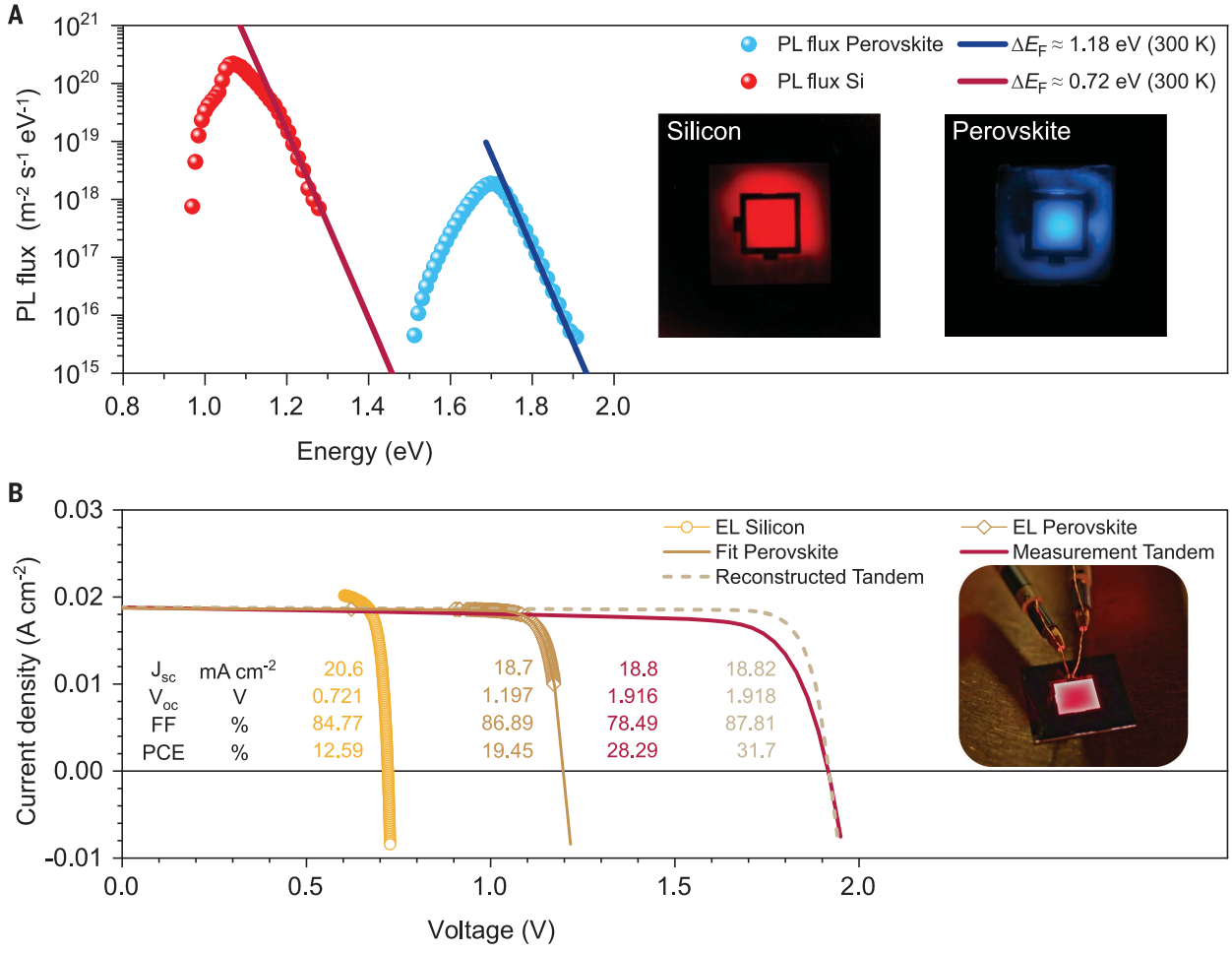


Fig. 5. Luminescence subcell analysis of a tandem solar cell with Me-4PACz and LiF interlayer. (**A**) Absolute PL spectra of the subcells recorded under 1-sun equivalent illumination. The excitation wavelengths are 455 nm and 850 nm for the perovskite and silicon subcell, respectively. PL images constructed from the integrated PL fluxes are also shown. The edge length of the active area (inner square) is 1 cm. (**B**) Reconstructed *J-V* curves calculated from injection-dependent electroluminescence (EL) measurements (open symbols) and shifted by the photogenerated current density. Furthermore, the perovskite subcell is fitted with a single-diode model (solid brown line). The reconstructed tandem *J-V* (dashed line) was calculated by adding the voltages of the subcells for each current density. The *J-V* measurement under simulated 1-sun illumination of this cell is shown as a solid red line. Furthermore, a photograph of the tandem solar cell at high injection current is shown. Due to a bandgap of 1.68 eV, the subcell emits light in the visible wavelength range and thus, the emission is visible by eye and with a regular digital camera.

and Me-4PACz, derived from comparisons of the pseudo–*J-V* curves to the measured *J-V* curves and radiative limits, as previously reported by Stolterfoht *et al.* (*54*). In addition to nonradiative losses at the PTAA interface (red), the film thickness (~10 nm, versus <1 nm with a SAM) and low conductivity of the PTAA led to greater transport losses than with Me-4PACz.

Integration into monolithic perovskite/silicon tandem solar cells

Efficient passivation in combination with fast hole extraction of Me-4PACz in perovskite single junctions could be transferred into monolithic tandem solar cells, which led to higher FF, V_{OC}, and stability. A schematic stack of this solar cell is shown in Fig. 4A. We used a silicon heterojunction solar cell as the bottom cell (*26*), based on a 260-μm-thick n-type float-zone Si wafer processed as described in the supplementary materials. The textured rear side enhanced the near-infrared (NIR) absorption, whereas the polished front side enabled the deposition of spin-coated perovskite. The 20-nm ITO recombination layer also served as the anchoring oxide for the SAMs (*7*). The top cell, with the same 1.68-eV perovskite bandgap and nominal precursor composition $Cs_{0.05}(FA_{0.77}MA_{0.23})_{0.95}Pb(I_{0.77}Br_{0.23})_3$ as analyzed above, formed the single-junction stack of ITO/HTL/perovskite/(LiF)/C_{60}/SnO_2/IZO/Ag/LiF. Figure 4B shows a scanning electron microscopy (SEM) cross-section image of a part of the tandem solar cell; no obvious differences were observed between perovskite films on the different HTLs (fig. S24). The molecular SAM cannot be resolved with SEM. Figure S25 shows a photograph and layout of the tandem device.

Figure 4C compares the PCE of tandem solar cells based on PTAA, MeO-2PACz, 2PACz, and Me-4PACz, with and without a LiF interlayer at the perovskite/C_{60} interface. With PTAA, the LiF interlayer led to rapid degradation of the cells (see fig. S26 for individual parameters). Without the interlayer, we achieved an average PCE of 25.25%. In contrast, the average efficiency of MeO-2PACz and 2PACz was 26.21% and 26.56%, respectively. The use of a LiF interlayer for Me-4PACz cells increased the V_{OC} but reduced the FF. Thus, both configurations reached a similar average PCE of 26.25% and 26.41%, respectively. However, with Me-4PACz the maximum PCEs (<29%) are higher than cells with 2PACz, mainly because of higher FF of up to 81%. These high FF values were achieved despite almost all cells being perovskite-limited (table S1). The statistics of all photovoltaic parameters

are shown in fig. S26. The *J-V* measurements of the champion cells of each configuration are shown in fig. S28; the PV parameters are summarized in table S2.

The tandem solar cells did not reach FF values comparable to those in single-junction cells because of the larger active area (1 cm^2) and a transparent conductive oxide (TCO) without grid fingers, leading to increased series resistance. The cells showed very high V_{OC} values of up to 1.92 V (fig. S30). With a V_{OC} of ~715 mV from the bottom cell at half illumination (fig. S31), the contribution of the perovskite subcell was ~1.2 V. Figure 4D shows a direct comparison between champion PTAA and Me-4PACz tandem cells; besides the 50-mV improvement in V_{OC}, the enhanced hole extraction boosted the FF by ~4% absolute.

We sent a tandem cell with Me-4PACz and a LiF interlayer to Fraunhofer ISE CalLab for independent certification (Fig. 4D; see fig. S32 for certificate). With a V_{OC} of 1.90 V, FF of 79.4%, and a short-circuit current density J_{SC} of 19.23 mA cm^{-2}, the cell had a PCE of 29.01% when measuring from J_{SC} to V_{OC}, similar to our in-house measurement (fig. S33) and was certified at the MPP with a PCE of 29.15% with a designated area of 1.064 cm^2. This PCE surpasses other monolithic (*10*, *27*) and four-terminal perovskite-based tandem solar cells (*57*) and is on par with the best GaAs single cell with the same active area (*27*).

Figure 4E shows the external quantum efficiency (EQE) of the certified tandem cell. Under AM1.5G-equivalent illumination conditions, the photogenerated current densities J_{ph} in the perovskite and silicon subcells were 19.41 mA cm^{-2} and 20.18 mA cm^{-2}, respectively, which agreed with the measured J_{SC} of 19.23 mA cm^{-2}. The tandem solar cell exhibited a nonideal current mismatch of 0.77 mA cm^{-2}, and even though the perovskite cell sets the slope around 0 V, the cell reached a FF of 79.5%. The cumulative photogenerated current density and loss caused by reflection were 39.59 mA cm^{-2} and 2.57 mA cm^{-2}, respectively. A comparison of EQEs and reflection losses between a cell of this work (planar front side) and a fully textured cell by Sahli *et al.* (*58*) is shown in fig. S34.

After the certification, we fabricated more Me-4PACz tandem solar cells without a LiF interlayer (fig. S26), which showed average performance similar to that with LiF. The champion cell showed a higher FF of 81% and lower V_{OC} of 1.87 V than without LiF. Together with a J_{SC} of 19.37 mA cm^{-2}, this led to a PCE of 29.29% and a stabilized efficiency of 29.32% (fig. S35).

We measured the stability of different nonencapsulated tandem solar cells (Fig. 4F). To track the degradation induced by either the top or the bottom cell more carefully, we developed a dichromatic LED setup using LEDs with center emission wavelengths of 470 nm and 940 nm (fig. S36) and with independent intensity calibration and recording. We adjusted the mismatch so that the J_{ph} in the individual subcells was equal to that measured under AM1.5G-equivalent illumination to maintain proper stability tracking of monolithic tandem solar cells (see below and supplementary text). The devices were measured under continuous MPP load (using voltage perturbation), at 25°C and in ambient air with 30 to 40% relative humidity. The photogenerated current densities of the subcells are given in table S3 and set which subcell is limiting. The degradation for a perovskite-limited tandem cell with Me-4PACz+LiF showed 75.9% of its initial efficiency (29.13%) after 300 hours. When we substituted the Me-4PACz with PTAA (perovskite-limited), the PCE decreased to 74.5% of its initial PCE (25.9%) after only 90 hours.

We additionally tracked a cell with Me-4PACz as HTL without a LiF interlayer to test the intrinsic stability of the HTL/perovskite combination. After 300 hours, the cell still operated at 95.5% of its initial PCE. Although the cells were current-matched, this track monitors a degradation of the perovskite, as it directly translates into the performance of the tandem cell and no degradation of the Si subcell is expected within these time scales. Our comparison strongly suggests that the use of a LiF interlayer reduces the stability. As described in other reports (*59*–*62*), the decrease in stability might be caused by deterioration of the electrodes and C_{60} interface upon migration of Li^+ and F^- ions. We note that it is important to declare the mismatch conditions because the use of a NIR-poor spectrum would lead to a Si-limited cell and thus to a higher stability (see supplementary text). Comparing this result to state-of-the-art stability tests of nonencapsulated tandem solar cells in ambient conditions, where the cells retained 90% of initial PCE after 61 hours (*58*) and 92% after 100 hours (*13*), our Me-4PACz tandem solar cell showed a superior operational stability.

In addition to the long-term stability measurements at 25°C, we conducted an MPP track of a Me-4PACz tandem cell at elevated temperatures. Following the procedure of Jošt *et al.*, the temperature was successively increased from 25° to 85°C and back to 25°C (*63*). There was no loss in PCE after this 200-min procedure, despite the high MA and Br amount of the wide-bandgap perovskite used here (fig. S39).

Subcell *J-V* characteristics of a monolithic tandem solar cell

One downside of monolithic multijunction solar cells is that the subcell characteristics are barely accessible. External quantum efficiency measurements are the only subcell-resolved measurements presented in almost all publications reporting multijunction solar cells. Here, we used absolute PL measurements in each subcell of a representative tandem solar cell (Me-4PACz + LiF). With this, we could estimate the QFLS, and thus the V_{OC} was accessible for both subcells independently. For this, we used hyperspectral absolute PL imaging at equivalent 1-sun conditions with an illumination spot larger than the area of the solar cells. The PL spectra and the integrated images are shown in Fig. 5A.

From the high-energy slope of the absolute PL spectra of the subcells, the individual implied V_{OC} values were calculated: 1.18 V for the perovskite subcell and 0.72 V for the Si subcell (*18*, *64*). From the PL spectra, we calculated the PLQY of both subcells, yielding values of 1.5% for Si and 0.02% for the perovskite. PLQY values exceeding 5% have already been demonstrated in perovskite single-junction devices for lower bandgaps (*19*).

To estimate the pseudo–*J-V* curves of the subcells, we performed absolute EL imaging, where the excess charge carriers are generated electrically to access the subcell characteristics (*65*–*68*). For each injected current, an EL image was recorded, from which the voltage of the subcells can be calculated from an average over the active area (fig. S40). With the reconstructed pseudo–*J-V* curves from injection current–dependent EL imaging, we analyzed the maximum possible efficiency of this cell stack with minimized charge transport losses (see supplementary materials for more details). We reconstructed both subcell *J-V* curves by calculating the implied voltage at each injected current, yielding a "pseudo" light–*J-V* (JV_{EL}) curve for each subcell after shifting it by the respective photogenerated current density J_{ph} calculated from EQE measurements; these J_{ph} values amounted to 18.7 and 20.6 mA cm^{-2} for the top and bottom cell, respectively. The open symbols in Fig. 5B show the measured EL data points averaged over the perovskite and silicon subcell and shifted by their respective J_{ph} values.

For the perovskite, we additionally fitted the data with a single-diode model to display the *J-V* curve over the whole voltage range, which was otherwise not accessible during the EL measurement. To obtain the tandem JV_{EL}, we added the voltages of the subcells for each current density. The dashed line shows the result. The reconstructed curve deviated from the electrically measured *J-V* curve under a solar simulator. This is mainly because EL gave access only to the internal voltage, whereas an electrical *J-V* curve displays the current density versus external voltage (which is affected by series resistances; see supplementary text). Hence, a high FF (87.8%) of the JV_{EL} can be regarded as the maximum achievable value for this particular tandem cell if the electrodes and all charge-selective layers were

free of series resistance losses. This would give a PCE of 31.7%, surpassing the theoretical PCE maximum of a silicon single cell (29.4%) (*69*). Thus, this cell stack has the capacity to overcome the 30% barrier through technical optimization of the contacts alone. However, by adjusting the mismatch conditions, even higher efficiencies are achievable.

To find the requirements for the highest efficiency, we fit the silicon subcell with a single-diode model. We conducted SPICE (Simulation Program with Integrated Circuit Emphasis) simulations to sweep the photogenerated current densities in the subcell. The single-diode models of the silicon and perovskite subcells were connected in series (schematically shown in fig. S42A), and the cumulative current density was fixed to 39.3 mA cm^{-2} (as calculated from EQE measurements for AM1.5G-equivalent illumination). Figure S42B shows the photovoltaic parameters as a function of the mismatch ($J_{ph,Si} - J_{ph,Pero}$). As shown in a previous publication, the V_{oc} is almost independent of the mismatch, whereas the FF is affected by it (*26*). A minimum FF occurs when the $J_{ph,Si}$ is 0.7 mA cm^{-2} below the $J_{ph,Pero}$. However, simultaneously the J_{SC} is highest under this condition. In a current-matching situation, the highest efficiency is 32.43%. This sets an estimation for the efficiency potential upon reduction of all transport losses. Decreasing the limitations set by the hole extraction speed, as shown in this work, is a viable path for exploring this potential.

REFERENCES AND NOTES

1. A. De Vos, *J. Phys. D* **13**, 839–846 (1980).
2. W. Shockley, H. J. Queisser, *J. Appl. Phys.* **32**, 510–519 (1961).
3. S. E. Sofia *et al*., *Sustain. Energy Fuels* **4**, 852–862 (2020).
4. Z. Li *et al*., *Joule* **2**, 1559–1572 (2018).
5. National Renewable Energy Laboratory, Best Research-Cell Efficiency Chart; www.nrel.gov/pv/cell-efficiency.html.
6. T. Todorov *et al*., *Adv. Energy Mater.* **5**, 1500799 (2015).
7. A. Al-Ashouri *et al*., *Energy Environ. Sci.* **12**, 3356–3369 (2019).
8. R. Lin *et al*., *Nat. Energy* **4**, 864–873 (2019).
9. G. E. Eperon *et al*., *Science* **354**, 861–865 (2016).
10. D. Kim *et al*., *Science* **368**, 155–160 (2020).
11. J. Xu *et al*., *Science* **367**, 1097–1104 (2020).
12. Y. Hou *et al*., *Science* **367**, 1135–1140 (2020).
13. B. Chen *et al*., *Joule* **4**, 850–864 (2020).
14. M. Jošt, L. Kegelmann, L. Korte, S. Albrecht, *Adv. Energy Mater.* **10**, 1904102 (2020).
15. Y. Jiang *et al*., *J. Mater. Chem. C* **4**, 5679–5689 (2016).
16. I. L. Braly *et al*., *Nat. Photonics* **12**, 355–361 (2018).
17. D. W. deQuilettes *et al*., *ACS Energy Lett.* **1**, 438–444 (2016).
18. T. Kirchartz, J. A. Márquez, M. Stolterfoht, T. Unold, *Adv. Energy Mater.* **10**, 1904134 (2020).
19. Z. Liu *et al*., *ACS Energy Lett.* **4**, 110–117 (2019).
20. P. Caprioglio *et al*., *Sustain. Energy Fuels* **3**, 550–563 (2019).
21. Q. Jiang *et al*., *Nat. Photonics* **13**, 460–466 (2019).
22. S. Yang *et al*., *J. Am. Chem. Soc.* **141**, 5781–5787 (2019).
23. L. Krückemeier, U. Rau, M. Stolterfoht, T. Kirchartz, *Adv. Energy Mater.* **10**, 1902573 (2020).
24. P. K. Nayak, S. Mahesh, H. J. Snaith, D. Cahen, *Nat. Rev. Mater.* **4**, 269–285 (2019).
25. A. Rajagopal, R. J. Stoddard, S. B. Jo, H. W. Hillhouse, A. K. Y. Jen, *Nano Lett.* **18**, 3985–3993 (2018).
26. E. Köhnen *et al*., *Sustain. Energy Fuels* **3**, 1995–2005 (2019).
27. M. A. Green *et al*., *Prog. Photovolt. Res. Appl.* **28**, 3–15 (2020).
28. P. Caprioglio *et al*., *Adv. Energy Mater.* **10**, 2000502 (2020).
29. M. A. Green, A. W. Y. Ho-Baillie, *ACS Energy Lett.* **4**, 1639–1644 (2019).
30. K. Jäger, L. Korte, B. Rech, S. Albrecht, *Opt. Express* **25**, A473–A482 (2017).
31. M. Jošt *et al*., *Energy Environ. Sci.* **11**, 3511–3523 (2018).
32. T. Leijtens, K. A. Bush, R. Prasanna, M. D. McGehee, *Nat. Energy* **3**, 828–838 (2018).
33. E. T. Hoke *et al*., *Chem. Sci.* **6**, 613–617 (2015).
34. D. J. Slotcavage, H. I. Karunadasa, M. D. McGehee, *ACS Energy Lett.* **1**, 1199–1205 (2016).
35. S. Mahesh *et al*., *Energy Environ. Sci.* **13**, 258–267 (2020).
36. M. Saliba *et al*., *Energy Environ. Sci.* **9**, 1989–1997 (2016).
37. M. Stolterfoht *et al*., *Energy Environ. Sci.* **12**, 2778–2788 (2019).
38. D. Luo *et al*., *Science* **360**, 1442–1446 (2018).
39. Q. Wang *et al*., *Adv. Energy Mater.* **9**, 1900990 (2019).
40. C. Huang *et al*., *J. Am. Chem. Soc.* **138**, 2528–2531 (2016).
41. A. Torres, L. G. C. Rego, *J. Phys. Chem. C* **118**, 26947–26954 (2014).
42. M. Saliba *et al*., *Nat. Energy* **1**, 15017 (2016).
43. Y. Lin *et al*., *Nat. Commun.* **9**, 4981 (2018).
44. B. Chen *et al*., *Adv. Mater.* **31**, e1902413 (2019).
45. N. Phung *et al*., *Adv. Energy Mater.* **10**, 1903735 (2020).
46. B. Krogmeier, F. Staub, D. Grabowski, U. Rau, T. Kirchartz, *Sustain. Energy Fuels* **2**, 1027–1034 (2018).
47. F. Staub *et al*., *Phys. Rev. Appl.* **6**, 044017 (2016).
48. M. Krumrey, G. Ulm, *Nucl. Instrum. Methods Phys. Res. A* **467–468**, 1175–1178 (2001).
49. D. Skroblin *et al*., *Rev. Sci. Instrum.* **91**, 023102 (2020).
50. K. A. Bush *et al*., *Nat. Energy* **2**, 17009 (2017).
51. J.-F. Guillemoles, T. Kirchartz, D. Cahen, U. Rau, *Nat. Photonics* **13**, 501–505 (2019).
52. J. Seo *et al*., *Energy Environ. Sci.* **7**, 2642–2646 (2014).
53. M. Stolterfoht *et al*., *Nat. Energy* **3**, 847–854 (2018).
54. M. Stolterfoht *et al*., *Adv. Mater.* **32**, 2000080 (2020).
55. N. Tessler, Y. Vaynzof, *ACS Energy Lett.* **5**, 1260–1270 (2020).
56. W. Wang *et al*., *Adv. Funct. Mater.* **30**, 1909755 (2020).
57. B. Chen *et al*., *Nat. Commun.* **11**, 1257 (2020).
58. F. Sahli *et al*., *Nat. Mater.* **17**, 820–826 (2018).
59. R. Quintero-Bermudez, J. Kirman, D. Ma, E. H. Sargent, R. Quintero-Torres, *J. Phys. Chem. Lett.* **11**, 4213–4220 (2020).
60. Z. Li *et al*., *Energy Environ. Sci.* **10**, 1234–1242 (2017).
61. S.-M. Bang *et al*., *ACS Energy Lett.* **5**, 1198–1205 (2020).
62. S. N. Habisreutinger *et al*., *Nano Lett.* **14**, 5561–5568 (2014).
63. M. Jošt *et al*., *Adv. Energy Mater.* **10**, 2000454 (2020).
64. T. Unold, L. Gütay, in *Advanced Characterization Techniques for Thin Film Solar Cells*, D. Abou-Ras, T. Kirchartz, U. Rau, Eds. (Wiley-VCH, 2011), pp. 151–175.
65. S. Chen *et al*., *Sci. Rep.* **5**, 7836 (2015).
66. D. Alonso-Alvarez, N. Ekins-Daukes, *IEEE J. Photovoltaics* **6**, 1004–1011 (2016).
67. S. Roensch, R. Hoheisel, F. Dimroth, A. W. Bett, *Appl. Phys. Lett.* **98**, 251113 (2011).
68. D. Hinken, K. Ramspeck, K. Bothe, B. Fischer, R. Brendel, *Appl. Phys. Lett.* **91**, 182104 (2007).
69. A. Richter, M. Hermle, S. W. Glunz, *IEEE J. Photovoltaics* **3**, 1184–1191 (2013).

ACKNOWLEDGMENTS

We thank M. Gabernig, C. Ferber, T. Lußky, H. Heinz, C. Klimm, and M. Muske at the Institute for Silicon Photovoltaics, Helmholtz-Zentrum Berlin (HZB), and T. Hänel, T. Henschel, M. Zelt, H. Rhein, K. Meyer-Stillrich, and M. Hartig at PVcomB (HZB) for technical assistance. A.A.-A. thanks A. Merdasa for his expertise during construction of the steady-state PL setup. Ei.K. and S.A. thank C. Wolff (University of Potsdam) and K. Brinkmann (University of Wuppertal) for fruitful discussion at the beginning of the project. A.M. acknowledges A. Drevilkauskaite for help with the synthesis of 4PACz and 6PACz materials. **Funding:** Supported by Federal Ministry for Education and Research (BMBF) grant 03SF0540 within the project "Materialforschung für die Energiewende"; the Federal Ministry for Economic Affairs and Energy (BMWi)–funded project ProTandem (0324288C); the HyPerCells graduate school; the Helmholtz Association within the HySPRINT Innovation lab project and TAPAS project; the Helmholtz Association via HI-SCORE (Helmholtz International Research School) (M.G., P.C., S.A., and D.N.); the European Union's Horizon 2020 research and innovation program under grant agreement 763977 of the PerTPV project; the Research Council of Lithuania under grant agreement S-MIP-19-5/SV3-1079 of the SAM project (A.M. and T.M.); Slovene Research Agency (ARRS) funding through research programs P2-0197 and J2-1727 (M.J., G.M., and M.T.); Deutsche Forschungsgemeinschaft projects 423749265 and 03EE1017C-SPP 2196 (SURPRISE and HIPSTER) (M.S., D.N., and S.A.); EPSRC and D. Lidzey for Ph.D. studentship funding via CDT-PV (EP/L01551X/1) (J.A.S.); and Erasmus+ (J.A.S.). **Author contributions:** A.A.-A., Ei.K., B.L., and S.A. planned the experiments, coordinated the work,and prepared the figures; Er.K., A.M., and T.M. designed and synthesized the Me-4PACz SAM and the (Me-)nPACz series; A.A.-A. and B.L. processed the single-junction cells and optimized the SAM deposition; Ei.K. and B.L. processed the tandem cells; A.B.M.V. processed the Si bottom cells; A.A.-A., H.H., and J.A.M. conducted and analyzed the PL experiments; J.A.M., A.A.-A., and Ei.K. performed the EL studies. H.H. recorded the terahertz measurements and performed the data analysis; P.C., M.G., and M.S. conducted the pseudo–*J-V* and FF-V_{OC} loss analysis (intensity-dependent V_{OC} and QFLS); D.M. performed the photoelectron spectroscopy; J.A.S., D.S., and N.P. performed crystallographic analysis; G.M., M.J., B.L., and Ei.K. designed and built the tandem aging setup and recorded the long-term MPP tracks; and S.A., V.G., M.S., T.U., T.M., C.G., R.S., M.T., La.K., A.A., D.N., B.S., and B.R. supervised the projects. All authors contributed to data interpretation and manuscript writing. **Competing interests:** HZB and Kaunas University of Technology have filed patents for the SAM molecules described above and their use in tandem solar cells. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1300/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S39
Tables S1 to S3
References (*70*–*85*)

19 June 2020; accepted 30 October 2020
10.1126/science.abd4016

REPORT

QUANTUM INFORMATION

Optically addressable molecular spins for quantum information processing

S. L. Bayliss[1]*, D. W. Laorenza[2]*, P. J. Mintun[1], B. D. Kovos[1], D. E. Freedman[2]†, D. D. Awschalom[1,3,4]†

Spin-bearing molecules are promising building blocks for quantum technologies as they can be chemically tuned, assembled into scalable arrays, and readily incorporated into diverse device architectures. In molecular systems, optically addressing ground-state spins would enable a wide range of applications in quantum information science, as has been demonstrated for solid-state defects. However, this important functionality has remained elusive for molecules. Here, we demonstrate such optical addressability in a series of synthesized organometallic, chromium(IV) molecules. These compounds display a ground-state spin that can be initialized and read out using light and coherently manipulated with microwaves. In addition, through atomistic modification of the molecular structure, we vary the spin and optical properties of these compounds, indicating promise for designer quantum systems synthesized from the bottom-up.

Optically addressable solid-state spins (*1*–*4*) are an important platform for quantum information science, with impressive demonstrations ranging from quantum teleportation (*5*) to the mapping of individual nuclear spins (*6*). The optical-spin interface of these solid-state systems is crucial for a diverse range of applications, from nanoscale sensing to long-distance quantum communication, as it enables straightforward single-spin readout and initialization. However, for this family of qubits, synthetic tunability of optical and spin properties, deterministic fabrication of multiqubit arrays, or translation of spin centers between different host materials and devices remain outstanding goals.

By contrast, chemical synthesis of molecular spin systems affords bottom-up qubit design (*7*, *8*). A chemical approach offers tunability through atomistic control over the qubit; scalability through chemical assembly of extended structures; and portability across different environments (e.g., solution, surface, solid-state), because the qubit is not confined to a specific host. These capabilities provide substantial control over the intrinsic and extrinsic environment of molecular qubits. Notably, with chemical synthesis, nuclear spins can be controllably placed around a molecular qubit (*9*), arrays of spins can be created in one-, two- and three-dimensional architectures (*10*, *11*), and molecular spins can be integrated into electronic and photonic devices (*12*, *13*). Molecular systems have shown impressive capabilities, including long spin coherence times (*14*, *15*), manipulation of photoexcited triplet states (*16*–*18*), and quantum optics with spin-singlet (S = 0) organic molecules (*19*). However, in contrast to spins in semiconductors, the ground-state spins of molecular systems have lacked an optical-spin interface for both qubit initialization and readout (*20*). Creating such an interface in a molecular platform would generate a class of qubits that can be engineered with atomic precision, with potentially transformative applications for bottom-up quantum technologies ranging from quantum sensors to hybrid quantum systems.

Here, we synthesize a series of molecular qubits with such an optically addressable ground-state spin. We show that these molecular spin qubits can be initialized and read out with light, and coherently manipulated with microwave fields. Notably, chemical modification of these qubits modulates the resulting optical-spin interface, illustrating the potential of bottom-up qubit creation for optically addressable spin systems.

To achieve the desired optical addressability, we target a molecular system consisting of a metal ion bonded to organic moieties (ligands), comprising a portable qubit of ~1-nm size. This metal-organic motif provides a well-defined qubit through the electronic spin of the central metal ion, and the surrounding ligands enable synthetic modification of the optical and spin properties.

The key requirements for such an optically addressable molecular spin qubit are (i) a ground-state spin that can be coherently manipulated and (ii) a spin-selective optical process to initialize and read out the spin. To achieve these functionalities, we selected a chromium ion (Cr^{4+}) coordinated by strong-field (aryl) ligands in a high-symmetry configuration, which gives rise to the energy-level structure shown in Fig. 1A. The d^2 electronic configuration of Cr^{4+} in a pseudo-tetrahedral environment produces a spin-triplet (S = 1) ground state with a small ground-state zero-field splitting, characterized by the parameters D and E (Fig. 1A and fig. S1), allowing for spin manipulation at readily available microwave frequencies.

A strong ligand-field environment ensures that the lowest-lying electronic excited state is a spin-singlet (S = 0) (*21*). This configuration leads to narrow optical transitions between the S = 1 ground state and the S = 0 excited state, which, when combined with the ground-state zero-field splitting, enables optical spin readout and initialization (i.e., spin polarization) through spin-selective resonant excitation. First, optical readout of the ground-state spin is possible because a probed spin sublevel (e.g., $|0\rangle$ in Fig. 1A) will give rise to more photoluminescence (PL) than the other spin sublevels (e.g., $|\pm1\rangle$ in Fig. 1A). Second, optical polarization of the ground-state spin results when selective excitation, combined with spontaneous emission, transfers population from the probed to the other spin sublevels (*22*). This is referred to as optical pumping or hole burning. To accumulate spin polarization over multiple excitation and emission cycles, the ground-state spin-lattice relaxation time (T_1) must be much longer than the excited-state lifetime (T_{opt}). These components are the key ingredients that we use to obtain the desired optical-spin interface.

With these criteria in mind, we synthesized three Cr^{4+} compounds (Fig. 1B), which differ by the placement of a single CH_3 (methyl group) on the coordinating ligands, through solution-phase chemistry. In brief, we react the appropriate aryl lithium species with $Cr^{3+}Cl_3(THF)_3$ at −78°C, which undergoes a disproportionation or auto-oxidation to yield the corresponding tetrahedral $Cr^{4+}R_4$ [R = *o*-tolyl, 2,3-dimethylphenyl, 2,4-dimethylphenyl; see (*23*) for further details] (*24*). We diluted each compound in their S = 0, isostructural tin analogs to form dilute molecular crystals (**1, 2, 3**), illustrated in Fig. 1C, thus reducing interactions between Cr^{4+} centers. The metal-center symmetry in these crystals varies from relatively high, S_4 for **1**, to low, C_1 for **3**, directly affecting the ground-state spin structure (*23*). All experiments were performed on **1**–**3** in an optical cryostat with microwave access [≈4 to 5 K at the sample mount, Fig. 1C and (*23*)] unless stated otherwise.

Under off-resonant excitation (785 nm), ground-state population is promoted to the first S = 1 excited state, undergoes fast intersystem crossing to the S = 0 state, and decays to

[1]Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA. [2]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA. [3]Department of Physics, University of Chicago, Chicago, IL 60637, USA. [4]Center for Molecular Engineering and Materials Science Division, Argonne National Laboratory, Lemont, IL 60439, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: awsch@uchicago.edu (D.D.A.); danna.freedman@northwestern.edu (D.E.F.)

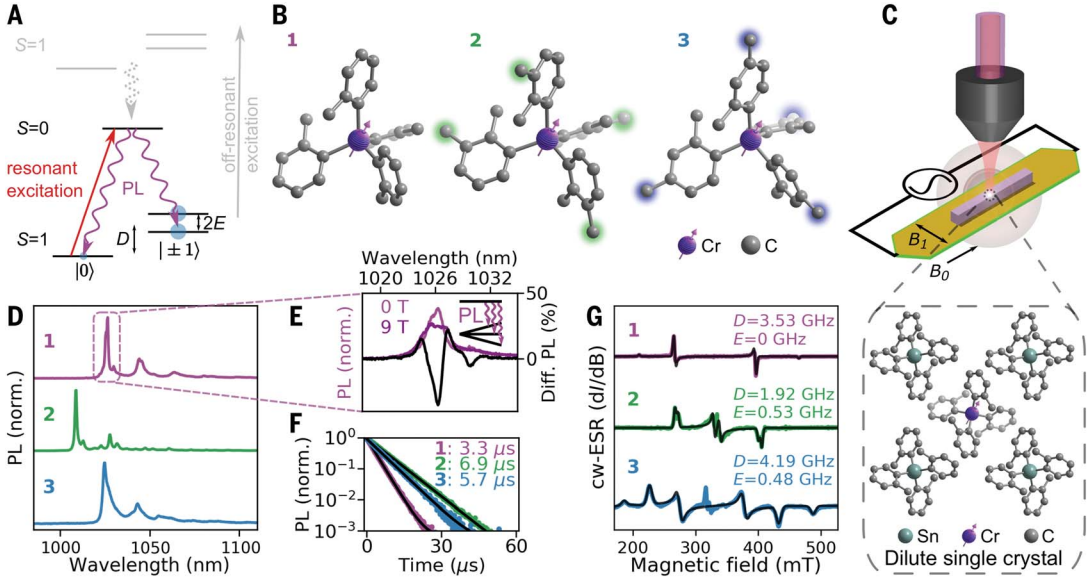


Fig. 1. Generating an optical interface for ground-state molecular spin qubits. (**A**) Energy-level diagram of Cr^{4+} in **1**–**3** depicting photoluminescence (PL) from the S = 0 state. (**B**) Molecular structures for **1**–**3** determined by single-crystal x-ray diffraction. Hydrogen atoms are omitted for clarity. Ligand modifications for **2** and **3** are highlighted in green and blue. Chromium and carbon atoms are shown in purple and gray, respectively. (**C**) Experimental schematic depicting optical excitation and PL collection for spin initialization and readout. Each Cr^{4+} compound is diluted in a single crystal (purple) of the isostructural S = 0 tin (Sn) analog. An illustrative structure is shown. A microwave field (B_1) from a waveguide (gold) is used for spin manipulation, and a static field (B_0) enables Zeeman splitting. (**D**) PL spectra for **1**–**3** at 4 K using off-resonant (785 nm) excitation. (**E**) Zeeman splitting of the zero-phonon line of **1** at 9 T. (**F**) Optical lifetimes for **1**–**3** measured using resonant excitation at the zero-phonon line. (**G**) X-band continuous-wave electron spin resonance (cw-ESR) spectra for **1**–**3** collected at 77 K. Simulations are shown in black, along with extracted D and E parameters. The central resonances at $g \approx 2$ to 2.1 are discussed in (*23*).

the S = 1 ground state, emitting near-infrared PL. For **1**–**3**, this emission comprises sharp zero-phonon lines (ZPLs) ranging from 1009 to 1025 nm (Fig. 1D), along with longer-wavelength phonon sidebands. The minor ligand modifications in **1**–**3** also result in distinctive ground-state spin structures, as observed in ground-state electron spin resonance (ESR) measurements (Fig. 1G), with extracted values of D and E lying in the readily addressable regime of <5 GHz for each compound [because the signs of D and E are not determined and only their magnitudes influence our experiments, we take D, $E > 0$ for clarity (*23*)]. These features, along with optical lifetimes (3.3 to 6.9 μs, Fig. 1F) that are much shorter than T_1 [see below and (*23*)], therefore suggest that **1**–**3** satisfy the above criteria for optically addressable molecular qubits while exhibiting variable optical and spin properties.

To further confirm the level structure in Fig. 1A, we measure the emission of **1** under a high magnetic field using off-resonant excitation (Fig. 1E). Owing to the S = 0 excited state, the Zeeman splitting of the ground state manifests directly as a shift in the optical emission energies. This effect is clearly shown by taking the difference in PL spectra at 9 and 0 T: Optical emission into the $|\pm1\rangle$ spin sublevels shift to lower and higher energies, giving characteristic peaks on either side of the zero-field ZPL in the differential spectrum (Diff. PL in Fig. 1E), along with a central dip [the feature at 1030 nm arises from the vibrational sideband (*23*)].

To demonstrate an optical-spin interface in these systems, we focus on **1** as an illustrative example before discussing **2** and **3**. Using a narrow-line laser, we resonantly excite the S = 1 ground state to the S = 0 excited state (Fig. 2A) and collect emission into the phonon sideband to remove excitation laser scatter. First, we characterize the emission as a function of the excitation wavelength (Fig. 2B), showing a ZPL at 1025 nm: We excite at this ZPL

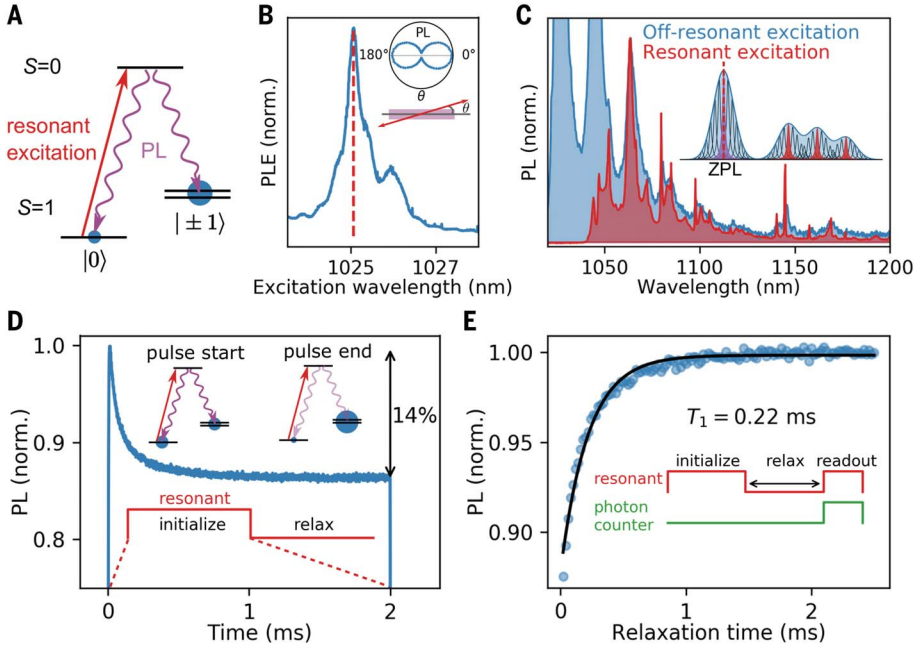


Fig. 2. All-optical ground-state spin initialization and readout of 1. (**A**) Energy-level structure showing optical spin initialization through spin-selective excitation. (**B**) Photoluminescence excitation (PLE) spectrum obtained by sweeping a narrow-line laser over the zero-phonon line. The dashed line shows the excitation wavelength used for all following experiments. (Inset) Dependence of the PL on laser polarization, defined by the angle θ from the crystal long axis. (**C**) Phonon sideband under resonant and off-resonant excitation showing emission line narrowing. (Inset) Schematic of subensemble excitation. (**D**) Time-resolved optical spin initialization. (**E**) All-optical measurement of the spin-lattice relaxation time (T_1).

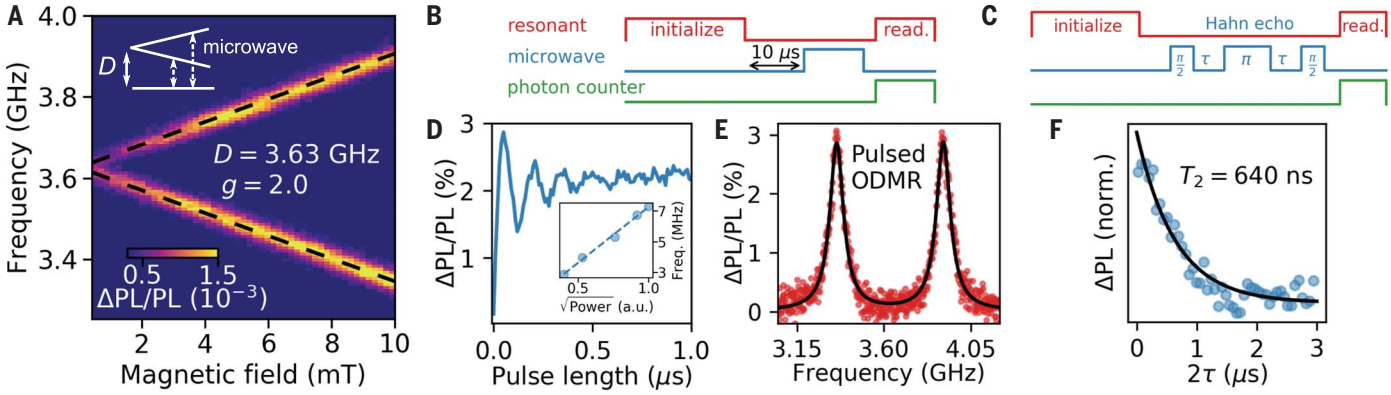


Fig. 3. Optically detected magnetic resonance (ODMR) and coherent spin manipulation of the ground state of 1. (**A**) ODMR as a function of magnetic field and microwave frequency using continuous-wave optical excitation. Dashed lines are a simulation with the stated g and D values. (**B**) Pulsed ODMR and (**C**) Hahn-echo sequences. (**D**) Rabi oscillations between the $|0\rangle$ and $|-1\rangle$ spin sublevels (B_0 = 10 mT). (Inset) Microwave-power dependence of the Rabi oscillation frequency. (**E**) Pulsed ODMR spectrum (B_0 = 10 mT) and double Lorentzian fit (black). (**F**) Optically detected ground-state spin coherence (B_0 = 2 mT) with exponential fit (black).

maximum (dashed line Fig. 2B) for all following experiments. To further maximize emission, we align the excitation polarization with the optical dipole transition, which is collinear with the long axis of the needle-like crystal (Fig. 2B, inset). The optical inhomogeneous linewidth of ≈150 GHz shown in Fig. 2B is >>D and thus appears prohibitive for spin-selective excitation; however, resonant excitation addresses only a subensemble of molecules from the inhomogeneous distribution. Whereas the inhomogeneous linewidth is broadened by static energy-level variations, likely caused by strain, the subensemble linewidth is limited only by dynamical processes (e.g., electron-phonon coupling) (*25*). To demonstrate that the subensemble linewidth is indeed much narrower than the inhomogeneous linewidth, we compare the phonon sidebands under resonant excitation and off-resonant excitation (Fig. 2C). The emission line narrowing (*26*) under resonant excitation indicates that the ensemble ZPL indeed consists of narrower subensembles, which we use for all following spin-selective experiments.

We next measure all-optical initialization and readout of the ground-state spin using hole burning and recovery. To initialize the spin, we apply the pulse sequence outlined in Fig. 2D, consisting of a long optical pulse (2 ms) followed by a wait time (2 ms) to equilibrate ground-state spin populations before the next pulse. The emission during the optical pulse shows the characteristic behavior of optical spin polarization: a gradual drop in emission as population is pumped from the probed ground-state spin sublevel (the "bright" state) and into the other ("dark") spin sublevels. The optical contrast between the start and the end of the pulse places a lower bound on the spin polarization of 14% (*23*).

Using this spin initialization, we measure the ground-state spin-lattice relaxation time, T_1, by performing the two-pulse experiment outlined in Fig. 2E. This sequence consists of an initialization pulse (300 μs), a variable relaxation time, and a readout pulse (20 μs). The initialization pulse transfers population to the "dark" spin sublevels. As ground-state spin population relaxes back to the "bright" sublevel, the emission increases. Measuring this emission at variable relaxation times yields T_1 = 0.22(1) ms. That T_1 is much longer than the optical lifetime (T_{opt} = 3.3 μs, Fig. 1F) confirms that many optical cycles can be used to accumulate ground-state spin polarization.

We next manipulate the ground-state spin of **1** using a microwave field. First, using continuous-wave (cw) optical excitation, we place a subensemble of spins in the "dark" state and monitor changes in emission (ΔPL) as we sweep the microwave frequency. When this microwave frequency matches the spin sublevel splitting, the "dark" and "bright" sublevels are mixed, resulting in increased PL. Figure 3A shows this optically detected magnetic resonance (ODMR) as a function of both the microwave frequency and an external magnetic field applied along the long axis of the crystal. The zero-field cw-ODMR spectrum provides D = 3.63 GHz, and the Zeeman splitting yields a g-factor of 2.0, in agreement with the ESR measurements (Fig. 1G).

To demonstrate coherent control over the ground-state spin, we drive Rabi oscillations (Fig. 3D) using the pulsed ODMR sequence outlined in Fig. 3B. This sequence consists of an optical initialization pulse, a wait time, a variable length microwave pulse, and an optical readout pulse. The inset shows the expected square-root dependence of the Rabi frequency on the applied microwave power. Next, using a π pulse calibrated from Fig. 3D, we perform pulsed ODMR at a fixed magnetic field, B_0 = 10 mT, while varying the microwave frequency (Fig. 3E). Finally, by replacing the single microwave pulse in Fig. 3B with a Hahn-echo sequence (Fig. 3C), we measure the spin coherence time T_2 = 640(60) ns (Fig. 3F, B_0 = 2 mT). The final π/2 pulse in the sequence projects the coherences onto spin populations for optical readout. In these pulsed ODMR experiments, the wait time (10 μs ~ $3T_{opt}$) between initialization and microwave manipulation ensures population is in the ground state prior to coherent control. This wait time, along with the above agreement between the ODMR and ESR spin parameters, shows that we coherently control the ground-state spin. Furthermore, the measured T_2, likely limited by the surrounding hydrogen nuclear spins, is comparable to other transition-metal–based molecular qubits in nuclear spin-rich environments (*15*, *27*), and we expect that ligand deuteration should greatly enhance T_2 (*27*, *28*). Thus, with **1**, we demonstrate optical initialization, microwave coherent control, and optical readout of the ground-state spin in an ensemble of molecular qubits.

Having demonstrated an optical-spin interface and coherent spin control for **1**, we highlight how this functionality is not specific to a single compound. In Fig. 4, we show optical initialization, microwave spin manipulation, and optical readout of **2** and **3** through cw-ODMR. As captured by the simulations, the variable peak intensities arise from ESR selection rules (*23*). Notably, the additional ligand methyl groups in **2** and **3** lead to substantial E by lowering the crystal symmetry, relative to **1**, indicating that noise-insensitive (i.e., clock-like) transitions could be used to enhance spin coherence in these compounds (*29*). These results also suggest that modification of the host material provides a further opportunity to manipulate ground-state spin structure. Compounds **1** to **3** illustrate that minor synthetic modifications provide variable

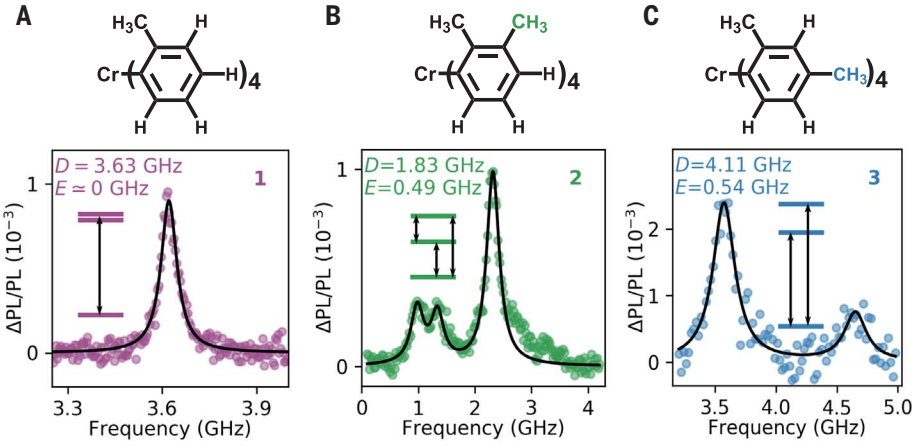


Fig. 4. Optical spin addressability with synthetic tunability. (**A** to **C**) cw-ODMR spectra and simulations (black) for **1-3,** with microwave transitions and ligand modifications depicted.

magnetic and electronic structures, indicating potential for designing compounds with multiple, separately addressable qubit sites.

Although small molecular ensembles may find applications for quantum sensing, extension of these results to the single-molecule level will be important for their implementation as qubits. We anticipate single-molecule sensitivity could be achieved by reducing radiative lifetimes through coupling to a photonic cavity and enhancing photon collection efficiency. Furthermore, we expect that initialization and readout fidelities could be improved by enhancing the spin selectivity of the optical excitation. Such enhancement could be achieved by engineering compounds with larger zero-field splittings, or narrower subensemble optical linewidths (*23*).

We envisage several promising avenues for optically addressable molecular spin systems, which may be more challenging to achieve with solid-state defects. First, multimetallic systems could be synthesized with interacting metal sites that exhibit distinct optical-spin interfaces, enabling both individual qubit addressability and multiqubit operations. Targeted nuclear spin placement around these metal centers could further provide long-lived quantum memories (*30*). Alternatively, ligand functionalization and extension of these results to biocompatible metal ions (*31*) may enable sensing of targeted analytes. Furthermore, self-assembly could facilitate integration with photonic or electronic devices (*32*). Finally, the highly tunable nature of molecules offers promise for rationally designed properties tailored to a specific purpose, e.g., telecom emission for long-distance quantum networks, or strong spin-phonon interactions for hybrid quantum systems. These results highlight pathways to design and create quantum technologies from the bottom-up.

REFERENCES AND NOTES

1. M. Atatüre *et al.*, *Science* **312**, 551–553 (2006).
2. D. D. Awschalom, R. Hanson, J. Wrachtrup, B. B. Zhou, *Nat. Photonics* **12**, 516–527 (2018).
3. A. Gottscholl *et al.*, *Nat. Mater.* **19**, 540–545 (2020).
4. N. Chejanovsky *et al.*, *Single spin resonance in a van der Waals embedded paramagnetic defect. arXiv:1906.05903 [cond-mat.mes-hall] (2019).*
5. W. Pfaff *et al.*, *Science* **345**, 532–535 (2014).
6. M. H. Abobeih *et al.*, *Nature* **576**, 411–415 (2019).
7. A. Gaita-Ariño, F. Luis, S. Hill, E. Coronado, *Nat. Chem.* **11**, 301–309 (2019).
8. J. Ferrando-Soria *et al.*, *Nat. Commun.* **7**, 11377 (2016).
9. C. E. Jackson, C. Y. Lin, S. H. Johnson, J. van Tol, J. M. Zadrozny, *Chem. Sci.* **10**, 8447–8454 (2019).
10. T. Yamabayashi *et al.*, *J. Am. Chem. Soc.* **140**, 12090–12101 (2018).
11. A. Urtizberea *et al.*, *Adv. Funct. Mater.* **28**, 1801695 (2018).
12. R. Vincent, S. Klyatskaya, M. Ruben, W. Wernsdorfer, F. Balestro, *Nature* **488**, 357–360 (2012).
13. M. Oxborrow, J. D. Breeze, N. M. Alford, *Nature* **488**, 353–356 (2012).
14. K. Bader *et al.*, *Nat. Commun.* **5**, 5304 (2014).
15. J. M. Zadrozny, J. Niklas, O. G. Poluektov, D. E. Freedman, *ACS Cent. Sci.* **1**, 488–492 (2015).
16. J. Wrachtrup, C. von Borczyskowski, J. Bernard, M. Orritt, R. Brown, *Nature* **363**, 244–245 (1993).
17. J. Köhler *et al.*, *Nature* **363**, 242–244 (1993).
18. V. Filidou *et al.*, *Nat. Phys.* **8**, 596–600 (2012).
19. D. Wang *et al.*, *Nat. Phys.* **15**, 483–489 (2019).
20. M. Atzori, R. Sessoli, *J. Am. Chem. Soc.* **141**, 11339–11352 (2019).
21. B. N. Figgis, M. A. Hitchman, in *Ligand Field Theory and Its Applications* (Wiley-VCH, 2000), pp. 131–141.
22. W. F. Koehl *et al.*, *Phys. Rev. B* **95**, 035207 (2017).
23. See the supplementary materials.
24. S. U. Koschmieder *et al.*, *J. Chem. Soc. Dalton Trans.*, 3427–3433 (1990).
25. H. Riesen, *Coord. Chem. Rev.* **250**, 1737–1754 (2006).
26. H. Riesen, E. Krausz, *J. Chem. Phys.* **97**, 7902–7910 (1992).
27. A. Ardavan *et al.*, *Phys. Rev. Lett.* **98**, 057201 (2007).
28. A. Ardavan *et al.*, *npj Quantum Inf.* **1**, 15012 (2015).
29. M. Shiddiq *et al.*, *Nature* **531**, 348–351 (2016).
30. M. V. G. Dutt *et al.*, *Science* **316**, 1312–1316 (2007).
31. M. K. Wojnar, D. W. Laorenza, R. D. Schaller, D. E. Freedman, *J. Am. Chem. Soc.* **142**, 14826–14830 (2020).
32. R. Chikkaraddy *et al.*, *Nature* **535**, 127–130 (2016).
33. S. L. Bayliss, D. W. Laorenza, P. J. Mintun, B. D. Kovos, D. E. Freedman, D. D. Awschalom, Dataset for: Optically addressable molecular spins for quantum information processing. *Zenodo* (2020), doi: https://doi.org10.5281/zenodo.3909747.

ACKNOWLEDGMENTS

We thank M. S. Fataftah for experimental suggestions and insightful discussions and C. P. Anderson, A. Bourassa, P. Deb, G. Smith, L. R Weiss, M. J. Amdur, K. A. Collins, and M. K. Wojnar for helpful comments on the manuscript. We thank P. H. Oyala for technical support with the ESR measurements and R. A. Sponenburg for technical support with ICP-OES experiments. **Funding:** We acknowledge funding from ONR N00014-17-1-3026 and the MRSEC Shared User Facilities at the University of Chicago (NSF DMR-1420709). This work made use of the Caltech EPR facility, which is supported by the NSF (NSF-1531940) and the Dow Next Generation Educator Fund, and IMSERC at Northwestern University, which has received support from Northwestern University, the State of Illinois, and the International Institute of Nanotechnology. Metal analysis was performed at the Northwestern University Quantitative Bio-element Imaging Center. D.E.F. and D.W.L. gratefully acknowledge the U.S. Department of Energy (DE-SC0019356) for total support on synthesis, design of molecules and ensemble ESR measurements. **Author contributions:** S.L.B., D.W.L., P.J.M., and B.D.K. performed the measurements and analyzed the data. D.W.L. synthesized the compounds. D.E.F. and D.D.A. advised on all efforts. All authors contributed to the data analysis and manuscript preparation. **Competing interests:** S.L.B., D.W.L., P.J.M., B.D.K., D.E.F., and D.D.A. are inventors on patent application no. 63008589 submitted by the University of Chicago that covers chemically tunable optically addressable molecular-spin qubits and associated methods. P.J.M. is a paid consultant to ARCH Venture Partners. **Data and materials availability:** The data can be accessed at Zenodo (*33*). Crystallographic data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif or by emailing data_request@ccdc.cam.ac.uk. CCDC codes: **1** (1992362), **2** (1992359), **3** (1992360), **1-Cr** (1992356), **2-Cr** (1992357), **3-Cr** (1992358), and **2-Sn** (1992361).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1309/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S6
References (*34*–*61*)

27 March 2020; accepted 2 November 2020
Published online 12 November 2020
10.1126/science.abb9352

BATTERIES

Reversible planar gliding and microcracking in a single-crystalline Ni-rich cathode

Yujing Bi[1], Jinhui Tao[1], Yuqin Wu[2,3], Linze Li[1], Yaobin Xu[1], Enyuan Hu[4], Bingbin Wu[1], Jiangtao Hu[1], Chongmin Wang[1], Ji-Guang Zhang[1], Yue Qi[2,3], Jie Xiao[1,5]*

High-energy nickel (Ni)–rich cathode will play a key role in advanced lithium (Li)–ion batteries, but it suffers from moisture sensitivity, side reactions, and gas generation. Single-crystalline Ni-rich cathode has a great potential to address the challenges present in its polycrystalline counterpart by reducing phase boundaries and materials surfaces. However, synthesis of high-performance single-crystalline Ni-rich cathode is very challenging, notwithstanding a fundamental linkage between overpotential, microstructure, and electrochemical behaviors in single-crystalline Ni-rich cathodes. We observe reversible planar gliding and microcracking along the (003) plane in a single-crystalline Ni-rich cathode. The reversible formation of microstructure defects is correlated with the localized stresses induced by a concentration gradient of Li atoms in the lattice, providing clues to mitigate particle fracture from synthesis modifications.

For long-range electrical vehicles, next-generation Li-ion batteries that use stable high-energy cathode materials are needed (*1*). Among different cathodes, Ni-rich NMC ($LiNi_xMn_yCo_{1-x-y}O_2$, $x > 0.6$), with >200 mA·hour/g capacity, high voltage (>3.8 V), and low cost, are considered as one of the most promising candidates (*2*). Traditional $LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$ cathodes are prepared using a coprecipitation method that aggregates nano-sized primary particles into micron-sized secondary particles. The agglomerated polycrystalline NMC have shortened diffusion lengths within their primary particles and an increased number of pores, which accelerates Li^+ transport. This has thus become the most popular morphology for conventional NMC particles (*3*).

Creating spherical secondary polycrystalline NMC particles reduces the surface/volume ratio, but pulverization along the weak internal grain boundaries are generally observed after cycling (*4*). These cracks are induced by the nonuniform volume change of primary particles during cycling and exacerbated by the anisotropy among individual particles and grains in the polycrystalline NMC (*5*). The intergranular cracking exposes new surfaces to electrolyte for side reactions, which accelerates cell degradation (*6*). As the Ni content increases to more than 0.6—e.g., $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$—usable capacity is further increased, but new challenges such as moisture sensitivity (*7*), aggressive side reactions, and gas generation (*8*) during cycling are found, all of which initiate from particle surfaces.

A feasible approach for stabilizing Ni-rich NMC is to synthesize single-crystalline cathodes with reduced surface areas, phase boundaries, and more integrated crystal structures, similar to $LiCoO_2$ (*9*). Single-crystalline NMC with Ni ≤ 0.6 that shows good performance has been demonstrated (*10*–*17*). However, the capacity of those single-crystalline NMC (Ni ≤ 0.6) is relatively low, owing to the limited Ni content. Single-crystalline Ni-rich cathodes with Ni > 0.6 substantially increase the capacity but are not much reported (*18*, *19*) because of the challenges of preparing high-performance Ni-rich NMC single crystals. Owing to low structure stability of Ni-rich cathode at high temperatures, higher-Ni-content cathodes require lower synthesis temperatures, in contrast to the high-temperature and time consuming calcination process required to grow single crystals. Synthesis of stoichiometric single crystals with competitive electrochemical properties is a challenge (*20*).

This work uses high-performance single-crystalline $LiNi_{0.76}Mn_{0.14}Co_{0.1}O_2$ (NMC76) as a model material to study how the potential triggers the structural changes of single crystals from the atomic to micron scale and its implications for the electrochemical properties of cathodes. In situ atomic force microscopy (AFM) and theoretical modeling are used to understand the coupled electro-chemo-mechanical behaviors of single crystals.

Synthesized NMC76 has an average particle size of 3 μm (Fig. 1A). A cross-sectional view (Fig. 1B) shows that NMC76 has a dense structure without cavities or grain boundaries. Pure phases of α-$NaFeO_2$–type layered structures are confirmed by both selected area electron diffraction (SAED; Fig. 1C) and x-ray diffraction (XRD; Fig. 1D). Lattice parameters *a* and *c* are 2.8756(1) and 14.2221(1) Å, respectively, from Rietveld refinement (see details in table S1). For comparison, polycrystalline NMC76 is found to contain many internal pores and intergranular boundaries along with surface films (fig. S1) formed from the reactions between NMC and air (*7*). By contrast, the surface of single-crystalline NMC76 is uniform (Fig. 1, E and F). Elemental mapping (Fig. 1, G and H) indicates a homogeneous distribution

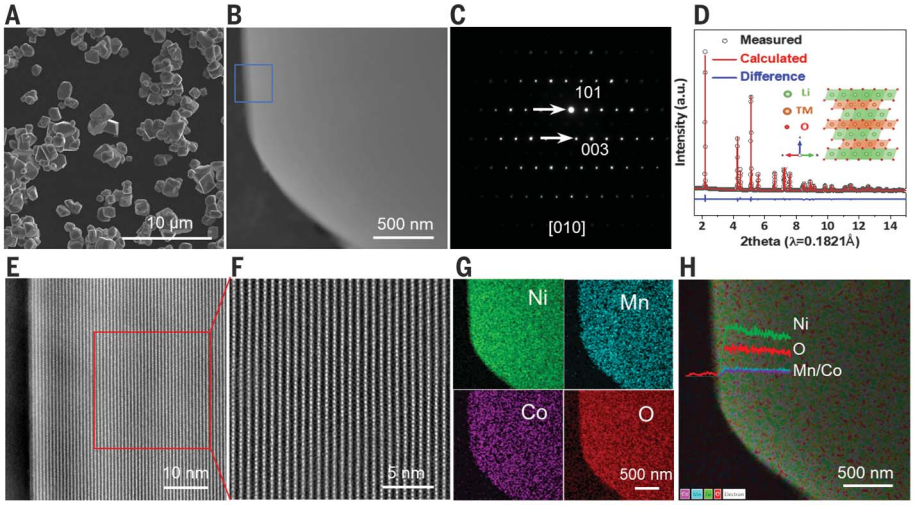


Fig. 1. Characterization of single-crystalline $LiNi_{0.76}Mn_{0.14}Co_{0.1}O_2$. (**A**) SEM image of single-crystalline NMC76. (**B**) Cross-sectional image of single-crystalline NMC76. (**C**) Electron diffraction pattern of a selected area of single-crystalline NMC76. (**D**) Synchrotron XRD and Rietveld refinement pattern. (**E**) High-resolution high-angle annular dark-field (HAADF)–STEM image of single-crystalline NMC76 [corresponding to the region enclosed by the blue square in (B)]. (**F**) Higher-magnification view corresponding to the region enclosed by the red square in (E). (**G**) Energy dispersive spectroscopy (EDS) elemental mapping of Ni, Mn, Co, and O. (**H**) EDS overlapped image and line scanning shows the elemental distribution intensity (inset).

[1]Pacific Northwest National Laboratory, Richland, WA 99352, USA. [2]School of Engineering, Brown University, Providence, RI 02912, USA. [3]Department of Chemical Engineering and Materials Science, Michigan State University, East Lansing, MI 48824, USA. [4]Chemistry Division, Brookhaven National Laboratory, Upton, NY 11973, USA. [5]Materials Science and Engineering Department, University of Washington, Seattle, WA 98195, USA.
*Corresponding author. Email: jie.xiao@pnnl.gov

of Ni, Mn, and Co with a stoichiometric ratio as designed (fig. S2 and table S2). Continuous phase transitions happen when potential changes (fig. S3), similar to polycrystalline NMC76.

Single-crystalline NMC76 is further tested in graphite/NMC full cells at realistic conditions. The typical loading of NMC76 cathodes is about 20 mg/cm^2 (4 mA·hour/cm^2) with about 32% porosity, which is needed to build a 250 W·hour/kg Li-ion cell (see table S3). Harlow *et al.* presented a wide range of testing results of Li-ion batteries based on graphite/$LiNi_{0.5}Mn_{0.3}Co_{0.2}O_2$ chemistry (*16*), which can serve as benchmarks for academic research. The electrode properties in terms of loading and so on in (*16*) are similar to ours because both are derived from realistic pouch cell parameters. At such a high cathode loading, Li metal will worsen the cycling stability (figs. S4 and S5) owing to the deepened stripping and deposition process of Li (*21*). Between 2.7 and 4.2 V (versus graphite), single-crystalline NMC76 delivers 182.3 mA·hour/g discharge capacity at 0.1 C and retains 86.5% of its original capacity after 200 cycles (Fig. 2A$_1$). With a cutoff of 4.3 V, single-crystalline NMC76 delivers 193.4 mA·hour/g capacity with 81.6% capacity retention after 200 cycles (Fig. 2A$_2$). Upon further increasing to 4.4 V, a 196.8 mA·hour/g discharge capacity is seen (Fig. 2A$_3$) along with a 72.0% capacity retention after 200 cycles. Note that 200 cycles at a C/10 charge rate and a C/3 discharge rate mean 2600 hours of cycling. The total testing time is equal to a cell undergoing 1300 cycles at 1 C. To evaluate the electrochemical properties of materials, it is important to use a relatively slow rate rather than a very fast rate, which produces seemingly "longer" cycles but may "hide" critical electrochemical information. It is the total time of charge-discharge that matters, not cycling number (*22*). Increased polarization (Fig. 2, B$_1$ to B$_3$) is observed when the cutoff voltage increases, which is presumably assigned to intensified electrolyte decomposition at increased voltages (figs. S6 to S8) and thus higher impedance resulting from cathode passivation films and single-crystal lattice change. Table S4 summarizes the electrochemical performances and testing conditions of all previously published single-crystalline Ni-rich NMC (Ni > 0.6) cathode materials.

Lattice gliding is observed in single-crystalline NMC76 at high voltages. Between 2.7 and 4.2 V (versus graphite), the entire single crystal is well maintained after 200 cycles (Fig. 2C$_1$). Upon increasing the cutoff voltage to 4.3 V, there are some gliding lines seen on the crystal surfaces after 200 cycles (Fig. 2C$_2$). Cycled to 4.4 V, single crystals appeared to be "sliced" (Fig. 2C$_3$ and fig. S9) in parallel, along the (003) plane and vertical to the *c* axis of the layered structure (Fig. 3C), which indicates a model II–type crack (in-plane shear) in fracture mechanics. Additionally, small cracks that

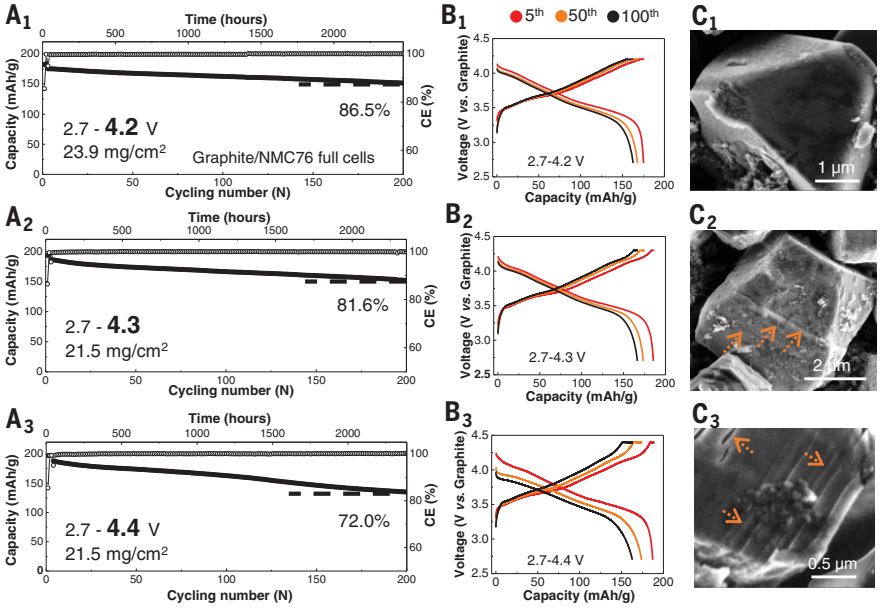


Fig. 2. Electrochemical performance of single-crystalline $LiNi_{0.76}Mn_{0.14}Co_{0.1}O_2$ and SEM images after cycling stability tests. (**A$_1$** to **A$_3$**) Cycling stability of single-crystalline NMC76 in full cells within different electrochemical windows. The top grayish lines and solid black lines represent coulombic efficiencies and discharge capacities at different cycles of the cells, respectively. The dashed partial lines indicate the capacities at the 200th cycles, with capacity retention shown underneath the dashed lines. (**B$_1$** to **B$_3$**) The corresponding charge-discharge curves of cells in (A$_1$) to (A$_3$). (**C$_1$** to **C$_3$**) SEM images of a single crystal after 200 cycles in (A$_1$) to (A$_3$). The orange arrows indicate gliding marks or microcracks found in cycled singe crystals. mAh, milliampere hour.

indicate a model I–type fracture (opening) are also observed at 4.4 V (Fig. 2C$_3$). All characterizations have been done by selecting various regions of NMC76 electrodes, and the same phenomenon is repeatedly found (figs. S10 to S12). Although single-crystalline NMC76 as a whole particle is still intact (Fig. 3A), gliding is the major mechanical degradation mode, especially when the cutoff voltage is above 4.3 V. Of note, the "gliding steps" formed in cycled crystals are quite different from cracking along intergranular boundaries of polycrystalline NMC particles (*23*). The scanning transmission electron microscopy (STEM) image for single-crystalline NMC76 (Fig. 3, B to D) confirms that on both sides of a gliding plane (yellow line in Fig. 3D), the *d*-spacing of the (003) plane (0.48 nm) is unchanged and the layered structure is well maintained after the gliding marks occur. The long-range lattice symmetry of the bulk material will therefore not be altered (fig. S13). Ni, Mn, Co, and O are still uniformly distributed in the vicinity of glided planes based on electron energy-loss spectroscopy (EELS) analysis (Fig. 3F and fig. S14). The uniform element distribution and intimately attached lattices across the gliding planes demonstrate that although planar gliding occurs, no new boundary is generated, and the sliced area maintains the same lattice structure and chemical conditions as in the bulk phase. The gliding line (or the slicing marks) cannot be observed on the cross section of bulk particles by SEM and are only visible by bright-field STEM on thinly sliced TEM samples. Although the internal lattice symmetry is well maintained after the gliding, the repeated gliding near the crystal surface eventually will evolve into microcracks, exposing new surfaces to the electrolyte (Fig. 2).

To further induce lattice gliding, the cutoff voltage of the NMC76 single crystal is increased to 4.8 V (versus Li^+/Li). Slicing marks and microcracks are present in almost every charged single crystal (fig. S15A). A slight deformation of individual single crystals is observed (Fig. 3G), probably because the gliding of each layer is equally likely to move toward symmetrically equivalent directions (*24*). Surprisingly, after discharging back to 2.7 V, the majority of single crystals revert to their original morphologies and the previously observed steps and microcracks disappear (fig. S15B). The glided layers within single crystals almost completely glided back to their original locations (Fig. 3H), fully recovering from the deformation (Fig. 3G), although some "traces" are visible (shown in Fig. 3H). Within the "regular" electrochemical window of 2.7 to 4.4 V (versus graphite), after extensive cycling, lattice gliding and microcracking are also seen within the crystal lattice at charged status (Fig.

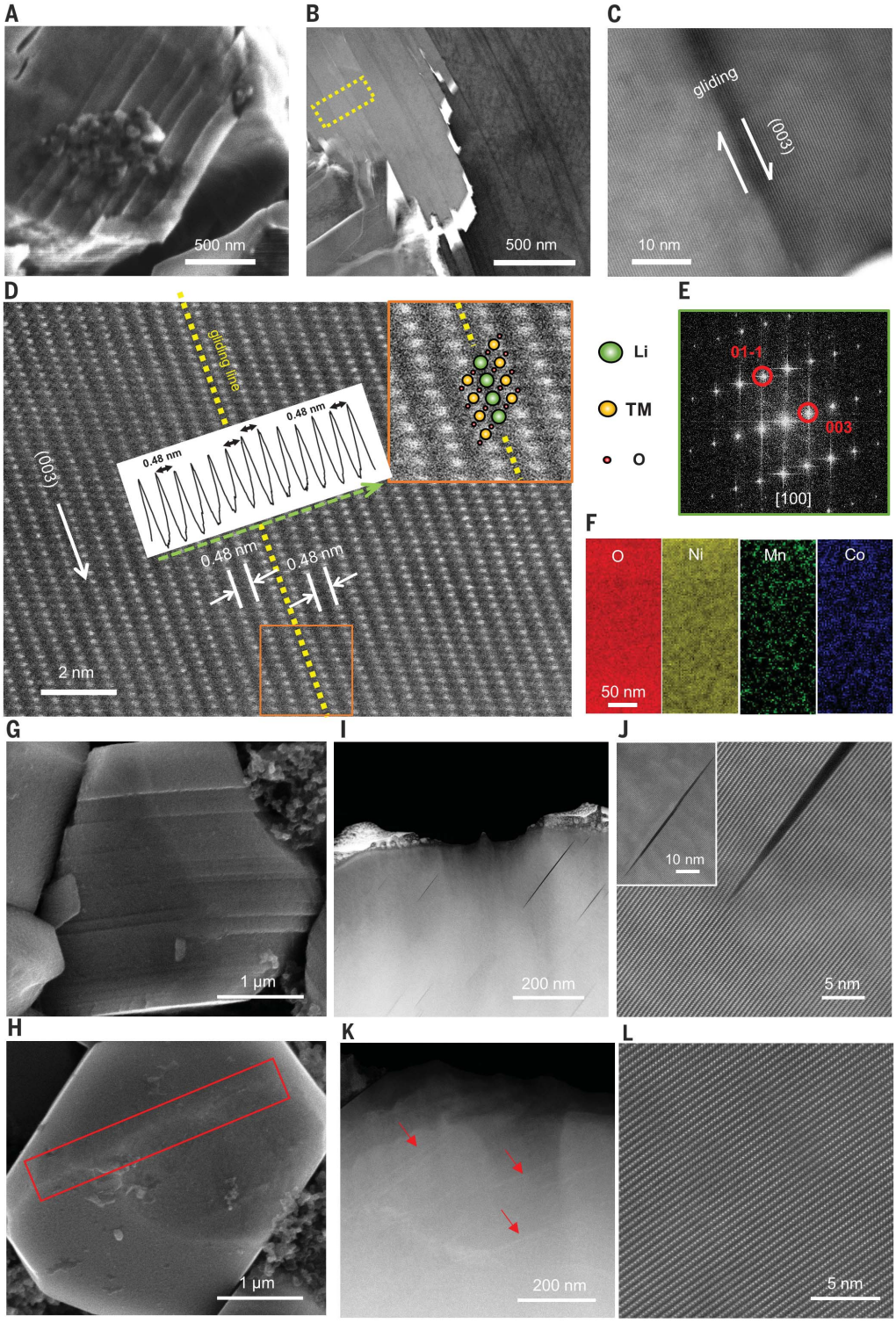


Fig. 3. Morphology and structure study of single-crystalline NMC76. (**A**) SEM image of single-crystalline NMC76 after 200 cycles. (**B**) Cross-sectional bright-field STEM image of single-crystalline NMC76 after 200 cycles. (**C**) Bright-field STEM image of internal slicing. (**D**) HAADF-STEM image from around the slicing area. The upper inset is a magnified image of the gliding area enclosed by the red square. (**E**) SAED of the gliding area. (**F**) EELS mapping of the region enclosed by the dashed rectangle in (B). (**G**) SEM image of single-crystalline NMC76 initially charged to 4.8 V (versus Li^+/Li). (**H**) SEM image of single-crystalline NMC76 discharged to 2.7 V (after being charged to 4.8 V versus Li^+/Li). The region enclosed by the red rectangle shows that the planes glided back with the traces seen. (**I** and **J**) STEM images of single-crystalline NMC76 at 4.4-V charge status (cycled in a full cell between 2.7 and 4.4 V for 120 cycles). (**K** and **L**) STEM images of single-crystalline NMC76 at discharge status (cycled in a full cell between 2.7 and 4.4 V for 120 cycles). The red arrows indicate the gliding marks inside the single crystal.

3I). STEM analysis of the NMC76 crystal (Fig. 3J) indicates that the microcracks initiate from inside of the crystal. At the discharge status of those cycled crystals (cut off at 4.4 V), few ridges or cracks are found on the crystals. Although not as visible as in charged crystals, STEM still uncovers some slicing marks (Fig. 3K) on discharged single-crystalline NMC76, which probably undergoes a reversible sliding process back and forth during 120 cycles. No microcracking is identified in those "self-healed" single crystals (Fig. 3L), suggesting that the lattice gliding and cracking in some of the crystals are still reversible after 120 cycles. As cycling continues, particle deformation will become dominant. Dislocation is also observed near the tip regions of the microcrack of single crystals charged at 4.4 V (fig. S16). The accumulation of dislocation is accompanied by the microcrack propagation (*25*). A trace amount of nano-sized NiO-like rock salt phase (fig. S17) is observed on the gliding exposure step area of single-crystalline NMC76 after cycling.

In situ AFM has been used to image the crystal surface in real time in an electrochemical cell. A ~3-μm-sized NMC76 single crystal is studied by in situ AFM during charge and discharge (Fig. 4). Regions B and C in Fig. 4A are enlarged in Fig. 4, B and C, respectively, to probe the origin and evolution of gliding steps and microcracks under the electrical field. The formation of nanosized crack domains is observed on the side surface from the open-circuit voltage (OCV) to 4.50 V (versus Li^+/Li) during charge, whereas these domains disappear in the discharge process (Fig. 4B). Moreover, planar gliding is characterized by the appearance of wide crystal steps on the side surface owing to the uneven movement between neighboring layers during polarization. Wider gliding steps are observed on the side surface starting at a 4.20 V charging process and lead to more and wider (~85 nm) gliding steps at 4.50 V (Fig. 4C). When the cell potential decreases to 4.19 V, a few wide gliding steps decrease in their width (fig. S18), indicating that the atomic layer recovers back to its original position (Fig. 4C). In situ AFM showing the reversible and continuous morphological changes during charge and discharge is shown in movie S1. The reversible gliding process is further illustrated in Fig. 4F. The observed lattice gliding is a direct observation of the lattice-invariant shear (LIS) (*24*). LIS should exist in many layered electrode materials, which experience stacking-sequence–change phase transformations owing to Li concentration change. It was also predicted that LIS will lead to particle deformation and ridges on the particle surface, but these signals are likely to be buried in the internal boundaries in a spherical secondary polycrystalline NMC particle. The micron-sized single crystal provides a clear platform to observe gliding or LIS-induced mechanical degradation.

Electrochemical potential difference is the driving force of Li-ion diffusion and the formation of the Li concentration gradient. Stress will be generated during Li^+ diffusion after establishing a Li concentration gradient in the lattice. An analytical cylindrical isotropic diffusion-induced stress model is applied to determine the stress generation when Li ions diffuse along the radial direction in the particle (fig. S19). The peak tensile stress along the tangential and axial directions occurs near the surface at the onset of delithiation during charge (fig. S20). Conversely, the peak tensile stress in all three directions occurs at the center of the particle (fig. S21) during lithiation. During charge (delithiation), the tensile stress along the axial and tangential directions is localized on the surfaces of single crystals, leading to microcrack opening normal to the (003) planes.

Local stress also has a shear component along other directions, which is solved numerically via COMSOL. The shear stress component along the *yz* direction that can trigger the gliding along the (003) planes is shown in Fig. 4, D and E. Although the signs of the shear stress during lithiation and delithiation are opposite, which explains the reversible gliding, the absolute values are not the same (fig. S22), because the elastic modulus is a function of Li concentration (*26*). Therefore, the gliding motion should be largely, but not completely, reversible. The irreversible gliding can generate small damages, being accumulated into the crack opening over a long cycling time, an analog of fatigue crack nucleation. These lead to the ridges and microcracks seen on the surfaces of single crystals after cycling.

The simple isotropic diffusion-induced-stress model can be used to predict if the cracks can be stabilized inside of the single crystal. Because the strain energy inside the particle reaches a maximum around the scaled time of $T_p = 0.1T$ during delithiation (fig. S20), its comparison with the fracture energy (2γ) is used as a criterion to evaluate the critical size of single-crystalline NMC76 (*26*). If the accumulated strain energy is not large enough to entirely cleave the crystal, the microcrack will be stabilized inside of the particle.

$$\Pi|_{T_p} = \int \frac{\sigma^2}{2E} dV = \pi * h * \left[\frac{\alpha * E_0 * (C_R - C_0)}{1-\nu}\right]^2 * \int_0^r \xi^2 \frac{1}{E} r dr < 2\gamma \quad (1)$$

where h is the height of the cylindrical particle, α is the concentration expansion coefficient, E_0 is Young's modulus of the nonlithiated particle, E is Young's modulus at a given Li-ion concentration, C_R is the Li-ion concentration at the surface, C_0 is the Li-ion concentration at the center, ν is Poisson's ratio, ξ represents the dimensionless stress (figs. S19 to S21), and γ is the surface energy; these values are listed in the supplementary materials. A lower-bound estimation of the critical size of the single crystal is predicted to be ~3.5 μm, below which cracks can be considered stable inside of the particle. The simulation result suggests that although fractures along the (003) direction appear in single crystals during cycling, the cracks are stable once formed and will not initiate catastrophic

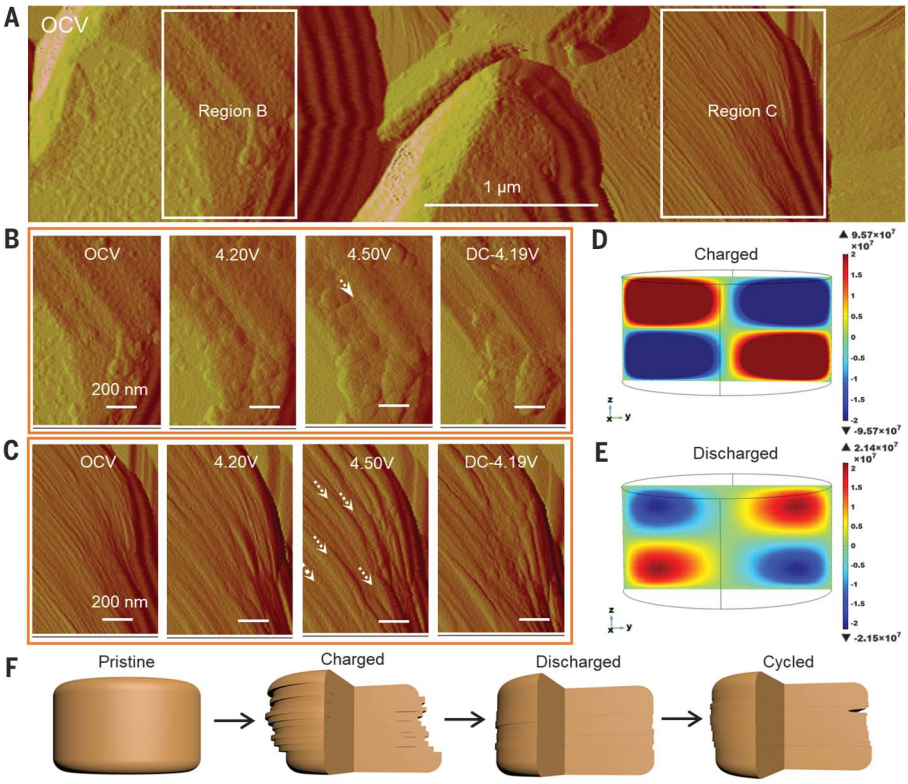


Fig. 4. Surface structure and morphology evolution by in situ AFM and mechanical analysis for single-crystalline NMC76. (**A**) AFM image at the OCV state. (**B** and **C**) Comparison of selected surface evolution during in situ AFM testing of regions B and C in (A), respectively. (**D**) COMSOL-simulated shear stress along the *yz* direction during charge (delithiation) at scaled time of 0.1*T*. (**E**) COMSOL-simulated shear stress along the *yz* direction during discharge (lithiation) at scaled time of 0.1*T*. (**F**) Schematic illustration of the structural evolution of single-crystalline NMC76 upon cycling.

reactions to produce a fracture zone that eventually pulverizes the entire single crystal. Increasing the applied current density will lead to a higher concentration gradient and higher stress generation. Increasing the cut-off voltage is equivalent to increasing (C_R – C_0) in Eq. 1. It means higher stress generation and large strain energies at increased voltages, which causes more gliding and cracking (Fig. 2). Our findings provide some strategies to stabilize single-crystalline Ni-rich NMC by either reducing the crystal size to below 3.5 μm, absorbing accumulated strain energy through modification of the structure symmetry, or simply optimizing the depth of charge without sacrificing much reversible capacity.

REFERENCES AND NOTES

1. M. S. Whittingham, *Chem. Rev.* **114**, 11414–11443 (2014).
2. J. B. Goodenough, K. S. Park, *J. Am. Chem. Soc.* **135**, 1167–1176 (2013).
3. A. van Bommel, J. R. Dahn, *J. Electrochem. Soc.* **156**, A362–A365 (2009).
4. P. Yan *et al*., *Nat. Commun.* **8**, 14101 (2017).
5. Y. W. Mao *et al*., *Adv. Funct. Mater.* **29**, 1900247 (2019).
6. M. M. Besli *et al*., *Chem. Mater.* **31**, 491–501 (2019).
7. R. Jung *et al*., *J. Electrochem. Soc.* **165**, A132–A141 (2018).
8. K.-W. Nam *et al*., *Adv. Funct. Mater.* **23**, 1047–1063 (2013).
9. Y. Lyu *et al*., *Adv. Energy Mater.* 10.1002/aenm.202000982 (2020).
10. H. Y. Li, J. Li, X. W. Ma, J. R. Dahn, *J. Electrochem. Soc.* **165**, A1038–A1045 (2018).
11. J. Li *et al*., *J. Electrochem. Soc.* **164**, A3529–A3537 (2017).
12. J. C. Garcia *et al*., *J. Phys. Chem. C* **121**, 8290–8299 (2017).
13. T. Kimijima *et al*., *J. Mater. Chem. A* **4**, 7289–7296 (2016).
14. G. Qian *et al*., *Energy Storage Mater.* **27**, 140–149 (2020).
15. Y. Liu, J. Harlow, J. Dahn, *J. Electrochem. Soc.* **167**, 020512 (2020).
16. J. E. Harlow *et al*., *J. Electrochem. Soc.* **166**, A3031–A3044 (2019).
17. J. Li, H. Y. Li, W. Stone, S. Glazier, J. R. Dahn, *J. Electrochem. Soc.* **165**, A626–A635 (2018).
18. J. Zhu, G. Y. Chen, *J. Mater. Chem. A* **7**, 5463–5474 (2019).
19. H. Li *et al*., *J. Electrochem. Soc.* **166**, A1956–A1963 (2019).
20. J. Xiao, "Understanding and modifying cathode/electrolyte interfaces," presented at the U.S. Department of Energy Vehicle Technologies Office Annual Merit Review, BAT407, 12 June 2019.
21. S. Chen *et al*., *Joule* **3**, 1094–1105 (2019).
22. M. B. Pinson, M. Z. Bazant, *J. Electrochem. Soc.* **160**, A243–A250 (2013).
23. H. H. Ryu, K. J. Park, C. S. Yoon, Y. K. Sun, *Chem. Mater.* **30**, 1155–1163 (2018).
24. M. D. Radin, J. Alvarado, Y. S. Meng, A. Van der Ven, *Nano Lett.* **17**, 7789–7795 (2017).
25. A. Ulvestad *et al*., *Science* **348**, 1344–1347 (2015).
26. R. Deshpande, Y. Qi, Y. T. Cheng, *J. Electrochem. Soc.* **157**, A967–A971 (2010).

ACKNOWLEDGMENTS

The authors thank Q. Li and D. Lu at Pacific Northwest National Laboratory (PNNL) for helping with electrode preparation and electrochemical tests. **Funding:** This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the U.S. Department of Energy (DOE) through the Applied Battery Research Program under contract no. DE-LC-000L053. E.H. is supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Vehicle Technology office of the DOE through the Advanced Battery Materials Research program under contract no. DE-SC0012704. Y.W. and Y.Q. thank the National Science Foundation for support through grant number DMR-1832808. This research used 28-ID-2 (XPD) of the National Synchrotron Light Source II, a DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0013004. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under contract DE-AC05-76RL01830. Part of this work was presented during the 2019 DOE Vehicle Technology Office Annual Merit Review Meeting. **Author contributions:** J.X. and Y.B. proposed the research and conceptualized the idea. Y.B. synthesized single-crystal samples. Y.B. performed the electrochemical measurements and SEM characterization with help from B.W. and J.H. J.T. conducted the in situ electrochemical AFM tests and analyzed data. C.W., L.L., and Y.X. collected the TEM data. E.H. collected the synchrotron data and structure refinement. Y.Q. and Y.W. carried out the mechanical modeling. J.-G.Z. participated in data analysis and discussion. Y.B. and J.X. drafted the manuscript with input and revision from all authors. **Competing interests:** The authors declare no competing interests. Part of this work has been filed under a patent application (#31837) through PNNL. **Data and materials availability:** All data are available in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1313/suppl/DC1
Materials and Methods
Figs. S1 to S22
Tables S1 to S4
References (*27*–*31*)
Movie S1

17 April 2020; accepted 27 October 2020
10.1126/science.abc3167

BIOPHYSICS

Protein condensates as aging Maxwell fluids

Louise Jawerth[1,2], Elisabeth Fischer-Friedrich[3,4], Supropriya Saha[1]*, Jie Wang[2]†, Titus Franzmann[2,4], Xiaojie Zhang[5], Jenny Sachweh[5]‡, Martine Ruer[2], Mahdiye Ijavi[2]§, Shambaditya Saha[6], Julia Mahamid[5], Anthony A. Hyman[2,3,7]¶, Frank Jülicher[1,3,7]¶

Protein condensates are complex fluids that can change their material properties with time. However, an appropriate rheological description of these fluids remains missing. We characterize the time-dependent material properties of in vitro protein condensates using laser tweezer–based active and microbead-based passive rheology. For different proteins, the condensates behave at all ages as viscoelastic Maxwell fluids. Their viscosity strongly increases with age while their elastic modulus varies weakly. No significant differences in structure were seen by electron microscopy at early and late ages. We conclude that protein condensates can be soft glassy materials that we call Maxwell glasses with age-dependent material properties. We discuss possible advantages of glassy behavior for modulation of cellular biochemistry.

Membraneless compartments in cells exhibit various material properties adapted to their biological purpose. Some compartments such as P-granules and stress granules are liquid-like, as shown by fusion and rapid molecular rearrangements (*1*, *2*). Others such as centrosomes are less dynamic and can exhibit properties of amorphous gels (*3*, *4*). Dynamic solids such as microtubules also exist (*5*). The study of rheological properties of biological compartments provides important information about the nature of molecular interactions and about the emergent, collective behaviors of the constituent molecules.

Under certain conditions, a protein solution in vitro can phase-separate to form protein-dense droplets (*6*, *7*). Such droplets can behave as liquids. However, in many cases, the dynamics of these liquid-like materials slow over time and behave more solid-like. This has been observed in vitro for a variety of evolutionarily distinct proteins involved in different biological functions (*2*, *3*, *6*–*10*) as well as in vivo (*4*, *9*, *11*–*14*). Such behavior has been termed hardening or aging (*2*, *10*).

We can distinguish several ways in which a liquid could become more solid-like (*15*–*17*). One way is gelation through the introduction of cross-links. When the cross-links reach a critical density, a connected polymer network is formed that extends through the whole system, and there is a sudden transition from a viscoelastic fluid to a solid (*17*, *18*). Another way for a liquid to become more solid-like is found in glass-forming systems, such as soft pastes and entangled polymer systems, which do not reach thermodynamic equilibrium (*16*, *19*–*21*). Glassy systems can be distinguished from gelation by the fact that there is no sudden change of material properties with age (*18*, *22*). Gels and glasses are typically amorphous. Liquids can also become more solid-like by the formation of fibers or crystal-like aggregates such as amyloids (*2*, *23*). Membraneless compartments in cells have been called condensates in part to emphasize their physical nature. However, in general, we do not understand their physical properties. Probing material properties of condensates requires understanding of the time-dependent rheology.

We began by characterizing the change in protein diffusion in condensates as a function of their age t_w (defined as time after formation of droplets), using fluorescence recovery after photobleaching (FRAP) (Fig. 1A and movie S1). Diffusion of nonbleached protein into the bleached region is a measure of molecular mobility. Figure 1B shows typical recovery curves at early and late age, for the *C. elegans* protein PGL-3 (fig. S1C) (*24*). At early age below 30 min, fluorescently labeled PGL-3 fluorescence recovers with a half-time of ~1 min. At age t_w = 46 hours, recovery is slow with a half-time of ~50 min (Fig. 1B). A second measure of dynamics is the coalescence time of two spherical condensates. Two condensates with liquid-like behavior merge into a final spherical condensate over a time that depends on the ratio of surface tension and viscosity (Fig. 1C) (*1*). After formation, coalescence of PGL-3 droplets occurs within 10 s, whereas at an age t_w = 46 hours, it takes tens of minutes (Fig. 1D and movie S2). Similar behaviors were seen for four other proteins of the FUS family [FUS, EWSR1, DAZAP1, and TAF15 (*10*)] (Fig. 1, A and C; fig. S1, A and C; and Table 1). These results are similar to previously published data and demonstrate the aging behavior of a wide variety of protein condensates (*2*, *3*, *6*, *7*, *9*, *25*).

To characterize the mechanism by which the material properties of condensates change between early and late times after their formation, we used both active (*26*) and passive microrheology. For active microrheology, condensates were deformed using two laser tweezers that acted on two beads linked to a droplet through attached protein (Fig. 1, E and F). Sinusoidal forces with a given frequency were exerted on one bead and the bead displacements were recorded, which served as a measure of droplet deformation (Fig. 2A). At an early age, the force-displacement curve is an ellipse showing a strong viscous component. At a later age, the force-displacement curve is almost a straight line indicative of predominantly elastic behavior (Fig. 1G). For passive microrheology, we followed the movement of fluorescent polystyrene beads immersed in droplets using a confocal microscope (Fig. 1H and movie S3) (*27*, *28*). We noticed that droplets decreased their size as they age (Fig. 1H). Furthermore, at early age, the diffusive bead motion reached larger distances as compared with later ages within the same time window (Fig. 1I). Despite the changes in material properties and droplet size for PGL-3 and FUS, cryo–electron tomography showed no detectable changes in structural features inside the drops, and the material appeared amorphous at all times. Occasional fibrous structures existed at the surface of late FUS droplets but not in the bulk (fig. S1B).

To further investigate the material properties, we measured the frequency-dependent complex shear modulus $G^* = G' + iG''$ as a function of the droplet age t_w after droplet formation using active microrheology (*26*). The storage modulus G' characterizes the elastic response of the droplets, and the loss modulus G'' characterizes the viscous component of the response. For PGL-3, at early ages and for low frequencies ω, the loss modulus is larger than the storage modulus, indicating predominantly liquid-like behavior (Fig. 2B, blue symbols). At high frequencies, the elastic modulus surpasses the loss modulus, indicating that at these frequencies, the mechanical response is predominantly elastic. The crossing-over between predominantly elastic and predominantly viscous behaviors occurs at the angular frequency ω_C. At this frequency, the storage modulus G' equals the loss modulus G''. The crossover frequency ω_C decreased as the sample aged (Fig. 2C, red symbols). The frequency-dependent responses obtained at early and late waiting times were different. However,

[1]Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany. [2]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstraße 108, 01307 Dresden, Germany. [3]Cluster of Excellence Physics of Life, Technische Universität Dresden, Dresden, Germany. [4]Biotec, TU Dresden, Tatzberg 47-49, 01307 Dresden, Germany. [5]EMBL Heidelberg, Meyerhofstraße 1, 69117 Heidelberg, Germany. [6]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), Dr. Bohrgasse 3, 1030 Vienna, Austria. [7]Center for Systems Biology Dresden, Pfotenhauerstr. 108, 01307 Dresden, Germany.
*Present address: Max Planck Institute for Dynamics and Self-Organization, Am Faßberg 17, 37077 Göttingen, Germany.
†Present address: Whitehead Institute for Biomedical Research 455 Main St, Cambridge, MA 02142, USA.
‡Present address: Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany.
§Present address: ETH Zürich, Vladimir-Prelog-Weg 1-5/10, 8093 Zürich, Switzerland.
¶Corresponding author. Email: hyman@mpi-cbg.de (A.A.H.); julicher@pks.mpg.de (F.J.)

they are both well described by a Maxwell fluid albeit with different characteristic relaxation times τ_C (Fig. 2, B and C, gray lines). For a Maxwell fluid the crossover frequency is $\omega_C = 1/\tau_C$, where τ_C is the Maxwell relaxation time. The time scale τ_C characterizes the droplet response to a deformation: For deformations on time scales shorter than τ_C, the droplet response is primarily elastic, whereas for deformations on time scales longer than τ_C, the response is primarily viscous and the droplet will flow. Therefore, our active microrheology data show that as the system ages, the time scale τ_C increases, indicating that the droplet exhibits elastic behaviors over an increasingly broad range of frequencies, corresponding to droplet aging. The scaling of frequency-dependent complex moduli is revealed by plotting the data with respect to a rescaled frequency axis. The data for different ages of PGL-3 droplets collapse on a single curve, which matches the behavior of a simple Maxwell fluid closely (Fig. 2D). We also performed active microrheology on protein condensates composed of FUS protein and again found rheology consistent with a Maxwell fluid. We repeated our measurements on PGL-3 under buffer conditions with varying salt concentrations (fig. S3, A to C) and with different temperatures (fig. S3, F to H, and supplementary materials). These data reveal that with increased temperature or increased salt concentration, droplet aging slows and droplets appear liquid-like for a longer period of time. At a late age, an additional short-time viscosity λ can be identified, which becomes

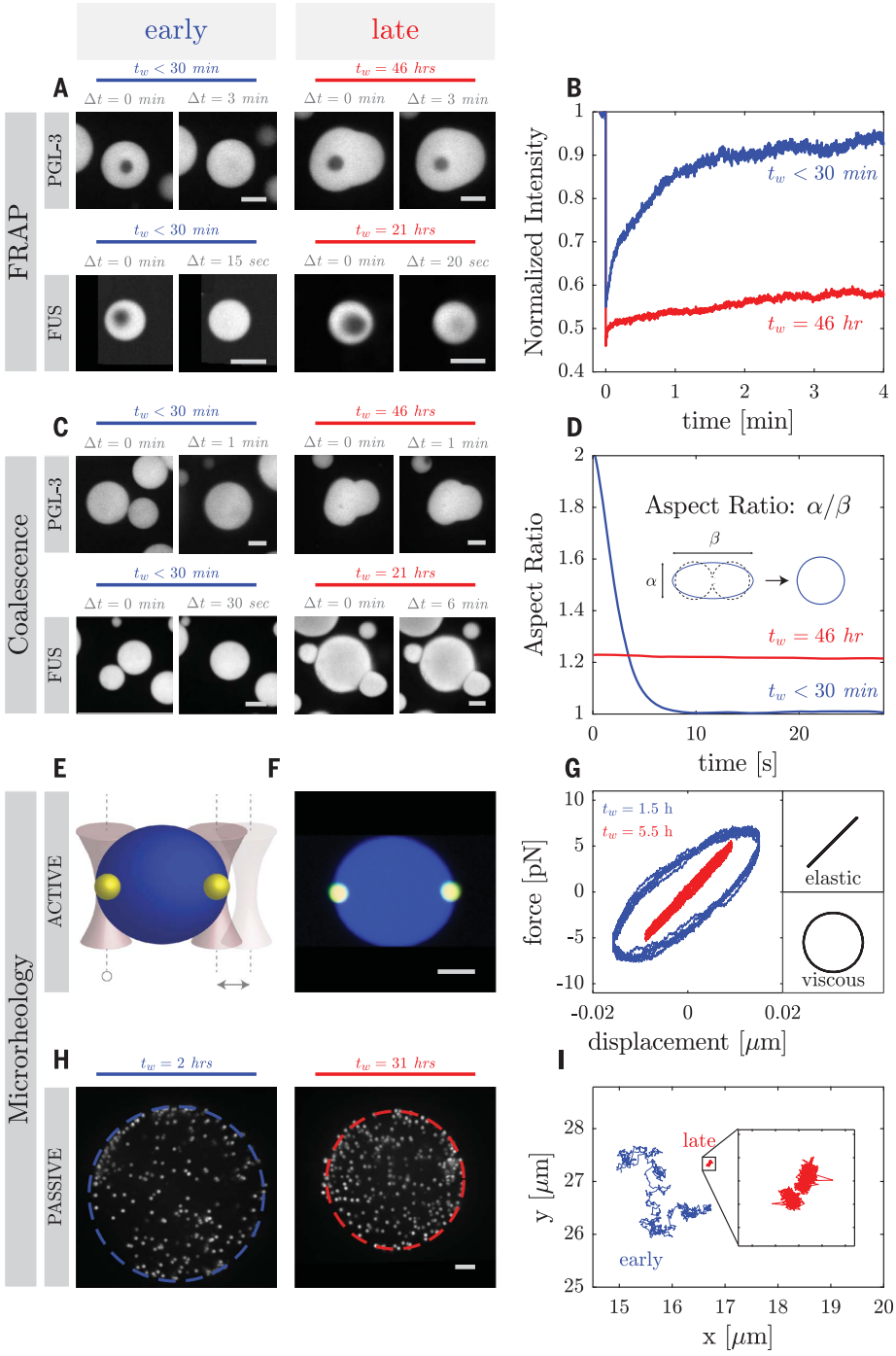


Fig. 1. Protein droplets exhibit age-dependent material properties. We compare the properties of droplets shortly after formation (t_w = 0 min, early droplets) with droplets observed after a time t_w of several hours of formation (late droplets). (**A**) FRAP of protein droplets composed of PGL-3–GFP and FUS-GFP, respectively. Shown are images of early and late droplets immediately after bleaching (Δt = 0 min) as well as after a time Δt after bleaching. (**B**) Fluorescence intensity of the photobleached area for the PGL-3–GFP droplets shown in (A) for both early (blue) and late (red) droplets. (**C**) Images of FUS-GFP and PGL-3–GFP droplets undergoing coalescence. Shown are images at the time of contact Δt = 0 as well as at a time Δt after coalescence. (**D**) Aspect ratio of the coalescing PGL-3–GFP droplets shown in (C) is shown as a function of time for early (blue) and late (red) droplets. The aspect ratio is defined as the ratio of the major to minor axes of a representative ellipse. (**E**) Schematic representation of the dual optical trap active microrheology experiment. Polystyrene beads are shown in yellow, and the droplet is shown in blue. One trap is fixed, and one trap oscillates. (**F**) Confocal image of a PGL-3–GFP droplet with two 2 μm–diameter beads in a dual optical trap. (**G**) Lissajous plot of force versus trap displacement for an early (blue) PGL-3 droplet indicating viscous behavior. The plot for the late (red) droplet indicating elastic behavior. (**H**) Confocal images of the midsection of a large PGL-3 droplet containing 1 μm–diameter fluorescent polystyrene beads early (blue) and late (red) after droplet formation. (**I**) Two examples of trajectories over 3 hours of embedded beads in a PGL-3 droplet, early (blue, t_w = 3 hours) and late (red, t_w = 36 hours) after droplet formation (length of inset box represents 0.5 μm). Scale bars in all panels represent 5 μm.

relevant at high frequencies and does not change significantly with age. This can be described by a Jeffreys fluid model (supplementary materials and fig. S8) (*29*). We determine λ values between 6 mPa·s and 1.3 Pa·s, which we attribute to the solvent. For FUS protein, we also observe an age-independent high-frequency viscosity in the same range (Fig. 4B and fig. S8, B and C). We conclude that at all conditions, the droplets exhibit Maxwell fluid rheology with age-dependent relaxation time.

Using passive microrheology, we measured the mean square displacement over time of beads immersed in the droplets (Figs. 1H and 3A and supplementary materials). We were concerned that the change in droplet radius would contribute to the mean square displacement of beads. Therefore, we first measured the droplet radius as a function of time (Fig. 3B, orange line). Droplet shrinking can also be estimated on the basis of net particle motion (Fig. 3, A and B, blue line). These two measures show good agreement, with small discrepancies at later ages (Fig. 3B, fig. S5A, and supplementary materials). The average particle displacement is proportional to distance from the droplet center, indicating that the droplet is shrinking uniformly in space (Fig. 3A). We quantified the mean square displacement

$$\langle \Delta r(\tau, t_w)^2 \rangle = \langle [r(t_w + \tau) - r(t_w)]^2 \rangle \quad (1)$$

of the beads over time τ for different droplet ages t_w, where r is the bead position in the image plane, after contributions from droplet

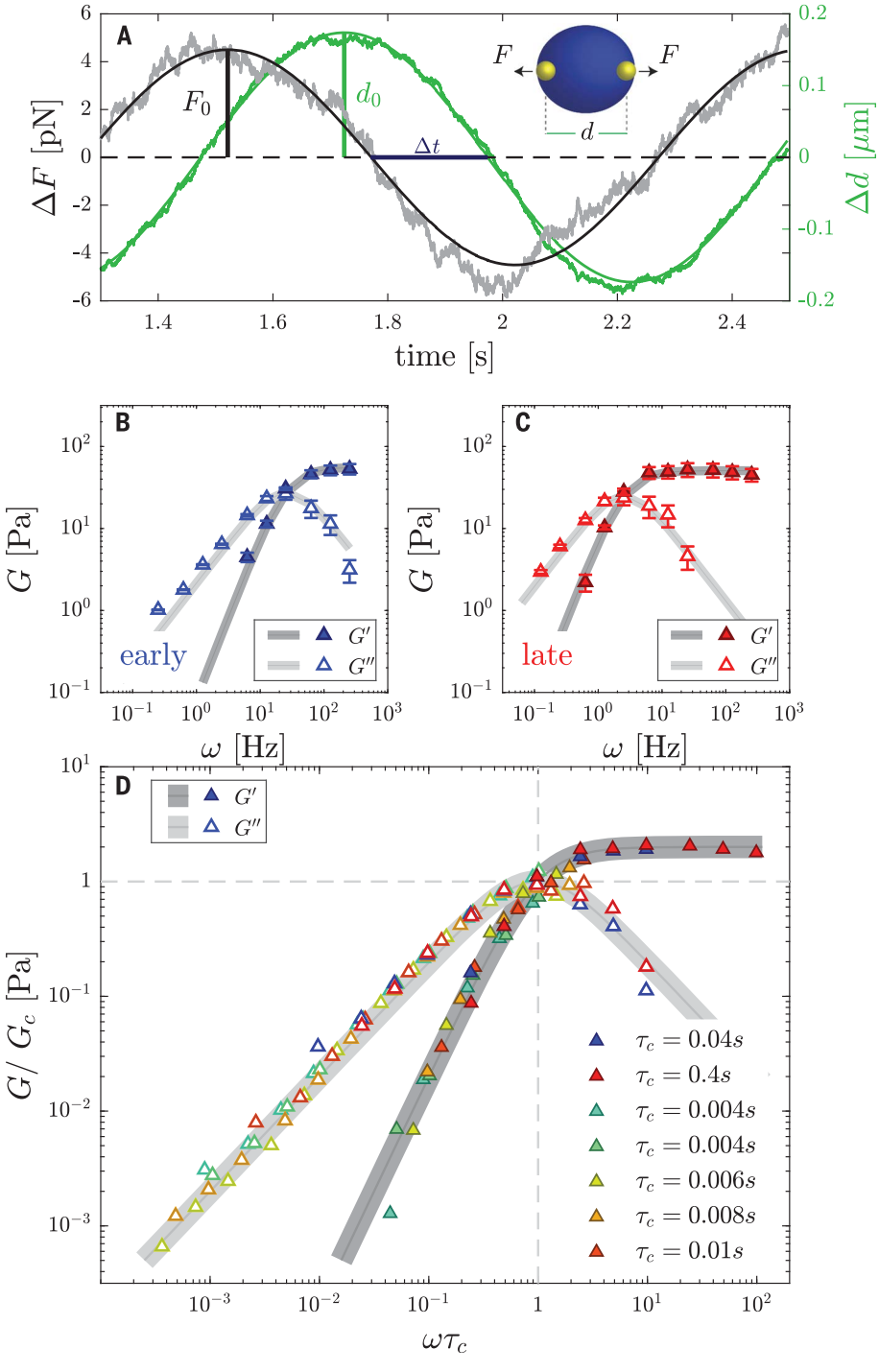


Fig. 2. Active rheology of early and late PGL-3 droplets. A droplet is trapped between two beads using a dual optical trap; one trap is undergoing oscillatory displacement while the other trap is at a fixed position. (**A**) Force ΔF is shown as a function of time (gray) together with a sinusoidal fit (black solid line). The corresponding droplet deformation Δd (green) is shown together with a sinusoidal fit (solid green line). We determine the force amplitude F_0, the deformation amplitude d_0, and the phase shift $\varphi = \omega\Delta t$ from the fits where ω is the imposed angular frequency. (**B**) Storage and loss moduli, G' and G'', are shown as a function of frequency for an early droplet (t_w = 0.5 hours with a surface tension of 4.5 μN/m). The corresponding frequency-dependent moduli of a Maxwell element with a characteristic time scale τ_C = 0.039 s and plateau modulus E = 56 Pa are shown in gray. Error bars represent measurement uncertainty (supplementary materials). (**C**) Same plot as in (B) for the same droplet late after formation (t_w = 1.5 hours with a surface tension of 19.3 μN/m). The complex elastic modulus of a Maxwell element with a characteristic time scale τ_C = 0.39 s and plateau modulus E = 50.7 Pa is shown in gray. Error bars represent measurement uncertainty (supplementary materials). (**D**) Scaled storage and loss moduli, G'/G_C and G''/G_C, are plotted as a function of scaled frequency $\omega\tau_C$ for the droplets shown in (B) and (C) (blue, red) and additional droplets (colors indicate characteristic time τ_C). G_C is the crossover modulus. A scaled Maxwell fluid is shown in gray. For all panels, the storage moduli G' are shown as solid triangles, and the loss moduli G'' are shown as open triangles. (B to D) Droplets were formed in a solution with a final KCl concentration of 75 mM.

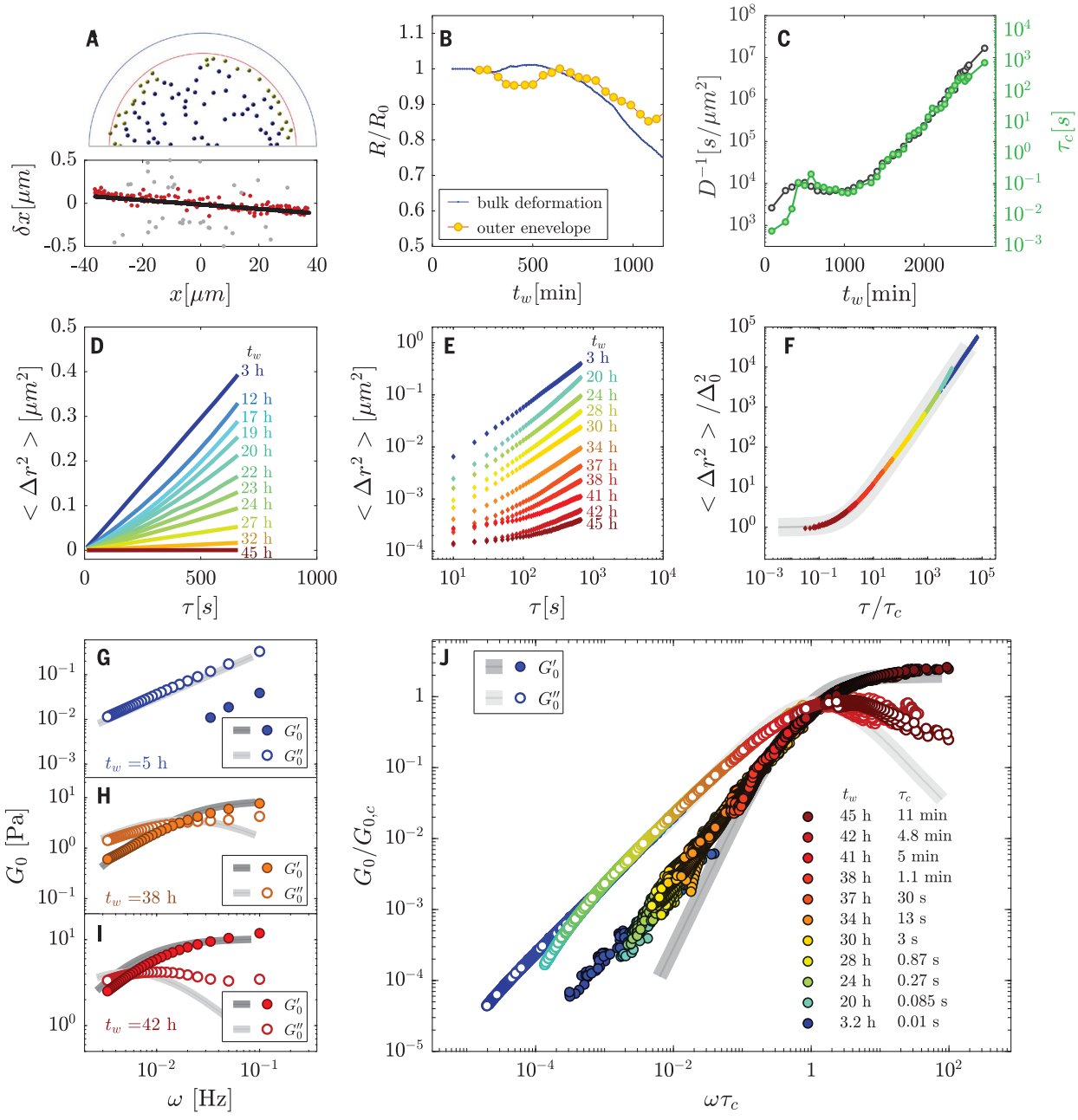


Fig. 3. Passive microrheology of early and late PGL-3 droplets. (**A**) Image of particles representing bead positions in the upper half of a droplet several hours after formation (t_w = 27 hours). Beads in yellow are considered boundary particles and are not considered for bead-diffusion analysis. The blue outline represents the droplet border shortly after droplet formation; the red outline represents the droplet border at t_w = 27 hours. Particle displacements δx in the x direction resulting from droplet shrinkage are shown below for a time interval of τ = 2.5 min at droplet age t_w = 8 hours plotted with respect to x position in the droplet. The slope of these particle displacements determined by a linear fit (black line) characterizes average particle motion due to droplet shrinkage. Data points within a standard deviation of the linear fit are shown in red, the remaining ones in gray. (**B**) Scaled droplet radius, R/R_0, is shown as a function of droplet age t_w, where R_0 is the initial droplet radius (orange). The scaled radius inferred from average particle displacements is shown in blue (supplementary materials). (**C**) Inverse of the diffusion coefficient D^{-1} of embedded beads calculated from the mean square displacement during τ = 20 min is shown as a function of waiting time t_w. The characteristic relaxation time τ_C is shown in green. (**D**) Mean square displacement $\langle \Delta r^2 \rangle$ as a function of lag time τ shown for different waiting times t_w. (**E**) Double logarithmic plot of $\langle \Delta r^2 \rangle$ as a function of lag time τ for different waiting times t_w. (**F**) Plot of normalized mean square displacement, $\langle \Delta r^2 \rangle / \Delta_0^2$, as a function of normalized lag time τ/τ_C, where Δ_0^2 is the offset and τ_C is a characteristic time. The values of τ_C that correspond to the mean square displacements shown in (E) are chosen such that the data for different waiting times t_w collapse onto a single curve, which is well described by $\langle \Delta r^2 \rangle = \Delta_0^2 \left(1 + \frac{\tau}{\tau_C}\right)$ (gray line). (**G** to **I**) Real and imaginary parts, G'_0 and G''_0, of the inferred complex shear moduli G_0^* for waiting times t_w = 5 hours (G), 38 hours (H), and 42 hours (I) (supplementary materials). (**J**) Normalized inferred shear moduli as a function of the normalized frequency $\omega\tau_C$ for all measurements of mean square displacement. The colors indicate the waiting time t_w, and the color code is the same as in (D) and (E). In addition, the value of characteristic times τ_C is indicated. The values of $G_{0,C}$ are inferred from the fit in (F) using the relationship $G_{0,C} = k_B T/(3\pi a \Delta_0^2)$. Inferred shear moduli shown in (G) to (J) were approximated using Eq. 14 of the supplementary materials. The gray lines in (G) to (J) correspond to the inferred complex moduli for a single-relaxation-time Maxwell fluid using the same approximation.

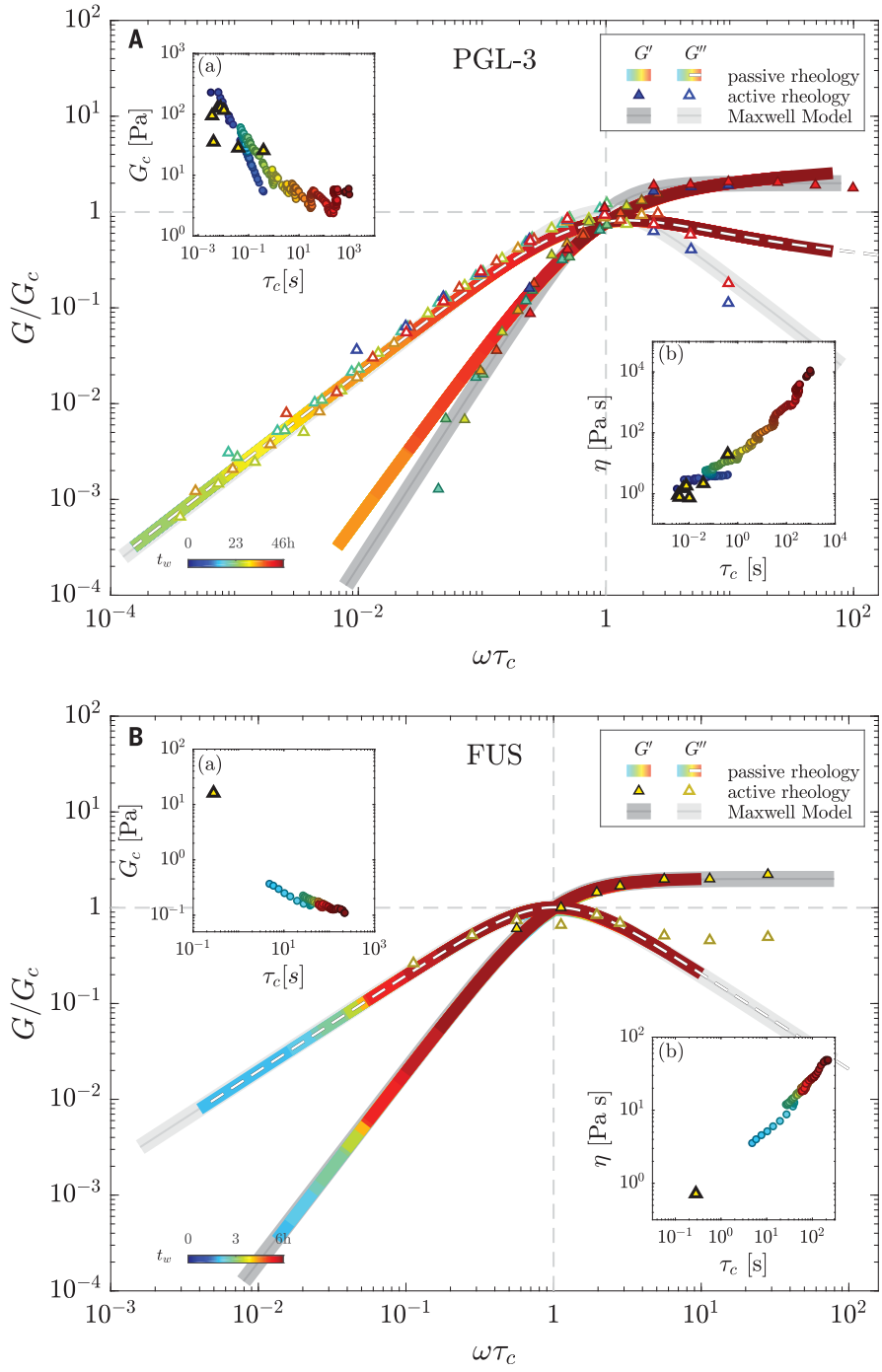


Fig. 4. Dimensionless loss and storage moduli obtained by active and passive microrheology collapse onto single curves. (**A**) Shear moduli G and G_0 scaled by the respective crossover moduli G_C and $G_{0,C}$ plotted as a function of dimensionless frequency $\omega\tau_C$. The active microrheology data are shown in open and solid triangles. The moduli inferred from passive microrheology as shown solid colored lines. These lines were obtained using passive microrheology data (Fig. 3F and supplementary materials). For active and passive microrheology, the same color code as in Figs. 2, B to D, and 3, E to J, is used. The thick gray lines represent the frequency response of a Maxwell fluid. (Inset a) Crossover moduli for active and passive microrheology G_C (yellow triangles) and $G_{0,C}$ (colored circles), respectively, are plotted as a function of characteristic time, τ_C. (Inset b) Viscosities for active and passive microrheology $\eta = 2G_C\tau_C$ (yellow triangles) and $\eta_0 = 2G_{0,C}\tau_C$ (colored circles), respectively, are plotted as a function of characteristic time, τ_C. Shear moduli inferred from passive microrheology were obtained using Eq. 18 of the supplementary materials. The gray lines correspond to the complex moduli of a single-relaxation-time Maxwell fluid. PGL-3 droplets were formed in a solution with a final KCl concentration of 75 mM. (**B**) Plot analogous to the upper plot with active and passive microrheology of FUS protein. FUS droplets were formed in a solution with final KCl concentrations of 75 mM and 250 mM for passive and active droplets, respectively.

shrinkage have been subtracted. We also corrected for the noise floor because of experimental noise (supplementary materials and fig. S7). The angle brackets denote an average over beads and over an age window of 1 hour (supplementary materials). We found that at early ages t_w, the mean square displacement increases linearly with time τ and reaches ~0.4 μm^2 within 10 min (Fig. 3, D and E, blue lines). For increasing droplet ages, the slope of the mean square displacement decreases (Fig. 3, D and E). The diffusivity can be determined through the slope of the mean square displacement as a function of time $\langle \Delta r(\tau, t_w)^2 \rangle \simeq \Delta_0^2 + 4D(t_w)\tau$, where Δ_0^2 is an offset that is independent of τ but can depend on age t_w. $D(t_w)$ is an age-dependent effective diffusivity; inverse $D(t_w)$ is shown in Fig. 3C, black line. As the system ages, $D(t_w)$ decreases steadily but remains almost constant with $D \simeq 10^{-4}$ $\mu m^2/s$ for t_w less than ~1000 min (15 hours). For passive, equilibrated samples, this value corresponds to a viscosity of ~1 Pa·s, which is consistent with previous estimates of PGL-3 droplet viscosity (*26*, *28*). For ages t_w greater than 1000 min, $D(t_w)$ decreases with age and reaches a value of $D \simeq 1.5 \cdot 10^{-7}$ $\mu m^2/s$ at ~2500 min (Fig. 3C). For passive liquids, this corresponded to a viscosity of ~10^3 Pa·s. When the mean square displacement curves are examined on a logarithmic scaling, the curve reaches a finite value at low τ, indicative of an elastic regime (Fig. 3, E and F). Consistent with the active microrheology, the mean square displacement data collapse to a single curve well described by a Maxwell fluid, when rescaling the time τ and $\langle \Delta r^2 \rangle$ by the age-dependent relaxation time τ_C and the offset Δ_0^2 (Fig. 3, D to F, and supplementary materials).

Passive microrheology assesses fluid properties from mean square displacement measurements using the fluctuation-dissipation relation (*27*, *30*). The fluctuation-dissipation relation applies to equilibrium systems. However, an aging system has not reached equilibrium. We reasoned that for sufficiently short time scales compared with the age t_w quasi-equilibrium of many degrees of freedom may hold (*31*), and the fluctuation-dissipation relation may provide useful estimates of material properties. We therefore determined the frequency-dependent complex shear modulus $G_0^* = G_0' + iG_0''$ inferred from the fluctuation-dissipation relation using the mean square displacement data (supplementary materials). The moduli G_0 obtained this way qualitatively exhibit the frequency-dependent behaviors of a Maxwell fluid with age-dependent Maxwell times (Fig. 3, G to I). The frequency-dependent moduli $G_0(\omega)$ as a function of frequency at different age t_w collapse on a single curve when the frequency axis is scaled with the time τ_C and the moduli are scaled by the value $G_{0,C}$, defined as the value of the real and imaginary parts of G_0^* at the point where both are equal

(Fig. 3J). Consistent with the active microrheology data, the passive microrheology data also exhibit at all ages the qualitative behaviors of a Maxwell fluid (compare Fig. 2D with Fig. 3J). We also performed passive microrheology on protein condensates composed of FUS protein and PGL-3–green fluorescent protein (GFP) (figs. S5 and S6). Our bead-tracking measurements on these condensates again reveal at all ages the qualitative properties of a single-time Maxwell fluid.

We highlight the similarity of active and passive microrheology by plotting the scaled shear moduli as a function of scaled frequency for both active (triangles) and passive (colored lines) microrheology in the same plot for both PGL-3 and FUS proteins (Fig. 4). The data collapse and are well captured by a single-time Maxwell fluid (Fig. 4, gray lines). When examining the moduli G_C, obtained for active microrheology taken from many aging droplets, we find values that vary over approximately one order of magnitude [Fig. 4, A and B, insets (a), yellow triangles]. The values of $G_{0,C}$ obtained by passive microrheology for one aging droplet exhibit a trend toward smaller values for larger τ_C [Figs. 3, G to I, and 4A, inset (a), colored circles]. We also estimated the long time viscosities, $\eta = 2G_C\tau_C$ and $\eta_0 = 2G_{0,C}\tau_C$, for active and passive microrheology, respectively [Fig. 4A, inset (b)]. The viscosity strongly increases with τ_C, and both active and passive microrheology share the same trend. To compare the inferred moduli determined by the active and passive microrheology, we define the ratio $\xi = \langle G_C \rangle / \langle G_{C,0} \rangle$, where brackets denote the average over several droplets and several values of τ_C. Using the window of $0.004\ s < \tau_C < 0.442\ s$, we find $\xi \approx 2.63 \pm 2.57$. One can interpret ξ as an effective noise strength that replaces the amplitude of fluctuations in the fluctuation-dissipation relation (*32*–*35*). The product ξT is sometimes called an effective temperature. This suggests that the fluctuation-dissipation relation is only weakly violated, at least, at early ages similar to previous reports on soft glasses (*32*, *36*). Our data further show that the elastic modulus does not tend to increase with time, suggesting that the term hardening is not appropriate for the change in material properties of condensates consisting of PGL-3 and FUS family proteins. Rather, we use the term aging.

A defining and unexpected rheological feature of the condensates we study is the self-similarity of their material properties as they age. This self-similarity is manifested by the fact that at all ages, they exhibit the same viscoelastic behavior of a Maxwell fluid. An ordinary Maxwell fluid does not age and reaches a thermodynamic equilibrium because it has a fixed relaxation time. However, the relaxation time of the protein condensates in this study increases with age. Age-dependent relaxation times are typical for glass-forming systems. We thus refer to a Maxwell fluid with age-dependent relaxation time as a Maxwell glass (supplementary materials). Our data suggest that the molecular components increasingly interact and lose their freedom to move as the condensates shrink and their density increases. Such behaviors are also seen in other glass-forming systems (*15*, *16*, *37*–*40*). The dependence of aging characteristics on salt concentration that we observe suggests that electrostatic interactions play a role in the aging of protein condensates (*41*).

A key question will be to relate glassy behavior to molecular configurations. The associative polymer ideas of Semenov and Rubenstein are a beautiful and elegant way to understand viscoelastic and Maxwell-like behaviors of polymer systems and describe the equilibrium properties and kinetics of such systems (*42*). However, they do not address aging. An example of how dynamics can slow is jamming (*43*). However, this is a simplified picture based on excluded volume and rigid objects such as spherical particles. For condensate-forming proteins, a more appropriate picture may be based on the fact that these proteins have large regions of disorder (*6*). Disordered proteins can take many different molecular configurations leading to complex energy landscapes with distributions of energy minima (*44*). Increasing relaxation times are likely related to the system exploring deeper energy minima that are increasingly inaccessible, but longer lived, as the system ages (*45*–*47*). Such ideas have been put forward in the context of protein folding (*44*, *48*, *49*) but still need to be explored in the context of higher-order protein assemblies.

In conclusion, we provide evidence that biological condensates made from a variety of proteins can best be described as Maxwell glasses. At any particular age, the rheology reveals the expected viscoelastic properties of a Maxwell fluid, well described by a single relaxation time. However, at different ages, the relaxation times themselves are different. Therefore, the term Maxwell glass incorporates the aging of Maxwell fluids and glasses. The role of glass-like behaviors have not been well characterized in biology, which has tended to stress gel-like characteristics of biological materials (*7*, *50*). This is because glass-like aging is a dynamic process and, therefore, more difficult to study because it requires a broad range of time scales. However, glass-like behavior offers a number of interesting possibilities for cellular function compared with a gel-like state. A soft gel will either respond elastically or break under stress and tends to require chemical modification for disassembly. On the other hand, because glasses can start to flow when mechanically stressed (*51*), glass-like soft pastes offer a tension sensor to cells that can couple mechanical stresses to biochemistry. The nature of glass-like states makes it easy to fluidize a biological glass simply by changing composition, temperature, or salt concentration. More generally, soft glasses allow cells to slow down biochemistry while maintaining a soft material, which can flexibly and rapidly respond to changing conditions. Little is known about the relationship between pathology and stiffness of aggregates, but stiffer structures likely could be more disruptive to the mechanical environment of a cell. Cells in tissues are constantly responding to mechanical perturbation from other cells in the tissue and must be able to respond as soft materials. Glass-like aging of protein condensates may, therefore, offer cells a way to flexibly modulate mechanical properties of membraneless organelles while allowing for rapid response to changing environmental cues.

Table 1. Characteristics of the proteins used in this study. The first four proteins are from the FUS family. Protein length is reported in number of amino acids. All molecular weights and sequence lengths are reported for untagged proteins. The addition of a GFP tag adds 29 to 31 kDa in molecular weight and 260 to 278 amino acids in length. The GFP tag is on the N terminus for DAZAP1, whereas all other proteins have the GFP tag on the C terminus.

Protein name	Length	Molecular weight [Da]	Plasmid name	Virus number	Gene
DAZAP1	407	43,384	TH1272	1424	DAZAP1
EWSR1	656	68,480	TH1259	1510	EWSR1
FUS	526	53,426	TH1204	3311	FUS
TAF15	589	61,559	TH1258	1480	TAF15
PGL-3	693	74,700	pSS2B	1522	PGL3

REFERENCES AND NOTES

1. C. P. Brangwynne *et al*., *Science* **324**, 1729–1732 (2009).
2. A. Patel *et al*., *Cell* **162**, 1066–1077 (2015).
3. J. B. Woodruff *et al*., *Cell* **169**, 1066–1077.e10 (2017).
4. M. Mittasch *et al*., *J. Cell Biol.* **219**, e201912036 (2020).
5. A. Desai, T. J. Mitchison, *Annu. Rev. Cell Dev. Biol.* **13**, 83–117 (1997).
6. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
7. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
8. Y. Lin, D. S. Protter, M. K. Rosen, R. Parker, *Mol. Cell* **60**, 208–219 (2015).
9. T. M. Franzmann *et al*., *Science* **359**, eaao5654 (2018).
10. J. Wang *et al*., *Cell* **174**, 688–699.e16 (2018).
11. A. R. Strom *et al*., *Nature* **547**, 241–245 (2017).
12. K. E. Kistler *et al*., *eLife* **7**, e37949 (2018).
13. S. J. Nair *et al*., *Nat. Struct. Mol. Biol.* **26**, 193–203 (2019).
14. S. Jang *et al*., *Biophys. J.* S0006-3495(20)30597-X (2020).
15. J. M. Hutchinson, *Prog. Polym. Sci.* **20**, 703–760 (1995).
16. C. B. Roth, *Polymer Glasses* (CRC, 2016).
17. M. Rubinstein, R. H. Colby, *Polymer Physics*, vol. 23 (Oxford Univ. Press, 2003).
18. H. H. Winter, M. Mours, in *Neutron Spin Echo Spectroscopy Viscoelasticity Rheology*, vol. 134 of *Advances in Polymer Science* (Springer, 1997), pp. 165–234.
19. L. Berthier, G. Biroli, *Rev. Mod. Phys.* **83**, 587–645 (2011).
20. H. Lindsay, P. Chaikin, *J. Chem. Phys.* **76**, 3774–3781 (1982).
21. A. Ikeda, L. Berthier, P. Sollich, *Soft Matter* **9**, 7669 (2013).
22. A. S. Negi, C. G. Redmon, S. Ramakrishnan, C. O. Osuji, *J. Rheol.* **58**, 1557–1579 (2014).
23. M. Kato *et al*., *Cell* **149**, 753–767 (2012).
24. S. Saha *et al*., *Cell* **166**, 1572–1584.e16 (2016).
25. S. Maharana *et al*., *Science* **360**, 918–921 (2018).
26. L. M. Jawerth *et al*., *Phys. Rev. Lett.* **121**, 258101 (2018).
27. M. L. Gardel, M. T. Valentine, D. A. Weitz, *Microrheology* (Springer, 2005).
28. S. Elbaum-Garfinkle *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7189–7194 (2015).
29. Y. L. Raikher, V. V. Rusakov, R. Perzynski, *Soft Matter* **9**, 10857 (2013).
30. T. G. Mason, *Rheol. Acta* **39**, 371–378 (2000).
31. L. C. E. Struik, thesis, Technische Hogeschool Delft (1977).
32. S. Jabbari-Farouji *et al*., *Europhys. Lett.* **84**, 20006 (2008).
33. L. Cipelletti, L. Ramos, *J. Phys. Condens. Matter* **17**, R253–R285 (2005).
34. L. F. Cugliandolo, J. Kurchan, L. Peliti, *Phys. Rev. E* **55**, 3898–3914 (1997).
35. L. F. Cugliandolo, *J. Phys. A* **44**, 483001 (2011).
36. B. Abou, F. Gallet, *Phys. Rev. Lett.* **93**, 160603 (2004).
37. R. L. Moorcroft, M. E. Cates, S. M. Fielding, *Phys. Rev. Lett.* **106**, 055502 (2011).
38. C. Derec, A. Ajdari, F. Lequeux, *Eur. Phys. J. E* **4**, 355–361 (2001).
39. G. B. McKenna, S. L. Simon, in *Polymer Glasses* (CRC, 2016), pp. 39–70.
40. B. Ruzicka, E. Zaccarelli, *Soft Matter* **7**, 1268 (2011).
41. J. van der Gucht, E. Spruijt, M. Lemmers, M. A. Cohen Stuart, *J. Colloid Interface Sci.* **361**, 407–422 (2011).
42. M. Rubinstein, A. N. Semenov, *Macromolecules* **34**, 1058–1068 (2001).
43. A. J. Liu, S. R. Nagel, *Annu. Rev. Condens. Matter Phys.* **1**, 347–369 (2010).
44. Y. Chebaro, A. J. Ballard, D. Chakraborty, D. J. Wales, *Sci. Rep.* **5**, 10386 (2015).
45. J. P. Bouchaud, *J. Phys. I* **2**, 1705–1713 (1992).
46. P. Sollich, F. Lequeux, P. Hébraud, M. E. Cates, *Phys. Rev. Lett.* **78**, 2020–2023 (1997).
47. C. Monthus, J.-P. Bouchaud, *J. Phys. Math. Gen.* **29**, 3847–3869 (1996).
48. J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
49. S.-H. Chong, S. Ham, *Sci. Rep.* **9**, 14927 (2019).
50. M. Hondele, S. Heinrich, P. De Los Rios, K. Weis, *Emerg. Top. Lif Sci*. (2020).
51. M. Cloitre, R. Borrega, L. Leibler, *Phys. Rev. Lett.* **85**, 4819–4822 (2000).

ACKNOWLEDGMENTS

We thank the members of MPI-CBG and MPI-PKS for useful discussions; in particular, we thank C. Duclut, P. Mccall, S. Muenster, L. Hubatsch, T. Harmon, and C. Weber. We also thank B. Schroth-Diez, J. Peychl, and the light microscopy facility as well as R. Lemaitre and the protein expression facility at the MPI-CBG. **Funding:** This project was supported by funding from the Max Planck Society. A.A.H. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (ZE1216/1-1; HY3/6-1). A.A.H., F.J., and E.F.-F. were supported by the Deutsche Forschungsgemeinschaft (DFG,German Research Foundation) under Germany's Excellence Strategy (EXC-2068; 390729961) Cluster of Excellence Physics of Life of TU Dresden. J.M. acknowledges funding from EMBL and the European Research Council (760067). **Author contributions:** L.J., A.A.H., and F.J. designed and coordinated the study. L.J. performed experimental work, image processing, and data analysis. E.F.-F. preformed optical tweezer data analysis. L.J., E.F.-F., F.J., A.A.H., and Su.S. performed theoretical work. Sh.S., M.R., and J.W. provided proteins and reagents. M.I. performed initial optical tweezer experiments. T.F., X.Z., J.S., and J.M. preformed cryo–electron microscopy experiments. All authors contributed to data analysis and interpretation. **Competing interests:** A.A.H. is a co-founder and SAB member of Dewpoint therapeutics. A.A.H. is also a co-founder of Caraway therapeutics. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1317/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
References (*52*–*59*)
Movies S1 to S3

View/request a protocol for this paper from *Bio-protocol*.

5 February 2019; resubmitted 18 June 2019
Accepted 16 October 2020
10.1126/science.aaw4951

ULTRACOLD CHEMISTRY

Resonant collisional shielding of reactive molecules using electric fields

Kyle Matsuda[1]*, Luigi De Marco[1], Jun-Ru Li[1], William G. Tobias[1], Giacomo Valtolina[1], Goulven Quéméner[2], Jun Ye[1]*

Full control of molecular interactions, including reactive losses, would open new frontiers in quantum science. We demonstrate extreme tunability of ultracold chemical reaction rates by inducing resonant dipolar interactions by means of an external electric field. We prepared fermionic potassium-rubidium molecules in their first excited rotational state and observed a modulation of the chemical reaction rate by three orders of magnitude as we tuned the electric field strength by a few percent across resonance. In a quasi–two-dimensional geometry, we accurately determined the contributions from the three dominant angular momentum projections of the collisions. Using the resonant features, we shielded the molecules from loss and suppressed the reaction rate by an order of magnitude below the background value, thereby realizing a long-lived sample of polar molecules in large electric fields.

Controlling chemical reactions and collisions has been a central focus of work on cold and ultracold molecules (*1*–*6*). Progress in cooling and trapping molecules has led to exciting advances in this area, including the precise characterization of scattering resonances (*7*–*9*), the observation of atom-molecule (*10*) and molecule-molecule (*11*–*13*) cold collisions, and the synthesis of new chemical species (*14*). In particular, ultracold polar molecules, for which both internal and external degrees of freedom are controlled, present distinctive opportunities (*15*–*26*). At ultralow temperatures, small perturbations to the long-range intermolecular potential, although negligible relative to chemical bonding energy scales, can vastly exceed the kinetic energy of the colliding molecules and thus strongly alter the rate of chemical reactions at close range (*3*). This sensitivity, combined with the rich structure of polar molecules and their tunability using external electromagnetic fields, suggests the exciting possibility of precisely controlling reactions (*1*, *4*–*6*). In addition to providing insights about fundamental chemical processes (*27*, *28*), such control would aid in the production of quantum-degenerate molecular gases (*29*–*31*) and could facilitate precision measurements (*32*) or studies of many-body physics (*33*) in these systems.

Applying an electric field $\mathbf{E}$ strongly modifies the reaction rates of ultracold polar molecules via dipolar interactions (*34*, *35*). In the ultracold regime, molecules collide predominantly in the lowest partial wave L allowed by quantum statistics ($L = 1$ for identical fermions). Owing to the anisotropy of the dipolar interaction, the likelihood of two molecules meeting at short range depends on how they approach each other relative to the direction of the induced dipole, or more formally, on the projection m_L of L onto the axis of $\mathbf{E}$. For fermionic molecules in three dimensions (3D), attractive head-to-tail ($m_L = 0$) collisions lead to rapid losses (*34*), which scale as d^6 in the induced dipole moment d (*36*). If the molecules are instead trapped in a quasi–two-dimensional (quasi-2D) geometry with $\mathbf{E}$ along the strongly confined direction, only repulsive side-to-side ($m_L = \pm 1$) collisions are allowed, suppressing losses and enhancing the elastic collision rate (*37*, *38*). For ground-state potassium-rubidium (KRb) molecules, we recently used this approach to achieve a ratio of elastic collisions to reactive collisions exceeding 100 (*31*).

Here, we experimentally demonstrate a striking effect of the electric field on molecular collisions: Chemical reaction rates in an ultracold gas of molecules are sharply varied by three orders of magnitude near particular values of the field strength $|\mathbf{E}|$. These values occur where higher rotationally excited states become degenerate with the initial collision channel, inducing resonant dipolar interactions that profoundly alter the long-range potential and hence the reaction rate (*39*). Although losses can also be resonantly enhanced, the most important effect is the substantial suppression of loss for an appropriate choice of $|\mathbf{E}|$. This shielding mechanism was first proposed by Avdeenkov *et al.* (*40*), and related theory was subsequently extended to a wide variety of bosonic and fermionic species of experimental interest (*39*, *41*, *42*). Alternatively, the use of microwave (*43*–*47*) or optical (*48*) dressing to suppress molecular loss has been proposed. However, microwave dressing has so far only led to an enhanced loss rate in experiments (*49*, *50*).

We prepared ultracold fermionic $^{40}K^{87}Rb$ molecules in the $|N, m_N\rangle = |1,0\rangle$ state, where N is the rotational angular momentum and m_N is its projection onto the axis of $\mathbf{E}$. Throughout, $|N, m_N\rangle|N', m'_N\rangle$ denotes the combined molecular state of a pair of molecules with one molecule in $|N, m_N\rangle$ and the other in $|N', m'_N\rangle$. We observed a drastic change in the two-body reactive loss rate near two field strengths, $|\mathbf{E}_1| = 11.72$ kV/cm and $|\mathbf{E}_2| = 12.51$ kV/cm, where the energies of $|0,0\rangle|2,\pm1\rangle$ and $|0,0\rangle|2,0\rangle$ (respectively) cross the energy of $|1,0\rangle|1,0\rangle$ (Fig. 1A). Near these crossings (Fig. 1B), the nearly degenerate states are resonantly coupled by dipolar interactions, becoming strongly mixed as the molecules approach to separations $r \sim r_0$ during a collision event, where $r_0 = 270a_0$ is the radius of the p-wave centrifugal barrier (*51*) and a_0 is the Bohr radius.

The consequence of this r-dependent state mixing is apparent in the adiabatic energy curves near $|\mathbf{E}_2|$ (Fig. 1C) (*52*). For $|\mathbf{E}| > |\mathbf{E}_2|$ (orange line), the energy of $|1,0\rangle|1,0\rangle$ is higher than that of $|0,0\rangle|2,0\rangle$. Hence, coupling between the states causes an increasing energy of $|1,0\rangle$ molecules as they approach, creating a repulsive barrier (with height ~ 300 μK) that is three orders of magnitude larger than the typical collision energy set by the temperature of the gas (250 nK). In this case, molecules are shielded from reactive losses, because they cannot meet at short range except by tunneling through the barrier (which occurs with a low probability). Conversely, for $|\mathbf{E}| < |\mathbf{E}_2|$ (green line), the energy of $|1,0\rangle|1,0\rangle$ is lower than that of $|0,0\rangle|2,0\rangle$, resulting in an attractive interaction and an enhanced loss rate. For comparison, the diabatic energy curve (black line) shows the behavior in the absence of resonant dipolar interactions. Although lossy collisions are suppressed for $|\mathbf{E}|$ just above resonance, elastic dipolar collisions for KRb are predicted to be nearly unaffected by the shielding (*39*). In general, the shielding is predicted to increase the ratio of elastic collisions to reactive collisions, potentially allowing for efficient evaporative cooling of molecules in 3D (*42*) and complementing the recently demonstrated evaporation in 2D (*31*).

This effect is akin to a Förster resonance—for example, between Rydberg atoms (*53*–*55*), in which $|\mathbf{E}|$ is tuned to create degeneracies between pairs of dipole-coupled states, resulting in resonant energy transfer. A key difference is the much smaller dipole moment of molecules relative to that of Rydberg atoms. Consequently, colliding molecules experience an adiabatic increase in the dipolar interaction energy as they approach. We stress that this is not a conventional scattering resonance arising from the presence of a molecule-molecule

[1]JILA, National Institute of Standards and Technology, and Department of Physics, University of Colorado, Boulder, CO 80309, USA. [2]Université Paris-Saclay, CNRS, Laboratoire Aimé Cotton, 91405 Orsay, France.
*Corresponding author. Email: kyle.matsuda@colorado.edu (K.M.); ye@jila.colorado.edu (J.Y.)

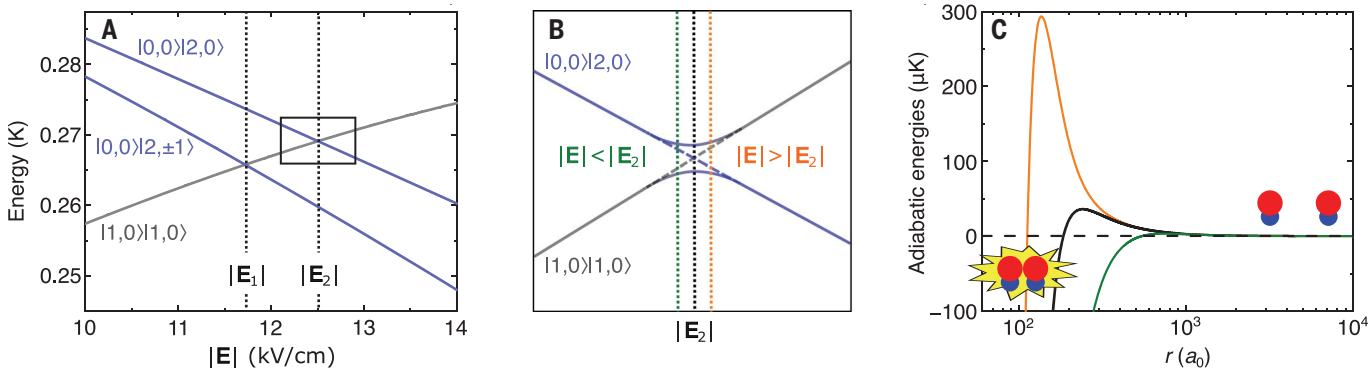


Fig. 1. Electric field–induced shielding. (**A**) Energies of the relevant combined molecular states as a function of $|\mathbf{E}|$, displaying the crossing of $|1,0\rangle|1,0\rangle$ with $|0,0\rangle|2,\pm1\rangle$ and $|0,0\rangle|2,0\rangle$. (**B**) Qualitative picture of the region near $|\mathbf{E}_2|$ (black dotted line). For large molecular separations $r >> r_0$ (dashed lines), the two states are not coupled. For separations $r \sim r_0$ (solid lines), an avoided crossing is opened as a result of dipolar interactions between the states. As two $|1,0\rangle$ molecules collide at a fixed $|\mathbf{E}| > |\mathbf{E}_2|$ (orange dotted line), the avoided crossing results in an effective repulsion; at a fixed $|\mathbf{E}| < |\mathbf{E}_2|$ (green dotted line), the avoided crossing results in an effective attraction. The same argument applies near $|\mathbf{E}_1|$. (**C**) Adiabatic energy curves for KRb molecules in $|1,0\rangle$ colliding with $m_L = \pm1$ at $|\mathbf{E}| = 12.504$ kV/cm (green curve) and 12.670 kV/cm (orange curve), corresponding to the values of $|\mathbf{E}|$ indicated by the green and orange vertical dotted lines in (B). The diabatic energy curve at 12.670 kV/cm (black solid curve) shows a barrier (36 μK) from the sum of the centrifugal (24 μK) barrier (*51*) and the semiclassical dipolar (12 μK) barrier (*31*, *37*, *38*) in the absence of resonant dipolar interactions. The dashed line shows the average collision energy of 250 nK. The red and blue spheres represent Rb and K, respectively.

bound state, but rather a resonance between two free scattering states enabled by the internal structure of the molecules (*39*).

The experimental setup has been described in detail previously (*31*). In brief, a degenerate mixture of ^{40}K and ^{87}Rb was prepared in six layers of a 1D optical lattice, with final trap frequencies $(\omega_x, \omega_y, \omega_z) = 2\pi \times (34, 17.7 \times 10^3, 34)$ Hz for KRb in each layer (gravity points along $-\hat{y}$). Weakly bound molecules were created with a magnetic field ramp through an interspecies Feshbach resonance at 546.62 G and were transferred to the ro-vibronic ground state by stimulated Raman adiabatic passage (STIRAP) at $|\mathbf{E}_{\mathrm{STIRAP}}| = 4.5$ kV/cm with the field along $+\hat{y}$. For studying the $|1,0\rangle$ state, it was advantageous to perform STIRAP at large $|\mathbf{E}|$ to bypass several avoided crossings at $|\mathbf{E}| <$ 1 kV/cm that arise from the hyperfine structure (*21*). Typical starting conditions were 2×10^4 molecules in the $|0,0\rangle$ state at a temperature $T = 250$ nK, corresponding to about 1.8 times the Fermi temperature. With nearly perfect occupancy of the lowest band ($k_B T/\hbar\omega_y \sim 0.3$, where k_B is the Boltzmann constant and $\hbar$ is the reduced Planck constant) and negligible tunneling between lattice sites, our system realized a stack of quasi-2D molecular gases.

To measure the reactive loss of the $|1,0\rangle$ state, we used the following protocol. Starting at $\mathbf{E}_{\mathrm{STIRAP}}$, we first applied a microwave π-pulse to transfer the molecules from $|0,0\rangle$ to $|1,0\rangle$ with a Rabi frequency of $2\pi \times 200$ kHz and a typical efficiency above 95%. Any remaining $|0,0\rangle$ population was quickly lost in a few milliseconds via *s*-wave reactive collisions with $|1,0\rangle$. Next, $\mathbf{E}$ was ramped to its target configuration in 60 ms. After a variable hold time t, $\mathbf{E}$ was ramped back to $\mathbf{E}_{\mathrm{STIRAP}}$ in 60 ms. To image $|1,0\rangle$ molecules, we applied another microwave pulse to transfer the molecules back to the $|0,0\rangle$ state, then used STIRAP to transfer to the Feshbach state before imaging the molecules in time-of-flight expansion. A constant magnetic field of 545.5 G was present during the measurements. We fitted the measured average density n as a function of t to the solution of the two-body loss rate equation $dn/dt = -\beta n^2$, where β is the two-body chemical reaction rate coefficient. In 2D, there is no temperature increase associated with the two-body loss (*31*, *56*), hence this rate equation is simplified in comparison to 3D (*29*).

To fully characterize the shielding effect, we first show how to tune the angular momentum character of the collisions by changing the orientation of $\mathbf{E}$ relative to the quasi-2D planes.

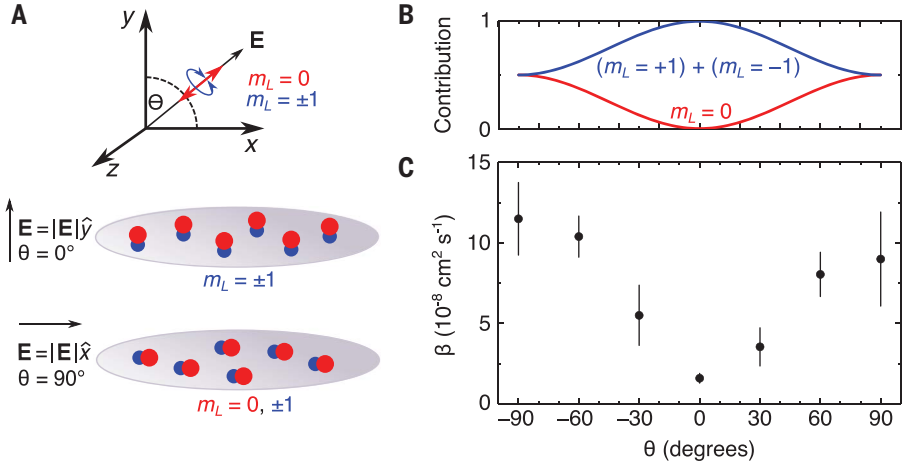


Fig. 2. Experimental setup and control of reactions by the electric field orientation. (**A**) Schematic of the experiment geometry. Owing to the strong confinement along $\hat{y}$, the m_L states participating in the collisions can be controlled with θ. (**B**) Contribution of the $m_L = 0$ (red) or $m_L = \pm1$ (blue) channels to the scattering as a function of θ. (**C**) β as a function of θ for a fixed bias strength $|\mathbf{E}| = 7.09$ kV/cm. Error bars are 1 SE from fits to the two-body rate equation.

Previous measurements in quasi-2D with $\mathbf{E}$ oriented along the tightly confined direction ($\hat{y}$) showed a suppression of β at moderately large values of d owing to repulsive $m_L = \pm1$ dipolar collisions (*31*, *37*). Here, we studied the dipolar anisotropy by tilting $\mathbf{E}$ away from the y axis by an angle θ, which enabled controlled mixing of the attractive $m_L = 0$ collisions into the scattering (Fig. 2A). Although the dipolar interaction in general mixes higher partial waves into the scattering, contributions from $|m_L| > 1$ are negligible for the relatively small values of d explored here (*52*).

Although the collisions always occurred along the $\hat{x}$ and $\hat{z}$ directions because of the strong confinement along $\hat{y}$, the angular momentum

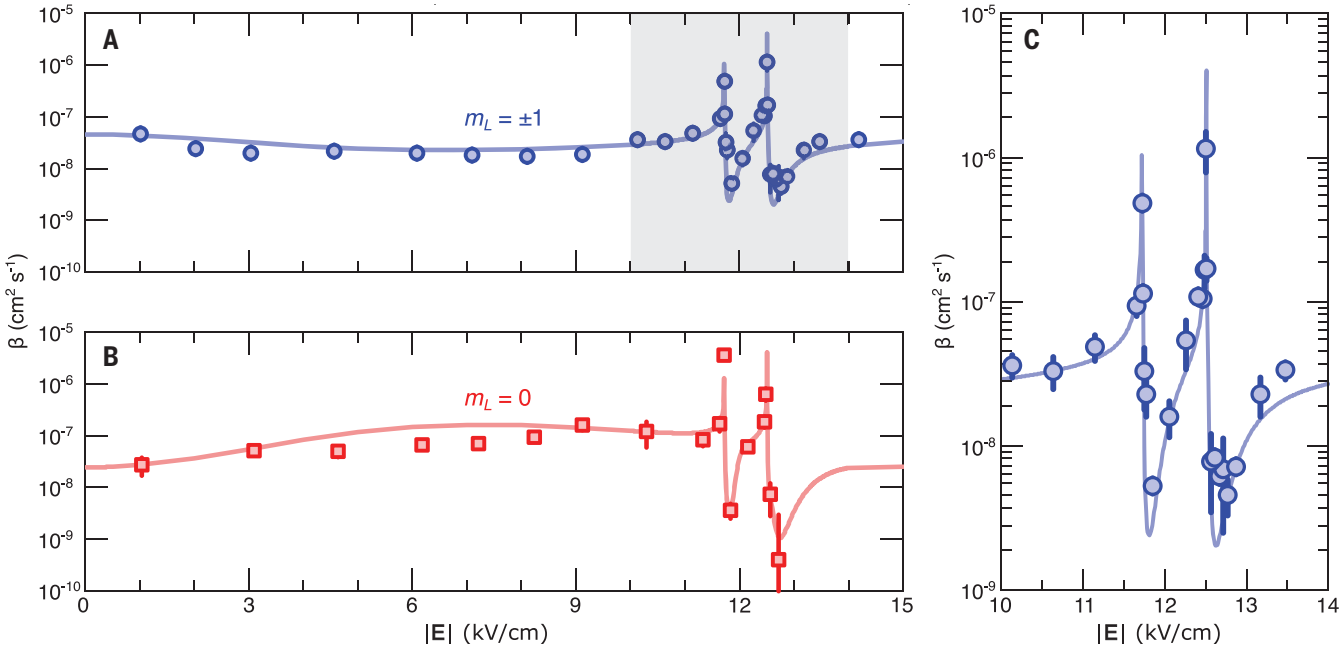


Fig. 3. 3D characterization of the shielding effect. (**A**) $\beta_{\pm1}$ (blue circles) versus $|\mathbf{E}|$ extracted from loss measurements at $\theta = 0°$. (**B**) β_0 (red squares) versus $|\mathbf{E}|$ extracted from loss measurements at $\theta = 90°$ and 0°. In (A) and (B), solid lines are theoretical predictions for the experimental $T = 250$ nK and $\omega_y = 2\pi \times 17.7$ kHz with no free parameters. Error bars are 1 SE from fits to the two-body rate equation. (**C**) Close-up of $\beta_{\pm1}$ in the region near $|\mathbf{E}_1|$ and $|\mathbf{E}_2|$ [gray shaded region of (A)].

character of the collisions with respect to $\mathbf{E}$ changed with θ. For $\theta = 0°$, the collisions decomposed equally into the $m_L = \pm1$ channels, which give equal contributions to the collision cross section as a result of the azimuthal symmetry of the dipolar interaction. For $\theta = 90°$, collisions along $\hat{x}$ corresponded to $m_L = 0$ scattering, and those along $\hat{z}$ were still an equal superposition of the $m_L = \pm1$ channels. Hence, by measuring β at $\theta = 0°$ and 90°, it was possible to extract the loss rate coefficients $\beta_{\pm1}$ and β_0 associated with the $m_L = \pm1$ and 0 channels, respectively. The full dependence on θ (Fig. 2B) was calculated by considering the mixing of the m_L states under rotations (*52*).

Figure 2C shows the measured β for $|1,0\rangle$ molecules at $|\mathbf{E}| = 7.09$ kV/cm as θ was varied over 180°. As expected, β increased with $|\theta|$ and reached a maximum at $\theta = \pm 90°$, where attractive $m_L = 0$ collisions dominated the loss rate. At 7.09 kV/cm, the relatively small value of $d = -0.12$ D limited the maximum increase of β to only an order of magnitude, in contrast to the much larger effect expected for larger $|d|$ (*57*). Our electrode geometry permitted excellent control of the curvature of $\mathbf{E}$ along $\hat{x}$, except near $\theta = \pm 90°$ (fig. S1), where we applied a small correction to the measured β to account for compression of the cloud attributable to the increased curvature in this configuration (*52*).

Having controlled the angular momentum channels participating in the collisions, we proceeded to explore the dependence of the $|1,0\rangle$ reaction rate on $\mathbf{E}$. We measured β at both $\theta = 0°$ and 90° to extract $\beta_{\pm1}$ and β_0 as a function of $|\mathbf{E}|$ (*52*), as summarized in Fig. 3, A and B, respectively. For both values of θ, we calibrated $|\mathbf{E}|$ to a few parts in 10^4 using spectroscopy on the $|0,0\rangle$ to $|1,0\rangle$ transition.

In the background region ($|\mathbf{E}| = 1$ to 11 kV/cm) away from resonance, we observed a slight decrease in $\beta_{\pm1}$ and a corresponding increase in β_0. $\beta_{\pm1}$ reached a minimum and β_0 reached a maximum near $|\mathbf{E}| = 7$ kV/cm, in agreement with theoretical predictions (Fig. 3, A and B, solid lines) (*52*). To understand the trends of $\beta_{\pm1}$ and β_0, we note that $|d|$ is nonmonotonic in the investigated range of $|\mathbf{E}|$ and reaches a maximum of 0.12 D at 7 kV/cm. Thus, the trends of $\beta_{\pm1}$ and β_0 are consistent with the semiclassical picture of repulsive side-to-side ($m_L = \pm1$) or attractive head-to-tail ($m_L = 0$) dipolar collisions modifying the loss rate, as previously measured for the $N = 0$ state (*31*, *37*, *38*). Away from resonance, our results illustrate the universal nature of this semiclassical effect, depending only on the value of $|d|$ and not on the rotational state of the molecule, when resonant dipolar effects are not important. In addition, the absence of any field-dependent Fano-Feshbach resonances is consistent with universal (unit probability) loss at short range for the $|1,0\rangle$ state (*51*), as previously measured with $N = 0$ molecules (*34*, *35*).

In the region near $|\mathbf{E}_1|$ and $|\mathbf{E}_2|$, resonant off-diagonal dipolar couplings to $|0,0\rangle|2,\pm1\rangle$ and $|0,0\rangle|2,0\rangle$ become the dominant contribution, instead of the diagonal dipolar interactions that determine β in the background region. Here, inelastic collisions, which result in transitions to the nearby combined molecular state, are possible owing to the long-range dipolar mixing of the states. Because only $|1,0\rangle$ molecules were detected in the experiment, the measured loss rate consisted of the sum of the inelastic and reactive rates, with the inelastic rate predicted to be negligible except for within a small range of $|\mathbf{E}|$ near resonance (*39*). The optimal shielding condition, where the overall loss is minimized, arises from a competition between the changes in the height of the repulsive barrier and the inelastic loss rate, with the inelastic rate falling off faster away from resonance (*39*).

We observed sharp features at $|\mathbf{E}_1|$ and $|\mathbf{E}_2|$ in both the $m_L = \pm1$ and $m_L = 0$ channels, in excellent quantitative agreement with the scattering theory predictions with no free parameters (Fig. 3, A and B, solid lines) (*52*). We measured a maximum variation of $\beta_{\pm1}$ by a factor of 300 ± 20, and a reduction in $\beta_{\pm1}$ by a factor of 8 ± 3 at the optimal shielding condition (12.77 kV/cm) relative to the value away from the features (10.13 kV/cm). In the $m_L = 0$ channel, we observed a variation of β_0 by a factor of 1000 ± 400 near the features. Comparing the measurements at the optimal shielding point (11.84 kV/cm) and away from the features (11.32 kV/cm), we observed a maximum suppression of β_0 by a factor of 23 ± 10 below its background value. (We excluded the point at 12.72 kV/cm from this analysis, because the extracted β_0 was consistent with zero within our measurement precision.)

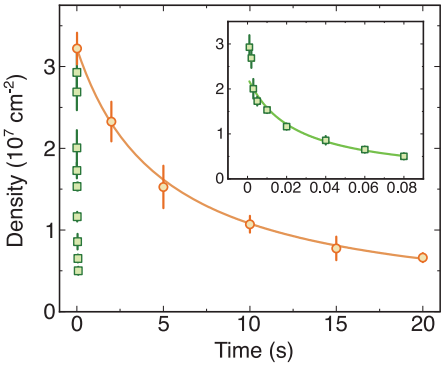


Fig. 4. Reaction shielding and enhancement near $|\mathbf{E}_2|$. Molecular loss measurements at $|\mathbf{E}|$ = 12.50 kV/cm (green squares) and 12.67 kV/cm (orange circles) for $\theta = 0°$ are shown. The inset shows the data for $|\mathbf{E}|$ = 12.50 kV/cm on an enlarged *x* axis. Solid lines are fits to the two-body rate equation; error bars are 1 SE of independent measurements.

As opposed to the semiclassical nature of the loss suppression when dipolar molecules are made to collide side-to-side (*31*, *37*), the presence of the resonances in both the $m_L = 0$ and $m_L = \pm 1$ channels highlights the quantum nature of the shielding mechanism, which is based on level repulsion between the two combined molecular states brought to degeneracy by $\mathbf{E}$. The effect occurs independently of the relative orientation of the dipoles, because approaching with either $m_L = 0$ or $m_L = \pm 1$ causes dipolar mixing of the states. Our quasi-2D measurements indicated that the shielding is present in all three of these channels at the same values of $|\mathbf{E}|$, even in the case of $m_L = 0$ where the dipoles approach in attractive head-to-tail collisions. We thus fully expect that the shielding will also be effective in 3D (*39*).

Figure 3C shows a close-up of $\beta_{\pm 1}$ near $|\mathbf{E}_1|$ and $|\mathbf{E}_2|$, emphasizing the narrow widths of the features. An intuitive explanation for the observed widths, which were on the order of tens of V/cm, comes from comparing the resonant dipolar interaction energy with the Stark shift of the two crossing states. Away from resonance, the reaction rate is controlled by the height of the *p*-wave centrifugal barrier, which occurs at r_0 (*51*). Near resonance, this barrier is modified by the dipolar coupling V_{dd} between the states, with an approximate energy scale of $V_{dd} \sim d_0^2/4\pi\varepsilon_0 r_0^3 = h \times (16\,\mathrm{MHz})$, where $d_0 = 0.574$ D is the permanent dipole moment, ε_0 is the permittivity of free space, and h is the Planck constant. The differential Stark shift of $|1,0\rangle|1,0\rangle$ and $|0,0\rangle|2,0\rangle$ near 12 kV/cm is roughly $\partial U/\partial|\mathbf{E}| = h \times 215$ kHz/(V/cm). This argument suggests a width on the order of $V_{dd}/(\partial U/\partial|\mathbf{E}|) = 75$ V/cm, in qualitative agreement with the exact result from scattering calculations.

To underscore the enormous change in β under a small variation of $|\mathbf{E}|$, Fig. 4 displays two molecular loss curves in the vicinity of $|\mathbf{E}_2|$ at $\theta = 0°$, corresponding to the largest measured difference in $\beta_{\pm 1}$. The values of $|\mathbf{E}|$ for the two curves, 12.50 kV/cm (green squares and inset) and 12.67 kV/cm (orange circles), correspond to the adiabatic energy curves in Fig. 1C. At $|\mathbf{E}| = 12.50$ kV/cm, the loss rate was strongly enhanced and the molecules were lost within 100 ms. In contrast, at $|\mathbf{E}| = 12.67$ kV/cm, the molecules were shielded from loss and ~20% of the initial density was still detected after 20 s of hold time, thereby realizing a long-lived gas of polar molecules in a large electric field.

We have demonstrated a method for controlling reactive losses using an external electric field and find excellent agreement with theoretical predictions. Our investigation of the $m_L = 0$ and $m_L = \pm 1$ collision channels strongly suggests that the shielding remains effective in 3D geometry, without the need for an optical lattice to protect the molecules. Indeed, we have made preliminary observations of long-lived molecules in a crossed optical dipole trap at 12.67 kV/cm. The shielding could be used to create a favorable ratio of elastic to reactive collisions for evaporative cooling, which would simplify future efforts to create quantum-degenerate molecular gases for species other than KRb (*41*, *42*). These results provide long-lived quantum gases of polar molecules in strong electric fields that are ready to be used to explore a wide range of exciting many-body phenomena and quantum information applications.

REFERENCES AND NOTES

1. R. V. Krems, *Phys. Chem. Chem. Phys.* **10**, 4079–4092 (2008).
2. L. D. Carr, D. DeMille, R. V. Krems, J. Ye, *New J. Phys.* **11**, 055049 (2009).
3. G. Quéméner, P. S. Julienne, *Chem. Rev.* **112**, 4949–5011 (2012).
4. M. Lemeshko, R. V. Krems, J. M. Doyle, S. Kais, *Mol. Phys.* **111**, 1648–1682 (2013).
5. N. Balakrishnan, *J. Chem. Phys.* **145**, 150901 (2016).
6. J. L. Bohn, A. M. Rey, J. Ye, *Science* **357**, 1002–1010 (2017).
7. S. Chefdeville *et al.*, *Science* **341**, 1094–1096 (2013).
8. A. Klein *et al.*, *Nat. Phys.* **13**, 35–38 (2017).
9. T. de Jongh *et al.*, *Science* **368**, 626–630 (2020).
10. M. T. Hummon *et al.*, *Phys. Rev. Lett.* **106**, 053201 (2011).
11. B. C. Sawyer *et al.*, *Phys. Chem. Chem. Phys.* **13**, 19059–19066 (2011).
12. X. Wu *et al.*, *Science* **358**, 645–648 (2017).
13. Y. Segev *et al.*, *Nature* **572**, 189–193 (2019).
14. P. Puri *et al.*, *Science* **357**, 1370–1375 (2017).
15. K. K. Ni *et al.*, *Science* **322**, 231–235 (2008).
16. T. Takekoshi *et al.*, *Phys. Rev. Lett.* **113**, 205301 (2014).
17. P. K. Molony *et al.*, *Phys. Rev. Lett.* **113**, 255301 (2014).
18. J. W. Park, S. A. Will, M. W. Zwierlein, *Phys. Rev. Lett.* **114**, 205302 (2015).
19. M. Guo *et al.*, *Phys. Rev. Lett.* **116**, 205303 (2016).
20. T. M. Rvachov *et al.*, *Phys. Rev. Lett.* **119**, 143001 (2017).
21. F. Seeßelberg *et al.*, *Phys. Rev. Lett.* **121**, 253401 (2018).
22. H. Yang *et al.*, *Science* **363**, 261–264 (2019).
23. E. S. Shuman, J. F. Barry, D. Demille, *Nature* **467**, 820–823 (2010).
24. S. Truppe *et al.*, *Nat. Phys.* **13**, 1173–1176 (2017).
25. L. W. Cheuk *et al.*, *Phys. Rev. Lett.* **125**, 043401 (2020).
26. S. Ding, Y. Wu, I. A. Finneran, J. J. Burau, J. Ye, *Phys. Rev. X* **10**, 021049 (2020).
27. S. Ospelkaus *et al.*, *Science* **327**, 853–857 (2010).
28. M. G. Hu *et al.*, *Science* **366**, 1111–1115 (2019).
29. L. De Marco *et al.*, *Science* **363**, 853–856 (2019).
30. W. G. Tobias *et al.*, *Phys. Rev. Lett.* **124**, 033401 (2020).
31. G. Valtolina *et al.*, *Nature* **588**, 239–243 (2020).
32. D. DeMille, J. M. Doyle, A. O. Sushkov, *Science* **357**, 990–994 (2017).
33. M. A. Baranov, M. Dalmonte, G. Pupillo, P. Zoller, *Chem. Rev.* **112**, 5012–5061 (2012).
34. K. K. Ni *et al.*, *Nature* **464**, 1324–1328 (2010).
35. M. Guo *et al.*, *Phys. Rev. X* **8**, 041044 (2018).
36. G. Quéméner, J. L. Bohn, *Phys. Rev. A* **81**, 022702 (2010).
37. M. H. G. de Miranda *et al.*, *Nat. Phys.* **7**, 502–507 (2011).
38. G. Quéméner, J. L. Bohn, *Phys. Rev. A* **83**, 012705 (2011).
39. G. Wang, G. Quéméner, *New J. Phys.* **17**, 035015 (2015).
40. A. V. Avdeenkov, M. Kajita, J. L. Bohn, *Phys. Rev. A* **73**, 022707 (2006).
41. G. Quéméner, J. L. Bohn, *Phys. Rev. A* **93**, 012704 (2016).
42. M. L. González-Martínez, J. L. Bohn, G. Quéméner, *Phys. Rev. A* **96**, 032718 (2017).
43. H. P. Büchler *et al.*, *Phys. Rev. Lett.* **98**, 060404 (2007).
44. A. V. Gorshkov *et al.*, *Phys. Rev. Lett.* **101**, 073201 (2008).
45. T. Karman, J. M. Hutson, *Phys. Rev. Lett.* **121**, 163401 (2018).
46. T. Karman, *Phys. Rev. A* **101**, 042702 (2020).
47. L. Lassablière, G. Quéméner, *Phys. Rev. Lett.* **121**, 163402 (2018).
48. T. Xie *et al.*, *Phys. Rev. Lett.* **125**, 153202 (2020).
49. Z. Z. Yan *et al.*, *Phys. Rev. Lett.* **125**, 063401 (2020).
50. X. Ye *et al.*, arXiv 2010.08685 [physics.atom-ph] (17 October 2020).
51. Z. Idziaszek, G. Quéméner, J. L. Bohn, P. S. Julienne, *Phys. Rev. A* **82**, 020703(R) (2010).
52. See supplementary materials.
53. G. Günter *et al.*, *Science* **342**, 954–956 (2013).
54. S. Ravets *et al.*, *Nat. Phys.* **10**, 914–917 (2014).
55. R. Faoro *et al.*, *Nat. Commun.* **6**, 8173 (2015).
56. B. Zhu, G. Quéméner, A. M. Rey, M. J. Holland, *Phys. Rev. A* **88**, 063405 (2013).
57. G. Quéméner, M. Lepers, O. Dulieu, *Phys. Rev. A* **92**, 042706 (2015).
58. K. Matsuda *et al.*, Data for "Resonant collisional shielding of reactive molecules using electric fields." Zenodo (2020); doi:10.5281/zenodo.4156490.

ACKNOWLEDGMENTS

We thank J. L. Bohn for stimulating discussions and careful reading of the manuscript. **Funding:** Supported by NIST, ARO MURI, DARPA DRINQS, NSF QLCI OMA–2016244, and NSF Phys-1734006; G.Q. received funding from Agence Nationale de la Recherche FEW2MANY-SHIELD Project ANR-17-CE30-0015. **Author contributions:** The experimental work and data analysis were done by K.M., L.D.M., J.-R.L., W.G.T., G.V., and J.Y. Theoretical calculations were done by G.Q. All authors contributed to interpreting the results and writing the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data presented in this work are available through Zenodo at (*58*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1324/suppl/DC1
Materials and Methods
Fig. S1

11 September 2020; accepted 9 November 2020
10.1126/science.abe7370

CANCER IMMUNOTHERAPY

Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer

Sri Krishna[1]*, Frank J. Lowery[1]*, Amy R. Copeland[1], Erol Bahadiroglu[2], Ratnadeep Mukherjee[2], Li Jia[3], James T. Anibal[2], Abraham Sachs[1], Serifat O. Adebola[2], Devikala Gurusamy[1], Zhiya Yu[1], Victoria Hill[1], Jared J. Gartner[1], Yong F. Li[1], Maria Parkhurst[1], Biman Paria[1], Pia Kvistborg[4], Michael C. Kelly[5], Stephanie L. Goff[1], Grégoire Altan-Bonnet[2], Paul F. Robbins[1]†, Steven A. Rosenberg[1]†

Adoptive T cell therapy (ACT) using ex vivo–expanded autologous tumor-infiltrating lymphocytes (TILs) can mediate complete regression of certain human cancers. The impact of TIL phenotypes on clinical success of TIL-ACT is currently unclear. Using high-dimensional analysis of human ACT products, we identified a memory-progenitor CD39-negative stem-like phenotype ($CD39^-CD69^-$) associated with complete cancer regression and TIL persistence and a terminally differentiated CD39-positive state ($CD39^+CD69^+$) associated with poor TIL persistence. Most antitumor neoantigen-reactive TILs were found in the differentiated $CD39^+$ state. However, ACT responders retained a pool of $CD39^-$ stem-like neoantigen-specific TILs that was lacking in ACT nonresponders. Tumor-reactive stem-like TILs were capable of self-renewal, expansion, persistence, and superior antitumor response in vivo. These data suggest that TIL subsets mediating ACT response are distinct from TIL subsets enriched for antitumor reactivity.

Cancer immunotherapies such as immune checkpoint blockade (ICB), adoptive T cell therapy (ACT), and chimeric antigen receptor (CAR) therapy rely on the targeted destruction of cancer cells by potent antitumor T cells (*1*–*6*). Studies of ACT using tumor-infiltrating lymphocytes (TILs) (termed TIL-ACT) and ICB indicate that antitumor responses can be mediated by T cells that target cancer cells through the recognition of mutated neoantigens presented on human leukocyte antigen (HLA) molecules (*7*–*11*). In multiple phase 2 clinical trials, TIL-ACT has been shown to mediate complete durable responses in certain patients with metastatic melanoma as well as epithelial cancers that are conventionally considered to be weakly immunogenic (*5*, *12*–*14*).

Successful immunotherapy is influenced by several tumor-intrinsic factors, including tumor mutational burden, neoantigen burden, HLA type and expression, DNA damage repair capacity, and expression of the programmed death ligand-1 (PD-L1) protein (*1*, *2*, *15*–*17*). Recently, T cell–intrinsic factors have also been associated with ICB response in murine and human studies (*18*–*21*). This subset of self-renewing stem cell–like TILs has been reported to exist in distinctive intratumoral structures and is characterized by the expression of transcription factor 7 (TCF7), along with the lack of cell surface inhibitory markers such as CD39 or TIM3 (*18*, *22*). However, analyses of tumor-reactive populations have concluded that antitumor neoantigen-specific TILs are enriched in subsets defined by the expression of PD-1 or CD39 (*23*–*29*). Thus, there is a lack of consensus regarding the tumor-reactive TIL subset that is directly responsible for successful immunotherapy including TIL-ACT.

Although the influence of the state of T cell differentiation on ACT has been evaluated in murine studies (*30*), the phenotypic fitness landscape of tumor-reactive TILs associated with the TIL-ACT response against human cancer is unclear. ACT using TILs in lymphodepleted human cancer patients provides an opportunity to investigate the T cell states responsible for cancer regression, because the T cell–enriched treatment product can be isolated and studied. We therefore compared the phenotypic differences that could distinguish TIL-ACT infusion products administered to patients who had complete response to therapy [termed complete responders (CRs), $n = 24$] from those whose disease progressed after ACT [termed nonresponders (NRs), $n = 30$] (Fig. 1A). These retrospective TIL-ACT infusion products were from a cohort of patients who had stage IV metastatic melanoma and had been treated with autologous in vitro–expanded TILs that had not been selected on the basis of tumor reactivity ("unselected" TILs). Additionally, patients in this cohort had not experienced prior immunotherapies in the form of genetically engineered T cell therapy or PD-1 blockade that might influence T cell phenotype (table S1 and Study Cohort Description section of the materials and methods) (*14*, *31*).

Initial single-cell analysis of a discovery set of 4.8 million infusion-product TILs from 7 CRs and 9 NRs by mass cytometry (CyTOF) revealed heterogeneous expression of 34 cell surface markers (fig. S1 and table S2). Supervised analysis of TIL CyTOF profiles identified clusters that appeared prevalent in either CRs or NRs (Fig. 1, B and C). Cluster 1, which was four times more abundant in CR infusion products relative to NR infusion products (corrected $P = 0.0264$; Fig. 1D), corresponded to $CD8^+$ T cells with high expression of CD44, CD27, and CD28 and low expression of TIM3, characterized in prior studies as memory-like (*18*) and stem-like (*22*) T cells (Fig. 1C). Notably, cluster 1 also had low expression of inhibitory marker CD39 and T cell activation marker CD69 (Fig. 1C). Machine learning–based unsupervised clustering of T cell activation and/or exhaustion states to classify patients according to their clinical outcome further confirmed CD69 and CD39 expression as two important features of most clinical relevance (fig. S2). Of note, cluster 2, with high levels of CD39 and CD69 ($CD8^+CD39^+CD69^+$) but lower expression of CD44, CD27, and CD28 (Fig. 1C), trended higher in NR infusion products than in CR infusion products, although this difference was not statistically significant (Fig. 1D).

Flow cytometric analysis of the 16 discovery samples defining $CD8^+CD39^-CD69^-$ cells as members of cluster 1 recapitulated CyTOF data with high confidence (fig. S3, A and B). An evaluation of an independent set of 38 infusion products (CRs $n = 17$, NRs $n = 21$) by multiparameter flow cytometry revealed that the frequency of the $CD8^+CD39^-CD69^-$ TIL population was 2.5-fold higher in CR infusion products relative to NR infusion products ($P = 0.0096$, percentage of CD3), supporting the association between this subset and ACT response (Fig. 1, E and F). Whereas the total number of infused T cells did not differ significantly between CRs and NRs in this set of 38 patient samples, the total number of infused $CD8^+CD39^-CD69^-$ cells was fourfold higher in CR infusion products than in NR infusion products ($P = 0.0031$; Fig. 1, G and H). When querying individual markers, there was a modest trend toward an association between $CD8^+$ $TIM3^+$ TILs and nonresponse to ACT (fig. S3C). Analyzing patient survival from all 54 patients, we found that whereas the absolute number of infused $CD39^+CD69^+$ T cells, which comprised most of the $CD8^+$ TILs in patient infusion products (fig. S3C), did not significantly affect melanoma-specific survival (MSS) or progression-free survival (PFS) in this cohort (fig. S4), higher numbers of infused $CD8^+CD39^-CD69^-$ cells were significantly associated with improved PFS [$P < 0.0001$,

[1]Surgery Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [2]Immunodynamics Group, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [3]National Institutes of Health Library, National Institutes of Health, Bethesda, MD 20892, USA. [4]Division of Immunology, The Netherlands Cancer Institute, Amsterdam, Netherlands. [5]Single Cell Analysis Facility, Cancer Research Technology Program, Frederick National Laboratory, Bethesda, MD 20892, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: sar@nih.gov (S.A.R.); paulrobbins@mail.nih.gov (P.F.R.)

hazard ratio (HR) = 0.255, 95% confidence interval (CI) 0.1257 to 0.5186; Fig. 1I] and MSS ($P < 0.0001$, HR = 0.217, 95% CI 0.101 to 0.463; Fig. 1J) in a dose-dependent manner (fig. S4 tertile analysis, bottom row). Furthermore, the ratio of $CD8^+CD39^-CD69^-$ to $CD8^+CD39^+CD69^+$ cells in the infusion products significantly affected MSS and PFS (fig. S5), suggesting that the ACT responses in this cohort were associated with the infusion of higher numbers of $CD8^+CD39^-CD69^-$ cells and not fewer $CD8^+CD39^+CD69^+$ TILs.

To further explore the potential importance of $CD8^+CD39^-CD69^-$ TIL [cluster 1, double-negative (DN)] and the infusion product–predominant $CD8^+CD39^+CD69^+$ TIL population [cluster 2, double-positive (DP)], we evaluated the transcriptome profile of these two subsets (fig. S6A). DN TILs had increased expression of quiescent T stem cell markers *KLF2*, *TCF7*, *S1PR1*, *LEF1*, *IL7R*, *CD27*, and *SELL* (CD62L),

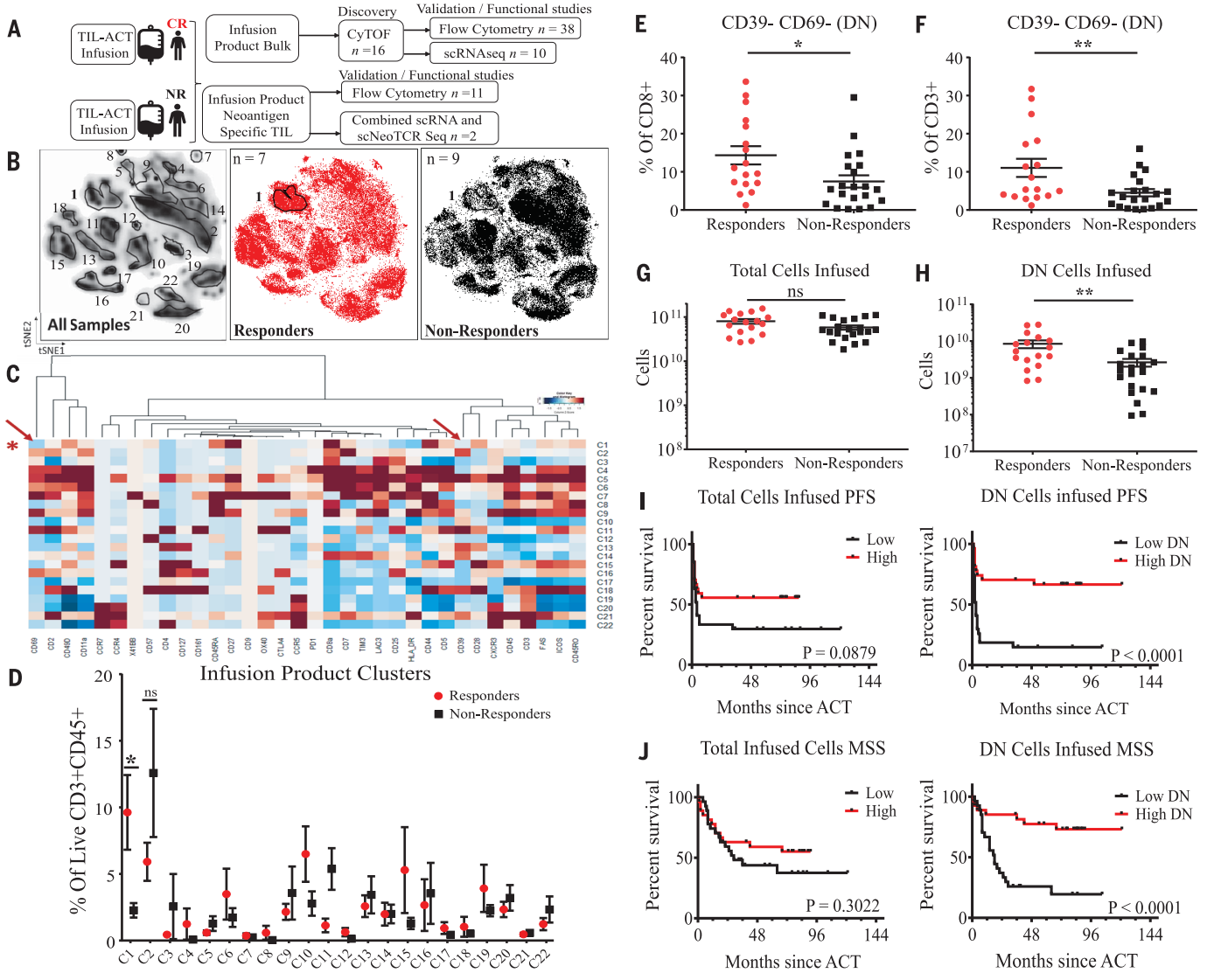


Fig. 1. Phenotypic landscape of TIL infusion products from patients with metastatic melanoma treated with ACT. (**A**) Melanoma cohort patient infusion products and schema used for this study. (**B**) t-SNE (t-distributed stochastic neighbor embedding) plots of live $CD45^+CD3^+$ cell clusters of infusion products from all patient infusion products (left), clusters from CR infusion products only (in red) (middle), and clusters from NR infusion products only (in black) (right). Cluster 1 represents the $CD39^-CD69^-$ DN population. (**C**) Heatmap of scaled protein expression (columns) per each cluster (rows); red arrows indicate CD69 and CD39 expression. The red asterisk indicates cluster 1. (**D**) Plot showing the percentage of CR and NR infusion-product cells in each cluster. Error bars indicate SEM. *P = 0.0264 from two-sided Wilcoxon rank sum test adjusted by Bonferroni correction for all clusters. (**E** and **F**) Flow cytometry–based independent validation of CR infusion products and NR infusion-product samples (n = 38) showing the percentages of $CD39^-CD69^-$ (DN) cells of total $CD8^+$ (E) and total $CD3^+$ (F). Error bars indicate mean ± SEM. (**G** and **H**) Scatterplot showing (G) the total number of infused cells and (H) total number of infused $CD8^+CD39^-CD69^-$ (DN) cells within CR infusion products and NR infusion products (n = 38). *$P < 0.05$ and **$P < 0.01$ by two-sided Wilcoxon rank sum test. Error bars indicate mean ± SEM. ns, not significant. (**I**) PFS of patients after infusion-product administration (n = 54) separated by median total cell numbers infused (left) and median $CD8^+CD39^-CD69^-$ (DN) cell numbers infused (right). (**J**) MSS of patients after infusion-product administration (n = 54) separated by median cell number infused (left) and median $CD39^-CD69^-$ cell number infused (right). "Low" indicates patients with infused cells less than median of the subgroup analyzed, and "high" indicates patients with cells infused greater than median of the subgroup analyzed. P values by the log-rank Mantel-Cox test are shown.

whereas DP TILs expressed *CD38*, *MKI67*, *TNF*, and *IFNG* found in differentiated, activated T cells (fig. S6B and table S3). Unsupervised clustering of 20,672 $CD8^+$ TILs by single-cell transcriptome analysis (scRNA) from 10 patient infusion products (5 CRs, 5 NRs) defined eight clusters (Fig. 2, A and B). Analysis of the distribution of TILs from CR and NR infusion products indicated that they segregated into a two-cluster solution (*18*) defined by S.Cluster.A (Super Cluster A, comprising clusters C0, C2, C5, C6, and C7) and S.Cluster.B (Super Cluster B, comprising clusters C1, C3, C4, and C8). Responder TILs were more frequent in S.Cluster.A (P = 0.0317), whereas S.Cluster.B largely comprised NR TILs (P = 0.03; Fig. 2, B and C). TILs in S.Cluster.A were enriched for DN gene signature (81% overlap; fig. S6C and table S4), whereas TILs in S.Cluster.B were enriched for DP gene signature (Fig. 2, D and E). Consistent with these analyses, CR

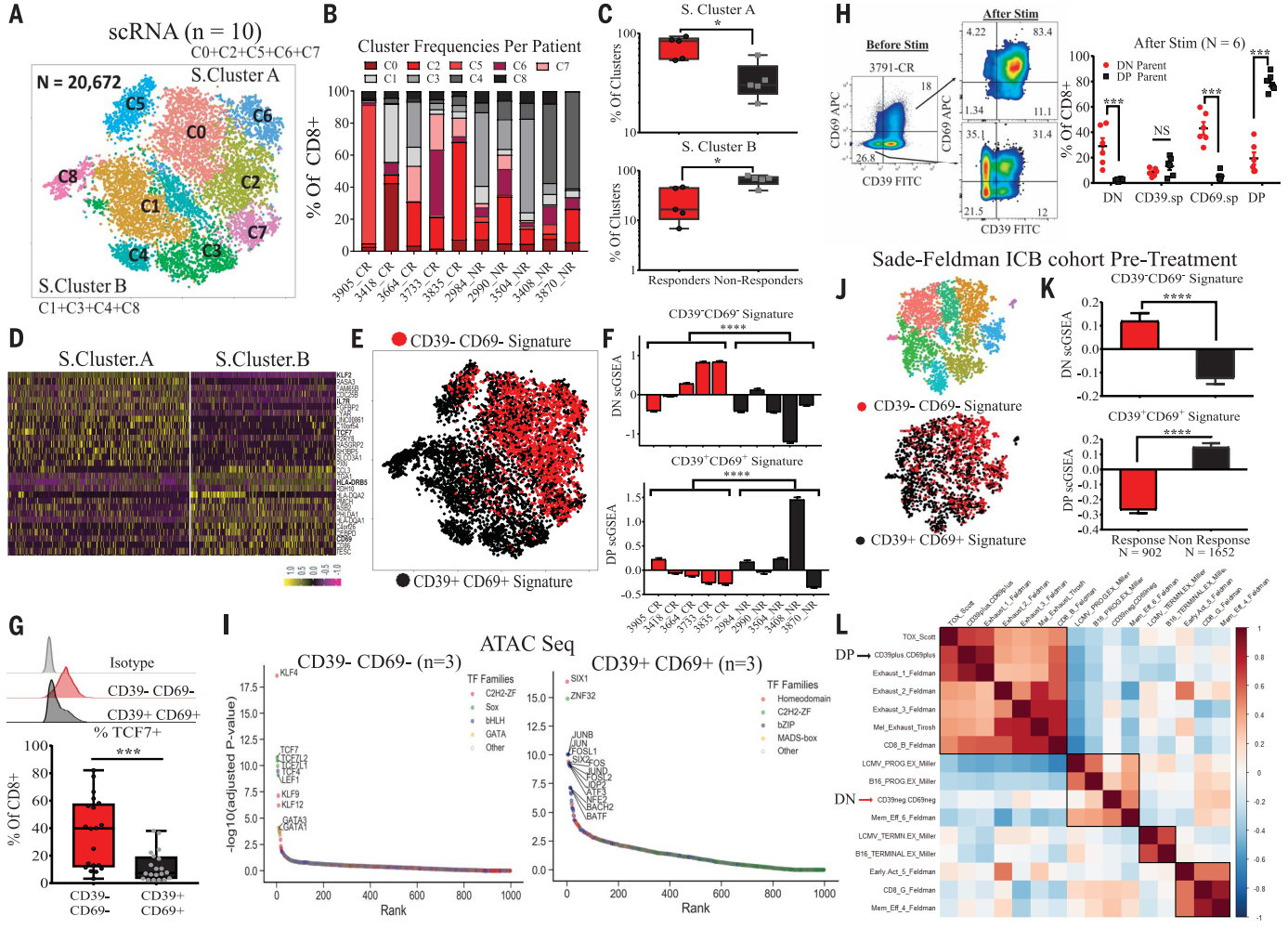


Fig. 2. $CD39^-CD69^-$ $CD8^+$ TILs in infusion products are in a memory progenitor stem-like state. (**A**) t-SNE plot of all $CD8^+$ TILs from CR and NR infusion products (5 CRs, 5 NRs). (**B**) Frequency distribution of the percentage of total $CD8^+$ T cells in each cluster (C1 through C8) for each patient infusion product. (**C**) Box plot showing median percentage of S.Cluster.A (top) and S.Cluster.B (bottom) within CR and NR infusion products with error bars showing minimum to maximum. *P < 0.05 by two-sided Wilcoxon rank sum test is shown. (**D**) Heatmap of the top 15 differentially expressed genes between S.Cluster.A and S.Cluster.B displayed for each cell. All discriminating genes are listed in table S4. (**E**) t-SNE plot displaying the top two quartiles of $CD39^-CD69^-$ (DN) and top two quartiles of $CD39^+CD69^+$ (DP) gene expression signature. (**F**) Each patient infusion product was scored by DN and DP gene signature scores, and their mean scGSEA scores are plotted on the y axis. ****P < 0.0001 by two-sided Wilcoxon rank sum test comparing the mean DN and DP signature scores between CRs and NRs. Error bars indicate mean ± SEM. (**G**) Intracellular TCF7 expression by flow cytometry in DN and DP subsets. The histogram shows a representative patient infusion-product sample (top), and the box plot shows the quantitation for 18 infusion products (bottom), with error bars showing minimum to maximum. ***P < 0.001 by two-sided Wilcoxon rank sum test. (**H**) Flow cytometry plots showing the phenotypes of DN, SP, and DP states before and after anti-CD3 and anti-CD28 stimulation (Stim) for 120 hours of a representative patient sample (left), and dot plots quantifying phenotypes of daughter cells after stimulation of FACS DN parent (red) or FACS DP parent (black) (right). ***P < 0.001 by two-sided Wilcoxon rank sum test, n = 6 infusion products. Error bars indicate mean ± SEM. APC, allophycocyanin; FITC, fluorescein isothiocyanate. (**I**) Enriched TF motifs in $CD39^-CD69^-$ (DN) and $CD39^+CD69^+$ (DP) TILs from patient infusion products in the validation set (n = 3 each) ranked by adjusted P values. The full list of significant TFs is shown in table S5. (**J**) t-SNE clusters of $CD8^+$ TILs from the melanoma ICB cohort (*18*) (top) and t-SNE plot colored by top two quartiles of DN and DP gene signatures (DN, red; DP, black) (bottom). (**K**) TILs from pre-ICB therapy were scored by the top DN and DP gene signature scores, and mean scGSEA scores are plotted on the y axis. Error bars indicate mean ± SEM. TILs from responding lesions are in red, and cells from progressing lesions are in black. Cell numbers and ****P < 0.0001 by two-sided Wilcoxon rank sum tests are shown. (**L**) Clustered correlation matrix of gene signatures from other studies (table S6) with DN and DP scGSEA scores on pre-ICB cells from the cohort.

TILs and NR TILs scored by single-cell gene set enrichment analysis (scGSEA) of DN and DP gene signatures indicated that CR TILs had higher DN scores on average, whereas NR TILs had higher DP scores ($P < 4.5 \times 10^{-12}$ for both; Fig. 2F).

We observed lower cell surface expression and coexpression of the exhaustion markers PD-1 and TIM3 in the DN subset relative to single-positive (SP) ($CD39^+CD69^-$ and $CD39^-CD69^+$) and DP subsets ($n = 39$ infusion products; fig. S7). Conversely, expression of memory markers CD62L and CD27 and progenitor marker SLAMF6 were increased in DN relative to DP subsets with variable expression within SP populations (fig. S7). Protein expression of the memory stem-like T cell marker TCF7 was fivefold higher in DN TILs relative to DP TILs ($P = 0.0002$, $n = 18$ infusion products; Fig. 2G) consistent with our scRNA analysis and other studies (*18*, *19*). Although DN TILs in infusion products expressed higher levels of both SLAMF6 and TCF7 transcripts (fig. S8A), TILs coexpressing TCF7 and SLAMF6 were not significantly enriched within DN subsets (fig. S8B), and addition of SLAMF6 to DN as a surface marker was comparable to DN alone in delineating TILs associated with survival benefit in this patient subset (fig. S8C). By contrast, DP TILs had significantly higher transcript and protein expression of the effector- and exhaustion-associated transcription factor TOX relative to DN TILs (fig. S8, A and B) (*32*). Notably, although the transcriptional profile of DN TILs resembled that of peripheral blood stem cell memory (SCM) T cells (*33*, *34*), cell surface expression of the canonical SCM markers CD95 and CCR7 was not associated with the DN phenotype in patient TIL infusion products (fig. S9).

To experimentally confirm that $CD39^-CD69^-$ infusion-product TILs were stem-like T cells, we isolated and stimulated the DN, SP, and DP subsets using anti-CD3 and anti-CD28 in vitro. Upon T cell receptor (TCR) stimulation, DN TILs underwent self-renewal and gave rise to both SP and $CD39^+CD69^+$ DP populations, whereas stimulated DP TILs largely remained in the same DP state, indicating a DN-progenitor state ($n = 6$ infusion products; Fig. 2H). Time-course kinetics of the daughter populations indicated that essentially all DN TILs underwent CD69 activation, but a subset of DN TILs (median = 17%) rested back to the $CD39^-CD69^-$ stem-like state 5 days after initial stimulation (fig. S10). By contrast, both $CD39^+$ parent populations (CD39 SP and DP) appeared to terminally differentiate into a stable $CD39^+CD69^+$ state after TCR stimulation (fig. S10). Although DP and DN TILs secreted similar levels of interleukin-2 (IL-2), interferon-γ (IFN-γ), granzyme A (GZMA), granzyme B (GZMB), and perforin in response to TCR stimulation, DN TIL stimulation resulted in higher levels of secreted IL-17A, tumor necrosis factor–α (TNF-α), IL-4, sFasL, and granulysin than DP TILs (fig. S11). Taken together, our results indicate that ACT response–associated $CD39^-CD69^-$ TILs are in a progenitor memory stem-like state capable of differentiation into other subsets, whereas the infusion product–dominant $CD39^+CD69^+$ TILs are in a terminally differentiated state.

To understand the key regulators of these two TIL states in the infusion products, we isolated matched DN and DP TILs from three patient infusion products and analyzed their epigenetic profiles using ATAC (assay for transposase accessible chromatin) sequencing. We queried the presence of 4314 human transcription factor (TF) motifs in open chromatin regions of DN and DP TILs (fig. S12 and table S5). Open chromatin regions within stem-like DN TILs were enriched for numerous binding sites for SOX and C2H2-ZF family TFs including *KLF4*, *TCF7*, *LEF1*, and *TCF7L1* (Fig. 2I and fig. S12) consistent with those of less differentiated human T cells (*18*, *35*). By contrast, chromatin regions within DP TILs were extensively enriched for multiple binding motifs of bZIP TFs *FOSL1*, *FOS*, *JUNB*, and *JUND*, indicating that epigenetic imprinting may be involved in maintaining the state of terminal differentiation resulting from chronic tumor cell activation (Fig. 2I and fig. S12) (*35*, *36*).

We inquired if TIL states in the infusion products corresponded to those within the fresh tumor by reanalyzing the transcriptomic profiles of $CD8^+$ T cells from a previous melanoma ICB response study (*18*) using DN and DP gene signatures described above (Fig. 2, J to L). We found that before immune checkpoint therapy, TILs from ICB-responding lesions had higher $CD39^-CD69^-$ signature scores ($P = 1.9 \times 10^{-10}$) than nonresponding lesions, whereas TILs from lesions that progressed during ICB had higher $CD39^+CD69^+$ signature scores ($P < 2.2 \times 10^{-16}$). Hierarchical clustering correlation analysis with T cell dysfunctional and progenitor gene signatures from other recent studies (listed in table S6) revealed that $CD39^-CD69^-$ TILs were most similar to ICB response–associated memory and progenitor-exhausted TILs, whereas $CD39^+CD69^+$ TILs were highly correlated with TOX^+ terminally exhausted TILs that are associated with poor ICB response (Fig. 2L) (*18*, *19*, *32*, *37*). We developed a phenotypic fitness score, defined as the difference between DN stem-like signature and DP differentiated signature (DN minus DP), and scored scRNA from ACT infusion products and melanoma ICB TILs. Single-cell fitness scores reconfirmed that ICB-responder TILs have on average more stem-like fitness compared with ICB-NR TILs (fig. S13). These data suggest that stem-like TILs found in ACT responder infusion products resemble those found in ICB responses.

Previous studies on stem-like and terminally differentiated TIL subsets were performed on bulk TIL populations lacking specific analyses of antitumor T cells (*18*, *22*). Given recent findings suggesting that neoantigen-specific TILs are nearly exclusively $CD39^+$, whereas $CD39^-$ TILs represent bystander cells (*23*, *24*, *27*), we sought to determine whether the $CD39^-CD69^-$ stem-like state associated with ACT response exists within the neoantigen-reactive T cells in infusion products. Using our previously described tandem minigene neoantigen screening platform (*12*, *13*), we defined 26 HLA class I–restricted neoantigens from infusion products for phenotypic evaluation ($n = 11$ patients; figs. S14 and S15 and table S7). Stem-like and differentiated TIL subsets defined by CD39 and CD69 expression were detectable within HLA-neoantigen $tetramer^+$ TILs as shown in representative CRs and NRs (Fig. 3, A and B). Combined analysis of all 26 neoantigen-specific T cell populations indicated that they largely existed in the $CD39^+CD69^+$ terminally differentiated state, consistent with other studies reporting CD39 enrichment of neoantigen-specific T cells (median DP = 62.7%; Fig. 3C) (*23*).

Stratification of the data by ACT response status, however, revealed that the frequency of DN cells was significantly higher in neoantigen-specific TILs in CR infusion products (median = 8.8%) than in NR infusion products (median = 0.5%, $P = 1.2 \times 10^{-4}$; Fig. 3D). CD39 as a single marker indicated that CD39-negative ($CD39^-$) neoantigen-reactive TILs were significantly higher in CR TILs relative to NR TILs ($P = 1.86 \times 10^{-5}$; Fig. 3E). Adjusting for tumor-specific cells significantly enhanced these differences between CR and NR infusion products, wherein neoantigen-specific DN TILs were 23-fold higher in CRs (median = 2.3×10^8 cells), and $CD39^-$ TILs were 40.4-fold higher in CR (median = 1.06×10^9 cells) (fig. S16). We found no differences in the total number of neoantigen-reactive TILs, or neoantigen-reactive $CD39^+$ or DP TILs between CR and NR infusion products (fig. S16). These results indicate that at least in the context of melanoma ACT infusion products, not all $CD39^-$ TILs are bystander T cells as previously reported (*23*) and that a subset of tumor neoantigen-reactive TILs exist in a $CD39^-$ stem-like state in CRs. Furthermore, these data indicate that enrichment of tumor-reactive neoantigen-specific TILs in differentiated subsets does not necessarily correspond to their frequency in stem-like states.

To explore the heterogeneity in stem-like and differentiated states of neoantigen-specific TILs at the single-cell transcriptomic level, we performed combined scRNA and scTCR sequencing on infusion products from a CR (patient 3713) and an NR (patient 4000), with defined neoantigen-reactive TCRs [Fig. 3, F to L, and table S8 (*38*–*41*)]. The infusion product from responder patient 3713 was dominated by two

major clusters, C0 and C1, in which C0 represented the stem-like DN state and C1 represented the differentiated DP state (Fig. 3F and fig. S17, A and B). Projection of the 20 neoantigen-specific TCR$^+$ (NeoTCR$^+$) clonotypes showed a broad distribution among the two clusters (Fig. 3, F and G). The immunodominant SRPX mutation–specific NeoTCRs (SRPX$_{mut}$) were enriched in cluster C0, whereas other NeoTCRs varied in their prevalence within the two clusters (Fig. 3G). CDR3β sequencing of fluorescence-activated cell–sorted (FACS) DN and DP infusion-product TILs confirmed this observation (fig. S17, C and D).

Finer analysis of SRPX$_{mut}$ NeoTCR$^+$ single cells indicated that 9 out of 10 of these clonotypes (and 15 out of 20 of all NeoTCR clonotypes) were enriched for positive phenotypic fitness

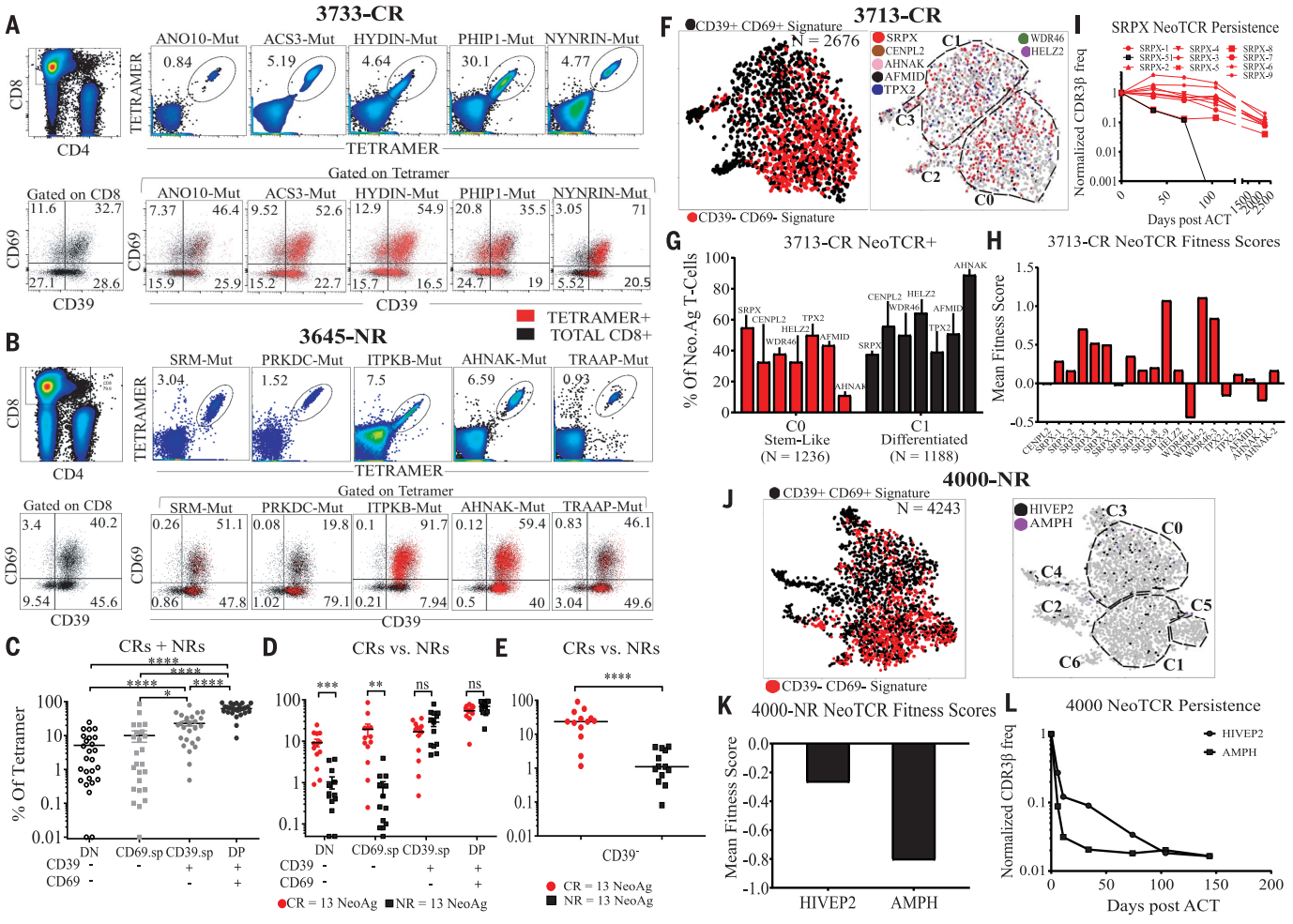


Fig. 3. Stem-like state within neoantigen-specific TIL populations from CR and NR infusion products. (**A**) Representative CR infusion product showing detection of neoantigen-specific tetramers (top) and CD39 and CD69 phenotypes of bulk CD8$^+$ TILs and CD8$^+$ tetramer$^+$ TILs (bottom). (**B**) NR infusion product with comparable numbers of neoantigens (top) and CD39/CD69 phenotypes of bulk CD8$^+$ TILs and CD8$^+$ tetramer$^+$ TILs (bottom). Numbers within quadrants for (A) and (B), listed under the tetramer category, represent the percentage of the subpopulation within each CD8$^+$ tetramer$^+$ gate. (**C**) DN, SP, and DP phenotypes within all 26 neoantigen-specific T cell populations expressed as a percentage of tetramer$^+$ cells of each individual tetramer from 11 patients. CRs and NRs are combined for this data analysis (CRs + NRs). *P < 0.05 and ****P < 0.0001 by two-sided Wilcoxon rank sum test followed by Bonferroni correction for multiple comparisons. Bar indicates median. (**D**) DN, SP, and DP phenotypes of 26 neoantigen-specific TILs subdivided by response status (CRs versus NRs). **P < 0.01 and ***P < 0.001 by two-sided Wilcoxon rank sum test between CRs and NR tetramer$^+$ cells for each subset. The bars indicate median. (**E**) CD39-negative neoantigen-specific tetramer$^+$ TILs (CD39$^-$ as a single marker) subdivided by response status (CRs versus NRs). ****P < 0.0001 by two-sided Wilcoxon rank sum test. Bar indicates median. (**F**) t-SNE plot of CD8$^+$ TILs from patient 3713-CR infusion product showing projection of stem-like DN and differentiated DP gene signatures (left) and projection of neoantigen-specific TCRs colored by antigen specificity (right). Dotted lines indicate the stem-like and differentiated clusters analyzed. (**G**) Patient 3713-CR NeoTCR$^+$ cells from stem-like C0 cluster and differentiated C1 cluster displayed as a percentage of total NeoTCR$^+$ cells for each neoantigen specificity. (**H**) Each NeoTCR$^+$ clonotype scored by their mean fitness score (value of DN minus DP scGSEA scores). Positive scores indicate clonotype enrichment in stem-like phenotypes, and negative scores indicate clonotype enrichment in differentiated state. (**I**) Post-ACT persistence of the immunodominant SRPX$_{mut}$ NeoTCR clonotypes from patient 3713-CR infusion product in peripheral blood normalized to their initial frequency in the infusion product (day 0). (**J**) t-SNE plot of CD8$^+$ TILs from patient 4000-NR infusion product showing projection of stem-like DN and differentiated DP gene signatures (left) and a projection of neoantigen-specific TCRs colored by antigen specificity (right). Dotted lines indicate the stem-like and differentiated clusters analyzed. (**K**) Each NeoTCR$^+$ clonotype (HIVEP2$_{mut}$, AMPH$_{mut}$) scored by their mean fitness scores as described in (H). (**L**) Post-ACT persistence of the HIVEP2$_{mut}$ and AMPH$_{mut}$ NeoTCR clonotypes from patient 4000-NR in peripheral blood normalized to their frequency in the infusion product. Patient follow-up was stopped at 150 days because of disease progression.

scores (Fig. 3H). Because prior studies have suggested that T cell–intrinsic differences potentially affect post-ACT clonotypic persistence (*33*, *42*), we analyzed the post-ACT peripheral blood of patient 3713. TCR repertoire sequencing revealed that most of the NeoTCR clonotypes (18 of 20) persisted up to 75 months, consistent with the idea of TIL expansion from the progenitor stem-like state (for $SRPX_{mut}$, see Fig. 3I; for other NeoTCRs, see fig. S17E). Among the 10 $SRPX_{mut}$-NeoTCRs targeting the same neo-epitope, one clonotype (TCR-51) with the lowest phenotypic fitness score declined in frequency and became undetectable only 3 months after treatment (Fig. 3, H and I).

Combined scRNA and scTCR sequencing of the NR patient 4000 infusion product showed two major stem-like clusters (C1 and C5) and one major differentiated cluster (C0) (Fig. 3J and fig. S18, A and B). However, $NeoTCR^+$ cells ($HIVEP2_{mut}$ and $AMPH_{mut}$) were largely concentrated in the differentiated C0 cluster (Fig. 3J and fig. S18, C to E). scGSEA scores indicated that both NeoTCR clonotypes had negative phenotypic fitness scores, suggesting terminal differentiation (Fig. 3K). In stark contrast to the 3713-CR persistent clonotypes, patient 4000 NeoTCRs did not persist after ACT infusion; rather, both infused NeoTCR clonotypes rapidly declined in peripheral blood after treatment (Fig. 3L). These results support our prior observations and other published reports that have linked cell therapy response to posttreatment TCR persistence (*33*, *42*, *43*).

Although our results indicate that T cell–intrinsic phenotypic differences are associated with persistence after ACT, additional factors such as TCR avidity against various tumor antigens might also play a profound role in T cell persistence. To address this issue, we performed an exploratory analysis of T cell persistence in a patient who had metastatic synovial cell sarcoma and experienced a complete response after the adoptive transfer of autologous peripheral blood mononuclear

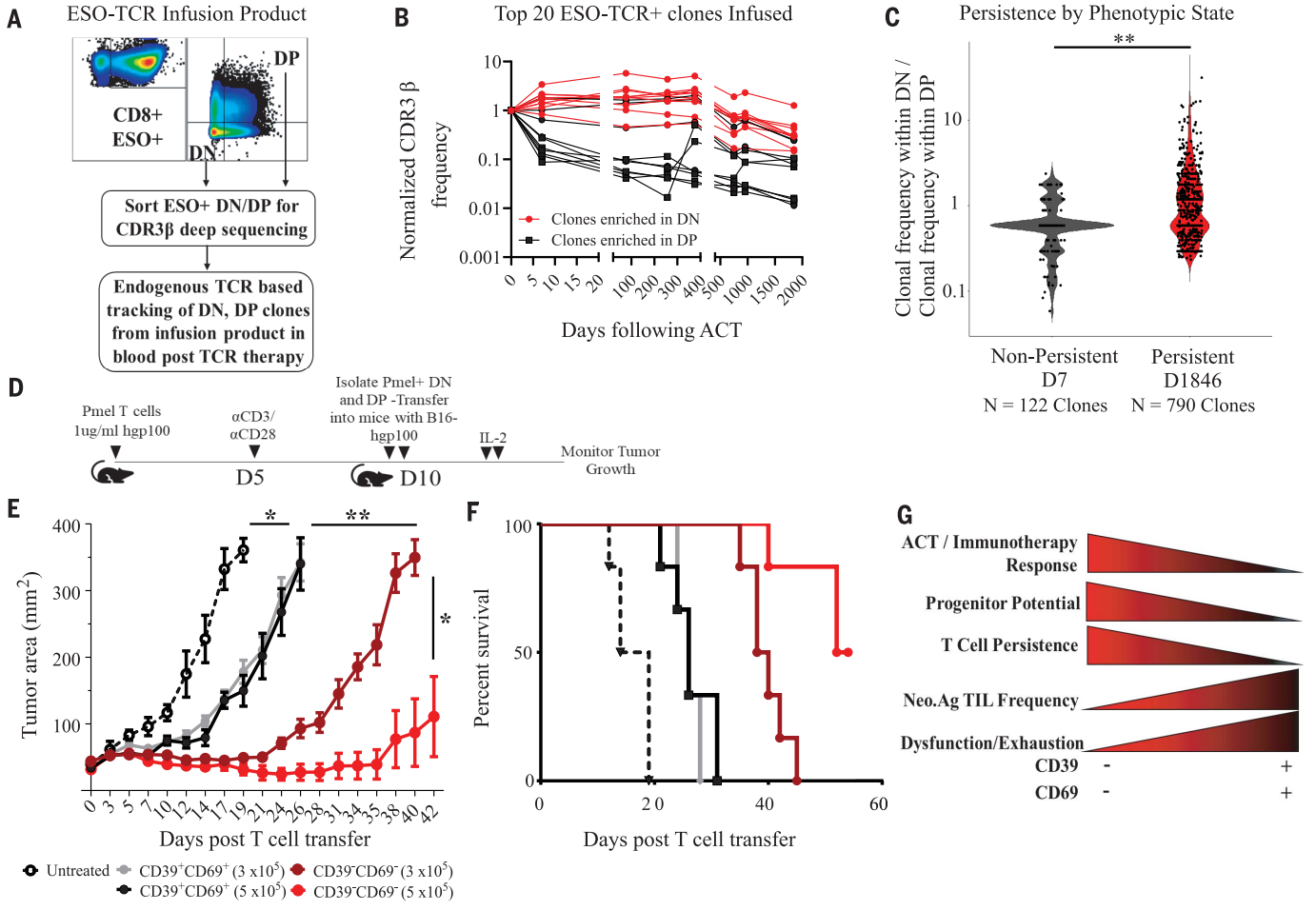


Fig. 4. Stem-like TILs mediate antitumor activity and TCR persistence. (**A**) Schema for using endogenous human TCR to track NY-ESO-1 TCR-transduced ($ESO.TCR^+$) infusion product in posttreatment peripheral blood of a CR patient to NY-ESO-1 TCR therapy (ESO.CR). (**B**) Posttreatment peripheral blood persistence of top 20 $ESO.TCR^+$ infusion product clones using endogenous human TCR according to enrichment in DN (in red) and DP (in black) phenotypes in the TCR infusion product. Clones with frequency in DN/frequency in DP > 1 are defined as enriched in DN state in the infusion product; clones with frequency in DN/frequency in DP < 1 are defined as enriched in DP state in infusion products. (**C**) Violin plot with median showing distribution of clones undetectable at day 7 after ACT (nonpersistent) (D7) compared with long-term persistent clones at day 1846 after ACT (D1846) by the ratio of their clonotypic frequency from the infusion product in the $CD39^-CD69^-$ stem-like state to their clonotypic frequency within the $CD39^+CD69^+$ differentiated state. $^{**}P < 0.01$ by two-sided Wilcoxon rank sum test. (**D**) Schema for adoptive cell transfer of sorted DN and DP Pmel-transgenic T cells into mice bearing established B16 melanoma tumors. (**E**) Tumor growth curves of mice bearing B16 tumors treated with Pmel DN or DP T cells in two doses. $n = 6$ mice per group. $^{*}P < 0.05$ and $^{**}P < 0.01$ for tumor growth kinetics calculated by Wilcoxon rank sum test. Data show mean ± SEM at indicated time points. (**F**) Tumor survival curves of mice bearing B16 tumors treated with Pmel DN or DP T cells in two doses. (**G**) Illustration of the role of stem-like T cells in immunotherapy and ACT success and paradoxical nature of tumor mutation–reactive T cells in stem-like and terminally differentiated states. Neo.Ag, neoantigen.

cells transduced with a TCR targeting a single HLA-A*02:01–restricted NY-ESO-1 epitope. We used endogenous genomic human TCRβ as barcodes to track stem-like DN clones and terminally differentiated DP clones within the TCR-transduced population in posttreatment blood (Fig. 4A) (*44*). Fourteen out of the top 20 ESO TCR$^+$ infusion-product clonotypes were enriched in the stem-like DN state, whereas six were enriched in the DP state. Stem-like clones demonstrated a gradual decrease in their frequency over a period of 5 years after TCR transfer therapy, whereas the differentiated clones declined rapidly to minor frequencies in posttreatment blood of the patient (Fig. 4B). Among 1311 top common clones, long-term persistent clones (detectable at 51 months after ACT, n = 790) were present at significantly higher frequencies in the stem-like state in the infusion product when compared with poorly persistent clones (i.e., those undetectable at 7 days after treatment, n = 122) (P = 0.0035; Fig. 4C). These analyses indicate that intrinsic stem-like phenotypes can modulate the behavior of T cell clonotypes after ACT, although larger patient cohort studies in TCR and CAR settings are needed to support these hypotheses.

Our findings suggested that treatment with CD39$^-$CD69$^-$ tumor-reactive TIL might result in superior tumor control. We first tested this hypothesis in vitro by isolating tumor-reactive DN and DP populations from a CR infusion product (patient 3733) by coculture with the autologous 3733-mel tumor line to select for tumor-reactive TILs, followed by multiple rounds of rapid expansion to evaluate for their proliferative potential and tumor recognition (fig. S19A). We found that the expansion of the tumor-reactive stem-like DN subsets was ~1000-fold higher than the differentiated DP subset (fig. S19B). The stem-like DN subset maintained tumor recognition, whereas the differentiated DP subset lost tumor reactivity after subsequent rounds of expansion (fig. S19, C and D).

To assess the in vivo antitumor effects of stem-like DN and differentiated DP subsets, we isolated CD39$^-$CD69$^-$ and CD39$^+$CD69$^+$ CD8$^+$ T cell populations from in vitro–expanded Pmel-transgenic TCR splenocytes and performed adoptive transfer of the isolated subsets into mice implanted with B16-F10 melanoma engineered to express the human gp100 antigen (Fig. 4D) (*45*). Whereas the adoptive transfer of either 3×10^5 or 5×10^5 differentiated DP Pmel T cells had a modest effect on tumor control and mouse survival, the transfer of the same numbers of stem-like DN T cells led to substantial tumor regression and improved survival in a dose-dependent manner (Fig. 4, E and F).

Harnessing antitumor T cell reactivity has accelerated immunotherapy treatment strategies against multiple human cancer types (*9*, *28*, *46*), but many disease settings currently exist outside the realm of approved immunotherapy modalities. In this study, we have explored the phenotypic diversity of antitumor neoantigen-specific T cells in a setting of successful immunotherapy. Our results support the hypothesis that responders to TIL-ACT received infusion products containing a pool of stem-like neoantigen-specific TILs that are able to undergo prolific expansion, give rise to differentiated subsets, and mediate long-term tumor control and T cell persistence, in line with recent murine ICB studies mediated by TCF$^+$ progenitor T cells (*18*, *20*, *21*). Our data also suggest that antitumor neoantigen-specific TIL subsets enriched for tumor reactivity (e.g., CD39$^+$) are likely terminally differentiated TILs with a relatively poor proliferative potential (Fig. 4G), likely owing to chronic antigenic stimulation (*35*, *47*). Although we cannot conclude that CD39$^+$ differentiated subsets are completely ineffective because 3 out of 24 patients experienced complete response with largely terminally differentiated DP TILs (<2% DN TILs of CD3$^+$ in the infusion products) (Fig. 1F), it is also possible that very low numbers of DN clones were sufficient to cause complete tumor regression in these patients (*48*). Our results argue that strategies aimed at the isolation and expansion of stem-like neoantigen-specific T cells, or the engineering of T cells to have stem-like attributes, might provide opportunities for future development of more effective T cell–based immunotherapies. Within tumor-reactive TIL populations, the lower frequency of stem-like TILs compared with their largely differentiated counterparts, even in CRs, highlights the paradox and challenges of mounting an effective antitumor immune response.

REFERENCES AND NOTES

1. A. Snyder *et al.*, *N. Engl. J. Med.* **371**, 2189–2199 (2014).
2. N. A. Rizvi *et al.*, *Science* **348**, 124–128 (2015).
3. S. L. Maude *et al.*, *N. Engl. J. Med.* **371**, 1507–1517 (2014).
4. N. McGranahan *et al.*, *Science* **351**, 1463–1469 (2016).
5. E. Tran *et al.*, *N. Engl. J. Med.* **375**, 2255–2262 (2016).
6. T. J. Fry *et al.*, *Nat. Med.* **24**, 20–28 (2018).
7. P. F. Robbins *et al.*, *Nat. Med.* **19**, 747–752 (2013).
8. M. R. Parkhurst *et al.*, *Cancer Discov.* **9**, 1022–1035 (2019).
9. E. Tran, P. F. Robbins, S. A. Rosenberg, *Nat. Immunol.* **18**, 255–262 (2017).
10. M. M. Gubin *et al.*, *Nature* **515**, 577–581 (2014).
11. N. van Rooij *et al.*, *J. Clin. Oncol.* **31**, e439–e442 (2013).
12. E. Tran *et al.*, *Science* **344**, 641–645 (2014).
13. N. Zacharakis *et al.*, *Nat. Med.* **24**, 724–730 (2018).
14. S. L. Goff *et al.*, *J. Clin. Oncol.* **34**, 2389–2397 (2016).
15. D. Chowell *et al.*, *Science* **359**, 582–587 (2018).
16. J. M. Taube *et al.*, *Clin. Cancer Res.* **20**, 5064–5074 (2014).
17. R. Mandal *et al.*, *Science* **364**, 485–491 (2019).
18. M. Sade-Feldman *et al.*, *Cell* **176**, 404 (2019).
19. B. C. Miller *et al.*, *Nat. Immunol.* **20**, 326–336 (2019).
20. I. Siddiqui *et al.*, *Immunity* **50**, 195–211.e10 (2019).
21. S. Kurtulus *et al.*, *Immunity* **50**, 181–194.e6 (2019).
22. C. S. Jansen *et al.*, *Nature* **576**, 465–470 (2019).
23. Y. Simoni *et al.*, *Nature* **557**, 575–579 (2018).
24. T. Duhen *et al.*, *Nat. Commun.* **9**, 2724 (2018).
25. A. Gros *et al.*, *J. Clin. Invest.* **124**, 2246–2259 (2014).
26. W. Scheper *et al.*, *Nat. Med.* **25**, 89–94 (2019).
27. Y. Simoni *et al.*, bioRxiv 2020.07.15.204172 [Preprint]. 16 July 2020; https://doi.org/10.1101/2020.07.15.204172.
28. S. Krishna *et al.*, *Cancer Res.* **78**, 6159–6170 (2018).
29. K. E. Kortekaas *et al.*, *Cancer Immunol. Res.* **8**, 1311–1321 (2020).
30. C. A. Klebanoff *et al.*, *Clin. Cancer Res.* **17**, 5343–5352 (2011).
31. M. E. Dudley *et al.*, *Clin. Cancer Res.* **16**, 6122–6131 (2010).
32. A. C. Scott *et al.*, *Nature* **571**, 270–274 (2019).
33. L. Gattinoni *et al.*, *Nat. Med.* **17**, 1290–1297 (2011).
34. E. Lugli *et al.*, *J. Clin. Invest.* **123**, 594–599 (2013).
35. R. C. Lynn *et al.*, *Nature* **576**, 293–300 (2019).
36. C. U. Blank *et al.*, *Nat. Rev. Immunol.* **19**, 665–674 (2019).
37. I. Tirosh *et al.*, *Science* **352**, 189–196 (2016).
38. T. D. Prickett *et al.*, *Cancer Immunol. Res.* **4**, 669–678 (2016).
39. A. Pasetto *et al.*, *Cancer Immunol. Res.* **4**, 734–743 (2016).
40. M. Parkhurst *et al.*, *Clin. Cancer Res.* **23**, 2491–2505 (2017).
41. C. J. Cohen *et al.*, *J. Clin. Invest.* **125**, 3981–3991 (2015).
42. S. A. Rosenberg *et al.*, *Clin. Cancer Res.* **17**, 4550–4557 (2011).
43. P. F. Robbins *et al.*, *J. Immunol.* **173**, 7125–7130 (2004).
44. S. P. D'Angelo *et al.*, *Cancer Discov.* **8**, 944–957 (2018).
45. K.-I. Hanada, Z. Yu, G. R. Chappell, A. S. Park, N. P. Restifo, *JCI Insight* **4**, e124405 (2019).
46. D. S. Chen, I. Mellman, *Nature* **541**, 321–330 (2017).
47. A. M. van der Leun, D. S. Thommen, T. N. Schumacher, *Nat. Rev. Cancer* **20**, 218–232 (2020).
48. J. A. Fraietta *et al.*, *Nature* **558**, 307–312 (2018).

ACKNOWLEDGMENTS

We thank D. White for curating the melanoma patient cohort and J. Panopoulos (Flowjo) for helpful discussions on high-dimensional analysis. We thank Z. Rae of the Single Cell Analysis Facility for single-cell sequencing support, R. Yosseph, S. Vodnala, K. Hanada, R. Islam, and the Surgery Branch TIL Laboratory and clinical team for generating and administering TIL therapy to patients. We thank the Frederick National Laboratory for Cancer Research sequencing facility for support for the Single Cell Analysis Facility (contract HHSN261200800001E) and ATAC sequencing. Sequencing was performed with the CCR Genomics Core. This work used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). **Funding:** This research was supported by the Center for Cancer Research intramural research program of the National Cancer Institute. S.K. acknowledges funding support from the NCI Director's Innovation Award. **Author contributions:** S.K., F.J.L., G.A.-B., P.F.R., and S.A.R. conceived the study and designed experiments; S.K., F.J.L., and S.L.G. curated patient samples for inclusion; S.K., F.J.L., A.R.C., R.M., A.S., S.O.A., D.G., Z.Y., V.H., Y.F.L., B.P., P.K., and M.C.K. performed experiments; S.K., F.J.L., E.B., L.J., J.T.A., D.G., V.H., J.J.G., M.P., M.C.K., S.L.G., G.A.B., P.F.K., and S.A.R. performed data analysis; and S.K., F.J.L., P.F.R., and S.A.R. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1328/suppl/DC1
Materials and Methods
Figs. S1 to S19
Tables S1 to S8
References (*49*–*62*)

View/request a protocol for this paper from *Bio-protocol*.

20 April 2020; accepted 9 November 2020
10.1126/science.abb9847

INDIAN MONSOON

Indian monsoon derailed by a North Atlantic wavetrain

P. J. Borah[1,2], V. Venugopal[1,2,3]*, J. Sukhatme[1,2], P. Muddebihal[1], B. N. Goswami[4]

The forecast of Indian monsoon droughts has been predicated on the notion of a season-long rainfall deficit linked to a warm equatorial Pacific. Here we show that nearly half of all droughts over the past century differ from this paradigm in that they (i) occur when Pacific temperatures are near-neutral and (ii) are subseasonal phenomena, characterized by an abrupt decline in late-season rainfall. This severe subseasonal rainfall deficit can be associated with a Rossby wave from mid-latitudes. Specifically, we find that the interaction of upper-level winds with an episodic North Atlantic vorticity anomaly results in a wavetrain that curves toward East Asia, disrupting the monsoon. This atmospheric teleconnection offers an avenue for improved predictability of droughts, especially in the absence of telltale signatures in the Pacific.

Embedded in the interannual and intraseasonal variability of Indian summer monsoon rainfall (ISMR) (*1*–*4*) are complex space-time patterns pertaining to its extreme states, namely floods and droughts (*3*, *5*, *6*). Given that the socioeconomic fabric of one-sixth of the world's population is intricately tied to the state of the monsoon (*7*), these extremes, especially droughts, have a devastating impact on agriculture and the economy. Against this backdrop, despite its high potential predictability, the forecast of ISMR one season in advance remains a big challenge (*8*–*10*). For instance, it is known that anomalously warm waters in the equatorial central and eastern Pacific (El Niño) are associated with some ISMR droughts (*11*–*15*). However, what drives droughts not associated with an El Niño has remained mostly unknown. Further, the prevailing notion is that ISMR droughts are characterized by large-scale season-long rainfall deficit. Indeed, a deeper insight into droughts by way of their subseasonal evolution is essential not only for a better understanding of these extremes but also, practically, for targeted improvement of general circulation models in realizing the potential for improved ISMR prediction.

Over the past century, India has experienced 23 droughts and, as identified in table S1, 13 of these occurred with, and 10 without, an El

[1]Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bangalore 560012, India. [2]Divecha Centre for Climate Change, Indian Institute of Science, Bangalore 560012, India. [3]Interdisciplinary Centre for Water Research, Indian Institute of Science, Bangalore 560012, India. [4]Department of Physics, Cotton University, Guwahati 781001, India.
*Corresponding author. Email: venu.vuruputur@gmail.com or venu@iisc.ac.in

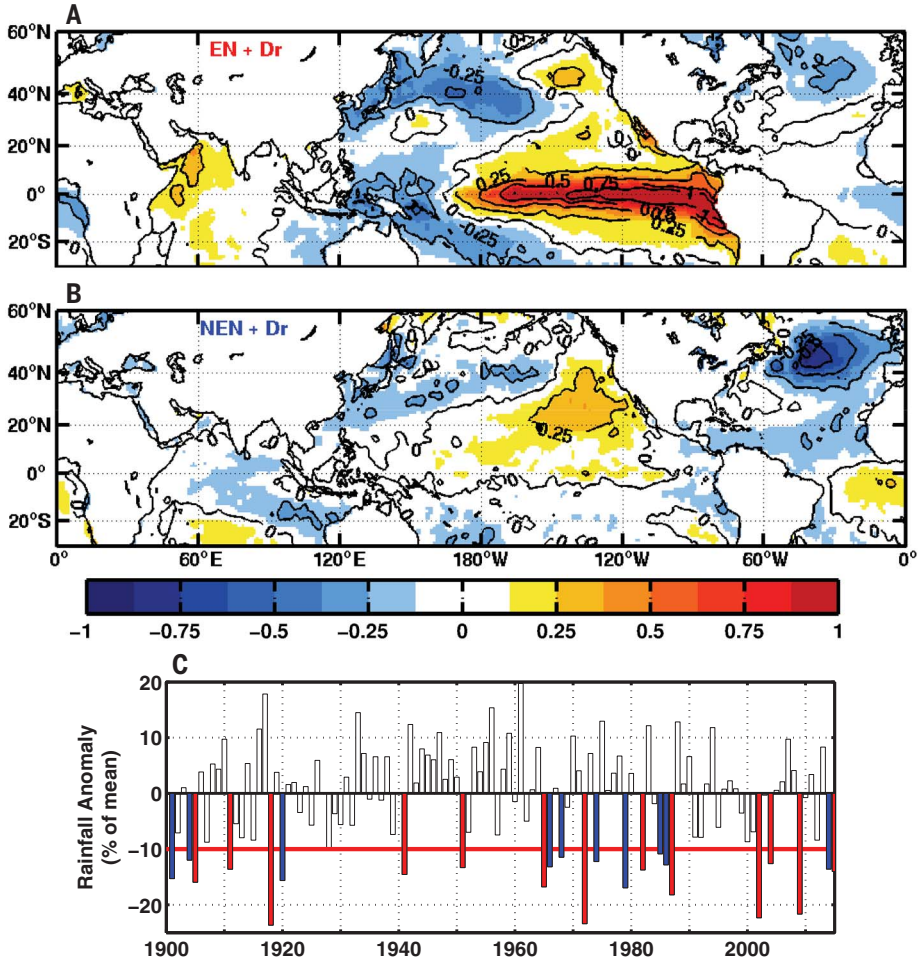


Fig. 1. Sea surface temperature and seasonal rainfall anomalies for the two types of Indian monsoon droughts. Spatial distribution of anomalies of detrended JJAS SST for the two types of Indian monsoon droughts. (**A**) EN + Dr and (**B**) NEN + Dr. The detrending is based on removal of a linear trend for the period 1901–2015 at each grid location, and the maps shown are the average of fluctuations around the trend, for the years listed in table S1. Based on one-degree monthly SST product from the Hadley Centre. Color bar is in degrees Celsius. (**C**) Interannual variation of anomalies of seasonal (JJAS) monsoon rainfall based on the Indian Institute of Tropical Meteorology homogeneous Indian monthly rainfall dataset for the period 1901–2015. The horizontal red line at −10% denotes the threshold for defining a drought, as per the India Meteorological Department (IMD). The departures marked in red (blue) indicate droughts that occurred during an El Niño (no El Niño) year. See methods in the SM.

Niño. These two types of droughts are henceforth referred to as EN + Dr and NEN + Dr (*16*), respectively. This sea surface temperature (SST)–based classification is confirmed in Fig. 1, A and B, which shows the mean of detrended SST anomalies during June through September (JJAS) for the years listed in table S1. Both types of droughts (Fig. 1C), although significantly different from normal years, are statistically indistinguishable from each other at a 5% significance level ($P = 0.09$) on a seasonal scale (fig. S1). This similar final seasonal state in the two categories of droughts, in the face of disparate oceanic conditions (warm versus neutral Pacific), prompts us to investigate their subseasonal evolution.

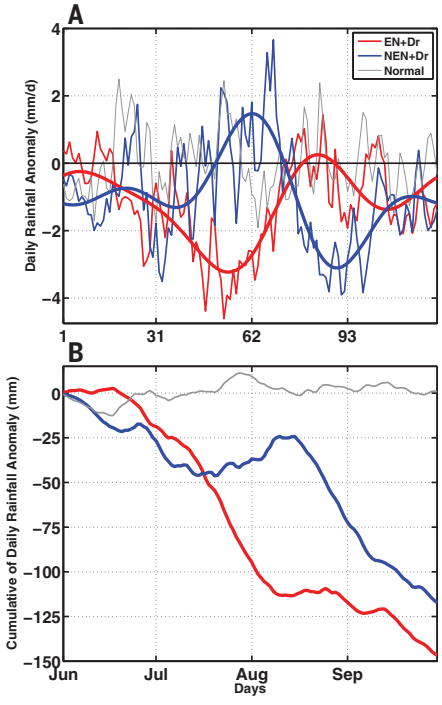


Fig. 2. Daily rainfall anomalies and their cumulative for the two types of droughts. (**A**) Temporal evolution of anomalies of daily area-averaged rainfall over central India (16.5°N to 26.5°N, 74.5°E to 86.5°E; land only) during JJAS, for the two categories of droughts, along with their respective leading harmonics (~120, 60, and 40 days; thick red and blue curves). The thin red and blue curves shown are an average of daily anomalies of 13 EN + Dr and 10 NEN + Dr years listed in table S1. (**B**) Cumulative daily rainfall shown in (A). The average daily rainfall anomalies for a normal year (based on 18 randomly chosen years with a seasonal rainfall anomaly between −5 and 5%) and its cumulative are shown as a gray line in both panels. Based on IMD gridded one-degree daily rainfall data from 1901 to 2015. See methods in the SM.

Figure 2A shows the temporal behavior of daily anomalies of area-averaged central India rainfall [16.5°N to 26.5°N, 74.5°E to 86.5°E; see the methods section in the supplementary materials (SM)] during JJAS, for the two drought types, in comparison to normal years. The daily anomalies for each of the drought years are shown in fig. S2, from which the averages shown in Fig. 2A are estimated. The difference in the evolution of the two drought types is immediately evident. Specifically, in the EN + Dr category (Fig. 2A, thin red curve), negative rainfall anomalies, that is, departures from the daily climatological mean, set in during mid-June and persist till early August. On the other hand, in NEN + Dr (Fig. 2A, thin blue curve), anomalies have more rapid and intense fluctuations culminating in a sharp drop (see long break highlighted by the blue ellipse in fig. S2B) during mid-August. It is this break in late August that drives the monsoon into a state of drought. The envelopes (Fig. 2A, thick curves) of these two types of droughts, represented by their respective leading harmonics (~120, 60, and 40 days), reinforces this distinction.

The difference in timing and duration of the breaks in these two categories is amplified in Fig. 2B, which shows the respective composites of the cumulative daily rainfall anomalies (see fig. S3 for individual years). A comparison of these composites also shows that NEN + Dr has a larger overall deficit than EN + Dr until early July. In sharp contrast, during the core of the monsoon season (July to mid-August), while the deficit persists and deepens in EN droughts, rainfall is near to above normal in NEN + Dr. This apparent recovery in the latter category is interrupted by the sudden break in mid-August. Although the timing of the deficits may be different, the rate of change (drop) in both categories is ~2 to 2.5 mm per day (0 to −100 mm in ~45 days during June and July for EN + Dr versus −25 to −75 mm in 20 days during mid to late August for NEN + Dr). The final state of the two curves in Fig. 2B is in agreement with the fact that, on average, the magnitude of the seasonal deficit is larger in EN + Dr (17%) than in NEN + Dr (13%), as can be inferred from table S1. While the discussion so far has revolved around the rainfall in central India, the contrast holds for all of India as well (fig. S4).

The temporal evolution of rainfall in the two categories of droughts has an associated spatial footprint. Figure 3 shows the spatial patterns of the composites of cumulative daily rainfall anomalies on a 20-day time scale for the two categories of droughts. What stands out here is that in EN + Dr, negative rainfall anomalies appear near the foothills of the Himalayas and Gangetic plains (also, Western Ghats) and then propagate southward. Once the deficit sets in, it persists and covers most of the country by the end of July. Despite the minor recovery during August in the north central plains—corresponding to positive rainfall anomalies in Fig. 2A or, equivalently, the flattening of the cumulative (red) curve in Fig. 2B—the deficit during June and July is sufficient for the season to end in a drought. In contrast, for NEN + Dr, below-normal rainfall in June is largely confined to the west central plains. This is followed by a period when most of the country experiences above-normal rainfall (Fig. 3, I and J, mid-July to mid-August). This path to a near-normal state is derailed by a deficit that appears in the central plains and rapidly expands to cover the entire country within 20 days (Fig. 3K, 20 August to 8 September). Our analysis thus clearly shows that while EN droughts are a consequence of a deficit that persists through the season, NEN droughts are subseasonal phenomena with a moderate early-season deficit compounded by an abrupt decline in late-season rainfall.

The apparent stranglehold of the equatorial Pacific on Indian monsoon variability (in particular, droughts) has been extensively documented (*13*, *14*, *17*). Here we focus on the droughts belonging to the latter category (NEN + Dr), which occur in the absence of anomalously warm waters in the Pacific (Fig. 1B). Specifically, we investigate the late-August break that derails the monsoon, by considering the period when the gradient of rainfall decline is steepest (Fig. 2, blue curves). The anticyclonic lower-level circulation over India confirms the break period (Fig. 4, C and D). Looking outward from India, a standout feature in the atmosphere is the presence of a strong, upper-level cyclonic circulation to the west, in the North Atlantic region (Fig. 4A). Note that this summertime feature is deep tropospheric (fig. S5) (*18*, *19*) and has a tendency to occur episodically in June and August with a duration of 2 to 3 weeks (fig. S6B, blue curve; see also the supplementary text section in the SM).

The interaction of upper-level winds and this vorticity anomaly, as seen in the 200 mbar meridional wind anomalies (shading in Fig. 4A), indicates a wavetrain to the east of the Atlantic that extends all the way to East Asia (*20*, *21*). At 500 mbar (Fig. 4B), the wavetrain curves toward the equator, with a prominent signature over the Indian region, apparently guided by the Tibetan plateau (the narrowing of the wind anomalies to the east of 60°E in Fig. 4B). Finally, the anticyclonic nature of wind anomalies over northern India, signs of which are evident at 500 mbar, becomes clearer at lower levels (700 and 850 mbar; Fig. 4, C and D). This circulation anomaly not only weakens the zonal flow over the Arabian Sea (i.e., disrupting the Somali jet) but also strengthens and spreads to cover all

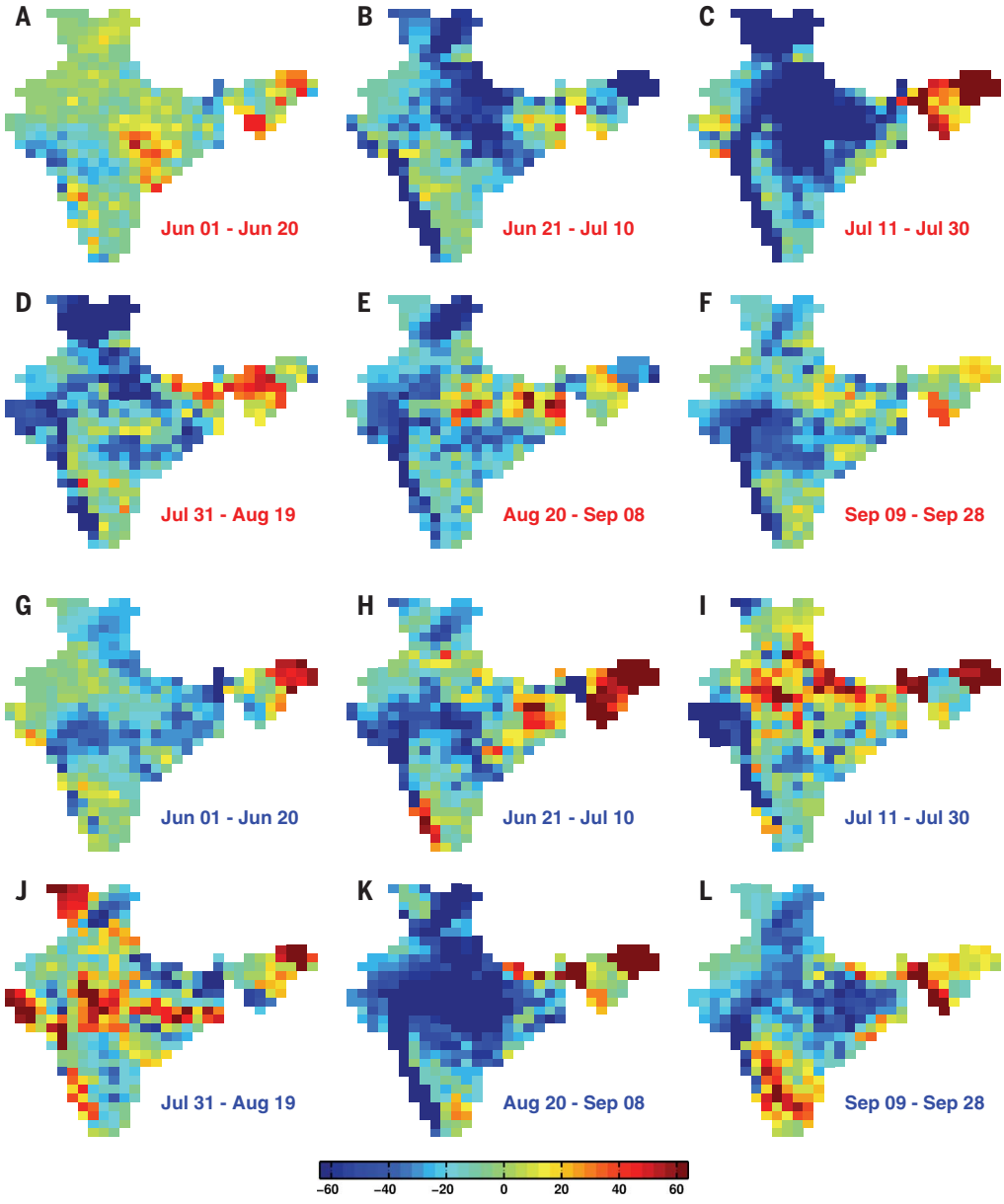


Fig. 3. Spatial distribution of cumulative rainfall anomalies on a 20-day time scale. (**A** to **F**) EN + Dr (top two rows) and (**G** to **L**) NEN + Dr (bottom two rows). The anomalies shown are composites based on 13 EN + Dr and 10 NEN + Dr years, respectively (see table S1). Color bar is in millimeters.

of India, as one descends from 700 to 850 mbar. Contrasting this with the low-level cyclonic circulation over India 20 days prior to the break (fig. S7) highlights the abruptness of this phenomenon (Fig. 3, J and K). Thus, the late-season decline in rainfall can be associated with a large-scale Rossby wave from the North Atlantic to the Indian subcontinent.

The wide-ranging influence of mid-latitudes on the tropics includes the initiation and variability of the Indian monsoon (*22*–*24*), monsoon bursts in Australia (*25*), and short-duration extremes in both hemispheres (*26*). The atmospheric teleconnection and its impact identified in our work are consistent with the recent recognition of the role played by mid-latitude Rossby waves in affecting tropical climate at seasonal and shorter-duration time scales (*26*–*28*). In addition, our diagnostic analysis brings forth two intriguing aspects that call for further scrutiny: First, most NEN droughts seem to coincide with years when the North Atlantic was anomalously cold (fig. S6A); and, second, there appears to be a significantly higher likelihood of episodes of cyclonic vorticity anomalies over cold summer waters (see supplementary text, fig. S6B, and table S2). Both of these associations are statistical in nature and do not imply causality, especially as it is difficult to excite a large-scale atmospheric response to SST anomalies in the mid-latitudes (*29*, *30*). Together, they raise the possibility that there might be an overarching common element behind the excitation and sustenance of cold SSTs and uncharacteristically long duration vorticity anomalies during the summer in the North Atlantic region.

The association of the Atlantic basin with the Indian monsoon in particular has been explored in great detail, albeit mostly on decadal and interannual time scales (*31*–*36*). However, the subseasonal atmospheric bridge identified in this study represents a potentially new pathway in our understanding of monsoon droughts, especially when the

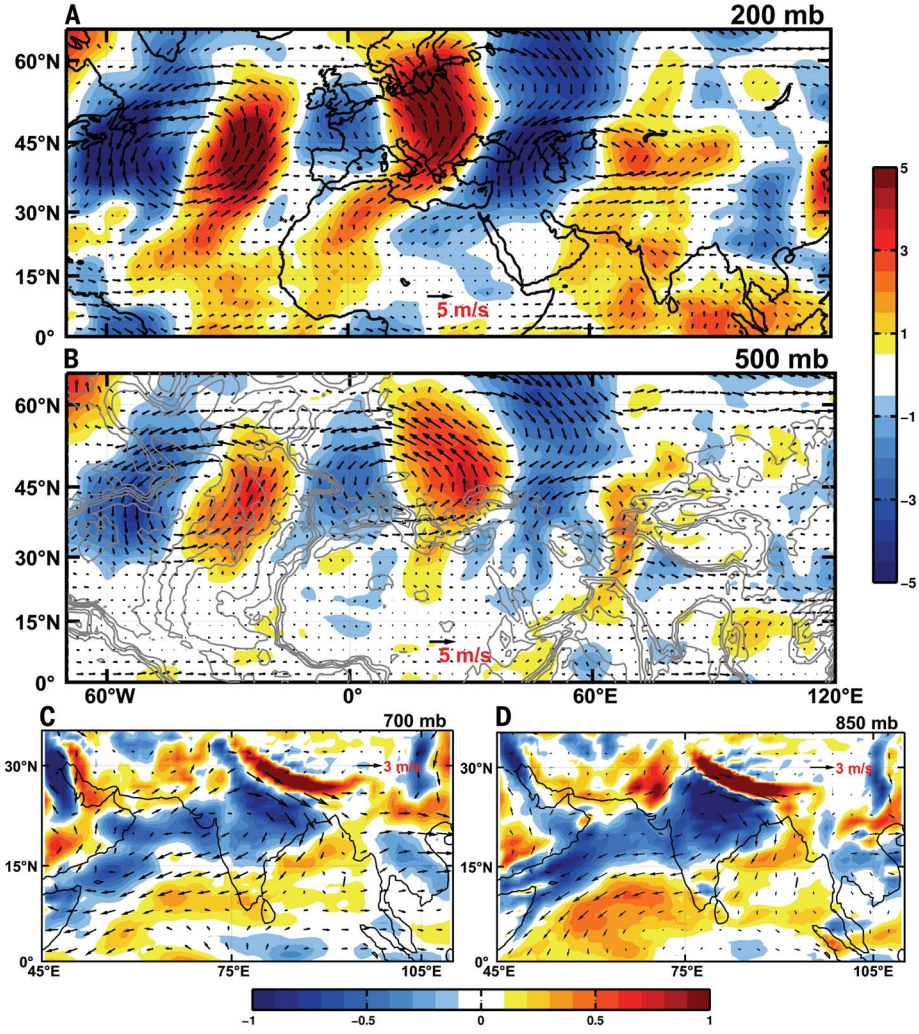


Fig. 4. Wind and vorticity anomalies during the late-August break in NEN droughts. Composites of anomalies of (top two panels) wind (vectors) and meridional velocity (shading) at (**A**) 200 mbar and (**B**) 500 mbar during the first half of the break (days 81 to 90; blue curve in Fig. 2B) in the Indian monsoon rainfall; and (bottom two panels) wind (vectors) and vorticity ($\times 10^5$; shading) at (C) 700 mbar and (D) 850 mbar, based on the respective break periods late in the season for each of the NEN + Dr years. The gray lines in the 500 mbar panel represent topography. Based on ERA 20th Century Reanalysis data. See methods in the SM.

equatorial Pacific temperatures are near-neutral. More pertinently, it points to the need for expanding the existing tropics-centered paradigm of monsoon droughts into a framework that includes extratropical teleconnections (*37*).

REFERENCES AND NOTES

1. B. Parthasarathy, D. A. Mooley, *Mon. Weather Rev.* **106**, 771–781 (1978).
2. V. Krishnamurthy, J. Shukla, *J. Clim.* **13**, 4366–4377 (2000).
3. S. Gadgil, *Annu. Rev. Earth Planet. Sci.* **31**, 429–467 (2003).
4. C. D. Hoyos, P. J. Webster, *J. Clim.* **20**, 4402–4424 (2007).
5. D. Sikka, "Monsoon drought in India," *Tech. Rep. 2*, COLA/CARE, Maryland, USA (1999).
6. B. N. Goswami, in *Intraseasonal Variability in the Atmosphere-Ocean Climate System*, W. K.-M. Lau, D. E. Waliser, Eds. (Springer Praxis Books, Springer, ed. 2, 2012), pp. 21–72.
7. S. Gadgil, S. Gadgil, *Econ. Polit. Wkly.* **41**, 4887–4895 (2006).
8. R. S. Nanjundiah, P. A. Francis, M. Ved, S. Gadgil, *Curr. Sci.* **104**, 1380–1393 (2013).
9. B. Wang *et al.*, *Nat. Commun.* **6**, 7154 (2015).
10. J. Li, B. Wang, *Clim. Dyn.* **46**, 2847–2861 (2016).
11. E. M. Rasmusson, T. H. Carpenter, *Mon. Weather Rev.* **111**, 517–528 (1983).
12. P. J. Webster *et al.*, *J. Geophys. Res. Oceans* **103**, 14451–14510 (1998).
13. K. K. Kumar, B. Rajagopalan, M. Hoerling, G. Bates, M. A. Cane, *Science* **314**, 115–119 (2006).
14. A. G. Turner, H. Annamalai, *Nat. Clim. Chang.* **2**, 587–595 (2012).
15. F. Fan *et al.*, *Atmos. Sci. Lett.* **18**, 175–182 (2017).
16. H. Varikoden, J. V. Revadekar, Y. Choudhary, B. Preethi, *Int. J. Climatol.* **35**, 1916–1925 (2014).
17. X. Li, M. Ting, *Geophys. Res. Lett.* **42**, 3502–3512 (2015).
18. T. Palmer, S. Zhaobo, *Q. J. R. Meteorol. Soc.* **111**, 947–975 (1985).
19. C. Wang, *Clim. Dyn.* **53**, 5119–5136 (2019).
20. B. J. Hoskins, G.-Y. Yang, *J. Atmos. Sci.* **57**, 1197–1213 (2000).
21. G. Branstator, J. Teng, *J. Atmos. Sci.* **74**, 1513–1532 (2017).
22. S. Bordoni, T. Schneider, *Nat. Geosci.* **1**, 515–519 (2008).
23. R. Krishnan, V. Kumar, M. Sugi, J. Yoshimura, *J. Atmos. Sci.* **66**, 553–578 (2009).
24. R. K. Yadav, *Clim. Dyn.* **32**, 549–563 (2009).
25. S. Narsey, M. J. Reeder, D. Ackerley, C. Jakob, *J. Clim.* **30**, 5377–5393 (2017).
26. N. Boers *et al.*, *Nature* **566**, 373–377 (2019).
27. Y.-K. Lim, *Clim. Dyn.* **44**, 3211–3222 (2015).
28. L. O'Brien, M. J. Reeder, *Q. J. R. Meteorol. Soc.* **143**, 2374–2388 (2017).
29. Y. Kushnir, I. M. Held, *J. Clim.* **9**, 1208–1220 (1996).
30. Y. Kushnir *et al.*, *J. Clim.* **15**, 2233–2256 (2002).
31. B. N. Goswami, M. S. Madhusoodanan, C. P. Neema, D. Sengupta, *Geophys. Res. Lett.* **33**, L02706 (2006).
32. R. Lu, B. Dong, H. Ding, *Geophys. Res. Lett.* **33**, L24701 (2006).
33. M. Rajeevan, L. Sridhar, *Geophys. Res. Lett.* **35**, L21704 (2008).
34. C. Wang, F. Kucharski, R. Barimalala, A. Bracco, *Meteorol. Z. (Berl.)* **18**, 445–454 (2009).
35. L. Krishnamurthy, V. Krishnamurthy, *Clim. Dyn.* **46**, 2269–2285 (2016).
36. C. T. Sabeerali, R. S. Ajayamohan, H. K. Bangalath, N. Chen, *Geophys. Res. Lett.* **46**, 4460–4467 (2019).
37. C. Stan *et al.*, *Rev. Geophys.* **55**, 902–937 (2017).

ACKNOWLEDGMENTS

We thank R. S. Nanjundiah, J. Srinivasan, D. Sengupta, and D. S. Battisti for useful discussions; J. M. Wallace, R. Roca, and A. H. Sobel for constructive suggestions on improving the narrative; and R. Matthew for providing seasonal anomalies of monsoon rainfall over the homogeneous regions. We thank IMD, UK Met Office, and ECMWF for daily gridded rainfall, monthly gridded SST, and ERA 20th Century Reanalysis datasets, respectively. **Funding:** P.J.B. thanks the Department of Science and Technology (DST), Government of India (GoI), for support under the INSPIRE PhD fellowship (IF160707). V.V. and J.S. thank DST, GoI, for financial support (DST/CCP/NCM/75/2017) under their Climate Change Programme. V.V. thanks the Ministry of Earth Sciences, GoI, for their support under the project (MOES/PAMC/H&C/41/2013-PC-II) "Advanced Hydrologic Research and Knowledge Dissemination." J.S. thanks the University Grants Commission, GoI, for support under the Indo-Israel Joint Research Programme (F 6-3/2018). B.N.G. thanks the Science and Engineering Research Board, GoI, for the SERB Distinguished Fellowship. **Author contributions:** V.V. formulated the initial question of dominant drought patterns. P.M. performed the preliminary set of computations with monthly rainfall. P.J.B. performed all subsequent calculations with daily data. V.V. and J.S. were involved in all of the analysis and interpretation and took the lead in writing the manuscript. B.N.G contributed to the interpretation and writing of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data sources and their access has been provided in the data and methods section in the SM.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1335/suppl/DC1
Data and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 and S2
References (*38*–*43*)

4 July 2019; accepted 21 October 2020
10.1126/science.aay6043

CORONAVIRUS

Preexisting and de novo humoral immunity to SARS-CoV-2 in humans

Kevin W. Ng[1]*, Nikhil Faulkner[1]*, Georgina H. Cornish[1]*, Annachiara Rosa[2]*, Ruth Harvey[3], Saira Hussain[3], Rachel Ulferts[9], Christopher Earl[4], Antoni G. Wrobel[5], Donald J. Benton[5], Chloe Roustan[6], William Bolland[1], Rachael Thompson[1], Ana Agua-Doce[7], Philip Hobson[7], Judith Heaney[13], Hannah Rickman[13], Stavroula Paraskevopoulou[13], Catherine F. Houlihan[13,14], Kirsty Thomson[13], Emilie Sanchez[13], Gee Yen Shin[13], Moira J. Spyer[13,15], Dhira Joshi[8], Nicola O'Reilly[8], Philip A. Walker[6], Svend Kjaer[6], Andrew Riddell[7], Catherine Moore[16], Bethany R. Jebson[17,19], Meredyth Wilkinson[17,19], Lucy R. Marshall[17,19], Elizabeth C. Rosser[17,18], Anna Radziszewska[17,18], Hannah Peckham[17,18], Coziana Ciurtin[17,18], Lucy R. Wedderburn[17,19], Rupert Beale[9], Charles Swanton[10], Sonia Gandhi[11], Brigitta Stockinger[12], John McCauley[3], Steve J. Gamblin[5], Laura E. McCoy[14]†, Peter Cherepanov[2]†, Eleni Nastouli[13,15]†, George Kassiotis[1,20]†

Zoonotic introduction of novel coronaviruses may encounter preexisting immunity in humans. Using diverse assays for antibodies recognizing SARS-CoV-2 proteins, we detected preexisting humoral immunity. SARS-CoV-2 spike glycoprotein (S)–reactive antibodies were detectable using a flow cytometry–based method in SARS-CoV-2–uninfected individuals and were particularly prevalent in children and adolescents. They were predominantly of the immunoglobulin G (IgG) class and targeted the S2 subunit. By contrast, SARS-CoV-2 infection induced higher titers of SARS-CoV-2 S–reactive IgG antibodies targeting both the S1 and S2 subunits, and concomitant IgM and IgA antibodies, lasting throughout the observation period. SARS-CoV-2–uninfected donor sera exhibited specific neutralizing activity against SARS-CoV-2 and SARS-CoV-2 S pseudotypes. Distinguishing preexisting and de novo immunity will be critical for our understanding of susceptibility to and the natural course of SARS-CoV-2 infection.

Immune cross-reactivity among seasonally spreading human coronaviruses (HCoVs) has long been hypothesized to provide effective but transient cross-protection against distinct HCoVs (*1*, *2*). To determine the degree of cross-reactivity between HCoVs and SARS-CoV-2, we developed a flow cytometry–based assay for SARS-CoV-2–binding antibodies. The main target for such antibodies is the spike glycoprotein (S), which is proteolytically processed into the S1 and S2 subunits, mediating target cell attachment and entry, respectively.

The S1-specific CR3022 antibody stained a smaller percentage of SARS-CoV-2 S–expressing human embryonic kidney (HEK) 293T cells and with lower intensity than COVID-19 convalescent sera (fig. S1), indicating that polyclonal immunoglobulin G (IgG) antibodies targeted a wider range of epitopes naturally processed and displayed on these cells. This assay also detected SARS-CoV-2 S–reactive IgM and IgA antibodies in COVID-19 convalescent sera (fig. S2). Indeed, the presence of SARS-CoV-2 S–reactive antibodies of all three Ig classes ($IgG^+IgM^+IgA^+$) distinguished COVID-19 sera from control sera with a high degree of sensitivity and specificity (Fig. 1A and fig. S3). All 156 seroconverted COVID-19 patients had contemporaneous IgG, IgM, and IgA responses to SARS-CoV-2 S throughout the observation period, with the exception of two patients who only had IgG antibodies (figs. S4 and S5). One of these patients was a bone marrow transplantation recipient who experienced HCoV infection 1 month before SARS-CoV-2 infection (fig. S6). Unexpectedly, a small proportion of SARS-CoV-2–uninfected patients sampled before or during the early spread of SARS-CoV-2 in the United Kingdom (table S1) also had SARS-CoV-2 S–binding IgG antibodies, but not IgM or IgA antibodies (Fig. 1A), suggesting the presence of cross-reactive immunological memory.

The S2 subunit exhibits a higher degree of homology among coronaviruses than S1 (fig. S7) and was likely the main target of cross-reactive antibodies. Competition with recombinant soluble S1 or S2 at doses that blocked binding of specific monoclonal antibodies (fig. S8) did not affect the frequency of cells stained with COVID-19 patient sera, although the intensity of staining was reduced by 31 and 37%, respectively (Fig. 1, B to D), indicating recognition of both S1 and S2. By contrast, soluble S2 completely abolished staining with SARS-CoV-2–uninfected patient sera, whereas soluble S1 had no effect (Fig. 1, B to D). Thus, SARS-CoV-2–uninfected patient sera cross-react with SARS-CoV-2 S2, and COVID-19 patient sera additionally recognize S1.

SARS-CoV-2 S–reactive IgG antibodies were detected by flow cytometry in five of 34 SARS-CoV-2–uninfected individuals with HCoV infection confirmed by reverse transcription–quantitative polymerase chain reaction, as well as in one of 31 individuals without recent HCoV infection (Fig. 2A and fig. S4A). This suggested that cross-reactivity may have persisted from earlier HCoV infections rather than having been induced by the most recent one.

To confirm antibody cross-reactivity using an independent assay, we developed enzyme-linked immunosorbent assays (ELISAs) using recombinant SARS-CoV-2–stabilized trimeric S ectodomain, S1, receptor-binding domain (RBD), or nucleoprotein (N). Rates of IgG seropositivity by SARS-CoV-2 S1–coated ELISA were congruent with, but generally lower than, those by flow cytometry (fig. S9). The three SARS-CoV-2–uninfected individuals with the highest cross-recognition of S by flow cytometry, plus an additional four individuals, had ELISA-detectable IgG antibodies against the SARS-CoV-2 S ectodomain, as well as N (Fig. 2A and fig. S4, B to D). By contrast, none of the control samples had ELISA-detectable IgG antibodies against the less-conserved SARS-CoV-2 S1 or RBD (Fig. 2A and fig. S4, B to D).

The prevalence of such cross-reactive antibodies was further examined in additional healthy donor cohorts (table S1). Among 50 SARS-CoV-2–uninfected pregnant women sampled in May of 2018, five showed evidence for SARS-CoV-2 S–reactive IgG antibodies, but not IgM or IgA antibodies (Fig. 2B and fig.

[1]Retroviral Immunology, The Francis Crick Institute, London NW1 1AT, UK. [2]Chromatin Structure and Mobile DNA Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [3]Worldwide Influenza Centre, The Francis Crick Institute, London NW1 1AT, UK. [4]Signalling and Structural Biology Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [5]Structural Biology of Disease Processes Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [6]Structural Biology STP, The Francis Crick Institute, London NW1 1AT, UK. [7]Flow Cytometry STP, The Francis Crick Institute, London NW1 1AT, UK. [8]Peptide Chemistry, The Francis Crick Institute, London NW1 1AT, UK. [9]Cell Biology of Infection Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [10]Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [11]Neurodegeneration Biology Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [12]AhRimmunity Laboratory, The Francis Crick Institute, London NW1 1AT, UK. [13]University College London Hospitals (UCLH) NHS Trust, London NW1 2BU, UK. [14]Division of Infection and Immunity, University College London (UCL), London WC1E 6BT, UK. [15]Department of Population, Policy and Practice, Great Ormond Street Institute for Child Health (ICH), UCL, London WC1N 1EH, UK. [16]Public Health Wales, University Hospital of Wales, Cardiff CF14 4XW, UK. [17]Centre for Adolescent Rheumatology Versus Arthritis at UCL, UCLH, Great Ormond Street Hospital (GOSH), London WC1N 3JH, UK. [18]Centre for Rheumatology Research, Division of Medicine, UCL, London, WC1E 6BT, UK. [19]UCL Great Ormond Street Institute for Child Health (ICH), UCL, London WC1N 1EH, UK. [20]Department of Medicine, Faculty of Medicine, Imperial College London, London W2 1PG, UK.
*These authors contributed equally to this work.
†Corresponding author. Email: george.kassiotis@crick.ac.uk (G.K.); e.nastouli@ucl.ac.uk (E.N.); peter.cherepanov@crick.ac.uk (P.C.); l.mccoy@ucl.ac.uk (L.E.M.)

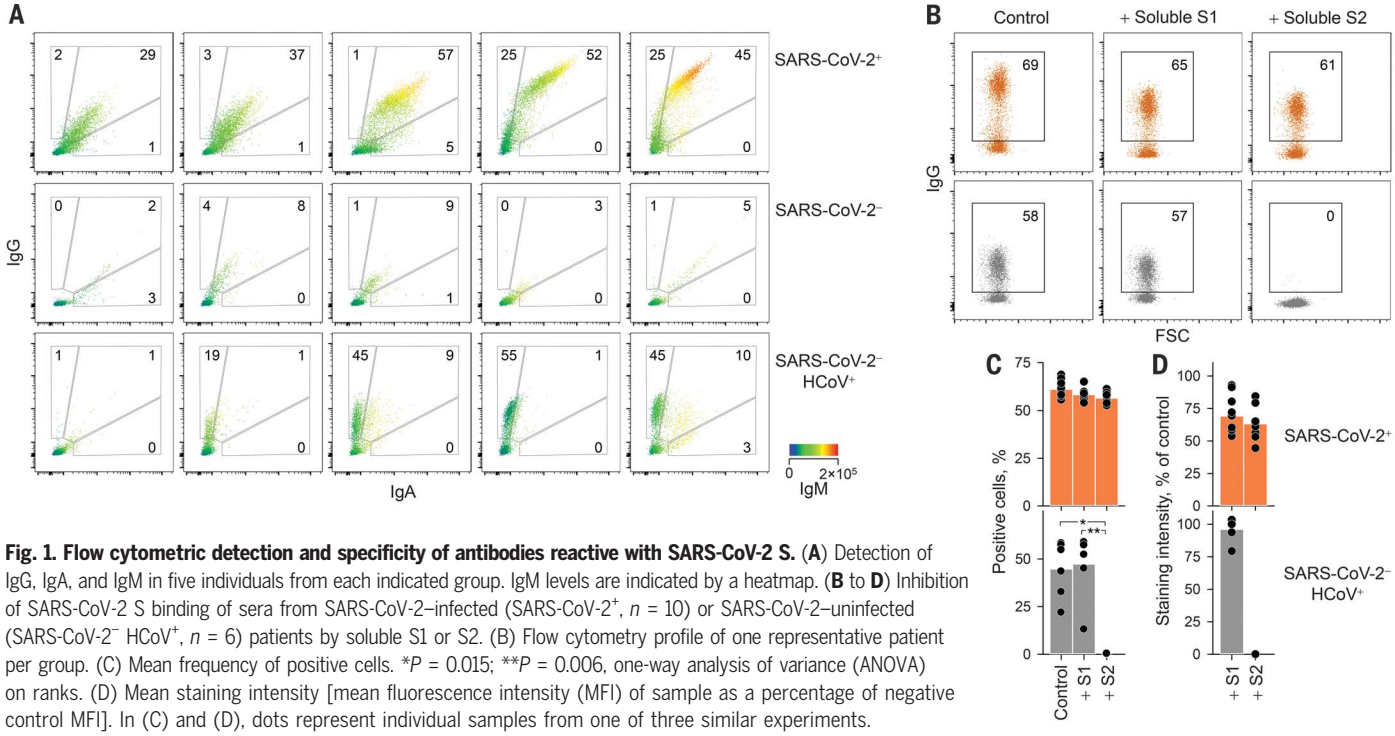


Fig. 1. Flow cytometric detection and specificity of antibodies reactive with SARS-CoV-2 S. (**A**) Detection of IgG, IgA, and IgM in five individuals from each indicated group. IgM levels are indicated by a heatmap. (**B** to **D**) Inhibition of SARS-CoV-2 S binding of sera from SARS-CoV-2–infected (SARS-CoV-2$^+$, *n* = 10) or SARS-CoV-2–uninfected (SARS-CoV-2$^-$ HCoV$^+$, *n* = 6) patients by soluble S1 or S2. (B) Flow cytometry profile of one representative patient per group. (C) Mean frequency of positive cells. **P* = 0.015; ***P* = 0.006, one-way analysis of variance (ANOVA) on ranks. (D) Mean staining intensity [mean fluorescence intensity (MFI) of sample as a percentage of negative control MFI]. In (C) and (D), dots represent individual samples from one of three similar experiments.

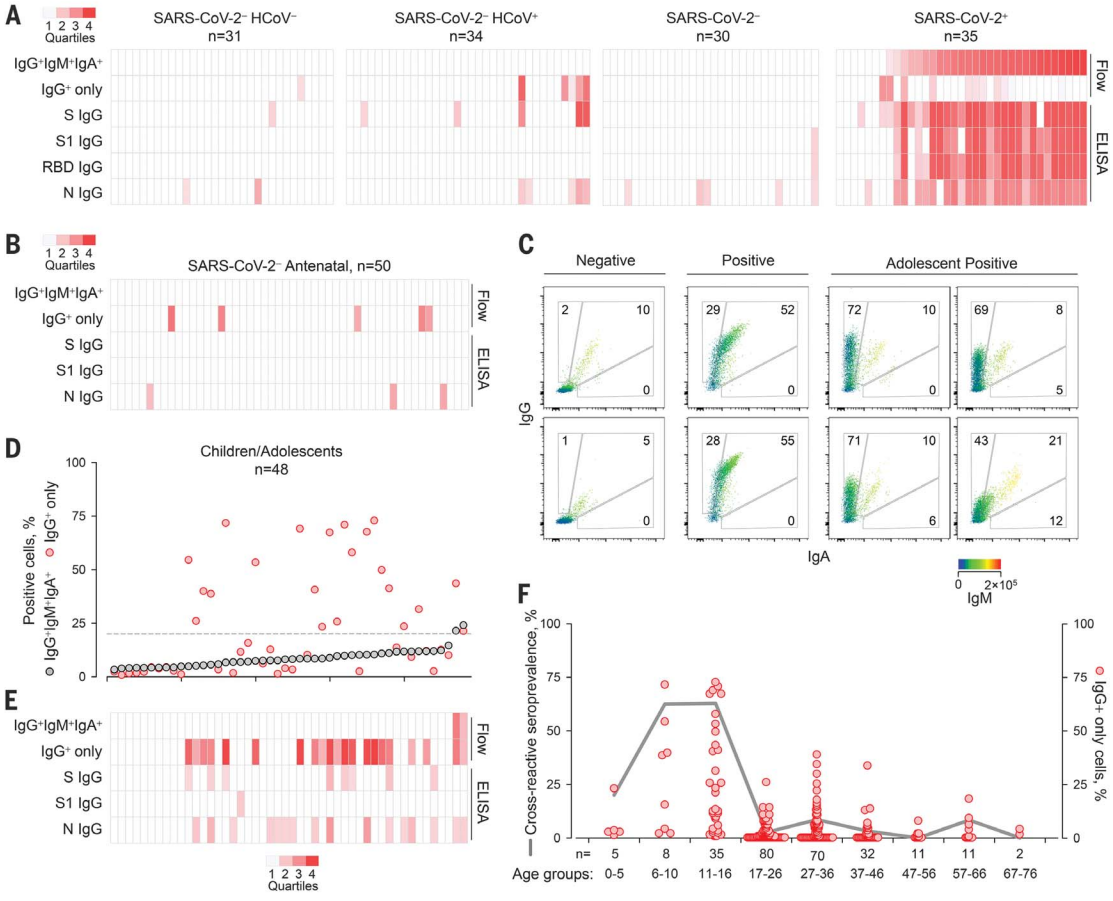


Fig. 2. Prevalence of SARS-CoV-2 S–cross-reactive antibodies detected by different methods. (**A**) Flow cytometry and ELISA results for each sample in cohorts A and C to E listed in table S1. (**B**) Flow cytometry and ELISA results for serum samples from SARS-CoV-2–uninfected pregnant women. (**C** to **E**) SARS-CoV-2 S–cross-reactive antibodies in healthy children and adolescents. (C) Representative flow cytometry profiles of seronegative donors (Negative) or COVID-19 patients (Positive) and of SARS-CoV-2–uninfected adolescents with SARS-CoV-2 cross-reactive antibodies. (D) Frequency of cells stained with all three antibody classes (IgG$^+$IgM$^+$IgA$^+$) or only with IgG (IgG$^+$) ranked by their IgG$^+$IgM$^+$IgA$^+$ frequency. The dashed line denotes the assay sensitivity cutoff. (E) Flow cytometry and ELISA results for each sample. (**F**) Prevalence of SARS-CoV-2 S–cross-reactive antibodies in the indicated age groups (line) and frequency of cells that stained only with IgG (dots) in all samples for which the date of birth was known. The heatmaps in (A), (B), and (E) represent the quartile values above each assay's technical cutoff.

S10). In a separate cohort of 101 SARS-CoV-2–uninfected donors sampled in May of 2019, three had SARS-CoV-2 S–reactive IgG antibodies (fig. S11) that did not correlate with antibodies to the diverse viruses and bacteria also present in several of these samples. SARS-CoV-2 S–reactive IgM and IgA were also detected in two of these donors, albeit at considerably lower levels than in COVID-19 patients (fig. S11), suggestive of recent or ongoing response. In an additional cohort of 13 donors recently infected with HCoVs, only one had SARS-CoV-2 S–reactive IgG antibodies, and these were at very low levels (fig. S12). This suggested that their emergence was not simply a common transient event after each HCoV infection in this age group (median age 51 years; table S1). Instead, given that HCoV-reactive antibodies are present in virtually all adults (*3*–*5*), the rarity of SARS-CoV-2 S cross-reactivity (16 of 302; 5.29%) indicates additional requirements such as random B cell receptor repertoire focusing or frequency of HCoV infection rather than time since the last HCoV infection. Indeed, the frequency of HCoV infection displays a characteristic age distribution, being the highest in children and adolescents (*1*, *4*–*8*). We therefore examined a cohort of younger SARS-CoV-2–uninfected healthy donors (age 1 to 16 years; table S1) sampled between 2011 and 2018. At least 21 of these 48 donors had detectable levels of SARS-CoV-2 S–reactive IgG antibodies (Fig. 2, C to E), whereas only one of an additional cohort of 43 young adults (age 17 to 25 years; table S1) had such antibodies (Fig. 2F). Staining with sera from SARS-CoV-2–uninfected children and adolescents was specific to HEK293T cells expressing SARS-CoV-2 S, but not the unrelated HERV-K113 envelope glycoprotein, and was outcompeted by soluble SARS-CoV-2 S2 (fig. S13). The prevalence of SARS-CoV-2 S–reactive IgG antibodies peaked at 62% between 6 and 16 years of age (Fig. 2F), when HCoV seroconversion in this age group also peaks (*3*, *4*, *6*, *7*), and was significantly higher than in adults ($P < 0.00001$, Fisher's exact test).

To determine the potential consequences of antibody cross-reactivity, we examined the ability of preexisting antibodies to inhibit SARS-CoV-2 entry into HEK293T cells (fig. S14 and supplementary text). Although not expected to directly inhibit RBD-mediated cell attachment, S2-targeting antibodies that can neutralize SARS-CoV-2 have recently been discovered (*9*, *10*). HEK293T cell infection with SARS-CoV-2 S pseudotypes was efficiently inhibited by sera from seroconverted (Ab^+) COVID-19 patients, but not from those who had not yet seroconverted (Ab^-) (Fig. 3A). Sera from SARS-CoV-2–uninfected donors with SARS-CoV-2 S–reactive antibodies also neutralized these pseudotypes, whereas none of the sera neutralized vesicular stomatitis virus (VSV) glycoprotein pseudotypes (Fig. 3A). Comparable neutralization of SARS-CoV-2 S pseudotypes was also observed with sera from SARS-CoV-2–uninfected adolescents (Fig. 3A). Moreover, most of the sera from SARS-CoV-2–uninfected donors with flow cytometry–detectable cross-reactive antibodies also neutralized authentic SARS-CoV-2 infection of Vero E6 cells, albeit on average less potently than COVID-19 patient sera (Fig. 3B). By contrast, sera from SARS-CoV-2–uninfected patients without cross-reactive antibodies exhibited no

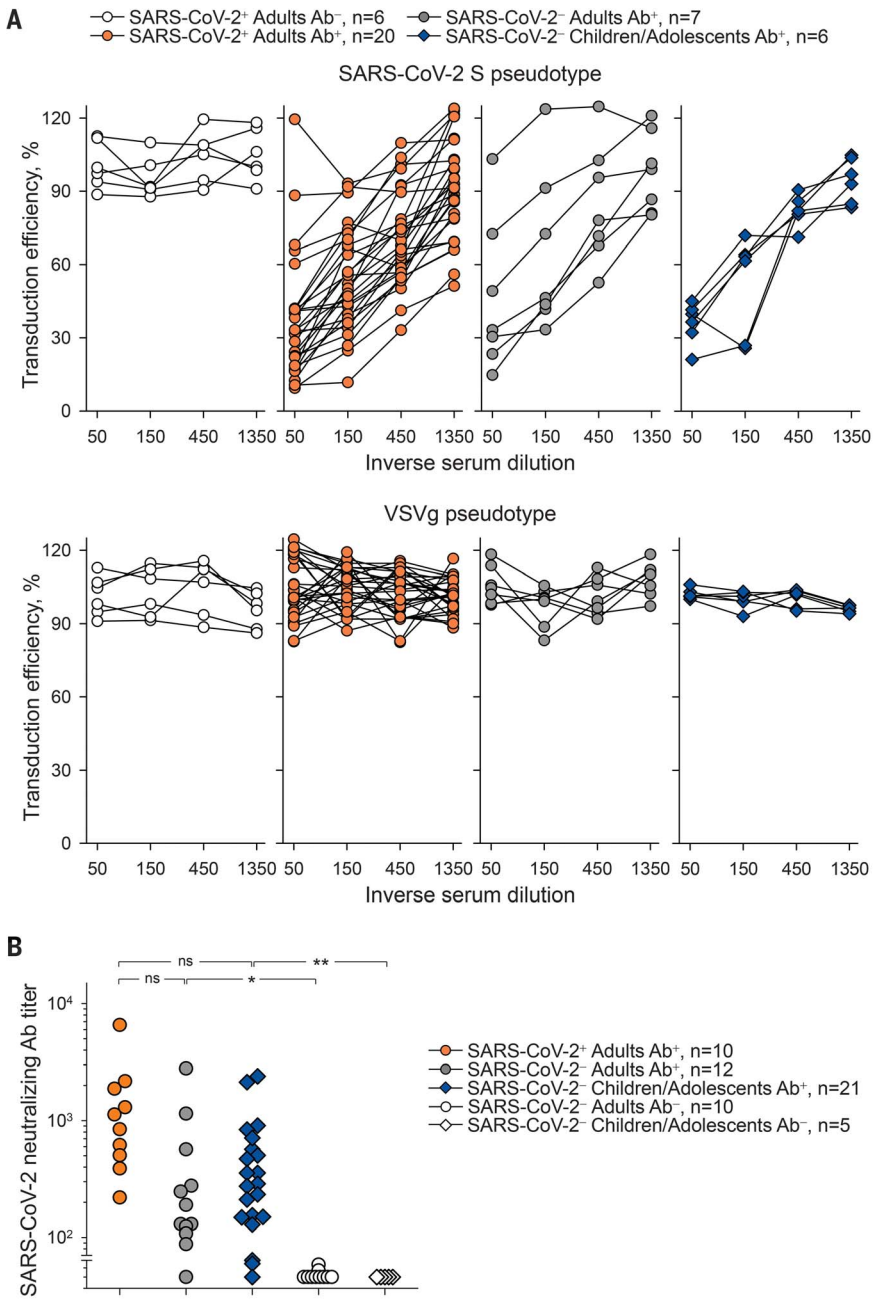


Fig. 3. Neutralization of SARS-CoV-2 S pseudotypes and authentic SARS-CoV-2 by SARS-CoV-2–infected and –uninfected patient sera. (**A**) Inhibition of transduction efficiency of SARS-CoV-2 S or VSVg pseudotypes by adult COVID-19 patients who seroconverted (SARS-CoV-2$^+$ Adults Ab$^+$) or not (SARS-CoV-2$^+$ Adults Ab$^-$) and SARS-CoV-2–uninfected adult donors (SARS-CoV-2$^-$ Adults Ab$^+$) or children and adolescent donors (SARS-CoV-2$^-$ Children/Adolescents Ab$^+$) with SARS-CoV-2 S–binding antibodies. Each line is an individual serum sample. (**B**) Authentic SARS-CoV-2 neutralization titers of sera from the same donors as in (A), as well as SARS-CoV-2–uninfected donors without SARS-CoV-2 S–binding antibodies (Ab$^-$). Dots represent individual samples. *P = 0.037; **P = 0.014; ns, not significant by one-way ANOVA on ranks.

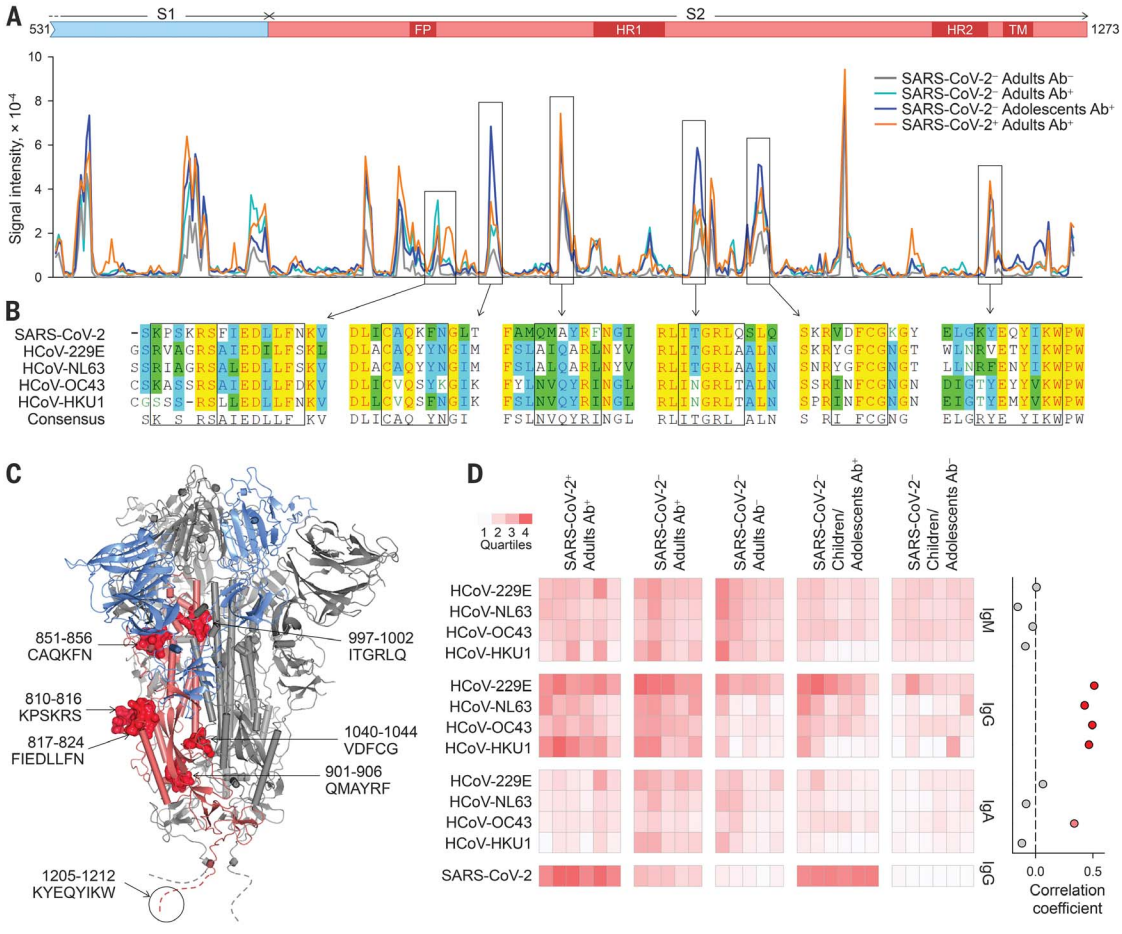


Fig. 4. Mapping of cross-reactive epitopes in SARS-CoV-2 S. (**A**) Signal intensity for each overlapping peptide along the length of SARS-CoV-2 S covered in the peptide arrays using pooled sera with (Ab^+) or without (Ab^-) flow cytometry–detectable SARS-CoV-2 S–reactive antibodies. Differentially recognized peaks are boxed. (**B**) Alignment of the amino acid sequences of SARS-CoV-2 and HCoV S glycoproteins. Boxes indicate predicted core epitopes. (**C**) Mapping of predicted epitopes targeted on the trimeric SARS-CoV-2 spike. The S1 (blue) and S2 (pink) subunits of one monomer are colored. Epitopes are shown for one monomer; the circled dashed line represent the membrane proximal region not present in the structure. (**D**) Left: Reactivity with the S glycoproteins of each HCoV of the indicated sera with (Ab^+) or without (Ab^-) flow cytometry–detectable SARS-CoV-2 S–reactive antibodies as determined by flow cytometry. Each column is an individual sample. Rows depict the staining for each antibody class. Right: Correlation coefficients between percentages of IgG staining for SARS-CoV-2 S and IgG, IgM, and IgA staining for each HCoV S glycoprotein.

neutralizing activity (Fig. 3B). Antiviral antibodies may also enhance viral entry by Fc receptor–mediated antibody-dependent enhancement. However, entry of SARS-CoV-2 S pseudotypes was not enhanced by either COVID-19 patient sera or SARS-CoV-2–uninfected patient sera in FcγRIIA-expressing K-562 cells (fig. S15).

Collectively, these findings highlight functionally relevant antigenic epitopes conserved within the S2 subunit. Over its entire length, SARS-CoV-2 S exhibits marginally closer homology with the S proteins of the betacoronaviruses HCoV-OC43 and HCoV-HKU1 than with the alphacoronaviruses HCoV-NL63 and HCoV-229E (fig. S16A). To probe shared epitopes, we constructed overlapping peptide arrays spanning the last 743 amino acids of SARS-CoV-2 S (fig. S16B). Multiple putative epitopes were differentially recognized by sera with cross-reactive antibodies (Ab^+), were reasonably conserved, and most mapped to the surface of S2 (Fig. 4, A and B, and table S2). An epitope overlapping the S2 fusion peptide was also recently identified as being cross-reactive with the corresponding peptides from HCoV-OC43 and HCoV-229E (*11*). Cross-reactivity with the identified epitopes was further supported by ELISAs coated with synthetic peptides (fig. S17).

As expected (*3*–*5*), reactivity with one or more HCoVs was detectable by flow cytometry in all sera (Fig. 4D and fig. S18). However, IgG and IgA reactivity against HCoVs was higher in SARS-CoV-2–uninfected adults with SARS-CoV-2–reactive IgG compared with those without ($P = 1.4 \times 10^{-6}$ for IgG and $P = 0.017$ for IgA, Student's t test) and in SARS-CoV-2–uninfected children or adolescents with SARS-CoV-2–reactive IgG compared with those without ($P = 0.010$ for IgG and $P = 0.021$ for IgA, Student's t test) (Fig. 4D), supporting a direct link between the two. Accordingly, IgG reactivity against each HCoV type was independently correlated with the presence of SARS-CoV-2–reactive antibodies (Fig. 4D).

Our results from multiple independent assays demonstrate the presence of preexisting antibodies recognizing SARS-CoV-2 in uninfected individuals. Identification of conserved epitopes in S2 targeted by neutralizing antibodies may hold promise for a universal vaccine protecting against current as well as future CoVs. Together with preexisting T cell (*12*–*14*) and B cell (*10*, *15*) memory, antibody cross-reactivity between seasonal HCoVs and SARS-CoV-2 may have important ramifications for natural infection. Epidemiological studies of HCoV transmission suggest that cross-protective immunity is unlikely to be sterilizing or long-lasting (*8*), which is also supported by repeated reinfection (*2*, *16*). Nevertheless, prior immunity induced by one HCoV can reduce the transmission of homologous and heterologous HCoVs and ameliorate the symptoms when transmission is not prevented (*1*, *2*). A possible modification of COVID-19 severity by prior HCoV infection may account for the age distribution of COVID-19 susceptibility, in which higher HCoV infection rates in children than in adults (*4*, *6*) correlate with relative protection from COVID-19 (*17*) and may also shape seasonal and geographical patterns of transmission. It is imperative that any effect, positive or negative, of preexisting

HCoV-elicited immunity on the natural course of SARS-CoV-2 infection be fully delineated.

REFERENCES

1. R. W. Aldridge *et al*., *Wellcome Open Res.* **5**, 52 (2020).
2. K. A. Callow, H. F. Parry, M. Sergeant, D. A. Tyrrell, *Epidemiol. Infect.* **105**, 435–446 (1990).
3. E. G. Severance *et al*., *Clin. Vaccine Immunol.* **15**, 1805–1810 (2008).
4. R. Dijkman *et al*., *J. Clin. Microbiol.* **46**, 2368–2373 (2008).
5. A. T. Huang *et al*., *Nat. Commun.* **11**, 4704 (2020).
6. N. Friedman *et al*., *Viruses* **10**, 515 (2018).
7. S. Nickbakhsh *et al*., *J. Infect. Dis.* (2020).
8. A. S. Monto *et al*., *J. Infect. Dis.* **222**, 9–16 (2020).
9. X. Chi *et al*., *Science* **369**, 650–655 (2020).
10. G. Song *et al*., bioRxiv 308965 [Preprint]. 23 September 2020; https://doi.org/10.1101/2020.09.22.308965.
11. E. Shrock *et al*., *Science* eabd4250 (2020).
12. A. Grifoni *et al*., *Cell* **181**, 1489–1501.e15 (2020).
13. J. Braun *et al*., *Nature* (2020).
14. N. Le Bert *et al*., *Nature* **584**, 457–462 (2020).
15. P. Nguyen-Contant *et al*., *mBio* **11**, e01991–e01920 (2020).
16. P. K. Kiyuka *et al*., *J. Infect. Dis.* **217**, 1728–1739 (2018).
17. R. Castagnoli *et al*., *JAMA Pediatr.* **174**, 882–889 (2020).

ACKNOWLEDGMENTS

We thank L. James and J. Luptak for the SARV CoV2 N expression construct and M. Pizzato for the SARS CoV2 S cDNA; the entire CRICK COVID-19 Consortium; the Cell Services and High Throughput Screening facilities at the Francis Crick Institute and UCLH Biochemistry (A. Goyale and C. Wilson); and M. Bennet and S. Caidan for training and support in the high-containment laboratory. **Funding:** This work was supported by the Centre of Excellence Centre for Adolescent Rheumatology Versus Arthritis (grant no. 2159), the Great Ormond Street Children's Charity, CureJM Foundation, the NIHR Biomedical Research Centres at GOSH and UCLH, and by the Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust. **Author contributions:** C.C., L.R.W., R.B., C.S., S.G., B.S., J.M., S.J.G., L.E.M., P.C., E.N., and G.K. designed the experiment. K.W.N., N.F., G.H.C., A.Ro., R.H., S.H., R.U., C.E., A.G.W., D.J.B., C.R., W.B., R.T., A.A.-D., P.H., and D.J. performed the investigations. J.H., H.R., S.P., C.F.H., K.T., E.S., G.Y.S., M.J.S., P.A.W., C.M., B.R.J., M.W., L.R.M., E.C.R., A.Ra., and H.P. provided reagents and samples. L.E.M., P.C., E.N., and G.K. wrote the manuscript with contributions from K.W.N., N.F., G.H.C., C.C., and L.R.W. N.O., S.K., A.Ri., C.C., L.R.W., R.B., C.S., S.G., B.S., J.M., S.J.G., L.E.M., P.C., E.N., and G.K. supervised the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1339/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S18
Tables S1 and S2
References (*18–40*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

31 July 2020; accepted 29 October 2020
Published online 6 November 2020
10.1126/science.abe1107

EVOLUTION

The evolution of a tropical biodiversity hotspot

Michael G. Harvey[1,2,3]*†, Gustavo A. Bravo[4,5,6]†, Santiago Claramunt[7,8,9], Andrés M. Cuervo[10,11], Graham E. Derryberry[3,12], Jaqueline Battilana[6], Glenn F. Seeholzer[9,12], Jessica Shearer McKay[9], Brian C. O'Meara[3], Brant C. Faircloth[12,13], Scott V. Edwards[4,5], Jorge Pérez-Emán[14,15], Robert G. Moyle[16], Frederick H. Sheldon[12,13], Alexandre Aleixo[17,18], Brian Tilston Smith[9], R. Terry Chesser[19,20], Luís Fábio Silveira[6], Joel Cracraft[9], Robb T. Brumfield[12,13]‡, Elizabeth P. Derryberry[3,11]‡

The tropics are the source of most biodiversity yet inadequate sampling obscures answers to fundamental questions about how this diversity evolves. We leveraged samples assembled over decades of fieldwork to study diversification of the largest tropical bird radiation, the suboscine passerines. Our phylogeny, estimated using data from 2389 genomic regions in 1940 individuals of 1287 species, reveals that peak suboscine species diversity in the Neotropics is not associated with high recent speciation rates but rather with the gradual accumulation of species over time. Paradoxically, the highest speciation rates are in lineages from regions with low species diversity, which are generally cold, dry, unstable environments. Our results reveal a model in which species are forming faster in environmental extremes but have accumulated in moderate environments to form tropical biodiversity hotspots.

Most biological diversity originated in tropical regions (*1*, *2*), but long-standing questions remain about how tropical species diversity forms and is maintained (*3*). Are tropical peaks in diversity a result of faster rates of evolution or greater time for species accumulation (*4*)? When species do form in the tropics, what is the primary driver: climatic instability (*5*), the narrow stratification of habitats along elevation gradients (*6*–*8*), or divergence caused by interactions among species (*9*, *10*)? Alternatively, recent evidence suggests that speciation occurs most frequently when few other species are present (*11*). Might this explanation apply even in the species-rich tropics? Addressing these questions requires detailed investigation of the dynamics of speciation and species diversity through time, among lineages, and across geographic space in tropical groups. This work is particularly urgent because the same tropical regions that harbor and generate biotic diversity are under rapidly intensifying human pressure (*12*).

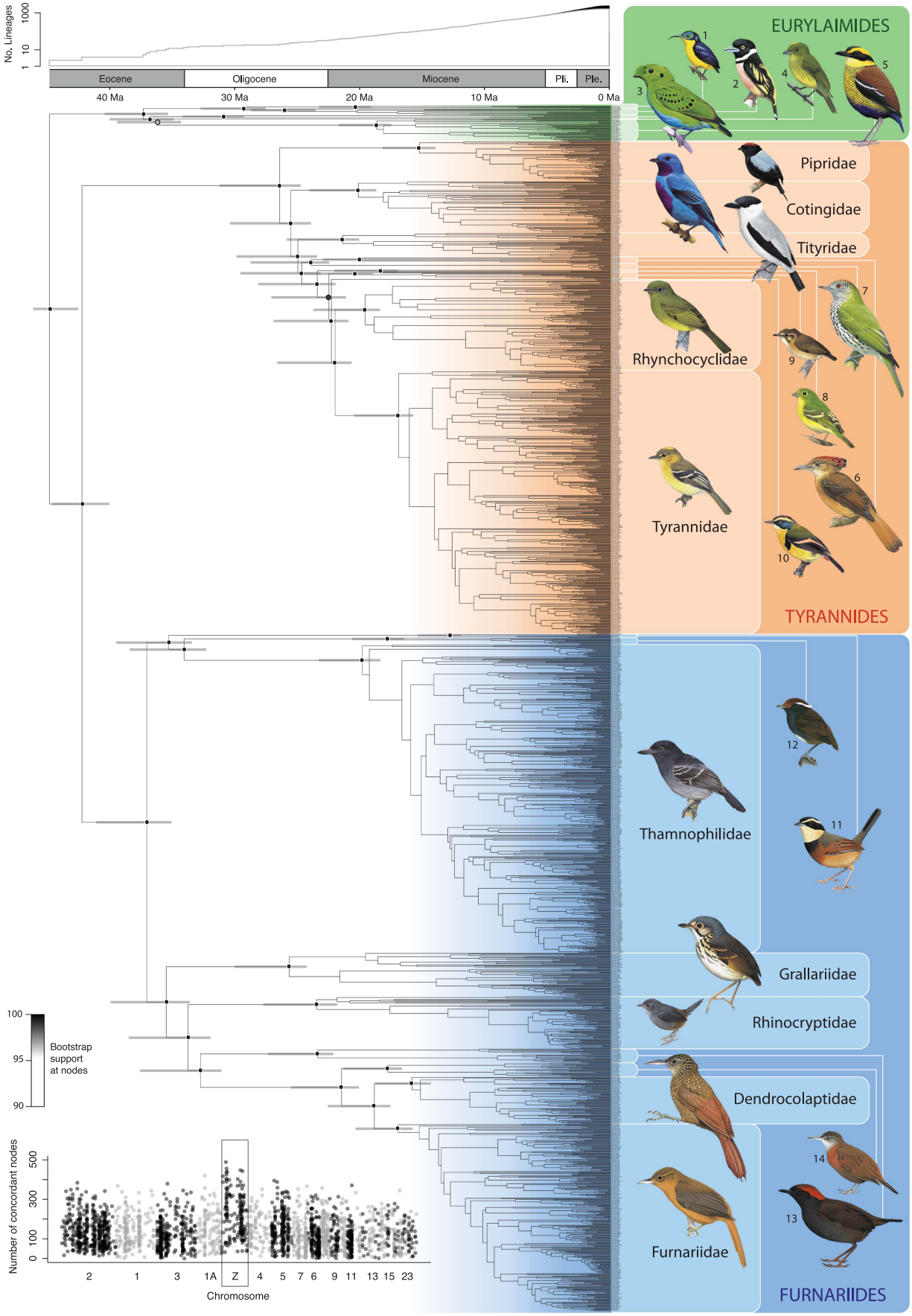
Through decades of expeditions and research in the tropics, our knowledge and sampling of tropical bird diversity are finally at a level of completeness sufficient for a detailed investigation of tropical diversification. Global avian diversity reaches its peak in the New World tropics, and a large portion of that diversity is attributable to a single massive radiation: the suboscine passerines (class Aves, suborder Tyranni). Roughly one in three Neotropical bird species (1250 of 4192 total) is a suboscine, even though the group is only 40 to 51 million years (Ma) old (*13*–*15*). Suboscines have varied phenotypes and ecologies and are the predominant avian group in essentially all terrestrial habitats in the Neotropics, from the Andean snow line to lowland Amazonia and from cloud forests to deserts (*16*). However, a comprehensive species-level estimate of suboscine phylogeny has not been available, precluding our understanding of diversification in this large tropical radiation and limiting identification of the broader drivers of tropical evolution.

We assembled a large, complete phylogenomic dataset [table S1 (*17*)] containing 1940 samples representing 1287 of 1306 suboscine species (98.5%) sequenced at 2389 orthologous genomic regions. Our maximum-likelihood phylogenetic tree (Fig. 1) was highly resolved. Relationships across the tree were supported by markers throughout the genome, with elevated support at sex-linked loci and high bootstrap support at most nodes. The tree topology and branch lengths were highly concordant between datasets with minimal and extensive filtering of missing data (fig. S1). A concatenated maximum-likelihood tree and a coalescent-based species tree accounting for heterogeneity in phylogenetic histories across the genome recovered highly similar relationships [figs. S2 and S3 (*17*)]. Time calibration of the phylogeny using existing fossil records within and outside suboscines indicates that the extant suboscine radiation began diversifying 44.5 Ma ago (range, 42.2 to 45.7 Ma ago) and individual suboscine families originated 18.7 to 36.5 Ma ago (stem ages; table S2). This time-calibrated tree provided the information needed for estimating the dynamics of suboscine diversification across time, lineages, and geography and for testing links between those dynamics and potential drivers of tropical diversity.

We found that suboscine diversification has been relatively steady since the origin of the group. Extant suboscine lineages have accumulated gradually, and our intraspecific sampling reveals continued accumulation through the present (Fig. 1). Diversification-through-time analyses did not find evidence for shifts in diversification rate over most of the history of the group [Fig. 2A (*17*)] aside from a drop within the past 2 Ma likely resulting from unsampled intraspecific diversity and unsorted ancestral polymorphism. Diversification models

Fig. 1. Species-level phylogeny of suboscine birds. The three infraorders are colored and the 24 families are outlined with white polygons. Smaller families are numbered as follows: (1) Philepittidae, (2) Eurylaimidae, (3) Calyptomenidae, (4) Sapayoidae, (5) Pittidae, (6) Onychorhynchidae, (7) Oxyruncidae, (8) Pipritidae, (9) Platyrinchidae, (10) Tachurisidae, (11) Melanopareiidae, (12) Conopophagidae, (13) Formicariidae, and (14) Scleruridae. Nodes that are at the crowns of families or older are annotated with circles that indicate bootstrap support (see gradient scale inset) and error bars depicting uncertainty in their ages. Above the phylogeny, a time axis and lineage-through-time (LTT) plots are depicted. The gray LTT plot depicts currently recognized species, and the black area near the top depicts the accumulation of additional lineages within currently recognized species. The inset in the bottom left is a Manhattan plot of gene tree support for the consensus topology at loci across the genome. Points are colored by chromosome (mapped to the zebra finch *Taeniopygia guttata*), alternating gray and black, and chromosomes are ordered by size (some smaller chromosomes are not labeled). Support is higher at sex-linked loci than at autosomal loci ($t = -10.3$, $P < 0.01$). Representative bird images for each family are illustrations reproduced by permission of Lynx Edicions.

[1]Department of Biological Sciences, The University of Texas at El Paso, El Paso, TX 79968, USA. [2]Biodiversity Collections, The University of Texas at El Paso, El Paso, TX 79968, USA. [3]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA. [4]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [5]Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA. [6]Museu de Zoologia da Universidade de São Paulo, 04263-000 Ipiranga, São Paulo, SP, Brazil. [7]Department of Natural History, Royal Ontario Museum, Toronto, Ontario M5S2C6, Canada. [8]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario M5S3B2, Canada. [9]Department of Ornithology, American Museum of Natural History, New York, NY 10024, USA. [10]Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá 111321, Colombia. [11]Department of Ecology and Evolutionary Biology, Tulane University, LA 70118, USA. [12]Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA. [13]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA. [14]Instituto de Zoología y Ecología Tropical, Universidad Central de Venezuela, Caracas, Venezuela. [15]Colección Ornitológica Phelps, Caracas, Venezuela. [16]Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA. [17]Finnish Museum of Natural History, University of Helsinki, 00014, Helsinki, Finland. [18]Department of Zoology, Museu Paraense Emílio Goeldi, CP 399, 66040-170 Belém, PA, Brazil. [19]US Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, USA. [20]National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA.
***Corresponding author. Email: mgh272@gmail.com**
†These authors contributed equally to this work.
‡These authors contributed equally to this work.


No. Lineages
1
10
1000
Eocene
Oligocene
Miocene
Pli.
Ple.
40 Ma
30 Ma
20 Ma
10 Ma
0 Ma
EURYLAIMIDES
1
2
3
4
5
Pipridae
Cotingidae
Tityridae
Rhynchocyclidae
7
9
8
Tyrannidae
6
10
TYRANNIDES
12
Thamnophilidae
11
Grallariidae
Rhinocryptidae
Dendrocolaptidae
Furnariidae
14
13
FURNARIIDES
100
95
90
Bootstrap support at nodes
Number of concordant nodes
0
100
300
500
2
1
3
1A
Z
4
5
7
6
9
11
13
15
23
Chromosome

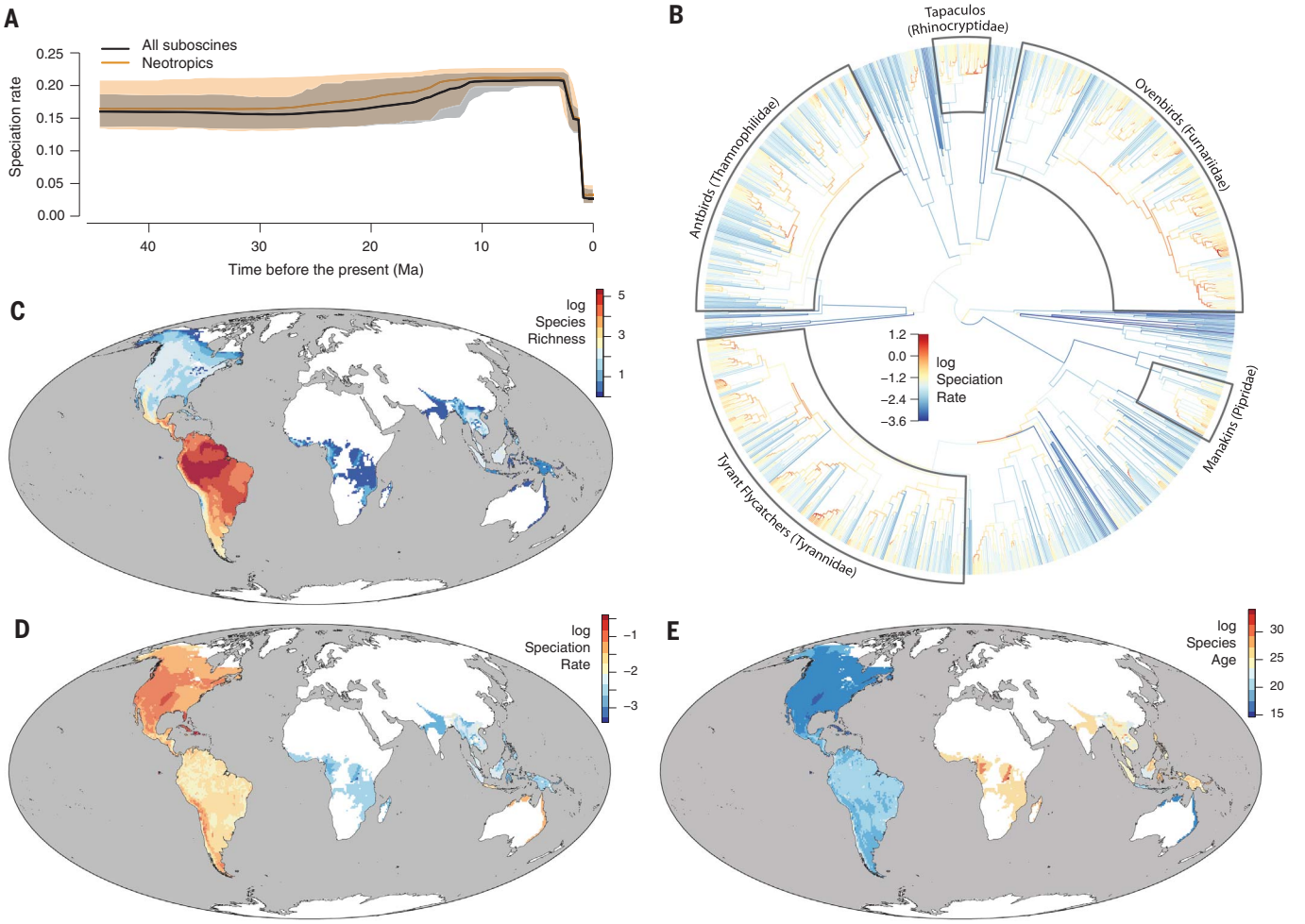


Fig. 2. Temporal, taxonomic, and geographic dynamics of suboscine diversification. (**A**) Suboscine speciation is relatively stable over most of the group's history, lacking an early burst. (**B**) Speciation rate varies widely among lineages, and five groups are consistently inferred to exhibit shifts to elevated speciation rates relative to background rates. Branches are colored according to the equal splits speciation rates estimator (ES). Groups outlined with polygons share an elevated speciation rate relative to the background regime in the best shift configuration from Bayesian analysis of macroevolutionary mixtures (BAMM) and are supported by other models with branch-specific diversification rates (see the supplementary text). The groups are labeled with the name of the most species-rich family contained therein. (**C**) Suboscine species richness is highest in the Neotropics. (**D**) Recent suboscine speciation rates (based on ES) do not peak in Neotropical centers of species richness. (**E**) Areas with high richness have moderate speciation rates and moderate diversity age, here measured using average node height. To avoid the excessive contribution of wide-ranging species to perceived spatial patterns, cell values in the maps of evolutionary statistics (D and E) represent averages weighted by the inverse of species range size.

fit to the suboscine phylogeny as a whole indicate that constant-rate models fit better than models with decreasing diversification rates through time [birth-death log likelihood = −3626.1; diversity-dependent log likelihood = −3658.7; change in corrected Akaike information criterion (ΔAICc) = 67.18]. Recent work has highlighted challenges associated with identifying alternative diversification histories in phylogenies of extant taxa (*18*), but even identifiable estimators do not show major shifts through time in suboscines (fig. S4). The steady dynamics in suboscines contrast with the pattern found in many other groups, in which an early burst of diversification is followed by a gradual slowdown (*19*). There are multiple potential causes for the absence of rate shifts (*20*, *21*). For example, tropical clades may have been less subject to bursts of speciation associated with episodic ecological opportunity, the suboscine radiation may be old enough that the signature of an early burst has eroded, or suboscines may represent an assemblage of smaller radiations with distinct dynamics that produce the overall trend [figs. S5 to S7 and tables S3 and S4 (*17*)].

Recent speciation rates varied >100-fold among different evolutionary lineages of suboscine birds (Fig. 2B). We focused on estimators of speciation rates that are most influenced by recent branching events and are thus robust to assumptions about extinction, but we also compared these with other model-based estimates (*17*). Five groups in particular were responsible for 77.4% of suboscine diversity and were consistently inferred to have experienced a shift to higher rates of diversification by a suite of methods [Fig. 2B (*17*)]. These include the lineages containing manakins (Pipridae), tyrant flycatchers (Tyrannidae), antbirds (Thamnophilidae), tapaculos (Rhinocryptidae), and ovenbirds (Furnariidae), groups that are highly distinct from one another in ecology and morphology (*16*). Future work, however, may reveal shared attributes that have contributed to elevated diversification across these groups.

When mapped onto geography, the Neotropical regions with the greatest suboscine species richness (Fig. 2C) do not contain lineages with the highest recent speciation

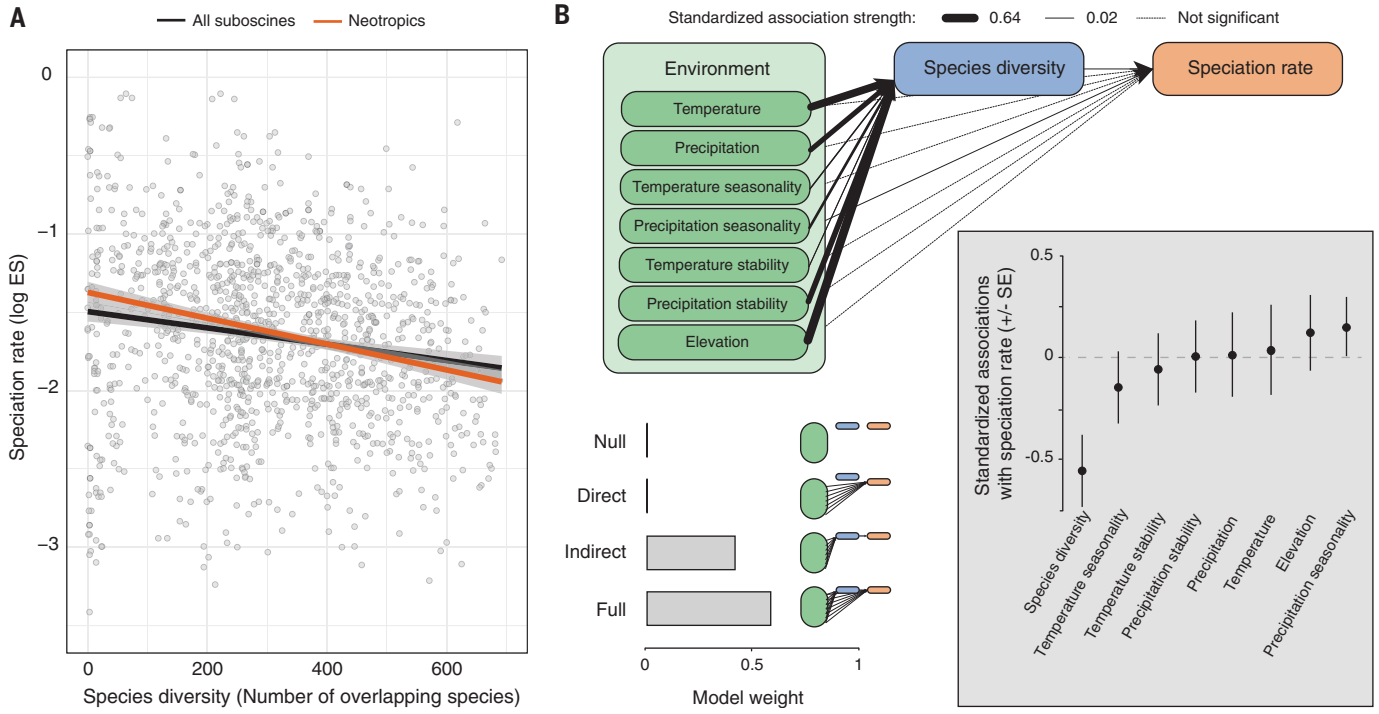


Fig. 3. Examination of the drivers of suboscine speciation rate variation reveals the role of low species diversity. (**A**) Plot of speciation rate (based on ES) versus species diversity for all suboscines showing that areas of high species diversity contain mostly lineages with low speciation rates. The negative speciation rate–diversity association is found both by examining all suboscines and by focusing on only those in the Neotropics. (**B**) The best-fitting model from phylogenetic path analysis includes strong associations between environmental variables and species diversity and then an impact of species diversity on speciation rate. Species diversity is the variable with the greatest impact (standardized association deviating from zero) on speciation rate (based on ES). The full model slightly outperforms a model in which the environment mediates speciation rate entirely through its impact on species diversity, and substantially outperforms a model in which environment impacts speciation rate directly.

rates (Fig. 2D) but instead exhibit moderate speciation rates and harbor species diversity that is older than in surrounding areas (Fig. 2E). This result is concordant with recent latitudinal studies of diversification in widespread samples of birds, mammals, and fishes that have found speciation rates in the tropics similar to or lower than those in the Temperate Zone (*22*–*24*). The spatial patterns in speciation rate and species age that we observed were consistent with the results of biogeographic modeling of the colonization history of suboscine lineages: Old World species were contained within a single, old clade of modest diversity, and Nearctic species derived from 15 to 20 recent (all <10.8 Ma ago) colonizations of the Temperate Zone [fig. S8 (*17*)]. Moreover, the transition of lineages into the Neotropics has only occurred four to nine times, which indicates that colonization is not a primary driver of elevated Neotropical diversity. Although the potential evolutionary importance of time has long been recognized (*25*, *26*), its contribution in suboscines further establishes a key role in the accumulation of tropical species. Species-rich regions in the tropics are not merely a locus of young diversity but contain diversity accumulated over a protracted evolutionary period.

We failed to find strong associations between speciation rates in the New World and environmental variables (table S5). Speciation rate was not correlated with elevation (ρ = 0.187; P = 0.100), nor did species from mountainous regions exhibit higher speciation rates than those from lowland regions (λ_0 = 0.176 species/Ma; λ_1 = 0.200; P = 0.933). This result is notable in light of evidence of positive associations between speciation rates and elevation in prior studies (*7*, *8*, *27*). Elevation may not be a strong and general predictor of speciation rate variation. We also found few strong associations between speciation rate and current climatic variables such as temperature, precipitation, seasonality, or average rates of climatic change since the last glacial maximum or across five time slices since the Pliocene (table S5).

A key result from our study is that the best predictor of elevated speciation rates in New World suboscines is low current standing species diversity (Fig. 3). Higher speciation rates were observed in lineages with ranges that overlapped few other suboscines (ρ = −0.247; P = 0.007; Fig. 3A) or few overall bird species (ρ = −0.223; P = 0.021). This result was robust to phylogenetic uncertainty, incomplete sampling, simulated random and nonrandom extinction, alternative taxonomic classifications, and the effects of spatial autocorrelation in species richness estimates [table S6 (*17*)]. Moreover, this result was also evident using state-dependent speciation and extinction models and traditional sister clade comparisons [table S7 and figs. S9 and S10 (*17*)]. The negative speciation–species richness association runs counter to the idea that greater potential for species interactions promotes speciation in the tropics (*8*). It instead supports recent arguments that the geographic dynamics of species formation at a given time are a response to rather than a cause of broad species diversity patterns (*11*, *28*). In this framework, new species form at higher rates in areas containing relatively few species or where there is a diversity debt relative to the environmental capacity for species.

A model including variable extinction rates suggested that areas of low diversity have experienced historically high levels of extinction (table S7), which may help explain the low levels of standing diversity in those areas today. This supports prior work suggesting that extinction may overwhelm speciation as a driver of patterns in species richness (*22*, *29*). In suboscines, areas of low diversity were associated with environments characterized by

low temperatures and precipitation, greater seasonality in temperature and precipitation, and larger changes in climate through geologic time (table S5). These may be areas in which the contemporary habitat is younger and lineages are recent colonists from elsewhere. Areas of low species richness and high speciation rates included parts of the central and southern Andes and Patagonia, the Caribbean, and the Nearctic (Fig. 2, C and D). These are also areas found to result in low simulated species persistence under recent (i.e., the past 800,000 years) paleoclimatic models (*30*). We used phylogenetic path analysis to simultaneously evaluate causal connections among different environmental and biotic variables. The best model contained causal links between climatic variables and species richness and then a link between species richness and speciation rates (Fig. 3B). This result indicates that environmental variables mediate the number of species in an area, which in turn drives speciation rate variation.

By characterizing the dynamics of diversity through time, among lineages, and across space in a major tropical group, we were able to address long-standing questions about tropical evolution. We found that diversity in a tropical group has accumulated over a protracted period, and that the hotpots of species diversity in the tropics are associated with time for species accumulation rather than exceptional speciation rates in those areas. We provide evidence that the environment mediates speciation dynamics through this filter of species richness. Overall, extreme environments appear to limit diversity but provide enhanced opportunities for speciation, whereas moderate environments reduce species formation within lineages but permit diversity to accumulate.

REFERENCES AND NOTES

1. D. Jablonski, K. Roy, J. W. Valentine, *Science* **314**, 102–106 (2006).
2. J. Rolland, F. L. Condamine, F. Jiguet, H. Morlon, *PLOS Biol.* **12**, e1001775 (2014).
3. A. Antonelli *et al*., *PeerJ* **6**, e5644 (2018).
4. G. L. Stebbins, *Flowering Plants: Evolution Above the Species Level* (Belknap, 1974).
5. J. Haffer, *Science* **165**, 131–137 (1969).
6. D. H. Janzen, *Am. Nat.* **101**, 233–249 (1967).
7. J. T. Weir, *Evolution* **60**, 842–855 (2006).
8. I. Quintero, W. Jetz, *Nature* **555**, 246–250 (2018).
9. D. W. Schemske, in *Speciation and Pattterns of Diversity*, R. Butlin, J. Bridle, D. Schluter, Eds. (Cambridge Univ. Press, 2009), pp. 219–239.
10. J. H. Brown, *J. Biogeogr.* **41**, 8–22 (2014).
11. D. Schluter, M. W. Pennell, *Nature* **546**, 48–55 (2017).
12. M. C. Hansen *et al*., *Science* **342**, 850–853 (2013).
13. S. Claramunt, J. Cracraft, *Sci. Adv.* **1**, e1501005 (2015).
14. R. O. Prum *et al*., *Nature* **526**, 569–573 (2015).
15. C. H. Oliveros *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7916–7925 (2019).
16. D. F. Stotz, J. W. Fitzpatrick, T. A. Parker III, D. Moskovits, *Neotropical Birds: Ecology and Conservation* (Univ. of Chicago Press, 1996).
17. Materials and methods are available as supplementary materials.
18. S. Louca, M. W. Pennell, *Nature* **580**, 502–505 (2020).
19. D. L. Rabosky, *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502 (2013).
20. J. J. Wiens, *Q. Rev. Biol.* **86**, 75–96 (2011).
21. D. L. Rabosky, A. H. Hurlbert, *Am. Nat.* **185**, 572–583 (2015).
22. J. T. Weir, D. Schluter, *Science* **315**, 1574–1576 (2007).
23. D. L. Rabosky *et al*., *Nature* **559**, 392–395 (2018).
24. J. D. Kennedy *et al*., *J. Biogeogr.* **41**, 1746–1757 (2014).
25. J. C. Willis, *Age and Area: A Study in Geographical Distribution and Origin in Species* (Cambridge Univ. Press, 1922).
26. P. V. A. Fine, R. H. Ree, *Am. Nat.* **168**, 796–804 (2006).
27. N. R. Polato *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12471–12476 (2018).
28. J. T. Weir, T. D. Price, *Am. Nat.* **177**, 462–469 (2011).
29. A. S. Meseguer, F. L. Condamine, *Evolution* (2020).
30. T. F. Rangel *et al*., *Science* **361**, eaar5452 (2018).
31. M. G. Harvey *et al*., Protocols, code, trees, and alignments for: The evolution of a tropical biodiversity hotspot, Zenodo (2020); https://doi.org/10.5281/zenodo.3976115.

ACKNOWLEDGMENTS

We thank the many field workers and museum staff and administrators who made this work possible, particularly those at the Field Museum of Natural History, the Instituto Alexander von Humboldt, the Instituto de Ciencias Naturales (Universidad Nacional de Colombia), the Instituto Nacional de Pesquisas da Amazônia, the Colección Ornitológica Phelps, the Instituto de Zoología y Ecología de la Universidad Central de Venezuela, the Museu de Zoologia da Universidade de São Paulo, the Museum of Southwestern Biology (University of New Mexico), the Australian National Wildlife Collection, the Peabody Museum of Natural History (Yale University), the Burke Museum (University of Washington), the Cornell University Museum of Vertebrates, the Academy of Natural Sciences of Philadelphia (Drexel University), the Museum of Vertebrate Zoology (UC Berkeley), the Museo de la Universidad de Costa Rica, the Laboratório de Genética e Evolução Molecular de Aves (Universidade de São Paulo), the Museum of Comparative Zoology (Harvard University), the University of Michigan Museum of Zoology, the Museu de Ciências e Tecnologia da Pontifícia Universidade Católica do Rio Grande do Sul, the Museu Paraense Emílio Goeldi, and the Museum of Natural Science (Louisiana State University). Fieldwork support by the Ornithology Section at MZUSP, especially by V. Piacentini and F. Schunck, was important to improving taxon sampling in Brazil. D. Lane, B. Whitney, J. V. Remsen Jr., S. S. Taylor, J. E. McCormack, C. H. Oliveros, S. P. Galeano, C. Sánchez, J. S. Schenck, M. Bornschein, R. Belmonte, M. Maldonado, the Derryberry and Brumfield laboratory members, the Ornithology Section at MZUSP, and the LSUMNS vertebrate group provided additional advice with study design and assistance with sampling. F. Raposo do Amaral, G. Thom, V. Piacentini, J. Weckstein, L. Musher, G. Del-Rio, and C. Miyaki generously provided sequence data for 51 individuals. R. Beco, S. Bolívar, F. Bocalini, L. Neves, and Rapid Genomics provided support with laboratory work. The Louisiana State University High Performance Computing Center, the University of Michigan Advanced Research Computing Technology Service, and the Harvard University FAS Research Computing Group provided computational support. D. L. Rabosky and his laboratory group, S. Mirarab, S. M. Smith, J. W. Brown, N. Upham, A. Cloutier, W. Jetz, and S. Orzechowski, provided advice and assistance with analyses. J. T. Weir and three anonymous reviewers provided helpful comments on an earlier version of this manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. **Funding:** This work was supported by U.S. National Science Foundation grants DEB-1146265 (to R.T.B.), DEB-1146423 (to E.P.D.), DEB-1146248 (to J.C.), DEB-1146345 (to R.G.M.), DEB-1011435 (to G.A.B and R.T.B), IOS-1210556 (to M.G.H. and R.T.B.), DBI-1523893 (to M.G.H.), DEB-1655624 (to B.C.F. and R.T.B.), and DEB-1241066 (to J.C.); Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2018-06747 (to S.C.); São Paulo Research Foundation–FAPESP grants 2012-23852-0 (to G.A.B) 56378-0 and 2017-23548-2 (to L.F.S.); and Brazilian Research Council (CNPq) grants 457491/2012-4 and 302291/2015-6 (to L.F.S); 457974-2014-1 (to G.A.B and L.F.S), and 306843/2016-1, 574008/2008-0, 563236/2010-8 and 471342/ 2011-4 (to A.A). **Author contributions:** E.P.D., R.T.B., J.C., L.F.S., R.T.C., A.A., F.H.S., R.G.M., and J.P.E. conceived of the study. M.G.H., G.A.B., S.C., A.M.C., G.E.D., J.B., G.F.S., J.S.M., B.C.F., and B.T.S. developed the methods and collected the data. M.G.H., G.A.B., S.C., G.E.D., and B.C.O. conducted the analyses. M.G.H., G.A.B., S.C., R.T.B., and E.P.D. wrote the paper with the assistance of S.V.E., B.C.O., B.C.F., B.T.S., J.P.E., and L.F.S. The manuscript reflects the contributions and ideas of all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw sequence data are in the NCBI Sequence Read Archive (PRJNA655842). Protocols, code, trees, and alignments are available on Zenodo (*31*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1343/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S72
Tables S1 to S47
References (*32–180*)
MDAR Reproducibility Checklist

1 October 2019; resubmitted 1 May 2020
Accepted 29 October 2020
10.1126/science.aaz6970

PALEOCLIMATE

Southern Ocean upwelling, Earth's obliquity, and glacial-interglacial atmospheric CO_2 change

Xuyuan E. Ai[1,2]*, Anja S. Studer[3], Daniel M. Sigman[1], Alfredo Martínez-García[2], François Fripiat[4], Lena M. Thöle[5,6], Elisabeth Michel[7], Julia Gottschalk[8]†, Laura Arnold[9], Simone Moretti[2,9], Mareike Schmitt[2], Sergey Oleynik[1], Samuel L. Jaccard[5]‡, Gerald H. Haug[2,9]

Previous studies have suggested that during the late Pleistocene ice ages, surface-deep exchange was somehow weakened in the Southern Ocean's Antarctic Zone, which reduced the leakage of deeply sequestered carbon dioxide and thus contributed to the lower atmospheric carbon dioxide levels of the ice ages. Here, high-resolution diatom-bound nitrogen isotope measurements from the Indian sector of the Antarctic Zone reveal three modes of change in Southern Westerly Wind–driven upwelling, each affecting atmospheric carbon dioxide. Two modes, related to global climate and the bipolar seesaw, have been proposed previously. The third mode—which arises from the meridional temperature gradient as affected by Earth's obliquity (axial tilt)—can explain the lag of atmospheric carbon dioxide behind climate during glacial inception and deglaciation. This obliquity-induced lag, in turn, makes carbon dioxide a delayed climate amplifier in the late Pleistocene glacial cycles.

The past ~1 million years of Earth history are characterized by large glacial-interglacial cycles with an average period of 100,000 years. The cycles originate from subtle variations in Earth's orbital parameters, which affect the seasonal and spatial distribution of solar insolation on Earth's surface (*1*). However, the direct effects of the orbital changes on global, annually averaged insolation are minor. Thus, feedbacks within Earth's climate system are thought to rectify and amplify the insolation changes into large, global climate changes. The concentration of the atmospheric greenhouse gas CO_2 varied by ~30% over the glacial-interglacial cycles (*2*). Accordingly, it is suspected that CO_2 was a key amplifier in the generation of these climate cycles; however, the causes of the CO_2 variation remain widely debated.

The ocean's biological carbon pump—the production of organic carbon by phytoplankton in the surface ocean and its sinking and remineralization—sequesters CO_2 in the deep ocean, which lowers the concentration of CO_2 in the atmosphere. The phytoplankton that underpin the biological pump require nutrients, namely nitrate and phosphate, which are supplied from the ocean interior and are returned to it in the sinking organic matter. In the Southern Ocean, where phytoplankton growth is limited by iron and/or light (*3*, *4*), the deeply stored nutrients and CO_2 brought to the surface by upwelling and surface-deep exchange are not fully consumed. This results in high concentrations of residual nutrients and the outgassing of CO_2 from the ocean to the atmosphere—i.e., a leak in the global ocean's biological pump (*5*). Different mechanisms have been proposed by which the Southern Ocean leak in the biological pump may have been stemmed during glacial intervals. First, increased sea ice cover could have lowered atmospheric CO_2 by preventing the degassing of CO_2 from underlying surface waters (*6*). Second, iron fertilization could have driven more-complete consumption of upwelled carbon and nutrients (*7*). Third, a slowing of deep-water ventilation in the Southern Ocean could have, in its own right, reduced the region's release of biological pump–sequestered CO_2 to the atmosphere while also encouraging more-complete nutrient consumption, which would have further decreased the CO_2 release (*8*).

During the ice ages, higher productivity coincided with high dust supply in the Sub-Antarctic Zone (SAZ) of the Southern Ocean (*9*–*11*), which together offer support for the Southern Ocean iron fertilization hypothesis (*7*). However, decreased productivity is found in the glacial Antarctic Zone (AZ), the more polar domain of the Southern Ocean (*10*–*12*). Information on surface ocean nutrient conditions is needed to determine the causes of these productivity changes and their implications for atmospheric CO_2.

When phytoplankton grow, they preferentially consume ^{14}N-bearing nitrate ($^{14}NO_3^-$) relative to ^{15}N-nitrate ($^{15}NO_3^-$), such that the consumption of nitrate leads to the progressive elevation of nitrate $\delta^{15}N$ {$\delta^{15}N$ [‰ (per mil)] = $[(^{15}N/^{14}N)_{sample}/(^{15}N/^{14}N)_{atmN2} - 1] \times 1000$}, which is in turn reflected in the $\delta^{15}N$ of plankton biomass (*13*). In the SAZ during ice ages, an increase in foraminifera-bound $\delta^{15}N$ ($\delta^{15}N_{fb}$) indicates more-complete nitrate consumption; its coincidence with the elevated dust supply and greater biological export production strengthens the evidence for iron fertilization in the SAZ (*14*). Records of bulk sediment $\delta^{15}N$ and diatom-bound $\delta^{15}N$ ($\delta^{15}N_{db}$) suggest that surface nitrate was also more fully consumed in the AZ during glacial periods (*15*–*17*). More-complete nitrate consumption, in the context of lower biogenic fluxes, points to a decline in gross nitrate supply to the surface waters of the glacial AZ (*17*). However, the cause of this change is still under discussion.

From the Indian sector of the AZ, we report two $\delta^{15}N_{db}$ records with (sub)millennial resolution over the last glacial cycle. MD12-3394 is near the northern edge of the AZ, and MD11-3353 is slightly poleward and closer to the Kerguelen Islands (Fig. 1). Chronology is on the basis of the alignment of organic biomarker (TEX_{86}^L)-based sea surface temperature (SST) to the Antarctic temperature stack (ATS) (*18*) or the European Project for Ice Coring in Antarctica (EPICA) Dome C temperature (*19*), with seven ^{14}C ages during the Holocene. The marine isotope stage (MIS) 3 interval of MD11-3353 is tuned by two additional $\delta^{15}N_{db}$ tie-points to MD12-3394 (see supplementary materials). The reliability of these chronologies is supported by comparison of planktonic $\delta^{18}O$ to the ATS (fig. S3).

During the last ice age, in both MD12-3394 and MD11-3353, $\delta^{15}N_{db}$ increased progressively from the penultimate interglacial to the Last Glacial Maximum (LGM), reaching values that are 3 to 4‰ higher than those of the interglacial (Fig. 1C). The records agree well with the previously published $\delta^{15}N_{db}$ record of PS75/72-4 from the Pacific sector of the AZ (*15*) (Fig. 1C) and with deep sea coral-bound $\delta^{15}N$ from the Drake Passage (*20*). Together, the datasets indicate a high degree of nitrate consumption [nitrate concentrations lower than present (*21*)] across a vast expanse of the glacial AZ.

The $\delta^{15}N_{db}$ records reveal three mechanistic modes of change in AZ nitrate supply, which point to changes in the Southern Westerly Winds (SWW) as their driver, as described below. In turn, the AZ changes can explain key features in the ice core record of atmospheric CO_2 and thus clarify the role of atmospheric CO_2 in glacial-interglacial cycles. We introduce each mode in terms of greater nitrate supply and thus lower AZ $\delta^{15}N_{db}$ (which is plotted upward in all figures).

Mode 1 is driven by global mean climate. The maxima in global warmth of MIS 5e, MIS 5c, MIS 5a, and MIS 1 (*22*, *23*) (Fig. 2C, pink bars)—which occur during 65°N summer insolation

[1]Department of Geosciences, Princeton University, Princeton, NJ 08544, USA. [2]Climate Geochemistry Department, Max Planck Institute for Chemistry, Mainz 55128, Germany. [3]Department of Environmental Sciences, University of Basel, Basel, Switzerland. [4]Department of Geosciences, Environment and Society, Université Libre de Bruxelles, Brussels, Belgium. [5]Institute of Geological Sciences and Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland. [6]Department of Earth Sciences, Utrecht University, Utrecht, Netherlands. [7]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Laboratoire CNRS-CEA-UVSQ, Gif-sur-Yvette, France. [8]Lamont-Doherty Earth Observatory, Palisades, NY 10964, USA. [9]Department of Earth Sciences, ETH Zürich, Zürich, Switzerland.

*Corresponding author. Email: xuyuana@princeton.edu

†Present address: Institute of Geosciences, Kiel University, Kiel 24118, Germany.

‡Present address: Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland.

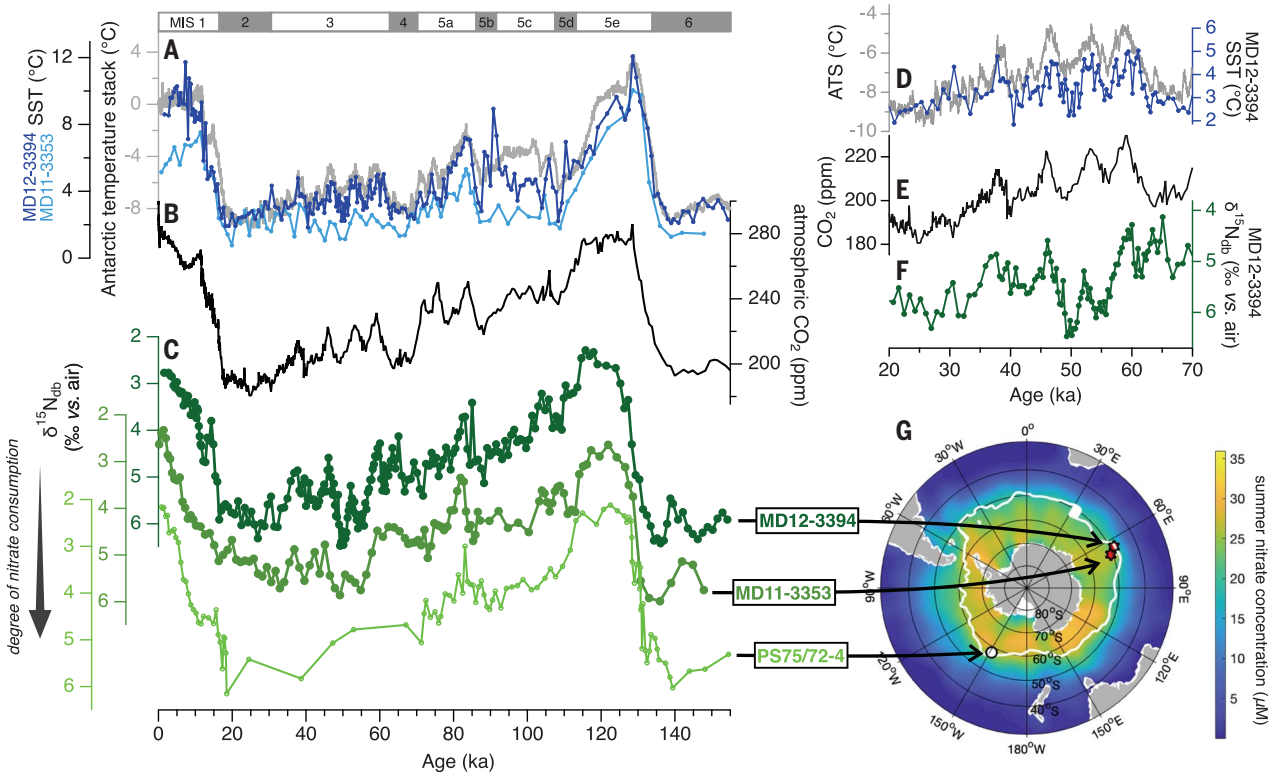


Fig. 1. High-resolution Antarctic diatom-bound $\delta^{15}N$ records through the last glacial cycle. (**A**) TEX^{L}_{86}–based SST of MD12-3394 (dark blue) and MD11-3353 (light blue) plotted over the ATS from ice cores (*18*) (gray). (**B**) Atmospheric CO_2 (*47*). (**C**) $\delta^{15}N_{db}$ of MD12-3394 (dark green), MD11-3353 (grass green), and PS75/72-4 (*15*) (light green). The age models (see text and supplementary materials) are based on correlation to Antarctic ice core temperature and are thus not fully independent. (**D**, **E**, and **F**) Expansion of MD12-3394 SST (dark blue), ATS (gray), CO_2 (black), and MD12-3394 $\delta^{15}N_{db}$ (dark green) from 20,000 to 70,000 years ago. (**G**) Map of austral summer surface nitrate concentration (*48*) and the Antarctic Polar Front (*49*) (white line) with core sites indicated. ka, thousand years ago. The marine isotope stages are inferred from the ATS.

maxima (Fig. 2A) (*1*)—correspond to low $\delta^{15}N_{db}$ and thus high AZ nitrate supply (Fig. 2D). For the northern part of the AZ, the region of our $\delta^{15}N_{db}$ records, the import of deep nitrate into the surface layer is associated with the upper cell of Southern Ocean overturning: the upwelling and northward transport of water at the surface caused by the SWW. In climate models, the SWW can respond to global warming with a poleward shift, and this may be associated with an increase in wind strength (*24*). Both changes lead to stronger wind stress at the latitudes of the Drake Passage, which in turn strengthens wind-driven northward surface transport and upwelling in the AZ (*25*). The wind-driven northward transport in the AZ is counteracted by southward eddy advection, such that the net flow of the upper cell is consistent with air-sea buoyancy fluxes. Nevertheless, the eddy compensation is partial (*26*), with models indicating that a change in the SWW-driven northward transport would cause a change in net AZ upwelling of the same sign (*27*). For the northern part of the AZ, where our $\delta^{15}N_{db}$ records were recovered, the observed amplitude of $\delta^{15}N_{db}$ elevation during the glacial maxima requires a >75% reduction in nitrate supply (*21*). Such a large change favors the interpretation that a weakening and/or equatorward displacement of the SWW caused upwelling to decline in the glacial AZ (*28*).

Mode 2 is driven by the bipolar seesaw. During MIS 3, $\delta^{15}N_{db}$ in MD12-3394 covaries with the millennial-scale Antarctic isotope maximum (AIM) events (Fig. 1, D to F, and Fig. 2, blue bars). Lower $\delta^{15}N_{db}$ is associated with higher Antarctic air temperature, which suggests that during the associated Heinrich events, while the Southern Ocean warms (*29*, *30*), AZ surface nitrate supply rises. Our data are consistent with the interpretation that a reduction of North Atlantic overturning during Heinrich events caused a southward shift and strengthening of the SWW (*31*), which lead to increased AZ upwelling (*32*).

For AZ upwelling to become stronger, both modes described above involve the warming of the Southern Ocean, which should be reflected in Antarctica. If the combination of the two modes is sufficient to explain the changes in the strength and position of the SWW and thus the changes in upwelling strength, then nutrient conditions in the AZ should be closely coupled with Antarctic ice core temperatures. We calculate the predicted $\delta^{15}N_{db}$ under the assumption that $\delta^{15}N_{db}$ varies linearly with the ATS (Fig. 2D) [(*18*); see materials and methods]. The difference in the measured $\delta^{15}N_{db}$ from this predicted $\delta^{15}N_{db}$ record is quantified as the $\delta^{15}N_{db}$ offset (Fig. 2G). The resulting $\delta^{15}N_{db}$ offset record leads us to propose a third mode of $\delta^{15}N_{db}$ change driven by changes in obliquity, as explained next.

Independent of the absolute temperature at high latitudes, ocean-only climate models indicate that a steeper temperature gradient between the middle and high latitudes results in a strengthening of the SWW (*24*). Low obliquity increases the insolation gradient, leading to a stronger temperature gradient between the middle and high latitudes (*33*, *34*), which appears to be reflected in the temperature difference between the moisture source region and ice core site as estimated from deuterium excess in the Antarctic ice cores (*35*, *36*) (Fig. 2F). In turn, a stronger temperature gradient yields stronger SWW, which enhance AZ upwelling (*34*). Consistent with the operation of a third, obliquity-driven mode of AZ upwelling change, the $\delta^{15}N_{db}$ offset covaries with obliquity and with the temperature difference between moisture source and site in Antarctic ice cores: Negative $\delta^{15}N_{db}$ offsets (higher nitrate supply than predicted) are associated with obliquity minima and larger temperature differences (Fig. 2, E to G, and fig. S6).

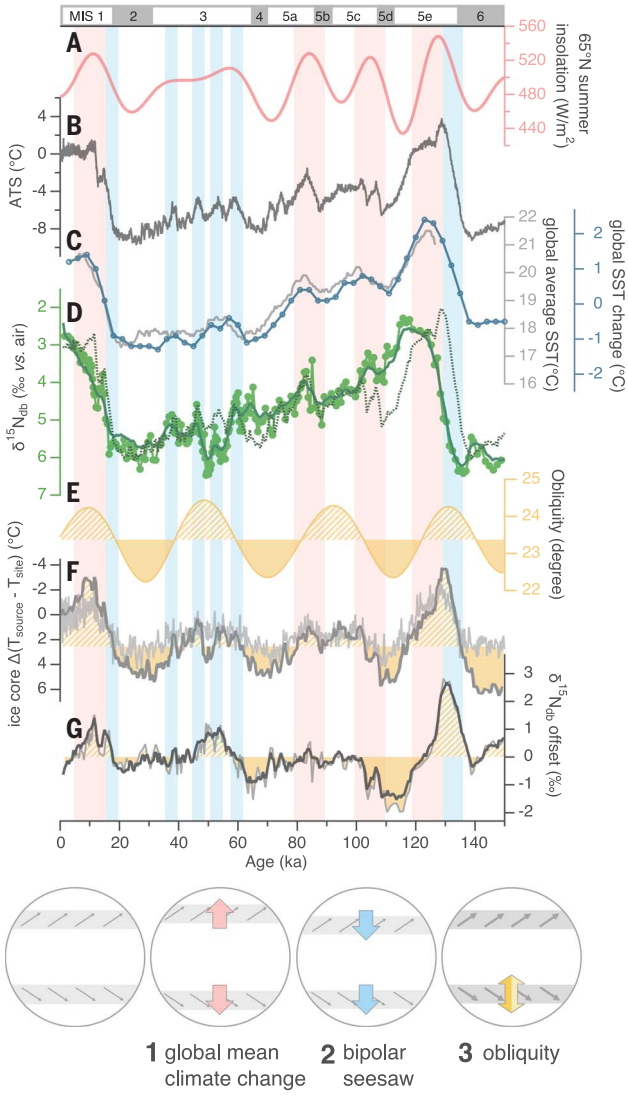


Fig. 2. Evidence for three modes of change in SWW-driven upwelling in the AZ. (**A**) Daily mean top-of-atmosphere insolation at 65°N summer solstice (*50*). (**B**) ATS (*18*). (**C**) Two reconstructions of global SST (*22*, *23*). (**D**) $\delta^{15}N_{db}$ from MD12-3394 (light green circles), the mean of $\delta^{15}N_{db}$ from MD12-3394 and MD11-3353 calculated using a Monte Carlo and Kalman filter method (*20*) with a resolution of 500 years (dark green line; fig. S1), and the predicted $\delta^{15}N_{db}$ if it strictly follows ATS (dotted line; see materials and methods). (**E**) Obliquity, with solid or hatched orange shading indicating lower or higher than the average of the past 150,000 years, respectively (*51*). (**F**) Changes in the difference between the air temperature at the moisture source (T_{source}) and the ice core site (T_{site}) of Vostok (*35*) (dark gray) and Dome Fuji (*36*) (light gray) relative to modern conditions, reconstructed from ice deuterium excess. Solid or hatched orange shading indicates higher or lower $T_{source} - T_{site}$ than the average of the past 150,000 years at Vostok, respectively. (**G**) $\delta^{15}N_{db}$ offset (measured minus predicted $\delta^{15}N_{db}$; see materials and methods) calculated using the $\delta^{15}N_{db}$ of MD12-3394 (gray line) and the mean $\delta^{15}N_{db}$ from the two records (black line). Solid or hatched orange shading indicates a negative or positive $\delta^{15}N_{db}$ offset, respectively, which implies greater than or less than predicted AZ upwelling, respectively. Illustration at the bottom shows the Westerly Wind belts and their changes under the three mechanistic modes. For modes 1 (pink) and 2 (blue), colors correspond to the shaded intervals in the plot, all shown for an increase in AZ upwelling (i.e., a decrease in $\delta^{15}N_{db}$). For mode 3, in the illustration and (G), solid orange or hatching, respectively, symbolizes stronger or weaker upwelling due to stronger or weaker meridional temperature gradient and thus stronger or weaker SWW.

The strongest negative $\delta^{15}N_{db}$ offset occurs at the end of the penultimate interglacial through the glacial inception at the MIS 5e–MIS 5d transition. From 125,000 to 115,000 years ago, global average SST declines by ~1°C (Fig. 2C) and extratropical SST declines by ~1.5°C (*37*), which should lead to less AZ upwelling by mode 1. In the high latitudes, the average SSTs of both 50 to 60°S and 50 to 60°N drop by 2 to 2.5°C (*22*); thus, there is no sign of a hemispheric asymmetry that might shift the SWW poleward and maintain AZ upwelling by mode 2. Yet, $\delta^{15}N_{db}$ remains at low interglacial values or even decreases slightly from 125,000 to 115,000 years ago, which indicates sustained AZ upwelling. Similarly, in MIS 5d, both Antarctic ice core temperatures and North Greenland Ice Core Project (NGRIP) temperature as well as AZ SST drop to values comparable to those observed in the LGM (*38*, *39*), and reconstructed global average SST shows an ~50% cooling toward the LGM level (*22*, *23*) (Fig. 2, B and C). However, $\delta^{15}N_{db}$ only increases ~30 to 35% toward its full glacial value (Fig. 2D). The maintenance of low $\delta^{15}N_{db}$ from 125,000 to 115,000 years ago and the relatively small $\delta^{15}N_{db}$ increase across MIS 5d (Fig. 2, D and G) suggest that SWW-driven upwelling during these intervals is stronger than the scenario in which AZ upwelling declines in lockstep with Antarctic cooling.

This stronger-than-expected upwelling can be explained by mode 3. Low obliquity (Fig. 2E), which steepens the meridional temperature gradient (Fig. 2F) and strengthens the SWW, appears to have counteracted the effect of global cooling on SWW-driven upwelling. Similar to the end of the penultimate interglacial, the current interglacial (the Holocene) is characterized by declining obliquity (Fig. 2E). Largely because of this trend, climate models simulate a SWW-driven increase in upwelling since 7000 years ago (*40*). Consistent with this, $\delta^{15}N_{db}$ and the $\delta^{15}N_{db}$ offset decline over the Holocene (*41*) (Fig. 1C and Fig. 2G), which affirms the proposed role of obliquity (i.e., mode 3) in Southern Ocean upwelling.

Due to mode 3, high obliquity often corresponds to intervals of positive $\delta^{15}N_{db}$ offset, especially during deglaciations. The deglacial warming of the Antarctic region due to modes 1 and 2 tends to cause the SWW to shift poleward and strengthen in the AZ. However, at the same time, high obliquity (mode 3) works against modes 1 and 2, weakening the SWW and depressing upwelling in the AZ, which reduces the decline in $\delta^{15}N_{db}$ and thus increases the $\delta^{15}N_{db}$ offset (Fig. 2, D and G), while it decreases the temperature difference between moisture source and ice core site (Fig. 2, E and F). Modes 1 and 2 appear to dominate the change in the SWW and thus drive stronger AZ upwelling during the deglaciations. Nevertheless, the counteracting effect of mode 3 is reflected in the late decline in $\delta^{15}N_{db}$ at termination 2 and the weak decline in $\delta^{15}N_{db}$ compared with the rise in Antarctic air temperature at termination 1 (Fig. 2D and fig. S8). Finally, other intervals during which changes in $\delta^{15}N_{db}$ deviate from those in Antarctic temperature—such as MIS 4 with its relatively low $\delta^{15}N_{db}$ and the interval of relatively high $\delta^{15}N_{db}$ at 48,000 to 57,000 years ago during MIS 3—can also be explained by mode 3 (Fig. 2, E to G, and the supplementary materials).

The AZ data, in combination with proxies from the SAZ, show that changes in Southern Ocean surface nutrient conditions can explain much of the ice age atmospheric CO_2 drawdown, with the AZ and SAZ having complementary correspondences with the CO_2 declines of the last glacial progression. During MIS 5, CO_2 is more closely correlated with AZ $\delta^{15}N_{db}$ than with Antarctic air temperature (Fig. 3, A, B, and I, and fig. S4). This is particularly apparent for the glacial inception (108,000 to 124,000 years ago; Fig. 3, F to I). The similar timing and proportionality of change in $\delta^{15}N_{db}$ and CO_2 at the MIS 5e–MIS 5d transition suggest that weaker upwelling and reduced CO_2 venting in the AZ play an important role in the

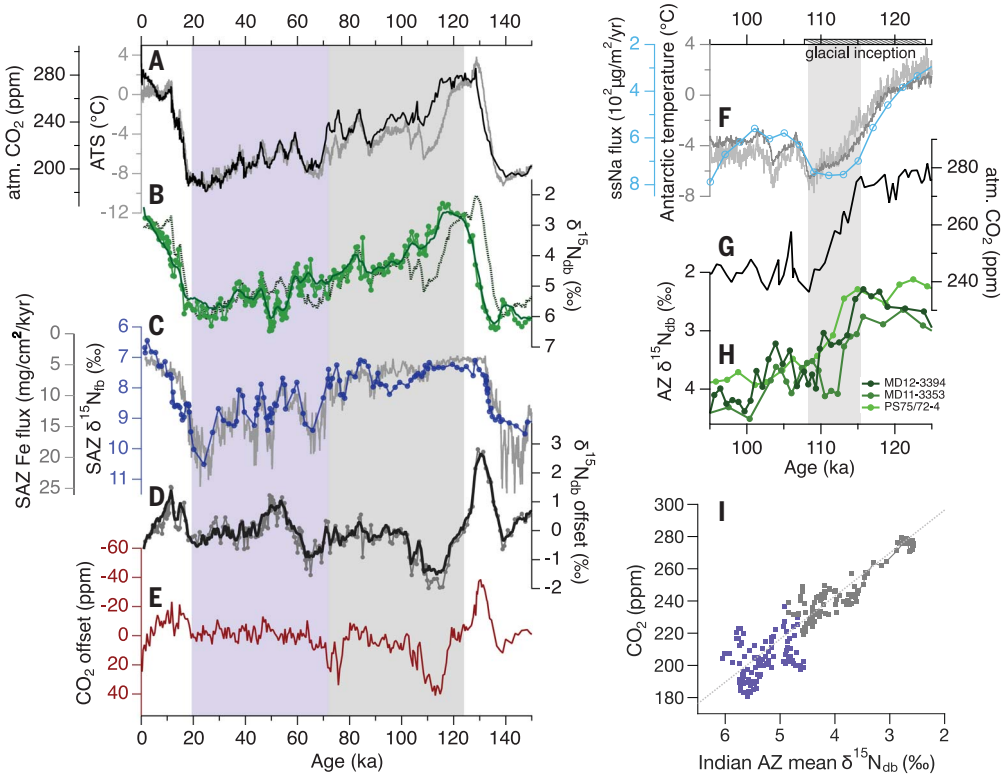


Fig. 3. Role of the AZ in atmospheric CO_2 change. (**A**) ATS (gray) (*18*) and atmospheric (atm.) CO_2 (black) (*47*). (**B**) Indian AZ mean $\delta^{15}N_{db}$ (dark green solid line), MD12-3394 $\delta^{15}N_{db}$ (light green circles), and the predicted $\delta^{15}N_{db}$ (dotted line). (**C**) $\delta^{15}N_{fb}$ (blue circles) and iron flux (gray line) at ODP Site 1090 in the SAZ (*14*). (**D**) $\delta^{15}N_{db}$ offset (measured minus predicted $\delta^{15}N_{db}$; see materials and methods) calculated using MD12-3394 $\delta^{15}N_{db}$ (gray circles) and the Indian AZ mean $\delta^{15}N_{db}$ (black line). (**E**) CO_2 offset from a linear regression between the atmospheric CO_2 composite and ATS (see materials and methods). The shaded intervals of (A) to (E) correspond to 72,000 to 124,000 years ago (gray) and 20,000 to 72,000 years ago (purple), respectively. (**F**, **G**, and **H**) An expansion of 95,000 to 125,000 years ago, covering the last glacial inception and including the three $\delta^{15}N_{db}$ records from Fig. 1. In (F), EPICA Dome C sea salt–Na flux (blue) (*42*) is plotted over ATS (dark gray) and EPICA Dome C air temperature (light gray) (*38*). The gray shading corresponds to the interval of CO_2 decline at the MIS 5e–MIS 5d transition. The hashed bar at the top indicates the glacial inception of 108,000 to 124,000 years ago. (**I**) Cross plot of CO_2 versus Indian AZ mean $\delta^{15}N_{db}$. Colors follow shaded intervals in (A) to (E). Connected markers are data for the glacial inception, indicated by the hashed bar in (F) to (H). Gray line is the trend fitted through the data of the interglacials (0 to 10,000 years ago and 115,000 to 126,000 years ago) and glacial maxima (20,000 to 27,000 years ago and 140,000 to 155,000 years ago). Deglacial intervals (126,000 to 140,000 years ago and 10,000 to 20,000 years ago) are excluded from the cross plot to focus on ice age inception and development (see fig. S8).

drawdown of CO_2 during the inception and early development of the last ice age. It has been proposed that sea ice drives the initial CO_2 decline at the MIS 5e–MIS 5d transition by acting as a barrier to ocean-atmosphere gas exchange (*22*). However, EPICA Dome C sea salt–Na flux, a proxy of AZ sea ice extent, indicates that the expansion of sea ice at the end of MIS 5e is in step with Antarctic air temperature and predates the CO_2 decline (Fig. 3F) (*42*).

The correlation of CO_2 with AZ $\delta^{15}N_{db}$ weakens after the MIS 5–MIS 4 transition. In particular, from 72,000 to 68,000 years ago, CO_2 declines sharply from 240 to 200 parts per million (ppm) even though AZ $\delta^{15}N_{db}$ remains between 4 and 5‰ (Fig. 3, A and B). The sharp drop in CO_2 may be caused by dust-borne iron fertilization in the SAZ: At this time in Ocean Drilling Project (ODP) Site 1090, $\delta^{15}N_{fb}$, iron flux, and biogenic fluxes all record sharp increases (*14*) (Fig. 3C and fig. S5). During MIS 3 through the LGM, both AZ $\delta^{15}N_{db}$ and SAZ $\delta^{15}N_{fb}$ exhibit millennial-scale variations that can explain CO_2 oscillations during that time interval (*14*) (Fig. 1F and Fig. 3, B and C). This suggests that the surface biogeochemical changes of the AZ and SAZ combine to explain much of the CO_2 variation during the second half of the ice age, including the AIM events and the deglacial CO_2 increase (*20*).

The effect of obliquity-driven mode 3 on Antarctic upwelling has substantial consequences for atmospheric CO_2. For comparison with the AZ $\delta^{15}N_{db}$ offset record, an analogous CO_2 offset record [as in (*36*)] is generated from the residuals of the linear correlation between CO_2 and Antarctic air temperature (Fig. 3E and materials and methods). The $\delta^{15}N_{db}$ and CO_2 offset records have similar features both between 0 and 20,000 years ago (the last deglaciation and the Holocene) and between 95,000 and 140,000 years ago (the penultimate deglaciation and early glacial progression of the last ice age) (Fig. 3, D and E). These features are consistent with mode 3 causing delays in the AZ contributions to both the deglacial CO_2 rise and the CO_2 decline associated with glacial inception and progression. Similarly, the parallel changes in the $\delta^{15}N_{db}$ and CO_2 offset records suggest that the 20-ppm increase in CO_2 during the late Holocene was caused by a strengthening of AZ upwelling (Fig. 3, A, B, D, and E) (*41*) due to falling obliquity (*40*) (Fig. 2E).

Atmospheric CO_2 change broadly works to amplify the glacial cycles. Our data point to the response of SWW-driven AZ upwelling to global climate (i.e., mode 1), along with SAZ iron fertilization, as central to this amplification. Moreover, the response of AZ upwelling to the bipolar seesaw (i.e., mode 2) causes the AZ to contribute to the CO_2 rises during deglaciations (*32*). However, CO_2 change appears to lag behind other components of the forcing-feedback system of the glacial cycles (*36*), with notable consequences. For example, extension of the period of strong greenhouse effect through the end of the penultimate interglacial and the beginning of the last ice age slows the global cooling, which contributes to the sawtooth pattern of the last glacial cycle (*43*). Similarly, gradually rising atmospheric CO_2 has been implicated in the persistent warmth of the Holocene, despite the orbital cues for preglacial cooling (*44*, *45*). If the glacial and deglacial lags in CO_2 change are owed to the effect of obliquity on AZ upwelling (i.e., mode 3), then this partly counteracts the long-recognized capacity of obliquity to affect global climate through the extent of polar ice (*46*). This counteraction may cause obliquity to cede control of glacial-interglacial transitions to precession and eccentricity, thus explaining the importance of the 20,000-year and 100,000-year periods in the glacial cycles of the past ~1 million years (*1*).

REFERENCES AND NOTES

1. J. Imbrie *et al*., *Paleoceanography* **7**, 701–738 (1992).
2. J. R. Petit *et al*., *Nature* **399**, 429–436 (1999).
3. J. H. Martin, S. E. Fitzwater, R. M. Gordon, *Global Biogeochem. Cycles* **4**, 5–12 (1990).
4. B. G. Mitchell, E. A. Brody, O. Holm-Hansen, C. McClain, J. Bishop, *Limnol. Oceanogr.* **36**, 1662–1677 (1991).
5. J. L. Sarmiento, J. R. Toggweiler, *Nature* **308**, 621–624 (1984).
6. B. B. Stephens, R. F. Keeling, *Nature* **404**, 171–174 (2000).
7. J. H. Martin, *Paleoceanography* **5**, 1–13 (1990).
8. D. M. Sigman, M. P. Hain, G. H. Haug, *Nature* **466**, 47–55 (2010).
9. N. Kumar *et al*., *Nature* **378**, 675–680 (1995).
10. K. E. Kohfeld, C. Le Quéré, S. P. Harrison, R. F. Anderson, *Science* **308**, 74–78 (2005).
11. S. L. Jaccard *et al*., *Science* **339**, 1419–1423 (2013).
12. R. A. Mortlock *et al*., *Nature* **351**, 220–223 (1991).
13. E. Wada, A. Hattori, *Geomicrobiol. J.* **1**, 85–101 (1978).
14. A. Martínez-García *et al*., *Science* **343**, 1347–1350 (2014).
15. A. S. Studer *et al*., *Paleoceanography* **30**, 845–862 (2015).
16. R. S. Robinson, D. M. Sigman, *Quat. Sci. Rev.* **27**, 1076–1090 (2008).
17. R. François *et al*., *Nature* **389**, 929–935 (1997).
18. F. Parrenin *et al*., *Science* **339**, 1060–1063 (2013).
19. L. M. Thöle *et al*., *Earth Planet. Sci. Lett.* **525**, 115716 (2019).
20. X. T. Wang *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 3352–3357 (2017).
21. P. C. Kemeny *et al*., *Paleoceanogr. Paleoclimatol.* **33**, 1453–1471 (2018).
22. K. E. Kohfeld, Z. Chase, *Earth Planet. Sci. Lett.* **472**, 206–215 (2017).
23. J. Shakun, D. Lea, L. Lisiecki, M. Raymo, *Earth Planet. Sci. Lett.* **426**, 58–68 (2015).
24. J. Lu, G. Chen, D. M. W. Frierson, *J. Atmos. Sci.* **67**, 3984–4000 (2010).
25. J. R. Toggweiler, B. Samuels, *Deep Sea Res. Part I Oceanogr. Res. Pap.* **42**, 477–500 (1995).
26. P. R. Gent, *Annu. Rev. Mar. Sci.* **8**, 79–94 (2016).
27. R. Abernathey, J. Marshall, D. Ferreira, *J. Phys. Oceanogr.* **41**, 2261–2278 (2011).
28. J. R. Toggweiler, J. L. Russell, S. R. Carson, *Paleoceanography* **21**, PA2005 (2006).
29. F. Lamy *et al*., *Earth Planet. Sci. Lett.* **259**, 400–413 (2007).
30. S. Barker *et al*., *Nature* **457**, 1097–1102 (2009).
31. S.-Y. Lee, J. C. H. Chiang, K. Matsumoto, K. S. Tokos, *Paleoceanography* **26**, PA1214 (2011).
32. R. F. Anderson *et al*., *Science* **323**, 1443–1448 (2009).
33. J. P. Sachs, R. F. Anderson, S. J. Lehman, *Science* **293**, 2077–2079 (2001).
34. A. Timmermann *et al*., *J. Clim.* **27**, 1863–1875 (2014).
35. F. Vimeux, K. M. Cuffey, J. Jouzel, *Earth Planet. Sci. Lett.* **203**, 829–843 (2002).
36. R. Uemura *et al*., *Nat. Commun.* **9**, 961 (2018).
37. J. S. Hoffman, P. U. Clark, A. C. Parnell, F. He, *Science* **355**, 276–279 (2017).
38. J. Jouzel *et al*., *Science* **317**, 793–796 (2007).
39. P. Kindler *et al*., *Clim. Past* **10**, 887–902 (2014).
40. V. Varma *et al*., *Clim. Past* **8**, 391–402 (2012).
41. A. S. Studer *et al*., *Nat. Geosci.* **11**, 756–760 (2018).
42. E. W. Wolff *et al*., *Nature* **440**, 491–496 (2006).
43. W. S. Broecker, J. van Donk, *Rev. Geophys.* **8**, 169–198 (1970).
44. W. F. Ruddiman *et al*., *Rev. Geophys.* **54**, 93–118 (2016).
45. Z. Liu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3501–E3505 (2014).
46. P. Huybers, C. Wunsch, *Nature* **434**, 491–494 (2005).
47. B. Bereiter *et al*., *Geophys. Res. Lett.* **42**, 542–549 (2015).
48. H. E. Garcia *et al*., *World Ocean Atlas 2018, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+nitrite, silicate)*, A. Mishonov, Ed. (NOAA Atlas NESDIS 84, 2019).
49. A. H. Orsi, T. Whitworth III, W. D. Nowlin Jr., *Deep Sea Res. Part I Oceanogr. Res. Pap.* **42**, 641–673 (1995).
50. J. Laskar, A. Fienga, M. Gastineau, H. Manche, *Astron. Astrophys.* **532**, A89 (2011).
51. A. Berger, M.-F. Loutre, *Quat. Sci. Rev.* **10**, 297–317 (1991).

ACKNOWLEDGMENTS

Cores MD12-3394 and MD11-3353 were retrieved during Indien Sud oceanographic cruises (E.M.), and we thank the crew of the R/V Marion Dufresne and the French Polar Institute (IPEV). We thank F. Rubach for technical support, A. Lin for help with $\delta^{15}N_{db}$ analysis, and N. Haghipour for assistance with radiocarbon analysis. We also thank X. T. Wang, J. Farmer, B. Zhang, and X. Zhang for discussions and N. Zhao for help in radiocarbon age calibration. **Funding:** This study was supported by U.S. NSF grant PLR-1401489 to D.M.S., ExxonMobil through the Andlinger Center for Energy and the Environment at Princeton University, Swiss NSF grant PBEZP2_145695 to A.S.S., Swiss NSF grants PP00P2_144811 and PP00P2_172915 to S.L.J., Swiss NSF grant P1BEP2_168625 to L.M.T., a Global Research Fellowship from the German Research Foundation (DFG grant GO 2294/2-1) to J.G., and the Max Planck Society. **Author contributions:** X.E.A., A.S.S., D.M.S., A.M.-G., and G.H.H. designed the study; X.E.A. and A.S.S. performed the $\delta^{15}N_{db}$ analyses; X.E.A., L.M.T., S.M., and M.S. performed the glycerol dibiphantanyl glycerol tetraether lipids analyses for TEX^{I}_{86} temperature reconstruction; X.E.A., L.M.T., S.L.J., and A.M.-G. constructed the age models; and E.M. provided access to the sediment cores and, with J.G. and L.A., contributed the planktonic foraminifera $\delta^{18}O$ measurements. X.E.A. and D.M.S. wrote the first draft of the manuscript. All authors contributed to the interpretation of the data and preparation of the final manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data in the analysis will be uploaded to the Pangaea database.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/1348/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 and S2
References (*52*–*75*)

11 June 2020; accepted 30 October 2020
10.1126/science.abd2115

Cell Culture Analyzer

Thrive Bioscience introduces CellAssist, which enables cell culture researchers to image, analyze, and document all cells, plates, reagents, and workflow details in a centralized database. CellAssist raises the bar for adherent cell culture by imaging and analyzing all the cells in a plate, not just cells chosen during typical manual inspection. Minutes after inserting a standard multiwell plate, CellAssist captures, analyzes, and displays hundreds of high-resolution images of cells at multiple magnifications, using phase-contrast or brightfield imaging. The visualization and analysis software allow researchers to comprehensively review images and metrics of the entire history of a project from their office or laboratory.

Thrive Bioscience
For info: 978-720-8044
www.thrivebiocellassist.com

Exosome Purification Kit

The Exo-spin 96 exosome purification kit from Cell Guidance Systems enables researchers to purify exosome samples in the standard 96-well plate format. The kit is based on the increasingly popular size-exclusion chromatography (SEC) method of exosome purification now used by most exosome research labs. Exo-spin 96 and Exo-spin mini columns have identical column dimensions and resin-bed length, allowing scale up from small to large sample numbers using the same protocol, which takes less than 1 h to complete and allows automation in the standard 96-well format. Advanced manufacturing techniques and stringent quality control systems used in the production of Exo-spin 96 kits ensure sample-to-sample consistency.

Cell Guidance Systems
For info: +44-(0)-1223-967316
cellgs.com/products/exo-spin-96.html

Durable Mirrors for High-Power Lasers

Optical Surfaces produces extremely durable, high-performance coated mirrors up to 600 mm in diameter for use with high-power femtosecond lasers. Employing the latest dielectric coating technology, these mirrors provide minimum pulse distortion and maximum usable bandwidth. When used with high-power femtosecond lasers operating at 750 nm to 850 nm, Optical Surfaces mirrors are designed for use at power densities exceeding 1 J/cm^2 for high-frequency 50-fs pulses without deterioration. Produced from materials including glass, ceramic, and silica, Optical Surfaces off-axis parabolic mirrors are designed to optimize the performance of rapid-pulsed, high-power lasers. All surface accuracies are checked by interferometer and are guaranteed to meet lambda/10 peak-to-valley (p-v) wave accuracy with low scatter. Typical applications for these extremely durable mirrors include plasma physics, laser manufacturing, and telecommunications.

Optical Surfaces
For info: +44-(0)-208-668-6126
www.optisurf.com

Bioactive Proteins

Abcam's new premium-grade bioactive proteins are highly active growth factors and cytokines designed to meet the specific needs of cell culture, including cell and gene therapy and regenerative medicine. Sourcing high-quality proteins with batch-to-batch consistency can be a challenge, as endotoxin levels and/or nonnative protein conformations and posttranslational modifications can affect safety and performance. These proteins are manufactured in mammalian cell lines under animal-free conditions, limiting endotoxins to ultralow levels [<5 endotoxin units (EU)/mg]. The use of mammalian cell lines produces proteins with the desired conformation and correct posttranslational modifications for optimal bioactivity. They resemble the proteins' native state, helping to ensure a safe, consistent outcome during drug discovery, development, and manufacturing. Developed by a team of industry-expert protein scientists, each batch is validated using state-of-the-art technology to ensure integrity, purity, concentration, monodispersity, and correct folding. Abcam can also design, manufacture, and validate proteins to custom needs and specifications.

Abcam
For info: 888-772-2226
www.abcam.com

Eight-Position Parallel Reaction Station

The DrySyn OCTO Mk II from Asynt is an eight-position reaction station that provides chemists with powerful magnetic stirring, heating, inert atmosphere, and reflux, all within a compact footprint. OCTO Mk II is an ideal tool for chemists performing syntheses and screening reactions in parallel. Accommodating individual reaction volumes of up to 6 mL, this reaction station has been designed to use standard low-cost reaction tubes to minimize your expenditure on consumable items. To further increase your parallel-synthesis productivity, up to three stations can be used together on a single standard magnetic hotplate stirrer, enabling you to perform up to 24 simultaneous reactions.

Asynt
For info: +44-(0)-1638-781709
www.asynt.com

Quantitative ELISA to Measure SARS-CoV-2 Antibodies

EUROIMMUN, a PerkinElmer company, has launched the Anti-SARS-CoV-2 QuantiVac ELISA (IgG) to quantify immunoglobulin G (IgG) antibodies against the SARS-CoV-2 S1 antigen. This quantitative test kit applies a recombinant S1 subunit of the SARS-CoV-2 spike protein, enabling detection of IgG antibodies, which have been the focus of vaccine development programs as well as discussions about possible immunity in previously infected individuals. This new ELISA assay supports careful evaluation of an important antibody subgroup enabling standard curve–based quantification of the anti-S1 IgG antibody concentration. EUROIMMUN has increased its ELISA production for COVID-19 diagnostics to more than 3 million tests per week and can supply large numbers of dried-blood spot collection sets for upcoming seroprevalence and vaccination development studies. Large international reference laboratories have implemented EUROIMMUN SARS-CoV-2 assays, and the tests are being applied in an increasing number of epidemiological studies worldwide.

PerkinElmer
For info: 800-762-4000
www.perkinelmer.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

INTRODUCTION TO CHANGCHUN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Prof. Huamin Yang
President of Changchun University of Science and Technology

"As the cradle of Chinese optical talents, CUST respects the talents, creates a friendly environment for talents, and offers spacious platforms and all-round resources as well as financial supports for talent cultivation. We sincerely welcome outstanding talents from home and abroad to work here to achieve your academic ambition."

Changchun University of Science and Technology (CUST), situated by the scenic South Lake of Changchun, Jilin Province, is originally known as Changchun Institute of Optics and Fine Mechanics which was founded by the Chinese Academy of Sciences in 1958. Professor Wang Daheng, the academician of the Chinese Academy of Sciences and the Chinese Academy of Engineering, was the first president. Professor Wang's motto of "being loyal and devoted, knowledgeable and professional, truth-seeking and innovative, modest and open-minded", and the school motto "Ambition, Perseverance, Practicality and Excellence" have exerted a profound and lasting influence on the development of the university during the last six decades.

Widely acknowledged as "the Cradle for Chinese Optical Talents" and the member of "National Basic Ability Construction Project of Western and Central China", CUST has cultivated over 100,000 graduates, among whom are academic elites, successful entrepreneurs and talents in all walks of life contributing themselves to the development of the country, the revitalization of science and technology, and the advance of the nation.

The university aims to cultivate innovative, versatile and practical talents, construct a comprehensive education system for undergraduates and graduates, with opto-elctronic technology being her outstanding feature, and optics, mechanics, electronics, computer as well as material science being its priority. The university coordinates the development of technology, science, arts, economics, management and law.

CUST actively conducts scientific research and enhances the researches with fruitful achievements in laser, optoelectronic instruments, advanced manufacturing, computer application, artificial intelligence, optoelectronic function material, nanometer material, as well as biomedical engineering.

The university always attaches great importance to international communication and has established substantial co-operations with more than 90 universities and research institutes from over 20 countries, including the United States, England, Germany and Russia, etc. More and more foreign scholars and students come to the university to conduct research or study, and international conferences are held here annually.

With the opportunity of promoting the cultivation of talents, CUST carries out the "Optical Talents Project" and launches a talent cultivating and recruiting program for the integrated development and prosperity.

Facing the chances of construction of World First Class Universities and Disciplines, CUST will always spare no efforts to become a top university in China with prestigious fame all over the world.

Sincerely welcome excellent talents to join CUST.

Tel: +86-431-8558-2300
Fax: +86-431-8538-3815
E-mail: zhaopin@cust.edu.cn

Add:
No. 7089, Weixing Road (East campus),
No. 7186, Weixing Road (South Campus),
No. 7989, Weixing Road (West Campus),
Changchun, Jilin Province, P. R. China
Post code: 130022

ADVERTISEMENT

Jiangsu Normal University (JSNU) is one of the top universities in Jiangsu Province, jointly sponsored and administrated by China's Ministry of Education and the People's Government of Jiangsu Province. Its origin could be dated back to 1952 in WuXi City of Jiangsu Province. In 1958, it was moved north to Xuzhou, and in the next year, the college was founded as Xuzhou Teachers' College, later was renamed as Xuzhou Normal University in 1996. From 2011 onwards, it has its present name of Jiangsu Normal University. Over the past years, as it attached priority to the building of high-level university and connotation enrichment, JSNU has enjoyed a rapid growth in many fields. In the 2021 US News Rankings, JSNU ranks the 1056th among global best universities, and the 102nd among China's mainland universities. And it ranks the 116th on the 2020 Chinese University Evaluation Rankings directed by Wu Shulian and his team.

JSNU has, in its university community, 22 colleges, CW CHU College, School of Continuing Education, International College and "The Joint Engineering Institute of Peter the Great St. Peterburg Polytechnic University and Jiangsu Normal University" (JSNU-SPbPU Institute of Engineering) which was the first non-independent legal entity for Sino-Russian jointly-running institute approved by the Ministry of Education, as well as an independent college — Kewen College. Nowadays, JSNU offers 59 undergraduate's programs, 35 master's programs at the first-layer of disciplines, 13 professional master programs, and 1 doctorate program targeted at China's special demands, covering 11 categories of disciplines. JSNU has a faculty population of 1,657 full-time professors, including 337 full professors and 584 associate professors. Among them, faculties with a doctorate degree account for 61.3%, and 43% of them went abroad for academic or research purposes. The university has 42 talents at state's top-level, including research fellows of Chinese Academy of Sciences and Chinese Academy of Engineering, as well as "Cheungkong Scholars". It has an enrollment of 19,300 full-time undergraduates, over 3,900 postgraduates and doctoral candidates, and over 440 international students in degree programs. At present, six of its disciplines were rated as provincial prestigious disciplines and 12 as provincial key disciplines. Its chemistry and engineering are listed among top 1% of ESI. The program of Chinese language and literature has blazed a trail in the field of Neurolinguistics with striking interdisciplinary features of arts, science and engineering. This program has won the titles of the national first-class program, the national characteristic program, the national teaching team and the national teaching demonstration center, which is globally influential. And its program of statistics has been granted one national key project and three excellent youth projects in recent years. And it is leading the researches of reliability theory and stochastic computing domestically, and it has made remarkable achievements in the national prevention against financial risks and the building of Smart Jiangsu Initiatives. And its biology program, entrusted by the state for special demands, has offered a lot to the national system of sweet potato industrial technology. JSNU has established one "2011" coordination innovation center, jointly sponsored by Jiangsu Province and the Ministry of Education. In the past five years, its faculties have been awarded 2 first prizes of National Teaching Achievements, 13 national key scientific research projects, 7 Outstanding Achievements Awards of the Ministry of Education. And its students have been awarded 7 awards for special contribution and 6 gold awards in national college student competitions, including "Internet +", "The Challenge Cup" and others. JSNU has been honored "National Model Campus" as the first group of winners and "National Model Organization" as the third group of winners. The university has been commended as "Jiangsu Model Organization" for ten consecutive times, "Jiangsu Model Campus" for five times, "Jiangsu Pacesetter Establishment" for five times by the CPC Jiangsu Committee and the Provincial People's Government. It also has been awarded "National Model Collective of National Solidarity and Progress" and other honors.

QUALIFICATIONS

1.Abide by national laws and regulations, devoted to Chinese higher education.
2.Have Ph.D. degree or equivalent plus postdoctoral experience.
3.Excellent record of research accomplishments.

APPLICATION

Please send the scanned copy of your resume and relevant support documents to email rsk@jsnu.edu.cn

CONTACT

E-mail: rsk@jsnu.edu.cn
Tel: (86516) 83500204
Address : No.101, Shanghai Road, Tongshan District, Xuzhou, Jiangsu Province P.R.China 221116

Follow us on Wechat

ADVERTISEMENT

Nanjing Forestry University (NFU) was founded in 1902 with the merger of the Forest Departments of Central University and of Jinling University. Located in the east of Nanjing, at the foot of beautiful Purple Mountain and east of Xuanwu Lake, NFU is a comprehensive university administered directly by Jiangsu Provincial Department of Education together with the Central Government. NFU has now been listed as one of the nation's top universities by China's Double First-class Initiative in Universities and Disciplines and is jointly built by Ministry of Education of China and Jiangsu People's Government.

Nanjing Forestry University is seeking top scholars and talents around the world to build a better tomorrow of NFU.

Eligibility Criteria

Metasequoia Scholar - Category C

(1) Willing to contribute to Chinese higher education; commitment to follow Chinese law and professional ethics.
(2) Graduated with PhD degree; minimum three years Postdoctoral research in overseas prestigious university/research institute, or obtained teaching or research position from overseas prestigious university/research institute/research department of companies.

Metasequoia Scholar - Category D

(1) Willing to contribute to Chinese higher education; commitment to follow Chinese law and professional ethics.
(2) Graduated with PhD degree; minimum two years Postdoctoral research in overseas prestigious university/research institute. The postdoctoral research and its duration can be discussed for those who have acquired PhD degree from overseas prestigious university and made significant academic achievements.

Supporting Policy

Category C

(1) Position: the qualified candidate will be offered a professorship and included in the public institutional system.
(2) Compensation: annual salary of RMB 600,000 for the first three years.
(3) Housing support: In addition to a temporary apartment on campus provided for three years, resettlement allowance of RMB 200,000 and housing subsidy of 1.5 million will be offered (excluding matching subsidy from local government).
(4) Research Fund: An amount of 2 million research fund will be provided (excluding matching fund from local government).
(5) Academic support: Sufficient space for office and lab; prioritized support to research team setup and master/PhD students enrollment.
(6) Facilitate kids to get enrollment and spouse to find a proper work.
(7) Prioritized support to apply for national young talent program and provincial talent program.

Category D

(1) Position: the qualified candidate will be offered a professorship and included in the public institutional system.
(2) Compensation: annual salary of RMB 300,000 for the first three years.
(3) Housing support: In addition to a temporary apartment(100 on campus provided for three years, resettlement allowance of RMB 200,000 and housing subsidy of RMB 800,000 will be offered (excluding matching subsidy from local government).
(4) Research Fund: An amount of 1 million research fund will be provided (excluding matching fund from local government).
(5) Academic support: Sufficient space for office and lab; prioritized support to research team setup and master/PhD students enrollment.
(6) Facilitate kids to get enrollment and spouse to find a proper work.
(7) Prioritized support to apply for provincial talent projects.

Contact Person from HR

For Chinese: Ms. Qian	Tel: 86-25-85420686	Email: grb@njfu.edu.cn
For foreigners: Ms. Tang	Tel: 86-25-85427310	Email: tzx@njfu.edu.cn

Disciplines open and contact information

ADVERTISEMENT

Tenure-track Positions at The HIT Center for Life Sciences, China

哈爾濱工業大學 HARBIN INSTITUTE OF TECHNOLOGY

The HIT Center for Life Sciences (HCLS) was founded in 2016. As the first special academic zone of Harbin Institute of Technology (HIT), HCLS is empowered by unparalleled freedom in research, graduate program, personnel employment and financial flexibility within the university. According to 2019, 2020, 2021US News Rankings of Best Global Universities, HIT's Biology and Biochemistry was ranked in the top ten in China.

HCLS is seeking outstanding PIs from all over the world. HCLS offers an internationally competitive start-up package and multidisciplinary innovative environment. The positions are open to all biological disciplines in Neuroscience, Microbiology and Immunology, Biophysics and Structural Biology (esp. Cryo-EM), Metabolomics, Chemical Biology, Nanotechnology and Tissue Engineering, Proteomics, DNA Damage and Repair, Genome Editing and translational research. HCLS is committed to being a world-class fundamental life science research institution.

QUALIFICATIONS

1.Ph.D. or M.D. degree or equivalent plus postdoctoral experience.
2.Excellent record of research accomplishments.
3.Great potentials in biomedical fields.

PACKAGE

1.International competitive salary:
500,000-750,000 CNY (tax included).
2.Sufficient Start-up fund:
10,000,000 CNY (2,000,000 CNY per annum for laboratory personnel, equipment, consumables and reagents as needed. No need to apply for additional funds during the appointment).
3.Better housing security:
A temporary furnished apartment (50-70 m2) and settlement allowance(Level A: 1,100,000 CNY, Level B: 600,000 CNY) will be provided.
4.Spacious Lab space and office space:
150 m2 or larger.
5.Guaranteed graduate student quota:
Each PI will allowed to accept up to 1-2 Ph.D students and 2-3 M.S. students. Postdoctoral and technical personnel can be recruited independently.
6.Advanced education resources:
HCLS will assist with arranging daycare or school for child and job for spouse.
7.Solid medical security services:
There are Harbin Institute of Technology Hospital and 2 affiliated hospitals to the University which are Heilongjiang Provincial Hospital, Harbin First Hospital.

Note: For candidates with exceptional accomplishment, the package can be negotiable.

EVALUATION

The initial appointment for PIs will be 5 years with full funding support from HIT. Permanent position will be obtained upon passing reviews by an international committee.

HOW TO APPLY

Applications in English as PDF files including a cover letter, a full CV, publication list, three referees and a statement of future research plans should be sent to Prof. Zhiwei Huang (**huangzhiwei@hit.edu.cn**), and indicate the "PI application" in the email subject. The application will open until filled.

Inner Mongolia University Warmly Welcomes Talents from Home and Abroad

Inner Mongolia University (IMU) is located in Hohhot, the capital of Inner Mongolia Autonomous Region and a famous historic and cultural city with a distance of 400 kilometers from Beijing. As the first comprehensive university in ethnic minority region after the founding of the People's Republic of China, IMU was established at a high starting point accomplished at one go in 1957 coinciding with the 10th anniversary of the founding of the Autonomous Region. With Ulanhu, the then-vice-premier of the State Council and governor of Inner Mongolia as the first president of IMU and a high-level faculty at the beginning stage of IMU consisting of discipline leaders and academic backbones selected by the CPC Central Committee and the central government from a dozen or so prestigious universities across China, including Peking University, who brought with them long-accumulated academic traditions and sublime spirits of universities, a solid foundation for running the university was laid, and healthy academic and campus atmosphere were introduced to IMU. IMU was approved to enroll graduate students in 1962, designated as a key university in China in 1978, authorized to confer doctoral degrees in 1984, included into Project 211 for the building of key universities in 1997, listed as a university co-sponsored by Inner Mongolia and the Ministry of Education in 2004, incorporated into the state program for building one high-level university in each of the provinces in the Mid-western Region in 2012, selected as one of the universities in China for the Double First-Class initiative in 2017, and appointed as a university of co-sponsorship by the Ministry of Education and Inner Mongolia, i.e. a university sponsored by the Ministry and included into the rank of universities under the administration of the Ministry, in 2018.

■ RECRUITMENT

To speed up the building process of the Double First-class Initiative, IMU has implemented the Steed Plan to enhance the introduction and cultivation of high-end academic talents and the Outstanding Talents Introduction Plan to introduce talents from home and abroad in large numbers. For the five years to come, IMU is ready to offer more than 400 vacancies for academic leaders, backbones and outstanding talents, provide first-class academic environment, resources, services and benefits for them and build with concerted efforts an ideal platform for career development, innovation and the realization of academic aspirations.

Tel: 0471-4992415 / 4995765 Email: ndszk@imu.edu.cn Website: www.imu.edu.cn

The recruitment of domestic and overseas talents for the School of Information Engineering, Jiangxi University of Science and Technology

Founded in 1958, Jiangxi University of Science and Technology (JXUST) was once attach to the Ministry of Metallurgical Industry China and China National Nonferrous Metals Industry Corporation. In 2013, it was jointly-sponsored by the Ministry of Education, the Ministry of Industry and Information Technology, and Jiangxi Provincial Government. JXUST has been authorized by the State Council of PRC to grant doctorate, master and bachelor degrees. The School of Information Engineering, JXUST was established in April 2001. JXUST is also one of the earliest universities in Jiangxi Province to conducting research in computer science and technology. The discipline of applications in computer technology in JXUST is now as a distinctive feature and brand at Jiangxi Province level. It was rated as one of the construction Projects for National First-class disciplines in 2009. Sixty years since the foundation, JXUST always keeps the school motto "Aim for the Highest and Take Responsibility as the First" in heart, and continuously promotes the construction of high-level talents team, consistently conducts the strategy of strengthening the university by talents, and try the best to achieve intensive development, characteristic development, and transformational development. The mission of JXUST aims to become an international well-known university in science and technology that has excellent academic level and encourages innovation and creativity.

Recruitment disciplines

optical engineering, electronic science and technology, information and communication engineering, computer science and technology, software engineering

Contacts: Staff Zhong Tel: 0086-797-8312029 Email: Z168@jxust.edu.cn

Nanjing Forestry University Special Job Fair Seeking "Metasequoia Scholars"

KEYWORDS

Forestry, Plant Protection, Horticulture, Grass Science, Forestry Engineering, Material Science and Engineering, Chemical Engineering and Technology, Biochemical Engineering, Applied Chemistry, Light Industry Technology and Engineering, Landscape Architecture, Garden Plants and Ornamental Horticulture, Mechanical Engineering , Control Science and Engineering, Civil Engineering, Transportation Engineering, Biology, Ecology, Environmental Science and Engineering, etc.

REQUIREMENTS

1. Graduated from a domestic "double first-class" university or a famous overseas university.
2. Two years of postdoctoral or work experience overseas.

HOW TO APPLY

CONTACT US

Wechat: 13810231645

2020 "Zhihui Chengyang" High-level Talent Forum & Qingdao University International Scholars Laoshan Forum

DISCIPLINES

History, Economics, Statistics, Law, Marxist Theory, Philosophy, Political Science, Sociology, Pedagogy, Psychology, Physical Education, Language and Literature, Drama and Film and Television, Management, Journalism and Communication, Art, Mathematics, computer science, mathematics, textile engineering, medicine, physics, oceanography, chemistry, biology, food science and engineering, mechanical engineering, etc.

QUALIFICATIONS

Outstanding scholars who have obtained a doctorate degree from a well-known university or academic institution at home and abroad, have the ability to independently conduct scientific research, and have representative results in the research field of the subject and development potential.

CONTACT US

Tel (wechat): 13810231645

For more info: Scan to apply

NUIST sincerely welcome high-level talents!

Nanjing University of Information Science & Technology (NUIST), founded in 1960 and renamed from Nanjing Institute of Meteorology in 2004, was designated as one of the key institutions of higher education in China in 1978, which enjoys the reputation as the cradle of meteorological talents and has worldwide influence in the field of Geoscience. In 2017, the university was listed in China's national Double First-class universities and disciplines and was also included in the provincial high-level universities in Jiangsu province. NUIST is co-constructed by the Jiangsu Provincial People's Government, the Ministry of Education of China (MOE), the China Meteorological Administration (CMA), the State Oceanic Administration (SOA) and the University of Chinese Academy of Sciences. NUIST is located in Nanjing Jiangbei New Area (National Level New Area).

After nearly 60 years of development, NUIST has established a comprehensive and coordinated discipline system of science, engineering, management, art and humanities, with atmospheric science as its core and information engineering as its basis. In recent years, the comprehensive rankings of NUIST in the four main international university rankings have promoted rapidly. In 2020, NUIST is ranked 40-58, 39 and 61 in US News, ARWU and QS rankings respectively. At present, NUIST has 6 doctoral programs, 22 master programs, 14 professional master programs and 71 bachelor programs. The discipline of Atmospheric Science ranks No. 1 in China. The four disciplines of Geoscience, Engineering, Computer Science and Environmental Science and Ecology rank top 1% ESI.

RECRUITMENT TYPES

(1) Type A

Academicians of Chinese Academy of Sciences or Chinese Academy of Engineering; Academicians from the National Academy of Sciences/Engineering from developed countries such as USA, UK, Canada, and Australia; Laureates of world renowned awards such as Nobel Prize, Carl-Gustaf Rossby Research Medal, Turing Award, and Field Medal; outstanding international scholars working in top universities or research institutions.

(2) Type B

Distinguished Professors of MOE, Winners of National Science Fund for Distinguished Young Scholars, National Distinguished Experts, Chiefs of NSFC Innovative Research Group, the first person of the 2nd Prize Project of the National Natural Science Awards/the 1st Prize Project of the State Technology Invention Awards/the 1st Prize Project of the National SCI-tech Progress Awards, National Teaching Masters, National key discipline leaders, Members of the Chinese Academy of Social Sciences, Tenured professors from renowned overseas universities. Applicants should be under the age of 50.

(3) Type C

National Young Scholars, Winners of National Science Fund for Outstanding Young Scholars, Young Scholars of MOE, Youth Top-notch Talent of Organization Department of the CPC Central Committee. Applicants should be under the age of 40.

(4) Type D (candidates for Type A, B and C)

The applicants who are not eligible for Type A/B/C but have potential to successfully apply for the four types of Youth Introduction Programs or other national talent introduction programs could be employed as backup talents.

SALARIES AND BENEFITS FOR HIGH-LEVEL TALENTS (UNIT: RMB)

Type	Annual salary	Settling-in allowance and housing subsidies	Start-up fund
Type A	1.5 million or more	5-20 million settling-in allowance; transitional housing or rental subsidy	1-10 million (0.5 million or more for humanities)
Type B	1 million or more	No less than 2 million settling-in allowance; transitional housing or rental subsidy	
Type C	800,000 or more	over 1.2 million settling-in allowance; a temporary accommodation	
Type D	300,000-600,000	Over 0.6 million settling-in allowance; a temporary accommodation	0.5 million or more

Contact information of HR of NUIST:

High-level talents:Tel: 86-25-58731027 (Ms.DING Ni), email: gdrc@nuist.edu.cn

Excellent teachers, Hong Kong, Macao and Taiwan teachers, and Professional foreign Teachers (non-ethnic-Chinese):

Tel: 86-25-58235197 (Ms. HUA Rong), email: zp@nuist.edu.cn

Northeast Agricultural University Welcomes Overseas Talents

Northeast Agricultural University (NEAU) is a national "211 Project" key construction university and a "World-class Discipline" construction university with agricultural science as its advantage, life science and food science as its characteristics, and multiple disciplines such as agriculture, industry, science, economics, management developing in a coordinated way.

NEAU is located in Harbin, Heilongjiang Province of China, and was established in 1948, which covers an area of 4.96 teaching faculties, possesses 3 national-level key disciplines and 3 national-level key (nurturing) disciplines, 4 disciplines stepped into the top 1% of ESI international disciplines ranking, 10 first-class subjects authorized to confer doctoral degree, and has 10 postdoctoral research mobile stations and 3 postdoctoral research stations.

In recent years, NEAU attaches great importance to the work of talents introduction and education,vigorously implements the strategy of "strengthening NEAU construction by emphasizing talents", and provides good development platforms and conditions for outstanding talents. Now we sincerely invite a group of leading talents of certain disciplines as well as other outstanding talents from overseas to join us.

I. Recruitment disciplines

Animal Husbandry, Veterinary Science, Food Science and Engineering, Agricultural Engineering, Biology, Crop Science, Horticulture, Agricultural Resources and Environment, Agricultural and Forestry Economic Management, Grass Science, Chemical Engineering and Technology, Water Conservancy Engineering, Ecology, Plant Protection, Landscape Architecture, Environmental Science and Engineering, etc.

II. Requirements

1. Support development of the People's Republic of China, abide by teachers' professional ethics, have the spirit of hard work and dedication, and ensure full-time work in NEAU.

2. Have more than 2 years of overseas study or research experience and have obtained a doctorate degree; professors and equivalents from famous overseas universities, research institutes or well-known enterprises research and development institutions will be preferred.

III. Remuneration

1. The recruited talents will be paid RMB 300,000-1 million/ year, according to their work performance or taking reference to their work situation abroad to determine via consultation, RMB 300,000-1 million as allowance for home settling-in, RMB 500,000-10 million as start-up funds for supporting scientific research, and be offered a turnover apartment. High-level talents will be assisted to arrange work for his/her spouse and children's admission to both pre-school and school.

2. The recruited talents will enjoy a set of supporting facilities for scientific research, independently determine the research direction, organize research teams, and use research funds within the specified scope of responsibility; high-level personnel will be equipped with 1-2 scientific research assistants.

3. The recruited talents will be given related certification of their professional and technical posts.

4. Other unmentioned issues can be further negotiated.

5.High-level leading talents can negotiate the above conditions one by one in details.

IV. Application methods

We invite qualified talents for interviews and discuss related matters.

Email: rsc@neau.edu.cn

Address: HR Department, Northeast Agricultural University, 600 Changjiang Road, Xiangfang District, Harbin City, Heilongjiang Province.

Postal code:150030

NEAU: www.neau.edu.cn

Contact persons: Mrs. Huang and Mr. Guan

Heilongjiang University of Science and Technology Is Recruiting Academic Faculty!

Heilongjiang University of Science and Technology is the only university in Heilongjiang, Jilin, and Inner Mongolia area which is characterized by mining engineering. Jointly sponsored by Heilongjiang Province People's Government and Ministry of Emergency Management (State Administration of Work Safety), it is praised as "outstanding university" by the higher education evaluation of Ministry of Education. It is the sole university which has doctoral-level talent program that respond to national special requirements"emergency rescue and influence control for coal mine accident", a university attached to post-doctoral research center for first-class discipline of safety science and engineering.

REQUIREMENTS

1. Candidates should have the nationality of the People's Republic of China, abide by discipline and law, have good ideological and political morality, have strong dedication and sense of responsibility, can devote themselves to teaching and research work.

2. For doctoral candidates, the first degree should be a full-time unified enrollment bachelor's degree from public university (excluding top-up), Master's degree and doctor's degree should be a full-time unified enrollment postgraduate. Graduates from schools in Hong Kong, Macao, Taiwan and foreign countries should have a certificate of academic degree issued by Chinese Service Center for Scholarly Exchange of the Ministry of Education.

RECRUITMENT PLAN

HLJUST plans to recruit 102 doctors in 2020. Recruiting categories include: Mineral Engineering, Surveying and Mapping, Geological Resources and Geological Engineering, Chemical Engineering and Technology, Electrical Engineering, Control Science and Engineering, Mechanical Engineering, Information and Communication Engineering, Electronics Science and Technology, Materials Science and Engineering, Architecture, Civil Engineering, Computer Science and Technology, Science of Business Administration, Science of Public Management, Applied Economics, Sociology, Marxist Theory, Foreign Languages and Literature, Mathematics, Physics and mechanics.

REMUNERATION AND BENEFITS

The employed candidates will be managed as public institution staff of Heilongjiang Province, and shall enjoy the corresponding remuneration as public institution staff and school faculty.

1. Leading talents (full-time)

(1) Settling-in allowance: 2,000,000 yuan RMB;

(2) Housing subsidy: 1,000,000 yuan RMB;

(3) Cost-of-living allowance: 10,000 yuan RMB per month;

(4)Start-up fund and annual salary: negotiable according to actual demand;

(5)Families: can be settled according to actual demand.

2. Outstanding talents (full-time)

(1) Settling-in allowance: 1,000,000 yuan RMB;

(2) Housing subsidy: 500,000 yuan RMB;

(3) Cost-of-living allowance: 5,000 yuan RMB per month for 5 consecutive years;

(4)Start-up fund and annual salary: negotiable according to actual demand;

(5)Families: can be settled according to actual demand.

3. Topnotch talents (full-time)

(1) Settling-in allowance: 500,000 yuan RMB;

(2) Housing subsidy: 300,000 yuan RMB;

(3) Cost-of-living allowance: 3,000 yuan RMB per month for 5 consecutive years;

(4) Start-up fund and annual salary: negotiable according to actual demand;

(5) Families: can be settled according to actual demand.

4. Provincial high-level talents (full-time)

(1) Settling-in allowance: 200,000 yuan RMB;

(2) Housing subsidy: 100,000 yuan RMB;

(3) Cost-of-living allowance: 2,000 yuan RMB per month for 5 consecutive years;

(4) Start-up fund and annual salary: negotiable according to actual demand;

(5) Families: can be settled according to actual demand.

5. Excellent Doctor (full-time)

(1) Settling-in allowance: 50,000-200,000 yuan RMB;

(2) Start-up fund (Science & Engineering): 30,000-100,000 yuan RMB;

(3)Start-up fund (Humanities & Social Science): 15,000-50,000 yuan RMB;

(5) Doctor's allowance: 600 RMB yuan per month.

HLJUST will give priority to those whose spouse needs to be settled in.The spouse will be recruited through open recruitment if he or she meets the requirements. It is negotiable according to actual demands.

HOW TO APPLY

E-mail address for application:
hljkjdxrsc@163.com

CONTACT

Mr. Wang, Mr. Han
Tel: +86(451) 88036082
Official Website: www.usth.edu.cn
Address: No. 2468, Puyuan Road, Songbei District, Harbin 150022

Multidisciplinary positions in Wenzhou Medical University

Wenzhou Medical University (WMU) is a higher institution co-developed by the Ministry of Education, National Health and Family Planning Commission and Zhejiang Provincial Government and a Priority Development University in Zhejiang Province. WMU offers 4 PhD Programs and 1 Professional PhD Programs in Clinical Medicine, Pharmacy, Biomedical Engineering and Basic Medicine, 12 Master Programs, 8 Professional Master's Programs, and 27 undergraduate programs in Medicine, Science, Engineering and Management. WMU has postdoctoral research centers under the Primary Category Discipline of Clinical Medicine and Pharmaceutical Science. WMU ranks 74th in the overall ESI ranking nationally. 7 programs including Clinical Medicine, Pharmacology and Toxicology, Biology and Biochemistry, Molecular Biology and Genetics, Materials Science, Chemistry, Neuroscience and Behavioristics rank within the global top 1% in ESI rankings, of which the Clinical Medicine Program ranks within the global top 0.2% in ESI rankings.

At present, the total enrollment of WMU amounts to 13,328, among which 890 are full-time international students. The graduates from the Clinical Medicine Program have consistently ranked within 10% nationwide in terms of their passing rate of National qualification examination for medical practitioners.

WMU has always been tapping its great potential for basic research and applied basic research to push itself to the forefront of international scientific research and to meet the needs of local development.WMU has high-level scientific research platforms and abundant clinical resources: 1 State Key Laboratory jointly built by Zhejiang Province and the Ministry of Health, 1 National Engineering Research Center, 1 National Clinical Research Center (for Ophthalmic Diseases), 2 China International Science and Technology Cooperation Bases, and 1 National Center for International Research; 5 directly-affiliated hospitals with more than 10 million outpatients and emergency patients visits per year and 3 of directly-affiliated hospitals are among the top 100 hospitals selected by Chinese Academy of Medicine in 2019 based on their technology impact. WMU has 25 non-directly affiliated hospitals (clinical colleges), which provide a solid foundation for medical teaching and research. For more information, please visit http://www.wmu.edu.cn/index.htm.

All-year-round recruitment

WMU is currently recruiting multidisciplinary faculties at several levels: discipline leading talents, discipline leaders, academic leaders, academic backbones and excellent doctoral and post-doctoral fellows. WMU provides a competitive salary, reward for scientific research, scientific research start-up funds, settling-in allowance, housing subsidies and medical insurance, etc. For more details of position description, please visit http://rsc.wmu.edu.cn/info/1043/2667.htm.

Research disciplines (including but not limited to)

Medical sciences, life science related disciplines, emerging interdisciplinary fundamental research, frontier research and translational research, etc.

1. Medical Sciences: clinical medicine, pharmaceutical science, basic medicine, public health and preventive medicine and medical technology, etc.
2. Sciences: life science, biology, physics, optics, chemistry and mathematics, etc.
3. Engineering: optical engineering, materials science and engineering, biomedical engineering, electronic science and technology, information and communication engineering, computer science and technology and environmental science and engineering, etc.

Application procedure

Please send your application including CV, research plan, cover letter and related supporting materials to **rscrck@wmu.edu.cn**. Please name your email after "name + University". Review of applications will start upon receipt.

CONTACT

Human Resource Department, Wenzhou Medical University
Tel. and fax: 0577-86689737
Email: rscrck@wmu.edu.cn
Official Website: www.wmu.edu.cn
Address: University Town, Wenzhou, Zhejiang, China

VAN ANDEL INSTITUTE®
INNOVATION UNLEASHED

A DIFFERENT KIND OF SCIENTIFIC HOME.

YOUR FUTURE STARTS HERE.

Are you an early career scientist looking for your first principal investigator role? Or mid-career searching for a close-knit institute that provides scientists exceptional support? If so, we want to hear from you!

Van Andel Institute is hiring! We are a global leader in cancer epigenetics and Parkinson's disease research, with complementary strengths in metabolism, structural biology and cell biology. An innovative hub for exceptional basic research, we engage the world's brightest minds to pursue bold ideas. Our scientists translate their discoveries through partnerships with academic, clinical, industry and philanthropic collaborators. Students in our research-intensive Graduate School complete their dissertation work in one of our 32 faculty-led labs.

We invite you to apply for one of six assistant professor positions in the departments of Epigenetics, Neurodegenerative Science, Metabolism and Nutritional Programming, Structural Biology and Cell Biology. If your research is in or related to one of these fields, we welcome your application. We also will consider strong applicants at the level of associate or full professor.

What we offer

A robust research environment: 32 labs, eight state-of-the-art cores, an accredited Graduate School and more than 400 dedicated scientists, educators and support staff.

Faculty support: 100% VAI-paid salary plus $65,000 annually in lab funds, a comprehensive start-up package, lab and home relocation services, administrative support, salaries for 1.5 graduate students, subsidized high-performance computing and Core Technologies and Services (including 60 hours/year of paid bioinformatics/biostatistics time), and internal funding programs.

Comprehensive benefits: Health, vision, dental, life and long-term care insurance, six weeks of parental leave, 401(k) with match and more.

One-of-a-kind operations and administrative support: Infrastructure designed to support science and give our investigators more bandwidth for research, and expert support teams including research operations, clinical collaboration management, high-performance computing, lab supply center, scientific events, grants/budget support, communications and marketing, philanthropy, on-site library services, glassware, business development and technology transfer.

A "Top 10" community: Located in the heart of the Medical Mile in downtown Grand Rapids, Michigan — a city consistently included on "Top 10" lists for raising a family, growth and much more.

Learn more and apply at vai.org/futurefaculty

CSIR-CENTRAL DRUG RESEARCH INSTITUTE
(Council of Scientific & Industrial Research)
Sector 10, Jankipuram Extension Sitapur Road, Lucknow-226031, India

Career Opportunities for Scientist
SCIENTIST RECRUITMENT-2020
Advertisement No 10/2020

A unique opportunity for research career in Science & Technology in the area of Drugs, Pharmaceuticals and Biomedical Research

CSIR-Central Drug Research Institute, Lucknow, a premier Institute under Council of Scientific and Industrial Research (CSIR), is involved in multidisciplinary R&D programs of both basic and applied nature.

With the aim to expand and strengthen the areas of Neurobiology, Virology, Parasite Biology, Endocrinology, Human Genetics, Analytical Biochemistry, Computational Chemistry, Medicinal Process Chemistry, Statistics, IPR & HR Management, the Institute is looking for enthusiastic, talented young researchers / professionals with brilliant academic record, proven scientific achievements and zeal to conduct research as per the mandate of the Institute aligned with National Missions. Due weightage will be given to the candidates involved in the product development/ technology innovation/applied technology / translation research. The positions offer exciting opportunities for career growth as per the Flexible Complementing Scheme for Scientists.

Applications are invited from motivated young Indian Nationals with requisite qualification and experience for the posts of Principal Scientist (1 post), Senior Scientist (8 posts) and Scientist (8 posts).

For detailed advertisement and submission of application, please visit https://recruit.cdri.res.in or access the link 'SCIENTIST RECRUITMENT-2020' on CSIR-CDRI Website https://cdri.res.in

Start date for registration and online application: 05/12/2020
Last date for registration and submission of application: 15/01/2021
Last date of receipt of physical copy of application at CSIR-CDRI: 01/02/2021

Who's the top employer for 2020?

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles at sciencecareers.org/topemployers

Science 2020 TOP EMPLOYER

Northeastern University College of Engineering

With over **185** tenured/tenure-track faculty, and **18** multidisciplinary research centers and institutes, and funding by eight federal agencies, Northeastern's College of Engineering is in a period of dynamic growth. Our emphasis on interdisciplinary, transformative and innovative research—tied to Northeastern's unique history of industry collaboration through the university's signature cooperative education program—enables partnerships with academic institutions, medical research centers, and companies near our centrally located Boston campus and around the globe.

Learn more and apply at coe.northeastern.edu/faculty-hiring

Consideration will be given to candidates at the assistant, associate, and full professor levels; successful applicants will be expected to lead internationally recognized research programs aligned with one or more of the college's strategic research initiatives. We are also seeking to recruit and support a broadly diverse community of faculty and staff and strives to foster an inclusive culture built on respect that affirms inter-group relations and builds cohesion. Applicants will be asked to submit a diversity statement discussing how they view their contributions to sustainment and improvement of diversity in the college and community at large.

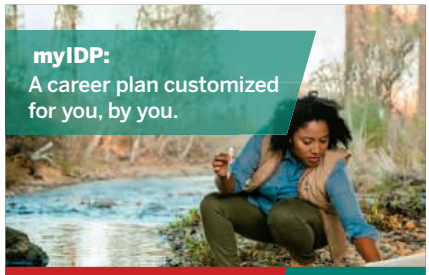
Northeastern University is an equal opportunity employer, seeking to recruit and support a broadly diverse community of faculty and staff. Northeastern values and celebrates diversity in all its forms and strives to foster an inclusive culture built on respect that affirms inter-group relations and builds cohesion.
All qualified applicants are encouraged to apply and will receive consideration for employment without regard to race, religion, color, national origin, age, sex, sexual orientation, disability status, or any other characteristic protected by applicable law.
To learn more about Northeastern University's commitment and support of diversity and inclusion, please see northeastern.edu/diversity.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Confused about your next career move?

Download Free Career Advice Booklets!

ScienceCcareers.org/booklets



There's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers In partnership with:

UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND
University of Massachusetts Medical School
FASEB Federation of American Societies for Experimental Biology

सत्यमेव जयते
Goverment Of India

APPOINTMENT OF DIRECTOR, IIT INDORE AND IIT MANDI

Ministry of Education
Department of Higher Education
Technical Section - I

Applications are invited for appointment to the post of **Director of Indian Institute of Technology (IIT)** at Indore and Mandi. The Director of an I IT is the academic and administrative head of the Institution. He/she is expected to have a minimum of 5 years' administrative experience and leadership qualities to head an Institute of National importance. The candidate/person should be a Ph.D. with first class or equivalent at the preceding degree, preferably in a branch of Engineering. In exceptional cases, candidates with Science, Mathematics or Management degrees may be considered. He/she should have an outstanding academic record throughout and a minimum of 10 years teaching experience as a Professor in a reputed Engineering or Technology Institute or University and should have guided Ph.D. students. The applicant should preferably be less than 60 years of age on the last date of receipt of the applications. The post carries a fixed pay of Rs. 2,25,000/- (Revised) per month, with allowances as per rules.

2. Interested individuals may apply giving their detailed resume in the prescribed format clearly bringing out research, teaching, industry-academia collaborations and administrative achievements, alongwith a two-page justification in support of their candidature, a two-page vision statement for the institution and contact details of at least two distinguished individuals well acquainted with their work. The application typed in the prescribed format along with enclosures may be sent by Registered/Speed Post to **The Under Secretary (TS.1), Department of Higher Education, Ministry of Education, Room No. 428 "C" Wing, Shastri Bhawan, New Delhi -110 001** so as to reach the Ministry on or before **31st December, 2020**. The detailed advertisement and the format of application is available on the website (**www.education.gov.in**).

Science Careers online @sciencecareers.org

By **Chee Kiang Ewe**

Dare to be different

Sitting in the giant lecture hall on my first day of university, I felt terribly out of place. As the first and only member of my family to attend college, leaving my home in rural Malaysia to study in New Zealand was exciting—and terrifying. I wished I was as confident and articulate as my classmates. In the years that followed, I coped by assimilating, relinquishing parts of my old self and remodeling my identity to blend in. I tuned my accent to sound like my classmates. I lied that my uncle was a doctor. I made friends, but consciously changing myself was mentally and emotionally draining. It took me a long time to see there could be another way.

When I came to the United States for graduate school, I changed my identity again. When my labmates made references to classic American films or books that I had never heard of, I nodded knowingly to hide my ignorance. I learned how to make small talk. I began to introduce myself exclusively as Ethan—my adopted American name—instead of my formal "ethnic-sounding" name, Chee Kiang. These acts took effort that I could have spent on my studies, but I didn't see any other option.

"I have stopped suppressing my identity as a first-generation, international student."

My perspective began to change when I was a teaching assistant (TA) in my second year of grad school. My students' evaluations of me were filled with negative comments about my accent while saying nothing about my teaching. I was embarrassed, but this time I resolved to handle the situation differently. My former impulse to focus on fitting in largely stemmed from my lack of self-confidence, but I had grown over the years. I had altered my identity twice; I was not going to do it again.

So, instead of changing my accent, I learned to slow the pace of my speech and to reiterate important points. Many of my students became more engaged. They started to show up to my office hours, seek me out for help, and strike up conversations when they saw me outside the classroom. I realized that I did not need to change my accent to be understood. Assimilation was not the only option.

Since then, I have stopped suppressing my identity as a first-generation, international student and instead celebrate it as a strength. Because of the extra miles I have covered, I am resourceful. Because of the additional hurdles I have overcome, I am persistent and resilient.

Throughout my academic career, I have felt I must work extra hard to be seen and heard. This pressure can be taxing. But it also spurs me to seize opportunities and take on responsibilities that help me grow as a scientist and professional. I watch research seminars on YouTube when I have time to fill, host additional office hours as a TA, and volunteer to organize departmental retreats and research symposiums. It isn't easy, but I have learned so much along the way.

I want underrepresented students like me to have a fighting chance in the academic world—and now I have the confidence to help make that happen. During the summer, U.S. Immigration and Customs Enforcement threatened to expel some international students amid the COVID-19 pandemic. Like many other international students and scholars in the United States, I felt unwanted and betrayed. So I took action. I sent an email to my department chair suggesting that this year's retreat highlight the importance of diversity—in race, ethnicity, gender, sexuality, nationality, socioeconomic status, and more. Much to my delight, my proposal was accepted. Over the 2-day meeting, students and faculty members had fruitful conversations about inequities in academia and celebrated the beautiful science conducted by diverse researchers from within and outside our campus.

I am still stumbling through the mysterious, labyrinthine world of academia, but now with more confidence than the insecure 20-year-old who thought he did not belong in that lecture hall. I am not the same as all of my peers—and that's OK. ■

Chee Kiang Ewe is a Ph.D. candidate at the University of California, Santa Barbara. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS