Vannevar Bush's report still resonates pp. 227 & 258

What it takes to safely reopen schools p. 241

Membrane fractionation of crude oil pp. 254 & 310

Science

$15
17 JULY 2020
sciencemag.org

AAAS



DEVELOPING BRAIN

Synapse complexity in mice changes with age pp. 253 & 270

CONTENTS

17 JULY 2020 • VOLUME 369 • ISSUE 6501



241

PHOTO: FUTIAN/CHINA NEWS SERVICE VIA GETTY IMAGES

NEWS

IN BRIEF

230 News at a glance

IN DEPTH

233 Famous psychologist faces posthumous reckoning
Dozens of papers on personality and health by Hans Eysenck have been retracted or are under suspicion *By C. O'Grady*

234 X-ray source gets a 100-fold boost in brightness
Rebuilt synchrotron is the first of more than a dozen of its kind in the works
By A. Cho

236 An optimist takes the helm at NSF
Sethuraman Panchanathan has spent his career finding "win-win" situations
By J. Mervis
EDITORIAL p. 229; PODCAST

237 'Huge hole' in testing data blurs racial, ethnic disparities
Researchers seek better data on pandemic's uneven burden
By K. Servick

238 The pandemic virus is slowly mutating. But does it matter?
Linking genomic changes to viral behavior proves difficult
By K. Kupferschmidt

239 Courses bring field sites and labs to the small screen
Some pandemic-induced changes to hands-on classes may be here to stay
By E. Pennisi

FEATURES

241 Not open and shut
School openings across the globe suggest ways to keep the coronavirus at bay, despite outbreaks
By J. Couzin-Frankel, G. Vogel, and M. Weiland
PODCAST

INSIGHTS

POLICY FORUM

246 The rotten apples of Brazil's agribusiness
Brazil's inability to tackle illegal deforestation puts the future of its agribusiness at risk
By R. Rajão

PERSPECTIVES

249 Can genomes predict coral bleaching?
Expansion of the coral genomic tool kit could facilitate more informed conservation
By R. A. Bay and L. Guerrero
RESEARCH ARTICLE p. 268

250 Nailing a Fe-rocious form of cancer
Disseminated cancer cells in brain fluids hijack iron from macrophages for metastatic survival *By L. Garzia and M. D. Taylor*
RESEARCH ARTICLE p. 276

252 Less can be more in functional materials
Nanopillar defects in $NaNbO_3$ induce a large piezoelectric response at high electric fields
By N. Bassiri-Gharb
REPORT p. 292

253 A synapse census for the ages
A molecular imaging survey reveals the evolution of mouse brain synapse populations from birth through old age
By K. D. Micheva et al.
RESEARCH ARTICLE p. 270

254 Reimagining petroleum refining
Rigid polymer membranes separate crude oil components without distillation
By J. F. Brennecke and B. Freeman
REPORT p. 310

256 Aging immunity may exacerbate COVID-19
Increased general inflammation in aging inhibits immunity and affects responses to infections
By A. N. Akbar and D. W. Gilroy

BOOKS *ET AL.*

258 Overseeing innovation
75 years after Vannevar Bush's impactful report, debate continues about directing science *By E. K. Gibson*
EDITORIAL p. 227

259 Color vision
An ambitious volume lends scientific and cultural context to an expansive concept
By M. Rossi

Published by AAAS

LETTERS

261 COVID-19 policies: Remember measles
By D. Guha-Sapir et al.

262 Restart science stronger after COVID-19
By A. I. Leshner

262 Uphold the nuclear weapons test moratorium
By J. King et al.

RESEARCH

IN BRIEF

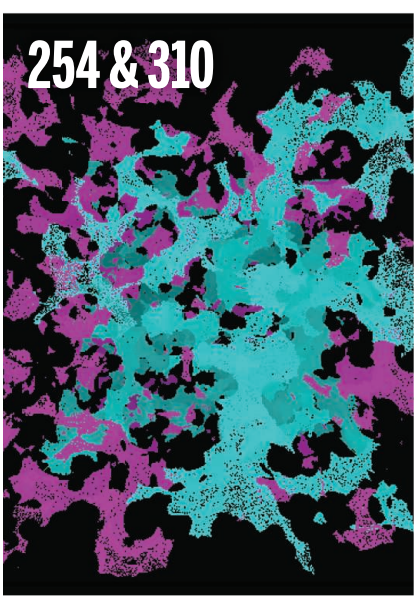
264 From *Science* and other journals

REVIEW

267 Metrology
Optical frequency combs: Coherently uniting the electromagnetic spectrum
S. A. Diddams et al.
REVIEW SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY3676

RESEARCH ARTICLES

268 Coral genomics
Population genetics of the coral *Acropora millepora*: Toward genomic prediction of bleaching *Z. L. Fuller* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABA4674
PERSPECTIVE P. 249



CREDITS (FROM LEFT): THOMPSON *ET AL.*; ROBERT NEUBECKER

269 Cancer
Tumor-initiating cells establish an IL-33–TGF-β niche signaling loop to promote cancer progression *S. Taniguchi* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY1813

270 Neurodevelopment
A brainwide atlas of synapses across the mouse life span *M. Cizeron* et al.
PERSPECTIVE p. 253

276 Cancer
Cancer cells deploy lipocalin-2 to collect limiting iron in leptomeningeal metastasis *Y. Chi* et al.
PERSPECTIVE p. 250

282 Human evolution
Ancient DNA indicates human population shifts and admixture in northern and southern China *M. A. Yang* et al.

REPORTS

289 Chemical physics
Second critical point in two realistic models of water *P. G. Debenedetti* et al.

292 Piezoelectrics
Giant piezoelectricity in oxide thin films with nanopillar structure *H. Liu* et al.
PERSPECTIVE p. 252

297 Coronavirus
Introductions and early spread of SARS-CoV-2 in the New York City area *A. S. Gonzalez-Reiche* et al.

302 Organic chemistry
Boron-enabled geometric isomerization of alkenes via selective energy-transfer catalysis *J. J. Molloy* et al.

307 Chemical physics
Molecular square dancing in CO-CO collisions *Z.-F. Sun* et al.

310 Organic separations
N-Aryl–linked spirocyclic polymers for membrane separations of complex hydrocarbon mixtures *K. A. Thompson* et al.
PERSPECTIVE p. 254; PODCAST

315 Coronavirus
Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic *R. E. Baker* et al.

320 Immunology
Autoreactivity in naïve human fetal B cells is associated with commensal bacteria recognition *J. W. Chen* et al.

325 Aging
A programmable fate decision landscape underlies single-cell aging in yeast *Y. Li* et al.

330 Coronavirus
Site-specific glycan analysis of the SARS-CoV-2 spike *Y. Watanabe* et al.

333 Genome editing
CRISPR-CasΦ from huge phages is a hypercompact genome editor *P. Pausch* et al.



DEPARTMENTS

227 Editorial
Science has always been political
By H. Holden Thorp
EDITORIAL p. 229; BOOKS p. 258

229 Editorial
Advice for the new NSF director
By Mary Sue Coleman
EDITORIAL p. 227; NEWS STORY p. 236

342 Working Life
When the jokes aren't funny *By Lala Tanmoy Das*

ON THE COVER



Synaptome maps of the developing mouse brain (relative brain size is shown to scale). The diversity of excitatory synapses (low to high: blue, green, yellow, red) changes with age (top to bottom: 1 day, 2 weeks, 3 weeks, 3 months, 12 months), peaking in the 3-month-old cortex and hippocampus and declining in old age, potentially mirroring life-span trajectories in brain function. See pages 253 and 270. *Image: Zhen Qiu, Mélissa Cizeron, and Seth Grant, Edinburgh University*

Science Staff 226
Science Careers 341

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), 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** Eva Frederick, Rodrigo Perez Ortega **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, 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), Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin (Acting 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

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 **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **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

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*
Judith Allen, *U. of Manchester*
Sebastian Amigorena, *Institut Curie*
James Analytis, *U. of California, 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, *U. of Texas at Dallas*
Peter Bearman, *Columbia U.*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Gabriele Bergers, *VIB*
Bradley Bernstein, *Mass. General Hospital*
Alessandra Biffi, *Harvard Med. School*
Peer Bork, *EMBL*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *U. of California, Santa Cruz*
Ron Brookmeyer, *U. of California, Los Angeles* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *U. of California, Irvine*
György Buzsáki, *New York U. School of Med.*
Blanche Capel, *Duke U.*
Annmarie Carlton, *U. of California, Irvine*
Nick Chater, *U. of Warwick*
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 Pittsburgh*
Roberta Croce, *VU Amsterdam*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *U. of North Carolina*
Tom Daniel, *U. of Washington*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Frans de Waal, *Emory U.*
Claude Desplan, *New York U.*
Sandra Díaz, *U. Nacional de Córdoba*
Ulrike Diebold, *TU Wien*
Hong Ding, *Inst. of Physics, CAS*
Jennifer Dionne, *Stanford U.*
Dennis Discher, *U. of Penn.*
Gerald Dorn, *Washington U. in St. Louis*
Jennifer A. Doudna, *U. of California, Berkeley*
Bruce Dunn, *U. of California, Los Angeles*
William Dunphy, *Caltech*
Christopher Dye, *U. of Oxford*
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*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Michael Feuer, *The George Washington U.*
Toren Finkel, *U. of Pittsburgh 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, *U. of California, Los Angeles*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *U. of South Florida*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Kip Guy, *U. of Kentucky College of Pharmacy*
Taekjip Ha, *Johns Hopkins U.*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. College 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*
Lora Hooper, *UT Southwestern Med. Ctr.*
Fred Hughson, *Princeton U.*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS and U. of Arizona*
Kai Johnsson, *EPFL*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *U. of California, Berkeley*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med. School*
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, *U. of Penn.*
Wendell Lim, *U. of California, San Francisco*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Jonathan Losos, *Washington U. in St. Louis*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Oscar Marín, *King's College London*
Charles Marshall, *U. of California, Berkeley*
Christopher Marx, *U. of Idaho*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth College*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Graham Medley, *LSHTM*
Jane Memmott, *U. of Bristol*
Baoxia Mi, *U. of California, Berkeley*
Edward Miguel, *U. of California, Berkeley*
Tom Misteli, *NCI, NIH*
Yasushi Miyashita, *U. of Tokyo*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *U. of California, 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*
Margaret Palmer, *U. of Maryland*
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. for Biological Studies*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Matthieu Piel, *Institut Curie*
Kathrin Plath, *U. of California, Los Angeles*
Martin Plenio, *Ulm U.*
Katherine Pollard, *U. of California, San Francisco*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Félix A. Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial College London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *U. of Texas at 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, *U. of California, Santa Barbara*
Rebecca Sear, *LSHTM*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *U. of California, Santa Cruz*
Jay Shendure, *U. of Washington*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *U. of California, San Francisco*
Robert Siliciano, *Johns Hopkins U. School of Med.*
Lucia Sivilotti, *U. College London*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina* **(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 College*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
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. of Sci.*
Jane-Ling Wang, *U. of California, Davis* **(S)**
David Waxman, *Fudan U.*
Jonathan Weissman, *U. of California, San Francisco*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *U. of California, Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *U. of Penn. School of Med.*
Jonathan Zehr, *U. of California, Santa Cruz*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

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.

Published by AAAS

EDITORIAL

Science has always been political



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

This month marks the 75th anniversary of Vannevar Bush's paper, "Science—The Endless Frontier." The report set out the rationale and structure of a system to fund science in the United States; it is the Magna Carta of American science. *Science* published a summary of Bush's paper on the same day it was submitted to President Harry Truman, and Bush's close relationship with the magazine is evident by the number of times he wrote and was quoted in *Science* during the 1940s.

Bush's report and advocacy seized on the historical moment. He was a driving force behind the Manhattan Project, and the scientists who developed the atomic bomb at Los Alamos were viewed as American heroes who won World War II with their minds (we now correctly regret the carnage of Hiroshima and Nagasaki and have more sophisticated views about the danger of nuclear weapons). Bush shrewdly recognized that it was a high-water mark for American trust in science. He used this political capital to create the federally funded scientific establishment in the United States.

> "'...The Endless Frontier' is first and foremost a masterwork of political persuasion."

Bush's document is widely celebrated by U.S. scientists, because it sets out many of the nation's most cherished principles. It is fundamentally based on the idea that professors should conduct basic research with federal dollars because universities offer a home for free inquiry and students can be trained in this environment of knowledge creation. It is one of the earliest and most persuasive cases for professors as teacher-scholars. He was passionate that university researchers should be free from government influence, and later tangled in the implementation of his vision with Senator Harley Kilgore, who felt that the government should force more applied research with more federal entitlements.

There is a disconnect, however, in how we remember this. Although Bush is often held up as a model for keeping politics out of science, his writing and efforts were unmistakably political. "Science—The Endless Frontier" is first and foremost a masterwork of political persuasion. Yes, it eventually ends up with professors carrying out curiosity-driven research without government intervention, but the exposition that leads to this conclusion is nothing short of a jeremiad—a prophecy of American deterioration without immediate investment in science. Perhaps deep in his psyche, basic research was important to Bush for its beauty and transcendent qualities, but in reading his three arguments for federal science, he believed it was crucial for American survival.

Bush begins with an argument that is still the most politically useful in American science—"For the War Against Disease." As science has gone in and out of favor in Washington, DC, medical research has almost always had bipartisan support, and Bush's brilliant framing ensured that fundamental understanding of biology always had support. The second argument—"For Our National Security"—leverages the support that Bush and his colleagues obtained in building the atomic bomb as well as the fear that Americans had of adversaries that could gain an advantage in military technology. "And for the Public Welfare" posits that scientific capability and a trained scientific workforce are necessary for a flourishing economy in the United States.

These arguments still work today. It would be hard to find Congressional testimony from scientific advocates over the intervening 75 years that didn't echo some of these themes. But Bush's paper is not without its critics. He was not an advocate for the humanities and social sciences; when he led the Carnegie Institution of Washington, he discontinued the archaeology program and he clearly states the primacy of natural science in "The Endless Frontier." Further, he sets up a syllogism by first stating that science is useful to the country, in order to argue for unfettered basic inquiry, which implies that basic knowledge must therefore also be useful. For many nonscience colleagues, this instrumentalism undermines the ability of academic institutions to make the case for knowledge as a public good.

Bush's paper is nothing if not political. Especially when viewed through the lens of coronavirus disease 2019 (COVID-19) and the dangerous antiscience movements of today, it is time to move away from proferring Vannevar Bush and his essay as symbols of nonpolitical science. His legacy is precisely the opposite: that science thrives when its advocates are shrewd politicians but suffers when its opponents are better at politics.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

10.1126/science.abd7628

Published by AAAS

EDITORIAL

Advice for the new NSF director

A new director, Dr. Sethuraman Panchanathan, takes the helm of the U.S. National Science Foundation (NSF) during a pivotal time for the future of scientific research in the United States. This month marks the 75th anniversary of the seminal report "Science—The Endless Frontier" by Vannevar Bush, director of the Office of Scientific Research and Development, submitted in 1945. The treatise laid the groundwork for the NSF and many other investments that have kept the United States at the forefront of scientific advancement and economic innovation. Today, American science and the United States face similar challenges that center around national security, economic strength, and social justice. To renovate the NSF in ways that address current issues, Dr. Panchanathan should use his position to build on the time-tested partnership between science and the U.S. government. To this end, I offer the following advice for improving the NSF in the coming years.

In his new role, the director should position himself as the nation's leading public advocate for science. Although American society relies on scientific research and investment more than ever before, the coronavirus disease 2019 (COVID-19) pandemic has exposed in the American populace a vast and dangerous ignorance about science. Because the NSF supports all science and engineering disciplines, the director is uniquely positioned to be America's spokesperson for the promise of science—and government investment in science—to solve vexing problems and grow the economy. The NSF director can also speak about the nature of scientific research and how both successes and failures create new knowledge and innovations worthy of taxpayer investments. To use this bully pulpit for effectively communicating with policy-makers and the public, the director should rely on the agency's communications resources.

Current events have refocused attention on the fact that all fields of scientific research in the United States have been far too White and male for far too long. The director should address the need for greater diversity in the scientific workforce. Drawing on a diverse talent pool can provide a key strategic advantage for the United States because scientists from diverse backgrounds ask different questions and in serving as role models, increase public support for science. The new director should seize this moment in our social consciousness to make science more inclusive by building on existing programs (such as the NSF ADVANCE and INCLUDES initiatives). He should also follow the recommendations of the National Science Board Vision 2030 report, which encourages institutional change and the adoption of proven science, technology, engineering, and mathematics (STEM) pedagogy and practices that boost diversity and inclusion. Such remodeling should prioritize inclusiveness and innovation in the way the NSF reviews and awards grants. The NSF's merit-based review system is the world's gold standard, but even gold needs occasional polishing. Streamlining the current grant application process will benefit applicants from all backgrounds by reducing structural barriers to entry that often hinder minority scientists and minority-serving institutions.

> "...the director is uniquely positioned to be America's spokesperson for... science..."

Innovations that result from NSF investments must serve the public good. Building on successful programs such as the NSF Innovation Corps, the new director should seek ways to help faculty and students move scientific discoveries into the marketplace efficiently for the public good. For example, providing researchers with relatively modest supplemental awards to support market analysis or other aspects of technology transfer could help science faculty advance NSF-funded research with potential applications in both the public and private sectors. And the director should guide the NSF to focus on emerging fields such as artificial intelligence, quantum mechanics, and cognitive neuroscience—all of which are crucial to America's national security and economic competitiveness.

These suggestions are starting points. Dr. Panchanathan should take advice not only from longtime scientists and university leaders like myself, but also from NSF career staff and the many university faculty members who rotate through NSF. Fortunately for all scientists and citizens, Dr. Panchanathan has a world-class track record of leadership on diversity, reform, and advancing emerging fields of science. For instance, at Arizona State University, Dr. Panchanathan developed assistive devices for the visually impaired, and demonstrably sharpened the university's focus on innovation and entrepreneurship. At this crucial juncture for America's scientific enterprise, NSF's future is in extremely capable hands.

–Mary Sue Coleman



Mary Sue Coleman
is president of the Association of American Universities, Washington, DC, USA.
marysuec@aau.edu

PHOTO: AUSTIN THOMASON/UNIVERSITY OF MICHIGAN

10.1126/science.abd7626

Published by AAAS

NEWS

IN BRIEF

Edited by **Catherine Matacic**



CONSERVATION

One-third of lemur species are critically endangered

Lemurs, whose populations have long been in decline, are now in even greater peril. Ninety-five percent of the 111 primate species—all native to Madagascar—are under threat of extinction from habitat loss and hunting, and 31% are critically endangered, according to a new review by the International Union for Conservation of Nature (IUCN). "They're hanging on by the skin of their teeth," says primatologist Russell Mittermeier of Global Wildlife Conservation, who led the analysis. During the last such review, conducted in 2008, 43% of all lemur species were under threat of extinction and 10% were critically endangered. (That review, however, lacked sufficient data for 55 species included in the new assessment.) The new assessment led IUCN to last week raise its Red List threat levels for 13 lemur species—bumping some from "endangered" to "critically endangered." They include Madame Berthe's mouse lemur (*Microcebus berthae*), the smallest of all primates with a body length of 9 to 11 centimeters (above). All told, 599 species had their threat levels revised, from West Africa's king colobus monkey to the North Atlantic right whale. Another 4260 species were added to the Red List for the first time.

Coronavirus origins explored

PUBLIC HEALTH | The World Health Organization (WHO) dispatched a two-member team to China on 10 July to lay the groundwork for an international mission to investigate the origins of the COVID-19 pandemic. The duo—an epidemiologist and an animal health expert whose names were not revealed—will discuss the scope of the mission with Chinese officials. The timeline for the larger investigation is unclear. Theories about the pandemic's origins have proliferated—including the possibility that the virus escaped from a lab or was created intentionally—but so far, there is little evidence for any of them. Many infectious disease experts suspect the virus jumped from bats into other animals before moving into humans, so studies of potential intermediary species could prove useful.

Project probes ice sheet collapse

GLACIOLOGY | The National Science Foundation revealed this week that it is funding a 5-year ice drilling project in Antarctica to seek evidence of the last time the West Antarctic Ice Sheet (WAIS) collapsed. Scientists have long suspected that 125,000 years ago, during the Eemian, meltwater from WAIS drowned a world not much warmer than today in 3 meters of rising tides. But signs of such a collapse—and the speed of the melt—are scarce. To find them, the more than $3 million project will drill a 2.5-kilometer-long core from Hercules Dome, a stretch of stable ice in the center of the continent, and analyze the frozen snow and gas bubbles trapped in the ice. Because Hercules sits at the saddle between the continent's main ice sheets, near the Transantarctic Mountains, a collapse of the WAIS would put it right "on the waterfront," says Eric Steig, the project's principal investigator and a glaciologist at the University of Washington, Seattle. The drilling, expected to begin in 2023, will inform projections of how quickly WAIS could melt under current global warming.

PHOTO: HOUDIN AND PALANQUE/MINDEN PICTURES

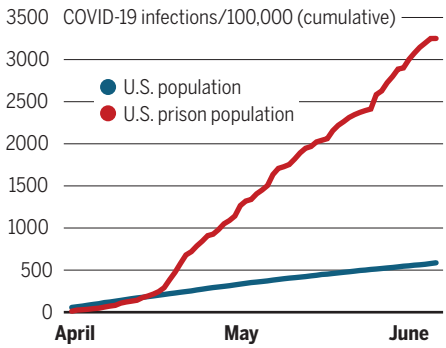
Published by AAAS

Coronavirus spikes in prisons

PUBLIC HEALTH | As COVID-19 cases soar in vulnerable communities across the United States, another group is seeing its numbers spike: inmates in state and federal prisons. A new study, using testing data from the U.S. Centers for Disease Control and Prevention, the Federal Bureau of Prisons, and state departments of corrections, found that—from 31 March to 6 June—prisoners were 5.5 times more likely to be diagnosed with the novel coronavirus than members of the general public, with cumulative infection rates of 3251 per 100,000 and 587 per 100,000, respectively (see chart, right). The study, published 8 July in *JAMA*, also found a higher rate of COVID-19 related deaths among U.S. prisoners—39 deaths per 100,000 compared with 29 deaths per 100,000 in the general public. Because prisoners tend to be younger, researchers say, the gap in deaths is lower than the gap in infections.

◀ Too close for comfort

Crowding, high rates of chronic health conditions, and limited access to personal protective equipment are likely contributors to cumulative infection rates in U.S. prisoners that are 5.5 times higher than in the general U.S. population.



Big bucks for antibiotics

BIOMEDICINE | Some of the world's largest pharmaceutical companies have jointly set up a $1 billion fund to help struggling biotech startups develop ways to combat the growing problem of antimicrobial resistance. The AMR Action Fund, which seeks to create two to four new antibiotics by the year 2030, was unveiled last week by giants such as Roche and Merck in partnership with the World Health Organization. Although the world is in dire need of new antibiotics, it has been hard to get investors to fund corporate efforts to find or synthesize them. That's because new antibiotics are generally reserved for resistant infections—and are given for short courses—making them less profitable than other drugs.

First full-sequencing success

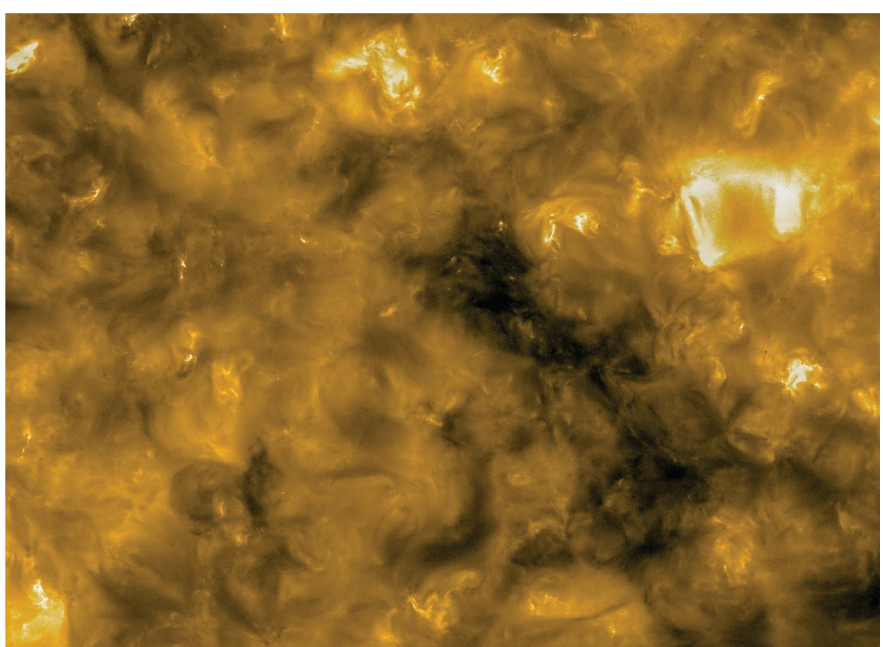
GENETICS | Even the most fully sequenced human genome is riddled with gaps, often because long stretches of repetitive DNA befuddle sequencing machines. By using diverse methods to sequence and arrange that DNA in the right order, researchers have closed all 29 gaps in the X chromosome sequence. That chromosome is now the first to have end-to-end, or telomere-to-telomere, coverage, an important and long-sought step toward pinning down the genetic basis of disease. "It was hard," says project coordinator Karen Miga from the University of California, Santa Cruz. Other chromosomes will soon follow, though a few will have to await technological advances, she and her team report this week in *Nature*. Next up for full sequencing: chromosomes eight and six.

Ice-measuring satellites align

GLACIOLOGY | The European Space Agency was set this week to raise the orbit of its aging ice measuring satellite, CryoSat-2, by nearly 1 kilometer to align it with NASA's ICESat-2 mission. Combining the satellites' data will offer unprecedented insight into the thickness—and vulnerability—of sea ice at the poles. CryoSat-2 uses reflected radar, which penetrates through snow to the ice beneath, whereas the newer ICESat-2, launched in 2018, uses reflected lasers that bounce off the top of the snow. CryoSat-2's raised orbit will allow the spacecraft to measure the same polar ice as ICESat-2 every 1.5 days. That will be especially useful in the Antarctic, where sea ice may be blanketed in a confounding layer of snow up to 2 meters high. The maneuver was set to begin on 16 July, and will be completed prior to the annual Arctic sea ice minimum in September.

Red alga imperils Hawaii's reefs

BIODIVERSITY | A mysterious red alga is rapidly spreading on coral reefs near the northwestern islands of Hawaii, blocking sunlight and smothering marine life below. Scientists first observed the alga—which they propose is a new species—in 2016 in the Pearl and Hermes Atoll, an uninhabited area more than 2000 kilometers northwest of Honolulu. Three years later, they discovered that thick, matlike layers of the seaweed had covered large expanses of the atoll, killing corals and other algae. The new alga, which they have named *Chondria tumulosa*, does not appear to



IN FOCUS There *is* something new under the Sun. When the European Space Agency's Sun-bound Solar Orbiter released its first photos this week, close-ups of the Sun's lower atmosphere, or corona, revealed tiny eruptions that mission scientists are calling "campfires." The flares, barely visible as bright specks in this image, are just 400 kilometers across—the smallest solar features ever seen in the extreme ultraviolet. Formed when magnetic field lines reconnect in bursts of energy, the flares could explain why the corona is so much hotter than the Sun's surface.

CREDITS: (GRAPHIC) C. MATACIC/*SCIENCE*; (DATA) UCLA LAW COVID-19 BEHIND BARS DATA PROJECT; (IMAGE) ESA/SOLAR ORBITER/EUI TEAM/CSL/IAS/MPS/PMOD/WRC/ROB/UCL/MSSL

Published by AAAS

be closely related to any native Hawaiian varieties, the team reported on 7 July in *PLOS ONE*. It's not clear where the seaweed came from, and the researchers say they are concerned that the species' rapid growth could disrupt the atoll's ecosystem.

White men dominate textbooks

DIVERSITY | Few nonwhite or female scientists appear in textbooks widely used in U.S. introductory undergraduate biology classes, a study has found. In seven of the most commonly used ones, only about 7% of more than 1100 scientists mentioned were nonwhite. Nonwhite students received 35% of U.S. bachelor's degrees in biology in 2015–16, and the lack of visible role models from underrepresented groups in textbooks might have discouraged more from entering the field, the research team suggests. What's more, the textbooks mentioned seven male biologists for every one female scientist; women received 60% of the bachelor's degrees in 2015–16. The share of the textbooks' references to women and people of color has grown only slowly, the researchers, from multiple institutions including Auburn University, reported in the 24 June issue of the *Proceedings of the Royal Society B*. To inspire undergrads, the authors suggest that universities highlight diversity in other ways—for example, by posting photographs in classrooms and other campus spaces of women and people of color doing research.

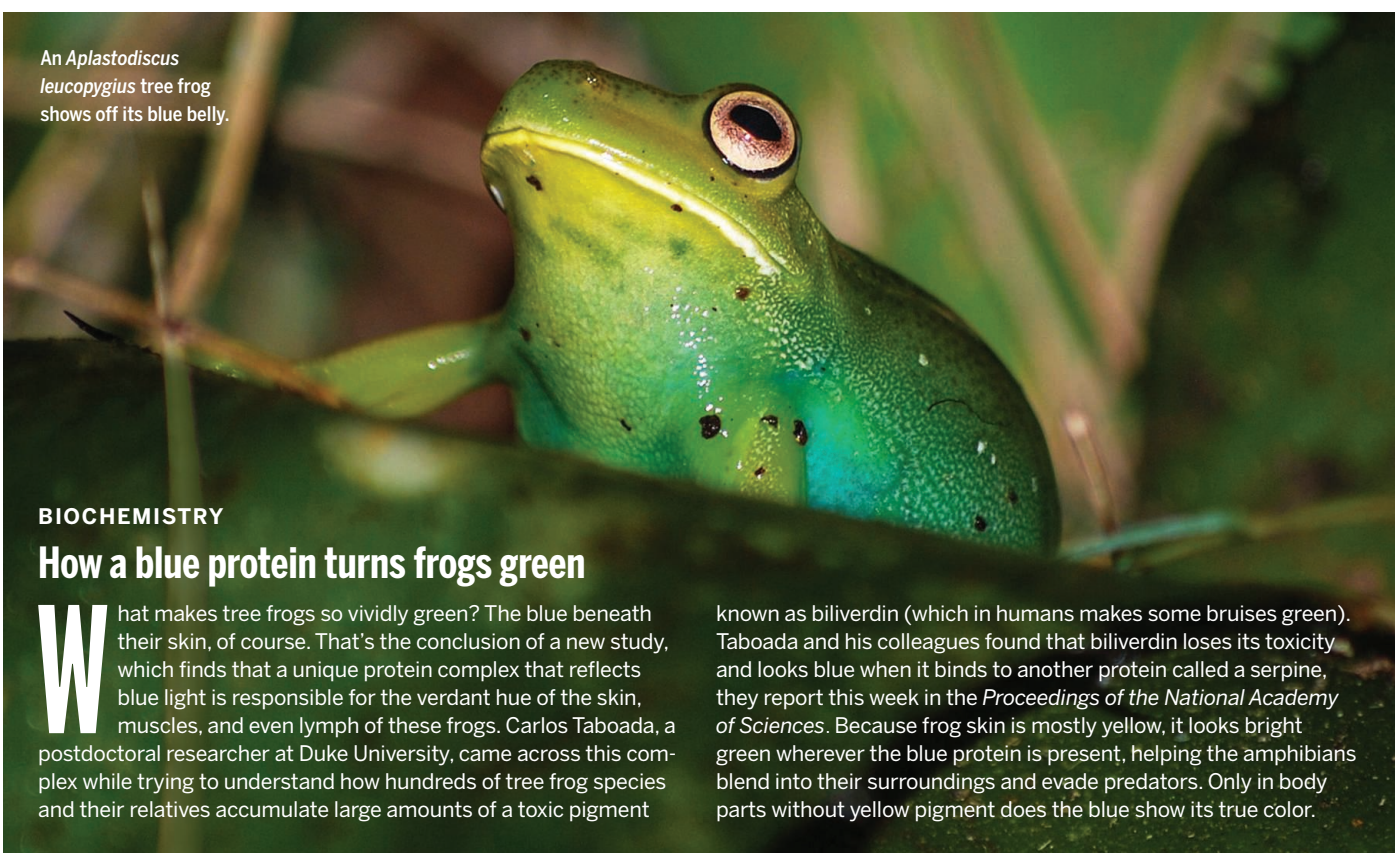
U.S. drops foreign student rule

HIGHER EDUCATION | The Trump administration agreed this week to drop its controversial proposal to prevent international students from staying in the United States if they are taking all their courses online. The government had argued the students didn't need to be in the country if all their coursework was offered remotely. But university officials protested the 6 July directive from the Department of Homeland Security, saying it would disrupt the education of hundreds of thousands of students and send a message that they weren't welcome on U.S. campuses. Last week, universities filed suit on the grounds that the administration had failed to follow federal laws in drawing up the rule, and on 14 July a federal judge in Boston announced that the government had agreed to rescind its directive. The settlement leaves intact a policy issued at the start of the coronavirus pandemic that allows foreign students to maintain their visas even if their courses are entirely online.

Who's to blame in 'Sharpiegate'

METEOROLOGY | The U.S. Department of Commerce succumbed to political interference from the White House and unnecessarily pressured the National Oceanic and Atmospheric Administration (NOAA) to rebuke its forecasters during 2019's Hurricane Dorian "Sharpiegate" scandal, Commerce's inspector general concluded in a report last week. The report details how Commerce's chief of staff, Michael Walsh, pushed NOAA's acting administrator, Neil Jacobs, to publish a critique of NOAA forecasters in Alabama who had correctly tweeted that Dorian posed no risk to the state, after President Donald Trump said otherwise. The letter damaged NOAA's apolitical reputation and could undermine public trust in its forecasts, the report found. Jacobs has since been nominated to serve as NOAA's administrator, and Walsh as Commerce's general counsel. The report is likely to increase Democratic scrutiny of Jacobs, making it possible that NOAA could go without a confirmed administrator for the entirety of the current presidential term.

SCIENCEMAG.ORG/NEWS
Read more news from *Science* online.



An *Aplastodiscus leucopygius* tree frog shows off its blue belly.

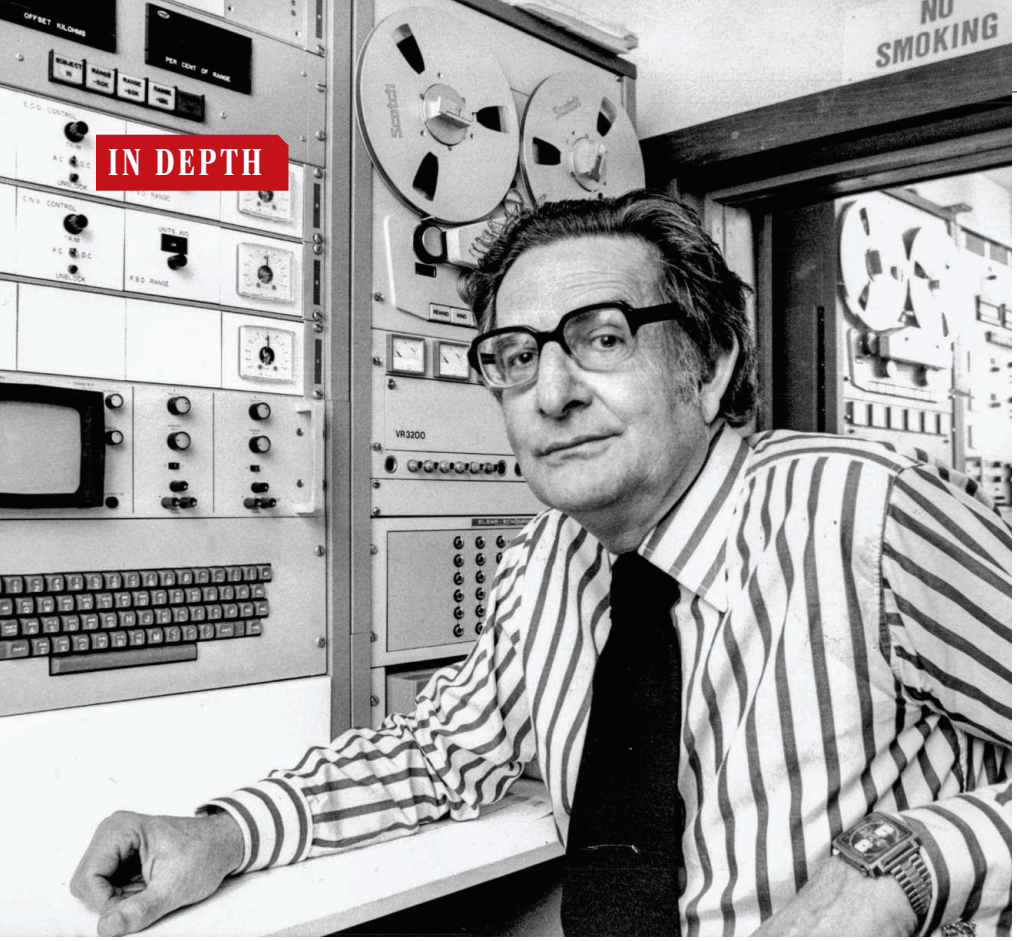
PHOTO: ANDRÉS E. BRUNETTI

BIOCHEMISTRY

How a blue protein turns frogs green

What makes tree frogs so vividly green? The blue beneath their skin, of course. That's the conclusion of a new study, which finds that a unique protein complex that reflects blue light is responsible for the verdant hue of the skin, muscles, and even lymph of these frogs. Carlos Taboada, a postdoctoral researcher at Duke University, came across this complex while trying to understand how hundreds of tree frog species and their relatives accumulate large amounts of a toxic pigment known as biliverdin (which in humans makes some bruises green). Taboada and his colleagues found that biliverdin loses its toxicity and looks blue when it binds to another protein called a serpine, they report this week in the *Proceedings of the National Academy of Sciences*. Because frog skin is mostly yellow, it looks bright green wherever the blue protein is present, helping the amphibians blend into their surroundings and evade predators. Only in body parts without yellow pigment does the blue show its true color.

Published by AAAS



Hans Eysenck reported that certain personality types have a 70-fold increased risk of dying from cancer.

PHOTO: NICK ROGERS/ASSOCIATED NEWSPAPERS/SHUTTERSTOCK

SCIENTIFIC INTEGRITY

Famous psychologist faces posthumous reckoning

Dozens of papers on personality and health by Hans Eysenck have been retracted or are under suspicion

By **Cathleen O'Grady**

One of Anthony Pelosi's most ambitious projects was on the back burner for more than 2 decades. In the early 1990s, Pelosi, a psychiatrist at Priory Hospital Glasgow, published two extensive critiques of the work of Hans Eysenck, a giant of 20th century psychology. Eysenck's papers contained questionable data and results so dramatic they beggared belief, Pelosi concluded. His critiques, and those by several others, were widely discussed in the field, but never led to formal investigations. Buried by the demands of clinical practice, research, and a young family, Pelosi never found the time to continue his effort. No one, he says, "picked up the baton."

More than a quarter-century later, Eysenck, who was celebrated for his theories of personality and individual differences, is finally falling from his pedestal. Last week, the *International Journal of Social Psychiatry* and the *Journal of the Royal Society of Medicine* issued expressions of concern for seven of his papers. Other journals have issued 64 such statements, as well as 14 retractions, over the past 6 months.

The renewed scrutiny comes in the wake of an inquiry by King's College London (KCL), where Eysenck was a psychology professor from 1955 to 1983 at what was then the Institute of Psychiatry. But Pelosi and others argue KCL failed to include many of Eysenck's other papers that also deserve a more thorough investigation in light of his lasting influence on the literature.

The case "throws up a lot of uncomfortable questions," says KCL neuroscientist Samuel Westwood. It's not clear whether the responsibility to investigate further lies with Eysenck's old institution, the journals that published his work, or a professional association, Westwood says. Still to be determined is the responsibility of Ronald Grossarth-Maticek, a physician and social scientist based in Heidelberg, Germany, with whom Eysenck co-authored the 25 papers KCL evaluated.

When he died in 1997, Eysenck was the third most cited psychologist in the world—behind Sigmund Freud and Jean Piaget. By then, he was already controversial, not just because of the criticisms by Pelosi and others, but also for espousing racist views on the genetics of intelligence.

Pelosi's interest was rekindled by an invitation to contribute to a 2016 special issue of the journal Eysenck founded, *Personality and Individual Differences*, to celebrate the centenary of his birth and his findings on personality and intelligence. "I think they thought I was going to write some kind of 'experts disagree' type article," Pelosi says. Instead, his manuscript summarized a litany of statistical and ethical criticisms he and others had raised. The journal deemed the paper inflammatory and did not include it in the special issue. In 2019, Pelosi found a new home for it in the *Journal of Health Psychology*, whose editor, David Marks, supported Pelosi's call for an investigation in an accompanying editorial.

Pelosi's critiques center on just one of Eysenck's many areas of research: the relationship between personality and health, specifically cancer and cardiovascular disease. This work, which mostly relied on data collected by Grossarth-Maticek in Germany and what was then Yugoslavia, showed "astonishing" evidence that "cancer-prone" and "heart-disease-prone" personality types exist, Pelosi writes. People with a cancer-prone personality had a risk of dying from cancer that was 40, 60, or even 70 times higher than that of people with a "healthy" personality, according to the duo. "These are unimaginably massive numbers in epidemiology," Pelosi says. Eysenck and Grossarth-Maticek also reported a clinical trial showing behavioral therapy could dramatically cut the death risk.

Other studies exploring the link between personality, stress, and health conditions have generally found that the various risk factors boost death risk by less than a factor of two. A large replication study in 2004 confirmed none of the links between personality and mortality reported in Eysenck's work, except for a modest association between cardiovascular disease and personality.

Several researchers have also reported evidence of errors and suspected data

Published by AAAS

manipulation in Eysenck's work with Grossarth-Maticek. Dutch medical psychologist Henk van der Ploeg reported in the 1990s that different versions of the data showed different dates and causes of death of research participants, suggesting they had been altered. Hermann Vetter, a statistician based in Germany, described data that show zero cases of lung cancer for "low-risk" personalities, with a rush of cancer cases appearing precisely at the point where the personality scores shift into a "high-risk" category. He concluded the data were "produced artificially … without pouring enough random error over it to make it appear more natural." Documents released in the course of litigation against tobacco companies—which funded some of Eysenck's work—show even some industry statisticians and researchers privately expressed doubt about the results.

In response to an interview request, Grossarth-Maticek, who is almost 80 but still offers counseling to people with cancer through his website, referred *Science* to a defense of his work posted on his website, which says the allegations are "untrue," "discriminatory," and "slander," and were made "without actual knowledge of the research program." He argues it would have been impossible to manipulate the data because they were given to other researchers to analyze before knowing the results, and denies that Eysenck's work was funded by the tobacco industry. He disputes their findings have defied replication, and claims KCL, a representative of "British and Jewish" psychology, didn't want "the little German Grossarth to dominate the scientific world stage."

Marks says there may be no "smoking gun," but the papers should be retracted anyway: "If they're so incredible, and have never been replicated, then we can dismiss those findings."

In response to Marks's editorial, KCL President Edward Byrne asked the university's Institute of Psychiatry, Psychology & Neuroscience to conduct an inquiry. In a report dated May 2019, a committee agreed with critics that the findings on personality and health were "incompatible with modern clinical science," designated 25 of Eysenck's papers as "unsafe," and asked Byrne to inform the relevant journals.

Since then, *Perceptual and Motor Skills* has retracted three of Eysenck's papers—two of which were not named in KCL's report—and *Psychological Reports* retracted 10. Both journals also added expressions of concern to dozens of Eysenck papers not included in the KCL inquiry. *Personality and Individual Differences* issued expressions of concern for three "unsafe" papers but declined to retract them because there was no evidence of "intentional deceit" and no admission of malpractice by the authors. But that's an "unrealistic bar" for retraction, says psychologist Simine Vazire of the University of Melbourne, editor-in-chief of *Collabra: Psychology*. Expressions of concern are "really ambiguous," whereas retracted papers are clearly no longer part of the scientific record, she says.

Marks and Pelosi say KCL's inquiry wasn't thorough enough. It only focused on papers Eysenck co-authored with Grossarth-Maticek, but missed work authored exclusively by Eysenck that relied on the same underlying data. Rod Buchanan, a historian who published a biography of Eysenck in 2010, has identified 87 publications he thinks should be retracted.

> **"These are unimaginably massive numbers in epidemiology."**
> **Anthony Pelosi**, Priory Hospital Glasgow

KCL's own researchers have also criticized a lack of transparency around their university's inquiry. The committee members were not named, their 2.5-page report lacked detail, and the university has rejected a freedom of information request by Westwood and other staff members. "It reflects badly on the institution," Westwood says. A KCL spokesperson says the committee investigated the publications that Marks's editorial designated as being of "immediate concern," and that it is standard policy to keep committee membership confidential.

One reason Pelosi says he wants a fuller investigation is that Eysenck's work still has an impact. A highly criticized 2008 meta-analysis that reported a link between stress and cancer included two dubious Eysenck papers; it has been cited more than 700 times, according to Google Scholar, including 50 times so far this year. Last month, a meta-analysis of studies investigating the effects of therapy on immune function in *JAMA Psychiatry* included one of the now-retracted papers. The idea that an upbeat personality can help people survive cancer permeates popular beliefs as well. "No epidemiologist takes [this work] seriously," Pelosi says, "but it does find its way into the scientific literature and I think it influences society."

A full investigation is warranted even if the work isn't particularly influential anymore, Vazire says. Otherwise, you can publish questionable research, "get superfamous, and then there are no consequences even when you get found out," she says. "That's terrible." ■

SCIENTIFIC FACILITIES

X-ray source gets a 100-fold boost in brightness

Rebuilt synchrotron is the first of more than a dozen of its kind in the works

By **Adrian Cho**

A brilliant new light shines in Grenoble, France, where officials at the European Synchrotron Radiation Facility (ESRF) last week announced the reopening of their completely rebuilt x-ray source. The ring-shaped machine, 844 meters around, generates x-ray beams 100 times brighter than its predecessor and 10 trillion times brighter than medical x-rays. The intense radiation could open up new vistas in x-ray science, such as imaging whole organs in three dimensions while resolving individual cells.

"The light is back at ESRF," said the lab's director general, Francesco Sette, at an 8 July online press conference. The reborn synchrotron, dubbed the Extremely Brilliant Source (EBS), will open to general users in late August, but since April, researchers have used its intense beams to study SARS-CoV-2, the virus responsible for the COVID-19 pandemic, and the disease's impact on the body. And the EBS is lighting the way for others, as the United States, Japan, and a dozen other countries develop similar machines.

A synchrotron is a ring-shaped accelerator that boosts charged particles such as electrons to high energies and near–light speed. Just as a wet rag flings droplets of water if you twirl it over your head, the circulating electrons radiate photons, including x-rays if the electrons have enough energy. In the 1950s, scientists began to siphon x-rays from electron accelerators built for particle physics experiments. Dedicated x-ray synchrotrons followed in the 1980s, employing magnets called wigglers to shake the electrons as they whirl around, causing them to produce more x-rays. In the 1990s, better synchrotrons debuted with magnets called undulators that shake the circulating electrons more harmoniously and effectively.

Published by AAAS



The ring-shaped synchrotron building in Grenoble, France, is unchanged, but physicists have replaced the machine inside with a radical new design.

Around the world dozens of synchrotrons of different energies now crank out light ranging from longer ultraviolet wavelengths to short "hard" x-rays, serving more than 55,000 users annually in fields ranging from chemistry and material sciences to geology and art history. Blasts of the ultraintense x-rays can probe the arrangement of atoms in a protein crystal even as they destroy it. Of the more than 166,000 3D protein structures biologists have deduced so far, more than 120,000 were determined at synchrotrons.

The trick to brightening ESRF's x-rays was to shrink the machine's already microscopic electron beam even further, says Pantaleo Raimondi, director of ESRF's accelerator and source division. The new machine will circulate a ribbonlike beam 2 micrometers high and 20 micrometers wide, one-thirtieth as wide as the old beam. To squeeze it, ESRF researchers followed a tack invented in the early 2000s and implemented at the lower energy MAX IV synchrotron in Sweden, which turned on in 2016.

In a synchrotron, magnets called dipoles sandwich the tubular vacuum chamber through which the electrons travel, supplying the vertical field that bends the particles' trajectory around the ring. The dipoles bend electrons slightly different amounts depending on their energies, causing the electron beam to spread. To keep it focused, more complex magnets called quadrupoles fit between the dipoles and act like lenses. But a quadrupole that focuses the electron beam horizontally spreads it vertically and vice versa, so the beam expands and contracts like an accordion as it circulates. MAX IV physicists realized they could reduce those oscillations by replacing longer dipoles with a larger number of shorter ones and more quads.

There was a catch, however. The quads also focus electrons differently depending on their energy. Even more complex magnets called sextupoles can correct for that effect. But if the beam is already narrow, the sextupoles work inefficiently and the scheme conks out before reaching maximal compression. MAX IV researchers sidestepped that problem by reducing the energy of their electrons, but ESRF couldn't do that and still produce hard x-rays. In 2008, Raimondi and colleagues found a way out of the dilemma: They would arrange the gauntlet of magnets so that the electron beam would spread out briefly as it passed through the sextupoles.

After shutting down their machine in December 2018, ESRF workers replaced almost all of its components in just 13 months at a cost of €150 million. Whereas the original machine had two long dipoles in each of its 32 segments, or arcs, the new one has seven, plus 24 other magnets. All told, more than 1000 new magnets were installed in the same doughnut-shaped hall as before. "The body of the car remains the same, but we took out the old motor and put in the engine of a Ferrari," Raimondi says.

The rebuilt machine should open up qualitatively new windows in x-ray science, says Harald Reichert, ESRF's director for research in the physical sciences. Hard x-rays can penetrate materials far more deeply than lower energy x-rays, and the new machine's intense x-ray beams will enable it to study samples up to 1 meter thick. So, scientists could scan an engine block and then zoom in on material defects with near atomic resolution, Reichert says.

Because the x-ray photons emerge from such a tiny electron beam, they should oscillate in unison like those in laser light, accentuating the wavelike nature of the x-ray beam. That enhanced coherence gives the new ESRF a big advantage for imaging. When researchers shine an x-ray beam through a sample, variations in the material will delay the wave front of the coherent x-rays to different degrees, creating a mottled intensity pattern on a distant detector. From many such patterns, researchers can extract a detailed 3D image of the sample.

For example, neuroscientists striving to map the individual neurons and their interconnections in the mouse brain must now cut a brain into fine slices, scan each slice with an electron microscope, and then use massive computer power to link the slices. With ESRF's penetrating, coherent beams, researchers may be able to do the same thing far faster in an intact mouse brain. "If they can do that, it's going to have a huge impact on the community," says Eva Dyer, a computational neuroscientist at the Georgia Institute of Technology. "It's supercool."

ESRF will have a few years to press its advantage. Its nearest competitor, the Advanced Photon Source (APS) at Argonne National Laboratory in Illinois, will undergo a similar yearlong rebuild in 2022, says APS Director Stephen Streiffer. "We see the two machines not so much as competitors but as sister facilities," Streiffer says. "There's plenty of discoveries to go around."

The RIKEN Spring-8 laboratory in Japan and the German Electron Synchrotron laboratory are also planning rebuilds, and China is designing a brand-new x-ray facility.

Researchers are already exploiting ESRF's new capabilities to help confront the COVID-19 pandemic. Several projects are studying the molecular structure of the SARS-CoV-2 virus, and another is using the new synchrotron to image lungs damaged by COVID-19, says Jean Susini, ESRF's director for research for life sciences. "In fact, the EBS's first light was used in COVID-19 research." ■

PHOTO: ESRF/P. JAYET

Published by AAAS

POLICY

An optimist takes the helm at NSF

Sethuraman Panchanathan has spent his career finding "win-win" situations

By **Jeffrey Mervis**

The new director of the National Science Foundation (NSF), Sethuraman "Panch" Panchanathan, knows that dark political clouds are hanging over the agency.

His boss, President Donald Trump, keeps proposing big cuts to NSF's budget. In Congress, Democrats and Republicans alike have lambasted NSF for not doing more to stop the Chinese government from harvesting the fruits of U.S.-funded research. New reports of sexual harassment, including some involving NSF grantees, seem to pop up weekly, and the national outcry over racial inequities has highlighted the chronic underrepresentation of minorities and women in science.

Yet the Indian-born computer scientist was unabashedly upbeat about the future of the U.S. academic research community—where he has spent most of his career—and NSF's role in supporting it during a 2 July remote interview with *Science*. He was just over a week into his 6-year term and was in the process of moving to the Washington, D.C., area.

"If I were advising a family member about whether or not to come here, I would tell them that the United States is still the land of opportunity and the best place to pursue science," says Panchanathan, 59, who has spent the past 23 years at Arizona State University (ASU), Tempe, most recently as executive vice president and head of research and innovation. "And I'm a living example of that."

His view of NSF's budget prospects is similarly optimistic. Asked about the president's repeated attempts to shrink NSF's budget, Panchanathan instead focused on how legislators have repeatedly come to its rescue. "The budget process involves both the administration and Congress, as I understand it," he says. "And I am very confident that we'll have a positive reception if we continue to communicate the value of federal investments in fundamental research." (This week, the House of Representatives's spending panel defied a 6% cut proposed by the White House for NSF in 2021 and approved a 3% increase.)

Panchanathan, who likes to be called Dr. Panch to avoid the otherwise almost inevitable mangling of his name by legislators, has followed a path taken by countless Asian-born scientists of his generation. He earned undergraduate degrees from top-ranked Indian universities before heading west, in his case to Canada, for his Ph.D. in electrical and computer engineering from the University of Ottawa. He spent 8 years as a faculty member there before ASU recruited him in 1997. In 2001 he founded the Center for Cognitive Ubiquitous Computing, which develops technology for people with disabilities. For the past decade he has helped ASU President Michael Crow aggressively promote more interdisciplinary research, partnerships with industry, and online courses to foster lifelong learning.



As a top official at Arizona State University, Sethuraman Panchanathan aided its rise to prominence.

"The thing about Panch is that he never makes enemies," says Morris Goldberg, his doctoral adviser at the University of Ottawa and a longtime collaborator at the center. "His goal is always to bring everything to a win-win situation."

The current debate in Congress over foreign influences on federally funded research will test those skills. Senator Rob Portman (R–OH), who has proposed more stringent oversight over that research, has accused NSF and other agencies of "being asleep at the switch."

Although he declined to comment on Portman's bill, Panchanathan envisions a way to satisfy both those calling for a crackdown and those worried that tighter controls could strangle U.S. innovation. "I hold dearly to the traditional scientific values of transparency and openness," he says. "But if there are folks who are not reciprocating in that way, we have to take appropriate measures to protect what we think is most important."

Panchanathan says he supports the conclusions of a report on foreign influences that his predecessor, France Córdova, commissioned last year from Jason, an independent group of scientists that advises the government on national security issues. It rejected the need for additional controls on fundamental research deemed sensitive and endorsed a 1985 presidential directive that relies on classifying any sensitive research. "The Jason report makes a lot of sense to me," he says.

So does a proposal by the Senate's top Democrat, Chuck Schumer (NY), that would supersize NSF's current $8 billion budget so it could pour money into the basic research needed to seed a handful of so-called "industries of the future." Although Panchanathan didn't embrace Schumer's plans for a new technology directorate to manage the huge investment, he dismisses the fears of many scientists that such a realignment would threaten NSF's traditional mission to support academic science.

"This approach is going to make the research enterprise very robust and very vibrant," he says. "When you engage in user-inspired research, the fundamental research questions that you ask yourself become different."

The Trump administration's efforts to restrict immigration pose a challenge that strikes particularly close to home. Panchanathan neither endorsed nor questioned those policies, which he says fall outside his purview. But he urged foreign scientists not to abandon their dreams of coming to the United States.

"When you do a Ph.D., it's understood that you're a smart person," he says. "You understand the context and why things have been done in a particular way. So you need to navigate [that system] to find the best way of getting to where you need to get to."

Panchanathan will now get a chance to test whether his advice to academics also works in a politically fraught environment. ■

PHOTO: ARIZONA STATE UNIVERSITY

Published by AAAS



PHOTO: DAVID RYDER/REUTERS

Researchers say testing sites in low-income areas, like this one in Interbay Village in Seattle, are crucial.

COVID-19

'Huge hole' in testing data blurs racial, ethnic disparities

Researchers seek better data on pandemic's uneven burden

By **Kelly Servick**

At a virtual meeting last month, cardiologist Garth Graham shared data that shine a narrow flashlight beam into a vast, shadowy crisis. As vice president of community health at CVS Health, he oversees nine centers offering free, rapid COVID-19 tests in low-income neighborhoods with high proportions of racial and ethnic minorities. From those sites, the view of the pandemic is dire. Roughly 35% of tests performed at a center in Phoenix and 30% in Atlanta had come back positive as of 23 June, he said. Nationwide, only 8.8% of tests were positive in the week ending 4 July, according to the U.S. Centers for Disease Control and Prevention (CDC), although Arizona reported a weekly 19% positive rate on 12 July.

The CVS test results suggest "the pandemic is unfolding very differently in Black and brown communities," Graham said.

"Those were extraordinary numbers," agrees Gary Gibbons, director of the U.S. National Heart, Lung, and Blood Institute, who also spoke at the meeting—a workshop on COVID-19 and Black communities organized by the National Academies of Sciences, Engineering, and Medicine (NASEM). The results are "a yellow flag that those areas ought to be targeted for more intense testing and tracing."

To understand which neighborhoods and communities face raging outbreaks and why, researchers need demographic information, including race and ethnicity, on who gets tested and who tests positive. But even as the U.S. testing infrastructure improves, testing remains sparse in many low-income and minority neighborhoods, and race and ethnicity information is missing for about half of reported COVID-19 cases nationwide.

"It's a huge hole in my research," says Sarita Shah, an epidemiologist at Emory University who studies racial disparities in severe COVID-19.

Hospital and death records show that Black people, Latinos, and Native Americans are disproportionately suffering and dying from severe disease. Nationwide, Black people are dying at 2.5 times the rate of whites. But without knowing the race of everyone tested, Shah can't say whether high death rates among Black people are driven primarily by more exposure and infection, diagnosis at a later stage of disease, or a higher risk of severe illness once infected.

Understanding these drivers might inform solutions, whether boosting accessibility to testing, improving public health messaging in hard-hit communities, or aiming for earlier diagnosis and more supportive care. "The lack of data speaks a lot to whose lives we value and whose lives we don't," says Richard Besser, president and CEO of the Robert Wood Johnson Foundation, who also spoke at the NASEM meeting.

Some of the missing testing data might be on the way. Federal guidance released last month will require laboratories to report race and ethnicity, along with age, sex, and ZIP code, to the government with COVID-19 test results starting 1 August. "Those data should have been collected from the start," says Nancy Krieger, a social epidemiologist at the Harvard T.H. Chan School of Public Health. She notes that the form used to report COVID-19 cases to CDC already has fields for race and ethnicity, but they're frequently left blank. It's not clear whether and how the new mandate will be enforced, she says.

Even when test sites do collect detailed demographic information, barriers to testing in vulnerable communities can be high, Besser notes. Low-income people who lack health insurance may fear unexpected fees from testing, or worry about losing income if they test positive and can't go to work.

And test sites may simply be scarce in their neighborhoods. Data from New York City suggest that early in the pandemic, testing failed to reach those who needed it most. Among the city's 177 ZIP codes, the number of tests done between 2 March and 6 April rose with the percentage of white residents in a ZIP code, whereas the proportion of positive tests decreased with the white population, epidemiologist Emanuela Taioli and colleagues at the Icahn School of Medicine at Mount Sinai reported on 25 June in the *American Journal of Preventive Medicine* (*AJPM*).

In unpublished follow-up analyses, Taioli has seen progress as the city set up more test sites in underserved areas. By mid-May, Taioli says, testing rates didn't correlate with the racial makeup of a ZIP code. But when her team looked to expand its analysis beyond New York, they learned that many states don't publish how many people they test—just how many test positive. That makes it harder to know whether groups are undercounted.

Many states now have efforts to bring mobile testing sites to neglected neighborhoods. And a National Institutes of Health initiative called RADx-UP will devote $300 million in its first phase to setting up and coordinating data collection across test sites in underserved and minority communities.

Other teams are finding ways to work around the data gaps, combining census data with whatever location data are available for

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

Published by AAAS

positive cases. On 18 June, a team at Emory released a "health equity interactive dashboard" that displays each U.S. county's case and death counts alongside existing government records, including the racial and ethnic makeup of the county, how many residents are uninsured or live in poverty, and how many are obese or have diabetes. Nationwide, the dashboard shows a strong correlation between Black populations and COVID-19 deaths; and between Latino populations and confirmed cases, says Shivani Patel, an Emory epidemiologist who leads the project.

The reasons for the disease's uneven burden trace back to systemic racism, which creates a web of disadvantages for Black and brown people. People of color face higher rates of diabetes, heart disease, and other conditions that worsen COVID-19 prognosis. They are also more likely to live in crowded neighborhoods and have low-wage, essential jobs where social distancing is difficult or impossible.

Vulnerable communities aren't all vulnerable in the same way, Ibraheem Karaye, an epidemiologist at the University of Delaware, found in his own county-by-county analysis. In a study published 26 June in *AJPM*, he and University of Delaware epidemiologist Jennifer Horney compared each U.S. county's COVID-19 case count with its score on CDC's 15-factor social vulnerability index. In northeastern states, residents' minority status and language best explained infection rates. In Gulf Coast states, in contrast, factors related to housing and transportation—such as lacking a vehicle or living in multiunit structures—predicted infections better than ethnic factors did. Karaye doesn't know why those differences exist, but says they hold clues about the kind of help communities in these regions might need.

Still, mashups of census data and test results can only go so far. Without knowing the race of the individuals tested, Patel notes, she can't say whether it's really the Latinos getting hit hardest in a majority-Latino area. Nor can she capture a situation where Black people living in a majority-white county are sickened disproportionately.

Shah, who has been volunteering with the local county's health department, sees the data collection problem firsthand. After doing nasal swabs at a drive-up testing site, she calls those who test positive to fill in personal information, including race. Even after multiple attempts, the team reaches only about half of these people. Shah says she'd love to note a person's race when they're sitting in front of her at the test site, but so far, the forms provided by labs that process the samples simply don't have a place to note it. "I wish it was something more complex than that," she says, "but it's not." ■



Data from Dutch mink farms hit by the coronavirus may shed more light on a mutation that has spread widely.

PHOTO: FABIAN BIMMER/REUTERS/NEWSCOM

COVID-19

The pandemic virus is slowly mutating. But does it matter?

Linking genomic changes to viral behavior proves difficult

By **Kai Kupferschmidt**

It's only a tiny change. At some point early in the pandemic, one of the 30,000 letters in the genome of SARS-CoV-2 changed from an A to a G. Today, that mutation, at position 23,403, has spread around the world. It is found in the vast majority of newly sequenced viruses and has become the center of a burning scientific question: Has the mutation become so common because it helps the virus spread faster? Or is it just coincidence?

More than 6 months into the pandemic, the virus' potential to evolve in a nastier direction—or, if we're lucky, become more benign—is unclear. In part that's because it changes more slowly than most other viruses, giving virologists fewer mutations to study. But some virologists also raise an intriguing possibility: that SARS-CoV-2 was already well adapted to humans when it burst onto the world stage at the end of 2019, having quietly honed its ability to infect people beforehand.

On average, the coronavirus accumulates about two changes per month in its genome. Sequencing SARS-CoV-2 genomes helps researchers follow how the virus spreads. Most of the changes don't affect how the virus behaves, but a few may change the disease's transmissibility or severity.

One of the earliest candidates was the wholesale deletion of 382 base pairs in a gene called *ORF8*, whose function is unknown. First reported by Linfa Wang and others at the Duke-NUS Medical School in Singapore in a March preprint, the deletion has since been reported from Taiwan as well. A deletion in the same gene occurred early in the 2003 severe acute respiratory syndrome (SARS) outbreak, caused by a closely related coronavirus; lab experiments later showed that variant replicates less efficiently than its parent, suggesting the mutation may have slowed the SARS epidemic. Cell culture experiments suggest the mutation does not have the same benign effect in SARS-CoV-2, Wang says, "but there are indications that it may cause milder disease in patients."

The mutation at position 23,403 has drawn the most attention—in part because it changed the virus' spike, the protein on its surface that attaches to human cells. The mutation changed the amino acid at position 614 of the spike from an aspartic acid (abbreviated D) to a glycine (G), which is why it's called G614.

Published by AAAS

In a *Cell* paper this month, Bette Korber and colleagues at Los Alamos National Laboratory showed that G614 has become more common in almost every nation and region they looked at, whereas D614 is virtually gone (see graphic, below). That might be a sign that it's outcompeted by G614, but it could also be a coincidence. "Any one mutation may rise to very high frequency across the world, just because of random chance," says Kristian Andersen, a computational biologist at Scripps Research. "This happens all the time."

Comparing the behavior of multiple viral variants carrying one mutation or the other could reveal a difference. The United Kingdom's COVID-19 Genomics Consortium has sequenced 30,000 SARS-CoV-2 genomes, allowing scientists to compare how fast 43 lineages carrying the G614 mutation and 20 with D614 spread. They estimated that the former grew 1.22 times faster than the latter—but the statistical significance was low. "Evidence for a difference is weak and if it does exist, the estimated effect is moderate," says evolutionary biologist Andrew Rambaut of the University of Edinburgh.

Researchers have also turned to cell culture experiments. When Korber's group engineered so-called viruslike particles to carry one spike protein or the other, the G614 variant appeared to be more efficient at entering cells. Jeremy Luban of the University of Massachusetts Medical School, who has found the same thing, explains that G614 causes a slight change in the shape of the spike, apparently making it easier for the protein to undergo the structural changes that cause the membranes of the virus and the cell to fuse. "Our data looks like it's somewhere between three and 10 times more infectious," Luban says. "That's a pretty enormous effect."

That does not mean the mutation has an effect in the real world, says virologist Emma Hodcroft of the University of Basel. In the past, she notes, "We have cases where we really thought that we had evidence for a mutation that was changing viral behavior and as more evidence came, it didn't seem to be the case." An increased ability to infect a laboratory cell line may not translate to the billions of diverse cells in a human body, adds Angela Rasmussen, a virologist at Columbia University: "Humans aren't Vero cells."

Animal experiments are another way to probe the effects of G614. One option, virologist Marion Koopmans of Erasmus Medical Center (EMC) says, would be to infect ferrets with it and D614 and look for differences in how much virus they shed. But infections in ferrets only last about 1 week, Koopmans notes. "The effect would have to be very big to show up in an experiment like that."

Another idea is to expose uninfected ferrets to animals carrying either of the two variants and see how well they transmit. An uncontrolled transmission experiment has already taken place on Dutch mink farms, where the new coronavirus jumped from humans to minks at least five separate times. Twice it was the D614 variant, and three times G614, Koopmans says. She hopes data from the outbreaks could show whether either one spread faster and wider than the other. But the experiment doesn't have the rigor of a lab study, she concedes. "We have a natural experiment here. The study design is not optimal."

The attention lavished on G614 may obscure a bigger question, however: With the virus having spread to at least 11 million people worldwide, why aren't more mutations that affect its behavior emerging?

Perhaps there's just little selection pressure on the virus as it races through millions of immunologically naïve people, scientists say. That could change with the advent of vaccines or new therapies, forcing the virus to evolve. But it could also indicate that the virus has been with people longer than we know, and was spreading before the first known cases in Wuhan, China, in December 2019. "The evolution of this virus to become a human pathogen may have already happened and we missed it," Rasmussen says.

Wang thinks a version of the virus may have circulated earlier in humans in southern Asia, perhaps flying under the radar because it didn't cause severe disease. "If it happens in a small or remote village, even with some people dying, nobody is going to know there's a spillover," Wang says. The virus could then have infected an animal that was brought to Wuhan and started the pandemic.

At Dutch mink farms, after all, the virus jumped not just from humans to animals, but also back from animals to humans, Wang says. "If that can happen in the Netherlands, surely it can happen in a village in Thailand, or in Yunnan province in southern China." ■

A change for the worse?

A one-letter mutation, rare in February but found in almost every newly sequenced genome today, may make SARS-CoV-2 more transmissible.



GRAPHIC: KORBER *ET AL*., *CELL*, DOI: 10.1016/J.CELL.2020.06.043. ADAPTED BY N. DESAI/*SCIENCE*

COVID-19

Courses bring field sites and labs to the small screen

Some pandemic-induced changes to hands-on classes may be here to stay

By **Elizabeth Pennisi**

In a normal summer, Appledore Island, a 39-hectare outcrop 12 kilometers off the coast of Maine and New Hampshire, becomes a classroom. Students from high school to graduate level live in close quarters, eat in a communal dining hall, and work shoulder to shoulder to explore the biology of the shore and waters in 18 courses organized by the Shoals Marine Laboratory. But this summer, with the pandemic surging, students have stayed home.

Instead, a skeleton staff on Appledore is streaming field trips and dissections of fish and invertebrates and setting up cameras to gather data for students. Rather than leading students around the island, coastal restoration ecologist Gregg Moore from the University of New Hampshire (UNH), Durham, hauls a backpack full of equipment: "a dual modem with two different cellular carriers, a signal-boosting directional antenna, and a large DC power source," he says. The equipment allows him to teach 12 remote students—twice the course's usual enrollment—basic techniques of coastal ecology.

Moore's is just one of hundreds of lab and field courses forced online by COVID-19—"a seismic shift for those who were not already involved in distance or online education," says Martin Storksdieck, a science education researcher at Oregon State University, Corvallis. Some researchers worry students will miss out on certain practical and problem-solving skills and won't be able to judge whether the hands-on work of a scientist is a good fit for them. But instructors are developing high-tech ways to simulate the field and lab experiences. "I would say [these courses] are *not* virtual," says Jennifer Seavey, director of the Shoals lab. "They are real." And some advantages are emerging. By lowering geographical and financial barriers, Seavey says, "Virtual

Published by AAAS

field courses are democratizing fieldwork."

The shift has taken ingenuity. "Professors must get creative and use a combination of what is available," including online videos and free or commercially available online labs, says Mildred Pointer, a physiologist at Howard University who is working on a fall course in general biology. No single tool meets all their needs, Pointer says.

As the pandemic gained momentum, emails flew among the leaders of the National Association of Geoscience Teachers. Many U.S. geology majors must take a "capstone" field course to graduate. The cancellation of more than three-quarters of these courses jeopardized graduation for many majors. So the association invited instructors to develop learning objectives that did not depend on students doing fieldwork. It also compiled online exercises to help the 29 field courses that have moved online this summer. Lessons range from "Orienteering in Minecraft" to "Geology of Yosemite Valley," which includes a 43-stop Google Earth tour with photos and embedded text.

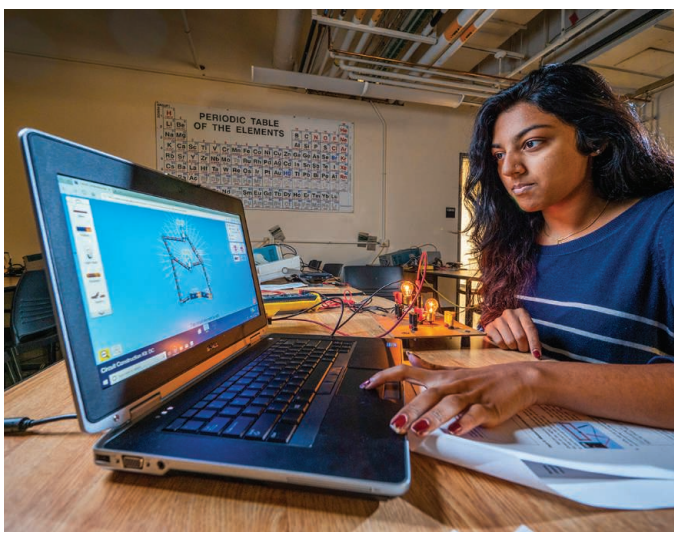
Like Moore, geoscientist Jim Handschy wanted to give remote students "as close to the real experience as possible." He runs Indiana University's Judson Mead Geologic Field Station in Montana, which had enrolled 60 students before classes were canceled in March. He and a few instructors visited each outcrop in their course plan, filmed the rocks and landscape, and captured magnified views of samples. Each week, the class delves deeper into the rock layers and their history. For their final project, students digitally map a 3100-hectare landscape.

Shannon Dulin, a geologist at the University of Oklahoma, Norman, who just finished teaching a field course, sees the value of learning how to survey a landscape without setting foot on it. On their class evaluations, her students said they gained unexpected skills. "And these are skills they are going to need on the job," she adds, as geologists are increasingly being asked to evaluate sites they don't visit.

In other fields, hands-on learning takes place in labs. Typically, students work in pairs and share equipment, "so there are a lot of issues about virus transmission," says Heather Lewandowski, a physicist at the University of Colorado (CU), Boulder. At her university this fall, lab exercises as diverse as building an electrical circuit or analyzing solar flare data will most likely be completely remote.

Luckily, physics already had a foot in the virtual lab world—especially at CU. There, back in 2002, Nobel laureate Carl Wieman developed the Physics Education Technology (PhET) Interactive Simulations project to provide "games" that teach students basic physics concepts. The PhET web portal now has 106 physics-based simulations and another 50 or so for other disciplines. It became a go-to place this spring for faculty shifting to online teaching; traffic increased fivefold, says Director Katherine Perkins.

In addition, several universities have adopted a handheld device called the iOLab that rents for $50 a semester. With it, students can measure magnetism, light intensity, acceleration, temperature, gravity, and atmospheric pressure, and do basic physics experiments at home. "They like that we trust them and are not just giving them instructions," says iOLab inventor and physicist Mats Selen at the University of Illinois, Urbana-Champaign.



A University of Colorado student uses an online simulator to design a circuit.

Lewandowski and her colleagues surveyed physics instructors and students about their experiences and posted their findings on arXiv, the physics preprint server, on 2 July. Respondents said online labs work best when projects are open-ended, and online class meetings are kept small. They complained about technical difficulties, students having unequal access to the internet and materials, and longer prep times for both students and instructors. But they reported they could meet most key learning objectives, Lewandowski says, even though "there are lots of things we can't replicate in remote experiments," such as such as building vacuum chambers or troubleshooting equipment.

Some institutions decided this spring that virtual just wouldn't do. The Marine Biological Laboratory (MBL) in Woods Hole, Massachusetts, simply canceled its summer courses. "MBL courses are world-renowned for the intensity of the hands-on nature of the lab work," says Director Nipam Patel. Students spend long hours with famous faculty and do their own projects using organisms collected locally. "We felt that it would be exceedingly difficult to replicate these experiences as a virtual lab course."

Other institutions will try for a mix of in-person and virtual labs. Suely Black, chemistry chair at Norfolk State University, expects only half of his students will be in lab each week this fall, while the other half will be in online classes analyzing data and writing reports. "The crisis has caused us to more critically evaluate what activities students must experience in the lab setting," he says.

Similarly, this fall, organic chemistry students at the University of Michigan (UM), Ann Arbor, will rotate into the lab in small groups, giving each a taste of the hands-on experience. Personal protection equipment is standard for this course and all the work is done in hoods with excellent air exchange, so "they are really fully protected," says UM biochemist Kathleen Nolta.

Storksdieck, an advocate of online learning, questions the value of smelling fumes or using a pipette. "We have to ask whether all the hands-on taught so far was all that great," he says. Dominique Durand, a biomedical engineer at Case Western Reserve University, says after he put a master's program in biomedical engineering completely online 5 years ago, he concluded that solving problems was more important than hands-on experience. And University of California, Santa Cruz, ecologist Erika Zavaleta thinks virtual courses will open fieldwork to far more students. "There are things you can do online that you can't do in person," she adds, such as visiting more places than possible by driving.

Even so, Handschy laments that his geology students will not have the 12-hour-a-day immersive interactions with each other and faculty that past classes have had. Natalie White, a rising junior at UNH who took Moore's course on Appledore last year, agrees: "You don't have all the time in between when you walk around the island and can ask impromptu questions." Appledore Island is the source of some her fondest memories. "I think they are missing out on the community." ■

PHOTO: GLENN J. ASAKAWA/UNIVERSITY OF COLORADO, BOULDER

Published by AAAS



In Turin, Italy, a teacher reads to students in a makeshift outdoor classroom.

PHOTO: NICOLÒ CAMPO/LIGHTROCKET VIA GETTY IMAGES

NOT OPEN AND SHUT

School openings across the globe suggest ways to keep the coronavirus at bay, despite outbreaks

By **Jennifer Couzin-Frankel**, **Gretchen Vogel**, *and* **Meagan Weiland**

In the spring, school gates around the world slammed shut. By early April, an astonishing 1.5 billion students were staying home as part of broader shutdowns to protect people from the novel coronavirus. The drastic measures worked in many places, dramatically slowing the spread of SARS-CoV-2, the virus that causes COVID-19.

However, as weeks turned into months, pediatricians and educators began to voice concern that school closures were doing more harm than good, especially as evidence mounted that children rarely develop severe symptoms from COVID-19. (An inflammatory condition first recognized in April, which seems to follow infection in some children, appears uncommon and generally treatable, although scientists continue to study the virus' effect on youngsters.)

Continued closures risk "scarring the life chances of a generation of young people," according to an open letter published last month and signed by more than 1500 members of the United Kingdom's Royal College of Paediatrics and Child Health (RCPCH). Virtual education is often a pale shadow of the real thing, and it left many parents juggling jobs and child care. Lower

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

Published by AAAS

income children who depend on school meals were going hungry. And there were hints that children were suffering increased abuse, now that school staff could no longer spot and report signs of it. It was time, a growing chorus said, to bring children back to school.

By early June, more than 20 countries had done just that. (Some others, including Taiwan, Nicaragua, and Sweden, never closed their schools.) It was a vast, uncontrolled experiment.

Some schools imposed strict limits on contact between children, whereas others let them play freely. Some required masks, whereas others made them optional. Some closed temporarily if just one student was diagnosed with COVID-19; others stayed open even when multiple children or staff were affected, sending only ill people and direct contacts into quarantine.

Data about the outcomes are scarce. "I just find it so frustrating," says Kathryn Edwards, a pediatric infectious disease specialist at the Vanderbilt University School of Medicine who is advising the Nashville, Tennessee, school system, which serves more than 86,000 students, on how to reopen. Her research assistant spent 30 hours hunting for data—for example on whether younger students are less apt to spread the virus than older ones, and whether outbreaks followed reopenings—and found little that addressed the risk of contagion in schools.

When *Science* looked at reopening strategies from South Africa to Finland to Israel, some encouraging patterns emerged. Together, they suggest that a combination of keeping student groups small and requiring masks and some physical distancing helps keep schools and communities safe, and that younger children rarely spread the virus to one another or bring it home.

But opening safely, experts agree, isn't just about the adjustments a school makes. It's also about how much virus is circulating in the community, which affects the likelihood that students and staff will bring COVID-19 into their classrooms.

"Outbreaks in schools are inevitable," says Otto Helve, a pediatric infectious disease specialist at the Finnish Institute for Health and Welfare. "But there is good news." So far, with some changes to schools' daily routines, he says, the benefits of attending school seem to outweigh the risks—at least where community infection rates are low and officials are standing by to identify and isolate cases and close contacts.



PHOTO: ISSOUF SANOGO/AFP VIA GETTY IMAGES

How likely are children to catch and transmit the virus?

Studies have found that people under age 18 are between one-third and one-half as likely as adults to contract the virus, with the risk lowest for the youngest children. In the town of Crépy-en-Valois, home to 15,000 people on the northern outskirts of Paris, researchers found that younger children also seem less likely to transmit it to others.

When two high school teachers developed minor respiratory symptoms in early February, no one suspected COVID-19. It was cold and flu season, and health officials still assumed the novel coronavirus was mostly confined to China. It wasn't until 25 February, after a contact was hospitalized in Paris, that the teachers realized they had been infected with SARS-CoV-2. For at least 12 days before the start of winter break on 14 February, and before France instituted precautions, the virus had been spreading freely at the school.

Arnaud Fontanet, an epidemiologist at the Pasteur Institute, and colleagues started an investigation in Crépy-en-Valois in late March to piece together the virus' reach in the town and its schools. In the high school, antibody testing showed that 38% of pupils, 43% of teachers, and 59% of nonteaching staff had been infected. (By then, several people associated with the school had been hospitalized with COVID-19 complications.) In six elementary schools, they found a total of three children who had caught the virus, likely from family members, and attended school while infected. But as far as the researchers could tell, those younger children didn't pass the virus on.

"It's still a bit speculative," says Fontanet, who shared results from the high school on 23 April and from the elementary schools on 29 June, both on the preprint server medRxiv. But high school students "have to be very careful. They have mild disease, but they are contagious." Children younger than 11 or 12, on the other hand, "probably don't transmit very well. They are close to each other in schools, but that is not enough" to fuel spread. At the same time, scientists note that children have more contacts than adults, especially at school, which could offset the lower odds they will spread the pathogen.

Other outbreaks also suggest elementary school pupils pose a smaller threat than older students. Among the worst schoolwide outbreaks was at Gymnasium Rehavia, a middle and high school in Jerusalem, where 153 students and 25 staff were infected in late spring. An outbreak at a New Zealand high school before the country's shutdown infected 96 people, including students, staff, and parents. Yet a neighboring elementary school saw few cases.

Published by AAAS



Children were ready as school reopened in Abidjan, Ivory Coast, in May.

But the picture is still blurry. Another Israeli outbreak was in an elementary school in Jaffa, with 33 students and five staff members affected. In an elementary school classroom in Trois-Rivières, Canada, nine of 11 students became infected.

Other data come from day care centers: In many countries, they stayed open for children of essential workers, and outbreaks appeared rare. Two flares in Canada—one in Toronto, and one outside of Montreal—led to temporary closures. In Texas, where overall cases have skyrocketed, at least 894 preschool staff and 441 children across 883 facilities have tested positive, according to news reports.

Tracing transmission through schools, one student at a time—as Fontanet and his colleagues did—should help show whether the virus tracks differently in children of different ages. Another clue about age-based spread came from Crépy-en-Valois's timeline of new infections. Among high school students and staff, new infections dropped off sharply once winter break started. But in the elementary schools, the (already low) rate of new cases held steady. Fontanet says that pattern suggests that whereas high schoolers were catching the virus at school, the younger pupils caught it from family members and not their classmates.

PHOTO: DAN KITWOOD/GETTY IMAGES

Should children play together?

The scenes hardly resemble typical school: preschool children instructed to spend recess playing alone inside a chalk square. Eight-year-olds told not to speak to their friends. Middle schoolers reminded to steer clear of classmates when entering or leaving the building.

As schools reopened, many embraced physical distancing for students to prevent viral spread. But although the strategy is effective, it is leaving more and more scientists, pediatricians, and parents deeply uncomfortable. They hunger for a compromise that protects communities from COVID-19 while supporting the mental health of young people. "There has to be a level of risk that we're willing to take if a child's in school," says Kate Zinszer, an epidemiologist at the University of Montreal.

Schools are "where our children run around, play and laugh and argue with each other. They need to return to that sort of a healthy normality as soon as possible," Russell Viner, RCPCH's president, said in a statement last month.

From the start, some countries bet on strands of research suggesting young children are unlikely to spread the virus: Schools in the Netherlands cut class sizes in half but didn't enforce distancing among students under age 12 when they reopened in April. Other schools adopted a "pod" model as a compromise. Denmark, the first country in Europe to reopen schools, assigned children to small groups that could congregate at recess. It also found creative ways to give those groups as much space and fresh air as possible, even teaching classes in a graveyard. Some classes in Belgium met in churches to keep students spread out. Finland has kept normal class sizes, but prevents classes from mixing with one another.

As spring wore on, many other countries began to rethink distancing in schools. The Canadian province of Quebec, which reopened many elementary schools in May with strict distancing, has announced fall plans that allow children to socialize freely in groups of six; each group must stay 1 meter away from other groups of students and 2 meters away from teachers. Although French preschoolers were photographed sitting inside their own recess squares in May, day cares there have now abandoned all distancing rules for children ages 5 and under. Older students are advised to stay at least 1 meter away from others while inside. But outside they can play freely with those in their class. The Netherlands recently announced that anyone under age 17 does not need to distance.

The changes are driven not just by pediatricians' advice, but also by practicalities: A full school building leaves little room for distancing. In Israel, pressure was intense to return everyone to school after a partial reopening on 3 May. Two weeks later, classrooms welcomed back all students, housing their usual 30 to 40 pupils. Distancing in class was impossible, says Efrat Aflalo of Israel's Ministry of Health (MOH). So the country embraced another protective strategy: wearing masks.



In London, children play separately on a school playground.

Published by AAAS

Should children wear masks?

Masks likely blunt COVID-19 spread at school, but children—even more than adults—find them uncomfortable to wear for hours and may lack the self-discipline to wear them without touching their faces or freeing their noses. Does discomfort override a potential public health benefit?

"For me, masks are part of the equation" for slowing the spread in schools, especially when distancing is difficult, says Susan Coffin, an infectious disease physician at the Children's Hospital of Philadelphia. "Respiratory droplets are a major mode of [virus] transmission," she says, and wearing a mask places an obstacle in those droplets' path.

In China, South Korea, Japan, and Vietnam—where masks are already widely accepted and worn by many during flu season—schools require them for almost all students and their teachers. China allows students to remove masks only for lunch, when children sit apart or are separated by glass or plastic partitions. Israel requires masks for children older than age 7 outside the classroom, and for children in fourth grade and above all day—and they comply, says Efrat Aflalo of Israel's MOH, who has 8- and 11-year-old boys. On the bus ride to school, "all the kids are sitting with masks on," she says. "They don't take them off. They listen to the orders."

Elsewhere, masks are less central. In some schools in Germany, students wear them in hallways or bathrooms, but can remove them when seated at their (distantly spaced) desks. Austria reopened with this approach but abandoned masks for students a few weeks later, when officials observed little spread within schools. In Canada, Denmark, Norway, the United Kingdom, and Sweden, mask wearing was optional for both students and staff.

Not all countries have the luxury of instituting a mask policy driven by science and comfort. Benin requires masks in public spaces, but because the cost can be prohibitive for families, schools do not turn maskless students away. Students in Ghana returned to school in May wearing masks—if they had them. South Africa, which faces a rising COVID-19 caseload, is racing to provide free masks to all students who need them.

For Aflalo, the potential value of masks was underscored after a record-setting heat wave struck Israel in mid-May. As temperatures rose to 40°C, masks became intolerable. With the health ministry's blessing, students and teachers largely put them aside for almost a week.

For 2 weeks—the typical COVID-19 incubation period—things seemed fine. Aflalo left to go camping in the desert with her family. But then, a crisis: While on vacation, "I started getting calls about the Gymnasium," Aflalo says, referring to Gymnasium Rehavia, the school in Jerusalem with the large outbreak. Aflalo can't say for sure that the outbreak was fueled by a lack of masks, but she believes the timing is suggestive.



Eating was a solitary affair in Beijing in early June, shortly after many schools reopened.

PHOTO: FU TIAN/CHINA NEWS SERVICE VIA GETTY IMAGES

What should schools do when someone tests positive?

The short answer: No one knows. That's largely because of a lack of data about how many silent cases might be brewing when an illness or two comes to light. "How do we best deal with infection?" Vanderbilt's Kathryn Edwards wonders. "Do we just close the classroom" or shutter the whole school?

Some schools have favored isolating only close contacts. In Germany, for example, classmates and teachers of an infected student are sent home for 2 weeks, but other classes continue. Until summer break, Quebec generally did the same; at least 53 students and teachers tested positive after many schools reopened in May, according to news reports, but officials believed many of those infections were contracted in the community.

Elsewhere, officials are more cautious. Taiwan, which has largely suppressed the virus, kept schools open after one case but said it would close them for two or more, a situation it hasn't yet faced. In Israel, schools closed for a single case, and close contacts of every infected individual were tested and quarantined, Efrat Aflalo of Israel's MOH says. By mid-June, 503 students and 167 staff had been infected, and 355 schools had closed temporarily. (That number is a small fraction of the 5000 schools across Israel.)

Widespread testing in schools, including of children without symptoms, could help officials choose the most effective policy. The U.K. government recently began a study of as many schools as it can recruit across England. The project will test students and staff at preschools, primary schools, and secondary schools several times over at least 6 months for both the virus and antibodies to it, in an effort to trace transmission patterns and viral prevalence. In Berlin, researchers from Charité University Hospital launched a study in 24 schools on 15 June—2 weeks before summer break—that will test a cohort of 20 to 40 pupils and five to 10 staff members from each school every 3 months for at least a year. The researchers will look for both active infections and antibodies, to map the extent of silent infections and the threat they pose to students and staff. A similar study started this week in 138 preschools and elementary schools across the German state of Bavaria.

Published by AAAS



In early June, workers disinfected the entrance to a middle and high school in Jerusalem that was the site of a large outbreak of COVID-19.

PHOTO: YONATAN SINDEL/FLASH90/REDUX

Do schools spread the virus to the wider community?

Because children so rarely develop severe symptoms, experts have cautioned that open schools might pose a greater risk to teachers, family members, and the wider community than to students themselves. Many teachers and other school staff are understandably nervous about returning to the classroom. In surveys of some U.S. school districts, as many as one-third of staff say they prefer to stay away. *Science* could find few reports of deaths or serious illnesses from COVID-19 among school staff, but information is sparse. Several teachers have died of COVID-19 complications in Sweden, where schools did not modify class sizes or make other substantive adjustments.

Early data from European countries suggest the risk to communities is small. At least when local infection rates are low, opening schools with precautions does not seem to cause a significant jump in infections elsewhere.

In Denmark, nationwide case numbers continued to decline after day care centers and elementary schools opened on 15 April, and middle and high schools followed in May. In the Netherlands, new cases stayed flat and then dropped after elementary schools opened part-time on 11 May and high schools opened on 2 June. In Finland, Belgium, and Austria, too, officials say they found no evidence of increased spread of the novel coronavirus after schools reopened. However, it's hard to be sure, because in most places, schools reopened in concert with other aspects of public life.

In a broader study of COVID-19 clusters worldwide, epidemiologist Gwen Knight at the London School of Hygiene & Tropical Medicine and her colleagues collected data before most school closings took effect. If schools were a major driver of viral spread, she says, "We would have expected to find more clusters linked to schools. That's not what we found." Still, she adds, without widespread testing of young people, who often don't have symptoms, it's hard to know for sure what role schools might play.

At the same time, open schools can change the overall balance of who becomes infected by adding cases among children. In Germany, the proportion of all new infections that were in children under age 19 ticked upward from about 10% in early May, when schools reopened, to nearly 20% in late June. But wider testing and a decline in cases among the elderly could also explain the increase. In Israel, infections among children increased steadily after schools opened. That paralleled a rise in cases nationwide, but it's not clear whether the country's rising caseload contributed to the increase within schools or vice versa.

"We try to focus the epidemiologic research and find the source, but it's hard," Efrat Aflalo of Israel's MOH says. "We cannot say right now this is because of this or that."

What lies ahead?

In much of the world, schools remained closed through summer, and fall will see a wave of reopenings. For millions of especially vulnerable children, however, the break may continue indefinitely. Many low-income countries are hesitant to reopen because they lack the resources to shrink class sizes or provide everyone with masks. In June, Bangladeshi Prime Minister Sheikh Hasina said schools will likely stay shut until the danger of COVID-19 has passed. Similarly, officials in the Philippines said in-person schooling will not resume until there is a vaccine to protect against the virus.

In other places, ranging from Mexico to Afghanistan to the United States, planning for the fall is underway. In the United States, school districts are releasing a patchwork of plans, which often include hybrid models that alternate distance learning with small in-person classes. Whether those plans sufficiently protect children, staff, and communities from COVID-19 will depend on how case numbers look as opening day approaches. This reality was thrown into stark relief late last month, when Arizona's governor announced he would delay the state's school reopening by at least 2 weeks, to 17 August at the earliest, because of a surge in cases.

The reopening experiment will continue. Yet scientists lament that as before, it may not generate the details they crave about infection patterns and paths of transmission. "There just isn't really a culture of research" in schools, Vanderbilt's Kathryn Edwards says. Gathering data from schoolchildren comes with layers of complexity beyond those of traditional pediatric research. In addition to seeking consent from families, it often requires buy-in from teachers and school administrators who are already overwhelmed by their new reality. Integrating research—the only sure way to gauge the success of their varied strategies—may be too much to ask. ■

Published by AAAS



POLICY FORUM

AGRICULTURE AND ENVIRONMENT

The rotten apples of Brazil's agribusiness

Brazil's inability to tackle illegal deforestation puts the future of its agribusiness at risk

By **Raoni Rajão,[1] Britaldo Soares-Filho,[1] Felipe Nunes,[1] Jan Börner,[2] Lilian Machado,[1] Débora Assis,[1] Amanda Oliveira,[1] Luis Pinto,[3] Vivian Ribeiro,[4] Lisa Rausch,[5] Holly Gibbs,[5] Danilo Figueira[1]**

In the increasingly polarized international political arena, it has become difficult to find common ground to solve Brazil's ongoing environmental crisis, which has global as well as local implications. International buyers of Brazil's agricultural commodities have raised concerns about products that are contaminated by deforestation (i.e., deforestation occurred during the process of producing the product) (text S12). European Union (EU) criticism of the Brazilian government bolsters demands to boycott Brazilian products and to withhold ratification of the trade agreement reached in 2019 between the EU and Mercosur, the South American trade bloc. Among the concerns is that increasing greenhouse gas (GHG) emissions from deforestation and forest fires in Brazil could cancel out EU climate change mitigation efforts. The Brazilian government and agribusiness contend that national laws ensure high conservation standards, and hence trading bans should not include legally authorized deforestation (*1*). Here, we address the interlinkage between illegal deforestation in the Amazon and Cerrado—the largest Brazilian biomes with the highest rates of deforestation—and EU imports of Brazil's soy and beef, the country's major agricultural commodities (table S9). Although most of Brazil's agricultural output is deforestation-free, we find that 2% of properties in the Amazon and Cerrado are responsible for 62% of all potentially illegal deforestation and that roughly 20% of soy exports and at least 17% of beef exports from both biomes to the EU may be contaminated with illegal deforestation. Raising awareness is important to press Brazil to conserve its environmental assets and to promote international political will for cutting telecoupled GHG emissions. This could be achieved, for example, through the environmental safeguards of the Mercosur-EU trade agreement, which require EU imports to comply with the export country's legislation.

Our study goes beyond previous assessments of soy and beef supply chain traceability and zero-deforestation commitments

[1]Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil. [2]Institute for Food and Resource Economics and Center for Development Research, University of Bonn, Bonn, Germany. [3]Escola Superior de Conservação Ambiental e Sustentabilidade–IPÊ, Nazaré Paulista 12960-000, Brazil. [4]Stockholm Environmental Institute, Stockholm, Sweden. [5]University of Wisconsin, Madison, WI, USA. Email: rajao@ufmg.br

PHOTO: ANDRE PENNER/AP PHOTO

Published by AAAS



Cattle walk near an illegally burnt deforested area in the northern Brazilian state of Para. Beef exports contaminated by illegal deforestation are a key concern among some Brazilian trade partners.

(*2*–*5*), because we explicitly link illegal deforestation on individual rural properties to their agricultural production and exports to EU countries (text S12). To do so, we compiled a comprehensive set of land-use and deforestation maps for Brazil; information on 815 thousand rural properties' boundaries from the Cadastro Ambiental Rural (CAR), the country's online environmental registry (*6*); TRASE (Transparency for Sustainable Economies) data; and GTA documents (cattle transport permits) that are issued when animals are traded between properties and to slaughterhouses (table S1). We also developed software to deal with the geospatial data challenge of calculating the level of law compliance for each individual property, so as to differentiate between its potentially legal and illegal deforestation alongside its production of cattle and soy (texts S3 to S7 and figs. S4, S6, and S7).

ENVIRONMENTAL COMPLIANCE AND ILLEGALITY

Many countries have national or regional environmental regulation to protect riparian forests, in addition to local zoning laws that limit deforestation and the expansion of agricultural and urban areas. What makes Brazil stand apart is its property-level Forest Code (FC) and national CAR registry system, designed to monitor environmental compliance of its rural properties (*6*) (fig. S3). Brazil's FC regulates conservation on rural private properties (*1*), establishing areas of permanent protection (APPs) along water streams and on hilltops as well as legal reserves (native vegetation in a section of the property) (text S4). These legal reserves range from 20% of the property in most of the country (including parts of the Cerrado) to 80% in the Amazon rainforest, the latter in recognition of its environmental importance and economic potential for forest products (*7*).

Of our CAR sample, roughly 162 thousand of 362 thousand properties (45%) in the Amazon, and 217 thousand of 452 thousand properties (48%) in the Cerrado, are noncompliant with the FC for deforesting APPs or failing to conserve their minimum legal-reserve areas up until 2008—the deadline year for granting amnesty to eligible past deforesters (text S4 and figs. S8 to S13). Although these noncompliance figures do not yet equate to illegality, they do entail the obligation to start a program of environmental regularization by 2020, whereby landowners must submit and follow a self-designed plan to attain FC compliance over a period of 20 years (*8*).

A more pressing issue is illegal deforestation. Roughly 120 ± 26 thousand properties (15% of our sample) in both biomes were deforested after 2008 (*1*). About 36 thousand of these properties in the Amazon (84%) and 27 thousand in the Cerrado (35%) carried out this deforestation, in all likelihood illegally (figs. S12 and S13), because these properties had no forest surplus (i.e., vegetation above legal-reserve conservation requirements) to be eligible for a deforestation permit (see the figure) (text S5).

A substantial share of this potentially illegal deforestation is linked to agricultural export commodities. Of 53 thousand properties growing soy in both biomes, 20% were deforested after 2008, about half of them in a potentially illegal way (text S7, figs. S15 to S17, and table S6). In the Cerrado, we find 9.3 ± 1.2 thousand properties with deforestation after 2008 (43% with potentially illegal deforestation). In the Amazon, 1.5 ± 0.3 thousand properties were deforested since 2008, 91% of which were potentially illegal, despite the soy moratorium that prevents the trading of soy grown on deforested lands in this biome (*5*). Although only 1% of newly deforested areas are being cropped with soy in the Amazon biome, in contrast to 5% in the Cerrado (table S7), even farmers complying with the soy moratorium are clearing the forest for pasture or other crops within their holdings, and hence are still profiting from deforestation.

Despite uncertainties related to mapping and geospatial data modeling (texts S5 to S7 and S11), this represents an area of about 3.7 Mha of soy out of 17.2 Mha planted within the CAR properties during the 2016–2017 season (text S7 and figs. S14 and S15). This figure, tantamount to a harvest of 11.3 ± 1.1 million metric tons (text S7 and table S6), represents a very high level of soy potentially contaminated with illegal deforestation, including sizable volumes to the EU. Roughly 41% of EU's soy imports come from Brazil: 13.6 million metric tons per year, of which 69% come from this region (table S8). Although it is not possible to trace back soy imports to individual properties, we calculate by using municipality export shares that a total of 1.9 ± 0.2 million metric tons of soy grown on properties with illegal deforestation may have reached EU markets annually during the period of analysis (table S1), of which 0.5 million metric tons came from the Amazon (text S7, table S11, and fig. S21). In sum, 18 to 22% of all soy exported from the region to the EU is potentially contaminated. Yet the level of contamination may exceed the upper bound of 22%, given that our CAR sample covers only 80% of soy planted in the region (text S11).

With respect to beef, Brazil provides between 25% and 40% of EU beef imports (table S15). By matching GTAs issued in the states of Pará and Mato Grosso in 2017 with CAR data, we identify the origin of 4.1 million heads traded to slaughterhouses. Of this total, we estimate that 12 ± 2% (0.5 ± 0.1 million heads) come directly from properties with potentially illegal deforestation (table S13). In addition, 48 ± 10% of all slaughtered heads may be contaminated with potentially illegal deforestation from indirect suppliers, as the cattle pass from one property to another before being slaughtered (text S7, fig. S25, and table S13). Although beef exports from Pará are negligible, Mato Grosso state is the third largest Brazilian source of EU imports (fig. S30). By tracing cattle between properties and slaughterhouses, and tracing beef exports from the latter to EU countries, we estimate that of 17.7 ± 1.2 thousand metric tons of beef exported from Mato Grosso and Pará in 2017, about 46 ± 7% may have been contaminated with potentially illegal deforestation, including both direct and indirect suppliers (text S7 and figs. S24 to S26).

GHG EMISSIONS, AND A WAY FORWARD

China and the EU, Brazil's major agricultural product trade partners, acquired 29% and 19% of the country's agricultural exports,

Published by AAAS

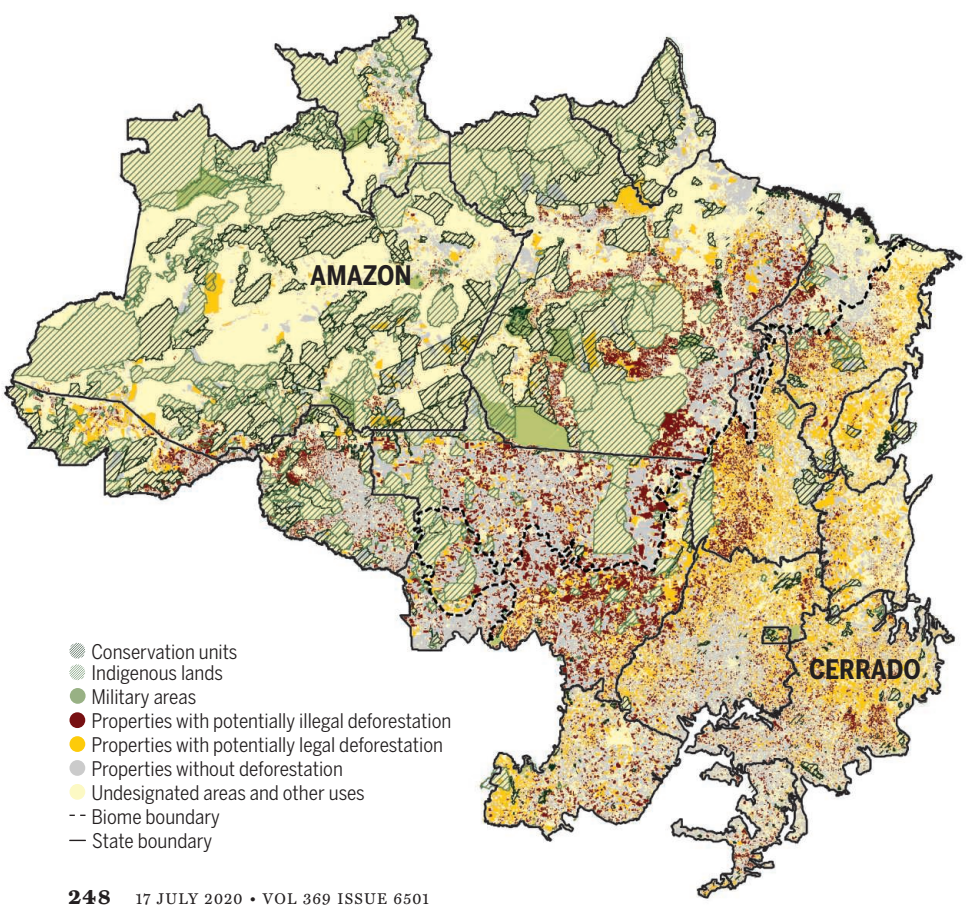
respectively, over the past 5 years (fig. S2) (*9*). All economic partners of Brazil should share the blame for indirectly promoting deforestation and GHG emissions by not barring imports and consuming agricultural products contaminated with deforestation, illegal or not. We calculate by superimposing a biomass map on deforestation maps (text S8) that EU soy imports alone could be responsible for the indirect emission of 58.3 ± 11.7 million metric tons of CO_2 equivalent ($MtCO_2e$) from both legal and illegal deforestation in the major Brazilian biomes between 2009 and 2017 (table S16), taking into account municipalities' export shares. Yet the EU share is likely to increase as a result of the Mercosur-EU and U.S.-China trade agreements. If implemented, these agreements will increase EU demand for Brazilian products (text S2) because of lower tariffs and to fill in the gap as U.S. exports to the EU could be redirected to China.

Most of Brazil's agricultural properties are deforestation-free. Of our CAR sample, 15% of properties were deforested after 2008, half of them potentially illegally. However, only 2% (17,557) of all properties in both biomes are responsible for 62% of all potentially illegal deforestation (text S10 and table S18). This small but very destructive portion of the sector poses a threat to the economic prospects of Brazil's agribusiness, in addition to causing regional and global environmental consequences. It is not enough to claim to be the world's most sustainable agriculture while a share of the sector fails to comply with the country's own environmental laws and supports the government's undoing of past environmental achievements (text S1). Instead, the government and agribusiness should take concrete steps to achieve countrywide environmental compliance. This is economically viable, given that about 4.1 Mha of legal-reserve debts in soy farms could be compensated by purchasing forest certificates from landowners with FC surplus (*10*). Additionally, the required restoration of 0.6 Mha of all riparian APPs together with 4.3 Mha of legal reserves on low-yield pasturelands in both biomes would remove 1.4 ± 0.3 $GtCO_2e$ (text S9, fig. S28, and table S17). This will greatly benefit agribusiness because its productivity depends on the rainfall regulated by the vast forests and other native vegetation (*7*) that still cover 60% of the Brazilian territory (*1*).

In the EU, public and private initiatives are building up to ensure agricultural imports free of tropical deforestation (*11*), and soon Chinese companies may follow suit (*12*). Yet so far there is a strong emphasis on private certification schemes that are costly, lack transparency, and encompass only specific farms and hence a small part of the sector. Here, we demonstrate that thanks to Brazil's already existing CAR registry (*6*), mapping and monitoring programs (*13*), and animal tracking system (GTA), it is possible to implement a national and public monitoring system that enforces environmental compliance at the property level to substantially reduce deforestation in the country's major agricultural supply chains. Brazil certainly has all the elements to feed the world with a responsible agricultural sector that tackles climate change and protects some of the world's most biodiverse regions. But to achieve this goal, the country and its international partners must acknowledge their shared environmental responsibilities as a main step to seek common solutions. ■

Linking rural property deforestation to agriculture and exports

Though most of Brazil's agricultural output is deforestation-free, 2% of properties are responsible for 62% of all potentially illegal deforestation. Roughly 20% of soy and at least 17% of beef exports from both biomes to the EU may be contaminated with illegal deforestation.



GRAPHIC: ADAPTED FROM RAJÃO *ET AL.* BY X. LIU/*SCIENCE*

REFERENCES AND NOTES

1. B. Soares-Filho *et al.*, *Science* **344**, 363 (2014).
2. J. Karstensen, G. Peters, R. Andrew, *Environ. Res. Lett.* **8**, 024005 (2013).
3. W. Carvalho *et al.*, *Perspect. Ecol. Conserv.* **17**, 122 (2019).
4. J. Alix-Garcia, H. Gibbs, *Glob. Environ. Change* **47**, 201 (2017).
5. F. Gollnow, L. Hissa, P. Rufin, T. Lakes, *Land Use Policy* **78**, 377 (2018).
6. Sistema Nacional de Cadastro Ambiental Rural, www.car.gov.br/publico/imoveis/index.
7. J. Strand *et al.*, *Nat. Sustain.* **1**, 657 (2018).
8. D. Rother *et al.*, *Trop. Conserv. Sci.* **11**, 10.1177/1940082918785076 (2018).
9. Ministry of Agriculture, Livestock and Food Supply, *AgroStat—Estatísticas de Comércio Exterior do Agronegócio Brasileiro* [Foreign Trade Statistics of Brazilian Agribusiness] (2020); http://indicadores.agricultura.gov.br/agrostat/index.htm.
10. B. Soares-Filho *et al.*, *PLOS ONE* **11**, e0152311 (2016).
11. Sustainable Trade Initiative, National Committee of the Netherlands, European Soy Monitor (IDH and IUCN NL, 2019); www.idhsustainabletrade.com/uploaded/2019/04/European-Soy-Monitor.pdf.
12. L. Jun, "We can feed the world in a sustainable way, but we need to act now" (World Economic Forum, 2019); www.weforum.org/agenda/2019/01/we-can-feed-the-world-in-a-sustainable-way-but-we-need-to-act-now/.
13. Instituto Nacional de Pesquisas Espaciais, *Prodes Project—Deforestation Monitoring of the Brazilian Amazon Rainforest and Cerrado Biome by Satellite* (2019); www.terrabrasilis.dpi.inpe.br/app/dashboard/deforestation.

ACKNOWLEDGMENTS

Supported by the Climate and Land Use Alliance, the Gordon and Betty Moore Foundation, Conselho Nacional de Desenvolvimento Científico e Tecnológico, and Fundação de Amparo à Pesquisa do Estado de Minas Gerais. Also supported by the Alexander von Humboldt Foundation (B.S.-F. and J.B.) and the German Federal Ministry of Education and Research (J.B.). Dataset, model, and results are available at csr.ufmg.br/radiografia_do_car.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/246/suppl/DC1

10.1126/science.aba6646

Published by AAAS



PERSPECTIVES

Thermal intolerance led to coral bleaching on a reef in American Samoa during a 2015 marine heatwave.

PHOTO: RACHAEL BAY

GENOMICS

Can genomes predict coral bleaching?

Expansion of the coral genomic tool kit could facilitate more informed conservation

By **Rachael A. Bay** *and* **Leslie Guerrero**

Coral reefs provide more than 1 billion people with food and income, yet these vital ecosystems are at risk because of rising ocean temperatures. Warming can induce coral bleaching, the breakdown of symbiosis between the coral animal and the photosynthetic algae that live within their cells, which can ultimately lead to coral death. Predicting which coral individuals are most tolerant to warming temperatures is essential to effective conservation plans. On page 268 of this issue, Fuller *et al.* (*1*) use genomic tools to predict bleaching outcomes for *Acropora millepora* corals. In doing so, they provide an integrative framework to advance predictions of climate change resilience on coral reefs, leading to more informed conservation and restoration efforts.

The frequency and severity of global coral bleaching episodes has been steadily increasing in recent decades (*2*, *3*), leading to mass mortality of corals and transformation of coral reef ecosystems (*4*). But not all corals are the same. In the 1970s, it was observed that corals from different oceanic regions had different upper thermal limits (*5*), leading to the general observation that warmer waters—for example, equatorial regions—harbor more heat-tolerant corals (*6*). Later observations of bleached corals immediately adjacent to unbleached corals opened the door to a role for intrinsic mechanisms of thermal tolerance variation. It is now known that thermal tolerance in corals is a function of the genetics of the coral, the algal symbiont and microbial community, and the environment that the coral has experienced (*7*–*10*).

Despite advancements in the use of genomic technology in corals over the past decade, the understanding of the genes responsible for determining thermal tolerance remains incomplete. This is partially due to the lack of high-quality genomic resources. To solve this problem, Fuller *et al.* produced a chromosome-level genome assembly from a single *A. millepora* coral. Genome assemblies serve as a "backbone" for genomic comparisons of individuals and populations, but assembling genomes in corals has proven difficult, in part because sequencing technology has been insufficient to resolve large repetitive regions. The genome Fuller *et al.* created using long-read

Department of Evolution and Ecology, University of California Davis, Davis, CA, USA. Email: rbay@ucdavis.edu

Published by AAAS

technology to cope with repetitive regions will be an enormous resource for future studies of coral ecology, evolution, cell biology, and development, facilitating previously untenable investigations in this ecologically important system.

With the aid of the new genome assembly, Fuller *et al.* endeavored to uncover the genetic variants associated with coral bleaching and test whether these variants could be used to predict individual bleaching susceptibility. One challenge is that bleaching is a complex trait. It is highly polygenic, encoded by many genes across the genome rather than a few key genes. This makes identifying the causative genetic variants difficult. Fuller *et al.* borrowed tools first used in human disease and agricultural applications to calculate "polygenic scores," combining the effects of many variants across the genome. In humans, this method was first used to predict the risk for schizophrenia and bipolar disease (*11*). Since, it has been used to predict traits as diverse as milk fat in dairy cows (*12*) and cancer risk in humans (*13*).

Fuller *et al.* sequenced genomes for 213 colonies of the coral *A. millepora* on the Great Barrier Reef. The colonies were collected during a bleaching event in 2017, so the authors also recorded the bleaching status for each colony. They found that although no individual genetic variants are significantly associated with bleaching, the combined polygenic score provides some predictive power. Furthermore, integrating this information with environmental variables (such as temperature and salinity) and symbiont species information, the authors were able to explain the majority (62%) of variation in bleaching among individuals.

The authors also scanned the genome for evidence of natural selection more generally. They found an unusually high amount of genetic diversity in a gene called *sacsin*. This gene encodes a co-chaperone that modulates heat shock protein 70 (Hsp70), which is involved in heat stress responses and is highly conserved throughout the tree of life. Hsp70 and associated proteins maintain proper protein folding during cellular stress before aggregations of misfolded proteins become toxic. Primarily studied within the context of human neurodegenerative disease, *sacsin* is important for the degradation of damaged mitochondria (*14*). These findings offer clues about coral cellular health mechanisms, especially considering that mitochondria play a role in response to thermal stress. Selection has maintained high genetic diversity of this locus for tens of millions of years, a time that long predates the split between *A. millepora* and other closely related species (*1*). This raises questions about the historical drivers of the genetic diversity of *sacsin* in *A. millepora* and the coral-specific role of *sacsin*.

Identifying genes under selection is useful for understanding coral biology, but this study's predictive framework could also advance current conservation and management efforts—for example, by identifying tolerant individuals for prioritization of protected areas, outplanting efforts, or selective breeding. Equally important to consider, however, are the potential consequences of limiting the future coral gene pool on the basis of a single trait. Trade-offs are common in nature, and although more thermally tolerant corals might survive the next heat wave, those same individuals might also carry undesirable traits such as disease susceptibility or low reproductive success. Experimentally testing for these trade-offs will increase the probability of success for conservation efforts. Ultimately, however, the only way to curb coral mortality and ensure future reef health is through climate change mitigation to reduce the incidence and intensity of heat wave events.

There is still much to learn about the genetic and environmental underpinnings of coral bleaching, and each step forward will enhance our predictive capacity. Fuller *et al.* show that genomic information improves the ability to explain individual variation in coral bleaching outcomes. It remains to be seen whether this polygenic score is reliable across independent samples and distant geographies. However, the authors provide a framework for integration of environmental and genomic data that could be used to predict the bleaching susceptibility of a coral even before a heat wave occurs and could be applied to any species of conservation concern. Such tools could enhance conservation and restoration practices by identifying individuals likely to be tolerant to future warming. ■

REFERENCES AND NOTES

1. Z. L. Fuller *et al.*, *Science* **369**, eaba4674 (2020).
2. J. M. Lough, K. D. Anderson, T. P. Hughes, *Sci. Rep.* **8**, 6079 (2018).
3. T. P. Hughes *et al.*, *Science* **359**, 80 (2018).
4. T. P. Hughes *et al.*, *Nature* **556**, 492 (2018).
5. S. L. Coles, P. L. Jokiel, C. R. Lewis, *Pac. Sci.* **30**, 159 (1976).
6. P. L. Jokiel, in *Coral Health and Disease*, E. Rosenberg, Y. Loya, Eds. (Springer, 2004), pp. 401–425.
7. S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, R. A. Bay, *Science* **344**, 895 (2014).
8. R. A. Bay, S. R. Palumbi, *Curr. Biol.* **24**, 2952 (2014).
9. G. B. Dixon *et al.*, *Science* **348**, 1460 (2015).
10. A. C. Baker, C. J. Starger, T. R. McClanahan, P. W. Glynn, *Nature* **430**, 741 (2004).
11. International Schizophrenia Consortium, *Nature* **460**, 748 (2009).
12. M. K. Hayr, T. Druet, D. J. Garrick, *J. Anim. Sci.* **94** (suppl_2), 13 (2016).
13. M. J. Machiela *et al.*, *Genet. Epidemiol.* **35**, 506 (2011).
14. F. Morani *et al.*, *Sci. Rep.* **9**, 11878 (2019).

10.1126/science.abc9342

CANCER

Nailing a Fe-rocious form of cancer

Disseminated cancer cells in brain fluids hijack iron from macrophages for metastatic survival

By **Livia Garzia**[1] *and* **Michael D. Taylor**[2]

Numerous cancer types can metastasize to a remote and inhospitable part of the human central nervous system (CNS): the subarachnoid or leptomeningeal space. This covers the surface of the nervous system and contains the cerebrospinal fluid (CSF). The CSF is largely hypocellular in healthy humans and has limited metabolic resources. It provides protective buoyancy for the brain and permits circulation of micronutrients and growth factors in the ventricular system. The CSF constitutes a relatively accessible proxy of the health of the CNS, and it is routinely sampled by physicians through a lumbar puncture. On page 276 of this issue, Chi *et al.* (*1*) reveal how human cancer cells thrive in this specialized anatomical location by hijacking a high-affinity iron transport system. This work highlights how studying tumor cell properties in the context of the challenges posed by the microenvironment can reveal unknown biology and possible approaches to therapy.

Lung and breast cancers can develop metastases in several organs, most commonly the bones, liver, and lungs. On rare occasions, disseminated cancer cells can reach the brain microvasculature, cross the blood-CSF barrier, and attach to the meninges, membranous layers that surround the CNS and form the subarachnoid space. Cancer cells then somehow thrive in the CSF-filled subarachnoid space and spread to the surface of the pia mater (a meningeal membrane). This leptomeningeal metastasis is an incurable condition; most patients succumb within months, and survival for 1 year is extremely rare (*2*). The CSF is a chal-

[1]Cancer Research Program, RI-MUHC and Department of Orthopedic Surgery, McGill University, Montreal, Canada. [2]Department of Neurosurgery, Hospital for Sick Children, Toronto, Canada. Email: mdtaylor@sickkids.ca; livia.garzia@mcgill.ca

Published by AAAS

lenging niche for metastatic cells to grow in because it contains very low concentrations of growth factors and micronutrients, and it is patrolled by immune cells whose composition is specific to the CSF (*3*, *4*).

Compared to the now consistent body of knowledge about the mechanisms enabling metastasis to the brain parenchyma, the understanding of the cascade of events leading to leptomeningeal metastasis is in its infancy. The brain parenchymal microenvironment is a major force in selecting the most fit metastatic cancer cells that can proficiently colonize the brain. Glia (stromal cells), neurons, and innate immune cells actively participate in creating an inflammatory microenvironment that is permissive to the extravasation (entry into tissue from the vasculature) of cancer cells and their invasion into the brain parenchyma (*5*, *6*). The brain parenchymal microenvironment favors the outgrowth of highly genetically heterogeneous tumor cells, resulting in brain metastases that display specific genetic features that prime them to grow in the brain parenchyma, such as mutations in the genes encoding phosphoinositide 3-kinase (PI3K) and the receptor tyrosine kinase HER2 (*7*).

Chi *et al*. hypothesized that single-cell transcriptomics of leptomeningeal metastases would reveal common pathways that confer to tumor cells the ability to survive and thrive in the subarachnoid space. Their analysis of five human lung and breast cancer leptomeningeal metastases reveals that distinct gene expression signatures are shared between these two cancer types, and that they converge on a high-affinity iron transport system. Iron, a high-demand micronutrient in the CSF, is sought after by various cell populations to sustain their iron-dependent cellular functions such as energy metabolism and DNA synthesis (*8*). In contrast to the parental, non–leptomeningeal-adapted primary tumors, tumor cells from leptomeningeal metastases express a high-affinity iron transporter, lipocalin 2 (LCN2), and its receptor, solute carrier family 22 member 17 (SLC22A17) (see the figure). This high-affinity iron-capturing system is best characterized as being expressed under physiological conditions by effectors of the innate immune response, with the purpose of efficiently binding bacterial siderophores (iron carriers), thereby outcompeting bacteria for the available iron (*9*).

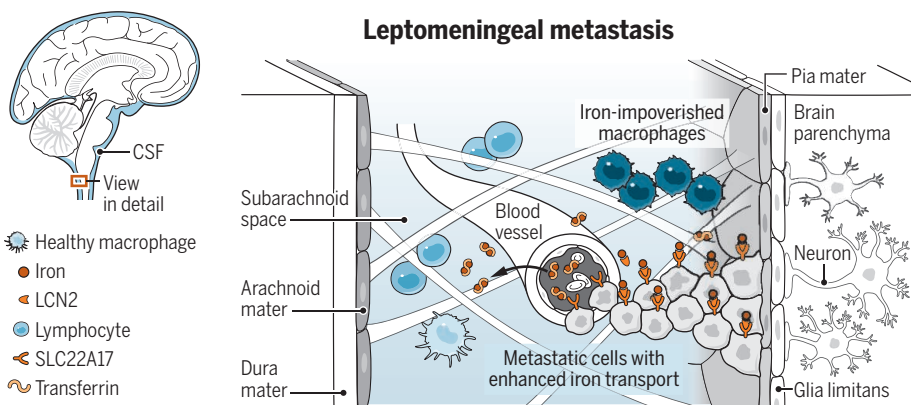
By applying comprehensive genetic and biochemical approaches in vitro and in immunocompromised mice grafted with human cells (xenografts), Chi *et al*. demonstrate that activation of the LCN2-SLC22A17 system is both necessary and sufficient for the survival of leptomeningeal metastases from human lung and breast cancers. Of note, this effect of the LCN2-SLC22A17 pathway on leptomeningeal metastases is mediated by its effects on iron metabolism, not by alternative cellular processes. Whether subclones of cells with high expression of LCN2 and SLC22A17 are preexisting in primary breast and lung cancers, and whether this iron-capturing system also confers increased survival of circulating tumor cells, requires further investigation.

Chi *et al*. also demonstrate that activated inflammatory macrophages from the CSF of mice and humans with leptomeningeal metastases increase the expression of LCN2 in tumor cells during coculture. This induction is mediated by the inflammatory cytokines interleukin-6 (IL-6), IL-8, and IL-1β. Few studies have focused on the biochemical composition of the CSF in the setting of leptomeningeal metastasis, and the study of Chi *et al*. contributes to bridging this knowledge gap. They report that concentrations of iron bound to LCN2 in clinical samples of CSF are elevated, and that this correlates with concentrations of IL-6 and hepcidin (a hormone that regulates iron absorption). In addition to providing tumor cells with the iron they require, LCN2 and SLC22A17 expression in metastatic cells represses local macrophage activation by depriving them of iron. Given the role of innate macrophage-dependent immune responses in primary tumors and brain parenchyma metastases (*10*, *11*), it is conceivable that decreased macrophage phagocytotic activity also translates to a more permissive microenvironment for metastatic outgrowth.

Because cancer therapies have improved and prolonged the survival of many cancer patients, the prevalence of leptomeningeal metastases has increased steadily over the past few decades. Although leptomeningeal metastases are a rare complication of advanced cancers in adults, they are common in primary brain cancers, particularly medulloblastoma (*12*, *13*). Chi *et al*. show that iron chelators, which reduce iron availability, confer a survival benefit to three different mouse models of leptomeningeal metastasis. Although iron chelation extended survival, it was not curative in a therapeutic or preventive setting. Understanding the alternative mechanisms by which leptomeningeal metastases circumvent iron deprivation will be valuable. The harsh microenvironment of the subarachnoid space likely imposes similar selection pressure on disseminated cells from various cancer types, and perhaps a treatment approach similar to the one tested by Chi *et al*. could be readily translated to a variety of patients with malignant leptomeningeal disease. ■

Iron promotes survival of metastatic cells

The meningeal membranes (dura mater, arachnoid mater, and pia mater) form the subarachnoid space, which is filled with nutrient-poor cerebrospinal fluid (CSF). Chi *et al*. find that leptomeningeal metastases, which grow along the pia mater, express lipocalin 2 (LCN2) and its receptor, solute carrier family 22 member 17 (SLC22A17), which promotes iron capture. This reduces iron availability for macrophages and promotes survival of metastases.



GRAPHIC: A. KITTERMAN/*SCIENCE*

REFERENCES AND NOTES

1. Y. Chi *et al.*, *Science* **369**, 276 (2020).
2. E. Le Rhun *et al.*, *Ann. Oncol.* **28**, iv84 (2017).
3. D. Schafflick *et al.*, *Nat. Commun.* **11**, 247 (2020).
4. H. Davson, K. Welch, M. B. Segal, *The Physiology and Pathophysiology of the Cerebrospinal Fluid* (Churchill Livingstone, 1987).
5. D. F. Quail, J. A. Joyce, *Cancer Cell* **31**, 326 (2017).
6. M. Valiente *et al.*, *Trends Cancer* **4**, 176 (2018).
7. P. K. Brastianos *et al.*, *Cancer Discov.* **5**, 1164 (2015).
8. J. Wang, K. Pantopoulos, *Biochem. J.* **434**, 365 (2011).
9. B. R. Wilson, A. R. Bogdan, M. Miyazawa, K. Hashimoto, Y. Tsuji, *Trends Mol. Med.* **22**, 1077 (2016).
10. S. Gholamin *et al.*, *Sci. Transl. Med.* **9**, eaaf2968 (2017).
11. L. Sevenich *et al.*, *Nat. Cell Biol.* **16**, 876 (2014).
12. R. Van Ommeren, L. Garzia, B. L. Holgado, V. Ramaswamy, M. D. Taylor, *Brain Pathol.* **30**, 691 (2020).
13. V. Ramaswamy *et al.*, *Lancet Oncol.* **14**, 1200 (2013).

ACKNOWLEDGMENTS

L.G. is supported by CIHR (PJT-162234), CRS (CRP-159384), and CCSRI. L.G. holds a FRQS-J2 career award.

10.1126/science.abb7041

Published by AAAS

PIEZOELECTRICS

Less can be more in functional materials

Nanopillar defects in $NaNbO_3$ induce a large piezoelectric response at high electric fields

By **Nazanin Bassiri-Gharb**

Ferroelectric crystals have a spontaneous polarization that can be reoriented with a sufficiently strong electric field, and they are piezoelectrically active, that is, an applied stress induces dielectric polarization. The elastic and electrostatic energies needed to distort the crystal during polarization are minimized by the formation of domains with near-uniform polarization. The motion of the resulting domain walls is a major contributor to the very large piezoelectric response of ferroelectric materials (*1*). On page 292 of this issue, Liu *et al*. (*2*) report on the exceptional electromechanical response of ferroelectric thin films of a simple ternary oxide, $NaNbO_3$. At low electric fields, the piezoelectric response is relatively limited and comparable to the industry-standard $Pb(Zr,Ti)O_3$. However, at increasing electric fields and decreasing frequencies, where the contributions from domain walls' motion become substantial, the effective piezoelectric response increases by more than an order of magnitude to >8000 pm/V.

In the past few decades, many advances in ferroelectrics have been achieved through materials processing, which enables precise control of structure, defects, and atomic distributions by leveraging molecular beam epitaxy and pulsed-laser deposition on single-crystalline substrates (*3*, *4*). Exceptional properties result from the design of epitaxial strains, boundary conditions at the interfaces, and superb control of the materials' chemistry at the single–unit cell level. In addition, chemical modification has made great strides in translating the high piezoelectric response of complex chemistries of bulk ceramics and single crystals—including chemically modified $Pb(Zr,Ti)O_3$, $Pb(Mg_{1/3}Nb_{2/3})O_3$-$PbTiO_3$, $(Na_{0.5}Bi_{0.5})TiO_3$, and $BaTiO_3$, or $(K,Na)NbO_3$-based chemistries—to thin films (*5*).

G.W. Woodruff School of Mechanical Engineering and School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Email: nazanin.bassirigharb@me.gatech.edu

Both approaches attempt to manipulate the material though modification of the elastic and electrostatic energies (*6*). Chemical pressures created by doping, and appropriate choice of substrates and electrodes, can modify local or global strain states (*7*). Similarly, charge manipulation can be implemented through some combination of creating internal interfaces (for example, within superlattices) (*8*), external interfaces (bilayers of polar and nonpolar materials) (*9*), or defect sites (*10*). Changes in the electrostatic or strain energies not only can result in changes in the overall domain configuration but also can modify the actual crystallographic structure (*11*, *12*).

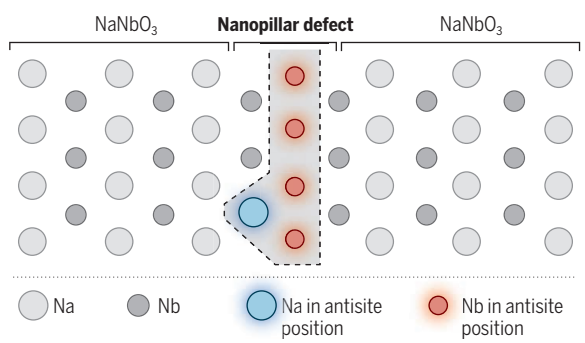
Liu *et al*. leverage all of these approaches in a simple yet highly effective self-assembly method. Sodium deficiency with respect to the stoichiometric and ferroelectrically active $NaNbO_3$ creates extended nanopillar defects within a perovskite-based matrix. Within the nanopillar regions, both Nb and Na atoms, which occupy the B and A perovskite lattice sites, respectively, can occupy antisite positions (see the figure, top). Additionally, the boundaries between the nanopillar regions and the matrix are characterized by edge-shared oxygen octahedra instead of corner-shared octahedra of the perovskite matrix.

The local strain fields associated with the nanopillars, possibly in addition to the uncompensated local charges, modify the ferroelectric polarization directions inside the material. Within the stoichiometric $NaNbO_3$ matrix, the spontaneous polarization directions can be oriented along any of the principal axes of a prototypical perovskite cell (a tetragonal distortion). Conversely, the strain-modified material in proximity to the nanopillars adopts a monoclinic distortion, where the polarization direction lies within a face or body-diagonal plane. The coexistence of monoclinic and tetragonal domains implies that the two distortions have similar energy levels. Hence, under an applied electric field, the polarization can rotate along multiple polarization directions belonging to either tetragonal or monoclinic distortions (see the figure, bottom). Such polarization rotations are often accompanied by a large enhancement of the piezoelectric response through mechanisms well understood in relaxor-ferroelectric compositions.

Liu *et al*.'s material offers practical advantages, in that the films are processed with a low-cost physical vapor deposition. Also, the chemical composition has limited environmental impact. High–piezoelectric-response materials containing Pb, Bi, or both have associated health and environmental concerns, whereas Na, Nb, and O are benign environmentally and nonhazardous

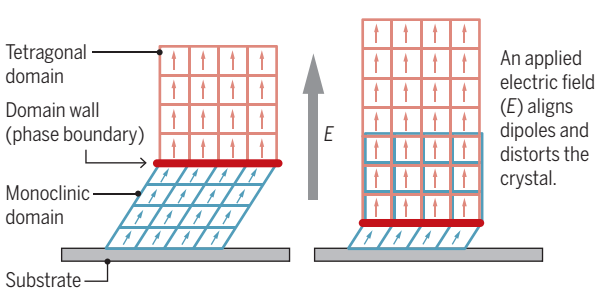
Self-assembling a piezoelectric

Piezoelectrics deform in response to an applied electric field. Liu *et al*. show that self-assembly of nanopillar structures in sodium niobate ($NaNbO_3$) leads to a strong piezoelectric effect at high electric fields.



Atomic rearrangements

Sodium deficiency in $NaNbO_3$ creates extended nanopillars of defects (outlined) in the perovskite crystal where the Nb and Na atoms occupy the other cation's site (an antisite position).



How electric fields move ferroelectrics

Ferroelectric crystals have a net polarization created by domains of several unit cells with aligned dipoles (arrows). Here, an electric field distorts the monoclinic form and grows the tetragonal form along a moving domain wall, which enhances piezoelectric effects.

GRAPHIC: C. BICKEL/*SCIENCE*

Published by AAAS

to humans (*3*). Finally, the nanopillars are formed through self-assembly in the material, rather than induced by exotic dopants or precisely controlled chemistries. During processing, the higher volatility of Na compared with Nb results in its stoichiometry deficiency (*13*) and ultimately creates Nb antisite occupancy that is critical to local structural modification. Such antisite cation occupancy has previously been reported in other ferroics and can substantially modify functionality. For example, Pb occupancy of Zr sites in $PbZrO_3$ thin films can induce ferroelectricity even at zero external fields, (*14*) in a material known as the prototypical antiferroelectric (one with antiparallel orientation of dielectric dipoles and zero net polarization at low electric fields).

Liu *et al.*'s approach for obtaining higher piezoelectric properties in simpler oxides through A-site cation deficiency could be explored with other ternary oxides. Both cation deficiency and size confinement are successful strategies for stabilizing ferroelectricity in otherwise paraelectric $SrTiO_3$ (*15*), and similar large piezoelectric enhancements might be obtained in thinner $NaNbO_3$ films, with coexisting tetragonal-monoclinic ferroelectric distortions. Ferroelectric thin films are useful for piezoelectrically actuated micro- and nanoelectromechanical systems, so it will be interesting to determine the minimum lateral size where the piezoelectric properties of the film are not affected by sample heterogeneity. Overall, neither perfection nor complexity was key to creating a material with an exceptional response, so in this case, less is indeed more. ■

REFERENCES AND NOTES

1. N. Bassiri-Gharb *et al.*, *J. Electroceram.* **19**, 49 (2007).
2. H. Liu *et al.*, *Science* **369**, 292 (2020).
3. J. A. Mundy *et al.*, *Nature* **537**, 523 (2016).
4. Q. Zhang *et al.*, *Adv. Mater.* **29**, 1702375 (2017).
5. S. Trolier-McKinstry, S. Zhang, A. J. Bell, X. Tan, *Annu. Rev. Mater. Res.* **48**, 191 (2018).
6. N. Domingo *et al.*, *Nanoscale* **11**, 17920 (2019).
7. D. G. Schlom *et al.*, *Annu. Rev. Mater. Res.* **37**, 589 (2007).
8. E. Bousquet *et al.*, *Nature* **452**, 732 (2008).
9. Y. Zhang *et al.*, *Nat. Nanotechnol.* **13**, 1132 (2018).
10. S. Das *et al.*, *Nature* **568**, 368 (2019).
11. X. Liu *et al.*, *Science* **364**, 264 (2019).
12. Y. Wei *et al.*, *Nat. Mater.* **17**, 1095 (2018).
13. N. Bassiri-Gharb, Y. Bastani, A. Bernal, *Chem. Soc. Rev.* **43**, 2125 (2014).
14. R. Gao *et al.*, *Chem. Mater.* **29**, 6544 (2017).
15. D. Lee *et al.*, *Science* **349**, 1314 (2015).

ACKNOWLEDGMENTS

The author acknowledges funding from the U.S. National Science Foundation under grant no. DMR-1255379, the Defense Threat Reduction Agency (DTRA) through grant no. HDTRA1-15-0035, and the Harris Saunders Jr. Chair at Georgia Tech.

10.1126/science.abc8007

GRAPHIC: N. CARY/*SCIENCE*; (PHOTO) R. J. WEINBERG

NEUROSCIENCE

A synapse census for the ages

A molecular imaging survey reveals the evolution of mouse brain synapse populations from birth through old age

By **Kristina D. Micheva[1], Richard J. Weinberg[2], Stephen J Smith[1,3]**

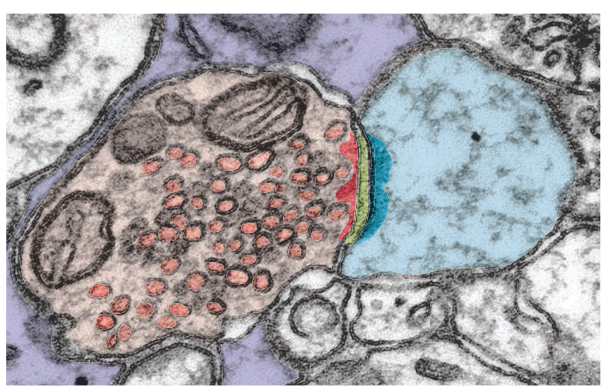
Brains can be viewed as vast ensembles of highly diverse and dynamic synapses that shape and store information as it travels through the networks of neurons that generate and interconnect those synapses. There are more than 100 trillion synapses per human neocortex (*1*), and each synapse is itself a highly complex entity, comprising thousands of diverse and cooperative signal-transduction proteins (*2*). Growing knowledge of these very large and heterogeneous populations of synapses, or "synaptomes," thus offers perhaps the most compelling glimpse to date of brains as machines deep and wide enough to support the richness and subtlety of the human mind. On page 270 of this issue, Cizeron *et al.* (*3*) describe the use of a mouse model to begin the task of systematically mapping synaptomes brain-wide, from birth to old age.

[1]Department of Molecular and Cellular Physiology, Stanford University, Stanford, CA, USA. [2]Department of Cell Biology and Physiology, University of North Carolina, Chapel Hill, NC, USA. [3]Allen Institute for Brain Science, Seattle, WA, USA. Email: stephens@alleninstitute.org

Cizeron *et al.* explored brains of transgenic mice fixed at 10 time points from birth through late adulthood. These mice expressed transgenes encoding two postsynaptic scaffold proteins of the PDZ domain–containing family—synapse-associated protein 102 (SAP102) and postsynaptic density protein 95 (PSD-95)—each genetically tagged with a different fluorophore. At least one of these proteins normally occurs in the postsynaptic density (PSD), a signaling complex common to all glutamatergic synapses (see the figure). Because the majority of central nervous system (CNS) synapses are glutamatergic, labeling SAP102 and PSD-95 provides an opportunity to detect most individual synapses. By imaging these two PDZ proteins and using machine-learning image-analysis tools (*4*), Cizeron *et al.* generated data representing intensities and morphologies of two vast sets of fluorescent

Census challenges: Synapse complexity and diversity

Glutamatergic synapses comprise hundreds of different proteins from both presynaptic and postsynaptic neurons. Synapse diversity reflects the various neurons as well as other influences, including electrical activity, neuromodulators, and glial cells.



Examples of differentially expressed synaptic signaling proteins

Presynaptic neuron			Pre ⇄ post	Postsynaptic neuron		Glia
Bouton	**Vesicles**	**Active zone**	**Synaptic cleft**	**PSD**	**Spine**	**Astrocyte**
Ion channels	v-SNAREs	t-SNAREs	Thrombospondins	SAP102, PSD-95	Ion channels	Aquaporins
Vesicle tethers	Synaptotagmins	Calcium channels	Cerebellins	Shanks	Kinases	Ion channels
Kinases	Synaptophysins	RIMS	Cadherins	AA receptors	Phosphatases	AA receptors
Phosphatases	Rab3s	Munc18S	Neuroligins	SynGAP	Calmodulins	AA transporters
Calmodulins	Synapsins	Munc13S	Neurexins	RhoGEFs	Cyclases	Thrombospondins
Cyclases	V-ATPase	CASK	LRRTMs	Kinases	Actin binders	Rho GTPases
GPCRs	SV2s	Bassoon/Piccolo	SynCAMs	Calmodulins	GPCRs	GPCRs

AA, amino acid; CASK, calcium/calmodulin-dependent serine protein kinase; GPCR, G protein–coupled receptor; GTPase, guanosine triphosphatase; LRRTMs, leucine-rich repeat transmembrane proteins; MUNC, mammalian homolog of Unc; PSD-95, postsynaptic density protein 95; RhoGEFs, Rho guanine nucleotide exchange factors; RIMS, regulating synaptic membrane exocytosis protein; SAP102, synapse-associated protein 102; SV2, synaptic vesicle protein 2; SynCAM, synaptic cell adhesion molecule; SynGAP, synaptic Ras GTPase-activating protein; t-SNARE, target SNAP receptor; v-SNARE, vesicle SNARE; V-ATPase, vacuolar-type H^+–adenosine triphosphatase.

Published by AAAS

puncta. On the basis of such data, this laboratory previously developed a classification scheme imputing individual puncta to three major synapse types, divided into a total of 37 subtypes (*4*). They now provide a view of synapse diversity as it unfolds from birth to late adulthood, across a fine-scaled anatomic parcellation.

SAP102 and PSD-95 play partially overlapping roles in organizing the PSD. Both of these PDZ proteins interact with postsynaptic glutamate receptors of the *N*-methyl-D-aspartate (NMDA) type to establish the core of a glutamatergic synapse. SAP102 and PSD-95 are expressed with different timing during mouse brain development, however, and appear to interact differentially with two receptor subunit isoforms (*5*). SAP102 interacts with the NMDA receptor subtype 2B (NR2B), which is characteristic of immature synapses, and PSD-95 interacts with the NR2A subtype associated with mature synapses. A progressive replacement of SAP102 and NR2B with PSD-95 and NR2A may underlie the transition from a highly plastic early stage in synaptogenesis to a more stable mature synapse.

Cizeron *et al*. argue that differing developmental trajectories of synaptic SAP102 and PSD-95 in functionally distinct brain regions provide a possible mechanistic basis for behavioral changes during a lifespan. Moreover, their work establishes a powerful and versatile template for further exploration of synaptomes in development and aging. Such explorations might probe larger subsets of the hundreds of other protein components of glutamatergic synapses, and proteins of the inhibitory synapses that use γ-aminobutyric acid (GABA) as their primary neurotransmitter. GABAergic synapses, the second-largest broad class of CNS synapses, are likely just as diverse and dynamic as glutamatergic synapses (*6*, *7*). Thus, extending the Cizeron *et al*. synapse census template from SAP102 and PSD-95 to other diversely expressed synaptic proteins might greatly deepen our knowledge of circuit development, homeostasis, modulation, and plasticity (*2*).

The same basic template for brain-wide developmental histology should also be adaptable to other protein-tagging and imaging modalities. Electron microscopy or super-resolution fluorescence microscopy would enable more reliable detection and mapping of the numerous tiny glutamatergic synapses that are critical to early synaptic network development and some forms of synaptic plasticity (*8*).

Because functional properties of individual synapses are also highly diverse and must somehow reflect diverse molecular properties (*2*), the Cizeron *et al*. histological results bring a need for parallel physiological studies of single synapses into the limelight. A recent in vivo calcium imaging study (*9*) provides an example of an approach that might fit this need. Perhaps the most promising avenue toward mapping functional synaptomes at scale would be to develop a "codebook" capable of predicting the difficult-to-measure function of an individual synapse from more readily accessible molecular and/or morphological measurements (*2*, *10*). The potential payoffs could be enormous, enhancing the much richer understanding of the brain's synaptic network architectures that are now beginning to emerge from electron microscopy connectomics (*11*) and viral trans-synaptic tracing methods (*12*).

The Cizeron *et al*. results on the intriguing patterning of CNS synapse diversity through development and aging also highlight the need to address more mechanistic questions about the underlying cell biology. The diverse phenotypes of individual synapses must somehow "inherit" their diversity largely from the proteomes of their two parent neurons. Such inheritance certainly involves trans-synaptic signaling interactions beyond mere juxtaposition of a stereotyped presynaptic part from one neuron with a stereotyped postsynaptic part from a second (*6*, *13*). There must also be experience-dependent environmental influences, including patterned electrical activity, paracrine neuromodulatory ligands, and adhesive interactions with extracellular matrix and glial cells (*14*). Brain-wide developmental synaptomes like those developed by Cizeron *et al*. may also be combined in the near future with transcriptomic neuron taxonomies (*15*) to reveal the origins of programmed synapse diversity in cell-type differentiation that must be fundamental to brain development, function, and plasticity. ■

REFERENCES AND NOTES

1. Y. Tang, J. R. Nyengaard, D. M. De Groot, H. J. Gundersen, *Synapse* **41**, 258 (2001).
2. N. A. O'Rourke, N. C. Weiler, K. D. Micheva, S. J Smith, *Nat. Rev. Neurosci*. **13**, 365 (2012).
3. M. Cizeron *et al*., *Science* **369**, 270 (2020).
4. F. Zhu *et al*., *Neuron* **99**, 781 (2018).
5. S. Won, J. M. Levy, R. A. Nicoll, K. W. Roche, *Curr. Opin. Neurobiol*. **43**, 94 (2017).
6. T. C. Südhof, *Cell* **171**, 745 (2017).
7. Z. J. Huang, A. Paul, *Nat. Rev. Neurosci*. **20**, 563 (2019).
8. A. Burette, F. Collman, K. D. Micheva, S. J Smith, R. J. Weinberg, *Front. Neuroanat*. **9**, 100 (2015).
9. R. Lu *et al*., *Nat. Methods* **17**, 291 (2020).
10. L. Groc, D. Choquet, *Science* **368**, eaay4631 (2020).
11. Y. Kubota, *Front. Neural Circuits* **13**, 64 (2019).
12. C. Saleeba, B. Dempsey, S. Le, A. Goodchild, S. McMullan, *Front. Neurosci*. **13**, 897 (2019).
13. S. Jang, H. Lee, E. Kim, *Curr. Opin. Neurobiol*. **45**, 45 (2017).
14. L. Privitera *et al*., *J. Neurosci*. **40**, 4644 (2020).
15. B. Tasic *et al*., *Nature* **563**, 72 (2018).

ACKNOWLEDGMENTS

S.J.S. is grateful to Allen Institute founder P. G. Allen for his vision, encouragement, and support. The work was also supported by NIH grants R01 MH111768, R01 NS094499 (K.D.M.), R01NS039444 (R.J.W.), and 1U24NS109113 (S.J.S.).

10.1126/science.abc9555

CHEMICAL ENGINEERING

Reimagining petroleum refining

Rigid polymer membranes separate crude oil components without distillation

By **Joan F. Brennecke** *and* **Benny Freeman**

Chemical separations involve often-complex mixtures of molecules into individual components or mixtures of components with similar properties. Such separations are integral to a variety of technologies, such as the production of liquid fuels and chemicals for transportation as well as everything from plastic bags to sterile medical equipment. On page 310 of this issue, Thompson *et al*. (*1*) describe the preparation and characterization of new polymeric membranes for crude oil fractionation, an extremely complex hydrocarbon separation process that is vital to the production of modern fuels and chemicals.

The traditional technique for separating crude oil into various fractions involves distillation, a process known since antiquity to separate more volatile from less volatile liquids. However, distillation is extremely energy intensive, relying on repeated evaporation and condensation cycles to achieve separation. Distillation dominates industrial chemical separations and consumes more than a quadrillion British thermal units (Btus) each year in the United States (~1% of total U.S. energy use). Thus, the U.S. National Academies of Sciences, Engineering, and Medicine (NASEM) pinpointed lower-energy alternatives to distillation as a critical need (*2*).

The use of membranes for separation of organic liquid mixtures is an attractive alternative because membranes do not rely on energy-intensive phase changes (such as boiling or condensation) to separate mixtures. The key to the development of new polymeric membranes by Thompson *et al*. was the synthesis of rigid spirobifluorene aryl diamine polymers (SADP). These poly-

McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, TX 78712, USA.
Email: jfb@che.utexas.edu; freeman@che.utexas.edu

Published by AAAS

mers are endowed with intrinsic microporosity, which permits high rates of molecule transport. Furthermore, the rigidity enables separation of molecules on the basis of size, shape, and affinity for the membrane material while helping to control swelling of the polymers when in contact with chemically challenging environments. Traditionally, polymers are constructed from organic building blocks, so they absorb these components and swell when in contact with organic mixtures (such as those found in crude oil). Swelling can be highly deleterious to a membrane's ability to separate molecules of similar shape and size. Therefore, although polymeric membranes are widely used for water purification (desalination) and gas separation (air separation, natural gas purification, and hydrogen recovery), they have not been commercialized for large-scale separations of complex organics, such as crude oil (*3*).



Energy-intensive distillation towers could one day be replaced with membrane separation units.

PHOTO: BIM/ISTOCK.COM

Early-stage experiments use dense, thick films and simple binary mixtures of organic liquids to acquire preliminary membrane separation data. This step is important for two reasons: developing structure-property relations to guide future membrane choices and elucidating fundamental details of penetrant transport mechanisms in such materials. However, to be of practical relevance, any interesting membrane material must also be manufactured as a thin-film asymmetric or thin-film composite membrane and be able to separate mixtures of interest. Whereas the former studies are quite common in the membrane literature, the latter thin-film studies are exceptionally rare. Thompson *et al.* report both types of data, including the successful preparation of thin-film (~0.2 μm) composite membranes by dip-coating the polymer onto a porous support. The success of these studies demonstrated the practical feasibility of the new approach.

In addition to showing excellent selectivity when challenged with model mixtures (for example, separation of 80% of 1,3,5-triisopropylbenzene from toluene), the thin-film membrane developed by Thompson *et al.* also demonstrated very good separation results for an actual sample of (extremely) light crude oil. The previously mentioned NASEM report also highlighted the lack of separation materials that can handle complex mixtures as a major gap in materials science (*2*). Crude oil is a quintessential example of a complex mixture of great commercial interest. Challenging the new membrane materials with real crude oil and demonstrating its successful separation into two streams (one comprising mostly higher boiling compounds and the other comprising lower boiling compounds) suggest that the membranes have practical promise. Thus, the new study moved all the way from rational molecular tailoring of new materials with attractive separation properties to the practical realization of a promising new thin separation membrane, a rare and extraordinary feat of great value to the field.

Another exciting aspect of the Thompson *et al.* study is its focus on one of the largest-scale chemical-separation challenges on the planet: the fractionation of crude oil into components needed for fuels and chemicals. Membranes are already being used for other important separation processes at scales comparable with that needed for crude oil fractionation. For example, as of 2019, 65.5 × 10^6 m^3/day (17,300 million gallons/day) of water are being desalinated by using reverse osmosis membranes (*4*). For comparison, the amount of global crude oil produced is ~4200 million gallons/day (*5*).

Desalination of Pacific Ocean seawater by using reverse-osmosis membranes requires about 2.5 kWh/m^3 to operate the membranes (*6*), with figures varying somewhat depending on location and salinity (*7*). Distillation-based desalination processes such as multieffect evaporation and multistage flash require 5.5 to 9 and 10 to 16 kWh/m^3, respectively (*8*). These numbers highlight the large reductions in energy use achievable by transitioning from distillation-based separations to membranes. Even hybrid distillation-membrane systems could yield dramatic energy savings and, in turn, markedly reduce the greenhouse gas production associated with such processes.

Challenges remain in the development of polymeric membranes for crude oil refining. The most important advance would be to increase a membrane's flux (the amount of material that permeates the membrane per unit membrane surface area), which determines the size (or number) of membrane modules required—and thereby the capital cost—to meet a particular refinery capacity. The flux measured by permeance through the less selective, standard polymers of intrinsic microporosity (PIMs) membranes is more than 10 times higher than that of the new SADP polymers, demonstrating a trade-off between permeability and selectivity. A reduction in flux over time (fouling) might well be another challenge, especially with heavier crude oils. The new work represents a major step toward a future in which citizens will drive by oil refineries without ever noticing the reduced energy demand, process intensification, and less carbon-intensive nature of this vital industrial process. ■

REFERENCES AND NOTES

1. K. A. Thompson *et al.*, *Science* **369**, 310 (2020).
2. U.S. NASEM, *A Research Agenda for Transforming Separation Science* (National Academies Press, 2019).
3. R. W. Baker, *Membrane Technology and Applications* (John Wiley & Sons, ed. 3, 2012), pp. 394–395.
4. E. Jones, M. Qadir, M. T. H. van Vliet, V. Smakhtin, S.-M. Kang, *Sci. Total Environ.* **657**, 1343 (2019).
5. U.S. Energy Information Administration (EIA), Short-term energy outlook (2020); www.eia.gov/outlooks/steo/report/global_oil.php.
6. N. Voutchkov, *Desalination* **431**, 2 (2018).
7. J. Kim, K. Park, D. R. Yang, S. Hong, *Appl. Energy* **254**, 113652 (2019).
8. S. Zhou, L. Gong, X. Liu, S. Shen, *Appl. Therm. Eng.* **159**, 113759 (2019).

ACKNOWLEDGMENTS

J.F.B. thanks the Welch Foundation for support (grant F-1945). B.F. thanks the Center for Materials for Water and Energy Systems for support (Award #DE-SC0019272).

10.1126/science.abd1307

Published by AAAS

VIEWPOINT: COVID-19

Aging immunity may exacerbate COVID-19

Increased general inflammation in aging inhibits immunity and affects responses to infections

By **Arne N. Akbar** and **Derek W. Gilroy**

Aging is associated with increased morbidity arising from a range of tissue dysfunctions. A common denominator of age-associated frailty is increased baseline inflammation, called inflammaging, that is present in older individuals. Recent studies have shown that the presence of excessive inflammation can inhibit immunity in both animals and humans and that this can be prevented by blocking inflammatory processes. This finding has important implications for the immunity of older individuals who are infected with pathogens such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that induce overwhelming inflammation, which can be fatal, particularly in older people. Reducing inflammation may be a therapeutic strategy for enhancing immunity in older people.

SARS-CoV-2 causes severe respiratory disease (coronavirus disease 2019, COVID-19) that mostly induces mild to moderate symptoms in younger individuals, but induces devastating morbidity and mortality in older individuals. A key hallmark of severe disease is exuberant inflammation in the respiratory tract of patients (*1*). Older healthy individuals (60 years and above) exhibit chronic low-grade sterile inflammation (not caused by a pathogen) characterized by high baseline serum concentrations of C reactive protein (CRP) and cytokines, including interleukin-6 (IL-6), and IL-8. This inflammaging predicts frailty and earlier mortality compared with individuals of the same age group who do not exhibit increased baseline systemic inflammation (*2*). Inflammaging may arise as a result of multiple mechanisms, including the accumulation of misfolded proteins, compromised gut barrier function, and obesity (*2*).

Defective resolution of immune responses in older individuals, due to the impaired clearance of dead and dying cells from sites of immune activity, may also lead to sustained inflammation in vivo (*3*). Senescent cells, which no longer divide, accumulate in every organ during aging and contribute to increased baseline inflammation (*4*). Senescent cell populations are found within diverse lineages, including lymphocytes, and are nonproliferative as a result of telomere erosion, damaged DNA, epigenetic changes, or mitochondrial dysfunction. Senescent nonlymphoid cells secrete inflammatory cytokines, chemokines, growth factors, and matrix metalloproteinases (MMPs). This is known as the senescence-associated secretory phenotype (SASP), and the secretion of these inflammatory molecules by any of these cell types may cause organ dysfunction during aging in humans and in animal models (*4*).

Inflammation is an essential component of the initiation phase of an immune response, but it is actively reduced by the clearance of dead and dying cells as well as other pro-resolution processes (*3*). The presence of excessive inflammation can inhibit antigen-specific immunity in vivo. This is illustrated by studies in both mice and humans showing that increased inflammation is detrimental for the efficacy of many vaccines, such as against influenza virus (*5*). The negative impact of inflammation on immunity during aging can be reversed in part by treatment with the mTOR (mammalian target of rapamycin) inhibitor rapamycin, which enhances influenza-specific antibody responses after influenza vaccination in treated individuals (*6*).

Furthermore, the short-term administration of an oral p38 mitogen-activated protein kinase (MAPK) inhibitor reduced baseline cutaneous inflammation in older people, in part by inhibiting the influx of inflammatory monocytes. This resulted in a significantly increased response to varicella zoster virus (VZV, which causes chicken pox in children and shingles in adults) recall antigen challenge (memory T cell response) in the skin in vivo (*7*). Therefore, reducing inflammaging with a short-term course of mTOR or p38 MAPK inhibitors and possibly other anti-inflammatory agents (e.g., steroidal drugs such as dexamethasone) may be a strategy for immune enhancement in older people. This may be of particular relevance for older patients with COVID-19 who have severe inflammation in the respiratory tract during disease progression that may hinder antiviral immunity.

Both senescent cell numbers and baseline inflammation increase during aging, suggesting that these events are interlinked. Because senescent cells secrete proinflammatory mediators and accumulate in every organ in the body during aging, another way to enhance immunity and reduce the inflammatory burden may be to eliminate these cells. Natural mechanisms exist for clearing senescent cells from tissues. These involve the expression of "kill me signals" [e.g., natural killer (NK) cell receptor (NKR) ligands] on senescent cell populations that activate receptors on cytotoxic T lymphocytes that enable them to kill senescent nonlymphoid cells in tissues (*8*).

The ability of T and NK cells to eliminate senescent cells raises the question of why nonlymphoid senescent cells accumulate during aging. One possibility is that the general multifactorial age-associated decline in immunity also extends to surveillance of senescent tissue cells in different organs. Alternatively, senescent cells may display evasion strategies that enable them to escape from immune clearance. These mechanisms include the shedding of decoy receptors that interfere with the ability of cytotoxic T cells to recognize them (*9*) and also the expression of inhibitory "don't kill me" signals [such as human leukocyte antigen E (HLA-E), a ligand for the inhibitory NKR, NKG2A] on their cell surface (*8*). The removal of senescent (p16INK4A-expressing) cells in transgenic mice by inducing activation of thymidine kinase can reverse age-associated organ dysfunction. The negative impact of senescent cells in aging has prompted the development of therapies to remove senescent cells with drugs called senolytics that are currently being tested in mice and humans (*10*).

Blocking inflammation directly or eliminating the cells that produce inflammatory mediators can improve health and immunity. The inflammatory nature of senescent cells in various tissues may inhibit immune responses of older individuals, and this must be factored in when considering the cause of hyperinflammatory responses during some SARS-CoV-2 infections. Because senescent cells increase in the lungs during aging (*11*), they are very likely to be present in the lungs of older COVID-19 patients and may participate in the initiation of an inflammatory cascade. Although the secretion of proinflammatory mediators by adipose tissue may contribute to inflammaging, it is not yet known

Division of Medicine, The Rayne Building, University College London, 5 University Street, London, UK.
Email: a.akbar@ucl.ac.uk

Published by AAAS

whether senescent cells within the adipose tissue or the adipocytes themselves produce these mediators. This is relevant in the context of COVID-19 because obesity is a comorbidity for severe disease.

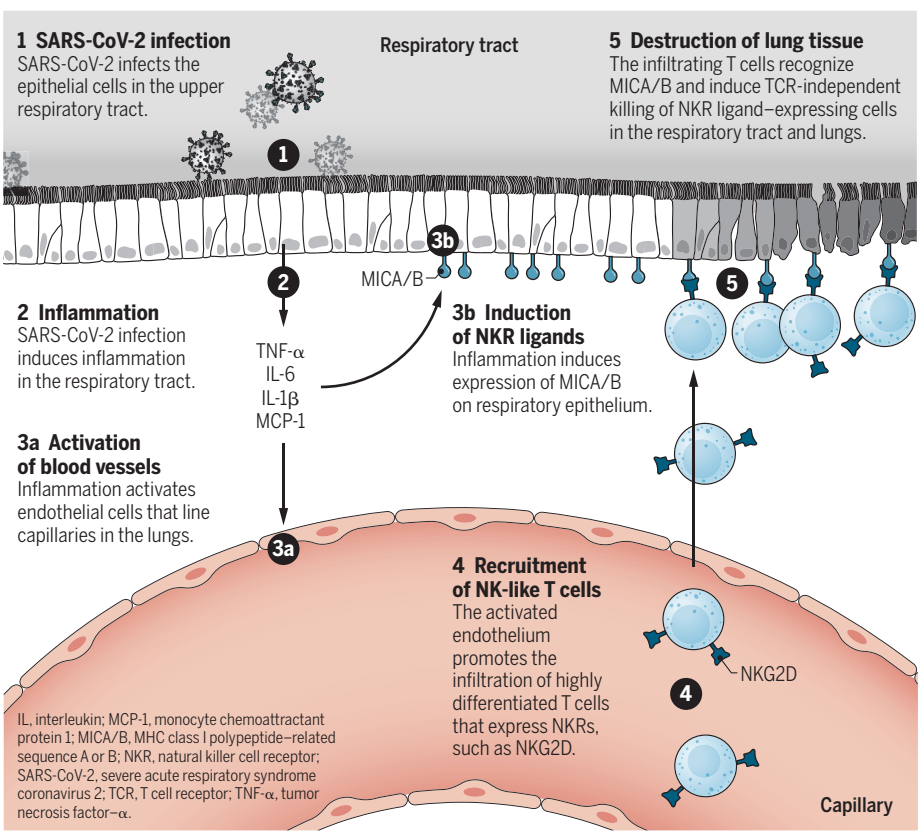
A key unknown issue is the relationship between high baseline inflammation and the hyperinflammation that occurs in older COVID-19 patients with severe disease. One hypothesis is that preexisting inflammatory cells, including senescent populations and adipocytes, create the inflammaging phenotype that amplifies subsequent inflammatory events. This would involve the recruitment of other inflammatory cells. For example, during an immune response, monocytes are recruited from the circulation into the skin of old individuals (*12*), and this may also occur in the lungs of COVID-19 patients with severe disease. Nevertheless, high amounts of inflammation alone do not explain the devastating tissue destruction that is observed in the lungs of COVID-19 patients with severe disease, and it may be that age-associated changes in T cells have a role in the immunopathology.

T lymphocytes that are highly differentiated and exhibit senescence-like characteristics accumulate in older individuals. These cells lose the capacity to proliferate after activation and express multiple markers of senescence, including DNA damage associated proteins [e.g., phosphorylated histone H2AX (γH2AX)] and cyclin-dependent kinase inhibitors (e.g., p16INK4A), but are highly efficient cytotoxic cells (*13*). However, these aged T cells express NKRs and can kill different cell types that express NKR ligands (*13*). Inflammation has been shown to increase the expression of NKR ligands by different cell types, including senescent nonlymphoid cells (*14*). Therefore, aged T cells that infiltrate the lungs of COVID-19 patients may not function well in an antigen-specific manner because of the inflammation present but also because they acquire NK cell activity and do not recognize specific antigens. NK cells are lymphocytes that do not require prior sensitization with antigens to mediate cytotoxic responses to infected and/or malignant cells. The inflammation in the lungs of COVID-19 patients, possibly resulting in part from the infiltration of inflammatory monocytes, would induce the expression of NKR ligands by nonlymphoid cells (*14*), and these may be targeted and killed by infiltrating T cells (see the figure). This would contribute to the reported collateral tissue damage in the respiratory tract of COVID-19 patients. There is an urgent need for immunohistological data from the lungs of these patients to address this possibility.

The propensity to mount inflammatory responses in tissues and changes in the behavior of different leukocyte populations have to be considered when addressing the immune responses of older individuals during infection. The baseline inflammation may not be detrimental in itself; instead, it may initiate an inflammatory cascade that amplifies the excessive inflammation that occurs in response to pathogens. In addition to senescent cells, other inflammatory cell types, such as adipocytes, may also trigger inflammation. Inflammaging has many implications for COVID-19 patients: The accumulation of senescent cells in the respiratory tract of older patients may be involved in the initiation of an inflammatory cascade that could inhibit T cell responses to virally infected cells that are present. Another consequence of the inflammation is the induction of NKR ligand expression by cells in the lung (*14*) that would make them susceptible to killing by infiltrating T cells that express NKRs. The search for an effective vaccine for COVID-19 should also consider the decreased vaccination efficacy in older individuals that may be associated with inflammaging (*15*). Therefore, the effective treatment of COVID-19 patients may require a combination of anti-inflammatory and antiviral regimes to complement vaccination against the virus. ■

Aged immune responses to SARS-CoV-2 infection

Highly differentiated T cells in older individuals may induce damage in SARS-CoV-2–infected lungs, as hypothesized in the diagram.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. M. Merad, J. C. Martin, *Nat. Rev. Immunol.* **20**, 355 (2020).
2. C. Franceschi, P. Garagnani, G. Vitale, M. Capri, S. Salvioli, *Trends Endocrinol. Metab.* **28**, 199 (2017).
3. R. P. H. De Maeyer *et al.*, *Nat. Immunol.* **21**, 615 (2020).
4. J. Campisi *et al.*, *Nature* **571**, 183 (2019).
5. A. Parmigiani *et al.*, *PLOS ONE* **8**, e79816 (2013).
6. J. Mannick *et al.*, *Sci. Transl. Med.* **10**, eaaq1564 (2018).
7. M. Vukmanovic-Stejic *et al.*, *J. Allergy Clin. Immunol.* **142**, 844 (2018).
8. B. I. Pereira *et al.*, *Nat. Commun.* **10**, 2387 (2019).
9. D. P. Muñoz *et al.*, *JCI Insight* **4**, e124716 (2019).
10. J. L. Kirkland, T. Tchkonia, Y. Zhu, L. J. Niedernhofer, P. D. Robbins, *J. Am. Geriatr. Soc.* **65**, 2297 (2017).
11. P. J. Barnes, J. Baker, L. E. Donnelly, *Am. J. Respir. Crit. Care Med.* **200**, 556 (2019).
12. E. S. Chambers *et al.*, *bioRxiv* 10.1101/2020.04.02.020081 (2020).
13. B. I. Pereira *et al.*, *Nat. Immunol.* **21**, 684 (2020).
14. A. Chauveau *et al.*, *J. Innate Immun.* **6**, 89 (2014).
15. A. Ciabattini *et al.*, *Semin. Immunol.* **40**, 83 (2018).

ACKNOWLEDGMENTS

Thanks to R. De Maeyer for help with the preparation of this article. A.N.A. and D.W.G. are funded by a Medical Research Council (U.K.) Grand Challenge in Experimental Medicine Grant (MR/M003833/1); A.N.A. receives an MRC grant (MR/P00184X/1).

10.1126/science.abb0762

Published by AAAS



NSF-funded investigators such as Harkamal Walia, shown here inspecting rice plants, aim to answer big questions facing science and society.

POLICY

Overseeing innovation

75 years after Vannevar Bush's impactful report, debate continues about directing science

By **Emily K. Gibson**

Since its publication in 1945, Vannevar Bush's *Science—The Endless Frontier* has elicited considerable examination and criticism. More often than not, contemporary discussions emphasize the report's shortcomings rather than its successes, using it as a foil for remarking on everything from the complexity of the innovation process to the immense growth of the American science and technology (S&T) enterprise over the past 75 years. Nevertheless, it persists.

What explains, then, the endurance of a document whose prescriptions offer questionable insight into a science and engineering base that has evolved far beyond what existed in 1945? The answer lies in the fact that it was an attempt to solve a policy debate that is still ongoing: What is the best way to organize and oversee scientific research in pursuit of innovation?

STATE-DIRECTED SCIENCE: WARTIME ORIGINS

Concerned about American S&T capabilities as World War II loomed on the horizon, the U.S. government created the National Defense Research Committee (NDRC) on 27 June 1940. Staffed by prominent U.S. scientists and headed by the engineer Vannevar Bush, the NDRC supplemented ongoing military research and development efforts by enlisting civilian scientists and industrial research laboratories.

A year after its creation, the committee was encompassed by the newly established Office of Scientific Research and Development (OSRD), which Bush was also tapped to lead. The OSRD expanded the NDRC's mandate to include medical research as well as new weapons development and testing capabilities, coordinating research that led to the production of penicillin, the tactical use of radar, and the development of the atomic bomb (*1*). Pleased with the OSRD's success, President Franklin D. Roosevelt asked Bush to propose a plan for continuing the wartime science effort into peacetime. Bush responded with *Science—The Endless Frontier*.



Vannevar Bush proposed a postwar plan that allowed researchers control over projects and priorities.

TOP-DOWN OR BOTTOM-UP? DIRECTING SCIENCE IN TIMES OF PEACE

In the report, Bush did not advocate for the continuation of OSRD. Rather, he called for the creation of a new organization—a national research foundation—that would primarily support fundamental research. In peacetime, he argued, the federal government needed to foster the acquisition of knowledge about the Universe from which future applications could be drawn (*2*).

Although few questioned that the federal government had an important role to play in supporting scientific research after WWII, not everyone agreed on how that support should be organized. Bush's plan described an organization that would encourage scientific advancement by awarding grants to scientists (primarily at universities) who would work on problems of their own choosing—a detail that elicited considerable criticism.

On 22 July 1945, just three days after the White House released Bush's report to the public, the *New York Times* ran an article penned by Waldemar Kaempffert, the paper's science editor (*3*). Long on praise for the OSRD, Kaempffert was skeptical of the engineer's vision for postwar science. The success of OSRD, he argued, was due to the efficiency of its mode of operation.

The reviewer is a science policy analyst at the National Science Foundation, Alexandria, VA 22314, USA. Email: egibson@nsf.gov

PHOTOS (TOP TO BOTTOM): UNIVERSITY COMMUNICATION/UNIVERSITY OF NEBRASKA-LINCOLN; EVERETT COLLECTION INC./ALAMY STOCK PHOTO

Published by AAAS

The office had sought out the best researchers and laboratories and organized, planned, and directed their research toward specific goals. "This procedure, which it has been frequently asserted produced brilliant results, is to be abandoned for a kind of laissez-faire system," Kaempffert explained. Although he urged Congress to give appropriate attention to Bush's plan, he concluded that "America needs something more like the OSRD, something more efficient than a body of higher-ups who wait for ideas to be submitted."

Kaempffert's article sparked a debate. James B. Conant—noted U.S. chemist, president of Harvard, and former NDRC member—responded in a letter published in the *New York Times* on 13 August (*4*). A veteran of the wartime science effort, Conant explained that OSRD's objective had not been to advance science generally but rather to create and improve weapons of war. The advancement of science in the postwar period, he argued, necessitated a different approach. Conant backed Bush's vision for a national research foundation, concluding, "There is only one proved method of assisting the advancement of pure science—that of picking men of genius, backing them heavily and leaving them to direct themselves."

Science—The Endless Frontier
Vannevar Bush
1945.

THE DEBATE MOVES TO CONGRESS AND GETS PERSONAL

The dispute that unfolded in the pages of the *New York Times* spilled over into the halls of Congress. While the disagreement was multifaceted, the issue of how much control such an agency should have over research priorities formed a key point of contrast between the dueling legislative proposals.

Bush was not the only individual with a vision for postwar American science. In fact, *Science—The Endless Frontier* was crafted, at least in part, to counter one particular vision with which he heartily disagreed: that of Harley Kilgore, a Democratic senator from West Virginia. While serving on the Military Affairs Committee, Kilgore had become convinced of the importance of scientific research to American national defense. He called for an organization that would support applied, as well as basic, research to ensure military preparedness after the war. But Bush felt Kilgore's view of what science and technology could offer society was exceedingly narrow. In contrast, Bush described science's primary aim as "increas[ing] the knowledge and the understanding of man…[and] extending his grasp of the environment in which he lives" (*5*).

Bush's vision for an agency devoted primarily to supporting undirected, fundamental science research won out, and the National Science Foundation (NSF)—established by President Harry Truman on 10 May 1950—cohered around the central principle that scientists themselves should determine the course of fundamental research. As a result, NSF adopted a decentralized organizational structure that would ensure that the scientific community at large exerted considerable influence over which projects and lines of research would receive federal support. But the creation of the NSF did not end the debate.

PRESSURE TO DIRECT RESEARCH PERSISTS

The fledgling agency immediately faced pressure from the Bureau of the Budget, the predecessor to the present-day Office of Management and Budget, to play a stronger role in directing research and establishing research priorities. Alan Waterman, the agency's first and longest-serving director, firmly resisted.

The debate reached a crescendo in 1968, when the agency's first congressional review culminated in an amendment to its founding charter that officially authorized the NSF to assume responsibility for supporting applied research. The mandate pushed the agency in new directions, which have served as vital sources of strength for the American S&T ecosystem, including lending greater support to engineering research and the creation of cross-disciplinary initiatives aimed at answering big questions facing science and society.

As *Science—The Endless Frontier* inspires reflection in its 75th year of existence, policy-makers continue to debate the proper role of the federal government in supporting fundamental research amid a shifting social, political, and economic landscape—as evidenced by the recent legislative proposals in the House and Senate that call for substantial organizational changes to the NSF. With a nod to Bush's report, the Endless Frontier Act represents the latest contribution to this debate. As long as science itself continues to evolve, however, such policy prescriptions will need constant refinement, and Bush's text will continue to feature as an important touchstone in those deliberations. ■

REFERENCES AND NOTES

1. L. Owens, *Bus. Hist. Rev.* **68**, 515 (1994).
2. V. Bush, *Science—The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research* (U.S. Office of Scientific Research and Development, 1945).
3. W. Kaempffert, *New York Times*, 22 July 1945, p. 71.
4. J. Conant, *New York Times*, 13 August 1945, p. 18.
5. M. Lomask, *A Minor Miracle: An Informal History of the National Science Foundation* (National Science Foundation, 1976), p. 40.

10.1126/science.abb4913

OPTICS

Color vision

An ambitious volume lends scientific and cultural context to an expansive concept

By **Michael Rossi**

The concept of "color" is tricky—but then so is "nature" and so is "history." It follows that to attempt a "natural history of color" is to wade into a combinatorial explosion of assumptions, definitions, methodologies, stories, and theories. Written by Rob DeSalle, a molecular biologist, and Hans Bachor, an optical physicist, *A Natural History of Color* accepts this challenge, endeavoring, as the authors state, to understand "[h]ow color has influenced… human existence [and] the entire natural world around us." This is a tall order, one that the authors interpret as requiring an accounting of color from the origins of the Universe through to the present day. How one evaluates the success of this project depends much upon what one wishes for going in.

A Natural History of Color is a companion piece to *The Nature of Color*, an exhibition at the American Museum of Natural History in New York City. Infelicitously, the exhibition was scheduled to open in March of 2020, which corresponded with the arrival of coronavirus disease 2019 (COVID-19). As of this writing, the museum is still closed. Nevertheless, the book stands on its own.

As the authors acknowledge, understanding color *tout court* is a vast, multidisciplinary project, involving physics, physiology, psychology, anthropology, and philosophy, among many other disciplines. To make sense of this potential chaos, the book takes evolution as its guiding principle, looking first at the evolution of light-sensitive molecules and their integration into, and usefulness to, living things; next, at the evolution of particular color schemes that can serve to attract, repel, mimic, and confuse other organisms that possess particular arrangements of light-sensitive molecules; and finally, at the evolution of distinct symbolic responses to color in human beings.

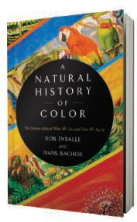
Clearly comfortable with this context of color, DeSalle and Bachor delve into topics such as the particle and wave natures of light, the different types of proteins used in different light-sensing organs, and the neurosci-

The reviewer is at the Department of History, University of Chicago, Chicago, IL 60637, USA. Email: michaelrossi@uchicago.edu

Published by AAAS

ence of mammalian color perception. Their explanations can be quite detailed and might be more than a casual reader needs, but the intricacy with which the authors weave the study of color in and out of a *longue-durée* history of life on Earth is impressive.

Perhaps to leaven their dips into the minutiae of physics, genetics, statistics, and molecular biology, the authors frequently reach for pop-culture references to illustrate (or augment) their points. A truncated list would include the movies *Predator* and *Animal House*; the TV shows *Game of Thrones*, *The Walking Dead*, and *Pee-wee's Playhouse*; and the rock groups Violent Femmes and the Police, as well as repeat appearances by Gary Larson's *The Far Side* cartoons. The effect is that of a stream-of-consciousness amble through the different disciplinary and personal predilections of the authors—a bit like *On the Road*, but with more rhodopsin than reefer and more genes than jazz.



A Natural History of Color
Rob DeSalle and Hans Bachor
Pegasus Books, 2020. 272 pp.

On questions of color and culture, the book feels thinner. Here, rather than the rich descriptions that characterize chapters about the evolutionary selection for opsins, for example, the reader finds a collection of anecdotes and factoids about color symbolism in predominantly Western civilization. One explanation for this shift in argumentation is that color symbolism and meaning in human culture are so vast as to defy easy characterization. This is, of course, accurate and acknowledged by the authors, but it is an admission that undercuts the enterprise. More to the point, the methodologies that prove so indispensable in studying photons and genes—entities that can be assumed to behave regularly across vast periods of time—are less helpful in studying people, who tend to be unpredictable, willful, and inveterate outlaws. As a result, the book's attempt to tell the natural history of color according to neat laws breaks down precisely when it matters most. This paradox is one of the most enchanting things about studying color but also one of the most frustrating.

This gap notwithstanding, the book does provide, unequivocally and generously, a long look not just at the evolutionary background of color perception but also at the ways in which science itself both draws from and contributes to an understanding of color. The writers are not simply practicing scientists and museum curators, they are enthusiasts eager to share their knowledge with their audience. One comes away from this book with a sense that both the scientific description of color and the science that is used to arrive at that description are mutually reinforcing: a feedback loop wherein a particular understanding of perception shapes a particular understanding of the world, which, in turn, shapes a particular understanding of perception, and so forth. In this way, the book is something of a meta-exercise in the history of perception—a way of viewing one's own perceptions in a historical mirror. If one approaches the book with this precept in mind, the experience will be rewarding. ■

10.1126/science.abd3644



An interactive exhibit allows visitors to use their bodies to paint vivid colors onto a virtual canvas.

PHOTO: ROB DESALLE

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



A child who has become malnourished as a result of a measles infection eats lunch in a measles isolation ward of a hospital in the Democratic Republic of the Congo.

PHOTO: HEREWARD HOLLAND/REUTERS/NEWSCOM

Edited by **Jennifer Sills**

COVID-19 policies: Remember measles

Nearly 120 million children in 37 countries are at risk of missing their measles-containing vaccine (MCV) shots this year, as preventive and public health campaigns take a back seat to policies put in place to contain coronavirus disease 2019 (COVID-19) (*1*). In March, the World Health Organization (WHO) issued guidelines indicating that mass vaccination campaigns should be put on hold to maintain physical distancing and minimize COVID-19 transmission (*2*). The disruption of immunization services, even for short periods, will lead to more susceptible individuals, more communities with less than the 95% MCV coverage needed for herd immunity, and therefore more measles outbreaks globally. A mere 15% decrease in routine measles vaccinations—a plausible result of lockdowns and disruption of health services—could raise the burden of childhood deaths by nearly a quarter of a million in poorer countries (*3*). Solutions for COVID-19, especially among the global poor, cannot include forgoing vaccinations.

Suspending mass vaccination campaigns is particularly pernicious in countries embroiled in conflict. Malnutrition among children and mothers in these regions is all too common—and increases mortality for those also infected by measles (*4*). In the conflict-affected regions of Ituri and North Kivu in eastern Democratic Republic of the Congo (DRC) alone, COVID-19 precautions are expected to delay measles vaccinations for nearly a million children (*1*).

COVID-19 mortality has so far been low in most poorer countries and trivial compared with the immediate risk of mortality due to measles. In the DRC, for example, the total number of deaths due to COVID-19 as of June was 37 (*5*), whereas the 2019 measles outbreak in the conflict region of eastern DRC killed more than 6000 people in a matter of months, mostly children (*6*). In some poor countries, more than 40% of the population is younger than 14 years old (*7*). This demographic group is the least affected by COVID-19 (in terms of symptoms, mortality, and transmission) (*8*) but at serious risk of measles (*9*). In addition, because many people in these countries live in cramped housing with unavoidable crowding, policies to minimize COVID-19 at the expense of measles prevention may not even be effective (*10*).

Postponement of measles campaigns prioritizes panic-driven policies for controlling COVID-19 without consideration of these policies' costs. If children are not immunized within the correct age window, they forgo benefits of lifelong immunity (*11*). Whole cohorts of children may be left unprotected. We must balance priorities of containing COVID-19 with efforts to control other high-transmission disease threats in poor countries, especially those affected by conflict.

Debarati Guha-Sapir[1]*, Maria Moitinho de Almeida[1], Mory Keita[2], Gregg Greenough[3], Eran Bendavid[4]
[1]Centre for Research on Epidemiology of Disasters, Research Institute on Health and Society, Université Catholique de Louvain, Brussels, Belgium. [2]World Health Organization, Regional Office for Africa, Dakar Hub, Dakar, Senegal. [3]Brigham and Women's Hospital, Department of Emergency Medicine, Boston, MA 02115, USA. [4]Stanford University, Center for Health Policy, Stanford, CA 94305, USA.
*Corresponding author.
Email: debarati.guha@uclouvain.be

REFERENCES AND NOTES

1. The Measles & Rubella Initiative, "More than 117 million children at risk of missing out on measles vaccine, as COVID-19 surges" (World Health Organization, Geneva, 2020).
2. WHO, "Guiding principles for immunization activities during the COVID-19 pandemic: Interim guidance" (2020); https://apps.who.int/iris/handle/10665/331590.
3. T. Roberton *et al.*, *Lancet* **8**, E901 (2020).
4. WHO, "Measles vaccines: WHO position paper—April 2017," *Weekly Epidemiological Record/Relevé Épidémiologique Hebdomadaire* **92**, 205 (2017).
5. Worldometer, Congo (2020); www.worldometers.info/coronavirus/country/congo/.
6. WHO, Africa, "Deaths from Democratic Republic of the Congo measles outbreak top 6000" (2020).
7. J. Ausubel, "Populations in Europe, Northern America are older than in other regions" (Pew Research Center, 2020).
8. CDC COVID-19 Response Team, *MMWR Morb. Mortal. Wkly. Rep.* **69**, 422 (2020); www.cdc.gov/mmwr/volumes/69/wr/mm6914e4.htm.
9. A. Minetti *et al.*, *PLOS Med.* **10**, e1001544 (2013).
10. G. Bhardwaj *et al.*, "Cities, crowding, and the coronavirus: Predicting contagion risk hotspots" (World Bank Group, Working Paper, Open Knowledge Repository, 2020); https://openknowledge.worldbank.org/handle/10986/33648.
11. S. Soucheray, "Reviews suggest first measles vaccination typically best after 9 months of age" (Center for Infectious Disease Research and Policy, 2019).

10.1126/science.abc8637

Published by AAAS

Restart science stronger after COVID-19

As the scientific enterprise and its supporting institutions rebuild from the disruptions caused by the coronavirus disease 2019 (COVID-19) pandemic, we must support the graduate students and young investigators whose developing careers have been interrupted or irretrievably damaged. Even before the pandemic, the next generation of the scientific workforce faced problems such as limited job opportunities and narrow career training, inadequate mentoring, and a less diverse and inclusive scientific establishment (*1*, *2*). Studies of these challenges have yielded recommendations for at least partial solutions, but few changes have been implemented (*1*, *2*). The pandemic has now increased the urgency of these issues, which must be addressed if science is to thrive in the future.

More than 60% of new Ph.D.s will not pursue careers in academic research. Graduate curricula must therefore be at least partially redesigned to accommodate their broader career interests (*3*). Mentorship quality can be improved with appropriate faculty training (*4*). The lack of diversity, equity, and inclusion in science must be taken seriously and ameliorated. Resources may have to be redistributed to ensure that younger scientists have independent careers earlier than has become common. Let us not waste the opportunity presented by this crisis to address long-standing problems. As we rebuild, we should look hard at our own enterprise, make some changes, and restart science stronger.

Alan I. Leshner
CEO Emeritus, American Association for the Advancement of Science and Former Executive Publisher, *Science*, Washington, DC 20005, USA.
Email: alan.i.leshner@gmail.com

REFERENCES AND NOTES

1. National Academies of Sciences, Engineering, and Medicine, *Graduate STEM Education for the 21st Century* (National Academies Press, 2018).
2. National Academies of Sciences, Engineering, and Medicine, *The Next Generation of Biomedical and Behavioral Sciences Researchers: Breaking Through* (National Academies Press, 2018).
3. Council of Graduate Schools, *Understanding Career Pathways for Program Improvement* (Council of Graduate Schools, 2015).
4. National Academies of Sciences, Engineering, and Medicine, *The Science of Effective Mentorship in STEMM* (National Academies Press, 2019).

10.1126/science.abd5780

Uphold the nuclear weapons test moratorium

The Trump administration is considering renewing nuclear weapons testing (*1*), a move that could increase the risk of another nuclear arms race as well as an inadvertent or intentional nuclear war. Following in the long tradition of scientists opposing nuclear weapons due to their harmful effects on both humanity and the planet (2), we ask the U.S. government to desist from plans to conduct nuclear tests.

During the Cold War, the United States conducted 1030 nuclear weapons tests, more than all other nuclear-armed nations combined (*3*). In 1996, the United States signed the Comprehensive Nuclear Test Ban Treaty (CTBT), agreeing not to conduct a nuclear weapons test of any yield (*4*). The United States has not yet ratified the CTBT but did spearhead the 2016 adoption of UN Security Council Resolution 2310, which calls upon all countries to uphold the object and purpose of the CTBT by not conducting nuclear tests (*5*).

Eight of the nine nuclear-armed states, including the five permanent members of the UN Security Council, have observed a moratorium on nuclear testing since 1998 (*3*, *4*). The ninth, North Korea, responding to international pressure, stopped testing warhead detonations (as opposed to missile flights) in 2017 (*6*). If the United States ratified the CTBT, joining the 168 countries who have already done so (*4*), there is a good chance that the other holdout countries would ratify the treaty as well (*7*).

In contrast, restarting U.S. nuclear weapons tests of any size, underground or aboveground, would give license to other countries, such as North Korea, India, and Pakistan, to resume testing. If the tests are underground, radioactive materials could leak into the local environment, including water supplies (*8*); if in the atmosphere, which is currently prohibited by the 1963 Limited Nuclear Test Ban Treaty (*9*), such tests would spread radioactivity, sometimes very widely (*8*). Once the United States breaches the treaty, there will be no way to prevent other nations from carrying out tests of larger warheads or to control leaks into the environment and atmosphere. Even a "limited" nuclear exchange between nuclear-armed nations can cause untold local death and destruction, as well as global climate and agricultural catastrophes stemming from the climate impacts of smoke from fires ignited by nuclear weapons (*10*).

The current U.S. arsenal includes thousands of warheads, together capable of obliterating every major city in any country on Earth. Yet the United States has embarked on a $1.7 trillion nuclear weapons enhancement program (*11*), of which the proposed testing would be one small—but dangerous—component. All nations, including the United States, should continue to reduce the number of nuclear warheads in the world's arsenals, not increase their efficacy or develop more lethal versions. Senator Edward Markey and Senate colleagues recently announced the Preserving Leadership Against Nuclear Explosives Testing (PLANET) Act, which would deny funding for and thereby prevent the renewal of testing (*12*). We urge the Senate to pass this bill and to ratify the CTBT immediately.

Jonathan King*, Martin Chalfie, Noam Chomsky, Joseph Cirincione, Sean Decatur, Melissa Franklin, Joseph Gerson, David P. Goldenberg, Gary Goldstein, William Hartung, Ira Helfand, Daniel Holz, Peter C. Kahn, Sheldon Krimsky, Edward Loechler, Val Moghadam, Stuart A. Newman, David Ozonoff, Prasannan Parthasarathi, William Phillips, H. David Politzer, Robert P. Redwine, Richard J. Roberts, Alan Robock, Catherine Ann Royer, Suzanne Scarlata, Elaine Scarry, George F. Smoot, Robert Socolow, Susan Solomon, Andrew Strominger, Eric J. Sundberg, Mriganka Sur, Max Tegmark, John F. Tierney, Cornelia van der Ziel, Michael VanElzakker, Frank N. von Hippel, Lawrence Wittner, Henry H. Wortis
*Corresponding author. Email: jaking@mit.edu
The full author affiliations are available online.

REFERENCES AND NOTES

1. J. Hudson, P. Sonne, "Trump administration discussed conducting first U.S. nuclear test in decades," *The Washington Post* (2020).
2. J. C. Jolly, *Endeavor* **26**, 149 (2002).
3. Arms Control Association, The Nuclear Testing Tally (2019); www.armscontrol.org/factsheets/nucleartesttally.
4. Comprehensive Nuclear Test Ban Treaty (CTBT) (2020); www.nti.org/learn/treaties-and-regimes/comprehensive-nuclear-test-ban-treaty-ctbt/.
5. Arms Control Association, UN Security Council Backs CTBT (2016); www.armscontrol.org/ACT/2016_10/News/UN-Security-Council-Backs-CTBT.
6. Arms Control Association, Chronology of U.S.-North Korean Nuclear and Missile Diplomacy (2020); www.armscontrol.org/factsheets/dprkchron.
7. D. Choubey, "The CTBT's importance for U.S. national security" (Carnegie Endowment for International Peace, 2009).
8. Comprehensive Nuclear Test Ban Treaty Organization, General Overview of the Effects of Nuclear Testing (www.ctbto.org/nuclear-testing/the-effects-of-nuclear-testing/general-overview-of-theeffects-of-nuclear-testing/).
9. John F. Kennedy Presidential Library and Museum, Nuclear Test Ban Treaty (www.jfklibrary.org/learn/about-jfk/jfk-in-history/nuclear-test-ban-treaty).
10. A. Witze, *Nature* **579**, 485 (2020).
11. Arms Control Association, U.S. Nuclear Modernization Programs (2018); www.armscontrol.org/factsheets/USNuclearModernization.
12. Ed Markey, U.S. Senator for Massachusetts, "Senator Markey, Senate colleagues announce legislation to prohibit restart of U.S. nuclear weapons testing" (2020).

COMPETING INTERESTS

D.H., R.S., and S. Solomon are members (in an unpaid advisory capacity) of the Science and Security Board of the Bulletin of the Atomic Scientists. J.F.T. is affiliated with the Center for Arms Control and Non-Proliferation.

SUPPLEMENTARY MATERIALS

Author affiliations
www.sciencemag.org/content/369/6501/262.2/suppl/DC1

10.1126/science.abd3313

Published by AAAS

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



A beetle outfitted with a minuscule, autonomous, steerable camera

PHOTO: UNIVERSITY OF WASHINGTON/MARK D. STONE

COMPUTER VISION

Beetlecam

Artificial, insect-scale systems are subject to weight, size, area, and power constraints. Iyer *et al.* report a lightweight vision system small enough to be mounted onto a live beetle. The vision system is integrated with energy-saving electronics that enable operation for up to 6 hours. The camera is triggered only when motion is detected with an accelerometer, and a wirelessly controlled piezo actuator steers the camera, enabling a broader field of view and the ability to capture panoramas. —MML

Sci. Robot. **5**, eabb0839 (2020).

CHEMICAL PHYSICS

Liquid-liquid critical point of water

Despite a broad range of experimental observations that indirectly point out the possible existence of a liquid-liquid phase transition in deeply supercooled water, no unambiguous experiment has shown this yet. The challenges of performing experiments under such conditions are associated with the inevitable rapid crystallization of the metastable liquid state, and therefore computational simulations become a crucial alternative. Using two accurate classical models of water and long (more than 40 microseconds) isobaric-isothermal molecular dynamics simulations, Debenedetti *et al*. provide strong computational evidence of the presence of a second metastable liquid-liquid critical point in the deeply supercooled region that has critical behavior consistent with the three-dimensional Ising universality class. —YS

Science, this issue p. 289

HUMAN EVOLUTION

A genetic history of China

The history of human movements into and within China has been difficult to determine solely from archaeological investigations or genetic studies of contemporary peoples. Yang *et al*. sequenced DNA from 26 individuals from 9500 to 300 years ago from locations within China. Analyses of these individuals, along with previously sequenced ancient individuals and present-day genomes representing global populations, show a split between ancient humans in northern and southern China. Neolithic northern Chinese individuals are closest to modern-day East Asians, whereas ancient individuals from southern China are most closely related to modern-day Southeast Asians and show an affinity to modern-day Austronesian populations. These results indicate that there was a southward movement and admixture of peoples during the Neolithic that gave rise to modern-day populations in East Asia. —LMZ

Science, this issue p. 282

ORGANIC CHEMISTRY

An E-Z boron swivel

Compounds with carbon-carbon double bonds can form two distinct isomers, depending on whether the heaviest substituents on both carbons lie on the same side (labeled Z) or diagonally across from each other (labeled E). Molloy *et al*. present a convenient method to reorient double bonds that bear boron and carbonyl substituents. When they are diagonally opposed, both substituents stay in plane, and the double bond is easily swiveled by photosensitization. However, once on the same side as the carbonyl, the boron rotates out of plane and further sensitization is inhibited. —JSY

Science, this issue p. 302

EDUCATION

Belonging improves student retention

Universities trying to increase inclusivity need to better retain students of color and first-generation students. Murphy *et al*. found that an hour-long classroom intervention focused on students' sense of belonging increased the grade point averages of disadvantaged students (especially the most struggling students) in the semester of treatment, improved feelings of

Published by AAAS

academic and social fit 1 year later, and increased retention in each of the following 2 years. Students read stories from upper-year students about the normalcy and temporariness of common challenges of college attendance and completed two writing assignments connecting their own experiences. The double-blind, randomized intervention, conducted through first-year writing courses at a broad-access Hispanic-serving institution shows that the correct short intervention can significantly affect retention. —JE

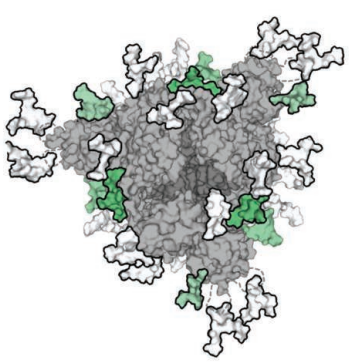
Sci. Adv. 10.1126/sciadv.aba4677 (2020).

CORONAVIRUS

SARS-CoV-2 spike protein, elaborated

Vaccine development for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is focused on the trimeric spike protein that initiates infection. Each protomer in the trimeric spike has 22 glycosylation sites. How these sites are glycosylated may affect which cells the virus can infect and could shield some epitopes from antibody neutralization. Watanabe *et al*. expressed and purified recombinant glycosylated spike trimers, proteolysed them to yield glycopeptides containing a single glycan, and determined the composition of the glycan sites by mass spectrometry. The analysis provides a benchmark that can be used to measure antigen quality as vaccines and antibody tests are developed. —VV

Science, this issue p. 330



Glycosylations, outlined in bold, extensively modify the surface of the SARS-CoV-2 spike protein.

IMMUNOLOGY

Prenatal antibodies are polyreactive

Extensive immunoglobulin gene rearrangements allow humans to recognize a diversity of potential pathogens. This antibody repertoire is more restricted during early life to prevent the generation of autoreactive B cells, though tolerance does not appear to be complete. Chen *et al.* examined the reactivities of antibodies cloned from individual human fetal B cells residing in the liver, bone marrow, and spleen. They observed the accumulation of autoreactive and polyreactive B cells, which were frequently cross-reactive to commensals in the absence of any somatic hypermutation. The generation of these reactive B cells before they are ever exposed to microbes may promote later beneficial commensal-host interactions and/or augmented host defense during the first weeks of life. —STS

Science, this issue p. 320

PIEZOELECTRICS

Simple strain enhancement

Piezoelectric materials are important as sensors and transducers for applications such as ultrasonics. Liu *et al*. discovered nanopillar regions in a sodium-niobium-oxide film that substantially improve the piezoelectric properties (see the Perspective by Bassiri-Gharb). These nanopillar regions reverse where the cations and anions are located in the crystal structure, with a distinctive boundary in between. This difference in structure results in a strain-sensitive polarity that enhances the piezoelectric properties in a chemically simple material. —BG

Science, this issue p. 292; see also p. 252

IN OTHER JOURNALS

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



Microbumps on insect and leaf surfaces generate tiny shock waves that shatter raindrops and make them rebound, thus reducing wetting.

BIOPHYSICS

The shattering effect on rain

Raindrops falling at terminal velocity are giant ballistic objects when viewed by a butterfly or felt by a leaf. So how do smaller organisms survive a rainy day? Many biological surfaces have nanoscale texture and show superhydrophobicity. Kim *et al*. found that high-speed water drops do not simply bounce off of these nanostructured surfaces. On impact, each drop spreads, shock-like waves form, and the drop shatters. As the liquid spreads across the "rough" surface, the shock waves nucleate holes in the upper liquid film, the holes coalesce, the liquid rebounds, and contact time is reduced by 70%. This is also a perfect mechanism to send waterborne fungal spores flying. Thus, superhydrophobicity means that an organism out in the rain gets neither wet nor cold and can carry on flying unperturbed, as birds and butterflies do. —CA

Proc. Natl. Acad. Sci. U.S.A. **117**, 13901 (2020).

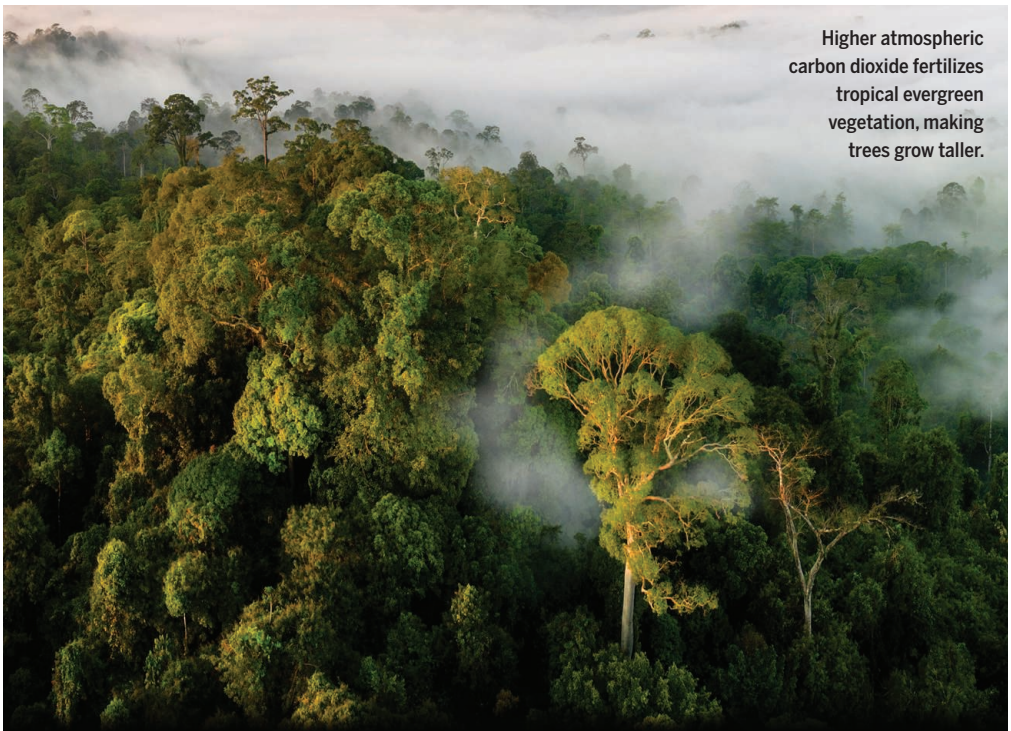
CHEMICAL PHYSICS

The amazing chemistry of three atoms

The elementary reaction $H + HD \rightarrow H_2 + D$ is currently of considerable interest, partially motivated by the unusual signatures observed experimentally and attributed to quantum interference and associated geometric phase effects. Using rigorous quantum reactive scattering calculations, Zhou *et al*. report the presence of another unexpected feature in this system, a narrow, long-lived Feshbach resonance at

CREDITS (FROM LEFT): WATANABE *ET AL.*; KIM *ET AL.*, *PROC. NATL. ACAD. SCI. U.S.A.* **117**, 13901 (2020)

Published by AAAS



Higher atmospheric carbon dioxide fertilizes tropical evergreen vegetation, making trees grow taller.

CLIMATE CHANGE

The tropical Asian carbon sink

Tropical vegetation plays a key role in the global carbon cycle, and much current research focuses on the extent to which it will continue to act as a carbon sink in the coming decades as the climate continues to warm. Scheiter *et al.* modeled the responses of tropical Asian vegetation under scenarios of future climate change. Their results indicate that woody biomass in the region is likely to increase and natural evergreen vegetation will get taller owing to the fertilization effect of increasing carbon dioxide, a result that is robust under different climate models until the end of this century. However, the resulting persistence of a carbon sink also depends on effective conservation measures to ensure that the forest vegetation persists too. —AMS

Glob. Chang. Biol. 10.1111/GCB.15217 (2020).

energies far below the reaction barrier, which has not been observed before. The present theoretical results are expected to stimulate further in-depth experimental study of this simple yet not completely explored three-atom system of fundamental importance. —YS

J. Phys. Chem. Lett. **11**, 4970 (2020).

QUANTUM MATERIALS

Strain for controlled emission

The optoelectronic properties of atomically thin two-dimensional (2D) materials make them attractive for quantum nanophotonic devices. The range of materials becoming available also provides a wide palette to choose from when selecting the operational wavelength for various applications. Chakraborty *et al*. exploited another property of 2D materials: that they can be reversibly stretched by as much as 10%. By combining an encapsulated monolayer of tungsten diselenide with a piezoelectric substrate, they show that application of local strain to the 2D layer can be used to tune and control emission wavelength as well as the emission axis. This technique could be used to develop an integrated and versatile platform of quantum optoelectronic device technology. —ISO

Optica **7**, 580 (2020).

IMMUNOLOGY

SAA1 sees PAMP in the mite

If you wake up with streaming eyes and nose, it could be that you are allergic to the house dust mite. The molecular mechanisms through which the molecular patterns of such organisms initiate allergic responses is poorly understood. Smole *et al*. report that the highly conserved acute-phase protein serum amyloid A1 (SAA1) acts as a soluble pattern recognition receptor for mite fatty acid–binding allergens Der p 13 and Blo t 13. These allergens disrupt lipid-stabilized SAA1 hexamers to adopt monomeric states that preferentially bind formyl peptide receptor 2 (FPR2). Activation of FPR2 in turn activates epithelial cell release of the potent allergy-inducing cytokine interleukin-33. This work opens new routes for investigating treatment of allergic disease while potentially offering insights into the pathogenesis of other SAA1-associated disorders such as atherosclerosis and Alzheimer's disease. —STS

Nat. Immunol. **21**, 756 (2020).

2D MATERIALS

Synthesizing graphitic sheets

Two-dimensional (2D) materials have a wide range of potential uses in electronics, chemistry, lubrication, and other applications. However, easy synthesis of these materials can be challenging. Mayyas *et al*. synthesized self-exfoliating, 2D graphitic materials with a range of organic solvent precursors using liquid gallium–based electrodes at room temperature. The strategy is flexible because several properties can be tuned, including thickness, porosity, and composition. Synthesis of other types of 2D materials should be possible using this interesting technique. —BG

Adv. Mater. 10.1002/adma.202001997 (2020).

PHYSIOLOGY

Sleeping on the fly

Sleep is more than just refreshing—it is critical for normal body function. Interfering with sleep results in problems with cognition and the gastrointestinal, immune, and circulatory systems; it even reduces life span. In studies of fruit flies and mice, Vaccaro *et al*. show that reactive oxygen species build up in the gut with sleep deprivation. If animals are allowed to sleep, then reactive oxygen species and oxidative stress are reduced. The negative effects of reduced sleep can be reversed with oral antioxidants or expression of antioxidant enzymes in the intestines, such that even animals allowed to have little or no sleep have a normal life span. —BAP

Cell **181**, 1307 (2020).

PHOTO: NATURE PICTURE LIBRARY/ALAMY STOCK PHOTO

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CORAL GENOMICS

Conservation help from genomics

Corals worldwide are under threat from rising sea temperatures and pollution. One response to heat stress is coral bleaching—the loss of photosynthetic endosymbionts that provide energy for the coral. Fuller *et al*. present a high-resolution genome of the coral *Acropora millepora* (see the Perspective by Bay and Guerrero). They were able to perform population genetic analyses with samples sequenced at lower coverage and conduct genome-wide association studies. These data were combined to generate a polygenic risk score for bleaching that can be used in coral conservation. —LMZ

Science, this issue p. 268; see also p. 249

METROLOGY

An optical timekeeper

Optical clocks, based on optical transitions of atoms, operate at much higher frequency than the microwave atomic clocks currently used as our timing standards. They have been shown to exhibit better stability and are poised to redefine the second. The development of stable, self-referenced optical frequency combs that span the microwave to optical wavelengths has been key to these efforts. Diddams *et al*. reviewed developments and refinements of these optical combs over the past 20 years and provide an overview of where they are finding application, from precision timing to high-resolution spectroscopy and imaging, ranging, and navigation. —ISO

Science, this issue p. 267

CORONAVIRUS

Aging immunity

Like most of the human body, the immune system deteriorates with age. This means that vaccination strategies are less effective in older people. These individuals are also prone to a phenomenon of increased general inflammation, termed inflammaging, which can suppress immunity and potentially increase susceptibility to infection. In a Perspective, Akbar and Gilroy discuss the potential causes of inflammaging, how it can affect immunity, and the implications for patients with coronavirus disease 2019 (COVID-19). Because severe COVID-19 seems to mostly occur in older people, the authors discuss whether inflammaging could contribute to the sometimes deadly hyperinflammation and lung damage in these patients. —GKA

Science, this issue p. 256

CANCER

Deconstructing a perilous loop

Cancer cells that give rise to solid tumors reside in a specialized microenvironment containing growth factors that support tumor growth and progression. The mechanisms by which tumor cells within this so-called "niche" attract and sustain these growth factors are poorly understood. Studying a mouse model of squamous cell carcinoma, Taniguchi *et al*. identified a signaling loop initiated by certain tumor cells within the niche that become more invasive on exposure to transforming growth factor β (TGF-β). They show that the tumor cells release the cytokine interleukin-33, which induces the differentiation of nearby myeloid cells into macrophages. These macrophages in turn release TGF-β, which feeds back to the tumor cells, promoting their malignant progression. —PAK

Science, this issue p. 269

NEURODEVELOPMENT

Brain synapses through the life span

Excitatory synapses connect neurons in the brain to build the circuits that enable behavior. Cizeron *et al*. surveyed synapses in the mouse brain from birth to old age and present the data as a community resource, the Mouse Lifespan Synaptome Atlas (see the Perspective by Micheva *et al*.). Molecular and morphological features defined 37 subtypes of synapses. Although synapse density generally increased in early development and declined in old age, the details differed in different brain areas. —PJH

Science, this issue p. 270; see also p. 253

CANCER

Ironing out a survival strategy

Metastatic cells display a notable ability to adapt to—and even thrive in—harsh microenvironments. One extreme example is leptomeningeal metastases (LM), cancer cells that enter a region of the central nervous system called the subarachnoid space, which is filled with cerebral spinal fluid (CSF). This anatomic site has a limited supply of micronutrients such as iron, and it harbors immune cells. Chi *et al*. used single-cell RNA sequencing to study CSF samples from cancer patients with LM (see the Perspective by Garzia and Taylor). They found that LM cells express and use components of a high-affinity iron-capturing system. Through this mechanism, the LM cells avoid the adverse effects of iron deprivation and potentially escape immune attack by limiting the supply of iron to macrophages. —PAK

Science, this issue p. 276; see also p. 250

CHEMICAL PHYSICS

Molecule-molecule forward scattering

The classical picture of nonreactive inelastic scattering implies that a considerable transfer of translational to rotational energy should be accompanied by backward scattering. Using cross–molecular beam scattering combined with velocity map imaging, Sun *et al*. provide evidence of an unusual pair-correlated channel in the collision of carbon monoxide (CO) molecules, specifically ^{13}CO with ^{12}CO, both initially in the ground rotational state. These two molecules were scattered into the forward direction, both finally in the high rotational states. The experimental work is supported by quasiclassical trajectory simulations that exhibit unexpected molecular dance moves and suggest that similar effects could be expected for other molecule-molecule scattering systems. —YS

Science, this issue p. 307

CORONAVIRUS

Blighted Gotham

Deaths caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in New York City (NYC) during the spring of 2020 have vastly exceeded those reported in China and many other countries. What were the early events that led to such a severe outbreak? Gonzalez-Reiche *et al*. sampled some of the early patients seeking assistance in February and March of 2020 at the Mount Sinai Health System. Phylogenetic analysis of virus sequences in these people, who were drawn from across NYC, showed that the virus had been independently introduced many times from Europe and elsewhere in the United States. Subsequent clusters of community transmission occurred. The focus of infection in NYC is a marker of the role this city plays

Published by AAAS

as a two-way hub for human movement. —CA

Science, this issue p. 297

ORGANIC SEPARATIONS

Separating organics without distillation

Hydrocarbon distillation is a widespread and energy-intensive process. Membranes might offer an alternative approach, but few can survive immersion in organic solvents nor are they able to extract relatively small molecules. Thompson *et al*. developed a series of polymers of intrinsic microporosity that they used for membrane-based separations of organic compounds in an organic solvent (see the Perspective by Brennecke and Freeman). The new membrane has a molecular weight cutoff of 253 daltons, far lower than existing ones closer to 600 daltons. The polymers were used to separate light shale crude oil and succeeded in fractionation of molecular weights of about 170 daltons. —MSL

Science, this issue p. 310; see also p. 254

CORONAVIRUS

Weathering the weather

In some quarters, it is hoped that increased humidity and higher temperatures over the Northern Hemisphere in the summer will snuff out the 2020 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. In reality, the situation is likely to be more complicated than that. Baker *et al.* used a climate-dependent epidemic model to simulate the SARS-CoV-2 pandemic, testing different scenarios of climate dependence based on known coronavirus biology. Levels of susceptibility among the population remain the driving factor for the pandemic, and without effective control measures, the pandemic will persist in the coming months, causing severe outbreaks even in humid climates. Summer will not substantially limit pandemic growth. —CA

Science, this issue p. 315

AGING

Programmed aging in yeast cells

Following the fate of individual yeast cells has revealed aging to be more of a programmable decision process rather than a simple accumulation of deleterious events. Li *et al*. combined single-cell studies and mathematical modeling to show that yeast cells showed two different forms of aging: one with more ribosomal DNA silencing, in which nucleoli were degraded, and another with more heme accumulation and heme-dependent transcription, in which mitochondria were more affected. Overexpression of the lysine deacetylase Sir2, which contributes to ribosomal DNA silencing, led to a third cell-aging fate in which the average life span was extended. If other cells age in similar ways, then this study may provide new ways to consider dynamics of aging and strategies to extend the health span. —LBR

Science, this issue p. 325

GENOME EDITING

Compact defense system in bacteriophages

The CRISPR-Cas system, naturally found in many prokaryotes, is widely used for genome editing. CRISPR arrays in the bacterial genome, derived from the genome of invading viruses, are used to generate a CRISPR RNA that guides the Cas enzyme to destroy repeat viral invaders. Recently, an unexpectedly compact CRISPR-Cas system was identified in huge bacteriophages. Pausch *et al*. show that even though this system lacks commonly found accessory proteins, it is functional. In addition to a CRISPR array, the only component of the system is an enzyme called CasΦ, which uses the same active site to process transcripts of the CRISPR arrays into CRISPR RNA and to destroy foreign nucleic acids. This system, which is active in human and plant cells, provides a hypercompact addition to the genome-editing toolbox. —VV

Science, this issue p. 333

ISCHEMIA

Optimizing antithrombotic therapy

Antiplatelet drugs are used for preventing and treating thrombotic diseases. However, they have serious adverse effects, including risk of excessive bleeding. mP6, a synthetic peptide targeting integrin signaling, has been shown to inhibit thrombosis without affecting bleeding, but its application has been hindered by inefficient and impractical intracellular delivery. Pang *et al*. developed lipid-stabilized high-loading peptide nanoparticles (HLPNs) for efficient in vivo delivery of the antithrombotic peptide. Injection of HLPNs loaded with engineered mP6 prevented thrombosis without affecting bleeding in mice and had therapeutic effects in a model of ischemia-reperfusion injury. This strategy might be successful in minimizing the side effects of current antiplatelet therapy. —MM

Sci. Transl. Med. **12**, eaaz7287 (2020).

PLANT BIOLOGY

Glutamate channels signal distress

Wounded plants generate Ca^{2+} waves and electrical signals that propagate from the wound site to distal tissues. Shao *et al*. found that root wounding or glutamate application to wounded roots induced these systemic signals in the model plant *Arabidopsis thaliana*. This required the glutamate receptor–like proteins GLR3.3 and GLR3.6 and inhibition of the proton pump AHA1 (see the Focus by Fichman *et al*.). GLR3.3 and GLR3.6 functioned as pH-sensitive, glutamate-gated Ca^{2+} channels. Glutamate leakage and increases in pH caused by wounding may lead to GLR activation and systemic wound signaling. —AMV

Sci. Signal. **13**, eaba1453, eabb9505 (2020).

IMMUNODEFICIENCY

Tempering lymphocyte activation

Diagnosis and treatment of inborn errors in immunity remain a challenge. Salzer *et al.* studied two siblings who were hospitalized for recurrent pediatric infections. Exome sequencing and functional analyses helped to identify homozygous loss-of-function mutations in genes encoding the protein HEM1 in both individuals. HEM1 is a component of actin-remodeling complexes that have a wide range of functions, from cell migration to cell division. By studying immune cells from these individuals and by generating HEM1-deficient mice, the authors established HEM1 as a key regulator of T and B cell functions. They characterized in depth the role of HEM1 in B cells and propose that HEM1 modulates signaling downstream of the B cell receptor. —AB

Sci. Immunol. **5**, eabc3979 (2020).

Published by AAAS

REVIEW SUMMARY

METROLOGY

Optical frequency combs: Coherently uniting the electromagnetic spectrum

Scott A. Diddams*, Kerry Vahala*, Thomas Udem*

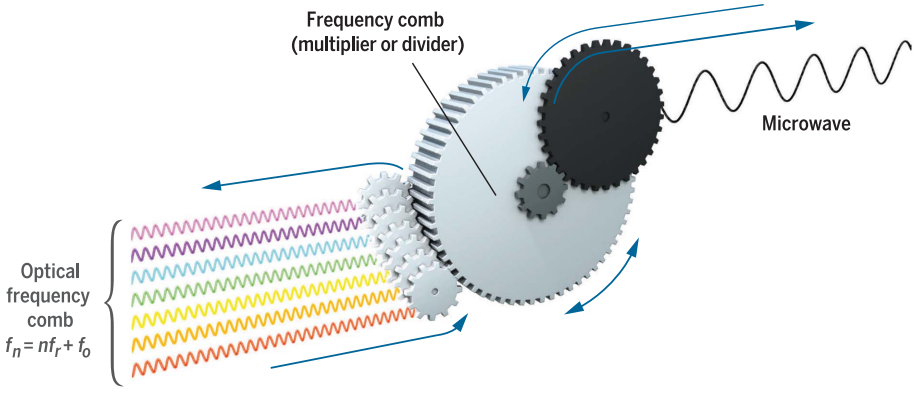
BACKGROUND: The generation and control of coherent electromagnetic waves, such as those provided by electronic oscillators in the radio frequency domain or lasers in the optical domain, have had an unparalleled impact on human society over the past century. For example, precise timing with radio waves referenced to atomic transitions undergirds navigation with the Global Positioning System. And modern communication systems are built around the properties of such waves to carry data through the air or within optical fibers that circumnavigate the globe. Such technical advances rely upon an exceptionally well developed and unified theoretical understanding of radio and optical waves, as well as the devices for their generation and control. Nonetheless and surprisingly, just 20 years ago, radio and optical technology realms remained largely distinct and isolated from one another. Although light waves could be modulated at radio rates and likewise electrical currents could be produced by demodulating optical signals, a simple coherent connection between radio and optical fields did not exist. As a result, many common technologies for the synthesis and control of radio frequencies, including those central to navigation, communications, and measurement, seemed futuristic for optics. Conversely, the full potential of optics for time standards, metrology, and science was not accessible, despite a decades-long recognition of the opportunities and attempts to harness their technological impact. This situation, which arose as a consequence of the enormously high frequencies of electromagnetic waves in the optical domain, thereby limited scientific progress and technical capabilities.

ADVANCES: The invention of the laser in 1960 represented an optical analog to radio oscillators, which were invented much earlier. This development motivated efforts to create a coherent bridge between the radio and optical realms with multiple oscillators of successively higher frequencies being chained together. However, the sheer complexity and size of such approaches made it clear that these systems would never be widely available and that their capability would be limited. Such efforts were upended after nearly four decades of work when an unanticipated breakthrough enabled by combined advances in femtosecond laser technology, nonlinear optics, and precision frequency metrology finally solved this problem. The key to overcoming these issues was a new approach to generate and control the spectrum of a mode-locked laser, which was called an optical frequency comb in accordance with the regularly spaced comb of frequencies it contained. Even though mode-locked lasers and optical frequency combs had existed previously, in 2000 it was demonstrated how their spectra could be expanded over an octave of optical bandwidth. This critical advance enabled the technique of self-referencing, by which the optical frequencies of the comb are locked in an absolute sense to a radio frequency reference. In a simple and elegant way, this produced an optical clockwork, analogous to a gearbox (see the figure), that provided the bidirectional coherent connection between the optical and radio frequency domains. Moreover, because this connection is simultaneously established to many thousands of optical frequencies within the comb, self-referencing makes the coherent translation of frequencies possible across hundreds of terahertz of the optical spectrum.

ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aay3676

OUTLOOK: In the two decades since the introduction of the frequency comb, entirely new scientific and technology vistas have been opened. New applications are leveraging an unprecedented sharing of the respective strengths of electronics and photonics, as well as a new freedom to work seamlessly across the broad optical frequency expanse. Frequency combs are used to realize and compare ultraprecise optical clocks at the 19th decimal place, which provides powerful approaches to test relativity and quantum theory as well as the search for physics beyond our present understanding. In addition, combs transfer this exceptional precision across the spectrum to perform massively parallel spectroscopy, generate the lowest-noise microwaves and few-cycle attosecond waveforms, and even help astronomers search for Earth-like exoplanets. Just as applications of frequency combs have expanded, there are new developments in frequency comb technology. Among these are combs that do not use mode locking for comb formation, as well as new approaches built on integrated photonics and microresonators that enable frequency combs on a silicon chip. Indeed, it is now clear that the once-enormous gap between optics and electronics will be further blurred as the coherent electromagnetic spectrum is united by frequency combs. ■



An optical clockwork. A frequency comb is a type of laser that functions in a manner analogous to a set of gears, as illustrated here, to link radio frequencies to a vast array of optical frequencies ("the comb") with values precisely determined by the expression $f_n = nf_r + f_0$, where f_r and f_0 are radio frequencies and n is an integer on the order of 10^5.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: scott.diddams@nist.gov (S.A.D.); vahala@caltech.edu (K.V.); thu@mpq.mpg.de (T.U.)
Cite this article as S. A. Diddams *et al.*, *Science* **369**, eaay3676 (2020). DOI: 10.1126/science.aay3676

REVIEW

METROLOGY

Optical frequency combs: Coherently uniting the electromagnetic spectrum

Scott A. Diddams[1,2]*, Kerry Vahala[3]*, Thomas Udem[4]*

Optical frequency combs were introduced around 20 years ago as a laser technology that could synthesize and count the ultrafast rate of the oscillating cycles of light. Functioning in a manner analogous to a clockwork of gears, the frequency comb phase-coherently upconverts a radio frequency signal by a factor of ≈10^5 to provide a vast array of evenly spaced optical frequencies, which is the comb for which the device is named. It also divides an optical frequency down to a radio frequency, or translates its phase to any other optical frequency across hundreds of terahertz of bandwidth. We review the historical backdrop against which this powerful tool for coherently uniting the electromagnetic spectrum developed. Advances in frequency comb functionality, physical implementation, and application are also described.

Electromagnetic waves, whether microwave signals transmitted by a cellular phone or visible light given off by a lamp, are governed by Maxwell's equations. However, this unified physical framework belies a technological disparity between the electronic and the optical bands of the electromagnetic (EM) spectrum: Optical frequencies are far too fast to count, measure, and phase-coherently control in a manner that has become routine for the EM signals at electronic rates (hereafter also called radio or microwave frequencies). This difference in frequency of a factor of 10^5 is more than an obvious fact; it has led to a scientific and technological gulf between the radio and optical realms over the history of human-generated and -controlled EM oscillators that began with the revolutionary experiments of Hertz (*1*) (Fig. 1).

This gap has limited scientific and technical capabilities, as there was not a straightforward means to accurately connect the frequency and phase of optical oscillators to those in electronic circuits. Notably, in the field of timekeeping and frequency standards, this meant that there was no counting mechanism for clock oscillators at hundreds of terahertz that were long predicted to be more accurate and precise than their microwave counterparts. Moreover, even though frequency is the most precisely measured physical quantity, optical spectroscopy of energy eigenstates in atoms and molecules had to rely on much less accurate interferometric measurements of wavelength. In some cases, this limited tests of fundamental physics. And although lasers existed since the 1960s and have proven to be exceptionally efficient carriers of data in long-haul fiber telecommunications networks, the agility with which optical waves can be synthesized, controlled, and manipulated falls short of the analog and digital radio frequency techniques that underlie modern cellular and computer networks.

This situation was radically changed around the turn of the century with the introduction of optical frequency combs that provided an elegant way to coherently unite the electromagnetic spectrum. Frequency combs provide a coherent and bidirectional link between radio frequency electronics and optics. The technologies belonging to these disparate realms underlie the key technical advancements of this age, and a sharing of their performance and other attributes has been enabled by the frequency comb linkage. Not surprisingly, the resulting unification of the EM spectrum has, in the brief 20 years since the introduction of the frequency comb, unleashed enormous new capabilities and created unforeseen technologies, the impact of which was both recognized and foreshadowed in awarding one-half of the 2005 Nobel Prize in Physics to Theodor Hänsch and John Hall (*2*, *3*).

Beginnings: Connecting the radio and optical domains

In 1886, Heinrich Hertz succeeded in artificially generating short bursts of radio waves and detecting them at a few meters' distance by observing sparks in an antenna-like structure (*1*, *4*). This discovery was inspired by Maxwell's equations (*5*) that had predicted the existence of electromagnetic waves 13 years earlier. It took several decades until oscillators based on vacuum-tube feedback were introduced for generating continuous and controllable coherent waves (*6*). The possibility of generating and receiving radio waves changed the world and is still evolving, as evidenced by the continued advances in wireless communications.

Until 1960, only incoherent sources were available at the much higher frequencies of light. The first important step in gaining coherent control over optical frequencies was a landmark paper by Charles Townes and Arthur Schawlow in which they pointed out how to extend the radio frequency technology of the maser, which Townes and co-workers had introduced earlier (*7*), into the optical realm (*8*). The first realization of this "optical maser," the device that is now called a "laser," was made by Theodore Maiman (*9*). And like Hertz's radio source, it emitted pulses of radiation. Shortly thereafter, Ali Javan and coworkers built the first continuously operating laser (*10*) and could demonstrate that it generated coherent light waves by observing a beat note between two waves (*11*). Unfortunately, the beat notes allowed only measurement of the frequency difference between two lasers. What was then called the "absolute" optical frequency remained orders of magnitude too large to measure directly, as was possible for radio frequencies. As a result, the radio and optical frequency realms remained separated, as sketched in the left part of Fig. 1.

At the same time, it was realized that control of the laser light waves would open up many possibilities in science and technology. An important example would be the optical version of radio frequency atomic clocks that were in those early days built in the form of a maser. Using the much higher frequencies of optical transitions in atoms held promise for a substantial boost in clock accuracy because the faster rate slices time into finer intervals. However, available electronic counting devices were orders of magnitude too slow, and they remain so even today. Therefore, the idea emerged to connect the optical and radio frequency realms by inching through the vast gap separating them by successive multiplication of a radio frequency reference, such as one derived from a radio frequency atomic clock. Early efforts to measure the frequency of laser light reached the CN laser at 964,312.3 MHz in 1967 (*12*). Through this and other work, the harmonic frequency chain was born (*13*, *14*) as a complex radio–optical system that foreshadowed the introduction of the vastly simpler optical frequency comb nearly four decades later. Figure 2A explains the concept of the harmonic frequency chain, and Fig. 2B shows its most refined implementation that phase-coherently connected a radio frequency input with a visible laser output (*15*). Other approaches have been proposed to count optical

[1]Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO, USA. [2]Department of Physics, University of Colorado, Boulder, CO, USA. [3]T. J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, CA, USA. [4]Max-Planck-Institut für Quantenoptik, Garching, Germany.
*Corresponding author. Email: scott.diddams@nist.gov (S.A.D.); vahala@caltech.edu (K.V.); thu@mpq.mpg.de (T.U.)

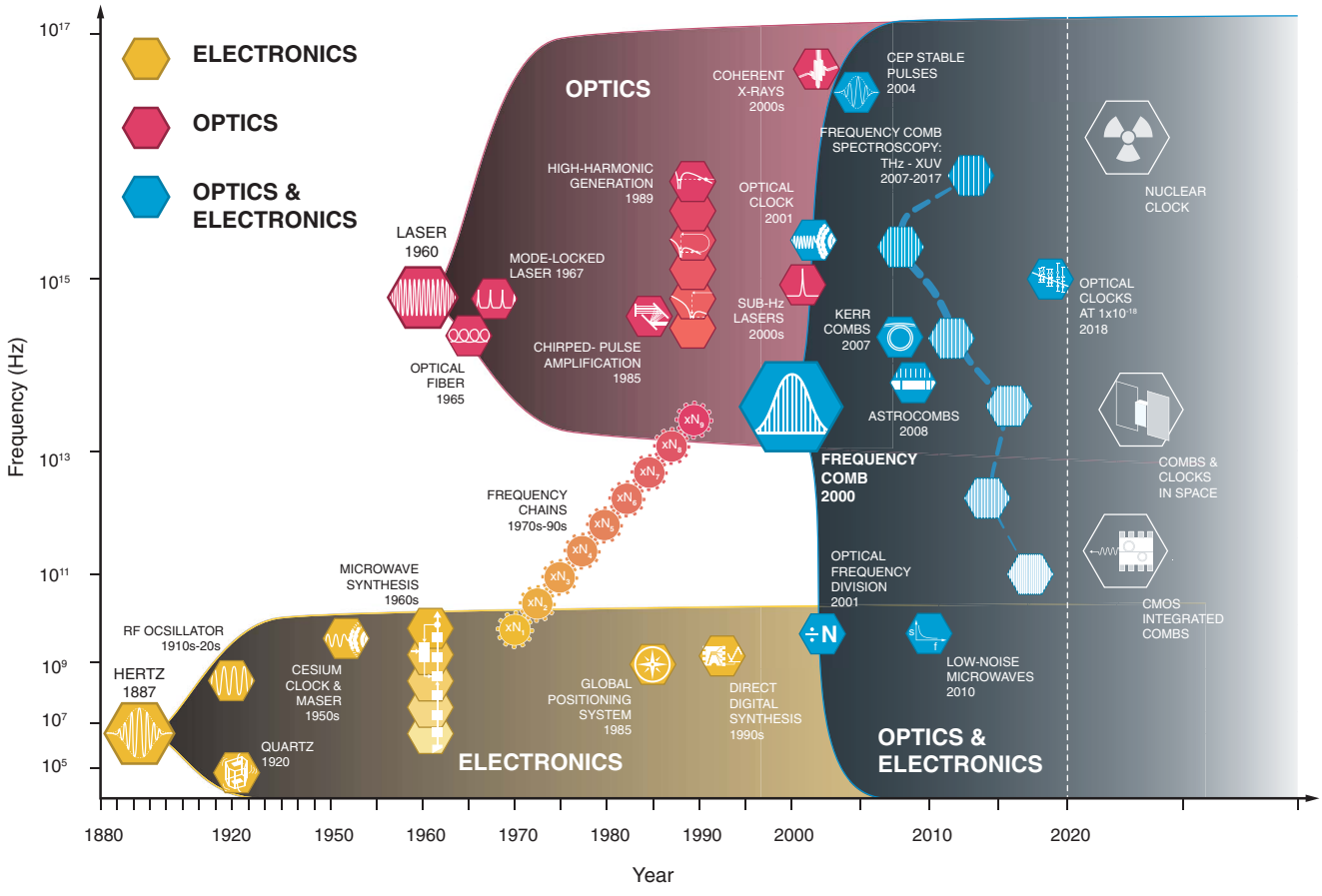


Fig. 1. The coherent electromagnetic spectrum and its development. In about 1887, Heinrich Hertz generated and detected artificial radio waves, which triggered a technological revolution that eventually gave birth to a plethora of systems for communications, navigation, and ranging (radar). Radio frequency (RF) masers and atomic clocks introduced in the 1950s were a technological game changer. The cesium atomic clock defines not only the second, but undergirds the International System of Units (SI) (*20*). With the introduction of the laser in 1960, the first source of coherent optical electromagnetic waves became available. This big leap in technology, as well as in frequency, introduced a gap of about 10^5 relative to the highly controllable radio frequency domain. Harmonic frequency chains bridged this gap, but the resulting systems were impractical. The frequency comb, which was introduced in its fully functioning form in 2000, coherently united the radio and optical realms in a simple and practical way. It has opened many new possibilities such as optical atomic clocks, ultralow-noise microwaves, carrier-envelope phase control of ultrashort pulses, and the coherent synthesis of radiation reaching from terahertz to extreme ultraviolet.

frequencies (*16*, *17*) but thus far have not been realized.

Among the early applications of these efforts was an experiment to refine the value for the speed of light (*18*). This was possible using $c = f\lambda$ where f, the optical frequency, was measured by using a harmonic frequency chain and λ, the wavelength, was determined by precision interferometry through comparison with the wavelength of a specific krypton transition (1960 definition of the meter). Ultimately, however, an improved realization of length was possible by defining the value of the speed of light to be exactly 299,792,458 m/s in the 1983 revision of the International System of Units (SI) (*19*, *20*). Because frequency is, and always has been, the most accurately measured quantity in all of physics, this definition implies that, in theory, frequency-like accuracy could be obtained for length. In practice, a laser stabilized to a frequency-calibrated optical transition (*21*, *22*) serves conclusively as a realization of the meter. Harmonic frequency chains were also used to determine transition frequencies in atomic hydrogen. These measurements tested quantum electrodynamics and determined the values of physical constants (*23*, *24*). Not much more could be done with the harmonic frequency chains because they were large and very complicated systems containing many microwave and laser subsystems that needed to operate in a highly coordinated fashion. Their sheer complexity made it clear that these systems would never be widely available and that their capability would be limited.

Nonetheless, harmonic frequency chains inspired new approaches aimed at simplification. A key initial step in this direction was the optical interval divider, as proposed by Hänsch in 1989 (*25*) (Fig. 2, C to E). This device works in a way analogous to a differential gear except for optical frequencies, precisely bisecting an arbitrarily large frequency interval (see Fig. 2D). The optical interval divider could be concatenated to successively bisect the interval between an optical frequency f and its own second harmonic $2f$ that can easily be obtained with nonlinear crystals. As described below, this idea was used later in a slightly different fashion in the operation of the modern optical frequency comb. An interval divider chain would have used roughly the same number of stabilized oscillators as a frequency chain. However, it offered an important advantage in that all the oscillators could have been of the same type, yielding an enormous reduction in complexity. At least two groups, at the Max Planck Institute of Quantum Optics, Germany (*26*), and the National Institute of Advanced Industrial Science and Technology, Japan (*27*), had started to construct such an interval divider chain, but these efforts were never completed.

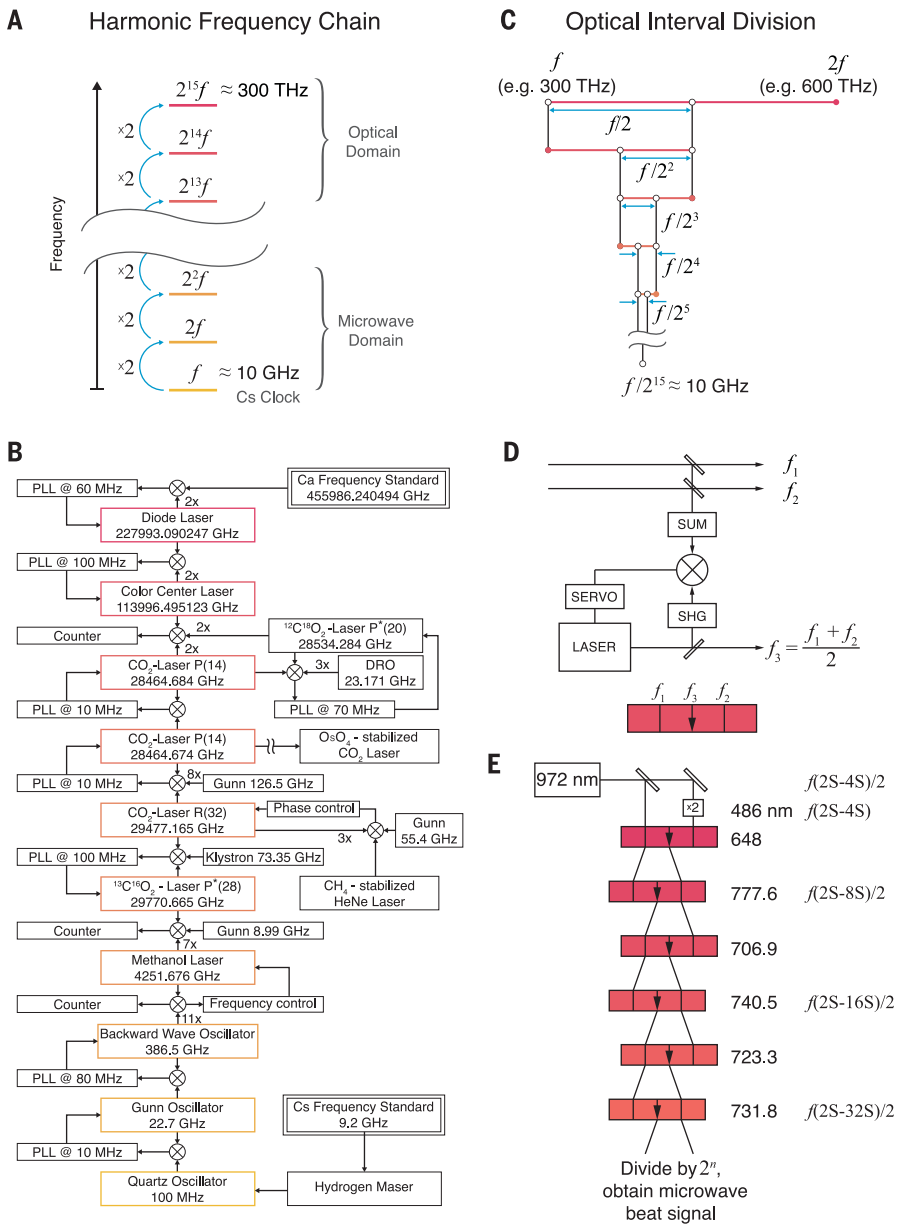


Fig. 2. Historical approaches to connect optical and microwave domains. (**A**) Successive nonlinear frequency multiplication with a harmonic frequency chain connecting radio and optical frequencies. For illustration, a factor of 2 in frequency separates each step. (**B**) Harmonic frequency chain implementation that existed at the Physikalisch Technische Bundesanstalt (PTB) at Braunschweig, Germany, and filled two large laboratories located in two different buildings [adapted from (*15*)]. Nonlinear devices generate harmonics in each step to which an intermediate oscillator is phase locked (PLL, phase-locked loop) to provide enough power for the next multiplication step. In this way one walks through the spectrum, from the microwave frequency of a cesium atomic clock all the way to the visible frequency of a calcium standard at 456 THz. The multiplicative factor between lower-frequency stages could be substantially larger than 2×. (**C**) As an alternative, one could converge to a radio frequency by successively bisecting the interval spanned by an optical frequency f and its own second harmonic $2f$. (**D**) An optical interval divider phase locks the second harmonic (SHG) of a laser at f_3 to the sum frequency (SUM) of two input lasers at f_1 and f_2, bisecting the frequency difference between the latter: $f_3 = (f_1 + f_2)/2$. (**E**) A proposed (*25*) but never realized implementation of an interval divider chain that would have been designed for application to atomic hydrogen. In agreement with the Rydberg formula, the transition frequencies are at small integer ratios, several of which (at the right hand side) agree with the frequencies in this interval divider chain. [(D) and (E) are adapted from (*25*).]

An optical clockwork

Around the turn of the millennium, the rapid advance in mode-locked lasers as optical frequency comb generators eliminated the need for harmonic frequency chains and optical frequency interval dividers. In a spectacular simplification, the frequency comb collapsed a dozen stages into a single direct link between an input radio frequency signal and a vast array of optically synthesized frequencies that are spaced on a perfectly even frequency grid extending over an octave of spectrum. Moreover, the comb phase coherently converts a radio frequency into an optical frequency with frequency multiplication by $\approx 10^5$, or in reverse, it divides by the same factor. In fact, in such a radio-to-optical conversion, both division and multiplication happen simultaneously. With the long-standing goal of frequency conversion readily at hand, one can now select the domain, radio or optical, in which the best oscillators exist and convert it to the other domain. And whereas the harmonic chains easily filled a few large laboratories and operated only occasionally, current frequency combs now fit in a shoe box and run continuously, thus finally realizing a straightforward means to unite the electromagnetic spectrum.

In a simple picture, the frequency comb can be associated with the longitudinal modes of a laser cavity that are separated by a frequency f_r. Comb formation relies on a nonlinear optical process, called mode-locking, to enforce equal frequency spacing and establish phase synchronization between the modes such that a short optical pulse is formed (Fig. 3A). This pulse propagates at its group velocity back and forth between the end mirrors of the laser resonator with a round trip time of f_r^{-1}. At each round trip, a copy of the pulse is transmitted through one of these mirrors that is partially transparent. The well-known Fourier transformation between time and frequency domain implies that the spectrum of the periodic pulse train generated in this way is in fact a discrete comb of frequencies. Moreover, the spectral width of the comb would approach the optical carrier frequency as the pulse duration is reduced to a single optical cycle.

Further technical details about the operation and generation of the frequency comb are provided in subsequent sections, but the range of ways that a frequency comb coherently connects the EM spectrum, whether from microwave to optical, optical to microwave, or optical to optical, is captured intuitively through the optical clockwork analogy illustrated in Fig. 3B. At its center is the frequency comb represented as a large gray gear that is enmeshed with a microwave source symbolized by the black gear. The frequency comb connects the microwave gear in a phase-coherent way, i.e., without slippage, to the smaller fast-spinning gears, each of them associated with a single optical

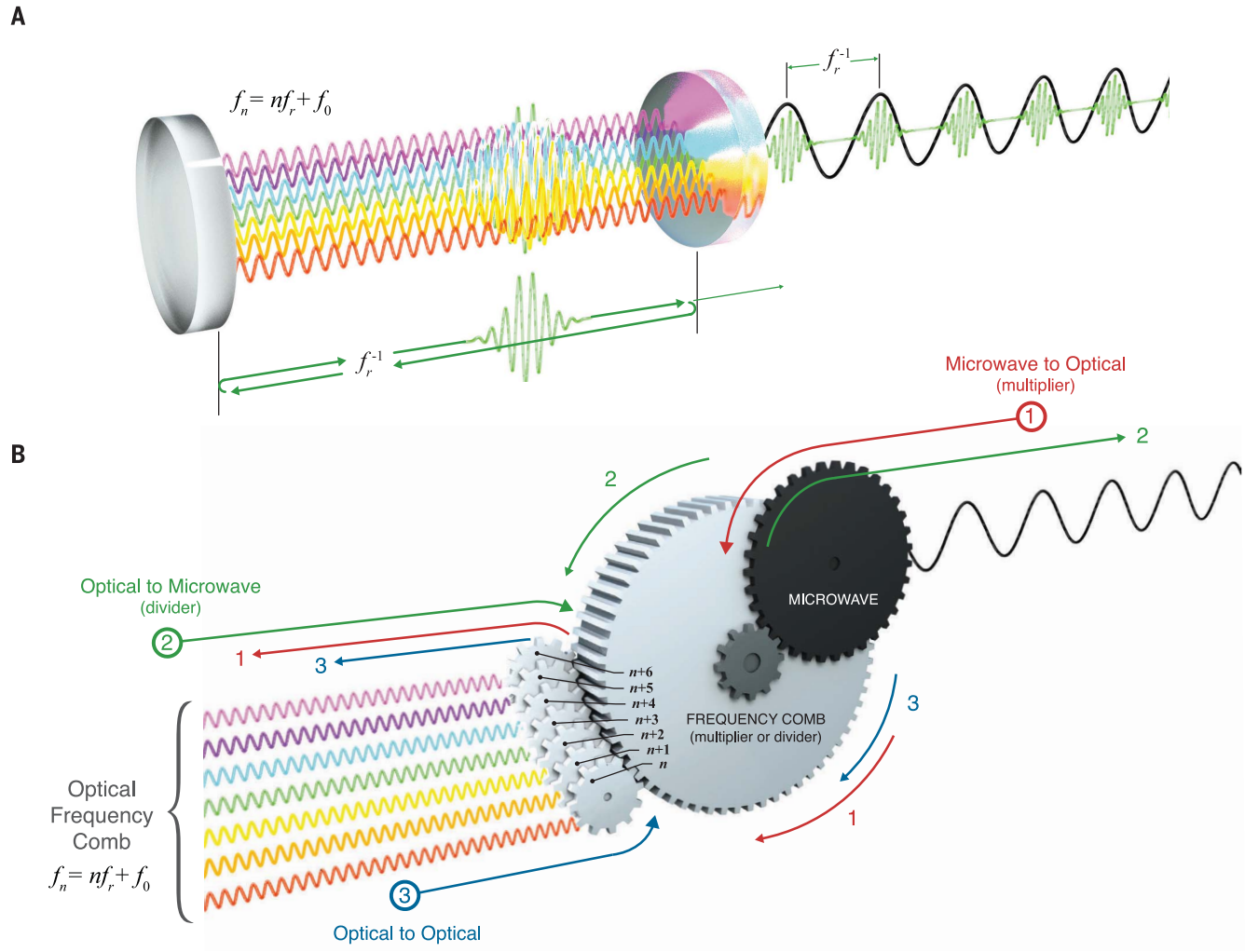


Fig. 3. The frequency comb as an optical clockwork. (**A**) An optical frequency comb, here illustrated as waves of different colors that are spatially separated for clarity, is composed of the longitudinal modes of an optical cavity made by two mirrors. The comb frequencies are given by $f_n = nf_r + f_0$. In reality the waves are spatially overlapped and phase synchronized such that their superposition produces a pulse circulating inside the cavity. The output of the cavity consists of a periodic train of replicas of the intracavity pulse separated by the round-trip time f_r^{-1}. (**B**) Mechanical analogy of the optical frequency comb. The frequency comb has three modalities in which it can operate to phase coherently connect electronic (microwave) and the optical realms of the EM spectrum: 1. An input microwave frequency is multiplied by a factor of ~10^5 to the optical domain, where it is output as an array of equally spaced light waves. 2. In the inverse of the first modality, an input high-frequency optical wave can be divided down to the microwave domain. 3. Spectral translation between the array of optical waves occurs when an input optical field controls the comb and is phase-coherently connected to the other frequencies in the optical domain. As shown, all modalities of operation are bidirectional and simultaneous.

frequency of the comb. Even though this clockwork shows only seven fast-spinning gears, an actual frequency comb would feature ≈10^5 of them covering a full octave. Also, the actual frequency ratios spanned by a comb are much larger than can be easily displayed in Fig. 3B. However, the clockwork illustrates the three major ways that the frequency comb enables coherence to be transferred or shared across the radio and optical spectra. Each of these operational modes has led to a broad application category of comb technology, as described below.

Optical frequency synthesis

The precise synthesis of new frequencies was required for the harmonic frequency chain (Fig. 2B) and more broadly for the advancement of radio communications. In a radio frequency synthesizer, new frequencies are generated through various processes, including electronic multiplication and division, phase-locked loops, and direct digital synthesis. Optical frequency combs operating in mode 1 (red arrows in Fig. 3B) can be viewed as a pronounced extension of the concept of frequency synthesis from the radio to optical realms. Here, the frequency comb transfers the radio frequency coherence (black sinusoid) to a very broad spectrum of equally spaced optical frequencies f_n represented by the small gears. These frequencies are given by the expression (*28*–*30*)

$$f_n = nf_r + f_0 \quad (1)$$

As noted, f_r is the pulse repetition rate, or equivalently, the frequency separation of the comb modes, whereas f_0 is an offset frequency that is common to all the modes and whose origin and measurement will be discussed in greater detail below. For many frequency combs, both f_r and f_0 are typically in the range of 0.1 to 10 GHz, such that $n = 10^4 \ldots 10^6$ is the required multiplier for f_n being in the near-infrared frequency band. As a result, if f_r and f_0 are stabilized to a known radio frequency reference, such as that provided by a cesium clock, then one obtains 100,000 stabilized lasers superimposed in one beam.

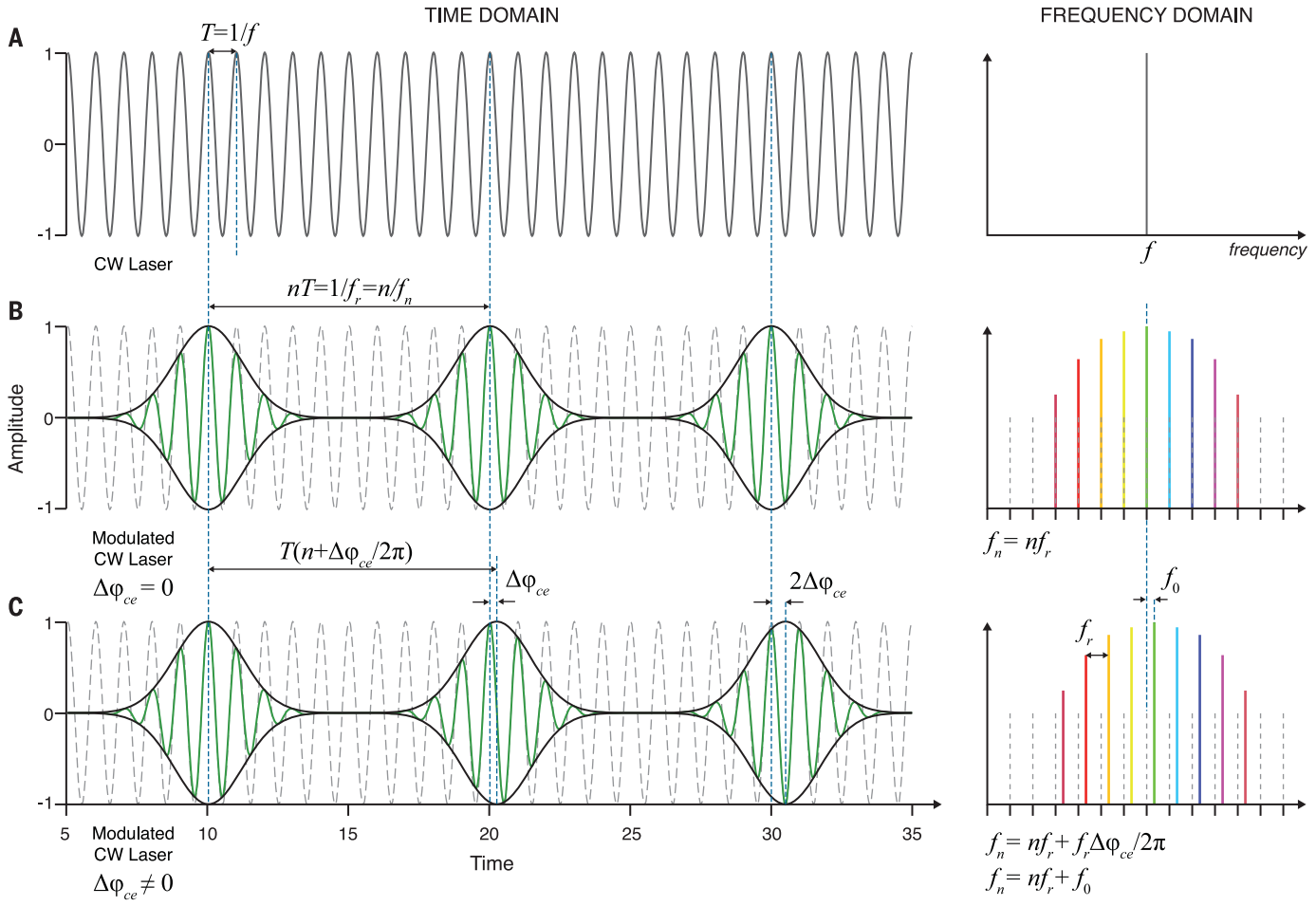


Fig. 4. Time- and frequency-domain representations of the optical frequency comb. (**A**) The continuous optical wave with frequency f and period T is represented as a single tone in the frequency domain. (**B**) An amplitude modulation on the optical wave at a lower frequency f_r that is an exact subharmonic of f would enable simple frequency division by an integer n. Here the modulation is shown at $f/10$. In reality, if $f = 500$ THz, the modulation at $f/10^5$ would yield a countable $f_r = 5$ GHz. (**C**) In general, the modulation frequency is not an exact subharmonic of f, but rather the modulation envelope "slips" with respect to the fixed carrier f from pulse to pulse. Here, for example, the slip is 1/4 of a cycle, so the modulation $f_r = f/(10 + 1/4)$. It is this noninteger portion of the denominator that causes the pulse-to-pulse carrier-envelope phase shift. In the frequency domain, the pulse-to-pulse phase shift results in a comb shifted by $f_0 = f_r\Delta\varphi_{ce}/(2\pi)$. By measuring f_0, one can determine the noninteger part of the divisor exactly.

The first demonstrations of synthesis with laser frequency combs were published around the turn of the 21st century (*31*–*34*). These works showed the power of measuring and controlling f_r and f_0, as well as the ability to generate a comb of optical frequencies traceable to the microwave SI second. Moreover, they opened the gates to a flood of optical frequency measurements that served to validate the painstaking efforts from previous decades with the frequency chain technology; see, for example, (*33*, *35*, *36*). Although this was important for the metrology community and the development of optical frequency and length standards, new opportunities of a more fundamental nature also emerged. For example, a comparison of microwave (cesium) and optical frequency standards by means of the laser frequency comb can be interpreted as a test of the constancy of fundamental constants, such as the electron-to-proton mass ratio and the fine-structure constant, either in space or time (*37*–*39*).

The ease with which any optical frequency could be measured has been extended to spectroscopy across the EM spectrum (*40*–*42*). Many such spectroscopy experiments use continuous-wave lasers that are simultaneously calibrated by a heterodyne measurement with the stabilized comb. But powerful approaches have also arisen in which the comb itself serves as an array of hundreds of thousands of individual lasers that simultaneously act as a spectroscopic probe (*43*–*46*). Coupling the frequency comb to another optical cavity is sometimes used to enhance the power of the relatively weak comb modes (often <100 nW in a single comb mode), thus improving the spectroscopic sensitivity (*47*, *48*). To read out the individual comb modes, clever detection schemes have been realized, including the use of a highly dispersive spectrometer (*49*) or a swept Vernier cavity filter (*50*). Alternatively, the addition of a second frequency comb, offset slightly in repetition rate, forms a dual-comb spectrometer in which the amplitude and phase of individual optical modes are uniquely mapped to a radio frequency beat recorded on a single photodetector (*51*–*57*). Frequency comb spectroscopy (*58*, *59*) is a rapidly growing subfield that is being applied to a wide range of basic and applied research topics. A noteworthy variation is the use of frequency combs to calibrate high-precision astronomical spectrographs aimed at identifying habitable exoplanets through the measurement of minute Doppler shifts in the stellar spectra (*60*–*62*).

Clocks and optical frequency division

A clock consists of an oscillator and a counting mechanism. This basic concept has been implemented throughout history with technology operating at ever-higher frequencies: beginning

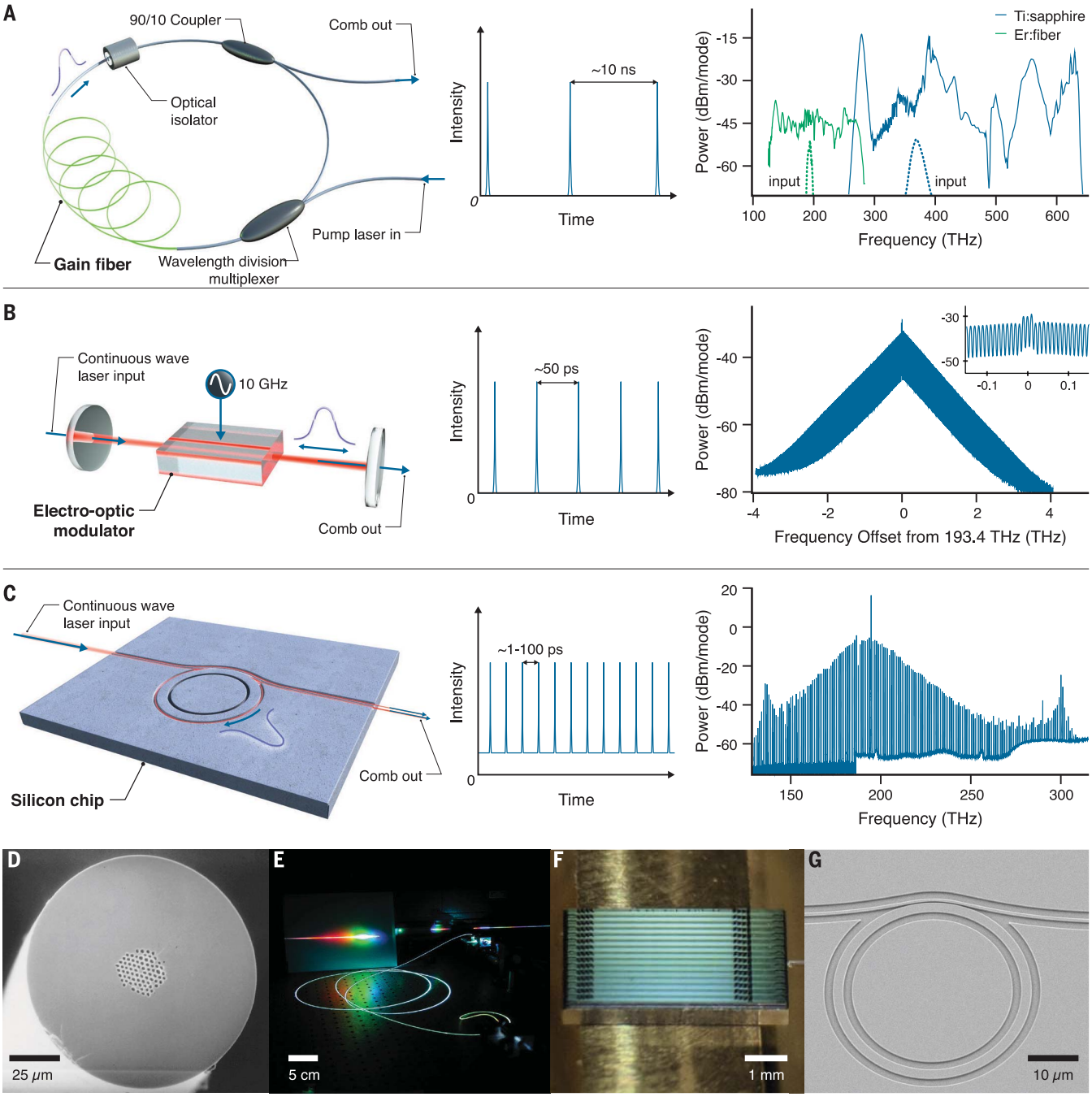


Fig. 5. Methods for optical frequency comb generation. (**A**) Mode-locked lasers. A pulse is stored within a cavity formed by an optical fiber loop. During each round trip, a fraction of the circulating pulse is coupled out, forming a pulse train with a typical period of 10 ns. Cavity losses are compensated by a gain medium, and an optical isolator ensures unidirectional operation. At right are typical spectra from such an Er:fiber laser, as well as a Ti:sapphire laser, both before and after spectral broadening to octave span in a nonlinear fiber. Comb modes are not resolved in these spectra. (**B**) Electro-optic comb generator. Microwave modulation creates an optical comb by adding sidebands to an injected carrier wave from a continuous-wave laser. The modulator is placed inside an optical cavity with free spectral range matched to the modulation frequency. The output pulse period is twice the modulation period, or 50 ps in this typical example. The spectrum shown at right has a 10-GHz mode spacing. (**C**) Microresonator frequency comb. A frequency comb is generated by the nonlinear Kerr effect in a passive microfabricated waveguide cavity. The balance of dispersion and nonlinearity yields a stable soliton that circulates in the ring cavity with a typical period in the range of 1 to 100 ps. At right is the spectrum of the associated frequency comb with a spacing of 1 THz (*172*). (**D** to **G**) The images in the lower panels show key nonlinear optical technologies used to generate octave-spanning frequency combs. (D) Nonlinear microstructure fiber (*97*), and (E) the visible-light spectrum such fiber can produce (*101*). (F) A millimeter-sized silicon chip with 15 nonlinear silicon-nitride waveguides (*173*). (G) A silicon-nitride ring resonator on a silicon chip for generating the comb with 1-THz mode spacing (*142*, *172*).

with the motion of Earth itself and progressing through pendulum clocks, quartz oscillators, and the microwave cesium clock. Along this historical path, a clear guiding principle in timekeeping has been to use the highest frequency possible for the clock oscillator, as this provides the opportunity to most precisely divide time into the smallest intervals. With amazing foresight, an optical time standard was already anticipated by Maxwell (*5*), even though he lived 100 years before the invention of the laser. And though a laser referenced to an atomic transition provides the pendulum of such an optical clock, until the frequency comb, there was no practical way to count the ultrafast cycles of optical light.

In the mechanical analogy of Fig. 3B, building an optical clock involves using the clockwork as indicated by the green arrows. In the actual laser frequency comb, the frequency of one of the comb modes f_n is controlled relative to the optical frequency reference, f_{opt} (e.g., a laser steered to an optical transition), such that Eq. 1 can be rewritten as

$$f_r = (f_{opt} - f_0)/n \qquad (2)$$

Here, the readout of the optical clock is the repetition rate of the pulse train of the frequency comb laser, which is essentially a very large subharmonic of f_{opt} because $f_0 \ll f_{opt}$. Within a year of the introduction of the self-referenced frequency comb, such optical clocks had been implemented (*63*, *64*), and their uncertainty was soon competitive with that of the best Cs microwave clocks. In addition to the frequency comb, there were two important scientific and technological advances at the end of the 1990s that accelerated the development of optical clocks and have elevated them to the peak of metrologic precision. The first was the maturation of laser cooling and trapping techniques proposed decades prior (*65*, *66*) that now hold atoms essentially motionless to eliminate Doppler effects and allow the realization of atomic-transition quality factors exceeding 10^{15}; see, for example, (*67*). The second was the development of a new class of hertz-linewidth lasers based on vibration-insensitive mounting of the reference resonator (*68*, *69*) that enabled operation at the fundamental thermodynamic limit (*68*, *70*, *71*).

The continued advancement of these technologies over the past 20 years has led to stabilized lasers with a linewidth of just a few tens of millihertz (*71*) and optical frequency standards and clocks with fractional frequency uncertainty pushing into the 10^{-19} range (*72*–*76*). Beyond counting the cycles of such ultraprecise optical frequency references, optical frequency division of a cavity-stabilized laser with the comb now provides microwave signals with short-term stability and phase noise that surpasses that of the lowest-noise microwave oscillators (*77*, *78*). Indeed, optically generated microwaves with a frequency comb leverage the ultrahigh optical quality factors and narrow-linewidth lasers mentioned above, the equivalent of which are unavailable directly in the microwave realm. Finally, combinations of optical clocks and free-running ultrastable optical oscillators together with the frequency comb clockwork have recently been applied to the generation of optical time scales (*79*–*82*), which are beginning to play a role in coordinated universal time (UTC) and will someday realize a new definition of the SI second (*83*, *84*).

Optical-to-optical coherence transfer

The superior performance of optical clocks and frequency standards has created a situation in which the SI definition of the second is now realized with a 100-times larger fractional uncertainty (1×10^{-16}) than the agreement found among the best optical clocks (1×10^{-18}) (*73*). This will ultimately be rectified with a future redefinition of the second in terms of an optical transition, but the present operational advantage for frequency synthesis is regularly realized by using an optical frequency to reference the comb. Not only does this provide optically stable microwaves, as just discussed, but it provides a means to synthesize other optical frequencies across hundreds of terahertz, or even petahertz, of bandwidth. This enormous range of synthesis bandwidth is displayed in Fig. 1, where it should be noted that the optical domain encompasses a bandwidth greater than the gap between radio frequency and optical realms. Previously, smaller gaps of a few terahertz in the optical realm could be bridged (*85*), but low-noise frequency comb sources have now connected the entire optical spectrum, with phase coherence being transferred between the multitude of comb modes (Fig. 3B, blue arrows).

One of the uses of optical coherence transfer is to compare the ratio of widely separated optical clock frequencies. Indeed, the most accurate and fastest means to evaluate uncertainty in the best optical clocks is to compare them, one against the other, by using the frequency comb. This process essentially amounts to measuring the ratio of optical frequencies (*86*). It was first demonstrated with the comparison of Hg^+ and Ca optical clocks (*63*) but has since been extended to practically all high-performance clocks. As with the microwave-to-optical mode of operation, the ratio of optical frequencies can also been interpreted to constrain physical models in which the fine-structure constant α is not in fact constant (*87*, *88*). In addition, the same ratios of optical frequencies are being used in searches for dark matter and other new physics (*89*, *90*). From a more practical standpoint, the coherent control of the frequency comb spectrum has enabled optical time transfer between remote clocks with femtosecond precision (*91*), and absolute ranging with nanometer-level precision over kilometer distances (*92*).

The carrier-envelope offset frequency

The mechanical analogy of the frequency comb presented in Fig. 3B is useful to understand its basic function and modes of operation, but it does not capture the physical origin and understanding of f_0, which is more generally called the carrier-envelope offset frequency. Indeed, if not for f_0, the synthesis from microwave to optical realms, and vice versa, would have been completed as depicted in Fig. 3 with the first mode-locked lasers in the 1960s. And though the origins of f_0 were understood as early as the mid-1970s (*28*), it took until the 1990s to realize how to measure and control it in the context of an optical frequency comb.

To understand the implications of f_0 for a frequency comb, it is helpful to first describe optical frequency division without it. Imagine the continuous light wave at a frequency f sketched in Fig. 4A to be amplitude modulated with a frequency f_r, as shown in Fig. 4B. In particular, the amplitude function is thought to produce a pulse of light exactly every n cycles of the continuous light wave. In this case, the lower (electronically countable) frequency f_r is an exact subharmonic such that $f_r = f/n$, and the associated optical frequency comb, would be $f_n = nf_r$. This time-frequency description of a modulated light wave was analogously depicted as the frequency-dividing clockwork in Fig. 3B.

However, the situation just described is not typically realized in frequency comb generators where dispersion in the optical cavity producing the comb gives rise to differing group and phase velocities. This causes the pulse envelope to be delayed relative to the carrier from pulse to pulse by the carrier-envelope phase slip $\Delta\varphi_{ce}$. This phenomenon has an intuitive geometrical interpretation for well-behaved pulse shapes, as shown in Fig. 4C, and was experimentally verified in 1996 (*93*). The consequence of the pulse-to-pulse carrier-envelope phase slip in the frequency domain is to offset the entire frequency comb by an amount $f_0 = f_r\Delta\varphi_{ce}/(2\pi)$ (*28*). Here we recognize that $\Delta\varphi_{ce}/(2\pi)$ is the fraction of the optical cycle by which the envelope slips from pulse to pulse, making it possible to realize division of the optical frequency to the modulated envelope as $f_r = f/(n + \Delta\varphi_{ce}/(2\pi))$, which is equivalent to Eq. 2.

Because f_r is the round-trip rate of pulses in the laser cavity, its measurement is straightforward by photo detection of the emitted laser pulses. A practical way to measure f_0 is,

however, more challenging. Eventually this so called "self-referencing" is achieved with a nonlinear interferometer that heterodynes the second harmonic of a set of modes with indices near n_L from the low-frequency side of the comb spectrum with another set of modes near index $n_H = 2n_L$ to generate the beat frequency $2f_{n_L} - f_{n_H} = f_0$. This "$f - 2f$" method is seen to be a variation of the optical interval divider (Fig. 2, C to E) and could be implemented with a combination of a frequency comb and interval dividers (*27*). However, the most simple and efficient implementation (*26*, *94*) requires an octave-spanning frequency comb that was not readily available in the 1990s. The solution to this issue was first presented by a new type of nonlinear optical fiber, commonly called microstructure or photonic crystal fiber (*95*, *96*), that could be used to generate octave-spanning frequency combs using nanojoule-level pulses from the existing narrowband mode-locked lasers (*97*). This provided the key missing technology for the measurement of f_0 and the closing of the clockwork that could then readily connect optical and microwave realms (*32*–*34*).

Frequency comb technologies

The history of frequency comb technologies involves a rich combination of laser physics, ultrashort optical pulse generation, nonlinear optics, and electro-optics. In most cases, the core technologies that are now central to frequency combs were developed for independent and unrelated reasons, which highlights the importance of cross-disciplinary research in the growth of the frequency comb field. Some of the most common techniques for generating optical frequency combs and extending their spectral bandwidth are shown in Fig. 5.

Mode-locked laser frequency combs

The mode-locked laser was introduced just a few years after the 1960 invention of the laser itself (*98*). By the early 1980s, progress in the field had led to techniques for the generation of femtosecond time-scale pulses (*99*) that were primarily used for measuring ultrafast dynamics in material systems. A key element of a mode-locked laser is a means to favor the generation of short pulses of light inside the laser cavity. Initially, active amplitude modulation was used, but in most mode-locked lasers, pulses are now generated passively with a saturable absorber or another nonlinear element that provides a net increase in gain for pulsed operation. Once generated, pulses are stored on a repetitive path, either within a cavity formed by mirrors or an optical fiber loop, as shown in Fig. 5A. In steady state, the circulating pulse is approximated by a soliton that maintains its shape indefinitely through a balance of nonlinearity and dispersion. This produces a train of pulses with identical envelope at the laser output and results in the perfectly regular comb structure.

The first self-referenced frequency combs used Kerr-lens mode-locked Ti:sapphire lasers that had been introduced in 1990 (*100*) and matured over the subsequent decade. Example spectra are shown in Fig. 5A. Typical Ti:sapphire lasers have peak emission around 800 nm and emit pulses with duration <100 fs. This corresponds to a Fourier-limited spectral bandwidth >5 THz, which falls considerably short of the octave spectrum required for self-referencing and detection of f_0. Broadening to octave span uses the previously mentioned microstructure fiber. This fiber has air channels running along its length that increase optical confinement and hence optical nonlinearity. Moreover, the channels allow for new freedom in dispersion design to substantially enhance the spectral broadening of the low-energy (nanojoule) pulses (*95*, *96*, *101*) (Fig. 5D). Such fibers immediately enabled the generation of octave-spanning frequency combs from the existing Ti:sapphire lasers (see spectra of Fig. 5A), as required for self-referencing. With improved engineering, Ti:sapphire lasers were also developed that could directly emit few-cycle pulses and octave-spanning spectra (*102*–*104*).

Although Ti:sapphire has proven to be a reliable approach in some frequency comb systems (*104*), the most widespread and mature frequency comb technology is now realized with mode-locked lasers using erbium-doped fiber as a gain medium (*105*). These leverage the robustness and fiber-integrated components provided by the telecommunications industry in the 1550-nm region of the spectrum (*106*–*111*). Er:fiber frequency combs employ solid-core nonlinear fibers to generate octave-span spectra; these comb systems, including gain fiber, pump lasers, amplifiers, and nonlinear fibers, can be entirely spliced together, making them extremely robust (*112*). Such frequency combs are commercially available from several companies, run autonomously in spectroscopy and metrology applications, and have even been operated on board a sounding rocket (*113*). Lasers using Yb- and Tm-doped fibers, as well as other solid state materials, have led to an expanding palette of frequency comb spectra (*114*, *115*). Additionally, nonlinear optical processes have pushed the extent of frequency combs across the terahertz (THz) (*116*, *117*) and mid-infrared (MIR) (*118*–*122*) domains and all the way into the extreme ultraviolet (XUV) (*123*–*126*) for a wide range of sensing, clock, and spectroscopic applications.

Electro-optic frequency combs

An early approach to generate a frequency comb employed an electro-optical (EO) modulator driven by a microwave signal to impose a grid of sidebands on a continuous-wave laser (*127*). Building on this core EO technological piece, in 1993 Motonobu Kourogi and co-workers introduced a very efficient modulator-based comb generator (*85*). The idea was to place the modulator inside an optical resonator and drive it with a microwave signal that matched the optical resonator mode spacing. As a result, the generated sidebands are resonantly enhanced and generate sidebands themselves (Fig. 5B). Although the input seed laser frequency might not be absolutely stable, these modulator-based comb generators could be used like a ruler to measure optical intervals of a few THz, limited ultimately by dispersion inside the resonant optical cavity (*128*). EO combs became an attractive part of the Hänsch interval division technique (Fig. 2, C to E) once the gap between successive subharmonics was reduced down a value on the order of the EO comb bandwidth (*27*). Efforts to increase the bandwidth of EO combs included spectral broadening in an optical fiber outside the EO cavity (*129*), or the introduction of parametric gain inside the EO cavity (*130*), and recent advances in integrated lithium-niobate waveguides promise still broader bandwidths (*131*) along the lines of Kourogi's original modulator-in-a-cavity design.

Although mode-locked laser combs and their octave-spanning frequency spectra supplanted EO combs, the simplicity and deterministic nature of the EO comb remained attractive. And thanks to improved lithium niobate waveguide modulators coming from the telecommunications industry, a different approach to EO combs reemerged around 2010 (*132*–*134*). Instead of placing the EO comb inside an optical cavity, modulators can be cascaded one after another, and with appropriately phased microwave drive, each modulator adds additional comb lines. This results in a broad bandwidth and spectrally flat frequency comb that can be compressed into a short pulse for additional nonlinear spectral broadening. However, owing to the multiplicative nature of this comb generator, the microwave phase noise can ultimately grow to overwhelm the comb teeth themselves. Recently, this problem was resolved with the introduction of an optical cavity filter, which resulted in an octave-span EO comb for which f_0 could be measured (*135*). Efficient spectral broadening using nanophotonic silicon nitride waveguides, such as shown in Fig. 5F, combined with techniques to further reduce the noise of the microwave source have now resulted in an EO comb with noise that is comparable to that of mode-locked laser combs (*136*).

Microresonator frequency combs

A third approach to the formation of frequency combs has attracted interest because it offers the prospect of miniature comb systems integrated on a semiconductor chip (*137*). Often called microcombs, these devices rely on access

to the Kerr nonlinearity at milliwatt power levels through resonant enhancement of light in whispering-gallery microresonators (*138*–*140*) or in ring-like resonators as shown in Fig. 5C (*141*, *142*). Key to their operation is the confinement of the input single-frequency laser light in a resonator mode with a cross-section of ≈1 μm^2 that has very low propagation loss, with material and dielectric-interface–dependent quality factors ranging from $Q \approx 10^6$ up to nearly 10^{11} being achieved. Under these conditions, milliwatts (or less) of input power are resonantly enhanced to as high as kilowatt levels, so that two sidebands are spontaneously generated with equal spacing about the pump frequency through parametric oscillation mediated by the Kerr nonlinearity (*143*, *144*). This process can cascade through four-wave mixing such that additional sidebands are generated (*143*), and in 2007 it was shown to lead to a frequency comb (*145*). Developments related to solitons in optical fiber cavities (*146*, *147*) combined with a deeper understanding of cascaded combs eventually led to operational modes where isolated solitons could be generated in microresonators (*148*–*152*). As with mode-locked lasers described above, a soliton pulse that propagates with unchanging shape around the cavity is the basis of a frequency comb.

Critical to the generation of solitons in microresonators is the control of the dispersion that the pulse experiences while traveling around the cavity. Here the lithographic control offered by modern fabrication techniques permits adjustment of geometric dispersion through nanometer-scale dimensional changes to the waveguide cross section (*153*, *154*). Most notably, this has led to microresonator designs that directly generate solitons with octave-spanning spectra (*142*, *155*). An example of such a resonator made from silicon nitride is shown in Fig. 5G, and its resulting comb spectrum with 1-THz mode spacing is shown in Fig. 5C. Other microresonator materials, such as aluminum nitride (*156*), lithium niobate (*157*), and III-V semiconductors (*158*), offer both third- and second-order nonlinearities that could potentially be used to both generate octave-span combs and then double the frequency of the long-wavelength parts for self-referencing.

There has been tremendous progress with microcomb system demonstrations, including optical frequency synthesis (*159*) and optical clocks (*160*, *161*). But there is still much work to be done before the full integration potential of microcombs is realized, as most still require off-chip lasers, amplifiers, or other optical fiber components. Nonetheless, repetition rates in the 10- to 100-GHz range make microcombs attractive for applications, including astronomical spectrograph calibration (*162*, *163*) and optical communications (*164*). Moreover, the high repetition rates can be advantageous in dual-comb measurements where high-speed acquisition is desirable, as shown in recent spectroscopy (*165*, *166*) and light detection and ranging (LIDAR) (*167*, *168*) demonstrations.

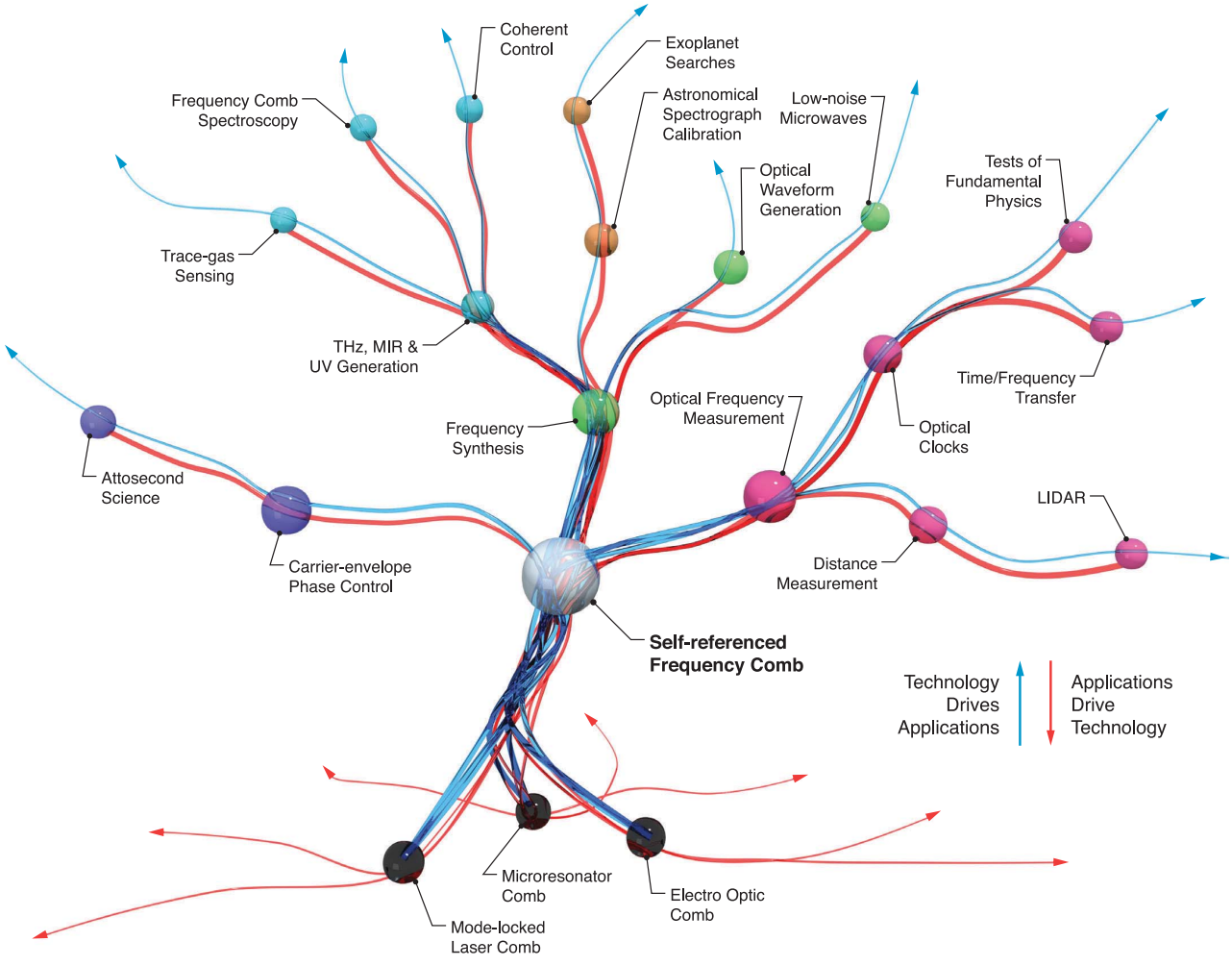


Fig. 6. The frequency comb evolutionary tree reenvisioned. Technology drives new applications, and applications feed the growth of new technologies.

The continuously evolving optical frequency comb

Optical frequency combs were originally developed to solve the problem of measuring and counting optical frequencies in terms of the SI cesium definition. As we have described, this was of critical importance for applications in precision spectroscopy and time and frequency metrology. And though spectacularly successful in that role, 20 years ago it was impossible to foresee the range of applications beyond metrology, spectroscopy, and clocks for which frequency combs are now being used. Indeed, one of the most surprising developments of optical frequency combs has been the diversity of applications that they have found and opened, as well as the intertwined technologies they have employed (Fig. 6). The symbiotic relationship between applications and technologies can be depicted as a kind of living and evolving tree (*114*), where new applications sprout upward as technological roots spread downward. The branches and roots find at their center the key aspect of self-referencing, which is the distinguishing feature of a frequency comb that coherently connects the optical and electrical realms.

The scientific depth and range of the technological roots that produce frequency combs have also evolved substantially, with several surprising innovations that rapidly advanced frequency comb capabilities. Far before the first self-referenced Ti:sapphire mode-locked lasers, nascent frequency comb spectroscopy was performed with a synchronously pumped dye laser that contained just six to eight frequency modes and a bandwidth that could only enable difference frequency measurements across a 1-GHz gap (*169*). This was expanded to the THz range with the introduction of EO comb technology in the 1990s. However, that did not forecast the serendipitous meeting of the Ti:sapphire laser with nonlinear fibers that led to the immediate expansion of frequency combs to hundreds of terahertz and the octave span required for self-referencing (*32*–*34*). Twenty years later, this self-referencing is now a routine part of the clockwork of the Er:fiber frequency combs that run autonomously as the most robust and widely used comb technology (*111*).

Nanoscale waveguide technology and nonlinear optics, that first appeared in the microstructure fiber, is another technology with significant evolutionary impact on frequency comb generation. Microresonator frequency combs (*137*, *145*, *152*) not only provide a comb-generating mechanism different from that of mode-locked lasers, but they can be built at wafer scale in a manner analogous to the fabrication of integrated electronic circuits. This is a first step along the path toward future self-referenced frequency combs that could be fully integrated on a chip in a robust package, perhaps along with other capabilities of advanced photonic integrated circuits (*170*). The same nanoscale waveguide fabrication also opens new opportunities by providing ultra-efficient nonlinear spectral broadening. For Er:fiber combs, this appears to be critical for applications such as space flight that would demand low electrical power (*171*). And for the older EO comb technology, this has made octave-span spectra at 10-GHz rates possible with a factor of 10 less power (*136*), which drives applications such as exoplanet astrophysics with a continuously running calibration comb (*61*). Thus, the connection between technology, applications, and science continues to grow.

In hindsight, perhaps the spread and interlacing of applications and technologies of Fig. 6 are a natural outcome, given the core synthesis and connectivity function of the frequency comb. And looking forward to the coming decades, we can only speculate on the new applications and technological roots that might emerge. It seems likely that as the technology becomes more integrated and tightly interfaced with applications, the boundary between electronics and optics will be further blurred, and the uniting aspects of the frequency comb on the electromagnetic spectrum will be more fully realized for advancing science.

REFERENCES AND NOTES

1. H. Hertz, *Untersuchungen über die Ausbreitung der elektrischen Kraft* (Johann Ambrosium Baath, Leipzig, 1894).
2. T. W. Hänsch, Nobel lecture: Passion for precision. *Rev. Mod. Phys.* **78**, 1297–1309 (2006). doi: 10.1103/RevModPhys.78.1297
3. J. L. Hall, Nobel lecture: Defining and measuring optical frequencies. *Rev. Mod. Phys.* **78**, 1279–1295 (2006). doi: 10.1103/RevModPhys.78.1279; pmid: 17086589
4. J. H. Bryant, The first century of microwaves-1886 to 1986. *IEEE Trans. Microw. Theory Tech.* **36**, 830–858 (1988). doi: 10.1109/22.3602
5. J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Macmillan and Co, London, 1873).
6. E. H. Armstrong, Some recent developments in the audion receiver. *Proc. Inst. Radio Eng.* **3**, 215 (1915).
7. J. P. Gordon, H. J. Zeiger, C. H. Townes, The maser—New type of microwave amplifier, frequency standard, and spectrometer. *Phys. Rev.* **99**, 1264–1274 (1955). doi: 10.1103/PhysRev.99.1264
8. A. L. Schawlow, C. H. Townes, Infrared and optical masers. *Phys. Rev.* **112**, 1940–1949 (1958). doi: 10.1103/PhysRev.112.1940
9. T. H. Maiman, R. H. Hoskins, I. J. D'Haenens, C. K. Asawa, V. Evtuhov, Stimulated optical emission in fluorescent solids. ii. spectroscopy and stimulated emission in ruby. *Phys. Rev.* **123**, 1151–1157 (1961). doi: 10.1103/PhysRev.123.1151
10. A. Javan, W. R. Bennett, D. R. Herriott, Population inversion and continuous optical maser oscillation in a gas discharge containing a he-ne mixture. *Phys. Rev. Lett.* **6**, 106–110 (1961). doi: 10.1103/PhysRevLett.6.106
11. A. Javan, E. A. Ballik, W. L. Bond, Frequency characteristics of a continuous-wave He–Ne optical maser. *J. Opt. Soc. Am.* **52**, 96 (1962). doi: 10.1364/JOSA.52.000096
12. L. O. Hocker, A. Javan, D. R. Rao, L. Frenkel, T. Sullivan, Absolute frequency measurement and spectroscopy of gas laser transitions in the far infrared. *Appl. Phys. Lett.* **10**, 147–149 (1967). doi: 10.1063/1.1754887
13. K. Evenson, J. Wells, F. Petersen, B. Danielson, G. Day, Accurate frequencies of molecular transitions used in laser stabilization: The 3.39-μm transition in CH_4 and the 9.33- and 10.18-μm transitions in CO_2. *Appl. Phys. Lett.* **22**, 192–195 (1973). doi: 10.1063/1.1654607
14. V. P. Chebotayev *et al*., Development of an optical time scale. *Appl. Phys. B* **29**, 63–65 (1982). doi: 10.1007/BF00694370
15. H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, G. Zinner, First phase-coherent frequency measurement of visible radiation. *Phys. Rev. Lett.* **76**, 18–21 (1996). doi: 10.1103/PhysRevLett.76.18; pmid: 10060423
16. D. J. Wineland, Laser-to-microwave frequency division using synchrotron radiation. *J. Appl. Phys.* **50**, 2528–2532 (1979). doi: 10.1063/1.326249
17. V. P. Chebotayev, Synchronization of Quantum Transitions. *JETP Lett.* **49**, 489 (1989).
18. K. M. Evenson *et al*., Speed of light from direct frequency and wavelength measurements of the methane-stabilized laser. *Phys. Rev. Lett.* **29**, 1346–1349 (1972). doi: 10.1103/PhysRevLett.29.1346
19. P. Giacomo, News from the BIPM: Some Changes in the SI. *IEEE Trans. Instrum. Meas.* **IM-34**, 116–117 (1985). doi: 10.1109/TIM.1985.4315282
20. M. Stock, R. Davis, E. de Mirandés, M. J. T. Milton, E. de Mirandés, M. Milton, The revision of the SI–The result of three decades of progress in metrology. *Metrologia* **56**, 022001 (2019). doi: 10.1088/1681-7575/ab0013
21. R. L. Barger, J. L. Hall, Pressure shift and broadening of methane line at 3.39 μm. *Phys. Rev. Lett.* **22**, 4–8 (1969). doi: 10.1103/PhysRevLett.22.4
22. S. N. Bagayev *et al*., A tunable laser at λ = 3.39 μm with line width of 7 Hz used in investigating a hyperfine structure of the $F_2^{(2)}$ line of methane. *Appl. Phys.* **13**, 291–297 (1977). doi: 10.1007/BF00882895
23. B. de Beauvoir *et al*., Absolute frequency measurement of the $2S - 8S/D$. *Phys. Rev. Lett.* **78**, 440–443 (1997). doi: 10.1103/PhysRevLett.78.440
24. T. Udem *et al*., Phase-Coherent Measurement of the Hydrogen 1S-2S Transition Frequency with an Optical Frequency Interval Divider Chain. *Phys. Rev. Lett.* **79**, 2646–2649 (1997). doi: 10.1103/PhysRevLett.79.2646
25. T. W. Hänsch, *The Hydrogen Atom*, G. F. Bassani, M. Inguscio, T. W. Hänsch, eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 1989), pp. 93–102.
26. T. Udem, Phasenkohärente optische Frequenzmessungen am Wasserstoffatom. Bestimmung der Rydberg–Konstanten und der 1S Lamb–Verschiebung, thesis, Ludwig-Maximilians-Universität München, Germany (1997).
27. K. Nakagawa, M. Kourogi, M. Ohtsu, Proposal of a frequency-synthesis chain between the microwave and optical frequencies of the Ca intercombination line at 657 nm using diode lasers. *Appl. Phys. B* **57**, 425–430 (1993). doi: 10.1007/BF00357386
28. J. N. Eckstein, High resolution spectroscopy using multiple coherent interactions, Ph.D. thesis, Stanford University (1978).
29. A. I. Ferguson, J. N. Eckstein, T. W. Hänsch, Polarization spectroscopy with ultrashort light pulses. *Appl. Phys. (Berl.)* **18**, 257–260 (1979). doi: 10.1007/BF00885511
30. T. Udem, R. Holzwarth, T. W. Hänsch, Optical frequency metrology. *Nature* **416**, 233–237 (2002). doi: 10.1038/416233a; pmid: 11894107
31. J. Reichert, R. Holzwarth, T. Udem, T. W. Hänsch, Measuring the frequency of light with mode-locked lasers. *Opt. Commun.* **172**, 59–68 (1999). doi: 10.1016/S0030-4018(99)00491-5
32. D. J. Jones *et al*., Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis. *Science* **288**, 635–640 (2000). doi: 10.1126/science.288.5466.635; pmid: 10784441
33. S. A. Diddams *et al*., Direct link between microwave and optical frequencies with a 300 THz femtosecond laser comb. *Phys. Rev. Lett.* **84**, 5102–5105 (2000). doi: 10.1103/PhysRevLett.84.5102; pmid: 10990877
34. R. Holzwarth *et al*., Optical frequency synthesizer for precision spectroscopy. *Phys. Rev. Lett.* **85**, 2264–2267 (2000). doi: 10.1103/PhysRevLett.85.2264; pmid: 10977987
35. T. Udem *et al*., Absolute frequency measurements of the Hg^+ and Ca optical clock transitions with a femtosecond laser Hg^+. *Phys. Rev. Lett.* **86**, 4996–4999 (2001). doi: 10.1103/PhysRevLett.86.4996; pmid: 11384404
36. J. Ye *et al*., Accuracy comparison of absolute optical frequency measurement between harmonic-generation synthesis and a frequency-division femtosecond comb. *Phys. Rev. Lett.* **85**, 3797–3800 (2000). doi: 10.1103/PhysRevLett.85.3797; pmid: 11041930
37. S. Bize *et al*., Testing the stability of fundamental constants with the $^{199}Hg^+$ single-ion optical clock $^{199}Hg^+$. *Phys. Rev. Lett.* **90**, 150802 (2003). doi: 10.1103/PhysRevLett.90.150802; pmid: 12732024

38. M. Fischer *et al*., New limits on the drift of fundamental constants from laboratory measurements. *Phys. Rev. Lett.* **92**, 230802 (2004). doi: 10.1103/PhysRevLett.92.230802; pmid: 15245149
39. T. M. Fortier *et al*., Precision atomic spectroscopy for improved limits on variation of the fine structure constant and local position invariance. *Phys. Rev. Lett.* **98**, 070801 (2007). doi: 10.1103/PhysRevLett.98.070801; pmid: 17359009
40. R. Holzwarth *et al*., Absolute frequency measurement of iodine lines with a femtosecond optical synthesizer. *Appl. Phys. B* **73**, 269–271 (2001). doi: 10.1007/s003400100633
41. V. Gerginov, C. E. Tanner, S. Diddams, A. Bartels, L. Hollberg, Optical frequency measurements of $6s^2\ s_{1/2} - 6p^2 p_{3/2}$ transition in a ^{133}Cs atomic beam using a femtosecond laser frequency comb. *Phys. Rev. A* **70**, 042505 (2004). doi: 10.1103/PhysRevA.70.042505
42. F.-L. Hong *et al*., Frequency measurements and hyperfine structure of the r(85)33–0 transition of molecular iodine with a femtosecond optical comb. *J. Opt. Soc. Am. B* **21**, 88 (2004). doi: 10.1364/JOSAB.21.000088
43. A. Marian, M. C. Stowe, J. R. Lawall, D. Felinto, J. Ye, United time-frequency spectroscopy for dynamics and global structure. *Science* **306**, 2063–2068 (2004). doi: 10.1126/science.1105660; pmid: 15550622
44. V. Gerginov, C. E. Tanner, S. A. Diddams, A. Bartels, L. Hollberg, High-resolution spectroscopy with a femtosecond laser frequency comb. *Opt. Lett.* **30**, 1734–1736 (2005). doi: 10.1364/OL.30.001734; pmid: 16075554
45. E. Baumann *et al*., Spectroscopy of the methane ν3 band with an accurate midinfrared coherent dual-comb spectrometer. *Phys. Rev. A* **84**, 062513 (2011). doi: 10.1103/PhysRevA.84.062513
46. A. M. Zolot *et al*., Broad-band frequency references in the near-infrared: Accurate dual comb spectroscopy of methane and acetylene. *J. Quant. Spectrosc. Radiat. Transf.* **118**, 26–39 (2013). doi: 10.1016/j.jqsrt.2012.11.024
47. M. J. Thorpe, K. D. Moll, R. J. Jones, B. Safdi, J. Ye, Broadband cavity ringdown spectroscopy for sensitive and rapid molecular detection. *Science* **311**, 1595–1599 (2006). doi: 10.1126/science.1123921; pmid: 16543457
48. B. Bernhardt *et al*., Cavity-enhanced dual-comb spectroscopy. *Nat. Photonics* **4**, 55–57 (2010). doi: 10.1038/nphoton.2009.217
49. S. A. Diddams, L. Hollberg, V. Mbele, Molecular fingerprinting with the resolved modes of a femtosecond laser frequency comb. *Nature* **445**, 627–630 (2007). doi: 10.1038/nature05524; pmid: 17287805
50. C. Gohle, B. Stein, A. Schliesser, T. Udem, T. W. Hänsch, Frequency comb Vernier spectroscopy for broadband, high-resolution, high-sensitivity absorption and dispersion spectra. *Phys. Rev. Lett.* **99**, 263902 (2007). doi: 10.1103/PhysRevLett.99.263902; pmid: 18233578
51. S. Schiller, Spectrometry with frequency combs. *Opt. Lett.* **27**, 766–768 (2002). doi: 10.1364/OL.27.000766; pmid: 18007926
52. F. Keilmann, C. Gohle, R. Holzwarth, Time-domain mid-infrared frequency-comb spectrometer. *Opt. Lett.* **29**, 1542–1544 (2004). doi: 10.1364/OL.29.001542; pmid: 15259740
53. I. Coddington, W. C. Swann, N. R. Newbury, Coherent multiheterodyne spectroscopy using stabilized optical frequency combs. *Phys. Rev. Lett.* **100**, 013902 (2008). doi: 10.1103/PhysRevLett.100.013902; pmid: 18232764
54. J. Roy, J.-D. Deschênes, S. Potvin, J. Genest, Continuous real-time correction and averaging for frequency comb interferometry. *Opt. Express* **20**, 21932–21939 (2012). doi: 10.1364/OE.20.021932; pmid: 23037343
55. T. Ideguchi *et al*., Coherent Raman spectro-imaging with laser frequency combs. *Nature* **502**, 355–358 (2013). doi: 10.1038/nature12607; pmid: 24132293
56. F. R. Giorgetta *et al*., Broadband phase spectroscopy over turbulent air paths. *Phys. Rev. Lett.* **115**, 103901 (2015). doi: 10.1103/PhysRevLett.115.103901; pmid: 26382677
57. I. Coddington, N. Newbury, W. Swann, Dual-comb spectroscopy. *Optica* **3**, 414 (2016). doi: 10.1364/OPTICA.3.000414
58. K. C. Cossel *et al*., Gas-phase broadband spectroscopy using active sources: Progress, status, and applications. *J. Opt. Soc. Am. B* **34**, 104–129 (2017). doi: 10.1364/JOSAB.34.000104; pmid: 28630530
59. N. Picqué, T. W. Hänsch, Frequency comb spectroscopy. *Nat. Photonics* **13**, 146–157 (2019). doi: 10.1038/s41566-018-0347-5
60. T. Steinmetz *et al*., Laser frequency combs for astronomical observations. *Science* **321**, 1335–1337 (2008). doi: 10.1126/science.1161030; pmid: 18772434
61. A. J. Metcalf *et al*., Stellar spectroscopy in the near-infrared with a laser frequency comb. *Optica* **6**, 233 (2019). doi: 10.1364/OPTICA.6.000233
62. R. A. Probst *et al*., A crucial test for astronomical spectrograph calibration with frequency combs, *Nature Astronomy* (2020).
63. S. A. Diddams *et al*., An optical clock based on a single trapped $^{199}Hg^+$ ion. *Science* **293**, 825–828 (2001). doi: 10.1126/science.1061171; pmid: 11452082
64. J. Ye, L. S. Ma, J. L. Hall, Molecular iodine clock. *Phys. Rev. Lett.* **87**, 270801 (2001). doi: 10.1103/PhysRevLett.87.270801; pmid: 11800866
65. D. Wineland, H. Dehmelt, Proposed $10^{14}\ \Delta\nu/\nu$ laser fluorescence spectroscopy on Tl^+ mono-ion oscillator. *Bull. Am. Phys. Soc.* **20**, 637 (1975).
66. T. W. Hänsch, A. L. Schawlow, Cooling of gases by laser radiation. *Opt. Commun.* **13**, 68–69 (1975). doi: 10.1016/0030-4018(75)90159-5
67. C. W. Chou, D. B. Hume, M. J. Thorpe, D. J. Wineland, T. Rosenband, Quantum coherence between two atoms beyond $Q = 10^{15}$. *Phys. Rev. Lett.* **106**, 160801 (2011). doi: 10.1103/PhysRevLett.106.160801; pmid: 21599347
68. B. C. Young, F. C. Cruz, W. M. Itano, J. C. Bergquist, Visible lasers with subhertz linewidths. *Phys. Rev. Lett.* **82**, 3799–3802 (1999). doi: 10.1103/PhysRevLett.82.3799
69. M. Notcutt, L.-S. Ma, J. Ye, J. L. Hall, Simple and compact 1-Hz laser system via an improved mounting configuration of a reference cavity. *Opt. Lett.* **30**, 1815–1817 (2005). doi: 10.1364/OL.30.001815; pmid: 16092355
70. K. Numata, A. Kemery, J. Camp, Thermal-noise limit in the frequency stabilization of lasers with rigid cavities. *Phys. Rev. Lett.* **93**, 250602 (2004). doi: 10.1103/PhysRevLett.93.250602; pmid: 15697887
71. J. M. Robinson *et al*., Crystalline optical cavity at 4 K with thermal-noise-limited instability and ultralow drift. *Optica* **6**, 240 (2019). doi: 10.1364/OPTICA.6.000240
72. T. L. Nicholson *et al*., Systematic evaluation of an atomic clock at 2×10^{-18} total uncertainty. *Nat. Commun.* **6**, 6896 (2015). doi: 10.1038/ncomms7896; pmid: 25898253
73. W. F. McGrew *et al*., Atomic clock performance enabling geodesy below the centimetre level. *Nature* **564**, 87–90 (2018). doi: 10.1038/s41586-018-0738-2; pmid: 30487601
74. S. M. Brewer *et al*., $^{27}Al^+$ Quantum-logic clock with a systematic uncertainty below 10^{-18}. *Phys. Rev. Lett.* **123**, 033201 (2019). doi: 10.1103/PhysRevLett.123.033201; pmid: 31386450
75. I. Ushijima, M. Takamoto, M. Das, T. Ohkubo, H. Katori, Cryogenic optical lattice clocks. *Nat. Photonics* **9**, 185–189 (2015). doi: 10.1038/nphoton.2015.5
76. N. Huntemann, C. Sanner, B. Lipphardt, C. Tamm, E. Peik, Single-ion atomic clock with 3×10^{-18} systematic uncertainty. *Phys. Rev. Lett.* **116**, 063001 (2016). doi: 10.1103/PhysRevLett.116.063001; pmid: 26918984
77. T. Fortier *et al*., Generation of ultrastable microwaves via optical frequency division. *Nat. Photonics* **5**, 425–429 (2011). doi: 10.1038/nphoton.2011.121
78. X. Xie *et al*., Photonic microwave signals with zeptosecond-level absolute timing noise. *Nat. Photonics* **11**, 44–47 (2017). doi: 10.1038/nphoton.2016.215
79. C. Grebing *et al*., Realization of a timescale with an accurate optical lattice clock. *Optica* **3**, 563 (2016). doi: 10.1364/OPTICA.3.000563
80. H. Hachisu, F. Nakagawa, Y. Hanado, T. Ido, Months-long real-time generation of a time scale based on an optical clock. *Sci. Rep.* **8**, 4243 (2018). doi: 10.1038/s41598-018-22423-5; pmid: 29523792
81. J. Yao *et al*., Optical-clock-based time scale. *Phys. Rev. Appl.* **12**, 044069 (2019). doi: 10.1103/PhysRevApplied.12.044069
82. W. R. Milner *et al*., Demonstration of a timescale based on a stable optical carrier. *Phys. Rev. Lett.* **123**, 173201 (2019). doi: 10.1103/PhysRevLett.123.173201; pmid: 31702265
83. H. Margolis, Timekeepers of the future. *Nat. Phys.* **10**, 82–83 (2014). doi: 10.1038/nphys2834
84. F. Riehle, P. Gill, F. Arias, L. Robertsson, The CIPM list of recommended frequency standard values: Guidelines and procedures. *Metrologia* **55**, 188–200 (2018). doi: 10.1088/1681-7575/aaa302
85. M. Kourogi, N. Ken'ichi, M. Ohtsu, Wide-Span Optical Frequency Comb Generator for Accurate Optical Frequency Difference Measurement. *IEEE J. Quantum Electron.* **29**, 2693–2701 (1993). doi: 10.1109/3.250392
86. J. Stenger, H. Schnatz, C. Tamm, H. R. Telle, Ultraprecise measurement of optical frequency ratios. *Phys. Rev. Lett.* **88**, 073601 (2002). doi: 10.1103/PhysRevLett.88.073601; pmid: 11863895
87. T. Rosenband *et al*., Frequency ratio of Al^+ and Hg^+ single-ion optical clocks; metrology at the 17th decimal place. *Science* **319**, 1808–1812 (2008). pmid: 18323415
88. R. M. Godun *et al*., Frequency ratio of two optical clock transitions in $^{171}Yb^+$ and constraints on the time variation of fundamental constants. *Phys. Rev. Lett.* **113**, 210801 (2014). doi: 10.1103/PhysRevLett.113.210801; pmid: 25479482
89. P. Wcisło *et al*., New bounds on dark matter coupling from a global network of optical atomic clocks. *Sci. Adv.* **4**, eaau4869 (2018). doi: 10.1126/sciadv.aau4869; pmid: 30539146
90. M. S. Safronova *et al*., Search for new physics with atoms and molecules. *Rev. Mod. Phys.* **90**, 025008 (2018). doi: 10.1103/RevModPhys.90.025008
91. F. R. Giorgetta *et al*., Optical two-way time and frequency transfer over free space. *Nat. Photonics* **7**, 434–438 (2013). doi: 10.1038/nphoton.2013.69
92. I. Coddington, W. C. Swann, L. Nenadovic, N. R. Newbury, Rapid and precise absolute distance measurements at long range. *Nat. Photonics* **3**, 351–356 (2009). doi: 10.1038/nphoton.2009.94
93. L. Xu *et al*., Route to phase control of ultrashort light pulses. *Opt. Lett.* **21**, 2008–2010 (1996). doi: 10.1364/OL.21.002008; pmid: 19881875
94. H. R. Telle *et al*., Carrier-envelope offset phase control: A novel concept for absolute optical frequency measurement and ultrashort pulse generation. *Appl. Phys. B* **69**, 327–332 (1999). doi: 10.1007/s003400050813
95. J. K. Ranka, R. S. Windeler, A. J. Stentz, *Technical Digest of Papers Presented at the Conference on Lasers and Electro-Optics. CLEO '99* (IEEE catalog no.99CH37013) (1999), pp. CPD8/1–CPD8/2.
96. W. J. Wadsworth *et al*., Soliton effects in photonic crystal fibres at 850 nm. *Electron. Lett.* **36**, 53 (2000). doi: 10.1049/el:20000134
97. J. K. Ranka, R. S. Windeler, A. J. Stentz, Visible continuum generation in air-silica microstructure optical fibers with anomalous dispersion at 800 nm. *Opt. Lett.* **25**, 25–27 (2000). doi: 10.1364/OL.25.000025; pmid: 18059770
98. L. E. Hargrove, R. L. Fork, M. A. Pollack, Locking of He-Ne laser modes induced by synchronous intracavity modulation. *Appl. Phys. Lett.* **5**, 4–5 (1964). doi: 10.1063/1.1754025
99. R. L. Fork, B. I. Greene, C. V. Shank, Generation of optical pulses shorter than 0.1 psec by colliding pulse mode locking. *Appl. Phys. Lett.* **38**, 671–672 (1981). doi: 10.1063/1.92500
100. D. E. Spence, P. N. Kean, W. Sibbett, 60-fsec pulse generation from a self-mode-locked Ti:sapphire laser. *Opt. Lett.* **16**, 42–44 (1991). doi: 10.1364/OL.16.000042; pmid: 19773831
101. P. Russell, Photonic crystal fibers. *Science* **299**, 358–362 (2003). doi: 10.1126/science.1079280; pmid: 12532007
102. U. Morgner *et al*., Nonlinear optics with phase-controlled pulses in the sub-two-cycle regime. *Phys. Rev. Lett.* **86**, 5462–5465 (2001). doi: 10.1103/PhysRevLett.86.5462; pmid: 11415276
103. T. M. Fortier, D. J. Jones, S. T. Cundiff, Phase stabilization of an octave-spanning Ti:sapphire laser. *Opt. Lett.* **28**, 2198–2200 (2003). doi: 10.1364/OL.28.002198; pmid: 14649940
104. T. M. Fortier, A. Bartels, S. A. Diddams, Octave-spanning Ti:sapphire laser with a repetition rate >1 GHz for optical frequency measurements and comparisons. *Opt. Lett.* **31**, 1011–1013 (2006). doi: 10.1364/OL.31.001011; pmid: 16599240
105. R. J. Mears, L. Reekie, I. M. Jauncey, D. N. Payne, Low-noise erbium-doped fibre amplifier operating at 1.54 μm. *Electron. Lett.* **23**, 1026 (1987). doi: 10.1049/el:19870719
106. J. W. Nicholson *et al*., All-fiber, octave-spanning supercontinuum. *Opt. Lett.* **28**, 643–645 (2003). doi: 10.1364/OL.28.000643; pmid: 12703927
107. F.-L. Hong *et al*., Broad-spectrum frequency comb generation and carrier-envelope offset frequency measurement by second-harmonic generation of a mode-locked fiber laser. *Opt. Lett.* **28**, 1516–1518 (2003). doi: 10.1364/OL.28.001516; pmid: 12956364
108. F. Tauser, A. Leitenstorfer, W. Zinth, Amplified femtosecond pulses from an Er:fiber system: Nonlinear pulse shortening and self-referencing detection of the carrier-envelope phase evolution. *Opt. Express* **11**, 594–600 (2003). doi: 10.1364/OE.11.000594; pmid: 19461769
109. B. R. Washburn *et al*., Phase-locked, erbium-fiber-laser-based frequency comb in the near infrared. *Opt. Lett.* **29**, 250–252 (2004). doi: 10.1364/OL.29.000250; pmid: 14759041
110. T. R. Schibli *et al*., Frequency metrology with a turnkey all-fiber system. *Opt. Lett.* **29**, 2467–2469 (2004). doi: 10.1364/OL.29.002467; pmid: 15584263

111. S. Droste, G. Ycas, B. Washburn, I. Coddington, N. Newbury, Optical Frequency Comb Generation based on Erbium Fiber Lasers. *Nanophotonics* **5**, 196–213 (2016). doi: 10.1515/nanoph-2016-0019
112. L. C. Sinclair *et al*., Invited Article: A compact optically coherent fiber frequency comb. *Rev. Sci. Instrum.* **86**, 081301 (2015). doi: 10.1063/1.4928163; pmid: 26329167
113. M. Lezius *et al*., Space-borne frequency comb metrology. *Optica* **3**, 1381 (2016). doi: 10.1364/OPTICA.3.001381
114. S. A. Diddams, The evolving optical frequency comb. *J. Opt. Soc. Am. B* **27**, B51 (2010). doi: 10.1364/JOSAB.27.000B51
115. M. E. Fermann, I. Hartl, Ultrafast fibre lasers. *Nat. Photonics* **7**, 868–874 (2013). doi: 10.1038/nphoton.2013.280
116. T. Yasui, Y. Kabetani, E. Saneyoshi, S. Yokoyama, T. Araki, Terahertz frequency comb by multifrequency-heterodyning photoconductive detection for high-accuracy, high-resolution terahertz spectroscopy. *Appl. Phys. Lett.* **88**, 241104 (2006). doi: 10.1063/1.2209718
117. I. A. Finneran *et al*., Decade-spanning high-precision terahertz frequency comb. *Phys. Rev. Lett.* **114**, 163902 (2015). doi: 10.1103/PhysRevLett.114.163902; pmid: 25955051
118. F. Adler *et al*., Phase-stabilized, 1.5 W frequency comb at 2.8–4.8 μm. *Opt. Lett.* **34**, 1330–1332 (2009). doi: 10.1364/OL.34.001330; pmid: 19412262
119. A. Schliesser, N. Picqué, T. W. Hänsch, Mid-infrared frequency combs. *Nat. Photonics* **6**, 440–449 (2012). doi: 10.1038/nphoton.2012.142
120. A. V. Muraviev, V. O. Smolski, Z. E. Loparo, K. L. Vodopyanov, Massively parallel sensing of trace molecules and their isotopologues with broadband subharmonic mid-infrared frequency combs. *Nat. Photonics* **12**, 209–214 (2018). doi: 10.1038/s41566-018-0135-2
121. G. Ycas *et al*., High-coherence mid-infrared dual-comb spectroscopy spanning 2.6 to 5.2μ. *Nat. Photonics* **12**, 202–208 (2018). doi: 10.1038/s41566-018-0114-7
122. A. S. Kowligy *et al*., Infrared electric field sampled frequency comb spectroscopy. *Sci. Adv.* **5**, eaaw8794 (2019). doi: 10.1126/sciadv.aaw8794; pmid: 31187063
123. C. Gohle *et al*., A frequency comb in the extreme ultraviolet. *Nature* **436**, 234–237 (2005). doi: 10.1038/nature03851; pmid: 16015324
124. R. J. Jones, K. D. Moll, M. J. Thorpe, J. Ye, Phase-coherent frequency combs in the vacuum ultraviolet via high-harmonic generation inside a femtosecond enhancement cavity. *Phys. Rev. Lett.* **94**, 193201 (2005). doi: 10.1103/PhysRevLett.94.193201; pmid: 16090171
125. A. Cingöz *et al*., Direct frequency comb spectroscopy in the extreme ultraviolet. *Nature* **482**, 68–71 (2012). doi: 10.1038/nature10711; pmid: 22297971
126. I. Pupeza *et al*., Compact high-repetition-rate source of coherent 100 eV radiation. *Nat. Photonics* **7**, 608–612 (2013). doi: 10.1038/nphoton.2013.156
127. T. Kobayashi, T. Sueta, Y. Cho, Y. Matsuo, High-repetition-rate optical pulse generator using a Fabry-Perot electro-optic modulator. *Appl. Phys. Lett.* **21**, 341–343 (1972). doi: 10.1063/1.1654403
128. M. Kourogi, B. Widiyatomoko, Y. Takeuchi, M. Ohtsu, Limit of optical-frequency comb generation due to material dispersion. *IEEE J. Quantum Electron.* **31**, 2120–2126 (1995). doi: 10.1109/3.477736
129. K. Imai, M. Kourogi, M. Ohtsu, 30-THz span optical frequency comb generation by self-phase modulation in an optical fiber. *IEEE J. Quantum Electron.* **34**, 54–60 (1998). doi: 10.1109/3.655007
130. S. A. Diddams, L.-S. Ma, J. Ye, J. L. Hall, Broadband optical frequency comb generation with a phase-modulated parametric oscillator. *Opt. Lett.* **24**, 1747–1749 (1999). doi: 10.1364/OL.24.001747; pmid: 18079922
131. M. Zhang *et al*., Broadband electro-optic frequency comb generation in a lithium niobate microring resonator. *Nature* **568**, 373–377 (2019). doi: 10.1038/s41586-019-1008-7; pmid: 30858615
132. I. Morohashi *et al*., Widely repetition-tunable 200 fs pulse source using a Mach-Zehnder-modulator-based flat comb generator and dispersion-flattened dispersion-decreasing fiber. *Opt. Lett.* **33**, 1192–1194 (2008). doi: 10.1364/OL.33.001192; pmid: 18516170
133. A. Ishizawa *et al*., Generation of 120-fs laser pulses at 1-GHz repetition rate derived from continuous wave laser diode. *Opt. Express* **19**, 22402–22409 (2011). doi: 10.1364/OE.19.022402; pmid: 22109116
134. A. J. Metcalf, V. Torres-Company, D. E. Leaird, A. M. Weiner, High-power broadly tunable electrooptic frequency comb generator. *IEEE J. Sel. Top. Quantum Electron.* **19**, 231–236 (2013). doi: 10.1109/JSTQE.2013.2268384
135. K. Beha *et al*., Electronic synthesis of light. *Optica* **4**, 406 (2017). doi: 10.1364/OPTICA.4.000406
136. D. R. Carlson *et al*., Ultrafast electro-optic light with subcycle control. *Science* **361**, 1358–1363 (2018). doi: 10.1126/science.aat6451; pmid: 30262499
137. T. J. Kippenberg, R. Holzwarth, S. A. Diddams, Microresonator-based optical frequency combs. *Science* **332**, 555–559 (2011). doi: 10.1126/science.1193968; pmid: 21527707
138. D. K. Armani, T. J. Kippenberg, S. M. Spillane, K. J. Vahala, Ultra-high-Q toroid microcavity on a chip. *Nature* **421**, 925–928 (2003). doi: 10.1038/nature01371; pmid: 12606995
139. K. J. Vahala, Optical microcavities. *Nature* **424**, 839–846 (2003). doi: 10.1038/nature01939; pmid: 12917698
140. I. S. Grudinin, V. S. Ilchenko, L. Maleki, Ultrahigh optical Q factors of crystalline resonators in the linear regime. *Phys. Rev. A* **74**, 063806 (2006). doi: 10.1103/PhysRevA.74.063806
141. J. S. Levy *et al*., CMOS-compatible multiple-wavelength oscillator for on-chip optical interconnects. *Nat. Photonics* **4**, 37–40 (2010). doi: 10.1038/nphoton.2009.259
142. Q. Li *et al*., Stably accessing octave-spanning microresonator frequency combs in the soliton regime. *Optica* **4**, 193–203 (2017). doi: 10.1364/OPTICA.4.000193; pmid: 28603754
143. T. J. Kippenberg, S. M. Spillane, K. J. Vahala, Kerr-nonlinearity optical parametric oscillation in an ultrahigh-Q toroid microcavityq. *Phys. Rev. Lett.* **93**, 083904 (2004). doi: 10.1103/PhysRevLett.93.083904; pmid: 15447188
144. A. A. Savchenkov *et al*., Low threshold optical oscillations in a whispering gallery mode CaF_2 resonator. *Phys. Rev. Lett.* **93**, 243905 (2004). doi: 10.1103/PhysRevLett.93.243905; pmid: 15697815
145. P. Del'Haye, *et al*., Optical frequency comb generation from a monolithic microresonator, *Nature* **450**, 1214 (2007).
146. S. Wabnitz, Suppression of interactions in a phase-locked soliton optical memory. *Opt. Lett.* **18**, 601 (1993). doi: 10.1364/OL.18.000601; pmid: 19802213
147. F. Leo *et al*., Temporal cavity solitons in one-dimensional Kerr media as bits in an all-optical buffer. *Nat. Photonics* **4**, 471–476 (2010). doi: 10.1038/nphoton.2010.120
148. T. Herr *et al*., Temporal solitons in optical microresonators. *Nat. Photonics* **8**, 145–152 (2014). doi: 10.1038/nphoton.2013.343
149. X. Yi, Q.-F. Yang, K. Y. Yang, M.-G. Suh, K. Vahala, Soliton frequency comb at microwave rates in a high-Q silica microresonator. *Optica* **2**, 1078 (2015). doi: 10.1364/OPTICA.2.001078
150. C. Joshi *et al*., Thermally controlled comb generation and soliton modelocking in microresonators. *Opt. Lett.* **41**, 2565–2568 (2016). doi: 10.1364/OL.41.002565; pmid: 27244415
151. P.-H. Wang *et al*., Intracavity characterization of micro-comb generation in the single-soliton regime. *Opt. Express* **24**, 10890–10897 (2016). doi: 10.1364/OE.24.010890; pmid: 27409909
152. T. J. Kippenberg, A. L. Gaeta, M. Lipson, M. L. Gorodetsky, Dissipative Kerr solitons in optical microresonators. *Science* **361**, eaan8083 (2018). doi: 10.1126/science.aan8083; pmid: 30093576
153. Y. Okawachi *et al*., Octave-spanning frequency comb generation in a silicon nitride chip. *Opt. Lett.* **36**, 3398–3400 (2011). doi: 10.1364/OL.36.003398; pmid: 21886223
154. V. Brasch *et al*., Photonic chip-based optical frequency comb using soliton Cherenkov radiation. *Science* **351**, 357–360 (2016). doi: 10.1126/science.aad4811; pmid: 26721682
155. M. H. P. Pfeiffer *et al*., Octave-spanning dissipative kerr soliton frequency combs in Si_3N_4 microresonators. *Optica* **4**, 684 (2017). doi: 10.1364/OPTICA.4.000684
156. Z. Gong *et al*., High-fidelity cavity soliton generation in crystalline AlN micro-ring resonators. *Opt. Lett.* **43**, 4366–4369 (2018). doi: 10.1364/OL.43.004366; pmid: 30211865
157. Y. He *et al*., Self-starting bi-chromatic $LiNbO_3$ soliton microcomb. *Optica* **6**, 1138 (2019). doi: 10.1364/OPTICA.6.001138
158. L. Chang *et al*., Ultra-efficient frequency comb generation in AlGaAs-on-insulator microresonators. *Nat. Commun.* **11**, 1331 (2020). doi: 10.1038/s41467-020-15005-5; pmid: 32165610
159. D. T. Spencer *et al*., An optical-frequency synthesizer using integrated photonics. *Nature* **557**, 81–85 (2018). doi: 10.1038/s41586-018-0065-7; pmid: 29695870
160. S. B. Papp *et al*., Microresonator frequency comb optical clock. *Optica* **1**, 10 (2014). doi: 10.1364/OPTICA.1.000010
161. Z. L. Newman *et al*., Architecture for the photonic integration of an optical atomic clock. *Optica* **6**, 680 (2019). doi: 10.1364/OPTICA.6.000680
162. M.-G. Suh *et al*., Searching for exoplanets using a microresonator astrocomb. *Nat. Photonics* **13**, 25–30 (2019). doi: 10.1038/s41566-018-0312-3; pmid: 30740138
163. E. Obrzud *et al*., A microphotonic astrocomb. *Nat. Photonics* **13**, 31–35 (2019). doi: 10.1038/s41566-018-0309-y
164. P. Marin-Palomo *et al*., Microresonator-based solitons for massively parallel coherent optical communications. *Nature* **546**, 274–279 (2017). doi: 10.1038/nature22387; pmid: 28593968
165. M.-G. Suh, Q.-F. Yang, K. Y. Yang, X. Yi, K. J. Vahala, Microresonator soliton dual-comb spectroscopy. *Science* **354**, 600–603 (2016). doi: 10.1126/science.aah6516; pmid: 27738017
166. M. Yu *et al*., Silicon-chip-based mid-infrared dual-comb spectroscopy. *Nat. Commun.* **9**, 1869 (2018). doi: 10.1038/s41467-018-04350-1; pmid: 29760418
167. M.-G. Suh, K. J. Vahala, Soliton microcomb range measurement. *Science* **359**, 884–887 (2018). doi: 10.1126/science.aao1968; pmid: 29472476
168. P. Trocha *et al*., Ultrafast optical ranging using microresonator soliton frequency combs. *Science* **359**, 887–891 (2018). doi: 10.1126/science.aao3924; pmid: 29472477
169. J. N. Eckstein, A. I. Ferguson, T. W. Hänsch, High-resolution two-photon spectroscopy with picosecond light pulses. *Phys. Rev. Lett.* **40**, 847–850 (1978). doi: 10.1103/PhysRevLett.40.847
170. M. J. R. Heck *et al*., Hybrid silicon photonic integrated circuit technology. *IEEE J. Sel. Top. Quantum Electron.* **19**, 6100117 (2013). doi: 10.1109/JSTQE.2012.2235413
171. P. Manurkar *et al*., Fully self-referenced frequency comb consuming 5 Watts of electrical power. *OSA Continuum* **1**, 274 (2018). doi: 10.1364/OSAC.1.000274
172. T. C. Briles *et al*., Interlocking Kerr-microresonator frequency combs for microwave to optical synthesis. *Opt. Lett.* **43**, 2933–2936 (2018). doi: 10.1364/OL.43.002933; pmid: 29905727
173. D. R. Carlson *et al*., Photonic-chip supercontinuum with tailored spectra for counting optical frequencies. *Phys. Rev. Appl.* **8**, 014027 (2017). doi: 10.1103/PhysRevApplied.8.014027

ACKNOWLEDGMENTS

We are grateful to T. Hänsch and J. Hall for their input, vision, and leadership in this field, as well as to N. Newbury, C. Oates, and J. Ye for their comments on this manuscript. We also thank B. Baxley, T. Briles, J. Knight, S. Papp, P. Russell, R. Windeler, and the University of Bath for their contributions to the figures. Finally, we are indebted to the many colleagues with whom we have had the privilege to learn about the ever-expanding field of optical frequency combs. **Funding:** This material is based on work supported by the National Institute of Standards and Technology (NIST); the Defense Advanced Research Projects Agency (DARPA) under awards in the DODOS (HR001 1-15-C-0055), APHi (FA9453-19-C-0029), SCOUT (W31P4Q-15-1-0011), and ACES programs; and the Air Force Office of Scientific Research (AFOSR) under awards FA9550-16-1-0016 and FA9550-18-1-0353. **Competing interests:** T.U. is an inventor on patents US6785303B1 and US20060268949 held by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften that cover optical frequency combs.

10.1126/science.aay3676

RESEARCH ARTICLE SUMMARY

CORAL GENOMICS

Population genetics of the coral *Acropora millepora*: Toward genomic prediction of bleaching

Zachary L. Fuller*, Veronique J. L. Mocellin, Luke A. Morris, Neal Cantin, Jihanne Shepherd, Luke Sarre, Julie Peng, Yi Liao, Joseph Pickrell, Peter Andolfatto, Mikhail Matz†, Line K. Bay*†, Molly Przeworski*†

INTRODUCTION: Coral reefs worldwide are suffering losses at an alarming rate as a result of anthropogenic climate change. Increased seawater temperatures, even only slightly above long-term maxima, can induce bleaching—the breakdown of the symbiotic relationship between coral hosts and their intracellular photosynthetic dinoflagellates from the family Symbiodiniaceae. Because these symbionts provide the majority of energy required by the coral host, prolonged periods of bleaching can eventually lead to the death of the colony. In the face of rapidly increasing temperatures, new conservation strategies are urgently needed to prevent future mass losses of coral cover, and these benefit from an understanding of the genetic basis of bleaching.

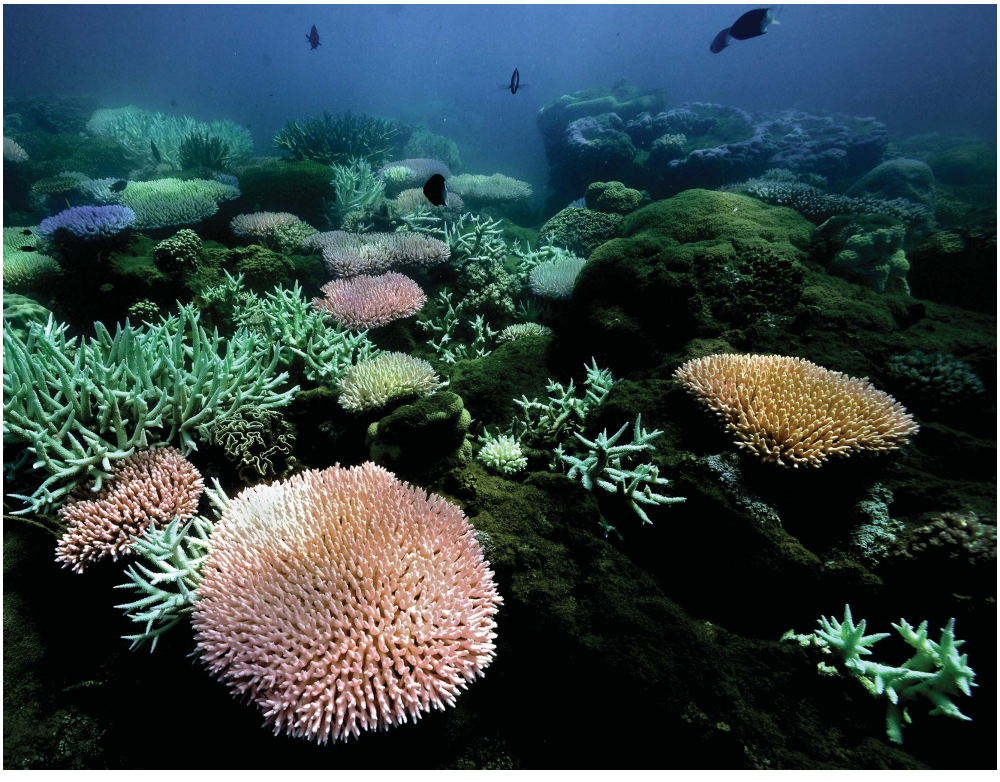
RATIONALE: Bleaching responses vary within and among coral species; in the reef-building coral *Acropora millepora*, a commonly distributed species across the Indo-Pacific, these differences have been shown to be at least partly heritable. In principle, therefore, interindividual differences in bleaching should be predictable from genomic data. Here, we demonstrate the feasibility of using a genomics-based approach to predict individual bleaching responses and suggest ways in which this can inform new strategies for coral conservation.

RESULTS: We first generated a chromosome-scale genome assembly as well as whole-genome sequences for 237 samples collected at 12 reefs distributed across the central Great Barrier Reef during peak bleaching in 2017. We showed that we can reliably impute genotypes in low-coverage sequencing data with a modestly sized reference haplotype panel, demonstrating a cost-effective approach for future large-scale whole-genome sequencing efforts. Very little population structure was detected across the sampled reefs, which was likely the result of the broadcast spawning mode of reproduction in *A. millepora*. Against this genomic background, we detected unusually old variation at the heat-shock co-chaperone *sacsin*, which is consistent with long-term balancing selection acting on this gene. Our genomic sequencing approach simultaneously provides a quantitative measure of bleaching and identifies the composition of symbiont species present within individual coral hosts. Testing more than 6.8 million variants for associations with three different measures of bleaching response, no single site reached genome-wide significance, indicating that variation in bleaching response is not due to common loci of large effect. However, a model that incorporates genetic effects estimated from the genome-wide association data, genomic data on relative symbiont species composition, and environmental variables is predictive of individual bleaching phenotypes.

ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aba4674

CONCLUSION: Understanding the genetics of heat and bleaching tolerance will be critical to predict coral adaptation and the future of coral reef ecosystems under climate change. This knowledge also supports both conventional management approaches and the development of new interventions. Our work provides insight into the genetic architecture of bleaching response and serves as a proof of principle for the use of genomic approaches in conservation efforts. We show that a model based on environmental factors, genomic data from the symbiont, and genome-wide association data in the coral host can help distinguish individuals most tolerant to bleaching from those that are most susceptible. These results thus build a foundation toward a genomic predictor of bleaching response in *A. millepora* and other coral species. ■



Bleached *A. millepora* colonies on the central Great Barrier Reef. *A. millepora* colonies presenting various severity of bleaching during March 2017 at Feather Reef on the Great Barrier Reef. Colonies with a greater severity of bleaching are those with the most pale colors. Prolonged periods of bleaching can lead to the eventual death of the coral host.

PHOTO: ERIC MASON/AUSTRALIAN INSTITUTE OF MARINE SCIENCE

The list of author affiliations is available in the full article online.
*Corresponding author. Email: zlf2102@columbia.edu (Z.L.F.); mp3284@columbia.edu (M.P.); l.bay@aims.gov.au (L.K.B.)
†These authors contributed equally to this work.
Cite this article as Z. L. Fuller *et al.*, *Science* **369**, eaba4674 (2020). DOI: 10.1126/science.aba4674

RESEARCH ARTICLE

CORAL GENOMICS

Population genetics of the coroal *Acropora millepora*: Toward genomic prediction of bleaching

Zachary L. Fuller[1]*, Veronique J. L. Mocellin[2], Luke A. Morris[2,3,4], Neal Cantin[2], Jihanne Shepherd[1], Luke Sarre[1], Julie Peng[5], Yi Liao[6,7], Joseph Pickrell[8], Peter Andolfatto[1], Mikhail Matz[6]†, Line K. Bay[2]*†, Molly Przeworski[1,9,10]*†

Although reef-building corals are declining worldwide, responses to bleaching vary within and across species and are partly heritable. Toward predicting bleaching response from genomic data, we generated a chromosome-scale genome assembly for the coral *Acropora millepora*. We obtained whole-genome sequences for 237 phenotyped samples collected at 12 reefs along the Great Barrier Reef, among which we inferred little population structure. Scanning the genome for evidence of local adaptation, we detected signatures of long-term balancing selection in the heat-shock co-chaperone *sacsin*. We conducted a genome-wide association study of visual bleaching score for 213 samples, incorporating the polygenic score derived from it into a predictive model for bleaching in the wild. These results set the stage for genomics-based approaches in conservation strategies.

Anthropogenic global warming is transforming ecosystems worldwide, pressing species to rapidly respond to changing environments (*1*–*4*). To what extent they will succeed is unclear; some may undergo range shifts or tolerate climate change through phenotypic plasticity. If there is sufficient heritable variation in traits that confer a fitness advantage in the new environment, others may adapt through evolutionary change (*5*–*9*). Particularly devastating has been the impact of increased seawater temperatures on marine ecosystems—notably on coral reefs, which cover less than 1% of the ocean floor but are home to more than a quarter of its total biodiversity (*10*). Ecological stress brought on by changes in temperature or salinity breaks down the symbiotic relationship between reef-building corals and their intracellular photosynthetic dinoflagellates (family Symbiodiniaceae) in a phenomenon known as bleaching (*11*). Because these symbionts provide most of the energy required to the coral host, prolonged periods of bleaching can lead to the death of the coral. Increased temperatures can also cause direct cellular damage and necrosis to bleached corals (*12*). Temperatures only slightly above long-term maxima can result in corals over vast areas losing their symbionts, decimating reefs worldwide (*13*–*15*).

Despite the vulnerability of corals to warming, there is phenotypic variation in the bleaching and heat stress response both within and among coral species (*16*–*18*). In the coral *Acropora millepora*, a common reef species across the Indo-Pacific, variation in heat tolerance among aposymbiotic larvae has been shown to be partly heritable (*19*). This finding implies both that a response to selection is possible and that bleaching should be partly predicted from genomic data. Simulation studies have suggested the potential for rapid adaptation in *A. millepora* populations (*20*), and gene expression analyses have associated genes related to oxidative, extracellular, transport, and mitochondrial functions with elevated heat tolerance (*19*). However, because previous studies rely on transcriptomics or reduced representation sequencing (*19*–*22*), they cannot distinguish causes from effects of bleaching, provide only a partial view of the genome, and are hampered by the lack of a high-quality, publicly available reference assembly. Thus, we remain far from a genetic understanding of bleaching and thermal tolerance and lack the ability to predict interindividual variation in the responses. In particular, the heritability of a trait does not specify its genetic architecture (*23*), so it remains unclear whether variation in heat tolerance is polygenic or influenced by a small number of common large-effect loci.

[1]Department of Biological Sciences, Columbia University, New York, NY, USA. [2]Australian Institute of Marine Science, Townsville, QLD, Australia. [3]AIMS@JCU, Australian Institute of Marine Science, College of Science and Engineering, James Cook University, Townsville, QLD, Australia. [4]College of Science and Engineering, James Cook University, Townsville, QLD, Australia. [5]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [6]Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA. [7]Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA, USA. [8]Gencove, New York, NY, USA. [9]Department of Systems Biology, Columbia University, New York, NY, USA. [10]Program for Mathematical Genomics, Columbia University, New York, NY, USA.
***Corresponding author. Email: zlf2102@columbia.edu (Z.L.F.); mp3284@columbia.edu (M.P.); l.bay@aims.gov.au (L.K.B.)**
†These authors contributed equally to this work.

Inferring population structure and demographic history from genome-wide patterns of variation

Heritable traits can be predicted from the results of genome-wide association studies (GWASs) alongside other predictors, an approach widespread in agriculture and increasingly in human genetics (*24*–*26*). We applied a similar approach to conservation biology, focusing on prediction of bleaching in wild populations of *A. millepora*. To this end, we assembled a highly contiguous chromosome-level de novo *A. millepora* genome and generated high-coverage, whole-genome resequencing data for samples collected across 12 reefs on the Great Barrier Reef (GBR).

First, we constructed a chromosome-scale assembly of the diploid *A. millepora* genome using a combination of PacBio and 10X Chromium barcoded Illumina reads generated from DNA extracted from a single adult colony (Fig. 1A, figs. S1 to S6, table S1, and materials and methods). We additionally sequenced two pools of aposymbiotic larvae and mapped these reads to our de novo assembly to remove symbiont contigs. In total, we assembled 475 Mb of sequence with chromosome-scale scaffolds for an N50 of 19.8 Mb (where N50 is the minimum scaffold length needed to cover 50% of the genome) (figs. S1 to S6). Using previously published transcriptome data (*27*, *28*), we predicted 28,188 gene models, which include >96% of core-eukaryotic single-copy orthologs. A contiguous reference genome facilitates complex trait mapping (*29*) and reliable demographic inference (*30*).

To investigate genetic variation associated with bleaching response, we collected branch fragments from a total of 253 individual *A. millepora* colonies during peak bleaching in March 2017 from 12 reefs (Fig. 1B and table S2) distributed along an ~300-km stretch of the central GBR. Ecological conditions were measured by use of 40 environmental and spatial variables [recorded at a resolution of 1 km^2 (*31*)] that varied among reefs, with the strongest axes of variation being longitude and latitude (Fig. 1C). In particular, mean monthly sea surface temperatures differed by more than 0.5°C. The depth of colonies from which fragments were collected varied from 1.2 to 7.2 m, representing the range of depths where *A. millepora* is commonly found on reef flats and upper reef slopes (*32*). For each colony, we recorded the bleaching level visually using the six-point coral health chart to the nearest half increment (*33*) then later estimated photosynthetic pigment concentration and symbiont cell density (Fig. 1D). Individual colonies sampled presented a range of bleaching phenotypes, both within and between reefs (fig. S7).

From this collection of samples, we resequenced 48 whole genomes of diploid

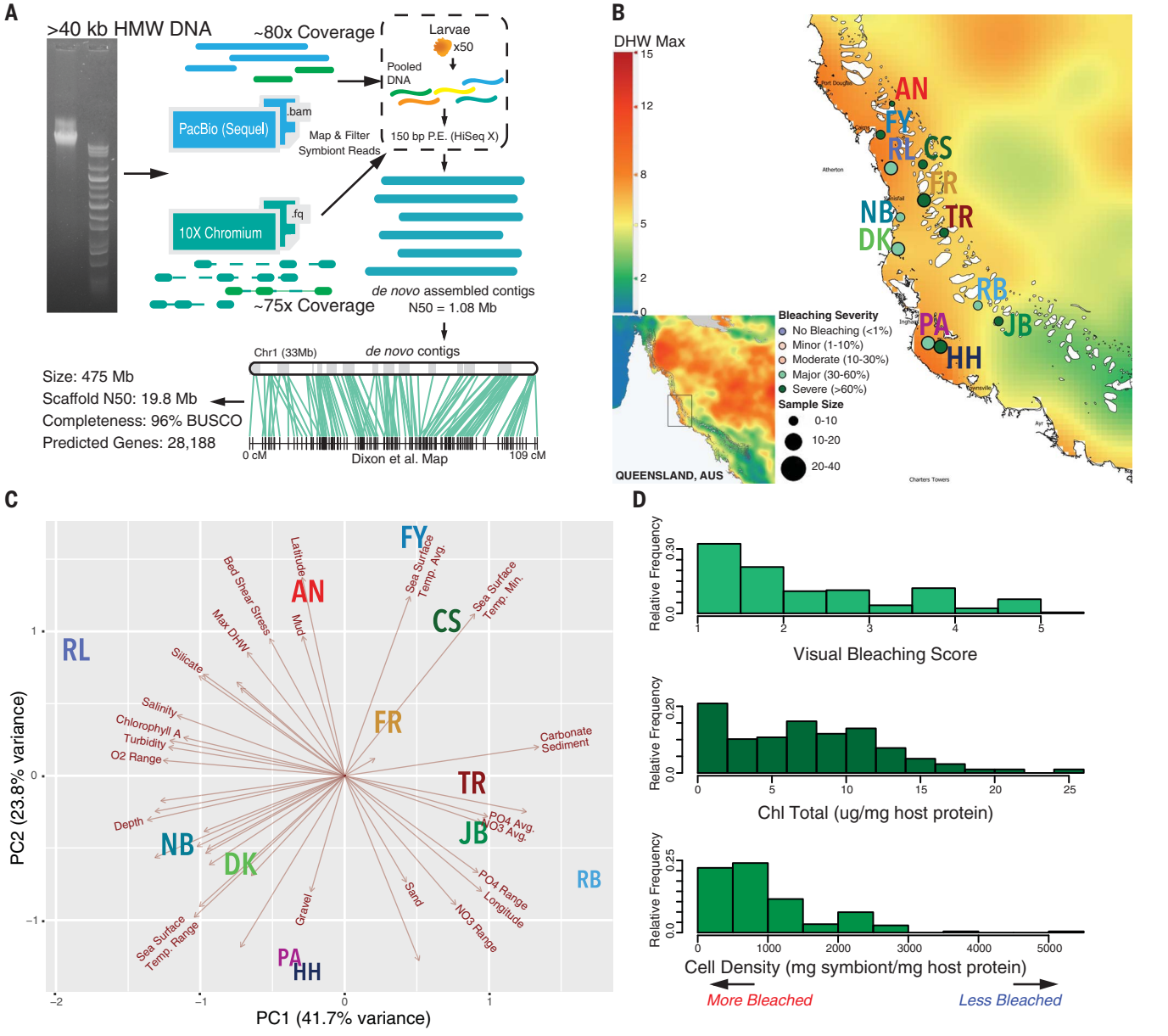


Fig. 1. Genome assembly and sample collection for *A. millepora*. (**A**) A de novo assembly for the *A. millepora* genome was constructed by using a hybrid sequencing approach. The alignment of reads generated from two pools of aposymbiotic larvae was used to filter out symbiont contigs (fig. S3). Assembled coral contigs were aligned to previously published linkage maps (*19*, *80*) to create a chromosome-scale assembly. (**B**) A total of 253 individuals were collected from across 12 reefs on the GBR in 2017. The size of the circles represents the number of individuals collected at each site, and the circles are colored according to the bleaching severity at each reef. The maximum degree heating weeks (DHW) is shown across the region from which samples were collected (and across the GBR in the inset). The spatial layer of DHW represents interpolated maximum values from the National Oceanic and Atmospheric Administration Coral Reef Watch (CRW) v3.1 satellite product at a resolution of 5 km. Each reef label is colored arbitrarily but consistent with labels presented in other figures. AN, Arlington Reef; FY, Fitzroy Reef; RL, Russell Island; CS, Coates Reef; FR, Feather Reef; NB, North Barnard Islands; TR, Taylor Reef; DK, Dunk Island; RB, Rib Reef; PA, Pandora Reef; JB, John Brewer Reef; HH, Havannah Island. (**C**) A PCA was performed for 40 environmental and spatial variables for each reef, with abbreviations in bold representing their location. The component loads for each environmental variable projected on the first two principal components are depicted with arrows. (**D**) The distribution of visual scores, chlorophyll abundance standardized by host coral protein content, and standardized symbiont cell densities among the collected individuals for which phenotype measurement was possible. A visual score of 1 indicates severe bleaching, whereas a score of 6 represents full pigmentation.

individuals (four from each of the 12 reefs) at high coverage (>20x for all, 36 samples >100x) (fig. S8 and table S3) to examine the extent of population structure, characterize the demographic history, and scan for signals of adaptation. We mapped reads from all individuals to our de novo genome assembly and called approximately 6.8 million biallelic single-nucleotide polymorphisms (SNPs) after applying stringent quality filters. We validated these filtered SNP calls by independently amplifying and sequencing 10 intergenic regions in eight samples at high coverage (>1400x) (fig. S10 and table S4), observing an overall genotype concordance of 99.6% between regions sequenced

by the two approaches. After an initial principal components analysis (PCA) (fig. S9) and test for outliers, we removed four individuals that were likely the result of sample misidentification at the time of collection (or conceivably represent rare diverged lineages) (fig. S16), bringing the total number of high-coverage resequenced genomes to 44. These 44 genomes are all approximately equidistantly related (fig. S11).

From this set of 44 unrelated individuals, we estimated the decay of pairwise linkage disequilibrium (LD), r^2, as a function of physical distance genome-wide. We observed that on average, r^2 falls below 0.05 after ~15 kb (Fig. 2A). We also estimated the nucleotide diversity (as π, the mean pairwise difference per base pair) in nonoverlapping intergenic regions of 1 kb, which yielded an average of $\pi = 0.363\%$ (Fig. 2B). Assuming a mutation rate of 4×10^{-9} per base pair per generation (*20*, *34*), this value of diversity corresponds to a long-term effective population size (N_e) of around 2.26×10^5. Applying the pairwise sequential Markovian coalescent (PSMC) (*35*) to each genome, we inferred how N_e has changed over the past ~10^6 generations (Fig. 2C). All samples display a highly similar demographic trajectory, with the inferred N_e showing a steady decline. Qualitatively similar results were obtained when demographic histories for the same samples were inferred jointly with the multiple sequentially Markovian coalescent (MSMC) (fig. S12) (*36*). The increase in estimated N_e back in time could reflect greater census sizes in the past or greater population structure in the ancestral population from which all current lineages similarly derive (*35*). A pattern of decline in estimated N_e has also been observed in other corals, including a number of *Acropora* and *Orbicella* species (*20*, *37*, *38*).

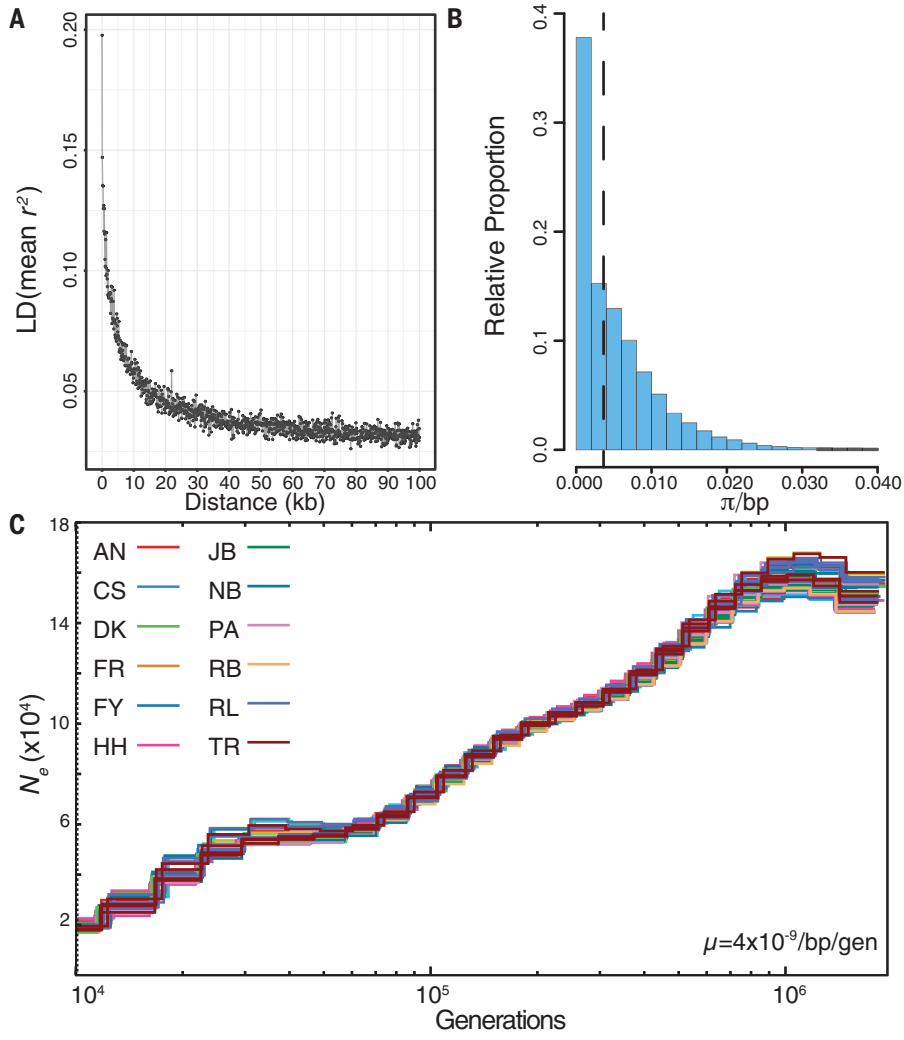


Fig. 2. Properties of genetic variation and inferred demographic history in sampled *A. millepora*. (**A**) The decay of LD, measured as the squared genotypic correlation coefficient (r^2), was estimated for a randomly chosen 1% of SNPs by using all sequenced samples. Each point represents the mean r^2 in bins of 100 bp. (**B**) Nucleotide diversity was measured by the average pairwise differences per base pair (π) (*86*) for SNPs in 1-kb windows, using intergenic regions genome-wide. (**C**) Effective population sizes (N_e) inferred from 44 sequenced *A. millepora* genomes by using PSMC (*35*), assuming a mutation rate μ of 4×10^{-9} per base pair per generation (*20*, *34*). Although an error in the estimate of μ will lead to a rescaling of both axes, it will not affect the qualitative conclusion of a decline in N_e toward the present. The *x* axis is scaled in terms of generations. This approach provides very little information about the recent past (here, less than ~10^4 generations).

The concordant long-term demographic histories inferred across all individuals suggest high connectivity among the 12 sampled reefs. Accordingly, there is no detectable relationship between geographic distance and fixation index (F_{ST}) between pairs of sampled reefs (Fig. 3A). A PCA of SNPs in approximate linkage equilibrium further reveals no obvious clusters (Fig. 3B). In particular, although environmental conditions differ between inshore and offshore reefs (Fig. 1C), F_{ST} between these samples is approximately 0 (-2.76×10^{-4}). More generally, no significant relationships were detected between either of the first two eigenvectors and any of the 40 environmental or spatial variables for each reef, indicating that any population structure that is present is at most very weakly associated with our measured environmental variables. The lack of population structure can be visualized by estimating and plotting relative effective migration rates. Using the program EEMS (*39*), we inferred homogenous migration rates and no obvious barriers to gene flow among sampled reefs (Fig. 3C). This pattern of gene flow across large geographic distances on the central GBR is presumably a result of its broadcast spawning mode of reproduction (*10*, *17*, *19*, *20*, *40*–*42*). These findings are consistent with previous inferences based on the SNP allele frequency spectrum in *A. millepora* (*20*) as well as analyses of microsatellites in *Acropora tenuis* (*43*). However, in other Indo-Pacific Acroporids, such as *Acropora cytherea* and *Acropora hyacinthus*, high genetic divergence, possibly representing cryptic species, has been observed among corals over smaller geographic scales (*44*). Thus, not all broadcast spawning coral species show low population structure at comparable geographic scales.

A scan for adaptation points to a heat-shock protein co-chaperone

The near absence of genetic differentiation across reefs indicates that there is little population structure over hundreds of kilometers. Yet, environmental conditions differ among reefs, notably in terms of thermal regimes (Fig. 1B). Because adults only reproduce with those individuals in their vicinity whereas larvae can disperse over much larger ranges (*16*, *20*), larvae may experience strong selection each generation, as they settle on a heterogeneous

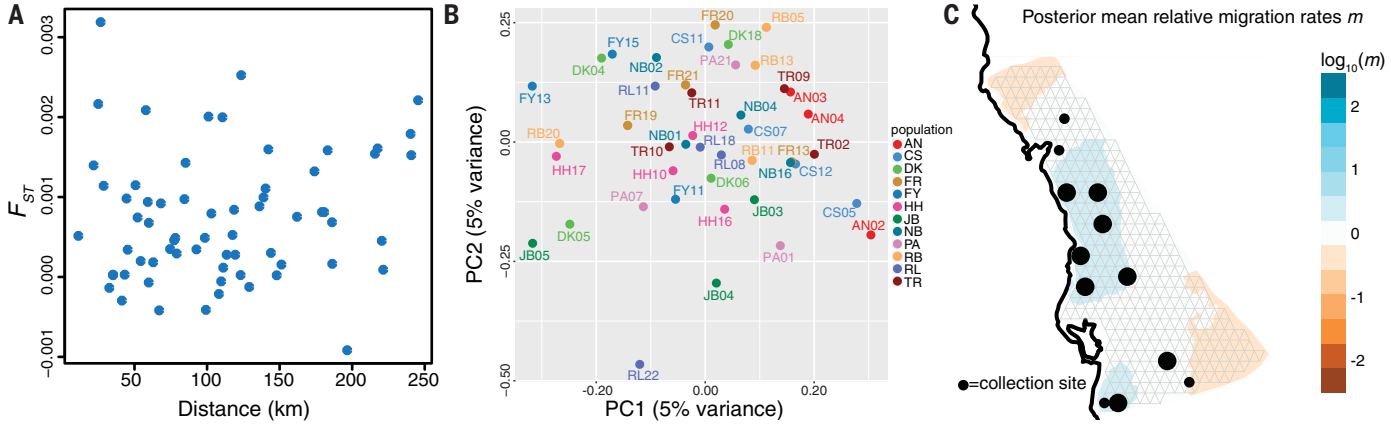


Fig. 3. Characterizing population structure and gene flow across 12 reefs. (**A**) There is no discernable relationship between geographic distance and genetic differentiation (measured as F_{ST}) across pairwise comparisons of sampled reefs. (**B**) A PCA by using LD-pruned genome-wide SNPs for the 44 resequenced high-coverage genomes. Each point is color-coded by the reef from which the individual was sampled. (**C**) EEMS (*39*) was used to estimate and visualize the relative effective migration surface by using LD-pruned and common (MAF > 5%) genome-wide SNPs for the 44 resequenced high-coverage genomes. The size of the dots reflects the number of individuals sequenced from each reef.

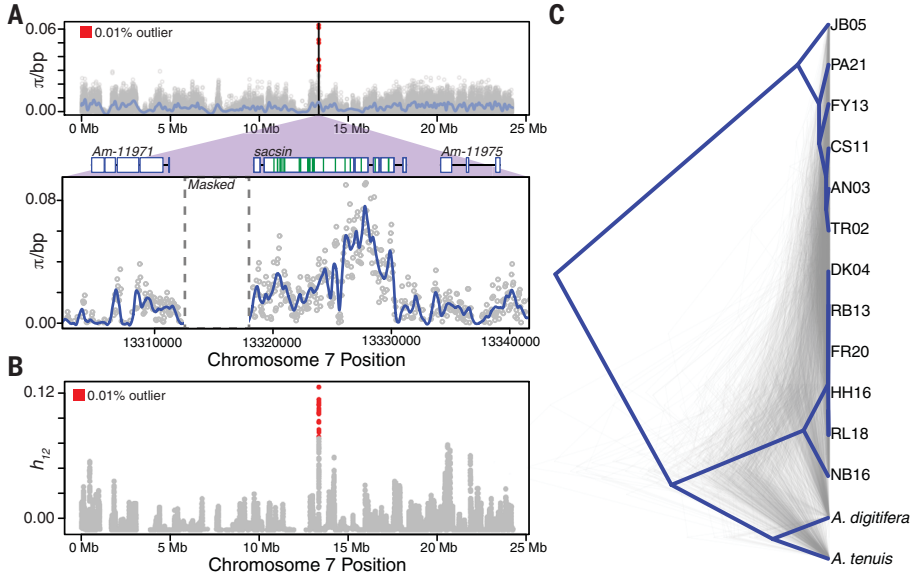


Fig. 4. Genomic scans for local adaptation detect a signal at *sacsin*. (**A**) (Top) Values of pairwise nucleotide diversity (π) (*86*) estimated in sliding windows of 1 kb and a step-size of 500 bp across chromosome 7, with points colored in red indicating regions in the top 0.01% of values genome-wide. A loess-smoothed trend line of π across windows is indicated in purple. The peak with the most extreme values falls in the *sacsin* gene region. (Bottom) A close-up view of π estimated across a region of chromosome 7 surrounding the *sacsin* gene, which includes all of the outlier windows indicated in red at top. π per base pair was calculated in sliding windows of 100 bp (gray dots); the loess-smoothed line is shown in blue. The predicted gene structure of *sacsin* is indicated above, as well as the flanking upstream and downstream genes. The 29 nonsynonymous differences fixed between samples of the two most common haplotypes in *sacsin* are denoted with green lines. There are also 20 fixed synonymous differences. Gray vertical dashed lines delimit a region that was masked from variant calling because of predicted repetitive elements. (**B**) The h_{12} summary statistic (*47*), which measures the frequency of the two most common haplotypes, is plotted across chromosome 7 (supplementary materials, materials and methods). Red dots represent h_{12} values in the top 0.01% genome-wide. The peak with the most extreme values falls in the *sacsin* gene region. (**C**) A gene tree for the central 1-kb region in *sacsin* is shown in dark blue; it was constructed for one randomly chosen haplotype from a randomly selected individual from each of the 12 sampled reefs. The gene tree is rooted to aligned sequences from the reference genomes for *A. digitifera* (*87*) and *A. tenuis* (draft assembly from www.reefgenomics.org). Shown in gray are gene trees for 1000 randomly sampled 1-kb regions from across the genome for the same individuals. Each gene tree was inferred by using maximum likelihood implemented in *dnaml* (*88*).

reefscape. Models suggest that this scenario of strong, spatially varying selection in the face of continual migration can lead to the maintenance of locally beneficial alleles over long periods of time (*45*, *46*). We therefore searched for signals of adaptation in the 44 resequenced genomes by scanning for genomic regions displaying the high levels of diversity that would result from such long-lived balancing selection. Genome-wide, the most extreme values (π > 5.0%, measured in 1-kb windows) (Fig. 4A) fell over a region containing a single gene, which is orthologous to *sacsin* in *Acropora digitifera*. The same genic region also has a highly elevated value of the h_{12} statistic (*47*), which measures the frequencies of the two most common haplotypes (Fig. 4B). Confirming that this pattern of relatedness is unusual, a PCA for the region stands out by comparison with the rest of the genome [as assessed with *lostruct* (*48*)] (fig. S13). These summary statistics point to an unusually deep genealogy with long internal branches at *sacsin*; the divergence of two haplotypes even predates the species split with *A. tenuis* and *A. digitifera* (Fig. 4C and fig. S15). There was no significant elevation in coverage surrounding *sacsin* for any sample to suggest the presence of a paralog (fig. S17). To further ensure that these findings were not artifactual, we confirmed the presence of the two diverged

haplotypes by independently sequencing and assembling overlapping 2-kb amplicons spanning the entire gene for a subset of four samples, which include homozygotes for both alternative haplotypes and a heterozygote (fig. S18 and tables S6 and S7). In a de novo assembly of sequences from both haplotypes across the coding region, each amplicon aligned in the expected order (fig. S19).

Sacsin is a co-chaperone for the heat-shock protein Hsp70 and contains multiple regions of homology to *Hsp90* (*49*, *50*). *Sacsin* has been shown to be significantly up-regulated when exposed to elevated temperatures in the laboratory in multiple coral species from the genus *Acropora* and *Pocillipora* (*51*–*54*), indicating that variation in the gene is plausibly associated with heat response in natural populations. Our analyses reveal that variation in *sacsin* has been maintained for an atypically long time (millions of generations), indicating that the gene is a target of balancing selection and hence of functional importance in the wild as well as in the laboratory. The source of balancing selection is unclear, however; there is no discernable association of the haplotype frequencies with genetic or current environmental principal components ($P > 0.15$) (table S5), although there may be an association with the dominant symbiont type ($P = 0.059$).

A GWAS for bleaching

Using the 44 high-coverage genomes, we used read-backed phasing (*55*) to construct a reference haplotype panel and impute genotypes in low-pass sequences of additional samples. We generated a mean of ~1.5x coverage whole-genome data (*56*) for 193 individuals after applying filters based on quality, read depth, and the proportion of missing data (fig. S21). Of these samples, 34 had also been sequenced at high coverage, allowing us to assess the accuracy of our imputation approach. In the 34 samples, the overall correlation between sequenced and imputed genotypes was >94% for SNPs with minor allele frequency >0.05 and genotype probability ≥0.95 (fig. S22), an imputation accuracy similar to what is observed in human data that use reference panels of similar size (*57*, *58*).

Combining the low-coverage genomes with the 44 high-coverage ones, we obtained genotype calls at ~6.8 million sites in a total of 237 total samples. After excluding outliers that were likely species misidentifications or possibly rare diverged lineages (*59*), individuals with high levels of relatedness, and those with missing phenotype data (figs. S23 to S24), our final sample size was 213. Consistent with our previous results, in the larger sample we again observed minimal population structure, evidence for high gene flow, and a peak of elevated genetic diversity at *sacsin* (fig. S25). Although the low-population structure is helpful for a GWAS, it makes precise estimates of SNP heritability unattainable without substantially larger sample sizes (*60*, *61*).

The sequencing also yielded data from the symbiont Symbiodiniaceae, of which multiple species are known to associate with *A. millepora* (*62*, *63*). Because intracellular symbiont DNA was extracted and sequenced simultaneously with the coral host, we used our genomic sequencing approach to characterize the symbiont species present in each sample (*11*, *59*). The abundance of symbiont reads in each sample was significantly different between quartiles of visual scores ($P = 2.6 \times 10^{-7}$ by means of a Kruskal-Wallis test) (Fig. 5A) as well as correlated with total chlorophyll content and symbiont cell density (fig. S27). We further determined the composition of symbiont types for each sample by calculating the relative proportion of reads that map to available draft genomes. In our samples, individuals dominated by *Durusdinium* had significantly higher visual scores on average ($P = 2.46 \times 10^{-7}$ by means of a Mann-Whitney *U* test) (Fig. 5B), which is consistent with previous findings that *Durusdinium* (formerly *Symbiodinium* Clade D) is associated with substantially higher thermal tolerance (but often lower growth rates) across diverse coral genera in the wild (*11*, *64*–*66*). Our results demonstrate that a low-coverage sequencing approach can be used to simultaneously obtain a quantitative measure of bleaching and determine the relative composition of symbiont types.

Given these data from corals and their symbionts, we used a linear mixed model (LMM) to test for additive effects of SNPs (with minor allele frequencies >5%) on the quantile-normalized visual score (*67*, *68*), including as covariates the top four principal components of environmental variables (which together account for >85% of the total environmental variance), the first two genetic PCs, batch effects (the collection date and sequencing batch), the collection depth (*69*), and the proportion of *Durusdinium* reads relative to all symbiont reads. We also conducted similar GWASs for total chlorophyll content and symbiont cell density (fig. S28). The three measures of bleaching are highly correlated (Spearman's correlation coefficient among pairs of phenotypes >0.6) (fig. S26); because we are missing phenotype data for chlorophyll content (so that $n = 190$ samples) and symbiont cell density ($n = 172$ samples), we focused on the standard visual score ($n = 213$ samples).

We determined a genome-wide significance threshold of $P = 4.3 \times 10^{-8}$ for the quantile-normalized visual score by considering the distribution of *P* values obtained from 10^5 permutation tests (fig. S34). With this approach, no single SNP in our GWAS is genome-wide significant, indicating the absence of common, major effect loci associated with variation in bleaching. In particular, the minimum *P* value in the *sacsin* gene is high (0.01); for related phenotypes of total chlorophyll content and symbiont cell density, however, SNPs in *sacsin* are among the top signals (minimal *P* values are 5.05×10^{-5} and 2.38×10^{-5}, respectively),

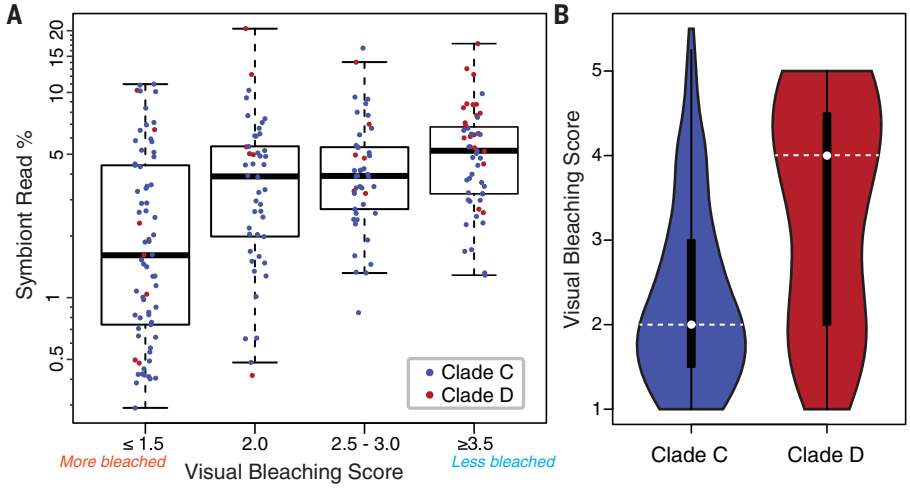


Fig. 5. Relationship between symbiont species and bleaching response revealed with high-throughput sequencing. (**A**) The total proportion of reads mapping to draft symbiont genomes was measured for each sample. Bins for the visual scores were chosen so that approximately equal numbers of samples were included in each. Points are color-coded according to the dominant symbiont genus found in each sample. The *y* axis is shown on the $\log_{10}$ scale. (**B**) Samples dominated by *Durusdinium* (formerly *Symbiodinium* clade D) have a significantly higher ($P = 2.46 \times 10^{-7}$ by means of a two-tailed Mann-Whitney *U* test) visual score as compared with samples dominated by *Cladocopium* (formerly *Symbiodinium* clade C). White dotted lines in each violin plot represent the median.

suggesting that the lack of signal in the visual score GWAS may reflect lack of power. For visual score, the top peaks were a region encompassing two genes on chromosome 13 and one with three genes on chromosome 14 [also detected with the software BIMBAM (fig. S33) (70)] (Fig. 6A). These associations remain to be confirmed, but among the three genes on chromosome 14 is malate synthase, a key component of the glyoxylate cycle, a pathway not found in metazoans outside of cnidarians (71). Malate synthase is up-regulated in response to thermal stress in *A. palmata* and implicated in the coral stress response (72, 73).

Predicting individual responses in natural populations

The GWAS results in 213 diploid individuals indicate that variation in bleaching is not due to common loci of large effect but instead is polygenic, as has also been suggested for thermal tolerance in *A. hyacinthus* (22), so that our study is underpowered to identify individual causal loci. Nonetheless, when combined, the estimated effects of variants across the genome may be predictive of phenotype (25). We therefore constructed a polygenic score [PGS; also called breeding value (23)] using LD-based clumping (for a *P* value threshold of 10^{-5}) (fig. S35) and the effect sizes estimated in the GWAS (74). Because we lacked a true out-of-sample validation set, we assessed the prediction accuracy of the PGS using a jackknife cross-validation procedure with an 85/15% training-test split in each partition; similar results were also obtained for 80/20% and 90/10% training-test splits (fig. S36).

Given that both environmental and genetic factors contribute to variation in bleaching, we combined our PGS with other predictors, including environmental effects and the dominant symbiont species type. Using a cross-validation procedure, we assessed the improvement in prediction accuracy gained from including these predictors in a linear model (Fig. 6B). Compared with a model including only batch effects and genetic principal components, we observed a large increase in prediction accuracy using four principal components for the environmental variables as well as a substantial improvement when we included data obtained from our genomic sequencing—namely, the proportion of *Durusdinium* reads out of all symbiont reads and the PGS (Fig. 6B). Including the PGS alone provided a small but significant increase in prediction accuracy [$P = 0.023$; mean incremental coefficient of determination (R^2) = 0.026]; the same was true in models that used the top six or eight environmental PCs instead (which in total explain 99% of the variance in environmental variables; $P \leq 0.029$) (fig. S37). In total, the model explains an estimated ~62% of the variance in quantile-normalized visual scores in our sample. However, moving forward it will be important to test its performance on a distinct sample rather than using cross-validation as we have of necessity done here.

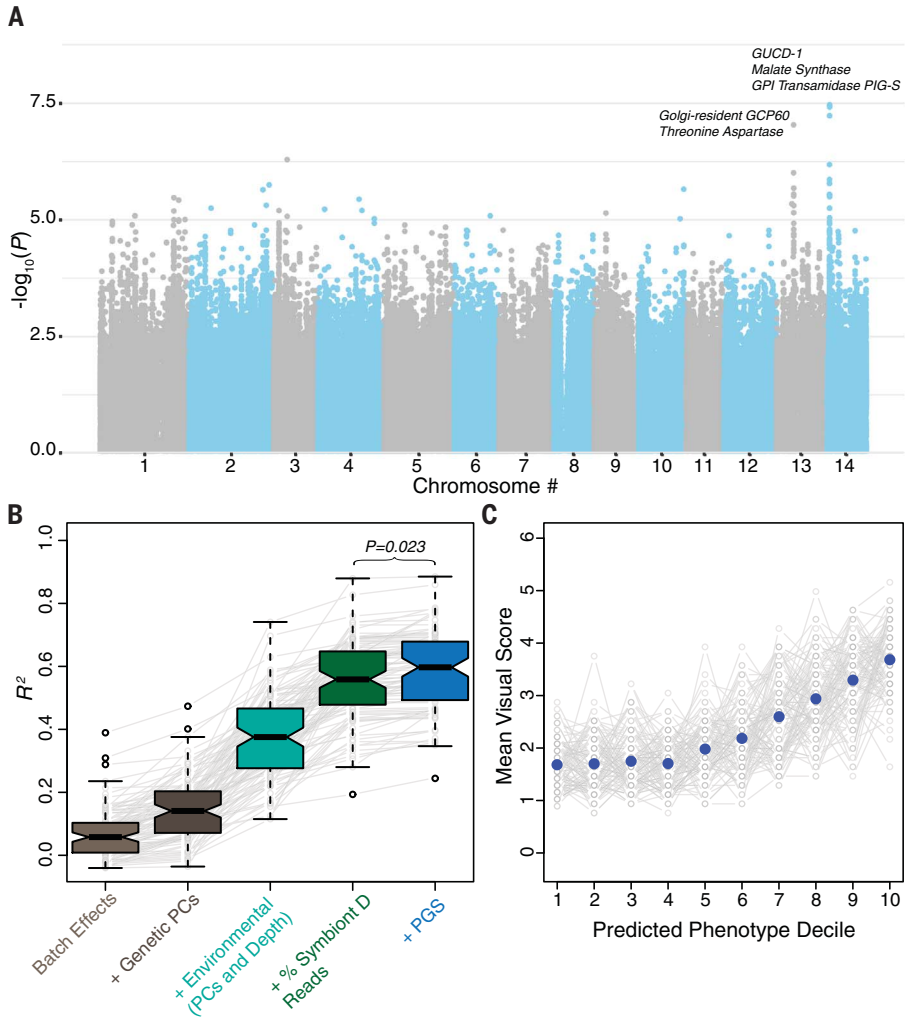


Fig. 6. A GWAS and prediction accuracy for bleaching response. (**A**) A Manhattan plot of *P* values obtained from testing for an association of each SNP with variation in visual score. Shown are the names of all genes in the two top peaks, on chromosome 13 and 14. (**B**) The increase in prediction accuracy (measured as R^2) for linear models with different sets of predictors. Each boxplot shows the distribution of R^2 obtained for 100 jackknife cross-validation partitions, with a random 15% of the total sample withheld as a test set in each. Batch effects include the collection date and the sequencing batch. Each boxplot shows the R^2 for sets of predictors in a linear model. Mann-Whitney *U* tests were performed to test for differences in the distribution of R^2 between models. The average number of SNPs used in the PGS in each jackknife partition is 15.32. (**C**) Using the full predictive model, individuals with the highest predicted trait values have the highest visual scores on average. Within each jackknife partition, the test set was divided into deciles according to the predicted phenotype. The blue dot for each decile indicates the mean visual score, and gray lines connect deciles in a given jackknife partition.

Implications

Understanding how species will respond to increasing temperatures is key to ensuring adequate protection through conventional management approaches and to supporting interventions aimed at restoration or facilitating adaptation to prevent future losses. In this work, we focused on bleaching in the coral *A. millepora*, a trait of central ecological importance, demonstrating the feasibility of predicting interindividual response from genomic and environmental data. The lack of genome-wide significant associations suggests that the differences in bleaching response are due to loci that individually explain only a small proportion of the variance. As has been the case in other species, we therefore expect that as sample sizes increase, so will the prediction accuracy of the PGS and hence of the model. With larger samples, it will also become possible to test for interactions between the coral genome and environmental factors or symbiont types

as well as to estimate the heritability of bleaching in the wild and hence learn more about the potential adaptive response.

In the development of effective conservation strategies, a predictive model of bleaching can help distinguish those individuals more likely to be tolerant on reefs across different shelf positions, latitudes, and environmental conditions, informing current spatial protection strategies and predictions about reef futures (*75*, *76*). As an illustration, using our cross-validation procedure and full predictive model, individuals in the top quartile of predicted phenotype values have significantly higher visual scores on average than those in the lowest quartile ($P = 8.8 \times 10^{-16}$ by means of a one-tailed Wilcoxon signed rank test) (Fig. 6C). The approach may also help to identify the most tolerant individuals to use for managed translocations and selective breeding for assisted evolution–based conservation strategies (*77*). In that regard, the lack of population structure in *A. millepora* over large geographic distances makes it an ideal system because one could conceivably breed corals with high predicted trait values from multiple locations. Alternatively, such predictions could be used to increase the abundance of such corals in local populations rather than moving them through assisted gene flow, which carries risks and ethical concerns.

Beyond the application to corals, our study demonstrates the feasibility of high-quality genome assembly, imputation, and GWASs in a nonmodel organism. Moreover, our work highlights the potential of combining genomic and environmental data to predict phenotypes of ecological importance—notably, traits that are likely to play a role in the response to climate change.

Materials and methods

Genome assembly and annotation

The de novo genome assembly of *A. millepora* was constructed by using a combination of PacBio reads and Illumina paired-end reads with 10X Genomics Chromium barcodes. Sequencing was performed by the New York Genome Center and Cold Spring Harbor Laboratory. We constructed PacBio-only assemblies using the software Canu (*78*) and Falcon (*79*) and used the 10X barcodes and paired-end reads for scaffolding and error correction. The high–molecular weight DNA used for the assembly was extracted from an adult specimen, which also contained symbiont DNA. To remove potential symbiont contigs from the coral genome assembly, we additionally generated paired-end reads from two pools of aposymbiotic larvae and mapped these reads to our draft assembly. Using an unsupervised clustering approach, we identified and removed assembled contigs that appeared to be derived from the symbiont and masked regions of the assembly with no reads from the larvae. To assemble chromosome-scale scaffolds, we used previously published linkage maps (*19*, *80*) inferred from SNP and restriction site–associated DNA sequencing (RAD-seq) markers to linearly arrange contigs. Previously published transcriptome data (*27*, *28*) were used to predict gene models and annotate the resulting assembly.

Sample collection and phenotyping

Coral samples were collected from 12 reefs in the central region of the GBR at the height of mass bleaching between 14 March and 1 April 2017. For each colony, the bleaching score was estimated visually by using the Coral Color Reference Card (*33*) to the nearest color, including half increments. During collection dives, samples were sent to the surface every 5 min and immediately fixed in liquid nitrogen for further analysis. The total chlorophyll content was estimated in three technical replicates by measuring absorbance at 665, 649, and 632 nm on a microplate reader (synergy H4, BioTek Instruments, USA). The total coral host protein was similarly measured in triplicate by using the DC Protein Assay (Bio-Rad Laboratories, USA) following the manufacturers microplate protocol. Ecological conditions at each reef were measured with environmental and spatial variables recorded at a resolution of 1 km^2 (*31*).

Whole-genome sequencing, variant calling, and genotype imputation

The libraries for samples sequenced at high coverage were prepared by using the Illumina TruSeq polymerase chain reaction (PCR)–free protocol and sequenced on an Illumina HiSeqX platform. Libraries for samples sequenced at low coverage were prepared for multiplexed shotgun genotyping with Tn5 transposase and tagmentation (*56*). The Tn5 libraries were sequenced on an Illumina HiSeq4000 platform across 10 lanes. Illumina adapters and polyadenylate [poly(A)] sequences were trimmed (*81*) and low-quality reads were removed before mapping to our *A. millepora* genome assembly with *bwa-mem* (*82*). Variants were called by using Genome Analysis Toolkit (GATK) and filtered according to GATK best practices (*83*). The algorithm we used for genotype imputation in the low-coverage sequencing data are based on the copying model of Li and Stephens (*84*) and implemented in the software *loimpute* developed by Gencove and available at www.gencove.com.

Population genetic analyses

We used the software PLINK (*74*) to estimate the decay of pairwise LD as a function of physical distance, to perform PCA, and to estimate F_{ST}. To estimate nucleotide diversity, we used a custom wrapper for vcftools (*85*) to take into account missing data and uncalled sites. We inferred historical changes in N_e using the software PSMC (*35*), for which we increased the stringency of the quality filter to 30 and used a bin size of 10 base pairs (bp). We used EEMS (*39*) to visualize effective migration rates across the 12 sampled reefs, setting the number of MCMC iterations to 5×10^6. The software lostruct (*48*) was used to identify localized genomic regions in which patterns of relatedness differ from what is typical for the genome, under default parameters. We used the read-aware phasing module implemented in SHAPEIT2 (*55*) to infer haplotypes in the high-coverage genomes. To identify localized genomic regions with high frequencies of the two most common haplotypes, we calculated the h_{12} summary statistic in window sizes of 100 SNPs using a custom Python script (*47*).

Genome-wide association tests and construction of a PGS

We performed GWASs for bleaching measured as either a visual score, a total chlorophyll content, or a symbiont cell density. For each GWAS, we removed samples with missing phenotype data. The set of covariates included in all GWAS models were the top four PCs of environmental and spatial variables, the first two genetic PCs, the collection date, the sequencing batch, collection depth, and the relative proportion of *Durusdinium* reads (out of all symbiont reads). All GWASs were performed by using standard linear regression implemented in PLINK (*74*), and phenotypes were quantile-normalized. Only common variants [minor allele frequency (MAF) ≥ 0.05] and sites with less than 10% of genotypes missing were considered. We additionally performed a GWAS for visual score using a LMM implemented in GEMMA (*67*, *68*) and in a Bayesian framework using the software BIMBAM (*70*). We set cutoffs for genome-wide significance for each phenotype by performing permutation tests. Specifically, we shuffled the values of each phenotype across samples 10,000 times; for each permutation, we performed a standard linear GWAS using the same covariates and assessed the minimal P value. We then set the genome-wide significance threshold as the 95th percentile of the distribution of these minimal permuted P values.

Because we lacked a true out-of-sample validation set to assess the accuracy of PGSs constructed from GWAS, we used a jackknife cross-validation (CV) procedure to subset the samples into 100 partitions of training and test sets. In each partition, we randomly withheld 15% of samples as the test set and selected the other 85% of individuals as the training set. We recalculated the genetic PCs for the training set and used the loadings from this PCA for the genetic PCs in the test set. For each jackknife partition, a standard GWAS for quantile-normalized visual score was performed on the training set. To build the PGS, we used a

P value threshold of 1×10^{-5} followed by LD-clumping to choose sets of approximately independent SNPs. For each set of selected SNPs, the PGS was calculated by summing the allelic dosages weighted by their estimated effect size. To assess prediction accuracy, a PGS was then estimated in the test set. We assessed the prediction accuracy of the PGS along with other factors that may influence bleaching (such as environmental principal components, depth, and proportion of *Durusdinium*) in linear models. Mann-Whitney *U* tests were performed to test for differences in predicted trait values across jackknife partitions across different sets of predictors.

REFERENCES AND NOTES

1. C. Bellard, C. Bertelsmeier, P. Leadley, W. Thuiller, F. Courchamp, Impacts of climate change on the future of biodiversity. *Ecol. Lett.* **15**, 365–377 (2012). doi: 10.1111/j.1461-0248.2011.01736.x; pmid: 22257223
2. C. Parmesan, G. Yohe, A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003). doi: 10.1038/nature01286; pmid: 12511946
3. S. A. Henson *et al*., Rapid emergence of climate change in environmental drivers of marine ecosystems. *Nat. Commun.* **8**, 14682 (2017). doi: 10.1038/ncomms14682; pmid: 28267144
4. M. Pacifici *et al*., Assessing species vulnerability to climate change. *Nat. Clim. Chang.* **5**, 215–224 (2015). doi: 10.1038/nclimate2448
5. A. A. Hoffmann, C. M. Sgrò, Climate change and evolutionary adaptation. *Nature* **470**, 479–485 (2011). doi: 10.1038/nature09670; pmid: 21350480
6. I.-C. Chen, J. K. Hill, R. Ohlemüller, D. B. Roy, C. D. Thomas, Rapid range shifts of species associated with high levels of climate warming. *Science* **333**, 1024–1026 (2011). doi: 10.1126/science.1206432; pmid: 21852500
7. G.-R. Walther *et al*., Ecological responses to recent climate change. *Nature* **416**, 389–395 (2002). doi: 10.1038/416389a; pmid: 11919621
8. J. Merilä, A. P. Hendry, Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evol. Appl.* **7**, 1–14 (2014). doi: 10.1111/eva.12137; pmid: 24454544
9. M. Kelly, Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philos. Trans. R. Soc. London B Biol. Sci.* **374**, 20180176 (2019). doi: 10.1098/rstb.2018.0176; pmid: 30966963
10. D. J. Ayre, T. P. Hughes, Climate change, genotypic diversity and gene flow in reef-building corals. *Ecol. Lett.* **7**, 273–278 (2004). doi: 10.1111/j.1461-0248.2004.00585.x
11. T. C. LaJeunesse *et al*., Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr. Biol.* **28**, 2570–2580.e6 (2018). doi: 10.1016/j.cub.2018.07.008; pmid: 30100341
12. N. Traylor-Knowles, Heat stress compromises epithelial integrity in the coral, *Acropora hyacinthus*. *PeerJ* **7**, e6510 (2019). doi: 10.7717/peerj.6510; pmid: 30828497
13. T. P. Hughes *et al*., Global warming and recurrent mass bleaching of corals. *Nature* **543**, 373–377 (2017). doi: 10.1038/nature21707; pmid: 28300113
14. T. P. Hughes *et al*., Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science* **359**, 80–83 (2018). doi: 10.1126/science.aan8048; pmid: 29302011
15. T. P. Hughes *et al*., Global warming transforms coral reef assemblages. *Nature* **556**, 492–496 (2018). doi: 10.1038/s41586-018-0041-2; pmid: 29670282
16. D. J. Ayre, T. P. Hughes, Genotypic diversity and gene flow in brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* **54**, 1590–1605 (2000). doi: 10.1111/j.0014-3820.2000.tb00704.x; pmid: 11108587
17. M. J. H. Van Oppen, L. M. Peplow, S. Kininmonth, R. Berkelmans, Historical and contemporary factors shape the population genetic structure of the broadcast spawning coral, *Acropora millepora*, on the Great Barrier Reef. *Mol. Ecol.* **20**, 4899–4914 (2011). doi: 10.1111/j.1365-294X.2011.05328.x; pmid: 22026459
18. J. R. Guest *et al*., Coral community response to bleaching on a highly disturbed reef. *Sci. Rep.* **6**, 20717 (2016). doi: 10.1038/srep20717; pmid: 26876092
19. G. B. Dixon *et al*., CORAL REEFS. Genomic determinants of coral heat tolerance across latitudes. *Science* **348**, 1460–1462 (2015). doi: 10.1126/science.1261224; pmid: 26113720
20. M. V. Matz, E. A. Treml, G. V. Aglyamova, L. K. Bay, Potential and limits for rapid genetic adaptation to warming in a Great Barrier Reef coral. *PLOS Genet.* **14**, e1007220 (2018). doi: 10.1371/journal.pgen.1007220; pmid: 29672529
21. K. E. Dziedzic, H. Elder, H. Tavalire, E. Meyer, Heritable variation in bleaching responses and its functional genomic basis in reef-building corals (*Orbicella faveolata*). *Mol. Ecol.* **28**, 2238–2253 (2019). doi: 10.1111/mec.15081; pmid: 30913323
22. N. H. Rose, R. A. Bay, M. K. Morikawa, S. R. Palumbi, Polygenic evolution drives species divergence and climate adaptation in corals. *Evolution* **72**, 82–94 (2018). doi: 10.1111/evo.13385; pmid: 29098686
23. D. S. Falconer, T. F. C. Mackay, *Introduction to Quantitative Genetics* (Pearson, ed. 4, 1996).
24. N. R. Wray, K. E. Kemper, B. J. Hayes, M. E. Goddard, P. M. Visscher, Complex trait prediction from genome data: Contrasting EBV in livestock to PRS in humans. *Genetics* **211**, 1131–1141 (2019). doi: 10.1534/genetics.119.301859; pmid: 30967442
25. E. S. Buckler *et al*., The genetic architecture of maize flowering time. *Science* **325**, 714–718 (2009). doi: 10.1126/science.1174276; pmid: 19661422
26. A. V. Khera *et al*., Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018). doi: 10.1038/s41588-018-0183-z; pmid: 30104762
27. A. Moya *et al*., Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO_2-driven acidification during the initiation of calcification. *Mol. Ecol.* **21**, 2440–2454 (2012). doi: 10.1111/j.1365-294X.2012.05554.x; pmid: 22490231
28. D. Bhattacharya *et al*., Comparative genomics explains the evolutionary success of reef-forming corals. *eLife* **5**, e13288 (2016). doi: 10.7554/eLife.13288; pmid: 27218454
29. J. Benevenuto, L. F. V. Ferrão, R. R. Amadeu, P. Munoz, How can a high-quality genome assembly help plant breeders? *Gigascience* **8**, giz068 (2019). doi: 10.1093/gigascience/giz068; pmid: 31184361
30. K. Nadachowska-Brzyska, R. Burri, L. Smeds, H. Ellegren, PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. *Mol. Ecol.* **25**, 1058–1072 (2016). doi: 10.1111/mec.13540; pmid: 26797914
31. S. A. Matthews *et al*., High-resolution characterization of the abiotic environment and disturbance regimes on the Great Barrier Reef, 1985-2017. *Ecology* **100**, e02574 (2019). doi: 10.1002/ecy.2574; pmid: 30645776
32. S. DiPerna, M. Hoogenboom, S. Noonan, K. Fabricius, Effects of variability in daily light integrals on the photophysiology of the corals *Pachyseris speciosa* and *Acropora millepora*. *PLOS ONE* **13**, e0203882 (2018). doi: 10.1371/journal.pone.0203882; pmid: 30240397
33. U. E. Siebeck, N. J. Marshall, A. Klüter, O. Hoegh-Guldberg, Monitoring coral bleaching using a colour reference card. *Coral Reefs* **25**, 453–460 (2006). doi: 10.1007/s00338-006-0123-8
34. Z. T. Richards, D. J. Miller, C. C. Wallace, Molecular phylogenetics of geographically restricted *Acropora* species: Implications for threatened species conservation. *Mol. Phylogenet. Evol.* **69**, 837–851 (2013). doi: 10.1016/j.ympev.2013.06.020; pmid: 23850500
35. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011). doi: 10.1038/nature10231; pmid: 21753753
36. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014). doi: 10.1038/ng.3015; pmid: 24952747
37. Y. Mao, E. P. Economo, N. Satoh, The roles of introgression and climate change in the rise to dominance of *Acropora* corals. *Curr. Biol.* **28**, 3373–3382.e5 (2018). doi: 10.1016/j.cub.2018.08.061; pmid: 30344117
38. C. Prada *et al*., Empty niches after extinctions increase population sizes of modern corals. *Curr. Biol.* **26**, 3190–3194 (2016). doi: 10.1016/j.cub.2016.09.039; pmid: 27866895
39. D. Petkova, J. Novembre, M. Stephens, Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* **48**, 94–100 (2016). doi: 10.1038/ng.3464; pmid: 26642242
40. P. Souter *et al*., Location and disturbance affect population genetic structure in four coral species of the genus *Acropora* on the Great Barrier Reef. *Mar. Ecol. Prog. Ser.* **416**, 35–45 (2010). doi: 10.3354/meps08740
41. S. W. Davies, E. A. Treml, C. D. Kenkel, M. V. Matz, Exploring the role of Micronesian islands in the maintenance of coral genetic diversity in the Pacific Ocean. *Mol. Ecol.* **24**, 70–82 (2015). doi: 10.1111/mec.13005; pmid: 25407355
42. W. Huang *et al*., Genetic diversity and large-scale connectivity of the scleractinian coral *Porites lutea* in the South China Sea. *Coral Reefs* **37**, 1259–1271 (2018). doi: 10.1007/s00338-018-1724-8
43. V. Lukoschek, C. Riginos, M. J. H. van Oppen, Congruent patterns of connectivity can inform management for broadcast spawning corals on the Great Barrier Reef. *Mol. Ecol.* **25**, 3065–3080 (2016). doi: 10.1111/mec.13649; pmid: 27085309
44. J. T. Ladner, S. R. Palumbi, Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Mol. Ecol.* **21**, 2224–2238 (2012). doi: 10.1111/j.1365-294X.2012.05528.x; pmid: 22439812
45. H. Levene, Genetic Equilibrium when more than one ecological niche is available. *Am. Nat.* **87**, 331–333 (1953). doi: 10.1086/281792
46. S. Yeaman, S. P. Otto, Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* **65**, 2123–2129 (2011). doi: 10.1111/j.1558-5646.2011.01277.x; pmid: 21729066
47. N. R. Garud, P. W. Messer, E. O. Buzbas, D. A. Petrov, Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLOS Genet.* **11**, e1005004 (2015). doi: 10.1371/journal.pgen.1005004; pmid: 25706129
48. H. Li, P. Ralph, Local PCA shows how the effect of population structure differs along the genome. *Genetics* **211**, 289–304 (2019). doi: 10.1534/genetics.118.301747; pmid: 30459280
49. D. A. Parfitt *et al*., The ataxia protein sacsin is a functional co-chaperone that protects against polyglutamine-expanded ataxin-1. *Hum. Mol. Genet.* **18**, 1556–1565 (2009). doi: 10.1093/hmg/ddp067; pmid: 19208651
50. M. Ménade *et al*., Structures of ubiquitin-like (Ubl) and Hsp90-like domains of sacsin provide insight into pathological mutations. *J. Biol. Chem.* **293**, 12832–12842 (2018). doi: 10.1074/jbc.RA118.003939; pmid: 29945973
51. R. Cunning, R. A. Bay, P. Gillette, A. C. Baker, N. Traylor-Knowles, Comparative analysis of the *Pocillopora damicornis* genome highlights role of immune system in coral evolution. *Sci. Rep.* **8**, 16134 (2018). doi: 10.1038/s41598-018-34459-8; pmid: 30382153
52. E. M. Hemond, S. T. Kaluziak, S. V. Vollmer, The genetics of colony form and function in Caribbean *Acropora* corals. *BMC Genomics* **15**, 1133 (2014). doi: 10.1186/1471-2164-15-1133; pmid: 25519925
53. R. Cunning, A. C. Baker, Not just who, but how many: The importance of partner abundance in reef coral symbioses. *Front. Microbiol.* **5**, 400 (2014). doi: 10.3389/fmicb.2014.00400; pmid: 25136339
54. A. B. Mayfield, Y.-J. Chen, C.-Y. Lu, C.-S. Chen, The proteomic response of the reef coral *Pocillopora acuta* to experimentally elevated temperatures. *PLOS ONE* **13**, e0192001 (2018). doi: 10.1371/journal.pone.0192001; pmid: 29385204
55. O. Delaneau, B. Howie, A. J. Cox, J.-F. Zagury, J. Marchini, Haplotype estimation using sequencing reads. *Am. J. Hum. Genet.* **93**, 687–696 (2013). doi: 10.1016/j.ajhg.2013.09.002; pmid: 24094745
56. S. Picelli *et al*., Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014). doi: 10.1101/gr.177881.114; pmid: 25079858
57. L. Huang *et al*., Genotype-imputation accuracy across worldwide human populations. *Am. J. Hum. Genet.* **84**, 235–250 (2009). doi: 10.1016/j.ajhg.2009.01.013; pmid: 19215730
58. B. L. Browning, Y. Zhou, S. R. Browning, A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* **103**, 338–348 (2018). doi: 10.1016/j.ajhg.2018.07.015; pmid: 30100085
59. D. P. Manzello *et al*., Role of host genetics and heat-tolerant algal symbionts in sustaining populations of the endangered coral *Orbicella faveolata* in the Florida Keys with ocean warming. *Glob. Change Biol.* **25**, 1016–1031 (2019). doi: 10.1111/gcb.14545; pmid: 30552831
60. P. M. Visscher *et al*., Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLOS Genet.* **10**, e1004269 (2014). doi: 10.1371/journal.pgen.1004269; pmid: 24721987

61. P. M. Visscher *et al*., Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLOS Genet.* **2**, e41 (2006). doi: 10.1371/journal.pgen.0020041; pmid: 16565746
62. J. E. Parkinson, I. B. Baums, The extended phenotypes of marine symbioses: Ecological and evolutionary consequences of intraspecific genetic diversity in coral-algal associations. *Front. Microbiol.* **5**, 445 (2014). doi: 10.3389/fmicb.2014.00445; pmid: 25202306
63. T. F. Cooper *et al*., Environmental factors controlling the distribution of symbiodinium harboured by the coral *Acropora millepora* on the Great Barrier Reef. *PLOS ONE* **6**, e25536 (2011). doi: 10.1371/journal.pone.0025536; pmid: 22065989
64. R. Berkelmans, M. J. H. van Oppen, The role of zooxanthellae in the thermal tolerance of corals: A 'nugget of hope' for coral reefs in an era of climate change. *Proc. Biol. Sci.* **273**, 2305–2312 (2006). doi: 10.1098/rspb.2006.3567; pmid: 16928632
65. M. K. Morikawa, S. R. Palumbi, Using naturally occurring climate resilient corals to construct bleaching-resistant nurseries. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10586–10591 (2019). doi: 10.1073/pnas.1721415116; pmid: 31061118
66. A. Klueter, J. Trapani, F. I. Archer, S. E. McIlroy, M. A. Coffroth, Comparative growth rates of cultured marine dinoflagellates in the genus *Symbiodinium* and the effects of temperature and light. *PLOS ONE* **12**, e0187707 (2017). doi: 10.1371/journal.pone.0187707; pmid: 29186143
67. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012). doi: 10.1038/ng.2310; pmid: 22706312
68. X. Zhou, A unified framework for variance component estimation with summary statistics in genome-wide association studies. *Ann. Appl. Stat.* **11**, 2027–2051 (2017). doi: 10.1214/17-AOAS1052; pmid: 29515717
69. P. R. Muir, P. A. Marshall, A. Abdulla, J. D. Aguirre, Species identity and depth predict bleaching severity in reef-building corals: Shall the deep inherit the reef? *Proc. Biol. Sci.* **284**, 20171551 (2017). doi: 10.1098/rspb.2017.1551; pmid: 29021175
70. B. Servin, M. Stephens, Imputation-based analysis of association studies: Candidate regions and quantitative traits. *PLOS Genet.* **3**, e114 (2007). doi: 10.1371/journal.pgen.0030114; pmid: 17676998
71. F. A. Kondrashov, E. V. Koonin, I. G. Morgunov, T. V. Finogenova, M. N. Kondrashova, Evolution of glyoxylate cycle enzymes in Metazoa: Evidence of multiple horizontal transfer events and pseudogene formation. *Biol. Direct* **1**, 31 (2006). doi: 10.1186/1745-6150-1-31; pmid: 17059607
72. N. R. Polato, N. S. Altman, I. B. Baums, Variation in the transcriptional response of threatened coral larvae to elevated temperatures. *Mol. Ecol.* **22**, 1366–1382 (2013). doi: 10.1111/mec.12163; pmid: 23331636
73. R. M. Wright, G. V. Aglyamova, E. Meyer, M. V. Matz, Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*. *BMC Genomics* **16**, 371 (2015). doi: 10.1186/s12864-015-1540-2; pmid: 25956907
74. S. Purcell *et al*., PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007). doi: 10.1086/519795; pmid: 17701901
75. I. B. Baums, A restoration genetics guide for coral reef conservation. *Mol. Ecol.* **17**, 2796–2811 (2008). doi: 10.1111/j.1365-294X.2008.03787.x; pmid: 18482262
76. J. E. Parkinson *et al*., Molecular tools for coral reef restoration: Beyond biomarker discovery. *Conserv. Lett.* 10.31219/osf.io/9k7ah (2019).
77. M. J. H. van Oppen, J. K. Oliver, H. M. Putnam, R. D. Gates, Building coral reef resilience through assisted evolution. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2307–2313 (2015). doi: 10.1073/pnas.1422301112; pmid: 25646461
78. S. Koren *et al*., Canu: Scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017). doi: 10.1101/gr.215087.116; pmid: 28298431
79. C.-S. Chin *et al*., Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* **13**, 1050–1054 (2016). doi: 10.1038/nmeth.4035; pmid: 27749838
80. S. Wang, L. Zhang, E. Meyer, M. V. Matz, Construction of a high-resolution genetic linkage map and comparative genome analysis for the reef-building coral *Acropora millepora*. *Genome Biol.* **10**, R126 (2009). doi: 10.1186/gb-2009-10-11-r126; pmid: 19900279
81. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014). doi: 10.1093/bioinformatics/btu170; pmid: 24695404
82. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010). doi: 10.1093/bioinformatics/btp698; pmid: 20080505
83. A. McKenna *et al*., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
84. N. Li, M. Stephens, Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213–2233 (2003). pmid: 14704198
85. P. Danecek *et al*., The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi: 10.1093/bioinformatics/btr330; pmid: 21653522
86. M. Nei, W. H. Li, Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5269–5273 (1979). doi: 10.1073/pnas.76.10.5269; pmid: 291943
87. C. Shinzato *et al*., Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* **476**, 320–323 (2011). doi: 10.1038/nature10249; pmid: 21785439
88. J. Felsenstein, G. A. Churchill, A Hidden Markov Model approach to variation among sites in rate of evolution. *Mol. Biol. Evol.* **13**, 93–104 (1996). doi: 10.1093/oxfordjournals.molbev.a025575; pmid: 8583911
89. Z. Fuller, zfuller5280/CoralGenomes: Public release. Zenodo (2020). doi: 10.5281/zenodo.3727666

ACKNOWLEDGMENTS

We thank members of the Australian Institute of Marine Science's coral bleaching group for assistance with sample collections and R. Durbin, A. Harpak, S. Myers, J. Pritchard, M. Schumer, and the Andolfatto, Przeworski, and Sella laboratories for helpful comments or discussions. We are particularly grateful to J. Sorkness from the Austin Reef Club for providing a live *A. millepora* fragment as a source of high-quality DNA for genome sequencing. We are also thankful to K. Dougan, A. Bellantuono, C. Granados-Cifuentes, and M. Rodriguez-Lanetty for making the *Durusdinium* draft genome sequence available. **Funding:** This project was funded by the Australian Institute of Marine Science, the Australian government's National Environmental Science Program, the Agouron Institute, and Columbia University. L.A.M. was supported by an AIMS@JCU Ph.D. scholarship and a National Environmental Science Program Tropical Water Quality Hub grant. M.M. was supported by National Science Foundation grant IOS1755277. We acknowledge computing resources from Columbia University's Shared Research Computing Facility project, which is supported by NIH Research Facility Improvement Grant 1G20RR030893-01, and associated funds from the New York State Empire State Development, Division of Science Technology and Innovation (NYSTAR) contract C090171. The coral collections were conducted under Great Barrier Reef Marine Park Authority collection permit G16/38488.1. **Author contributions:** Z.L.F., L.K.B., M.M., and M.P. designed the project; Z.L.F. and M.P. drafted the manuscript, alongside L.K.B. and M.M.; V.J.L.M., L.A.M., N.C., M.M., and L.K.B. collected samples and performed phenotyping; Z.L.F., Y.L., J.S., and L.S. performed analyses; P.A. and J.P. provided expertise and technical support; J.P. and L.S. conducted the sequencing validation; Z.L.F., L.K.B., and M.P. funded the research; and L.K.B., M.M., and M.P. jointly supervised the work. **Competing interests:** The authors declare no competing interests. **Data availability:** The reference genome assembly is available under BioProject PRJNA633778. All paired-end reads are deposited on the Sequence Read Archive under accession nos. SAMN13447100-13447355 and under BioProject PRJNA593014. The phased reference haplotype panel and imputation pipeline are freely available on www.gencove.com. All scripts and associated data are available through Zenodo (*89*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/eaba4674/suppl/DC1
Materials and Methods
Figs. S1 to S37
Tables S1 to S8
References (*90*–*133*)

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

5 December 2019; accepted 1 June 2020
10.1126/science.aba4674

RESEARCH ARTICLE SUMMARY

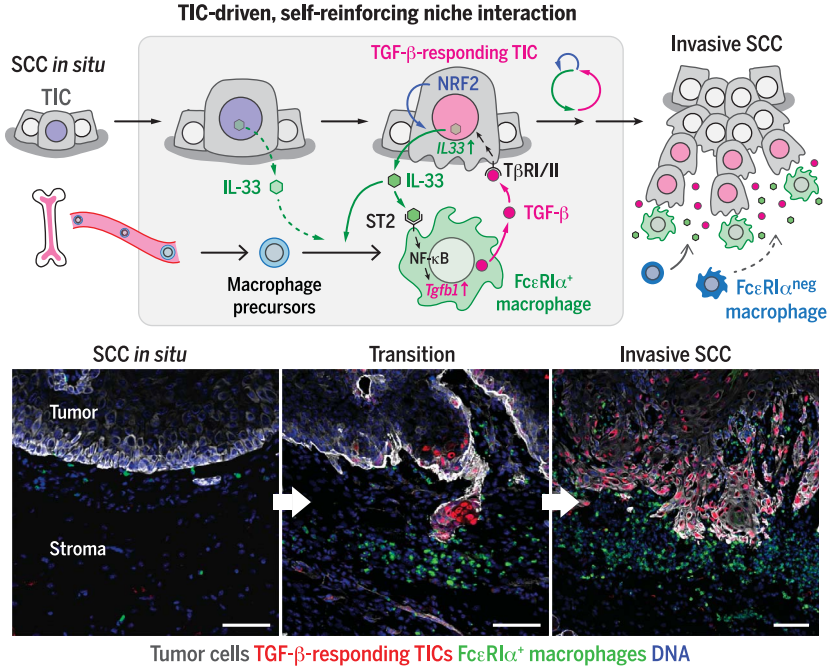
CANCER

Tumor-initiating cells establish an IL-33–TGF-β niche signaling loop to promote cancer progression

Sachiko Taniguchi, Ajit Elhance, Avery Van Duzer, Sushil Kumar, Justin J. Leitenberger, Naoki Oshimori*

INTRODUCTION: A small subset of tumor cells with long-term tumorigenic capacity, known as tumor-initiating cells (TICs), play a pivotal role in cancer development and therapy resistance. However, the development of effective TIC-targeted therapies is moving at a restricted pace due to the lack of identification of TIC vulnerabilities. Just as normal stem cells are regulated by external cues derived from specialized microenvironments, or stem cell niches, the stem-like state of TICs and the malignant phenotypes of their progeny are controlled by various factors emanating from the TIC-associated tumor microenvironment, the so-called TIC niche. Therefore, a mechanistic understanding of the cross-talk between TICs and the niche could accelerate the development of durable cancer therapeutics. Although the TIC niche is thought to evolve through reciprocal interactions with TICs, the mechanism by which the TIC–niche interaction emerges in the course of tumor development is poorly understood. Solid tumors are known to recruit immune cells in the stroma and create favorable conditions for their growth and survival. However, not much is known about how TICs regulate the localization and function of TIC-supportive immune cells in their spatial proximity.

RATIONALE: Using a mouse model of squamous cell carcinoma (SCC), we previously showed that transforming growth factor β (TGF-β) induces a subset of drug-resistant TICs that give rise to invasive, poorly differentiated progeny. We observed that these TGF-β–responding tumor cells are spatially associated with localized TGF-β expression in the adjacent stroma. Therefore, the mechanisms that lead to "TGF-β–rich" tumor microenvironments may underlie the development of TIC–niche interactions and potentially be exploited as a new target for destabilizing TICs. Because normal stem cells coordinate their niches by sending short-distance signals, we hypothesized that TICs might send a specific signaling molecule to the adjacent stroma to induce a TIC-supporting niche.



Signaling loop in the TIC niche that promotes cancer progression in mice. TGF-β–responding TICs release IL-33 via the NRF2-mediated antioxidant response. This induces differentiation of immature myeloid cells into FcεRIα$^+$ macrophages in close proximity to the TICs. In turn, FcεRIα$^+$ macrophages send reciprocal paracrine TGF-β signals to TICs to promote invasive progression and drug resistance of SCC and further induce the release of IL-33, thereby establishing a self-reinforcing signaling loop.

RESULTS: Focusing on the cytokine milieu and immune cells in the proximity of TGF-β–responding TICs, here, we address how TICs generate a spatially distinct niche microenvironment that is required for invasive progression and drug resistance of SCC. In a search for potential paracrine regulators of the neighboring tumor microenvironment, we identified interleukin-33 (IL-33) as the most highly upregulated cytokine in TGF-β–responding TICs. Whereas IL-33 is stored in the nucleus under normal conditions, we found that it is released into the extracellular space in response to the NRF2-mediated antioxidant response, a hallmark of TGF-β–responding TICs. This TIC-derived IL-33 was required for invasive progression and drug resistance of SCC. Mechanistically, IL-33 induces the accumulation of a subset of tumor-associated macrophages expressing the IL-33 receptor ST2 and the high-affinity IgE receptor (FcεRIα) in close proximity to TICs (i.e., within a 50-μm radius). These previously unappreciated FcεRIα$^+$ macrophages were differentiated and alternatively activated from bone marrow–derived cells and created a TGF-β–rich niche microenvironment through the IL-33–ST2–NF-κB pathway, inducing paracrine TGF-β signaling to TICs and further upregulating IL-33 expression. The abrogation of the pathway or the depletion of FcεRIα$^+$ macrophages reduced the number of TGF-β–responding TICs, the rate of invasive tumor progression, and chemotherapy resistance.

ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aay1813

CONCLUSION: Therapy-resistant TICs are considered to be major culprits in cancer treatment failure. Studying a mouse model, we unveil the cellular and molecular basis for the formation of a TIC niche that promotes malignant progression and drug resistance of SCC. The discovery of the IL-33–TGF-β niche signaling loop between TICs and FcεRIα$^+$ macrophages provides mechanistic insights into self-reinforcing TIC–niche interactions, which could be a potential target for destabilizing TICs to improve cancer treatment efficacy. ■

The list of author affiliations is available in the full text online.
*Corresponding author. Email: oshimori@ohsu.edu
Cite this article as S. Taniguchi *et al.*, *Science* **369**, eaay1813 (2020). DOI: 10.1126/science.aay1813

RESEARCH ARTICLE

CANCER

Tumor-initiating cells establish an IL-33–TGF-β niche signaling loop to promote cancer progression

Sachiko Taniguchi[1], Ajit Elhance[1], Avery Van Duzer[1], Sushil Kumar[1], Justin J. Leitenberger[2], Naoki Oshimori[1,2,3,4]*

Targeting the cross-talk between tumor-initiating cells (TICs) and the niche microenvironment is an attractive avenue for cancer therapy. We show here, using a mouse model of squamous cell carcinoma, that TICs play a crucial role in creating a niche microenvironment that is required for tumor progression and drug resistance. Antioxidant activity in TICs, mediated by the transcription factor NRF2, facilitates the release of a nuclear cytokine, interleukin-33 (IL-33). This cytokine promotes differentiation of macrophages that express the high-affinity immunoglobulin E receptor FcεRIα and are in close proximity to TICs. In turn, these IL-33–responding FcεRIα$^+$ macrophages send paracrine transforming growth factor β (TGF-β) signals to TICs, inducing invasive and drug-resistant properties and further upregulating IL-33 expression. This TIC-driven, IL-33–TGF-β feedforward loop could potentially be exploited for cancer treatment.

A small subset of tumor cells with long-term tumorigenic capacity, known as tumor-initiating cells (TICs), play a pivotal role in cancer development and therapy resistance (*1*, *2*). Just as normal stem cells are regulated by external cues derived from specialized microenvironments or stem cell niches (*3*), the stem-like state of TICs and the malignant phenotypes of their progeny are controlled by various factors emanating from the TIC-associated tumor microenvironment, the so-called TIC niche (*4*, *5*). Therefore, a mechanistic understanding of the interaction between TICs and their niche would likely accelerate the development of effective cancer therapeutics. Although the TIC niche is thought to evolve through reciprocal cross-talk with TICs (*6*), the mechanism by which the TIC–niche relationship emerges in the course of tumor development is poorly understood. Solid tumors are known to recruit immune cells in the stroma and create favorable conditions for their growth and survival (*7*, *8*). In fact, the spatial distribution of tumor-infiltrating immune cells can be predictive of poor patient survival and therapeutic responses (*9*, *10*). However, not much is known about how TICs regulate the localization and function of TIC-supportive immune cells in their spatial proximity.

Squamous cell carcinoma (SCC), the second most prevalent cancer worldwide, harbors TICs at the tumor–stroma interface (*11*, *12*), an ideal anatomical location for paracrine interactions with potential niche cells. One of the critical regulators of stem cells and SCC development is transforming growth factor β (TGF-β), which acts as a tumor suppressor or promoter at different stages of cancer development (*13*–*16*). To investigate the stage-specific role of TGF-β in tumorigenesis, we previously devised a mouse model of SCC that allowed us to label and lineage trace tumor cells responding to TGF-β (*17*). Using this approach, we showed that TGF-β induces slow-cycling tumor cells that give rise to invasive, poorly differentiated progeny. Moreover, TGF-β–responding tumor cells survived chemotherapy through a cytoprotective mechanism activated by NRF2, the master transcription factor for antioxidant responses, and their progeny repopulated the recurrent tumors (*17*), indicating that TGF-β–responding tumor cells function as drug-resistant TICs. We observed that these TGF-β–responding tumor cells are spatially associated with localized TGF-β expression in the adjacent stroma. In several types of human cancers, abundant TGF-β expression in the stroma is known to be associated with poor prognosis (*18*, *19*). Therefore, the molecular and cellular mechanisms that lead to "TGF-β–rich" tumor microenvironments may underlie the development of TIC–niche interactions and could potentially be exploited as a new target for destabilizing TICs.

Here, by focusing on the cytokine milieu and immune cells in the proximity of TGF-β–responding TICs, we unveiled the cellular and molecular basis of a TIC–niche interaction that promotes malignant progression and drug resistance in a mouse model of SCC. We identified a TIC-derived nuclear cytokine, interleukin-33 (IL-33), a member of the IL-1 family that has a pleiotropic role in tumorigenesis (*20*, *21*), and IL-33–responding macrophages as key components of a signaling loop in the TIC niche that emerges during the early stages of SCC development and regulates tumor progression.

Results

IL-33 expression is up-regulated in TGF-β–responding tumor cells

To study TIC–niche interactions emerging in an early tumor microenvironment, we developed an oncogenic, HRAS-driven SCC mouse model using an in utero, epidermis-specific lentiviral (LV) transduction method (*17*, *22*). Briefly, epidermal progenitor cells of *TetO-HRAS^{G12V}* (*23*); *Rosa-LSL-EYFP* (*24*) transgenic mouse embryos were transduced with LV vectors expressing rtTA and Cre constitutively and mScarlet fluorescent protein in response to TGF-β–SMAD2/3 signaling (fig. S1, A and B). Successful transduction was monitored by the expression of yellow fluorescent protein (YFP) (fig. S1C). Postnatally, the LV-rtTA induced HRAS^{G12V} expression in a doxycycline (Dox)–dependent manner in YFP$^+$ cells, which initiated SCC formation (fig. S1D). We observed histopathologically distinct regions manifesting well-differentiated and invasive SCC (fig. S1E), and the LV fluorescent reporter (as well as SMAD2 phosphorylation; fig. S1F) illuminated TGF-β–responding tumor cells in invasive SCC (Fig. 1A, pink signal). Moreover, the frequency of TGF-β–responding tumor cells was positively correlated with the distribution of TGF-β ligand in the adjacent stroma (Fig. 1A, green signal, and 1B). This raised the question of whether TGF-β–responding TICs establish specific interactions with adjacent cells in the stroma, potentially niche cells, to maintain TGF-β–rich microenvironments.

Normal stem cells coordinate their niches by sending short-distance cues (*25*, *26*) and subsequently receive reciprocal signals, including TGF-β, to maintain stem cell properties (*14*, *16*). Therefore, we hypothesized that TICs may send a specific signaling molecule to the adjacent stroma to directly induce a TGF-β–rich, TIC-supporting niche. To identify candidate molecules, we purified TGF-β reporter$^+$ and reporter^{neg} tumor basal (YFP$^+$, integrin α6^{hi}, β1^{hi}) cells by fluorescence-activated cell sorting (FACS) and compared their transcriptomes by RNA sequencing (RNA-seq) (fig. S1G). By combining current and previously published data (*17*), we obtained a highly reliable list of genes differentially expressed in TGF-β–responding tumor cells (Fig. 1C, significance cutoff, adjusted $P < 0.01$), including a battery of NRF2 target genes involved in antioxidant responses (Fig. 1C, triangles, and fig. S1H).

We then focused on genes encoding cytokines and chemokines (Fig. 1C, diamonds, and

[1]Department of Cell, Developmental & Cancer Biology, Oregon Health & Science University, Portland, OR 97239, USA. [2]Department of Dermatology, Oregon Health & Science University, Portland, OR 97239, USA. [3]Department of Otolaryngology – Head & Neck Surgery, Oregon Health & Science University, Portland, OR 97239, USA. [4]Knight Cancer Institute, Oregon Health & Science University, Portland, OR 97239, USA.
*Corresponding author. Email: oshimori@ohsu.edu

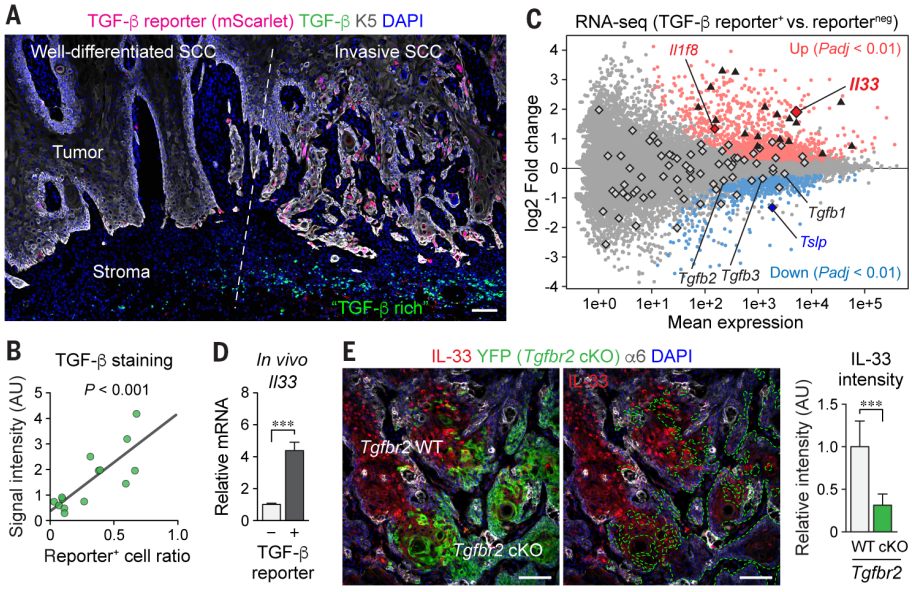


Fig. 1. IL-33 is up-regulated in tumor cells adjacent to TGF-β–rich microenvironments in a mouse model of SCC. (**A**) Fluorescent reporter illuminates higher TGF-β signaling activity in the invasive region of SCC. Note that increased TGF-β immunostaining (green) in the stroma correlates with the presence of TGF-β reporter$^+$ cells (pink) at the invasive tumor front. (**B**) The proportion of TGF-β reporter$^+$ cells and the intensity of TGF-β immunolabeling were measured by fluorescence microscopy of tumor sections with active and inactive TGF-β signaling (n = 7). (**C**) MA plot presentation of RNA-seq data comparison between TGF-β–responding and -nonresponding tumor basal cells. Biological replicates, n = 5. Red and blue dots indicate significantly up- and down-regulated genes, respectively (adjusted P < 0.01). Triangles indicate up-regulated NRF2-regulated genes involved in antioxidant responses (fig. S1H). Diamonds indicate genes encoding cytokines or chemokines (fig. S1I). (**D**) qPCR analysis of *Il33* mRNA in FACS-isolated TGF-β reporter$^+$ and reporter^{neg} tumor basal cells in vivo (the sorting strategy is shown in fig. S1G; n = 3). Data are expressed as relative mean expression with SEM and were analyzed with unpaired t test, ***P < 0.001. (**E**) Immunodetection of IL-33 protein in mosaic tissue with *Tgfbr2* cKO (YFP$^+$) and WT (YFP^{neg}) tumor cells. Mosaic tissues were induced by a partial activation of K14-CreER by limited dose of tamoxifen. (Right) Dotted lines denote the boundaries of WT and cKO tumor cell regions. (Graph) Quantification of IL-33 signal intensity (n = 3). Data are expressed as relative mean intensity with SD and were analyzed with unpaired t test, ***P < 0.001. Scale bars, 50 μm.

fig. S1I). There was no significant difference in *Tgfb1*, *Tgfb2*, and *Tgfb3* expression between TGF-β reporter$^+$ and reporter^{neg} tumor cells, supporting the notion that a paracrine signal from the adjacent stroma, but not an autocrine signal, may play a major role in generating TGF-β–responding TICs. We identified *Il33* as the most significantly up-regulated and highly expressed gene in TGF-β–responding tumor cells. We verified elevated *Il33* mRNA expression by quantitative PCR (qPCR) analysis of independent in vivo samples (Fig. 1D). To assess the association between TGF-β signaling and IL-33 expression in vivo, we locally deleted the *Tgfbr2*, the gene encoding the type II TGF-β receptor (*27*), in tumor epithelial cells by activating K14-CreER using a limiting dose of tamoxifen (*28*). The side-by-side comparison of YFP$^+$ *Tgfbr2* conditional knockout (cKO) and YFP^{neg} wild-type (WT) mosaic tissues revealed that the lack of the TGF-β receptor (Fig. 1E, green) was significantly correlated with reduced IL-33 immunolabeling (Fig. 1E, red). These results strongly suggest that TGF-β signaling is involved in the up-regulation of IL-33 in SCC.

NRF2-mediated antioxidant response triggers the extracellular release of IL-33

We next sought to identify what triggers IL-33 release from TGF-β–responding tumor cells. IL-33 is stored in the nucleus under normal conditions but is released into the extracellular space as an alarmin cytokine upon cell injury, especially by necrotic cell death (*20*). It remains unclear how IL-33 can be released from living tumor cells, particularly TICs. Immunostaining confirmed the nuclear localization of IL-33 in normal epidermal cells (fig. S2A). However, as tumors progressed to invasive SCC, cytoplasmic IL-33 staining became evident in keratin 5–positive (K5$^+$) tumor basal cells, whereas IL-33 tended to remain in the nucleus of K5^{dim} suprabasal cells (Fig. 2A). Cytoplasmic IL-33 localization was more frequently observed in TGF-β–responding tumor cells compared with their nonresponding counterparts (Fig. 2B), suggesting that IL-33 may be preferentially released from TGF-β–responding TICs in the adjacent stroma.

It has been reported that nonlethal mechanical stress can induce the secretion of nuclear IL-33 (*29*). We explored the possibility of nonlethal oxidative stress and the NRF2-mediated antioxidant response as a potential mechanism of IL-33 release, because the latter is a hallmark of TGF-β–responding TICs (*17*). The activity of NRF2 is primarily controlled at the level of protein stability (*30*). K5$^+$ tumor basal cells in invasive SCC accumulated stabilized NRF2 protein in the nucleus, and these NRF2$^+$ cells exhibited cytoplasmic localization of IL-33 (Fig. 2C, arrowheads). Such a correlation was also observed in invasive SCC of a chemical carcinogenesis mouse model (fig. S2B) and human SCC (Fig. 2D). It is noteworthy that stabilized NRF2 was detected explicitly in tumor epithelial cells, but not in stromal cells, and that IL-33 protein expressed in stromal cells was sequestered in the nucleus (Fig. 2C, arrows), further supporting the potential link between NRF2 activity and IL-33 release.

In vitro, H_2O_2-induced NRF2 stabilization in HRAS^{G12V}-expressing cells was correlated with the reduction of intracellular IL-33 protein levels (Fig. 2E) without changing *Il33* mRNA levels (fig. S2C). At the same time, IL-33 protein became detectable in the conditioned medium of H_2O_2-treated cells (Fig. 2F), suggesting that IL-33 is released from the cells. To assess the functional relevance of NRF2 in the subcellular localization of IL-33, we depleted NRF2 or KEAP1 (an E3 ubiquitin ligase that leads to NRF2 degradation) (*30*). Short hairpin RNA (shRNA)–mediated *Nfe2l2* (*Nrf2*) or *Keap1* knockdown (KD) did not affect *Il33* mRNA levels (fig. S2D). However, *Nrf2* KD hampered H_2O_2-induced reduction of intracellular IL-33 (Fig. 2G). Conversely, *Keap1* KD stabilized NRF2 and decreased intracellular IL-33 levels even without H_2O_2 (Fig. 2H).

As we previously reported, prolonged TGF-β treatment of HRAS^{G12V}-expressing cells induced NRF2 stabilization and the expression of NRF2 target genes such as *Hmox1* and *Me1* (Fig. 2I and fig. S2E) (*17*). TGF-β treatment also up-regulated *Il33* mRNA levels (fig. S2E), and IL-33 protein was detected in both the nucleus and cytoplasmic space. This TGF-β–induced cytoplasmic IL-33 signal was suppressed by *Nrf2* KD, resulting in sharp nuclear IL-33 localization. By contrast, *Keap1* KD facilitated both NRF2 stabilization and cytoplasmic IL-33 localization (Fig. 2I). Together, the data from mouse and human tissues and results from in vitro experiments strongly suggest that the NRF2-mediated antioxidant response activated in TGF-β–responding TICs facilitates the extracellular release of IL-33.

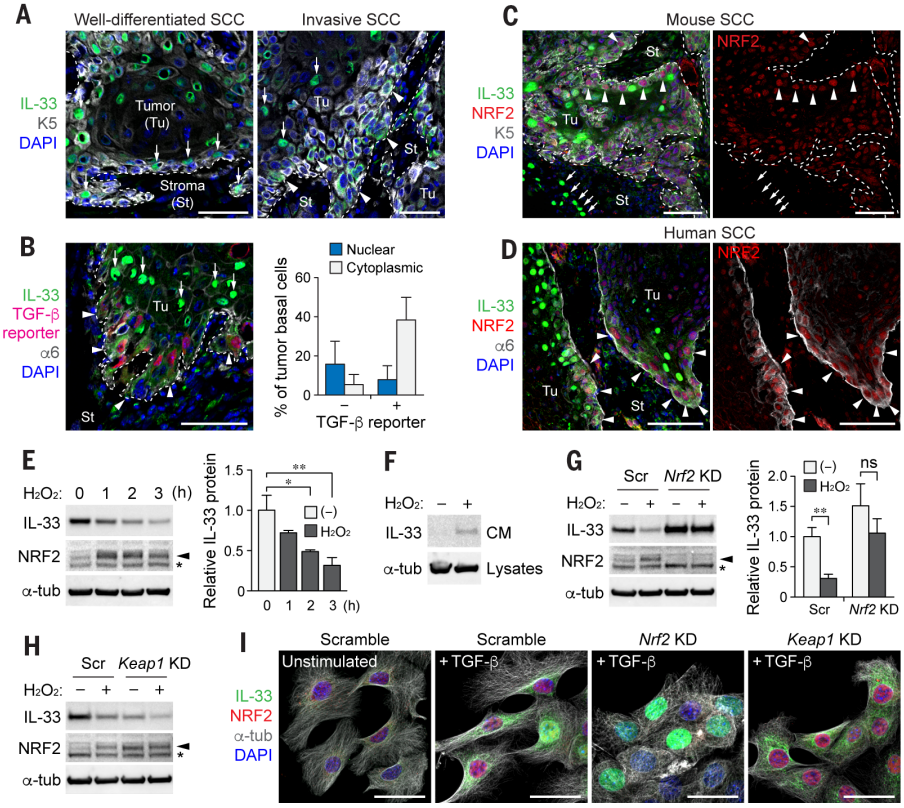


Fig. 2. The NRF2-mediated antioxidant response induces extracellular release of IL-33. (**A**) Immunolabeling of tumor tissues showing a range of IL-33 expression patterns. In well-differentiated SCC, IL-33 nuclear staining can be observed in both $K5^+$ basal (arrows) and $K5^{dim}$ suprabasal cells. In invasive SCC, cytoplasmic IL-33 staining can be observed in $K5^+$ basal cells (arrowheads). (**B**) TGF-β $reporter^+$ tumor cells at the invasive front exhibited cytoplasmic IL-33 localization (arrowheads), whereas suprabasal cells had nuclear IL-33 (arrows). (Graph) Quantification of IL-33 nuclear versus cytoplasmic localization in TGF-β $reporter^+$ and $reporter^{neg}$ tumor basal cells. (**C**) Immunostaining of invasive SCC tissue showing the correlation between nuclear NRF2 expression and cytoplasmic IL-33 localization (arrowheads). Note that nuclear IL-33–expressing $K5^{neg}$ stromal cells are negative for NRF2. (**D**) Immunolabeling of human SCC section showing the correlation between cytoplasmic IL-33 localization and nuclear NRF2 expression (arrowheads). (**E**) Western blots showing that nonlethal oxidative stress induced by 500 μM H_2O_2 reduces IL-33 in total cell lysates over 3 hours. Arrowhead indicates an NRF2-specific band. Asterisk indicates a nonspecific band. (Graph) Quantification of IL-33 protein (n = 3). The relative intensity of IL-33 bands normalized based on the α-tubulin (α-tub) loading control is shown. Data are shown as mean with SD and were analyzed with unpaired t test, $^*P$ = 0.0195, $^{**}P < 0.01$. (**F**) Western blot analysis of IL-33 protein in the concentrated conditioned medium (CM) of cells with and without H_2O_2 treatment. (**G**) *Nrf2* (*Nfe2l2*) KD prevents H_2O_2-induced IL-33 protein reduction in total cell lysates. (Graph) Quantification of IL-33 protein (n = 4). Data are shown as mean with SD. (**H**) Western blots show that *Keap1* KD cells increase NRF2 and decrease intracellular IL-33 protein levels regardless of oxidative stress. (**I**) Immunolabeling of IL-33 and NRF2 in scramble control, *Nrf2* KD, or *Keap1* KD cells with or without TGF-β treatment (50 pM for 36 hours). Note that *Nrf2* KD results in an exclusive nuclear localization of IL-33, whereas TGF-β–treated and *Keap1* KD cells show granular nuclear and cytoplasmic IL-33 localization. Dotted lines denote the tumor–stroma boundaries. Scale bars, 50 μm.

Tumor-derived IL-33 promotes paracrine TGF-β signaling, invasive progression, and drug resistance

To determine the role of tumor-derived IL-33 in SCC development, we generated LV vectors carrying potent *Il33* shRNA, which did not cause noticeable differences in cell morphology or proliferation rates in vitro (fig. S3, A to D). We then depleted IL-33 from tumor epithelial cells in vivo through the transduction of the LV vectors in utero (fig. S1, A and B). Immunostaining confirmed the efficient depletion of IL-33 in YFP^+ tumor cells, whereas infrequent nuclear IL-33 expression remained in YFP^{neg} cells in the stroma (Fig. 3A). To validate *Il33* KD efficacy in vivo, we purified tumor basal cells (YFP^+, integrin $\alpha6^{hi}$, $\beta1^{hi}$), $CD45^+$ immune cells, $CD140a^+$ fibroblasts, and $CD31^+$ endothelial cells by FACS (fig. S3E). qPCR analysis of FACS-purified cells verified the tumor cell–specific *Il33* KD and indicated that IL-33 is most highly expressed in tumor epithelial cells (fig. S3F), among which the TGF-β–responding TIC population is the dominant source (Fig. 1D).

Transduced mice developed 8.2 ± 2.8 tumors (n = 46) and 6.1 ± 3.4 tumors (n = 31) in scramble control and *Il33* KD conditions, respectively. The average size of the five largest tumors indicated that control tumors grew larger over time, whereas *Il33* KD tumors were stunted in growth (Fig. 3B). Histopathological analysis showed that control tumors progressed to invasive, poorly differentiated states when assessed by a differentiation marker, keratin 10 (K10) (Fig. 3, C and D). By contrast, *Il33* KD tumors showed smoother tumor edges and well-differentiated states. Despite smaller tumor size, a short 5-ethynyl-2′-deoxyuridine (EdU) pulse revealed that *Il33* KD tumor cells are more proliferative than those in control tumors (Fig. 3E).

We previously demonstrated that TGF-β–responding tumor cells are a slow-cycling TIC population that give rise to invasive, poorly differentiated progeny (*17*), raising the intriguing possibility that stunted progression of *Il33* KD tumors may be due to the reduction of paracrine TGF-β signaling toward tumor cells. To determine whether tumor-derived IL-33 is responsible for evoking TGF-β signaling, we quantified TGF-β–responding tumor cells by flow cytometry. In YFP^+ tumor epithelial cells (Fig. 3F, green gates), the frequency of TGF-β $reporter^+$ ($mScarlet^+$) cells (Fig. 3F, gray versus orange gates) were significantly decreased in *Il33* KD tumors compared with control tumors (Fig. 3F, graph). In vitro, TGF-β–induced SMAD2 phosphorylation and cytostatic responses were still intact in *Il33* KD cells (fig. S3, D and G), strongly suggesting that the reduction of TGF-β signaling in *Il33* KD tumors is likely due to a non-cell-autonomous effect, potentially the lack of TGF-β–rich tumor microenvironments.

Because TGF-β can exert drug resistance effects on TICs in vivo (Fig. 3G) (*17*), we next assessed the efficacy of cisplatin chemotherapy in *Il33* KD tumors. Consistent with the reduction of TGF-β–responding tumor cells, $K5^+$ basal cells of *Il33* KD tumors underwent apoptosis more frequently (Fig. 3G). After cisplatin treatment, most tumors shrank regardless of *Il33* status; however, control tumors typically recurred in 3 weeks. By contrast, almost all *Il33* KD tumors failed to regrow (Fig. 3H). Our data suggest that tumor-derived IL-33 plays an important role in generating TGF-β–responding TICs that promote invasive progression and drug resistance in SCC.

FcεRIα$^+$ macrophages accumulate in close proximity to TGF-β–responding tumor cells

We previously observed dense $CD11b^+$ and $Ly6C^+$ cell localizations at TGF-β–rich stroma

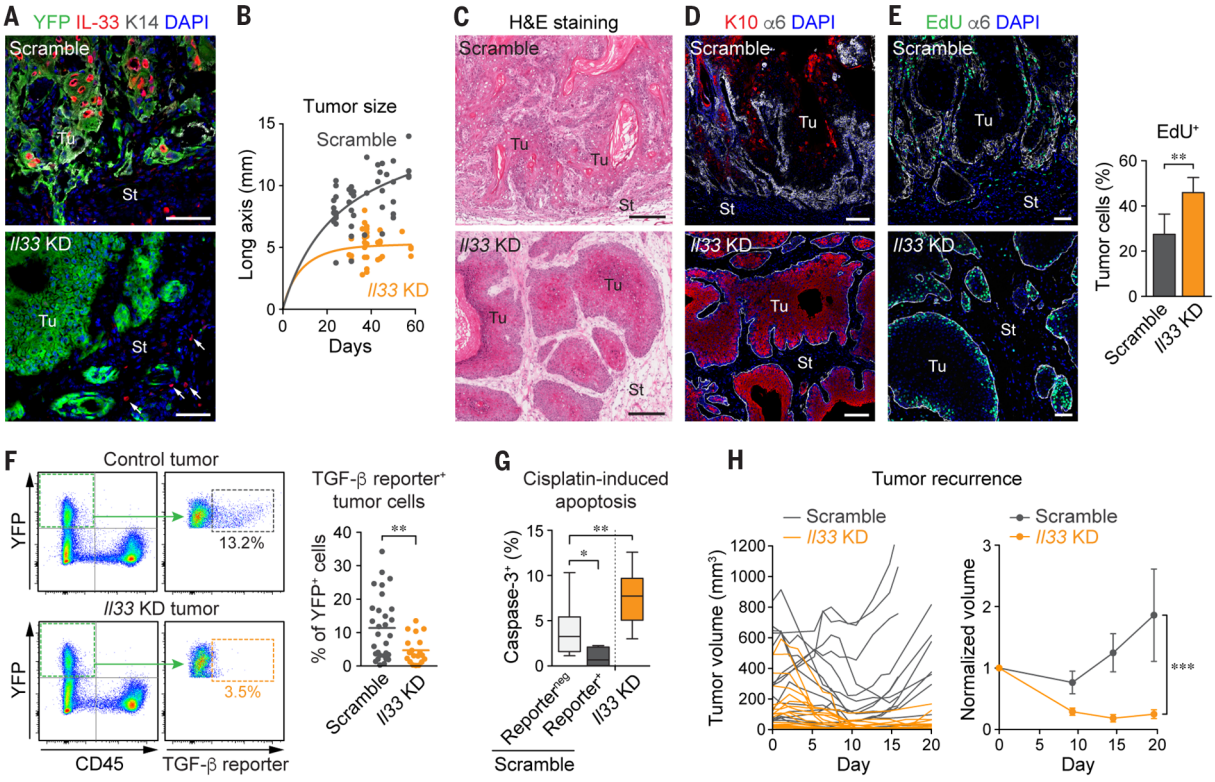


Fig. 3. Depletion of tumor-derived IL-33 impairs tumor progression and paracrine TGF-β signaling. (**A**) Immunolabeling of tumor sections showing efficient IL-33 protein depletion in *Il33* shRNA-transduced (YFP$^+$) cells. Note that some YFP^{neg} stromal cells maintain nuclear IL-33 (arrows). (**B**) HRAS^{G12V}-driven tumors in control and *Il33* KD mice were sized at the time of euthanasia. Scramble control, n = 46; *Il33* KD, n = 31. Approximation curves were drawn by applying the Michaelis-Menten kinetics curve. (**C**) Hematoxylin and eosin (H&E) staining of tumor sections showing smoother edges in *Il33* KD tumors compared with control tumors, suggestive of reduced invasive capacity. Scale bars, 100 μm. (**D**) Immunolabeling of K10 showing a loss of differentiation property in control tumors, but not in *Il33* KD tumors. (**E**) Detection of EdU incorporated 4 hours before euthanasia. (Graph) Quantification of EdU$^+$ cells in integrin α6$^+$ tumor basal cells of control and *Il33* KD tumors (n = 3) (n = 806 to 1747 cells). Data are shown as mean with SEM and were analyzed with unpaired t test, $^{**}P$ = 0.004. (**F**) Flow cytometry analysis of TGF-β reporter (mScarlet) expression in YFP$^+$ tumor epithelial cells. YFP$^+$ cells in *Il33* KD tumors show fewer and reduced TGF-β reporter activity compared with those in control tumors. (Graph) Proportion of TGF-β reporter$^+$ tumor cells. Each dot indicates values from an individual tumor. Scramble control, n = 30; *Il33* KD, n = 19. Data are shown as mean with individual values and were analyzed with unpaired t test, $^{**}P$ = 0.007. (**G**) Tumor-bearing mice were treated with cisplatin (10 mg/kg), and cells undergoing apoptosis (cleaved caspase-3$^+$) in K5$^+$ tumor basal cells were quantified 2 days after treatment (n = 3). Data are shown in box-and-whisker plots (midline, median; box, 25th and 75th percentiles; whiskers, min and max) and were analyzed with unpaired t test, $^{**}P$ = 0.0061, $^{*}P$ = 0.0334. (**H**) Volume of tumors after cisplatin administration. (Left) Spider plots showing the changes for each tumor. (Right) Normalized tumor volume of each condition. Scramble control, n = 3 (15 tumors), *Il33* KD, n = 5 (24 tumors). Tu, tumor. St, stroma. Scale bars, 50 μm [except for (C)].

(*17*), raising the possibility that CD11b$^+$Ly6C$^+$ immature myeloid cells (*31*) differentiate into TGF-β–producing immune cells in response to the cytokine milieu around TICs. Therefore, we investigated whether TGF-β–responding TICs are associated with a specific subset of immune cells and, if so, how TIC-derived IL-33 is involved in the development of spatially distinct microenvironments. To address this, we measured the frequency of TGF-β reporter$^+$ tumor cells and the density of immune cells within a 50-μm radius of the tumor edges by immunofluorescence analysis of tumor sections (fig. S4, A and B). Our data showed that the peritumoral areas were dominated by CD11b$^+$ myeloid cells rather than CD3$^+$ T cells. Despite these lower cell numbers, the density of FoxP3$^+$ regulatory T (Treg) cells was positively correlated with the frequency of TGF-β–responding tumor cells. This is not surprising because TGF-β promotes the differentiation and function of Treg cells (*32*), but it suggests the presence of a TGF-β gradient in the stroma. The strongest positive correlation was found in cells expressing the high-affinity IgE receptor FcεRIα (P = 0.0018) (fig. S4B). FcεRIα$^+$ cells accumulated in the proximity (~50-μm radius) of TGF-β–responding tumor cells that showed nascent invasive phenotypes (Fig. 4A, i). Furthermore, their close spatial relationship was sustained during invasive progression at the tumor–stroma interface (Fig. 4A, ii and iii, and B).

Mast cells make up most of the FcεRIα$^+$ cells in the normal dermis, but it is known that their density declines in invasive SCC and mast cells become dispensable for tumor progression (*33*, *34*). We found most of the FcεRIα$^+$ cells in invasive SCC to be negative for mast cell tryptase, an indication that these are not mast cells but another myeloid cell lineage (fig. S4C). To identify the major FcεRIα$^+$ cells in SCC, we analyzed a panel of immune cell markers for FcεRIα expression by flow cytometry (fig. S4D). This comprehensive analysis revealed that a fraction of CD11b$^+$Ly6C$^+$Ly6G^{neg} immature myeloid cells/monocytes and F4/80$^+$ macrophages express FcεRIα (Fig. 4C). In particular, a subset of macrophages with low major histocompatibility complex (MHC)-II surface expression showed the highest FcεRIα levels. Immunofluorescence analysis validated that FcεRIα$^+$ cells are mostly F4/80$^+$ macrophages, particulary

those in the ~50-μm peritumoral areas (Fig. 4D, arrowheads).

Low MHC-II expression suggests that FcεRIα^{+} cells may resemble alternatively activated macrophages, also known as M2-like tumor-associated macrophages (*35*, *36*). To characterize FcεRIα^{+} macrophages, we purified FcεRIα^{+} and FcεRIα^{neg} macrophages from SCC by FACS (Fig. 4E) and analyzed their transcriptomes by RNA-seq. Compared with FcεRIα^{neg} macrophages, FcεRIα^{+} macrophages up-regulated the expression of multiple genes, including *Arg1*, *Mrc1* (*Cd206*), *Cd163*, *Stab1*, and *Il10*, which are known to be elevated in alternatively activated macrophages (*37*) (Fig. 4F). By contrast, FcεRIα^{+} macrophages significantly down-regulated the expression of classically activated (M1-like) macrophage markers such as *Tnf*, *Cd40*, and *MHC-II*. We further validated RNA-seq results by qPCR analysis of marker genes. Although an established M1 marker, *Nos2*, was similarly expressed regardless of FcεRIα expression, the M2 markers *Arg1* and *Tgfb1* were significantly up-regulated in the FcεRIα^{+} macrophages compared with their FcεRIα^{neg} counterparts (Fig. 4G). Indeed, immunostaining showed that most FcεRIα^{+} cells were overlapped with prominent TGF-β localization in the stroma (Fig. 4H and fig. S4E).

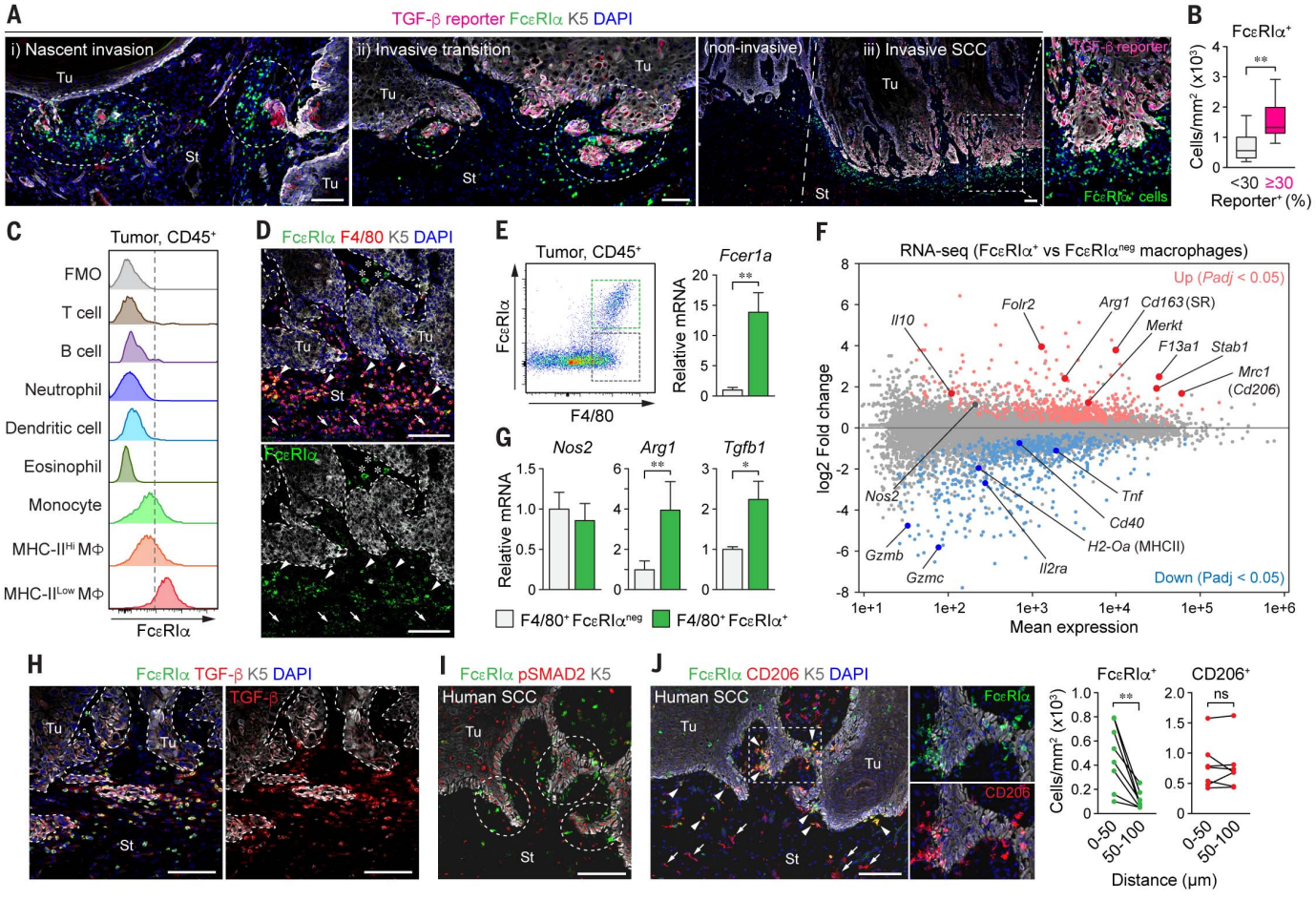


Fig. 4. FcεRIα^{+} macrophages accumulate in the proximity of TGF-β–responding invasive tumor cells. (**A**) Immunolabeling of tissue sections from different SCC stages showing clusters of FcεRIα^{+} cells accumulated within a 50-μm radius of TGF-β reporter^{+} invasive tumor cells (dotted circles and box). (**B**) Quantification of FcεRIα^{+} cells in the stroma adjacent to the tumor leading edge with TGF-β reporter high (≥30%) versus low (<30%) showing higher density of FcεRIα^{+} cells near TGF-β–responding cells. n = 4 mice, total 12 images that included both TGF-β reporter high and low areas and were analyzed with unpaired t test, $^{**}P$ = 0.0015. (**C**) Flow cytometry analysis of FcεRIα expression in different immune cell types in SCC. Gating strategy is shown in fig. S4D. (**D**) Immunolabeling of invasive SCC section showing that most FcεRIα staining overlaps with F4/80, an established murine macrophage marker, and present within a 50-μm radius of tumor leading edges. Note that F4/80^{+} cells farther than the 50-μm radius from tumor edges were negative for FcεRIα. Asterisks indicate rare FcεRIα^{+}F4/80^{neg} cells. Dotted lines denote the tumor–stroma boundaries. (**E**) FACS plots of F4/80^{+}FcεRIα^{+} (green dotted box) and F4/80^{+}FcεRIα^{neg} (gray dotted box) macrophages in SCC. (**F**) MA plot presentation of RNA-seq data comparison between FcεRIα^{+} and FcεRIα^{neg} macrophages in SCC. Biological replicates, n = 3. Red and blue dots indicate significantly up- and down-regulated genes, respectively (adjusted P < 0.05). Dark red and dark blue dots highlight some of the "M2" and "M1" marker genes, respectively. (**G**) qPCR analysis of FACS-isolated macrophage populations in SCC (n = 3). Data are shown as mean with SD. *Fcer1a*, $^{**}P$ = 0.0019. *Arg1*, $^{**}P$ = 0.0032. *Tgfb1*, $^{*}P$ = 0.0266. (**H**) Immunolabeling of invasive SCC section showing that most FcεRIα^{+} cells overlap with TGF-β protein localization in the stroma. Magnified images are shown in fig. S4E. (**I**) Immunolabeling of human SCC section showing that FcεRIα^{+} cells cluster around invasive phospho-SMAD2 (pSMAD2)$^{+}$ tumor cells (dotted circles). (**J**) Immunolabeling of human SCC section showing that most FcεRIα^{+} cells overlap with CD206, an established marker of alternatively activated macrophages, at the tumor-stroma interfaces, whereas CD206^{+}FcεRIα^{neg} cells present in the stroma farther than a 50-μm radius from tumor edges. (Graph) Quantification of the density of FcεRIα^{+} cells and CD206^{+} cells 0 to 50 or 50 to 100 μm away from tumor edges. Data were analyzed with paired t test, $^{**}P$ = 0.0035. Tu, tumor. St, stroma. Scale bars, 50 μm.

These previously unappreciated FcεRIα$^+$ macrophages were also found in invasive SCC induced by DMBA–TPA chemical carcinogenesis (*38*) and in *K14-HPV16*–transgenic models (*33*), where they accumulated around phospho-SMAD2$^+$ tumor cells (fig. S4, F and G). Moreover, we detected FcεRIα$^+$CD206$^+$ tumor-associated macrophages in human SCC patient samples, which were also clustered around phospho-SMAD2$^+$ TGF-β–responding tumor cells (Fig. 4I). FcεRIα$^+$CD206$^+$ macrophages in human SCC were also significantly enriched within a 50-μm radius from the tumor edges (Fig. 4J, arrowheads), whereas FcεRIα^{neg}CD206$^+$ macrophages were distributed more evenly in the stroma (Fig. 4J, arrows and graph). Collectively, these results strongly suggest that FcεRIα$^+$ cells are a subset of tumor-associated macrophages that create spatially distinct, TGF-β–rich tumor microenvironments in the proximity of TICs.

IL-33 induces FcεRIα$^+$ macrophages that activate proinvasive TGF-β signaling

Macrophages are known pro-tumorigenic immune cells and a major component of the TIC niche (*39*, *40*), and previous studies showed that IL-33 can induce macrophage recruitment and alter their phenotype to resemble alternatively activated macrophages (*41*, *42*). We next sought to determine whether TIC-derived IL-33 is involved in FcεRIα$^+$ macrophage differentiation. We found that FcεRIα$^+$ macrophages in SCC expressed high levels of the IL-33 receptor ST2 (*43*) (Fig. 5A). By contrast, ST2 was barely detected in TGF-β reporter$^+$ and reporter^{neg} tumor cells (fig. S5A). Therefore, IL-33 released from TICs may act as a short-distance paracrine signal to induce the differentiation of FcεRIα$^+$ macrophages.

To address this possibility, we used an in vitro culture system utilizing bone marrow–derived hematopoietic progenitor cells. Similar to macrophage colony-stimulating factor (CSF1), IL-33 treatment induced macrophage differentiation in 6 days (fig. S5B). Flow cytometry, qPCR, and immunofluorescence analyses showed that FcεRIα and ST2 expression were most robustly elevated in IL-33–induced macrophages (Fig. 5, B and C, and fig. S5C). Moreover, CSF1-induced macrophages could acquire FcεRIα expression by subsequent IL-33 stimulation (fig. S5, C and D). These results suggest that IL-33–ST2 signaling directly induces the differentiation of FcεRIα$^+$ macrophages from immature myeloid cells and preexisting macrophages.

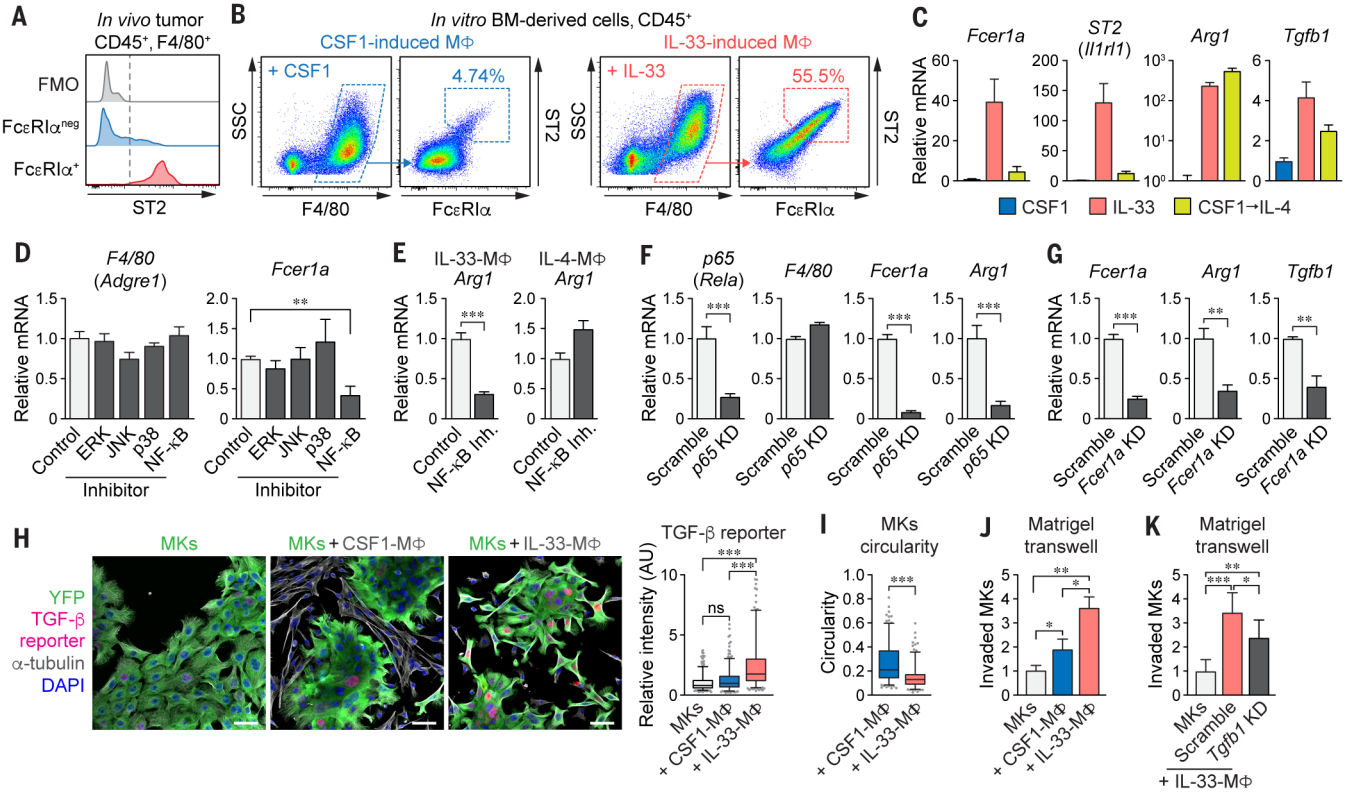


Fig. 5. IL-33-induced FcεRIα$^+$ macrophages activate paracrine TGF-β signaling and epithelial cell invasion in vitro. (**A**) Histogram presentation of ST2 expression in FcεRIα$^+$ and FcεRIα^{neg} macrophages in SCC. (**B**) Flow cytometry analysis of FcεRIα and ST2 expression in CSF1- and IL-33–induced F4/80$^+$ macrophages from bone marrow (BM)–derived cells (see fig. S5B). (**C**) qPCR analysis of CSF1-induced, IL-33–induced, and CSF1-induced→IL-4–treated macrophages in vitro (n = 3). Data are shown as relative mean expression with SD. (**D**) qPCR analysis of IL-33–induced macrophages treated with a MEK inhibitor (U0126, 0.5 μM), a JNK inhibitor (SP600125, 0.5 μM), a p38 inhibitor (SB202190, 0.5 μM), or an NF-κB inhibitor (BAY11-7082, 1 μM) during macrophage differentiation (n = 3). Data are shown as relative mean expression with SD and were analyzed with unpaired t test, $^{**}P$ = 0.0022. (**E**) qPCR analysis of IL-33–induced and IL-4–activated macrophages treated with the NF-κB inhibitor (n = 3). Data are shown as relative mean expression with SD and were analyzed with unpaired t test, $^{***}P < 0.001$. (**F** and **G**) qPCR analysis of IL-33–induced macrophages from ex vivo expanded hematopoietic progenitor cells that were LV transduced with scramble control, (F) *Rela* (NF-κB *p65*) shRNA, or (G) *Fcer1a* shRNA (see fig. S5E). (**H**) LV-transduced YFP$^+$ MKs were cocultured with CSF1- or IL-33–induced YFP^{neg} macrophages (MΦs) for 24 hours. (Graph) Quantification of TGF-β fluorescent reporter intensity in YFP$^+$ MKs. Data are shown in box-and-whisker plots (midline, median; box, 25th and 75th percentiles; whiskers, 5th and 95th percentiles with outliers); $^{***}P < 0.001$. ns, not significant. Scale bars, 50 μm. (**I**) Quantification of the circularity of YFP$^+$ MKs cocultured with CSF1- or IL-33–induced MΦs. Data are shown in box-and-whisker plots, $^{***}P < 0.001$. (**J** and **K**) Quantification of YFP$^+$ MKs invaded through Matrigel-coated membranes. (J) MKs were cultured with CSF1- or IL-33–induced MΦs or with (K) IL-33–induced MΦs transduced scramble control or *Tgfb1* shRNA (n = 3). Data are shown as mean with SD, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Scale bar, 50 μm.

Similar to FcεRIα$^+$ macrophages in vivo, IL-33–induced FcεRIα$^+$ macrophages in vitro also expressed several markers of alternative activation, e.g., *Arg1* and *Tgfb1*, at higher levels comparable to those of macrophages activated by IL-4 (Fig. 5C). However, IL-4–activated macrophages expressed significantly less FcεRIα and ST2, suggesting that IL-33 may induce FcεRIα$^+$ M2 macrophages through a specific signaling pathway. To understand the mechanism of FcεRIα$^+$ macrophage differentiation, we examined the impact of small-molecule inhibitors that target ST2 downstream signaling pathways, including extracellular signal-regulated kinase (ERK), c-Jun N-terminal kinase (JNK), p38 mitogen-activated protein kinase (MAPK), and nuclear factor κB (NF-κB) (*44*). When we used inhibitor concentrations that permit bone marrow–derived cells to differentiate into macrophages, an NF-κB inhibitor, BAY 11-7082, significantly down-regulated FcεRIα expression (Fig. 5D). The NF-κB inhibitor also repressed *Arg1* expression up-regulated by IL-33, but not by IL-4 (Fig. 5E), suggesting that the ST2–NF-κB pathway is critical for FcεRIα$^+$ macrophages. Moreover, we transduced hematopoietic progenitor cells expanded ex vivo (*45*) with LV-puro vectors containing NF-κB *p65* (*Rela*) shRNA. After puromycin selection, we induced macrophage differentiation in vitro (fig. S5E). Again, *p65* KD significantly repressed *Fcer1a* and *Arg1* expression induced by IL-33 without affecting *F4/80* expression (Fig. 5F). Furthermore, *Fcer1a* KD reduced the expression of *Arg1* and *Tgfb1* (Fig. 5G), suggesting that FcεRIα may be involved in alternative activation of macrophages induced by IL-33.

Next, we tested whether IL-33–induced macrophages could activate paracrine TGF-β signaling by coculture with YFP$^+$ mouse keratinocytes (MKs) harboring the TGF-β reporter (fig. S5F). TGF-β reporter expression (Fig. 5H) and SMAD2 phosphorylation (fig. S5G) in keratinocytes were significantly induced by IL-33–induced macrophages, and this effect was mitigated by the addition of anti–TGF-β–neutralizing antibody in the culture medium (fig. S5H). TGF-β reporter$^+$ keratinocytes cultured with IL-33–induced macrophages exhibited an elongated cell shape (Fig. 5I), suggesting an induction of invasive properties. We previously showed that TGF-β–responding tumor cells in vivo show partial epithelial-mesenchymal transition (EMT) phenotypes (*17*, *46*). Such phenotypes, including the redistribution of E-cadherin from cell–cell junctions to cytoplasmic granules and the expression of an EMT transcription factor ZEB2 (fig. S5, I and J), were recapitulated in keratinocytes cultured with IL-33–induced macrophages. To functionally assess invasive properties, we performed a transwell assay by seeding YFP$^+$ keratinocytes and YFP^{neg} macrophages on a layer of Matrigel. The number of keratinocytes that invaded to the other side of the transwell significantly increased in the presence of IL-33–induced macrophages compared with those with CSF1-induced macrophages (Fig. 5J), and this effect was mitigated by *Tgfb1* KD in IL-33–induced macrophages (Fig. 5K and fig. S5K). On the basis of these data, we conclude that IL-33 can induce the differentiation of immature myeloid cells into FcεRIα$^+$ macrophages that are capable of activating paracrine TGF-β signaling and invasive properties in epithelial cells.

TIC-derived IL-33 accumulates FcεRIα$^+$ macrophages to create a TGF-β-rich, tumor-promoting niche

Our in vitro experiments provided the mechanistic basis of potential TIC–niche interactions that drive invasive tumor progression. We then sought to determine the impact of tumor-derived IL-33 on stromal cells and the IL-33–related mechanisms on SCC progression in vivo. Flow cytometry analysis showed that the overall proportions of YFP$^+$ tumor cells, CD45$^+$ immune cells, and the double-negative cells (including fibroblasts and endothelial cells) did not significantly differ between *Il33* KD and control tumors (fig. S6A). Additionally, the frequency of F4/80$^+$ macrophages in the CD45$^+$ population was not affected by *Il33* KD in vivo (Fig. 6A). However, the proportion of FcεRIα$^+$ST2$^+$ cells in the F4/80$^+$ macrophages was significantly lower in *Il33* KD tumors compared with control tumors (Fig. 6B). Moreover, immunofluorescence analysis of *Il33* KD tumors revealed a significant reduction of FcεRIα$^+$TGF-β$^+$ cells in the stroma compared with control tumors (Fig. 6C). These results suggest that IL-33 released from the TIC population increases the density of FcεRIα$^+$TGF-β$^+$ macrophages in the adjacent stroma.

To further substantiate the function of IL-33 released from TICs toward FcεRIα$^+$ macrophage differentiation and tumor progression, we generated LV vectors expressing soluble ST2 (sST2), a decoy receptor for IL-33, to sequester extracellular IL-33 and to block ST2 signaling (*43*) (fig. S6B). In vitro, cells transduced with this LV vector secreted sST2 in the culture medium in a Dox-dependent manner (fig. S6C), and the conditioned medium of sST2-expressing cells suppressed macrophage differentiation induced by IL-33, but not by CSF1 (fig. S6D). Using this LV vector, we induced HRAS^{G12V}-driven tumors that ectopically expressed sST2 in vivo and detected a significant reduction of FcεRIα$^+$ cells in the stroma (Fig. 6D). Moreover, the growth of sST2-overexpressing tumors was stunted (Fig. 6E), and they maintained a well-differentiated status similar to that of *Il33* KD tumors (Fig. 6F), further supporting the function of IL-33 in FcεRIα$^+$ macrophage differentiation and tumor progression in vivo. We also evaluated the impact of the NRF2-mediated IL-33 release on the accumulation of FcεRIα$^+$ macrophages in the stroma by depleting *Nrf2* or *Keap1* in vivo. Consistent with our in vitro data (Fig. 2I), *Nrf2* KD tumors largely maintained nuclear IL-33 (Fig. 6G). By contrast, *Keap1* KD tumors showed cytoplasmic localization of IL-33 in tumor cells and increased numbers of FcεRIα$^+$ macrophages in the stroma.

It has been reported that elevated IL-33 serum levels are associated with various inflammatory diseases (*20*, *47*). Therefore, tumor-derived IL-33 might systemically induce myelopoiesis or differentiation of immature myeloid cells in hematopoietic organs such as the spleen and bone marrow (*48*). However, tumor-induced splenomegaly was observed at similar levels in control and *Il33* KD tumor-bearing mice (fig. S6E). Likewise, myelopoietic states in the spleen and bone marrow were similarly exacerbated in the two groups (fig. S6F). Additionally, FcεRIα$^+$ cells remained scarce (< 0.3%) in the spleen and bone marrow of tumor-bearing mice in both groups (fig. S6G). These results suggest that immature myeloid cells could be recruited into tumor tissues through IL-33–independent mechanisms, and that IL-33 acts as a short-distance cue in the TIC microenvironment to generate FcεRIα$^+$ macrophages. To test these possibilities, we injected ex vivo expanded hematopoietic progenitor cells marked by cell membrane–localized tdTomato into the circulation of tumor-bearing mice. At 7 days after injection, transplanted tdTomato$^+$ cells were detected at the tumor stroma and a fraction of them expressed FcεRIα (Fig. 6H). Moreover, *ST2* KD in tdTomato$^+$ cells (fig. S6H) did not prevent their infiltration into the stroma but rather suppressed FcεRIα expression (Fig. 6H and fig. S6I), supporting the idea that recruited immature myeloid cells are influenced by TIC-derived IL-33 to differentiate into FcεRIα$^+$ macrophages.

Finally, we sought to determine whether FcεRIα$^+$ macrophages are a critical mediator of IL-33–induced paracrine TGF-β signaling and tumor progression in vivo. We injected the anti-FcεRIα antibody into tumor-bearing mice to deplete FcεRIα$^+$ macrophages through antibody-dependent cellular cytotoxicity. At 3 days after injection, efficient depletion was confirmed by immunofluorescence analysis of tumor sections using an anti-ST2 antibody (Fig. 6I). Both the frequency and intensity of TGF-β reporter expression were significantly decreased in FcεRIα$^+$ cell-depleted tumors compared with control (Fig. 6J). By 7 days after injection, the growth of FcεRIα$^+$ cell-depleted tumors was stunted compared with control tumors (Fig. 6K). To exclude the potential effect of FcεRIα$^+$ mast cells, we also injected the anti-FcεRIα antibody into tumor-bearing, mast cell–deficient (*Kit$^{w\text{-}sh}$*) mice (*49*). We confirmed that FcεRIα$^+$ macrophages exist around

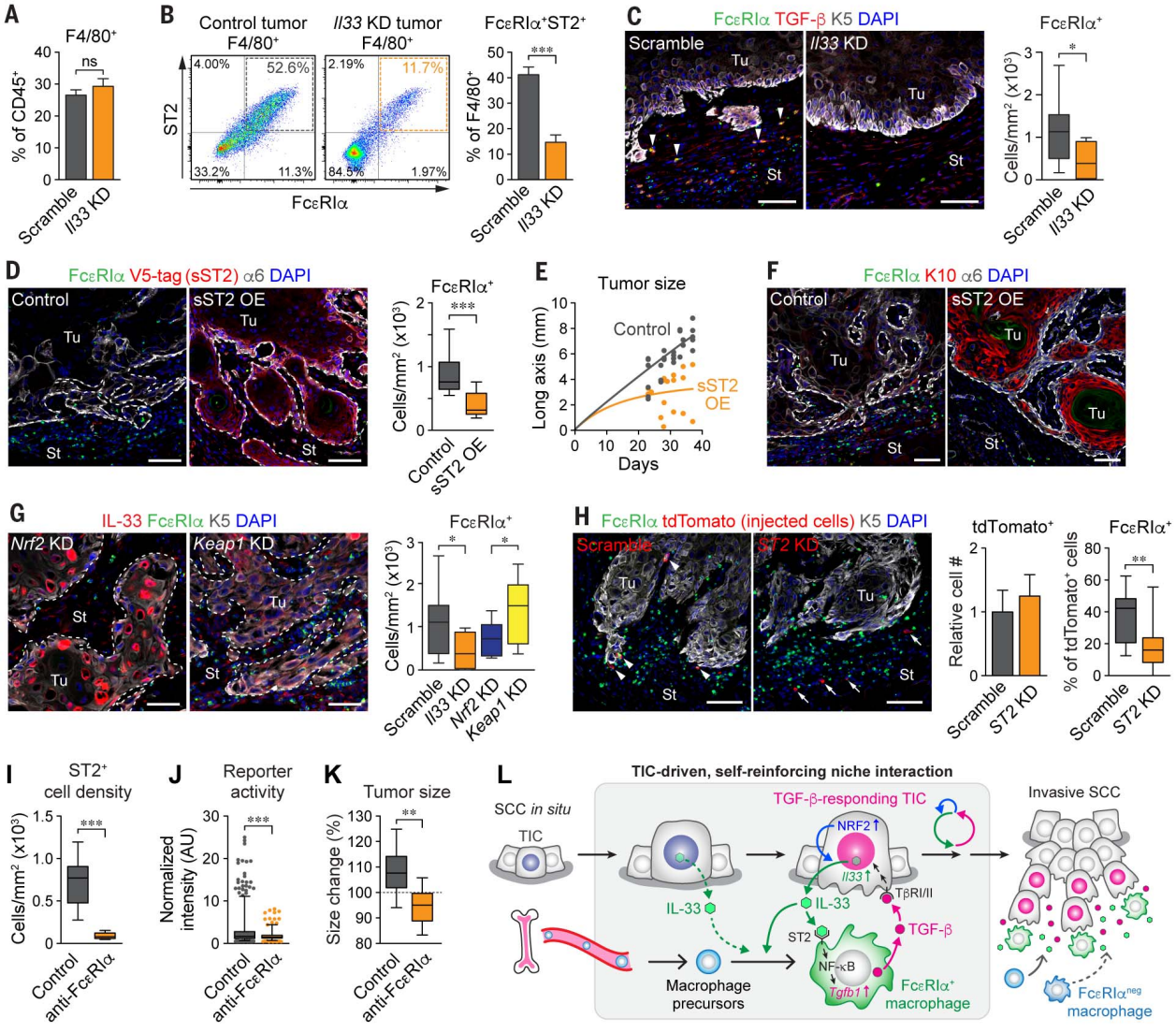


Fig. 6. FcεRIα$^+$ macrophages mediate IL-33–induced paracrine TGF-β signaling and invasive tumor progression. (**A**) No difference is observed in the frequency of F4/80$^+$ macrophages in live, CD45$^+$ cells in control and *Il33* KD tumors. Scramble control, $n = 30$; *Il33* KD, $n = 20$. Data are shown as mean with SEM. ns, not significant. (**B**) F4/80$^+$ macrophages in *Il33* KD tumors have a smaller FcεRIα$^+$ST2$^+$ double-positive population than control tumors. (Graph) Quantification of FcεRIα$^+$ST2$^+$ macrophages. Scramble control, $n = 27$; *Il33* KD, $n = 20$. Data are shown as mean with SEM and were analyzed with unpaired *t* test, ***$P < 0.001$. (**C**) Immunolabeling of tumor sections showing fewer FcεRIα$^+$TGF-β$^+$ cells in the stroma of *Il33* KD tumors compared with control. (Graph) Quantification of FcεRIα$^+$ cells in the stromal area within a 50-μm radius of tumor edges analyzed with unpaired *t* test, *$P = 0.0394$. (**D**) Immunolabeling of tumor sections showing that sST2-overexpressing (OE) tumors have fewer FcεRIα$^+$ cells in the adjacent stroma. (Graph) Quantification of FcεRIα$^+$ cells in the stromal area within a 50-μm radius of tumor edges analyzed with unpaired *t* test, ***$P < 0.001$. (**E**) HRAS^{G12V}-driven tumors overexpressing sST2 or control were sized. Control, $n = 6$; sST2 OE, $n = 6$. Approximation curves were drawn by applying the Michaelis-Menten kinetics curve. (**F**) Immunolabeling of K10 showing a sustained differentiation property in sST2-overexpressing tumors. (**G**) Immunolabeling of IL-33 and FcεRIα in *Nrf2* (*Nfe2l2*) or *Keap1* KD tumors. Whereas *Nrf2* KD tumors show mostly nuclear IL-33, *Keap1* KD tumors show cytoplasmic IL-33 localization. (Graph) Quantification of FcεRIα$^+$ cells in the stromal area within a 50-μm radius of the tumor leading edges ($n = 3$-4). Data are shown in box-and-whisker plots, *$P = 0.0394$ (control versus *Il33* KD) or 0.0372 (*Nrf2* KD versus *Keap1* KD). (**H**) Immunolabeling of tumor sections showing tdTomato$^+$ immune cells infiltrated from the circulation at 7 d after injection. Note that control tdTomato$^+$ cells expressed FcεRIα (arrowheads) but *ST2* KD cells were largely negative (arrows). Magnified images are shown in fig. S6l. (Graph) Quantification of tdTomato$^+$ cells in the stroma and FcεRIα-expressing tdTomato$^+$ cells ($n = 4$) analyzed with unpaired *t* test, **$P = 0.0019$. (**I**) FcεRIα$^+$ cell depletion by anti-FcεRIα antibody injection in vivo. Quantification of cells expressing ST2 (a surrogate marker for FcεRIα) within the stromal area in a 50-μm radius of tumor basal cells ($n = 3$), ***$P < 0.001$. (**J**) Quantification of TGF-β fluorescent reporter intensity in K5 basal cells. Intensity values in each image were normalized by the median values (set as 1) ($n = 3$) and analyzed with unpaired *t* test, ***$P < 0.001$. (**K**) Tumors were sized before and 1 week after anti-FcεRIα antibody injection. Isotype control, $n = 3$, 10 tumors; anti-FcεRIα, $n = 3$, 10 tumors. **$P = 0.0016$. (**L**) Model of TIC-driven feedforward mechanism of invasive SCC progression. TGF-β–responding TICs release IL-33 through the NRF2-mediated antioxidant response, which induces differentiation of immature myeloid cells into FcεRIα$^+$ macrophages in their close proximity. In turn, FcεRIα$^+$ macrophages send reciprocal paracrine TGF-β signaling to TICs to promote invasive progression and drug resistance of SCC, and further induce the release of IL-33, establishing a self-reinforcing niche signaling loop between TICs and FcεRIα$^+$ macrophages. Dotted lines denote the tumor–stroma boundaries. Scale bar, 50 μm.

invasive, TGF-β–responding cells in a mast cell–deficient condition and were depleted by the anti-FcεRIα antibody (fig. S6, J to L). Moreover, FcεRIα$^+$ cell depletion under this condition still exhibited growth-suppressing effects, similar to the effect seen in control tumors (fig. S6M).

Our studies provide evidence of a feedforward TIC–niche interaction model. The TIC-intrinsic, NRF2-mediated antioxidant response triggers the extracellular release of IL-33 from TGF-β–responding TICs, which induces TGF-β–rich tumor microenvironments by accumulating FcεRIα$^+$ macrophages in their close proximity. This self-reinforcing, TIC–niche signaling loop may prove to be a critical driver of invasive cancer progression and drug resistance in human SCC (Fig. 6L).

Discussion

The development of effective TIC-targeted therapies is moving at a restricted pace because of the lack of identification of TIC vulnerabilities. Some of these vulnerabilities may stem from the cross-talk between TICs and their niches, which is critical in maintaining, or potentially inducing, stem-like properties in tumor cells. Here, focusing on TGF-β–responding tumor cells in a mouse model of SCC, a previously characterized TIC population (*17*), we uncover a mechanism by which TICs generate the niche microenvironment that is required for invasive progression and drug resistance in SCC.

We identified IL-33 as the most significantly up-regulated and highly expressed cytokine in TGF-β–responding tumor cells. IL-33 is known to play multiple and sometime opposing roles in tumorigenesis (*21*). It has been reported that overexpression of IL-33 inhibits tumor growth by enhancing cytotoxic T and NK cell activation (*50*, *51*), whereas mice lacking ST2 also attenuated tumor progression by increasing T helper 1 (T_H1)/T_H17 cytokines and cytotoxic activity of NK cells (*52*). Here, we demonstrate that IL-33 has a dominant role in driving tumor progression in SCC and describe how this cytokine evokes pro-tumorigenic changes in the tumor microenvironment. In addition to IL-33, other cytokines might be regulated by TGF-β–responding tumor cells to modulate tumor-promoting and immunosuppressive microenvironments. For example, TSLP, a significantly down-regulated cytokine in TGF-β–responding cells (Fig. 1C), is known to reduce the incidence of chemically induced skin squamous carcinogenesis (*53*), and the loss of TSLP signaling accelerates malignancy by recruiting CD11b$^+$Gr1$^+$ myeloid cells (*54*). Therefore, it is likely that the function of IL-33 in tumorigenesis is tightly regulated by the cytokine milieu where IL-33–responding cells are located.

We have demonstrated that the NRF2-mediated antioxidant response activated in TGF-β–responding TICs is involved in the extracellular release of IL-33. It was previously reported that mechanical stress triggers the extracellular secretion of nuclear IL-33 (*29*), and that mechanical stress also induces reactive oxygen species (ROS) and activates the NRF2-mediated antioxidant responses (*55*, *56*). Our previous work showed that TGF-β–responding tumor cells maintain lower ROS levels compared with their nonresponding counterparts, most likely because of enhanced NRF2 activity (*17*). Although the molecular mechanism of IL-33 release remains to be clarified, our study has revealed an interesting link between this TIC-intrinsic, stress-resistant activity and spatially restricted cytokine release as a mechanism for TIC niche formation.

The cellular components and molecular pathways that are essential for building the TIC niche remain unclear; however, tumor-associated macrophages are considered to be crucial players (*40*, *57*, *58*). Thus, the mechanisms underlying the emergence of a spatial and functional relationship between TICs and tumor-associated macrophages may provide a critical target for destabilizing TICs. The discovery of FcεRIα$^+$ macrophages and the IL-33–TGF-β signaling loop that connects TICs and those macrophages is of great importance because it adds to the understanding of heterogeneity in tumor-associated macrophages and provides mechanistic insights into the development of the TIC niche. Therefore, targeting IL-33- or FcεRIα$^+$ macrophage–relevant pathways may be an attractive avenue to develop durable cancer treatments, especially when conventional approaches to modulating tumor-associated macrophages as a whole are ineffective. Recently, it has been reported that a subset of macrophages differentiated from human peripheral blood monocytes express FcεRIα and markers of alternative activation (*59*) and that IgE plays a dual role in tumor suppression and promotion (*60*, *61*). Therefore, it is becoming increasingly important to understand tumor stage–specific regulation of FcεRIα$^+$ macrophages by IgE-related immune mechanisms.

Finally, normal stem cell niches not only protect tissue-resident stem cells from differentiation and apoptotic stimuli, but also act as a safeguard against the excessive stem cell production that can lead to cancer. The TIC niche may differ in that it can be expanded to accommodate an increasing number of TICs. The TIC-driven, self-reinforcing TIC–niche interactions may underlie the ability of cancers to expand and progress. The question of whether IL-33 or TGF-β can trigger the inception of the niche signaling loop remains open. Given that IL-33 is stored in the nucleus of normal and premalignant cells, the release of IL-33 through either NRF2-dependent or -independent mechanisms might precede the deposition of TGF-β in the stroma. Moreover, IL-33 and TGF-β are initially released in less potent and latent forms, respectively, and both are activated in the extracellular space by various mechanisms (*20*, *62*). Therefore, it is likely that the establishment of a robust IL-33–TGF-β signaling loop is regulated by another layer of cellular and molecular interplay, which could also be a potential target to destabilize TIC–niche interactions.

Materials and Methods

LV construction, production, and concentration

For generation of an LV rtTA-2A-Cre construct, we replaced the puromycin cassette of pLKO.1 between BamHI and NsiI sites with an open reading flame containing synthesized cDNA of the reverse tetracycline transactivator (rtTA3), a P2A linker sequence, and a PCR-amplified cDNA of codon-improved Cre recombinase (iCre). To generate an LV vector containing TGF-β reporter, we upgraded the construct with a brighter version of nuclear red fluorescent protein, NLS-mScarlet-I (Addgene plasmid 85044) (*46*) and inserted the chicken hypersensitivity site 4 insulator (cHS4) (*47*) that improves sensitivity and transgene expression. To generate an LV-rtTA-2A-Cre–TGF-β reporter construct, we created a multicloning site sequence containing PmeI-NheI-BmtI-XbaI-SalI-MluI sites between NsiI and KpnI, and then inserted a DNA fragment containing 12× repeated SMAD-binding elements (SBE, 5′-AGCCAGACA-3′), a 113-bp minimal CMV promoter, mScarlet-I, and cHS4 insulator sequences between KpnI and NheI sites. Depending on the experiments, we also used the LV-rtTA–TGF-β reporter construct (no Cre), particularly when the Rosa-YFP expression was not ideal and when we used *K14-CreER* transgenic mice. For gene KD experiments, we used LV-puro-shRNA vectors that allowed us to select transduced cells by puromycin selection. For Dox-inducible sST2 expression, we cloned the sST2 coding sequence under the tetO promoter and inserted this fragment by replacing the TGF-β reporter cassette. DNA plasmid mini/midi/maxi-prep, gel extraction, and PCR clean-up kits from Zymo Research were used. Large-scale production of VSV-G pseudotyped LV was performed by calcium phosphate transduction of 293T cells (Clontech/Takara, 632180) plated on poly-L-lysine–coated cell culture dishes with pLKO.1 and helper plasmids: pMD2.G and psPAX2 (Addgene plasmid, 12259 and 12260). Viral supernatant was collected 30 to 36 hours after transfection, filtered through a 0.45-μm polyvinylidene difluoride (PVDF) filter (Millipore), concentrated first by passing through a Vivacell 70/100 centrifugal concentrator (100-kDa molecular weight cut-off, Sartorius, VS6041), and then by ultracentrifugation in an Ultra-Clear tube (Beckman Coulter, 344057) for 90 min in an MLS-50 rotor at 45,000 rpm or a SW-55 Ti rotor at 42,000 rpm (Beckman Coulter).

Mice, ultrasound-guided in utero microinjection, and tumor formation analysis

The following mice were used: *Tg(TetO-HRAS)65Lc/Nci* (*TetO-Hras*^{G12V}, NCI Mouse Repository, donated by L. Chin) (*23*), *Gt(ROSA)26Sor*$^{tm(EYFP)Cos}$*/J* (*Rosa26-lox-stop-lox(LSL)-EYFP*, The Jackson Laboratory, donated by F. Costantini) (*24*), *B6;129-Tgfbr2*^{tm1Karl}*/J* (*Tgfbr2-floxed*, shared by S. Karlsson), *Tg(KRT14-cre/ERT)20Efu/J (K14-CreER*, shared by E. Fuchs), *Gt(ROSA)26Sor*$^{tm4(ACTB-tdTomato,-EGFP)Luo}$*/J* (*mT/mG*, The Jackson Laboratory, donated by L. Luo), and *B6.Cg-Kit*$^{W-sh}$*/HNihrJaeBsmJ* (*Kit*$^{Wsh/Wsh}$, The Jackson Laboratory). Mice were backcrossed into the FVB/NJ for more than five generations or the C57BL/6J for 10 generations. Ultrasound-guided in utero microinjection was performed as described previously (21). Briefly, females at embryonic day 9.5 of gestation were anesthetized with isoflurane on a heated stage before injection. Then, 300 to 500 nl of LV suspension was injected into each embryo's amniotic sac. Surgical procedures were limited to 30 min to maintain high survival rates. Postnatally, the LV-transduced mice were given mouse chow containing Dox (2 mg/kg, 5AKR, TestDiet), which activates rtTA to induce the expression of oncogenic HRAS^{G12V} from the *TetO-Hras* transgene. Cisplatin (Sigma-Aldrich, P4394) was dissolved in saline at 0.5 mg/ml and administrated by intraperitoneal injection (10 mg/kg). For cell proliferation assessment in vivo, EdU (25 μg/g, Invitrogen, A10044) was injected intraperitoneally 4 hours before lethal administration of CO_2. For FcεRIα$^+$ macrophage depletion, an in vivo grade anti-FcεRIα antibody (Tonbo Biosciences, MAR-1) was subcutaneously injected twice a week (100 μg/injection), and an isotype control antibody (Bio X Cell, Armenian hamster IgG) was used in the control cohorts. For intradermal orthotropic injection, 1×10^6 cells (*TetO-Hras*^{G12V}) transduced with the LV-rtTA–TGF-β reporter with Matrigel (Cultrex, type 3, R&D systems) were injected. For inoculation of macrophage progenitor cells in the circulation, ex vivo expanded tdTomato$^+$ hematopoietic progenitor cells (2×10^5 cells) were suspended in phosphate-buffered saline (PBS, 50 μl) and injected into the retro-orbital sinus of tumor-bearing mice. The Oregon Health & Science University (OHSU) Animal Care and Use Committee approved the animal experimentation protocols used in this study.

Tissue harvest, sectioning, and immunofluorescence microscope imaging

Because of humane end points based on tumor size, number, and location on the body surface, mice were euthanized at different time points between 23 and 58 days after Dox-induced HRAS^{G12V} expression. Before tumor collection, dorsal hairs were removed with either an electric hair clipper or hair removal cream. Mice were then photographed, and tumor size was measured with a ruler. Tumors were dissected and fixed with 4% paraformaldehyde (16% solution, Electron Microscopy Sciences) in PBS for 15 min at room temperature. After washing with PBS overnight, tumor tissues were embedded and cryo-frozen in Tissue-Tek OCT compound (Sakura, 4583) in a cryomold (Sakura, 4557). Then, 5- to 10-μm cryosections were cut with a Cryostat (Leica, CM1850), mounted on SuperFrost Plus slides (Fisherbrand), and permeabilized for 30 min in 0.3% Triton X-100 in PBS. Tissues to be stained using the TSA Plus signal amplification kit (PerkinElmer, NEL744001KT) were treated with 1% H_2O_2 for 30 min at room temperature. When immunolabeling with mouse antibodies, sections were incubated with the M.O.M. blocking kit according to the manufacturer's instructions (Vector Laboratories, BMK-2202). The following primary antibodies were used: chicken anti-GFP (to detect YFP) (Abcam, ab13970), chicken anti-keratin 5 (Covance, SIG-3475), rabbit anti–TGF-β (R&D Systems, AB-100-NA), rabbit anti–phospho-SMAD2 (Cell Signaling Technologies, 138D4), goat anti–IL-33 (R&D Systems, AF3626), rabbit anti-NRF2 (MBL International, PM069), rat anti-CD45 (BD Biosciences, 30-F11), rat anti-CD11b (BD Biosciences, M1/70), rat anti-CD3 (BD Biosciences, 17A2), rat anti-CD8a (eBioscience, 4SM15), rat anti-CD4 (eBioscience, 4SM95), rat Ly6G (BioLegend, HK1.4), rat anti-F4/80 (Bio-Rad, Cl:A3-1), Armenian hamster anti-FcεRIα (BioLegend, MAR-1), rat anti-integrin α6 (BD Biosciences, GoH3), mouse anti–α-tubulin (Sigma-Aldrich, DM1A), rabbit anti–keratin 14 (Abcam, EPR17350), rabbit anti–keratin 10 (Abcam, EP1607IHCY), goat anti-CD31 (R&D Systems, AF3628), goat antitryptase/MCP-6 (R&D Systems, AF3736), goat antihuman IL-33 (R&D Systems, AF3025), rabbit anti-CD206 (Abcam, ab64693), rabbit anti–E-cadherin (Cell Signaling Technologies, 24E10), rabbit anti SIP1/ZEB2 (Bethyl, IHC-00691), and rabbit anti–V5-tag (Abcam, ab9116). Sections were treated with primary antibody mixes and incubated at 4°C overnight. After washing with PBS with 0.1% Triton X-100 (PBST), sections were treated for 30 min at room temperature with secondary antibodies conjugated with Alexa Fluor 488, 546, or 647 (Jackson ImmunoResearch and ThermoFisher Scientific). Slides were washed, counterstained with 4'6'-diamidino-2-phenylindole (DAPI), and mounted in Prolong Gold (ThermoFisher Scientific). Imaging was performed on a Zeiss Axio Observer.Z1 microscope equipped with Apotome.2 using ×10/0.45, ×20/0.8, or ×40/0.95 air. Images were collected using Zeiss ZEN software and processed using an image-processing package, Fiji (https://fiji.sc/). For reporter intensity analysis, nuclei were selected and mean intensity per area was calculated after background subtraction. For immune cell quantification, images were framed to capture regions of the tumor–stroma interface containing high (>30%)- and low (<30%)-reporter–positive areas of the tumors. This was done to minimize the effects of variation in image quality or settings between regions being compared. Using ImageJ's ruler and measurement tools, a continuous area extending 50 μm radially from the border of the tumor into the stroma was manually marked out for quantification for both the high- and low-reporter regions. For reporter-positive regions, the area began at the first reporter-positive cell and ended once there were no more reporter-positive cells within 50 μm. For reporter-negative regions, the area was set to capture as much area as possible without coming within 50 μm of the reporter-positive region. The Alexa Fluor 647/Cy5 channel was used for most of the immune cell markers to minimize autofluorescence. Immune marker–positive cells were quantified by counting stromal cells that stained with both the immune cell marker and had clear, single-nucleus DAPI staining. To analyze circularity, a mask was created for the YFP channel, individual cells were selected with the "analyze particles" function, and then circularity [4π(area/perimeter^2)] was measured.

In vitro cell culture experiments

Keratinocytes isolated from neonatal *TetO-Hras*^{G12V}*; Rosa26-LSL-YFP* transgenic mouse epidermis were maintained in E medium with 10% fetal bovine serum (FBS) and 50 μM $CaCl_2$ at 37°C with 7.5% CO_2. For LV transduction in culture, cells were plated in six-well plates at 1.0×10^5 cells per well, incubated with viruses in the presence of polybrene (10 μg/ml) for 30 min at 37°C and 7.5% CO_2, and then plates were spun at 1100*g* for 30 min at 37°C in an Eppendorf 5810R centrifuge. After centrifugation, virus-containing medium was removed and replaced with fresh growth medium. For H_2O_2-induced oxidative stress, 100 to 500 μM H_2O_2 was added to the culture medium. For bone marrow–derived cell culture, bone marrow cells in femurs from adult mice (7 to 10 weeks old) were collected by flushing with Dulbecco's modified Eagle's medium (DMEM) with 10% FBS, and filtered through a 40-μm cell strainer. The EasySep mouse CD117 (c-Kit) positive selection kit (STEMCELL Technologies, 18757) was used to enrich c-Kit$^+$ hematopoietic stem/progenitor cells that were cultured in DMEM with 10% FBS at 37°C and 7.5% CO_2. For ex vivo expansion of c-Kit$^+$ cells, we followed the culture protocol described in (*45*). For stimulation experiments, medium was supplemented with either recombinant murine TGF-β1 (50-100 pM, R&D Systems, 7666-MB), recombinant murine IL-33 (50 ng/ml, PeproTech, 210-33), recombinant murine M-CSF/CSF1 (10 ng/ml, PeproTech, 315-02), or recombinant murine IL-4 (10 ng/ml, PeproTech, 214-14). The anti–TGF-β–neutralizing

antibodies (R&D Systems, AB-100-NA) were used for blocking experiments. For immunolabeling and visualization of in vitro culture cells, cells were seeded onto coverslips and fixed with prewarmed 4% paraformaldehyde for 10 min at 37°C. Cells were permeabilized with 0.3% Triton X-100 in PBS and incubated with the primary antibodies listed above at 4°C overnight. After washing with PBST, cells were treated for 30 min at room temperature with secondary antibodies conjugated with Alexa Fluor 488, 546, or 647. Coverslips were washed, counterstained with DAPI, and mounted in Prolong Gold.

FACS and flow cytometry

Tumors were dissected from mice, minced by scalpels in a chilled dish on ice, and treated with 0.2 mg/ml collagenase (Sigma-Aldrich C2674 or Roche 10103586001) in Hank's balanced salt solution with 10 mM Hepes for 30 min at 37°C or overnight at 4°C. Mouse back skin (the dermis side) was also digested in the same collagenase solution. The collagenase-digested tissue fraction was vigorously pipetted up and down by adding ice-cold PBS and pelleted at 400*g* for 15 min at 4°C. Precipitated tissue fractions were resuspended in 10× TrypLE (Invitrogen, A12177) to dissociate into single cells at 37°C for 15 min, and filtered through 100-, 70-, and 40-μm cell strainers. Cells were pelleted at 400*g* for 10 min, and precipitated cells were washed with staining buffer (0.1% bovine serum albumin in PBS) and pelleted again at 400*g* for 5 min. Total cell number was counted with a Countess Automated Cell Counter (Invitrogen), and cell density was adjusted to 2 million cells per 100 μl. Cells were blocked with 1 μg/ml TruStain FcX (anti-mouse CD16/32) antibody in staining buffer for 15 min at 4°C. To stain for ST2, biotinylated anti-ST2 antibody (1:1000, MD Biosciences, DJ8) was used. After removing unbound antibodies, cells were resuspended in staining buffer and incubated with antibodies for cell surface markers and Live/Dead Aqua (1:250, Invitrogen, L34957) for 30 min at 4°C. The following antibodies were used: CD49f/integrin α6-BV421 (1:250, BD Biosciences, clone GoH3), CD45-PerCP/Cy5.5 (1:100, BD Biosciences, clone 30-F11), F4/80-APC (1:200, BioLegend, clone BM8), FcεRIα-PE/Cy7 (1:2000, BioLegend, clone MAR-1), CD326/EpCAM-FITC (1:200, eBioscience, clone G8.8), MHCII-PerCP (1:200, BioLegend, clone M5/114.15.2), Ly6C-PerCP/Cy5.5 (1:400, eBioscience, clone HK1.4), Ly6G-APC/Cy7 (1:400, BD Biosciences, clone 1A8), CD11c-BV605 (1:200, BioLegend, clone N418), CD11b-BV650 (1:1000, BD Biosciences, clone M1/70), Siglec-F-BV711 (1:250, BD Biosciences, clone E50-2440), CD45-BV786 (1:2000, BD Biosciences, clone 30-F11), and CD117/c-Kit-PE (1:500, BioLegend, clone ACK2). Cell isolations were performed on a BD Influx cell sorter equipped with BD FACS software (BD Biosciences). Sorted cells were used for RNA preparation, RNA-seq, and real-time qPCR. Flow cytometry analysis was performed on a BD LSRFortessa Cell analyzer (BD Biosciences).

RNA purification, RNA-seq, and real-time qPCR

Total RNA from FACS-purified cells directly sorted into TRI Reagent (Zymo Research) was purified using the Direct-zol RNA MicroPrep Kit (Zymo Research) per the manufacturer's instructions. The quality of the total RNA for sequencing was determined with a bioanalyzer, and all samples used had RNA Integrity Score numbers above the standard required value. Library preparation using the Illumina TruSeq mRNA sample preparation kit was performed at the OHSU Massively Parallel Sequencing Shared Resource (MPSSR), and RNAs were single-end sequenced on Illumina HiSeq 2500 machines. Alignment of reads was done using the STAR aligner with the mm10 build of the mouse genome. Transcript assembly and differential expression was determined using DESeq2 with Refseq mRNAs to guide assembly. For reverse transcription (RT)–qPCR, equivalent amounts of RNA were reverse-transcribed using the SuperScript IV VILO cDNA synthesis kit (Invitrogen, 11756050). cDNAs were mixed with the indicated primers and PowerUp SYBR Green Master Mix (Applied Biosystems, A25742), and RT-qPCR was performed on a ViiA 7 Real-time PCR System (Applied Biosystems). cDNAs were normalized to equal amounts using primers against *Actb*, *Gapdh*, *or Ppib*. The following primer sequences were used (5′-3′): *Actb* forward: GGCTGTATTCCCCTCCATCG, *Actb* reverse: CCAGTTGGTAACAATGCCATGT; *Fcer1a* forward: GAGTGCCACCGTTCAAGACA, *Fcer1a* reverse: GTAGATCACCTTGCGGACATTC; *Nos2* forward: ACATCGACCCGTCCACAGTAT, *Nos2* reverse: CAGAGGGGTAGGCTTGTCTC; *Arg1* forward: CTCCAAGCCAAAGTCCTTAGAG, *Arg1* reverse: AGGAGCTGTCATTAGGGACATC; *Tgfb1* forward: GAGCCCGAAGCGGACTACTATG, *Tgfb1* reverse: CAGCCACTGCCGTACAACTCC; *Il33* forward: ACTGCATGAGACTCCGTTCTG, *Il33* reverse: CCTAGAATCCCGTGGATAGGC; *Adgre1* (*F4/80*) forward: CTTTGGCTATGGGCTTCCAGTC, *Adgre1* (*F4/80*) reverse: GCAAGGAGGACAGAGTTTATCGTG; *Il1rl1* (*ST2*) forward: TGACACCTTACAAAACCCGGA, *Il1rl1* (*ST2*) reverse: AGGTCTCTCCCATAAATGCACA; *Nfe2l2* (*Nrf2*) forward: TCTTGGAGTAAGTCGAGAAGTGT, *Nfe2l2* (*Nrf2*) reverse: GTTGAAACTGAGCGAAAAAGGC; *Keap1* forward: TGCCCCTGTGGTCAAAGTG, *Keap1* reverse: GGTTCGGTTACCGTCCTGC; *Rela* (*p65*) forward: ACTGCCGGGATGGCTACTAT, *Rela* (*p65*) reverse: TCTGGATTCGCTGGCTAATGG; *Hmox1* forward: AAGCCGAGAATGCTGAGTTCA, *Hmox1* reverse: GCCGTGTAGATATGGTACAAGGA; *Me2* forward: TCAACAAGGACTTGGCTTTTACT, *Me2* reverse: TGCAGGTCCATTAACAGGAGAT; *Krt14* forward: AGCGGCAAGAGTGAGATTTCT, *Krt14* reverse: CCTCCAGGTTATTCTCCAGGG; *Ptprc* (*CD45*) forward: GTTTTCGCTACATGACTGCACA, *Ptprc* (*CD45*) reverse: AGGTTGTCCAACTGACATCTTTC; *Pdgfra* (*CD140a*) forward: TCCATGCTAGACTCAGAAGTCA, *Pdgfra* (*CD140a*) reverse: TCCCGGTGGACACAATTTTTC; and *Pecam1* (*CD31*) forward: ACGCTGGTGCTCTATGCAAG, *Pecam1* (*CD31*) reverse: TCAGTTGCTGCCCATTCATCA.

shRNA sequences for gene KD

For gene KD experiments, DNA oligos of following target sequences were synthesized (5′-3′): the scramble shRNA: CAACAAGATGAAGAGCACCAA; *Il33* sh1: GCATCCAAGGAACTTCACTTT; *Il33* sh2: GCCAATAGAATGGGATCTCAT; *Nfe2l2* (*Nrf2*) sh1: CCAAAGCTAGTATAGCAATAA; *Nfe2l2* (*Nrf2*) sh2: GCCTTACTCTCCCAGTGAATA; *Keap1* sh1: GCCCAATTCATGGCTCACAAA; *Keap1* sh2: GGATGATCACACCGATGAATA; *Rela* (*p65*) sh: AGAAGACATTGAGGTGTATTT; *Fcer1a* sh: CCCACCATGGATTAGAATATT; *Tgfb1* sh: GCTGCGCTTGCAGAGATTAAA; and *Il1rl1* (*ST2*) sh: GCTGCAATATCCCTGATTATT.

Protein gel electrophoresis and Western blotting

Total cell lysates were prepared using radioimmunoprecipitation assay buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.5% deoxycholate, 0.1% SDS) supplemented with protease inhibitors (Halt Protease Inhibitor, Thermo Scientific, 78430), or used directly with 1X Bolt LDS Sample Buffer (Invitrogen, B0007). The protein concentrations of clarified supernatants were measured with the BCA Protein Assay Kit (Pierce, 23227). Proteins in lysates were separated by Bolt 4 to 12% Bis-Tris Plus gels (Invitrogen) and transferred to a PVDF membrane (Millipore, Immobilon-FL, IPFL00010) using a Pierce Power Blotter XL System (Invitrogen, PB0013). Transferred membranes were blocked for 1 hour in Odyssey blocking buffer (LI-COR, 927-50000) diluted 1:1 in Tris-buffered saline (TBS, containing 20 mM Tris-HCl, pH 7.4, 150 mM NaCl), and then incubated with primary antibodies in the blocking buffer or TBS overnight at 4°C. After washing with TBS, membranes were incubated with secondary antibodies in Odyssey blocking buffer diluted 1:1 in TBS for 1 hour at room temperature. Membranes were washed in TBS and imaged on an Azure c600 Infra-Red Imaging System (Azure Biosystems). Primary antibodies used were as follows: anti–IL-33 (1:500, goat polyclonal, R&D Systems, AF3626), anti–phospho-SMAD2 (1:2000, rabbit monoclonal, Cell Signaling Technology, 3108), anti-SMAD2/3 (1:2000, mouse monoclonal, BD Biosciences, 610842), anti-NRF2 (1:3000, rabbit polyclonal, MBL International, PM069), and anti–α-tubulin (1:3000, mouse monoclonal, Sigma-Aldrich, DM1A). IRDye680- or IRDye800-conjugated secondary antibodies were used (1:10,000, LI-COR).

REFERENCES AND NOTES

1. D. Nassar, C. Blanpain, Cancer stem cells: Basic concepts and therapeutic implications. *Annu. Rev. Pathol.* **11**, 47–76 (2016). doi: 10.1146/annurev-pathol-012615-044438; pmid: 27193450
2. E. Batlle, H. Clevers, Cancer stem cells revisited. *Nat. Med.* **23**, 1124–1134 (2017). doi: 10.1038/nm.4409; pmid: 28985214
3. K. A. Moore, I. R. Lemischka, Stem cells and their niches. *Science* **311**, 1880–1885 (2006). doi: 10.1126/science.1110542; pmid: 16574858
4. A. Kreso, J. E. Dick, Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–291 (2014). doi: 10.1016/j.stem.2014.02.006; pmid: 24607403
5. M. R. Junttila, F. J. de Sauvage, Influence of tumour microenvironment heterogeneity on therapeutic response. *Nature* **501**, 346–354 (2013). doi: 10.1038/nature12626; pmid: 24048067
6. B. C. Prager, Q. Xie, S. Bao, J. N. Rich, Cancer stem cells: The architects of the tumor ecosystem. *Cell Stem Cell* **24**, 41–53 (2019). doi: 10.1016/j.stem.2018.12.009; pmid: 30609398
7. S. I. Grivennikov, F. R. Greten, M. Karin, Immunity, inflammation, and cancer. *Cell* **140**, 883–899 (2010). doi: 10.1016/j.cell.2010.01.025; pmid: 20303878
8. D. Hanahan, L. M. Coussens, Accessories to the crime: Functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**, 309–322 (2012). doi: 10.1016/j.ccr.2012.02.022; pmid: 22439926
9. J. Galon *et al*., Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960–1964 (2006). doi: 10.1126/science.1129139; pmid: 17008531
10. D. G. DeNardo *et al*., Leukocyte complexity predicts breast cancer survival and functionally regulates response to chemotherapy. *Cancer Discov.* **1**, 54–67 (2011). doi: 10.1158/2159-8274.CD-10-0028; pmid: 22039576
11. M. Schober, E. Fuchs, Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-β and integrin/focal adhesion kinase (FAK) signaling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10544–10549 (2011). doi: 10.1073/pnas.1107807108; pmid: 21670270
12. G. Driessens, B. Beck, A. Caauwe, B. D. Simons, C. Blanpain, Defining the mode of tumour growth by clonal analysis. *Nature* **488**, 527–530 (2012). doi: 10.1038/nature11344; pmid: 22854777
13. W. Cui *et al*., TGFbeta1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. *Cell* **86**, 531–542 (1996). doi: 10.1016/S0092-8674(00)80127-0; pmid: 8752208
14. N. Oshimori, E. Fuchs, The harmonies played by TGF-β in stem cell biology. *Cell Stem Cell* **11**, 751–764 (2012). doi: 10.1016/j.stem.2012.11.001; pmid: 23217421
15. M. Pickup, S. Novitskiy, H. L. Moses, The roles of TGFβ in the tumour microenvironment. *Nat. Rev. Cancer* **13**, 788–799 (2013). doi: 10.1038/nrc3603; pmid: 24132110
16. C. J. David, J. Massagué, Contextual determinants of TGFβ action in development, immunity and cancer. *Nat. Rev. Mol. Cell Biol.* **19**, 419–435 (2018). doi: 10.1038/s41580-018-0007-0; pmid: 29643418
17. N. Oshimori, D. Oristian, E. Fuchs, TGF-β promotes heterogeneity and drug resistance in squamous cell carcinoma. *Cell* **160**, 963–976 (2015). doi: 10.1016/j.cell.2015.01.043; pmid: 25723170
18. A. Calon *et al*., Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. *Cancer Cell* **22**, 571–584 (2012). doi: 10.1016/j.ccr.2012.08.013; pmid: 23153532
19. E. Richardsen, R. D. Uglehus, S. H. Johnsen, L.-T. Busund, Immunohistochemical expression of epithelial and stromal immunomodulatory signalling molecules is a prognostic indicator in breast cancer. *BMC Res. Notes* **5**, 110 (2012). doi: 10.1186/1756-0500-5-110; pmid: 22353218
20. F. Y. Liew, J.-P. Girard, H. R. Turnquist, Interleukin-33 in health and disease. *Nat. Rev. Immunol.* **16**, 676–689 (2016). doi: 10.1038/nri.2016.95; pmid: 27640624
21. M.-H. Wasmer, P. Krebs, The role of IL-33-dependent inflammation in the tumor microenvironment. *Front. Immunol.* **7**, 682 (2017). pmid: 28119694
22. S. Beronja, G. Livshits, S. Williams, E. Fuchs, Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nat. Med.* **16**, 821–827 (2010). doi: 10.1038/nm.2167; pmid: 20526348
23. L. Chin *et al*., Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468–472 (1999). doi: 10.1038/22788; pmid: 10440378
24. S. Srinivas *et al*., Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4 (2001). doi: 10.1186/1471-213X-1-4; pmid: 11299042
25. M. B. Baghdadi *et al*., Reciprocal signalling by Notch-Collagen V-CALCR retains muscle stem cells in their niche. *Nature* **557**, 714–718 (2018). doi: 10.1038/s41586-018-0144-9; pmid: 29795344
26. Y. Ge, E. Fuchs, Stretching the limits: From homeostasis to stem cell plasticity in wound healing and cancer. *Nat. Rev. Genet.* **19**, 311–325 (2018). doi: 10.1038/nrg.2018.9; pmid: 29479084
27. P. Levéen *et al*., Induced disruption of the transforming growth factor β type II receptor gene in mice causes a lethal inflammatory disorder that is transplantable. *Blood* **100**, 560–568 (2002). doi: 10.1182/blood.V100.2.560; pmid: 12091349
28. V. Vasioukhin, L. Degenstein, B. Wise, E. Fuchs, The magical touch: Genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8551–8556 (1999). doi: 10.1073/pnas.96.15.8551; pmid: 10411913
29. R. Kakkar, H. Hei, S. Dobner, R. T. Lee, Interleukin 33 as a mechanically responsive cytokine secreted by living cells. *J. Biol. Chem.* **287**, 6941–6948 (2012). doi: 10.1074/jbc.M111.298703; pmid: 22215666
30. M. Rojo de la Vega, E. Chapman, D. D. Zhang, NRF2 and the hallmarks of cancer. *Cancer Cell* **34**, 21–43 (2018). doi: 10.1016/j.ccell.2018.03.022; pmid: 29731393
31. V. Bronte *et al*., Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nat. Commun.* **7**, 12150 (2016). doi: 10.1038/ncomms12150; pmid: 27381735
32. J. C. Marie, J. J. Letterio, M. Gavin, A. Y. Rudensky, TGF-β1 maintains suppressor function and Foxp3 expression in CD4+CD25+ regulatory T cells. *J. Exp. Med.* **201**, 1061–1067 (2005). doi: 10.1084/jem.20042276; pmid: 15809351
33. L. M. Coussens *et al*., Inflammatory mast cells up-regulate angiogenesis during squamous epithelial carcinogenesis. *Genes Dev.* **13**, 1382–1397 (1999). doi: 10.1101/gad.13.11.1382; pmid: 10364156
34. S. M. Ghouse *et al*., Although abundant in tumor tissue, mast cells have no effect on immunological micro-milieu or growth of HPV-induced or transplanted tumors. *Cell Rep.* **22**, 27–35 (2018). doi: 10.1016/j.celrep.2017.12.010; pmid: 29298428
35. S. Aras, M. R. Zaidi, TAMeless traitors: Macrophages in cancer progression and metastasis. *Br. J. Cancer* **117**, 1583–1591 (2017). doi: 10.1038/bjc.2017.356; pmid: 29065107
36. A. Mantovani, F. Marchesi, A. Malesci, L. Laghi, P. Allavena, Tumour-associated macrophages as treatment targets in oncology. *Nat. Rev. Clin. Oncol.* **14**, 399–416 (2017). doi: 10.1038/nrclinonc.2016.217; pmid: 28117416
37. P. J. Murray *et al*., Macrophage activation and polarization: Nomenclature and experimental guidelines. *Immunity* **41**, 14–20 (2014). doi: 10.1016/j.immuni.2014.06.008; pmid: 25035950
38. E. L. Abel, J. M. Angel, K. Kiguchi, J. DiGiovanni, Multi-stage chemical carcinogenesis in mouse skin: Fundamentals and applications. *Nat. Protoc.* **4**, 1350–1362 (2009). doi: 10.1038/nprot.2009.120; pmid: 19713956
39. W. Zhou *et al*., Periostin secreted by glioblastoma stem cells recruits M2 tumour-associated macrophages and promotes malignant growth. *Nat. Cell Biol.* **17**, 170–182 (2015). doi: 10.1038/ncb3090; pmid: 25580734
40. V. Plaks, N. Kong, Z. Werb, The cancer stem cell niche: How essential is the niche in regulating stemness of tumor cells? *Cell Stem Cell* **16**, 225–238 (2015). doi: 10.1016/j.stem.2015.02.015; pmid: 25748930
41. M. Fang *et al*., IL33 promotes colon cancer cell stemness via JNK activation and macrophage recruitment. *Cancer Res.* **77**, 2735–2745 (2017). doi: 10.1158/0008-5472.CAN-16-1602; pmid: 28249897
42. Z. He *et al*., Epithelial-derived IL-33 promotes intestinal tumorigenesis in Apc $^{Min/+}$ mice. *Sci. Rep.* **7**, 5520–5524 (2017). doi: 10.1038/s41598-017-05716-z; pmid: 28710436
43. R. Kakkar, R. T. Lee, The IL-33/ST2 pathway: Therapeutic target and novel biomarker. *Nat. Rev. Drug Discov.* **7**, 827–840 (2008). doi: 10.1038/nrd2660; pmid: 18827826
44. J. Schmitz *et al*., IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. *Immunity* **23**, 479–490 (2005). doi: 10.1016/j.immuni.2005.09.015; pmid: 16286016
45. A. C. Wilkinson *et al*., Long-term *ex vivo* haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* **571**, 117–121 (2019). doi: 10.1038/s41586-019-1244-x; pmid: 31142833
46. M. A. Nieto, R. Y.-J. Huang, R. A. Jackson, J. P. Thiery, EMT: 2016. *Cell* **166**, 21–45 (2016). doi: 10.1016/j.cell.2016.06.028; pmid: 27368099
47. A. Mitsui *et al*., Serum IL-33 levels are increased in patients with psoriasis. *Clin. Exp. Dermatol.* **41**, 183–189 (2016). doi: 10.1111/ced.12670; pmid: 25976154
48. D. I. Gabrilovich, S. Ostrand-Rosenberg, V. Bronte, Coordinated regulation of myeloid cells by tumours. *Nat. Rev. Immunol.* **12**, 253–268 (2012). doi: 10.1038/nri3175; pmid: 22437938
49. M. A. Grimbaldeston *et al*., Mast cell-deficient *W-sash c-kit* mutant *Kit* $^{W\text{-}sh/W\text{-}sh}$ mice as a model for investigating mast cell biology in vivo. *Am. J. Pathol.* **167**, 835–848 (2005). doi: 10.1016/S0002-9440(10)62055-X; pmid: 16127161
50. X. Gao *et al*., Tumoral expression of IL-33 inhibits tumor growth and modifies the tumor microenvironment through CD8$^+$ T and NK cells. *J. Immunol.* **194**, 438–445 (2015). doi: 10.4049/jimmunol.1401344; pmid: 25429071
51. J. Kim *et al*., Intratumorally establishing type 2 innate lymphoid cells blocks tumor growth. *J. Immunol.* **196**, 2410–2423 (2016). doi: 10.4049/jimmunol.1501730; pmid: 26829987
52. I. Jovanovic *et al*., ST2 deletion enhances innate and acquired immunity to murine mammary carcinoma. *Eur. J. Immunol.* **41**, 1902–1912 (2011). doi: 10.1002/eji.201141417; pmid: 21484786
53. S. Demehri *et al*., Elevated epidermal thymic stromal lymphopoietin levels establish an antitumor environment in the skin. *Cancer Cell* **22**, 494–505 (2012). doi: 10.1016/j.ccr.2012.08.017; pmid: 23079659
54. M. Di Piazza, C. S. Nowell, U. Koch, A.-D. Durham, F. Radtke, Loss of cutaneous TSLP-dependent immune responses skews the balance of inflammation from tumor protective to tumor promoting. *Cancer Cell* **22**, 479–493 (2012). doi: 10.1016/j.ccr.2012.08.016; pmid: 23079658
55. R. Aikawa *et al*., Reactive oxygen species in mechanical stress-induced cardiac hypertrophy. *Biochem. Biophys. Res. Commun.* **289**, 901–907 (2001). doi: 10.1006/bbrc.2001.6068; pmid: 11735132
56. S.-K. Lee *et al*., Mechanical stress activates proinflammatory cytokines and antioxidant defense enzymes in human dental pulp cells. *J. Endod.* **34**, 1364–1369 (2008). doi: 10.1016/j.joen.2008.08.024; pmid: 18928848
57. J. B. Mitchem *et al*., Targeting tumor-infiltrating macrophages decreases tumor-initiating cells, relieves immunosuppression, and improves chemotherapeutic responses. *Cancer Res.* **73**, 1128–1141 (2013). doi: 10.1158/0008-5472.CAN-12-2731; pmid: 23221383
58. H. Lu *et al*., A breast cancer stem cell niche supported by juxtacrine signalling from monocytes and macrophages. *Nat. Cell Biol.* **16**, 1105–1117 (2014). doi: 10.1038/ncb3041; pmid: 25266422
59. G. Pellizzari *et al*., IgE re-programs alternatively-activated human macrophages towards pro-inflammatory anti-tumoural states. *EBioMedicine* **43**, 67–81 (2019). doi: 10.1016/j.ebiom.2019.03.080; pmid: 30956175
60. G. Crawford *et al*., Epithelial damage and tissue γδ T cells promote a unique tumor-protective IgE response. *Nat. Immunol.* **19**, 859–870 (2018). doi: 10.1038/s41590-018-0161-8; pmid: 30013146
61. M. D. Hayes *et al*., Inflammation-induced IgE promotes epithelial hyperplasia and tumour growth. *eLife* **9**, e51862 (2020). doi: 10.7554/eLife.51862; pmid: 31931959
62. I. B. Robertson, D. B. Rifkin, Regulation of the Bioavailability of TGF-β and TGF-β-Related Proteins. *Cold Spring Harb. Perspect. Biol.* **8**, a021907 (2016). doi: 10.1101/cshperspect.a021907; pmid: 27252363

ACKNOWLEDGMENTS

We thank E. Fuchs for sharing the original LV vectors and *K14-CreER* mice; L. Chin for *TetO-Hras* mice; S. Karlsson for *Tgfbr2-floxed* mice; L. Coussens for *K14-HPV16* mice; F. Costantini for *Rosa-YFP* mice; L. Luo for *mTmG* mice; The Jackson Laboratory for *Kit*$^{w\text{-}sh}$ mice; L. Coussens for helpful suggestions; W. Anderson for editing the manuscript; OHSU's Department of Comparative Medicine (an AAALAC facility) for care and housing of our mouse colony; the OHSU Flow Cytometry Shared Resource for assisting with FACS (P. Streeter, Director); the MPSSR for Illumina RNA-seq library preparation and sequencing (R. Searles, Director); and the KCVI Epigenetics Consortium for RNA-seq data analysis (L. Carbone, Director, and B. Davis). **Funding:** This work was supported by an NIH K99-R00 Pathway to Independence Award (4R00CA178197-03), the Medical Research Foundation of Oregon, and the Collins Medical Trust (all to N.O.). **Author contributions:** S.T., S.K., and N.O. designed the experiments. S.T., A.E., A.V.D., and N.O. performed experiments and analyzed data. J.J.L. assessed tissue samples and provided clinical samples for histopathological analysis. S.T., S.K., and N.O. wrote the manuscript. All authors provided intellectual input and approved the final manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw gene expression data and processed counts are available on GEO, accession no. GSE151783.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/eaay1813/suppl/DC1
Figs. S1 to S6

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

27 May 2019; accepted 29 May 2020
10.1126/science.aay1813

RESEARCH ARTICLE

NEURODEVELOPMENT

A brainwide atlas of synapses across the mouse life span

Mélissa Cizeron[1,2]*, Zhen Qiu[1]*, Babis Koniaris[1,3], Ragini Gokhale[1], Noboru H. Komiyama[1,4], Erik Fransén[5,6], Seth G. N. Grant[1,4]†

Synapses connect neurons together to form the circuits of the brain, and their molecular composition controls innate and learned behavior. We analyzed the molecular and morphological diversity of 5 billion excitatory synapses at single-synapse resolution across the mouse brain from birth to old age. A continuum of changes alters synapse composition in all brain regions across the life span. Expansion in synapse diversity produces differentiation of brain regions until early adulthood, and compositional changes cause dedifferentiation in old age. The spatiotemporal synaptome architecture of the brain potentially accounts for life-span transitions in intellectual ability, memory, and susceptibility to behavioral disorders.

Excitatory synapses are the main class of brain synapse, and their postsynaptic proteins regulate both innate and learned behaviors (*1*–*6*). Mutations in these proteins cause >130 brain diseases (*7*), including disorders that characteristically arise in childhood, in adolescence, and in young or elderly adults.

Using synaptome mapping (*8*), we mapped excitatory synapse diversity and spatiotemporal synaptome architecture in >100 brain regions from birth until 18 months of age in mice (fig. S1). Synapses were labeled using fluorescent tags on endogenous PSD95 (PSD95-eGFP) and SAP102 (SAP102-mKO2) (*8*), two postsynaptic scaffold proteins that assemble multiprotein signaling complexes (*9*–*12*) necessary for synaptic plasticity and innate and learned behaviors (*2*, *3*, *5*–*7*, *12*). Disrupting the normal expression of these scaffold proteins or their associated proteins results in human neurodevelopmental and psychiatric disorders including autism, schizophrenia, and intellectual disability (*6*, *7*, *9*, *12*).

Our findings reveal a spatiotemporal program of synapse diversity across the brain, which we call the life-span synaptome architecture (LSA). The LSA shows how synapse diversity is generated as brain regions become dissimilar and how the synaptome architecture changes through development to adulthood and old age. The LSA provides a framework for understanding stereotypical life-span trajectories of behavioral changes and psychological functions (*13*–*15*) and why gene mutations characteristically result in synaptic pathology in certain brain areas and at certain ages. The Mouse Lifespan Synaptome Atlas and interactive visualization and analysis tools (*16*) provide a community resource for investigation of synapse function across all brain regions and the life span.

[1]Genes to Cognition Program, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK. [2]Institut NeuroMyoGène, Université de Lyon, Université Claude Bernard Lyon 1, CNRS UMR-5310, INSERM U-1217, 69008 Lyon, France. [3]School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, UK. [4]Simons Initiative for the Developing Brain (SIDB), Centre for Discovery Brain Sciences, University of Edinburgh, Edinburgh EH8 9XD, UK. [5]Department of Computational Science and Technology, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 10044 Stockholm, Sweden. [6]Science for Life Laboratory, KTH Royal Institute of Technology, SE-171 65 Solna, Sweden.
*These authors contributed equally to this work.
†Corresponding author. Email: seth.grant@ed.ac.uk

Life-span synaptome mapping pipeline and data resource

Parasagittal brain sections from cohorts of PSD95-eGFP; SAP102-mKO2 male mice were collected at 10 postnatal ages: 1 day (1D), 1 week (1W), 2 weeks (2W), 3 weeks (3W), 1 month (1M), 2 months (2M), 3 months (3M), 6 months (6M), 12 months (12M), and 18 months (18M) (Fig. 1A and figs. S2 and S3). Whole-brain sections were imaged at single-synapse resolution (*17*) on a spinning disc confocal microscope (pixel resolution 84 nm and optical resolution ~260 nm), and the density, intensity, size, and shape parameters of individual puncta were acquired using computer vision methods described previously (*8*). Synapses were classified into three types: type 1 express PSD95 only, type 2 express SAP102 only, and type 3 express both PSD95 and SAP102 (*8*). Thirty-seven subtypes were defined on the basis of molecular and morphological features (*8*). Supervised synaptome maps were generated by registering the data to the Allen Reference Atlas (*18*). Data were delineated into 109 anatomical subregions within 12 overarching regions, comprising isocortex, olfactory areas, hippocampal formation, cortical subplate, striatum, pallidum, thalamus, hypothalamus, midbrain, pons, medulla, and cerebellum (table S1).

All data and analysis tools are available in the Mouse Lifespan Synaptome Atlas (*16*). Synaptome Explorer enables in-depth exploration of raw and processed image data in single sections at single-synapse resolution, and the Synaptome Homology Viewer enables comparison of brain regions within and between mice of different ages.

The synaptome continuously changes across the life span

Raw images at low and high magnification revealed that each synaptic protein has a distinct spatiotemporal pattern and that the synaptome changes with age (Fig. 1A and figs. S2 and S3). To quantify the spatiotemporal differences in the synaptome, the life-span trajectories of PSD95 and SAP102 puncta density, intensity, and size were plotted as graphs and heatmaps revealing characteristic patterns for the whole brain, 12 regions, and 109 subregions (Fig. 1B, figs. S4 and S5, and table S1). Each parameter continuously changed across the life span. Synapse density rapidly increased during the first month in all brain areas and then fluctuated before declining in old age (Fig. 1, A to C, and figs. S3 to S5) (whereas the adult brain size remained unchanged; fig. S6), consistent with previous studies of synapse number quantified using electron microscopy in the rat brain (*19*–*22*). Each brain area underwent a specific program of synapse development, maturation, and aging. For example, the density of synapses peaked in the brainstem before cerebrum structures, potentially reflecting the requirement for the brainstem in early postnatal functions (fig. S7). The two synapse proteins showed different spatiotemporal trajectories (Fig. 1, B and C, and figs. S4 and S5), with SAP102 puncta density peaking before that of PSD95 in most brain areas (fig. S7), consistent with previous literature (*23*). Although, together, PSD95 and SAP102 label most excitatory synapses, additional markers would be required for an assessment of total excitatory synapse number in all brain regions and at all developmental stages.

Between 3M and 18M, most brain regions and subregions showed significantly ($P < 0.05$, Bayesian test with Benjamini-Hochberg correction) reduced synapse density (Fig. 1C, top panel, 70/109 subregions for PSD95; 78/109 for SAP102; fig. S8) and increased size (Fig. 1C, bottom panel, 56/109 subregions for PSD95, 80/109 for SAP102; fig. S8). Examination of the size distribution of the synapse populations showed a shift toward larger synapses with age (effect size >0.25 with $P < 0.01$, Kolmogorov-Smirnov test), consistent with previous electron microscopy studies in the aging macaque dorsolateral prefrontal cortex (*24*–*26*).

Life-span changes in the synaptome architecture can be divided into three broad phases. During the first phase (LSA-I), from birth to 1M, the numbers of puncta increased rapidly. The second phase (LSA-II) began as the rate of

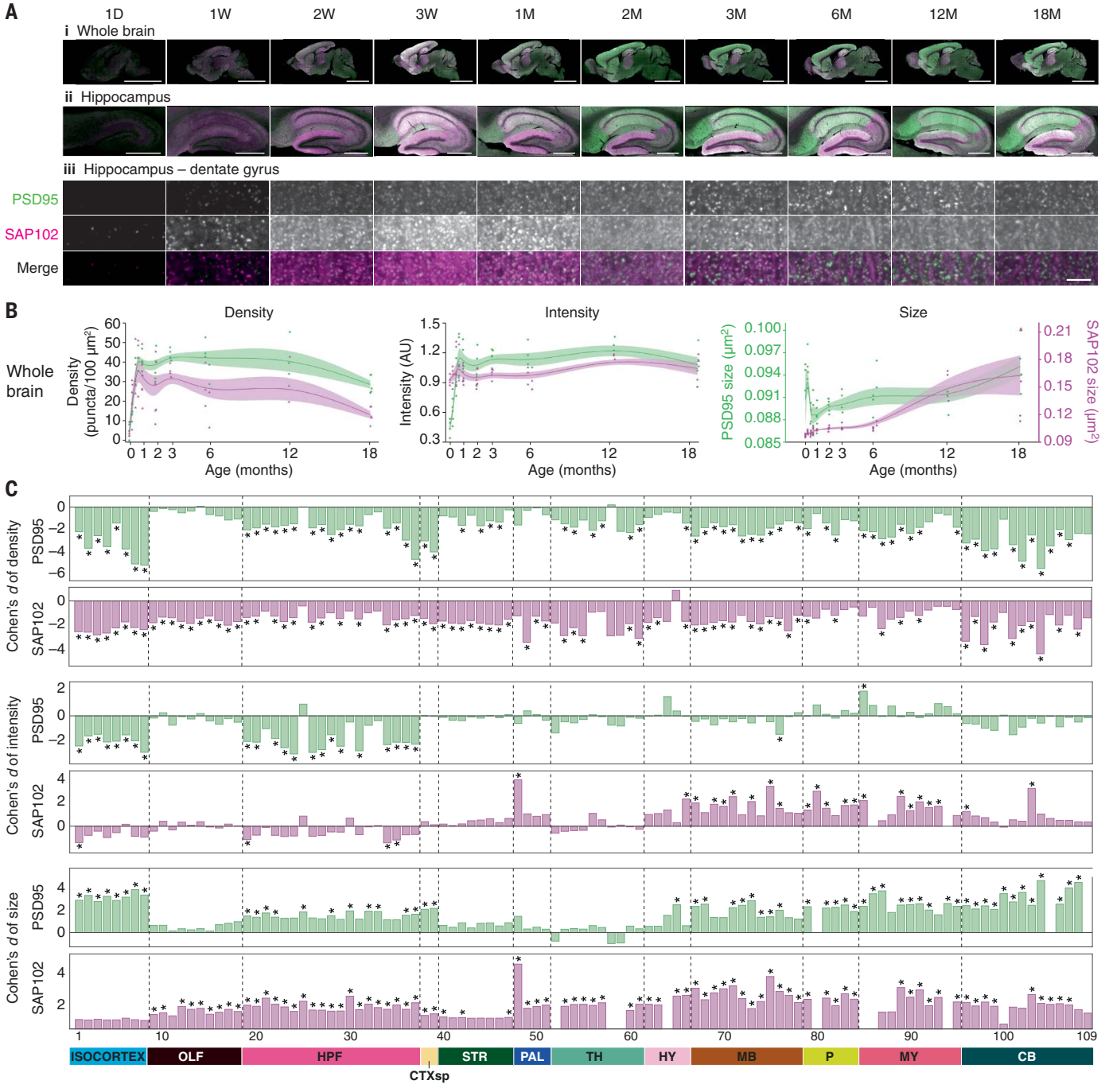


Fig. 1. Life-span trajectories of synapse parameters. (**A**) PSD95-eGFP (green) and SAP102-mKO2 (magenta) expression acquired at low [20× (i and ii)] and high [100× (iii)] magnification in the whole brain (i), hippocampus (ii), and molecular layer of the dentate gyrus (iii) at 10 ages across the mouse postnatal life span. Scale bars, 4 mm (i), 500 µm (ii), 3.5 µm (iii). D, day; W, week; M, month. (**B**) Life-span trajectories of synapse density, intensity [normalized to the mean intensity, arbitrary units (AU)], and size in the whole brain. Shown are PSD95-eGFP (green) and SAP102-mKO2 (magenta). Points represent individual mice, with beta-spline-smoothed curve of mean values and SEM shown. (**C**) Differences (Cohen's *d*) in synapse parameters between 3M and 18M in brain subregions (numbered; see table S1). *$P < 0.05$, Bayesian test with Benjamini-Hochberg correction. CB, cerebellum; CTXsp, cortical subplate; HPF, hippocampal formation; HY, hypothalamus; MB, midbrain; MY, medulla; OLF, olfactory areas; P, pons; PAL, pallidum; STR, striatum; TH, thalamus.

increase in puncta density slowed and was characterized by relative stability until 6M (adulthood). The third phase (LSA-III), late adult life, was characterized by a decline in puncta density and an increase in synapse size (Fig. 1C and fig. S8).

Synapse diversity across the life span

Each synapse type (Fig. 2, A and D, and figs. S9 and S10) and subtype (Fig. 2, B, C, and E, and figs. S11 and S12) had a specific trajectory in each brain region and subregion, reaching their peak values at different ages. Thus, the synapse composition of brain regions continued to change throughout the life span and was not restricted to LSA-I, when synapse density increased. Moreover, the presence of more than one peak at different ages (e.g., subtypes 17 and 18, $P < 0.05$, paired *t* test and Kolmogorov-Smirnov

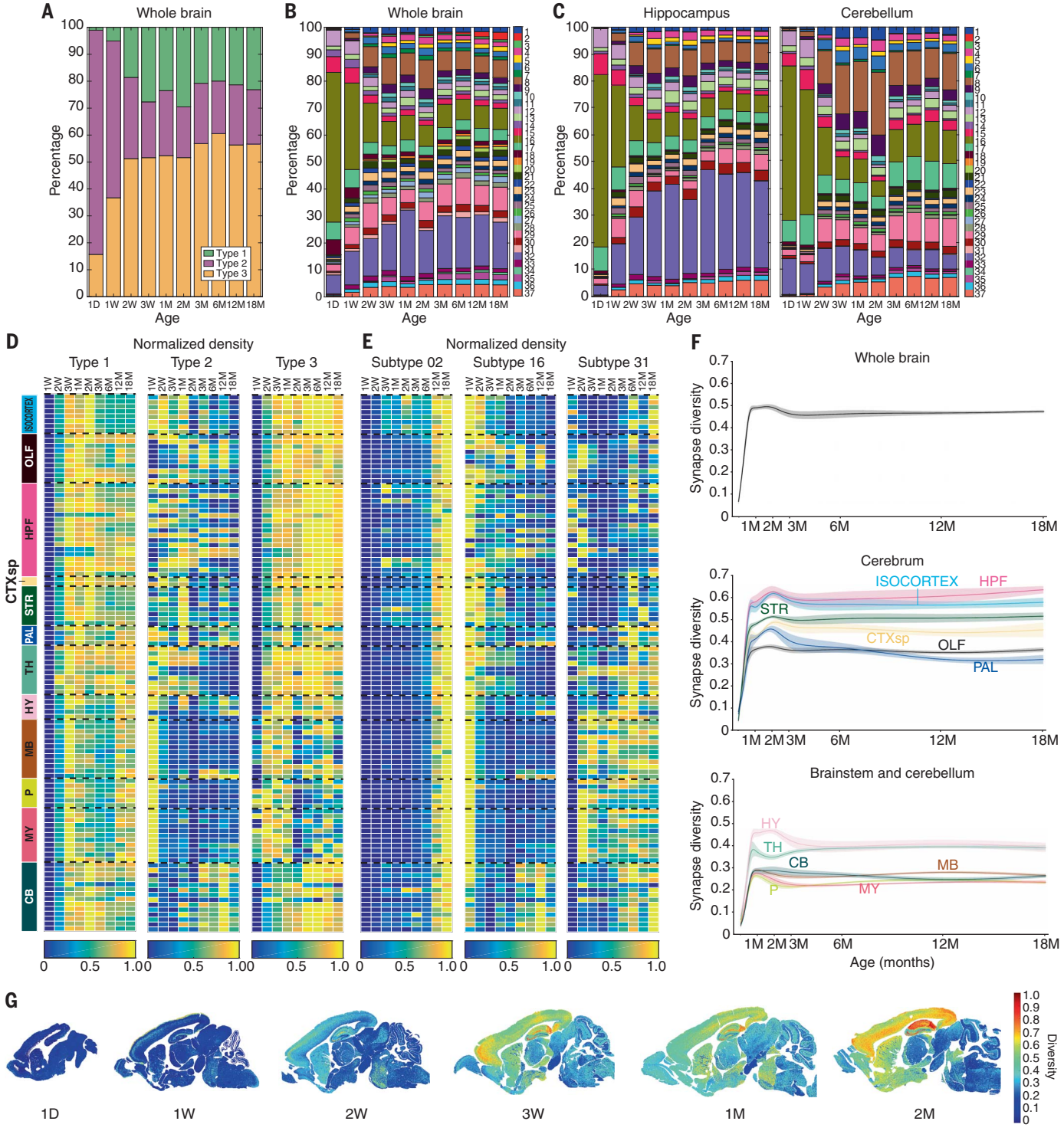


Fig. 2. Life-span trajectories of synapse types, subtypes, and diversity. (**A**) Stacked bar plot of percentage of synapse type density (type 1, PSD95 only; type 2, SAP102 only; type 3, colocalized PSD95+SAP102) in the whole brain across the life span. (**B**) Percentage of synapse subtype density in the whole brain across the life span. Key: synapse subtypes (1 to 37). (**C**) Percentage of synapse subtype density in hippocampus and cerebellum across the life span. (**D**) Life-span trajectories of synapse type density in 12 regions and 109 subregions (rows; see table S1). Density in each subregion was normalized (0 to 1) to its maximal density across the life span (columns). Twelve brain regions are shown (abbreviations as in Fig. 1C). (**E**) Life-span trajectories of three representative synapse subtypes (2, 16, and 31) in each of 109 subregions (rows; see table S1). Density in each subregion was normalized (0 to 1) to its maximal density across the life span (columns). (**F**) Life-span trajectories of synapse diversity (Shannon entropy) for whole brain (top) and main regions from the cerebrum (middle) and brainstem and cerebellum (bottom). Beta-spline-smoothed curve of mean and SEM are shown. (**G**) Unsupervised synaptome maps showing the spatial patterning of synapse diversity (Shannon entropy) per area (pixel size 21.5 × 21.5 μm) in representative parasagittal sections [for all ages, see fig. S15 and (*8*)].

test, Cohen's $d > 1.2$; figs. S10 to S12) suggests that shaping synapse composition (through processes such as transcriptional regulation, synapse pruning, and growth) is an ongoing process. In LSA-III, some subtypes were decreased ($P < 0.05$, Bayesian test with Benjamini-Hochberg correction), whereas others were increased ($P < 0.05$, Bayesian test with Benjamini-Hochberg correction), with differing specificity to brain regions (fig. S13). For example, subtypes 2, 27, and 34, which are large synapses, increased in many brain regions (Fig. 2E and figs. S12 and S13), whereas subtypes 12, and 14 to 16, which are small synapses, were lost in olfactory areas and thalamus in the old brain (Fig. 2E and figs. S12 and S13). Thus, subtypes of excitatory synapses are selectively gained or lost with aging, and different regions of the brain age in distinct ways.

LSA-I was initially dominated by a small subset of synapse types and subtypes, and these were overtaken by expanding populations of other types and subtypes (Fig. 2, A to C, and figs. S9 and S11). For example, type 2 and subtype 16 synapses dominated in the first postnatal week (Cohen's $d > 2$ with $P < 0.01$, two-way ANOVA with post hoc multiple-comparisons test) but were reduced by 1M (Cohen's $d < -2$, $P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test). We next quantified synapse diversity and found that all regions and subregions showed a rapid initial increase in the first 3 postnatal weeks (Fig. 2F and fig. S14). Brain areas responsible for higher cognitive functions (isocortex, cortical subplate, hippocampus, striatum) continued

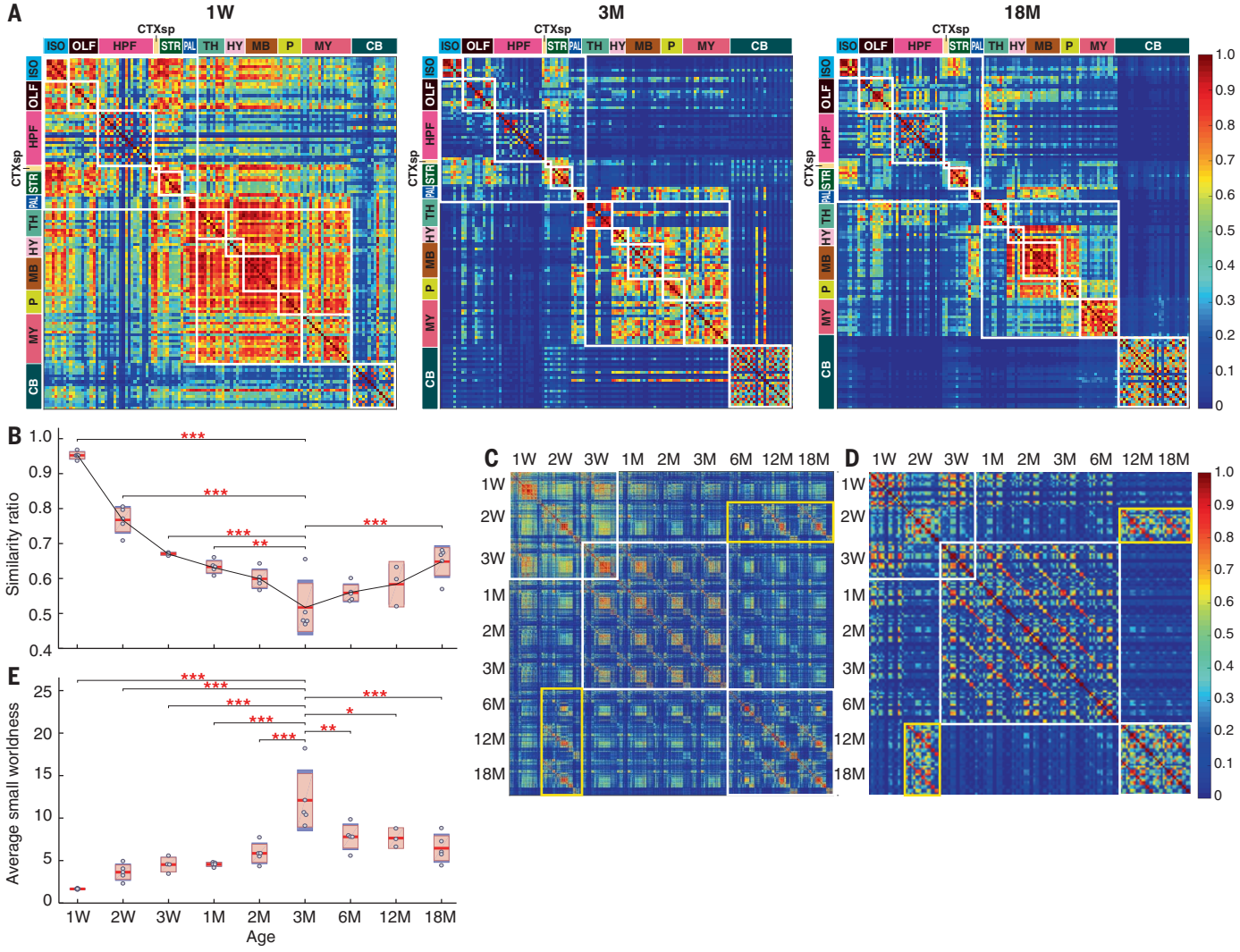


Fig. 3. Life-span synaptome architecture. (**A**) Matrix of similarities between pairs of subregions (rows and columns) at 1W, 3M, and 18M (for all ages, see fig. S16). Small white boxes indicate the subregions that belong to the same main brain region (see color code, left and top) and larger white boxes indicate main clusters: cerebrum, brainstem, and cerebellum. Note the reduction in similarity from 1W to 3M and the increase to 18M. Iso, isocortex; other abbreviations as in Fig. 1C. (**B**) Similarity ratio comparing the relative similarity of the synaptome in each main brain region with that of every other region (*8*). See the materials and methods for details. There were significant differences in the ratio between 3M and other ages: $^{**}P < 0.01$, $^{***}P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test. (**C**) Whole-brain hypersimilarity matrix showing the similarity between pairs of subregions at all ages. White boxes indicate the three main clusters corresponding to LSA-I, LSA-II, and LSA-III. Yellow boxes indicate the increased similarity of the old brain with the young brain (see higher-magnification image in fig. S18). (**D**) Hippocampus hypersimilarity matrix showing the similarity of pairs of hippocampal subregions at all ages. White boxes indicate the three main clusters corresponding to LSA-I, LSA-II, and LSA-III. Yellow boxes indicate the increased similarity of the old brain with the young brain. (**E**) Average small worldness across the life span. Scatter plots indicate the average small worldness per mouse brain section at different ages. There were significant differences in small worldness between 3M and other ages: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test.

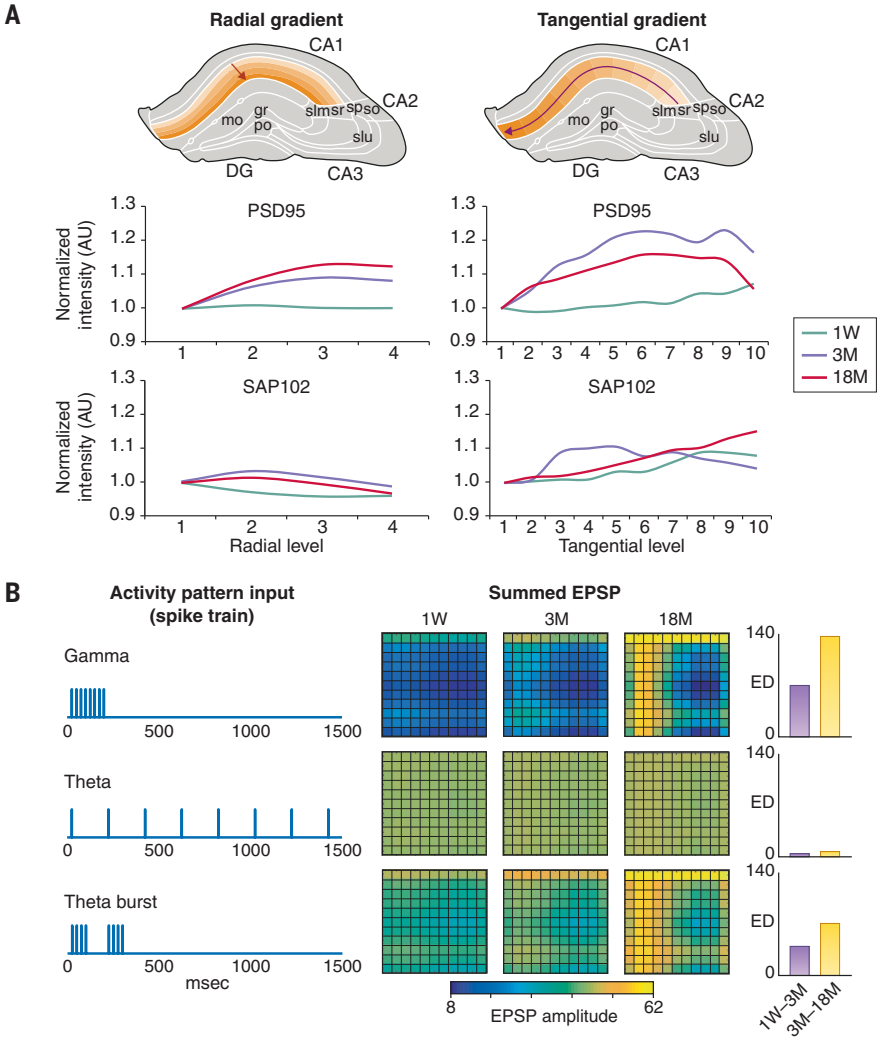


Fig. 4. Life-span changes in hippocampus architecture and electrophysiological properties. (**A**) Schematics of the hippocampus showing radial and tangential gradients in CA1sr subfield. Graphs show gradients of normalized synapse intensity (AU) of PSD95 and SAP102 at 1W, 3M, and 18M. CA1, cornu ammonis 1; CA2, cornu ammonis 2; CA3, cornu ammonis 3; DG, dentate gyrus; gr, granular layer; mo, molecular layer; po, polymorphic cell layer; slm, stratum lacunosum-moleculare; slu, stratum lucidum; so, stratum oriens; sp, stratum pyramidale; sr, stratum radiatum. (**B**) Summed response (EPSP amplitude) to three patterns (gamma, theta, and theta burst) of 20 action potentials of the 11 × 11 matrix of hippocampus synapses at three ages. Histograms show changes (summed Euclidean distance, ED) between 1W and 3M (purple) and 3M and 18M (yellow). (**C**) Schematic of the flow of information (arrows) in the trisynaptic hippocampal circuit connecting the DG molecular layer (DGmo), CA3 stratum radiatum (CA3sr), and CA1 stratum radiatum (CA1sr) and the life-span trajectory of synapse subtype density (normalized) in each region. EC, entorhinal cortex.

to expand their excitatory synaptic diversity (as reflected by the markers PSD95 and SAP102) after LSA-I, reaching a peak at 2M, whereas brain areas serving basal neurophysiological functions (midbrain, pons, medulla) peaked at 3W to 1M during LSA-I (Fig. 2F and fig. S14). Synapse diversity plateaued from 3M on in most brain areas (Fig. 2F). Unsupervised synaptome maps of the mouse brain, which visualize the anatomical distribution of synapse diversity (Fig. 2G and fig. S15), clearly show the increase in diversity in LSA-I, with the emergence of layers in the isocortex and subregional differentiation in the hippocampus.

Synaptome architecture first specializes, then dedifferentiates

To reveal how changes in synapse composition might contribute to differences between brain areas, we plotted similarity matrices of brain subregions at each age (Fig. 3A and fig. S16) (matrices were nonrandom, $P < 0.05$, Cohen's $d > 2$, permutation test). Similarity across all brain areas was highest in the first postnatal week and diminished until 3M ($P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test) (Fig. 3, A and B, and fig. S16). As the brain aged beyond 3M, there was a progressive increase in the similarity between brain areas ($P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test) (Fig. 3, A and B, and fig. S16). Individual brain regions also showed a reduction in similarity with the rest of the brain (i.e., differentiation) during the first 3 months ($P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test) (fig. S17), and from 3M to 6M on all regions except the cerebellum and medulla showed an increase in similarity (i.e., dedifferentiation) ($P < 0.01$, two-way ANOVA with post hoc multiple-comparisons test) (fig. S17).

We investigated whether dedifferentiation in LSA-III represents a return to a synaptome resembling that of a young brain or to a distinct, elderly-specific synaptome architecture. Using a hypersimilarity matrix that compares all subregions at all ages, we found that the 18M brain, in contrast to the 3M brain, was more similar to the 2W brain (Cohen's d = 1.5, $P < 0.01$, Bayesian test) (yellow boxes in Fig. 3C and figs. S18 and S19A). The hypersimilarity matrix also revealed three major blocks corresponding to the LSA phases (white boxes in Fig. 3C and fig. S18), with a transition between LSA-I and LSA-II at 3W, which corresponds to the behavioral transition from dependence on maternal care to independent living. The hypersimilarity matrix of hippocampal subregions showed a similar pattern (Cohen's d = 1.7, $P < 0.01$, Bayesian test) (Fig. 3D and fig. S19B).

To identify the synapse types and subtypes contributing to the differentiation-dedifferentiation trajectory, we correlated the abundance of each synapse subtype with the similarity ratio using brainwide data and regional data (fig. S20),

revealing a role for all three synapse types and a subset (21/37) of synapse subtypes ($r > 0.5$ or $r < -0.5$, $P < 0.05$, Mantel test with Benjamini-Hochberg correction).

The functional connectivity between brain areas, measured using resting-state functional magnetic resonance imaging, correlates with the topology (small worldness) of the synaptome network (*8*). Small worldness increased from birth to 3M ($P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test) and then declined to 18M ($P < 0.001$, two-way ANOVA with post hoc multiple-comparisons test) (Fig. 3E), suggesting that the differentiation-dedifferentiation trajectory influences the integrative property of brain circuits.

Life-span synaptome changes alter functional outputs

To explore how the age-dependent changes in synaptome architecture may cause changes in cognitive functions, we focused on the hippocampal formation, which is key for spatial navigation, learning, and memory (*27*). In the CA1 stratum radiatum of the adult mouse, there are orthogonal (radial and tangential) spatial gradients in PSD95 and SAP102 synaptic parameters that produce a local architecture of molecularly diverse synapses (*8*, *28*). Quantification of these gradients at 1W, 3M, and 18M showed age-dependent changes for PSD95 intensity in both radial and tangential directions (Fig. 4A and fig. S21). Conversely, for SAP102 no radial gradient was observed, and the tangential gradient was established by 1W and thereafter remained unchanged (Fig. 4A and fig. S21). This shows that these two closely related synaptic proteins undergo distinct spatiotemporal changes within the dendrites of CA1 pyramidal cells, producing a changing two-dimensional synaptome map across the life span. Using a computational simulation approach that tests the response (excitatory postsynaptic potential, EPSP) of CA1 synaptome maps to patterns of neural activity (*8*), we found that gamma and theta-burst patterns produced differential responses between 1W to 3M and 3M to 18M ($P < 0.05$, paired t test, Kolmogorov-Smirnov test), in contrast to theta trains, which produced a stable response at all ages (Fig. 4B and fig. S22). This illustrates how life-span synaptome changes affect synaptic responses to distinct temporal patterns of neural activity.

Different subregions of the hippocampal formation contribute distinct cognitive functions, which together produce an integrated behavioral output (*27*, *29*, *30*). This integrated function is exemplified by the trisynaptic circuit, in which axons project from neurons in the dentate gyrus to CA3 neurons, which project to CA1 neurons (*27*). Our data showed that each subregion in the trisynaptic circuit undergoes a different life-span trajectory of synaptic subtype composition, indicating that the memory functions controlled by these hippocampal subregions are highly likely to change with age (Fig. 4C and fig. S23).

Discussion

The dynamic temporal trajectories of excitatory synapse number, protein composition, morphology, and type and subtype diversity in >100 brain areas reveal a life-span synaptome architecture for the mouse brain. The continuum of changes in the synaptome architecture is divided into three epochs that broadly correspond to childhood and adolescence, early adulthood, and late adulthood. Synapse diversity expands between birth and early adulthood, driving the differentiation of brain regions, before changes in synapse composition progressively dedifferentiate brain regions in old age. These changes alter brain network and hippocampal physiological properties, and are potentially relevant to the trajectory of cognitive functions described in life-span studies of human behavior (*31*–*34*) and changes in the behavioral repertoire of animals across the life span (*13*–*15*).

The LSA reveals how factors that modify the expression of synaptic proteins (including genetic mutations, toxic proteins, inflammation, and drugs) can target particular synapses and brain regions at different ages and lead to behavioral changes. Expanding our approach to other synaptic proteins labeling greater synapse diversity, examining the synaptome of neuron types and dendritic morphology, and linking these brainwide synapse resolution data to transcriptional mechanisms that control brain gene expression across the life span (*35*) should enhance the effort to uncover the mechanisms and impacts of brain development, aging, and disease. Our highly scalable synaptomic methods and the Mouse Lifespan Synaptome Atlas described here provide new tools for addressing these issues.

REFERENCES AND NOTES

1. S. G. Grant *et al.*, *Science* **258**, 1903–1910 (1992).
2. N. H. Komiyama *et al.*, *J. Neurosci.* **22**, 9721–9732 (2002).
3. M. Migaud *et al.*, *Nature* **396**, 433–439 (1998).
4. A. J. Silva, R. Paylor, J. M. Wehner, S. Tonegawa, *Science* **257**, 206–211 (1992).
5. P. C. Cuthbert *et al.*, *J. Neurosci.* **27**, 2673–2682 (2007).
6. J. Nithianantharajah *et al.*, *Nat. Neurosci.* **16**, 16–24 (2013).
7. A. Bayés *et al.*, *Nat. Neurosci.* **14**, 19–21 (2011).
8. F. Zhu *et al.*, *Neuron* **99**, 781–799.e10 (2018).
9. E. Fernández *et al.*, *Mol. Syst. Biol.* **5**, 269 (2009).
10. R. A. Frank *et al.*, *Nat. Commun.* **7**, 11264 (2016).
11. R. A. W. Frank, F. Zhu, N. H. Komiyama, S. G. N. Grant, *J. Neurochem.* **142**, 504–511 (2017).
12. H. Husi, M. A. Ward, J. S. Choudhary, W. P. Blackstock, S. G. Grant, *Nat. Neurosci.* **3**, 661–669 (2000).
13. C. Darwin, *The Expression of the Emotions in Man and Animals* (Murray, 1872).
14. W. James, *The Principles of Psychology* (Holt, 1890).
15. N. Tinbergen, *The Study of Instinct* (Clarendon, 1951).
16. R. Gokhale, Z. Qiu, B. Koniaris, S. G. N. Grant, *The Mouse Lifespan Synaptome Atlas* (2020); www.brain-synaptome.org.
17. X. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11551–11556 (2005).
18. H.-W. Dong, *Allen Reference Atlas: A Digital Color Brain Atlas of the C57black/6J Male Mouse* (Wiley, 2008).
19. I. Adams, D. G. Jones, *Neurobiol. Aging* **3**, 179–186 (1982).
20. G. K. Aghajanian, F. E. Bloom, *Brain Res.* **6**, 716–727 (1967).
21. K. D. Micheva, C. Beaulieu, *J. Comp. Neurol.* **373**, 340–354 (1996).
22. B. Crain, C. Cotman, D. Taylor, G. Lynch, *Brain Res.* **63**, 195–204 (1973).
23. G. M. Elias, L. A. B. Elias, P. F. Apostolides, A. R. Kriegstein, R. A. Nicoll, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20953–20958 (2008).
24. D. Dumitriu *et al.*, *J. Neurosci.* **30**, 7507–7515 (2010).
25. J. H. Morrison, M. G. Baxter, *Nat. Rev. Neurosci.* **13**, 240–250 (2012).
26. A. Peters, C. Sethares, J. I. Luebke, *Neuroscience* **152**, 970–981 (2008).
27. J. Basu, S. A. Siegelbaum, *Cold Spring Harb. Perspect. Biol.* **7**, a021733 (2015).
28. M. J. Broadhead *et al.*, *Sci. Rep.* **6**, 24626 (2016).
29. D. M. Bannerman *et al.*, *Nat. Rev. Neurosci.* **15**, 181–192 (2014).
30. I. Soltesz, A. Losonczy, *Nat. Neurosci.* **21**, 484–493 (2018).
31. P. B. Baltes, S. W. Cornelius, A. Spiro, J. R. Nesselroade, S. L. Willis, *Dev. Psychol.* **16**, 625–635 (1980).
32. S. C. Li *et al.*, *Psychol. Sci.* **15**, 155–163 (2004).
33. G. Reinert, "Comparative factor analytic studies of intelligence throughout the human life-span," in *Life-Span Developmental Psychology*, L. R. Goulet, P. B. Baltes, Eds. (Elsevier, 1970), pp. 467–484.
34. E. M. Tucker-Drob, *Dev. Psychol.* **45**, 1097–1118 (2009).
35. N. G. Skene, M. Roy, S. G. Grant, *eLife* **6**, e17915 (2017).
36. M. Cizeron *et al.*, Mouse Lifespan Synaptome Atlas dataset (2020); https://doi.org/10.7488/ds/2796.
37. Z. Qiu, S. G. N. Grant, Matlab code for: A brain-wide atlas of synapses across the mouse life span, Zenodo (2020); https://doi.org/10.5281/zenodo.3825214.

ACKNOWLEDGMENTS

We thank C. McLaughlin and K. Elsegood for mouse colony and laboratory management; D. Kerrigan and D. Fricker for genotyping; N. G. Skene for statistical advice; S. Munni, O. Kealy, and H. Taczynski for image calibration; D. Maizels for artwork; and C. Davey for editing. **Funding:** This work was supported by the Wellcome Trust (Technology Development grant no. 202932), the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (695568 SYNNOVATE), and the Simons Foundation Autism Research Initiative (529085). **Author contributions:** M.C. designed the animal cohort; collected, prepared, and imaged brain samples; delineated brain regions and subregions; calibrated synapse detection; analyzed synapse parameters and hippocampal gradients; and interpreted the data. Z.Q. developed the methodology and optimized the life-span SYNMAP pipeline; analyzed data of the whole life-span mouse cohort; segmented images and quantified and classified puncta; performed unsupervised and supervised mapping of synapse parameters, types, and subtypes; and analyzed diversity and network topology. E.F. performed statistical analysis and computational modeling of synaptome physiology. B.K. constructed the Synaptome Explorer and the Synaptome Homology Viewer. R.G. constructed the website. N.H.K. provided advice and supervision. S.G.N.G. conceived the project, performed analyses, supervised the project, and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available at the Mouse Synaptome Atlas (*16*) and Edinburgh DataShare (*36*), and code is available at Zenodo (*37*). Requests for materials should be addressed to S.G.N.G.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/270/suppl/DC1
Materials and Methods
Figs. S1 to S23
Table S1
References (*38*–*43*)
MDAR Reproducibility Checklist

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

23 November 2019; accepted 19 May 2020
Published online 11 June 2020
10.1126/science.aba3163

CANCER

Cancer cells deploy lipocalin-2 to collect limiting iron in leptomeningeal metastasis

Yudan Chi[1], Jan Remsik[1], Vaidotas Kiseliovas[2], Camille Derderian[1], Ugur Sener[3], Majdi Alghader[1], Fadi Saadeh[1], Katie Nikishina[1], Tejus Bale[4], Christine Iacobuzio-Donahue[1,4], Tiffany Thomas[5], Dana Pe'er[2,6], Linas Mazutis[2,6], Adrienne Boire[1,3,7]*

The tumor microenvironment plays a critical regulatory role in cancer progression, especially in central nervous system metastases. Cancer cells within the cerebrospinal fluid (CSF)–filled leptomeninges face substantial microenvironmental challenges, including inflammation and sparse micronutrients. To investigate the mechanism by which cancer cells in these leptomeningeal metastases (LM) overcome these constraints, we subjected CSF from five patients with LM to single-cell RNA sequencing. We found that cancer cells, but not macrophages, within the CSF express the iron-binding protein lipocalin-2 (LCN2) and its receptor SCL22A17. These macrophages generate inflammatory cytokines that induce cancer cell LCN2 expression but do not generate LCN2 themselves. In mouse models of LM, cancer cell growth is supported by the LCN2/SLC22A17 system and is inhibited by iron chelation therapy. Thus, cancer cells appear to survive in the CSF by outcompeting macrophages for iron.

Spread of cancer cells into the cerebrospinal fluid (CSF)–filled leptomeninges is known as leptomeningeal metastasis (LM). This form of metastasis has become increasingly common (*1*) and is typically fatal within months (*2*). Under normal physiological conditions, the leptomeningeal space is isolated from the systemic circulation by the blood-CSF barrier. This anatomic compartment is hypoxic and contains sparse amounts of metabolic intermediates and micronutrients (*3*). In the setting of LM, the normally acellular CSF contains cancer cells as well as lymphocytes, macrophages, and neutrophils. Cancer cells within this microenvironment must therefore cope with oppressive metabolic constraints while evading immune responses. To elucidate the mechanisms by which cancer cells overcome these constraints, we first applied single-cell transcriptomic techniques to CSF collected from cancer patients harboring LM and then undertook functional studies in mouse models of LM.

Cancer cells within human spinal fluid express LCN2 and SLC22A17

Cancer cells in the CSF disseminate throughout the central nervous system. Within this compartment they are vastly outnumbered by immune cells, primarily macrophages and lymphocytes (Fig. 1A) (*4*). To explore cancer and immune cell responses to the nutritionally sparse CSF, we applied single-cell RNA sequencing (scRNA-seq) to cellular material collected from the CSF of five patients with LM (Fig. 1B and figs. S1 and S2). LM was secondary to breast cancer primaries in three patients and to non–small cell lung cancer primaries in two patients (table S1). After scRNA-seq, the proportions of each cell type within the CSF remained consistent with those identified by clinical analysis (fig. S3, A and B).

We found that all CSF cells showed upregulated expression of iron transport genes, consistent with functional iron deficiency within this anatomic space (fig. S3C). Hepcidin, a protein implicated in both inflammation and iron deficiency (*5*), was detected at higher concentrations in the CSF of cancer patients harboring LM relative to those without LM (fig. S3D). In contrast, CSF levels of the ubiquitous iron transporter transferrin were equivalent in the two groups of patients (fig. S3E). We found that whereas immune cells expressed canonical iron transporter transcripts, cancer cells expressed a diverse array of genes associated with iron binding and transport (fig. S3C). Of these, transcripts for a single iron-binding and receptor pair were expressed exclusively within the cancer cell population in all five patients: lipocalin-2 (LCN2) and solute carrier family 22 member 17 (SLC22A17) (Fig. 1, C and D). Protein expression corresponding to these single-cell transcriptional data was confirmed by enzyme-linked immunosorbent assay (ELISA) and flow cytometry of human CSF (fig. S3, F to H). Finally, in autopsy tissues, LCN2 and SLC22A17 protein expression was detected by immunofluorescence in the cancer cell population but not in the macrophage or monocyte populations (Fig. 1, E to H, and fig. S3, I to O).

LCN2, also known as neutrophil gelatinase-associated lipocalin (NGAL), is a β-barrel secreted protein that binds siderophore-complexed ferric iron with high affinity (*6*–*8*). SLC22A17 is a LCN2 transporter expressed in a various cell types, including cancer cells (*9*). To investigate the functional consequences of LCN2/SLC22A17 expression, we used three mouse models of LM generated through iterative in vivo selection (table S2). Unselected parental cells are nonspecifically metastatic; the LM subpopulation of cells (LeptoM) readily enter into and grow within the leptomeningeal space (*10*). The mouse lung adenocarcinoma (LLC-Par, LLC-LeptoM), human lung adenocarcinoma (PC9-Par, PC9-LeptoM), and human breast adenocarcinoma (MDA231-Par, MDA231-LeptoM) models all share key features of human LM (*11*).

Flow cytometry of CSF from the immunocompetent LLC-LeptoM mouse model revealed infiltrating cancer cells accompanied by lymphocytes, macrophages, and neutrophils (Fig. 2A). Concentrations of hepcidin and LCN2 were elevated in the CSF of mice with LM (fig. S4, A and B). LCN2 staining by immunohistochemistry was specific to cancer cells within the leptomeningeal space (fig. S4C). We also found higher levels of LCN2 mRNA and protein in LeptoM cells than in their parental (Par) counterparts (fig. S2, D and E). Although previous studies have documented LCN2 expression in activated macrophages within other anatomic sites (e.g., mammary fat pad) (*11*, *12*), we found that LCN2 was expressed by cancer cells and not macrophages within the leptomeningeal space (fig. S4F). LCN2 binds productively to three known receptors (SLC22A17, LRP2, and MC4R) (*13*, *14*). In both humans and mouse models, SLC22A17 constituted the major LCN2 receptor (Fig. 1, F and H, and fig. S4, G to J).

LCN2 promotes cancer cell growth within the leptomeningeal space in mice

We next investigated the functional relevance of LCN2/SLC22A17 in the CSF. We found that short hairpin RNA (shRNA)–mediated loss of LCN2 expression in LeptoM cells inhibited their growth within the leptomeninges in all three mouse models and conferred a survival benefit to the animals (Fig. 2, B and C, and fig. S5, A to G). In contrast, LCN2 knockdown did not alter the growth of LeptoM cells within iron-replete anatomic sites in vivo (fig. S5, H to J). Knockdown of the LCN2 receptor SLC22A17 phenocopied these results (Fig. 2, D and E, and fig. S5, K and L). Conversely, overexpression of LCN2 in parental cells with no propensity for growth in the leptomeninges promoted the growth of these cells in the leptomeningeal space and hastened the death of these animals (Fig. 2, F and G, and fig. S6). Together, these results support a mechanistic role for the LCN2/SLC22A17 axis in leptomeningeal cancer cell growth.

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [2]Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [3]Department of Neurology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [4]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [5]Department of Pathology and Cell Biology, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA. [6]Parker Institute for Cancer Immunotherapy, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [7]Brain Tumor Center, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
*Corresponding author. Email: boirea@mskcc.org

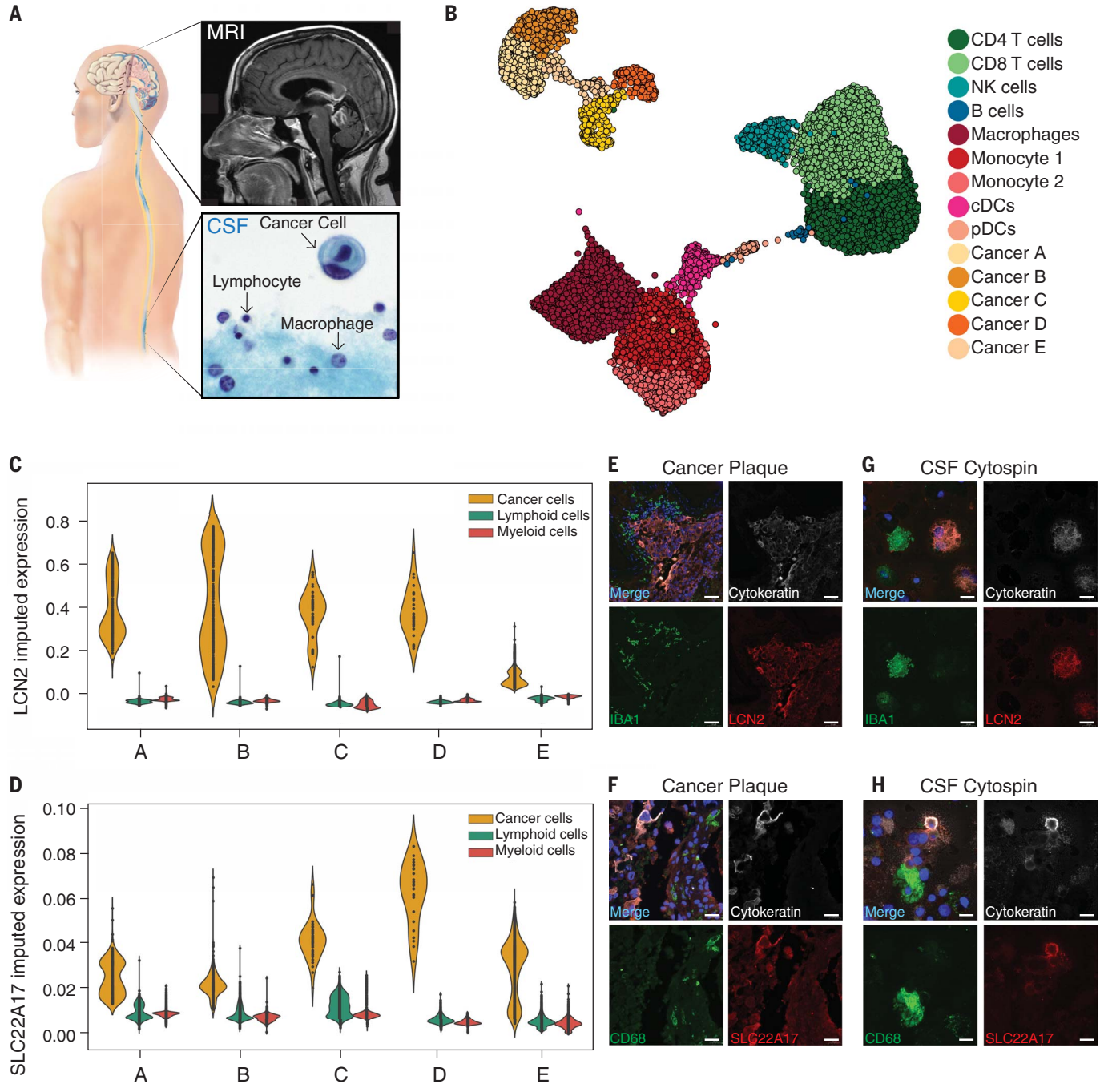


Fig. 1. Sparse CSF iron promotes LCN2 expression in cancer cells. (**A**) Human LM from breast cancer (patient B). White plaques of LM were visualized by magnetic resonance imaging (top right) and Giemsa-stained cytospin of CSF (bottom right). Major cell populations are indicated; the cytospin area pictured is 274 μm × 258 μm. (**B**) Single-cell transcriptional map of cancer and immune cells present in CSF of five patients, projected with uniform manifold approximation and projection (UMAP). Each dot represents a cell, colored by PhenoGraph cluster; major cell types are manually annotated according to fig. S1. Each individual patient is projected in fig. S2. (**C** and **D**) *LCN2* and *SLC22A17* gene expression in LM patients. Violin plots of *LCN2* (C) and *SLC22A17* (D) show imputed gene expression in cells from individual patients. All cells are grouped into three compartments: cancer cells (cancer patients A to E), lymphoid cells ($CD4^+$ T cells, $CD8^+$ T cells, B cells, NK cells), and myeloid cells [monocyte 1, monocyte 2, macrophage, conventional dendritic cells (cDCs), plasmacytoid dendritic cells (pDCs)]. (**E** and **F**) LCN2 (E) and SLC22A17 (F) detection by immunofluorescent staining of leptomeninges collected at autopsy from patients harboring LM. Cytokeratin (white) indicates cancer cells. Macrophages are indicated by the macrophage-specific proteins CD68 or ionized calcium-binding adaptor molecule 1 (IBA1) in green; LCN2 and SLC22A17 are shown in red. Images are representative of three cancer patients. Scale bars, 50 μm. (**G** and **H**) LCN2 (G) and SLC22A17 (H) expression assessed by cytospin staining of cancer cells and macrophages in CSF collected post mortem from patients in (E) and (F). Scale bars, 20 μm.

Inflammatory cytokines induce LCN2 expression in cancer cells

We next explored the mechanism leading to LCN2 expression in cancer cells within the LM. Downstream of both STAT and NF-κB transcriptional promoters, expression of LCN2 may be induced by a variety of inflammatory stimuli (*15*, *16*). To assess the inflammatory state of cells during LM, we queried the scRNA-seq dataset and found high expression levels of transcripts downstream of JAK-STAT and NF-κB promoters in the macrophage population (Fig. 3A and fig. S7A). Consistent with this, inflammatory cytokine (IL-6, IL-8, and IL-1β) concentrations were significantly higher in the CSF of cancer patients with LM than in control patients without LM (fig. S7B). This was also observed in mouse models (fig. S7, C and D). We hypothesized that macrophage-generated cytokines stimulate cancer cell LCN2 expression in the CSF.

To test this hypothesis, we cocultured LeptoM cancer cells with supernatant collected from macrophages freshly isolated from either the CSF or spleen of mice harboring LLC-LeptoM or LLC-Par tumor cells (Fig. 3B and fig. S7E). We found that macrophages from the CSF of mice harboring LM strongly induced LCN2 expression in LeptoM cancer cells (Fig. 3C). Coculture of these CSF macrophages with LeptoM cancer cells induced expression of LCN2 in cancer cells but not in macrophages (Fig. 3D). Outside the leptomeninges, LCN2 expression by macrophages/monocytes and neutrophils is typically induced by inflammation (*17*). In our models, although LCN2 expression was increased in extracranial splenic monocytes and neutrophils after lipopolysaccharide (LPS) treatment, LCN2 expression remained unchanged in these cells in the

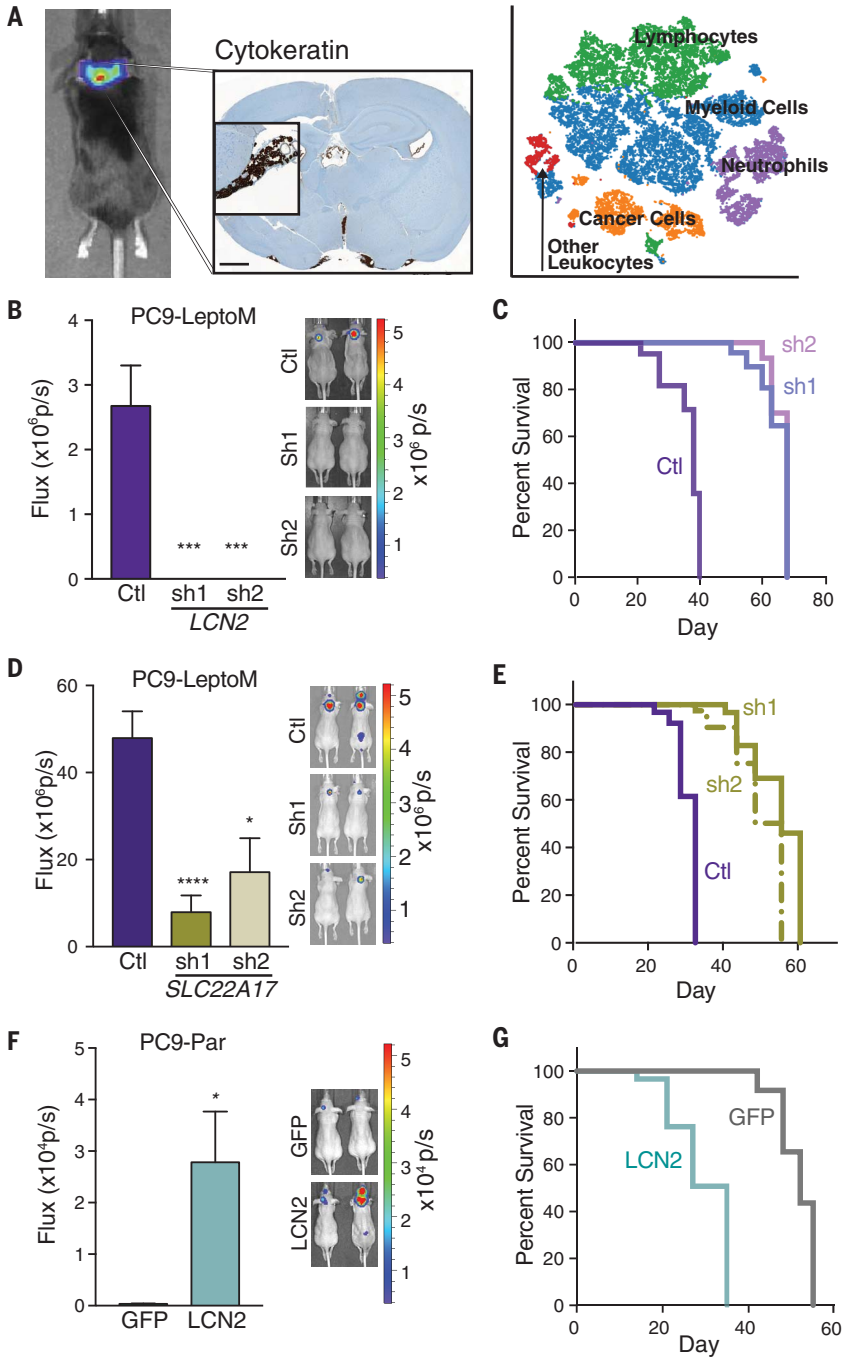


Fig. 2. LCN2 supports cancer cell growth in the leptomeninges. (**A**) Mouse model of LM. Left: In vivo bioluminescence imaging (BLI) of the LLC model at day 14 after intracardiac dissemination. Center: Immunohistochemistry for pan-cytokeratin. Scale bar, 1000 μm for larger image, 50 μm for inset. Right: t-distributed stochastic neighbor embedding (tSNE) of multicolor flow cytometry of CSF collected from syngeneic LLC mouse model (day 14). Major cell types: orange, cancer cells ($CD45^-$); purple, neutrophils ($CD45^+CD11b^+Ly6G^+$); green, lymphocytes ($CD45^+CD3^+$); blue, myeloid cells ($CD45^+CD11b^+$ excluding neutrophils); red, other leukocytes ($CD45^+CD11b^-Ly6G^-CD3^-$). A representative sample is shown; n = 5 mice. (**B**) Tumor growth in mice injected with PC9-LeptoM cells expressing either shCtl or shLCN2. Two independent shRNAs (sh1 and sh2) were used to target human *LCN2*. Left: Histogram represents in vivo BLI imaging after inoculation at day 28; n = 8 to 10 per group in each of two independent experiments. Right: Representative BLI images. ***$P < 0.001$ (unpaired t test). Data represent mean ± SEM. (**C**) Kaplan-Meier survival curve of mice shown in (B). $P < 0.0001$. (**D**) Tumor growth in mice injected with PC9-LeptoM cells expressing either shCtl or shSLC22A17. We used two independent shRNAs (sh1 and sh2) to target human *SLC22A17*. Left: Histogram represents in vivo BLI imaging at day 28 after inoculation. n = 8 to 10 per group in each of two independent experiments. Right: Representative BLI images. *$P < 0.05$, ****$P < 0.0001$ (unpaired t test). Data are means ± SEM. (**E**) Kaplan-Meier survival curve of mice shown in (D). $P < 0.0001$. (**F**) Tumor growth in mice injected with PC9-Par cells expressing either green fluorescent protein (GFP) or LCN2. Left: Histogram represents in vivo BLI imaging after inoculation at day 21. n = 8 to 10 per group in each of two independent experiments. Right: Representative BLI images. *$P < 0.05$ (unpaired t test). Data represent mean ± SEM. (**G**) Kaplan-Meier survival curve of mice shown in (F). $P < 0.0001$.

CSF (figs. S3G, S4F, and S7F). Moreover, although LCN2 may induce influx of neutrophils to extracranial sites (*18*), we did not observe LCN2-dependent changes in CSF leukocyte composition (fig. S7G).

To identify the relevant cytokines in human disease, we immune-depleted select inflammatory cytokines from LM-positive human CSF (Fig. 3E). Whereas whole CSF from LM patients strongly induced LCN2 expression in cancer cells, immune depletion of IL-6 and/or IL-8 or inhibited this biological effect in both PC9 and MDA231 model systems (Fig. 3F and fig. S8A). Conversely, addition of recombinant IL-6, IL-8, and/or IL-1β to artificial CSF induced expression of LCN2 from PC9, MDA231, or LLC (fig. S8, B to D) LeptoM cells. Together, these data indicate that macrophage-generated cytokines induce expression of LCN2 in cancer cells within the LM.

To gain more insight into the properties of these proinflammatory leptomeningeal macrophages, we examined the transcriptome of the macrophage/monocyte population in our scRNA-seq dataset. Consistent with the known transcriptional heterogeneity of these cells (*19*), we observed nonreciprocal expression of both M1 (proinflammatory) and M2 (anti-inflammatory) polarization transcriptional gradients (fig. S8E). Unexpectedly, this analysis revealed the hypoxia transcriptional signature as a major feature of this macrophage population.

LCN2 supports cancer cell growth in the hypoxic leptomeninges in mice

Hypoxia signaling is intimately linked to iron homeostasis (*20*). We found that freshly isolated CSF cancer cells and immune cells from human LM samples showed evidence of hypoxia-induced transcriptional changes as well as a correlation between hypoxia and iron metabolic signatures (Fig. 4, A and B, and fig. S9A). In mouse models, bulk RNA-seq of LeptoM derivative cell lines with and without LCN2 knockdown revealed LCN2-dependent expression of the hypoxia transcriptional signature, suggesting a role for LCN2 in hypoxia (Fig. 4C and fig. S9, B and C).

The biological relevance of hypoxia signaling in LM remains undetermined. In the absence of disease, the CSF is hypoxic, with partial pressure of O_2 ranging from 65 to 130 ± 49 mmHg. However, in inflammatory injury (*21*) or an impaired blood-CSF barrier (*11*), CSF oxygenation is improved. To study hypoxia in our LM mouse models, we generated a dual-luciferase reporter system consisting of constitutive firefly luciferase and nano-luciferase downstream of hypoxia response elements. The PC9-LeptoM and LLC-LeptoM models demonstrated NLuc activation (an indication of hypoxia) upon inoculation into the CSF, and this activation remained stable over the disease course (Fig. 4D). Despite this hypoxia, the

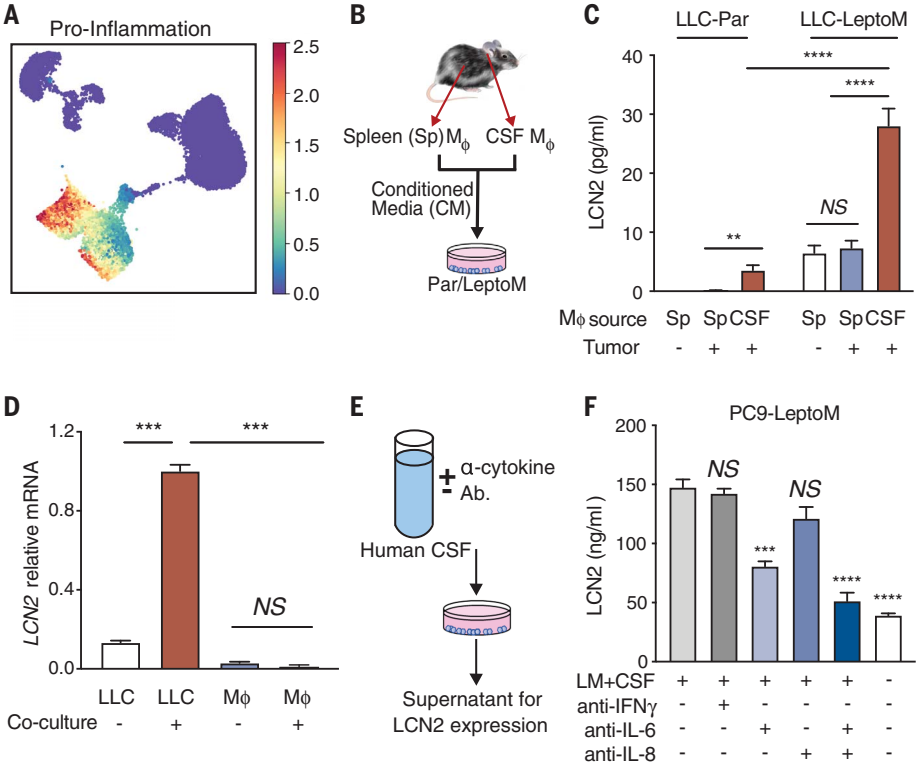


Fig. 3. Cancer cells generate LCN2 in response to inflammatory cytokines. (**A**) Mean expression of canonical proinflammatory cytokines from single-cell transcriptomics. Markov affinity-based graph imputation of cells (MAGIC)–imputed cytokine expression was standardized to zero mean and unit of standard deviation. See the proinflammatory cytokine dataset in table S5. (**B**) Schematic for cancer cell coculture with supernatant from macrophages in LLC-LeptoM model. Conditioned media was collected from macrophages, freshly sorted from CSF or spleen of C57Bl/6 mice harboring LLC-LeptoM cells or no cancer by fluorescence-activated cell sorting (FACS), and added to cultures of Par cells or LeptoM cells for 14 days, exchanging fresh conditioned media every 3 days. (**C**) Evaluation of LCN2 levels by ELISA in conditioned media generated as in (B); $n = 3$ in each of two independent experiments. $^{**}P < 0.01$, $^{****}P < 0.0001$ (unpaired t test); NS, not significant. Data are means ± SEM. (**D**) *LCN2* mRNA detection by quantitative polymerase chain reaction in CSF-derived macrophages and LLC-LeptoM cells after coculture at day 14; $n = 3$ in each of two independent experiments. $^{***}P < 0.001$ (unpaired t test). Data are means ± SEM. RLU, relative luminescence unit. (**E** and **F**) CSF from patients harboring LM was treated with neutralizing antibodies to IL-6, IL-8, or IFN-γ and added to PC9-LeptoM cells for 12 hours. LCN2 was quantified by ELISA in conditioned media 24 hours after removing CSF; $n = 3$ in each of two independent experiments. $^{***}P < 0.001$, $^{****}P < 0.0001$ (unpaired t test). Data are means ± SEM.

LeptoM cells continued to grow and expressed hypoxia-inducible factors HIF-1α and HIF-2α (Fig. 4, D and E, and fig. S9, D and E). In vitro, LeptoM derivatives demonstrated robust growth under hypoxic conditions in an LCN2-dependent manner: Knockdown of LCN2 expression with shRNA inhibited LeptoM growth and promoted apoptosis in hypoxia (fig. S9, F to P), and overexpression of LCN2 provided resistance to hypoxic stress (fig. S9, Q and R). From these experiments, we conclude that cancer cell LCN2 expression supports cancer cell growth within hypoxic CSF in LM.

Cancer cells use LCN2 to collect sparse extracellular iron in the CSF

In the absence of disease, CSF contains minimal extracellular iron. To investigate the levels of this micronutrient in LM, we assayed total iron levels by mass spectroscopy in the CSF of cancer patients. We found that total iron concentration and the proportion of iron bound to LCN2 was increased in the CSF of patients with LM relative to patients without LM (fig. S10, A and B). In addition, patient CSF LCN2 levels correlated with iron concentration (table S3). Reflecting the importance of inflammatory signaling, CSF iron levels also correlated with IL-6 and hepcidin.

Turning to our mouse models, we found that LCN2 gene expression was up-regulated in LeptoM cells relative to Par cells (figs. S4D and S10C). In vivo, LCN2 knockdown by shRNA was partially rescued through addition of iron-loaded transferrin (Fig. 5, A and B, and fig. S10, D to F), which suggests that iron transport plays a key role in LCN2-dependent cancer cell

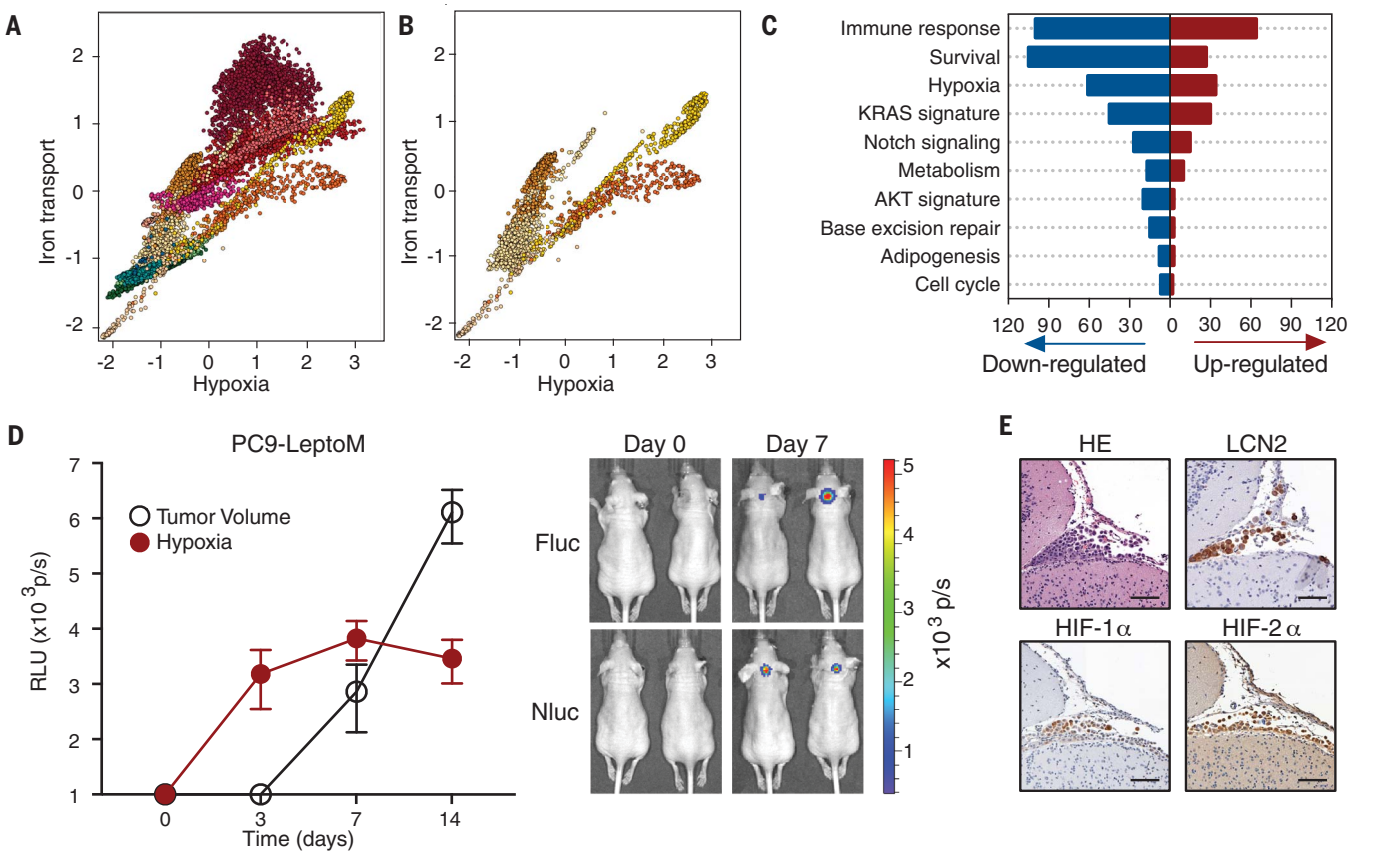


Fig. 4. LCN2 supports cancer cell growth in the hypoxic CSF microenvironment. (**A**) Correlation between hypoxia and iron ion transport gene signatures. Axis value represents mean gene signature expression per cell, standardized to have zero mean and unit of standard deviation. Each dot represents a cell, colored by cell type as in Fig. 1B. (**B**) Cancer cell populations from (A) alone. (**C**) Gene set enrichment (GSA) analysis of PC9 and LLC models by bulk RNA sequencing of LeptoM shCtl and shLCN2. Numbers of significantly up-regulated (red) and down-regulated (blue) genes in GSA analysis are indicated on the *x* axis; $n = 2$ per group. $P < 0.05$. See also fig. S7B. (**D**) Leptomeningeal tumor growth in a dual-reporter in vivo system. PC9-LeptoM cells express Firefly luciferase (Fluc) constitutively; NanoLuc (Nluc) is induced downstream of the hypoxia response element. Left: Fluc and Nluc are assayed by BLI at indicated time points. Solid circles indicate Nluc; open circles indicate Fluc. Right: Representative BLI images at day 0 and day 7. $n = 8$ to 10 per group in each of two independent experiments. Data are means ± SEM. (**E**) Leptomeningeal tissue sections stained with hematoxylin and eosin (HE) and immunohistochemistry for LCN2 and hypoxia-inducible factors HIF-1α and HIF-2α in PC9-LeptoM model. Scale bars, 50 μm.

growth in the CSF. To investigate this further, we examined iron uptake in these LeptoM cells in vitro. Inhibition of either SLC22A17 or LCN2 expression in LeptoM cells by shRNA in vitro inhibited iron accumulation and cell growth (fig. S10, G to I). Intracellular iron accumulation and cell growth were rescued by addition of exogenous transferrin (fig. S10, G to I).

Activated macrophages cannot generate reactive oxygen species unless iron concentrations are sufficiently high (*22*). We hypothesized that macrophages within the CSF, particularly in LM, have lower iron stores than macrophages circulating in the blood. To address this, we studied the syngeneic LLC-LeptoM mouse model (table S2). We collected macrophages from either the CSF or the spleen in the setting of LM or after challenge with LPS (Fig. 5F). We found that the intracellular iron content of CSF macrophages declined significantly in the setting of LM when compared with LPS treatment (Fig. 5C and fig. S10J). We also found that shRNA-mediated knockdown of cancer cell LCN2 expression increased macrophage iron content (Fig. 5C and fig. S10K). Finally, we found that the impairment of iron uptake in the CSF has functional consequences for these macrophages: Both respiratory burst and phagocytosis were impaired in the setting of LM in a LCN2-dependent fashion (Fig. 5, D and E). Together, these observations are consistent with a model in which inflammatory signals promote cancer cell LCN2 production. This in turn allows cancer cells to acquire the iron that is present in limiting amounts in CSF, which not only supports their own growth but inhibits iron uptake and iron-dependent functional activities of macrophages.

Analysis of iron chelation therapy for LM in mice

Because iron is limiting in the CSF, we reasoned that iron chelation might impair cancer cell growth in the CSF and tested this hypothesis in our mouse models (table S2). We inoculated recipient mice with either MDA231-LeptoM or PC9-LeptoM cells and treated them intracisternally with vehicle, the iron chelator deferoxamine (DFO), or the copper chelator D-penicillamine (D-Pen) on day 0 or day 7 after engraftment and every 3 days thereafter (fig. S11, A and B). DFO treatment substantially suppressed iron levels within the LeptoM cells as well as their growth (Fig. 5F and fig. S11, C to F). Notably, DFO treatment conferred a survival benefit to the MDA231 and PC9 mouse models relative to vehicle control (Fig. 5G, fig. S11, G and H, and table S4). As expected for this chelator treatment, we found that iron concentration in the CSF from DFO-treated mice was decreased at day 28 relative to that in control mice (fig. S11, I and J). The number of CSF macrophages was slightly reduced by the D-Pen or DFO treatment, whereas the numbers of neutrophils, T cells, and monocytes were not affected (fig. S11, K and L).

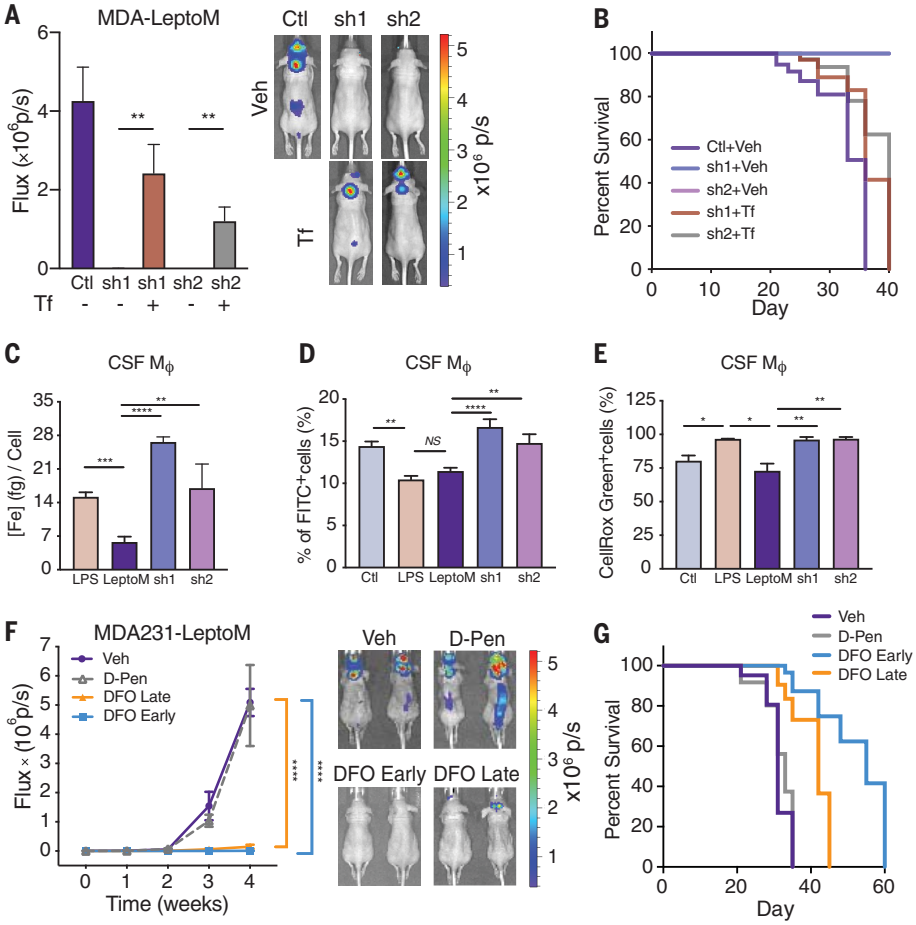


Fig. 5. LCN2 transports iron to support cancer cell growth. (**A**) Tumor growth in mice injected with MDA231-LeptoM cells expressing shCtl or shLCN2. Treatment with either vehicle or diferric holotransferrin (Tf; 5 mg/ml) began on day 1 and continued every third day for a total of seven doses. Left: Histogram represents in vivo BLI imaging after inoculation at day 21; n = 8 to 10 per group in each of two independent experiments. Right: Representative BLI images. $^{**}P < 0.01$ (unpaired t test). Data are means ± SEM. (**B**) Kaplan-Meier survival curve of LeptoM groups from (A). sh1 + Veh versus sh1 + Tf, $P < 0.0001$. sh2 + Veh versus sh2 + Tf, $P < 0.0001$. (**C**) Soluble iron concentration from macrophages collected from CSF in LPS-stimulated, LeptoM, or LeptoM-shLCN2 (sh1 and sh2) mouse model at day 14. Soluble iron was measured by mass spectroscopy. n = 4 per group. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ (unpaired t test). Data are means ± SEM. (**D** and **E**) Phagocytosis (D) and reactive oxygen species (ROS) generation (E) in CSF-derived macrophages by flow cytometric analysis after LeptoM (LLC model), shLCN2 LPS, or vehicle (Ctl) treatment. Phagocytosis was detected by fluorescein isothiocyanate (FITC)–labeled *Escherichia coli*; CellROX Green$^+$ cells represent percent of live cells with ROS. n = 12 to 15 (Ctl), n = 4 or 5 (LPS), n = 11 to 18 (LeptoM), n = 8 (shLCN2). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{****}P < 0.0001$ (unpaired t test). Data are means ± SEM (two independent experiments). (**F**) Tumor growth in MDA231-LeptoM model after chelator treatment: deferoxamine (DFO) or D-penicillamine (D-Pen) on day 0 (early) or day 7 (late); see also fig. S11. Left: Results from in vivo BLI imaging on day 28 after inoculation; n = 8 to 10 per group in each of two independent experiments. Right: Representative BLI images. $^{****}P < 0.0001$. Data are means ± SEM. (**G**) Kaplan-Meier survival curve of mice treated in (F). $P < 0.0001$.

Discussion

Circulating immune cells routinely encounter nutritionally sparse environments as they migrate from the circulation into the tissues. Indeed, immune cells use a discrete set of transcriptional programs to make use of limited resources, including iron and oxygen. In the case of inflammatory macrophages, interferon-γ signaling promotes the generation of nitric oxide (*23*), impairing oxidative phosphorylation (*24*) and allowing for generation of reactive oxygen species (*25*).

Cancer cells cope with challenging environmental constraints differently from immune cells. The genetic heterogeneity of cancer cells provides these cells with a selective advantage over other cell types. To study cellular competition for sparse nutrients at human scale, we applied single-cell RNA-seq to patient-derived samples of LM, a lethal complication of cancer. In doing so, we uncovered a specific example of how the dynamic transcriptional heterogeneity of cancer cells can confer a functional selective advantage (*26*–*28*). We found that cancer cells make use of LCN2/SLC22A17, a high-affinity iron collection system that enables them to effectively outcompete other cells in the leptomeninges for sparse environmental iron. The other major iron-utilizing cell in the CSF, the macrophage, is rendered iron-deficient by this process, resulting in impaired respiratory burst and phagocytosis.

The leptomeninges pose unusual constraints on both the infiltrating immune system and cancer cells; we found that intercellular signaling is substantially altered within the leptomeningeal space. Investigators studying other experimental systems focused on extracranial sites have observed macrophage LCN2 generation (*29*), which promotes cancer cell migration and invasion (*13*, *30*). In LM, CSF macrophages do not produce LCN2. Rather, the inflammatory microenvironment promotes the generation of LCN2 in cancer cells. As obligate partners for cancer cell migration, invasion, and metastasis, tumor-associated macrophages are well known to alter the behavior of cancer cells (*31*). The evolutionary dynamics that we have discovered between malignant and nonmalignant cells within the leptomeninges reveal both the robust nature of cancer's transcriptional plasticity and microenvironmental vulnerabilities ripe for therapeutic exploitation.

REFERENCES AND NOTES

1. S. Kesari, T. T. Batchelor, *Neurol. Clin.* **21**, 25–66 (2003).
2. K. Oechsle, V. Lange-Brock, A. Kruell, C. Bokemeyer, M. de Wit, *J. Cancer Res. Clin. Oncol.* **136**, 1729–1735 (2010).
3. R. Spector, S. Robert Snodgrass, C. E. Johanson, *Exp. Neurol.* **273**, 57–68 (2015).
4. D. Subirá *et al.*, *Neuro-oncol.* **14**, 43–52 (2012).
5. E. Nemeth *et al.*, *J. Clin. Invest.* **113**, 1271–1276 (2004).
6. J. Yang *et al.*, *Mol. Cell* **10**, 1045–1056 (2002).
7. D. H. Goetz *et al.*, *Mol. Cell* **10**, 1033–1043 (2002).
8. L. Kjeldsen, D. F. Bainton, H. Sengeløv, N. Borregaard, *Blood* **83**, 799–807 (1994).
9. C. Mertens *et al.*, *PLOS ONE* **11**, e0166164 (2016).
10. C. Mertens *et al.*, *OncoImmunology* **7**, e1408751 (2017).
11. A. Boire *et al.*, *Cell* **168**, 1101–1113.e13 (2017).
12. M. Jung *et al.*, *Sci. Signal.* **9**, ra64 (2016).
13. L. R. Devireddy, C. Gazin, X. Zhu, M. R. Green, *Cell* **123**, 1293–1305 (2005).
14. I. Mosialou *et al.*, *Nature* **543**, 385–390 (2017).
15. N. Hamzic, A. Blomqvist, C. Nilsberth, *J. Neuroendocrinol.* **25**, 271–280 (2013).
16. Q. Liu, M. Nilsen-Hamilton, *J. Biol. Chem.* **270**, 22565–22570 (1995).
17. T. H. Flo *et al.*, *Nature* **432**, 917–921 (2004).
18. A. R. Moschen *et al.*, *Cell Host Microbe* **19**, 455–469 (2016).
19. D. M. Mosser, J. P. Edwards, *Nat. Rev. Immunol.* **8**, 958–969 (2008).
20. S. V. Torti, F. M. Torti, *Nat. Rev. Cancer* **13**, 342–355 (2013).
21. J. H. Cha *et al.*, *Nat. Commun.* **5**, 4952 (2014).
22. G. Cairo, S. Recalcati, A. Mantovani, M. Locati, *Trends Immunol.* **32**, 241–247 (2011).

23. R. J. Arts *et al*., *Cell Metab.* **24**, 807–819 (2016).
24. J. Van den Bossche *et al*., *Cell Rep.* **17**, 684–696 (2016).
25. E. L. Mills *et al*., *Cell* **167**, 457–470.e13 (2016).
26. E. Azizi *et al*., *Cell* **174**, 1293–1308.e36 (2018).
27. I. Tirosh *et al*., *Science* **352**, 189–196 (2016).
28. Y. Lavin *et al*., *Cell* **169**, 750–765.e17 (2017).
29. M. Jung *et al*., *Mol. Cell. Biol.* **32**, 3938–3948 (2012).
30. X. Duan *et al*., *Int. J. Physiol. Pathophysiol. Pharmacol.* **10**, 105–114 (2018).
31. J. Condeelis, J. W. Pollard, *Cell* **124**, 263–266 (2006).
32. A. Gayoso, J. Shor, Doublet Detection (version v2.4). Zenodo, DOI:10.5281/zenodo.2658730 (2018).
33. F. A. Wolf, P. Angerer, F. J. Theis, *Genome Biol.* **19**, 15 (2018).
34. L. Haghverdi, A. T. L. Lun, M. D. Morgan, J. C. Marioni, *Nat. Biotechnol.* **36**, 421–427 (2018).
35. J. H. Levine *et al*., *Cell* **162**, 184–197 (2015).
36. L. McInnes, J. Healy, J. Melville, arXiv 1802.03426 [stat.ML] (6 December 2018).
37. D. van Dijk *et al*., *Cell* **174**, 716–729.e27 (2018).
38. M. D. Turner, B. Nedjai, T. Hurst, D. J. Pennington, *Biochim. Biophys. Acta* **1843**, 2563–2582 (2014).
39. G. Finak *et al*., *Genome Biol.* **16**, 278 (2015).
40. R. R. Coifman, M. Maggioni, S. W. Zucker, I. G. Kevrekidis, *Curr. Opin. Neurobiol.* **15**, 576–584 (2005).
41. M. Setty *et al*., *Nat. Biotechnol.* **37**, 451–460 (2019).
42. A. Subramanian *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
43. J. C. Reijneveld, M. J. Taphoorn, E. E. Voest, *J. Neurooncol.* **42**, 137–142 (1999).
44. Y. Ushio, N. L. Chernik, J. B. Posner, W. R. Shapiro, *J. Neuropathol. Exp. Neurol.* **36**, 228–244 (1977).
45. D. X. Nguyen *et al*., *Cell* **138**, 51–62 (2009).
46. O. Butovsky *et al*., *Nat. Neurosci.* **17**, 131–143 (2014).
47. M. I. Love, W. Huber, S. Anders, *Genome Biol.* **15**, 550 (2014).
48. V. K. Mootha *et al*., *Nat. Genet.* **34**, 267–273 (2003).

ACKNOWLEDGMENTS

We are deeply grateful to the patients and families who donated clinical samples for this research. **Funding:** Supported by National Cancer Institute (NCI) grants P30 CA008748 and R35 CA220508 (C.I.-D.) and NCI grant 1U2CCA233284 (D.P.), Pew Charitable Trusts grant GC241069 (A.B.), Damon Runyon Cancer Research Foundation grant GC240764 (A.B.), Pershing Square Sohn Cancer Research Alliance grant GC239280 (A.B.), the Baker Family Foundation (A.B.), American Brain Tumor Association grant BRF1900019 (J.R. and A.B.), the Robert J. Kleberg Jr. and Helen C. Kleberg Foundation (C.I.-D.), and the Alan and Sandra Gerry Metastasis and Tumor Ecosystems Center (D.P.). **Author contributions:** A.B. and Y.C. conceived the project; A.B., Y.C., and J.R. designed the experiments; Y.C., J.R., and C.D. performed mouse work; V.K., L.M., and D.P. performed and analyzed scRNA-seq experiments; U.S. collected and clinically annotated human spinal fluid; T.B. and C.I.-D. collected and analyzed autopsy tissues and spinal fluid; C.D., M.A., F.S., and K.N. performed ELISAs and immunohistochemistry; T.T. performed all mass spectroscopy work; and Y.C., J.R., and A.B. wrote the manuscript. **Competing interests:** A.B. is an inventor on U.S. provisional patent application 62/258,044, "Modulating Permeability of the Blood Cerebrospinal Fluid Barrier," filed by Memorial Sloan Kettering Cancer Center. A.B. is an unpaid member of the Scientific Advisory Board of EVREN Technologies. **Data and materials availability:** Single-cell and bulk RNA-sequencing data have been deposited to NCBI GEO as GSE150681 SuperSeries. Cell lines LLC-LeptoM, PC9-LeptoM, and MDA231-LeptoM are available from A.B.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/276/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S7
References (*32*–*48*)

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

21 August 2019; accepted 1 June 2020
10.1126/science.aaz2193

HUMAN EVOLUTION

Ancient DNA indicates human population shifts and admixture in northern and southern China

Melinda A. Yang[1,2,3], Xuechun Fan[4,5], Bo Sun[6], Chungyu Chen[7], Jianfeng Lang[8], Ying-Chin Ko[9], Cheng-hwa Tsang[10], Hunglin Chiu[10], Tianyi Wang[1,2,11], Qingchuan Bao[12], Xiaohong Wu[13], Mateja Hajdinjak[14], Albert Min-Shan Ko[1], Manyu Ding[1,2,15], Peng Cao[1,2], Ruowei Yang[1,2], Feng Liu[1,2], Birgit Nickel[13], Qingyan Dai[1,2], Xiaotian Feng[1,2], Lizhao Zhang[1,2], Chengkai Sun[16], Chao Ning[17], Wen Zeng[18], Yongsheng Zhao[18], Ming Zhang[1,2,15], Xing Gao[1,2,15], Yinqiu Cui[17], David Reich[19,20,21,22], Mark Stoneking[14], Qiaomei Fu[1,2,15]*

Human genetic history in East Asia is poorly understood. To clarify population relationships, we obtained genome-wide data from 26 ancient individuals from northern and southern East Asia spanning 9500 to 300 years ago. Genetic differentiation in this region was higher in the past than the present, which reflects a major episode of admixture involving northern East Asian ancestry spreading across southern East Asia after the Neolithic, thereby transforming the genetic ancestry of southern China. Mainland southern East Asian and Taiwan Strait island samples from the Neolithic show clear connections with modern and ancient individuals with Austronesian-related ancestry, which supports an origin in southern China for proto-Austronesians. Connections among Neolithic coastal groups from Siberia and Japan to Vietnam indicate that migration and gene flow played an important role in the prehistory of coastal Asia.

East Asia, which consists of China, Mongolia, the Korean peninsula, and nearby islands, is home to almost one-quarter of the world's population and harbors a diverse array of ethnic groups and linguistic backgrounds (*1*). However, the genetic history of East Asia, especially in China, is poorly understood. Patterns of genetic relatedness among present-day East Asians run along a north-to-south cline (*2*–*4*), and high levels of genetic drift in East Asia suggest that East Asian populations underwent strong population bottlenecks before the Holocene and to a greater degree than European populations (*5*). Ancient DNA studies have identified how East Asian ancestry affected populations in Southeast Asia (*6*, *7*), the Eastern Steppe (*8*), and northeastern Siberia (*9*). These studies also indicate that ancient individuals from the Southwest Pacific Islands share a close relationship with present-day Taiwanese Austronesians, who in turn share a close relationship with mainland East Asians (*10*).

In most studies to date, present-day East Asian populations (e.g., Han or Dai) have been used to represent East Asian ancestry in modeling studies. However, the archaeological record suggests that East Asians may have been more diverse in the past than they are today (*11*, *12*). This genetic diversity is not well studied, largely owing to a lack of sampling, which makes it difficult to characterize past population structure in northern and southern East Asia and limits inferences of how past populations affected extant East Asians.

A craniometric study on past and present humans has suggested that human history in Asia is characterized by two so-called layers of ancestry: a first layer composed of pre-Neolithic hunter-gatherers with a second layer of northern East Asians who spread across Asia from the Early Neolithic to the present (*13*), contributing ancestry to many East Asians today. Obtaining genetic data from Neolithic East Asians, particularly those from China, would help resolve the role they played in forming the genetic patterns of present-day East Asians.

[1]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences (CAS), Beijing 100044. [2]Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing 100044. [3]Department of Biology, University of Richmond, Richmond, VA 23173. [4]International Research Center for Austronesian Archaeology, Pingtan 350000. [5]Fujian Museum, Fuzhou 350001. [6]Shandong Provincial Institute of Cultural Relics and Archaeology, Jinan 250012. [7]Institute of History and Philology, Academia Sinica, Taipei 11529. [8]School of History and Culture, Shandong University, Jinan 250100. [9]Environment-Omics-Disease Research Center, China Medical University and Hospital, Taichung 40402. [10]Institute of Anthropology, National Tsinghua University, Hsinchu 30013. [11]School of Cultural Heritage, Northwest University, Xi'an 710069. [12]The Institute of Cultural Relics and Archaeology, Inner Mongolia Autonomous Region 010011. [13]School of Archaeology and Museology, Peking University, Beijing 100871. [14]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig. [15]University of Chinese Academy of Sciences, Beijing 100049. [16]Shandong Museum, Jinan 250014. [17]School of Life Sciences, Jilin University, Changchun 130023. [18]Institute of Cultural Heritage, Shandong University, Qingdao 266237. [19]Department of Genetics, Harvard Medical School, Boston, MA 02115. [20]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138. [21]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142. [22]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115.
*Corresponding author. Email: fuqiaomei@ivpp.ac.cn

Genome-wide data from Neolithic East Asians shows their close relationship to East Asians today

To gain insights into the genetic history of East Asians during the Neolithic, we sampled genetic material from ancient individuals across East Asia dating to 9500 to 300 calibrated years before the present (cal yr B.P.) (Fig. 1, A and B) (*14*). In northern East Asia (defined as north of the Qinling-Huaihe line; Fig. 1A), we sampled from the northern Chinese provinces of Inner Mongolia and Shandong. In southern East Asia, we sampled from the southern Chinese province of Fujian in mainland East Asia as well as two Taiwan Strait islands (Fig. 1A, Table 1, and table S1). We used large-scale ancient nuclear DNA capture techniques (*15*) to enrich for endogenous DNA at 1.2 million single-nucleotide polymorphisms (SNPs) (*16*) in 26 individuals. Of these, 24 individuals passed our analysis filters. These filters include the following: (i) identification of characteristic ancient DNA damage signatures that suggest the presence of endogenous DNA (*17*); (ii) assessment of modern human contamination rates to determine for each sample whether to include all fragments (≤3% contamination) or only those with the characteristic damage signature (>3% contamination) (*18*, *19*); and (iii) exclusion of individuals where a contamination rate could not be estimated (*14*). In total, we obtained genetic information from 24 individuals from 11 sites, sequenced to between 0.01- and 7.60-fold coverage at the targeted SNPs (Table 1), with 16 individuals sampled from southern East Asia and 8 individuals sampled from northern East Asia.

To determine whether any individuals had ancestry that deeply diverged from that of present-day East Asians, we first investigated the extent to which they shared ancestry with previously sampled ancient Asians who separated early from the common ancestor of East Asians, referred to as early Asians [e.g., 8000- to 4000-year-old Hòabìnhians from Laos and Malaysia in Southeast Asia (*7*), the 3000-year-old Ikawazu individual from Japan (*7*), and the 40,000-year-old Tianyuan individual from Beijing, China (*20*)] (table S1). In a principal components analysis (PCA) that included ancient and present-day Asians (*14*), all Neolithic East Asians cluster with populations of East Asian ancestry (Fig. 1C and fig. S1), which include present-day East Asians and Neolithic Asians from Siberia, Tibet, Southeast Asia, and the Southwest Pacific who have primarily East Asian–related ancestry (*8*–*10*, *21*). Notably, this includes Early Neolithic southern East Asians (Qihe and Liangdao), who have cranial morphology that clustered with that of early Asians (*13*). Thus, our results fail to support the version of the two-layer model in which these individuals are included in the

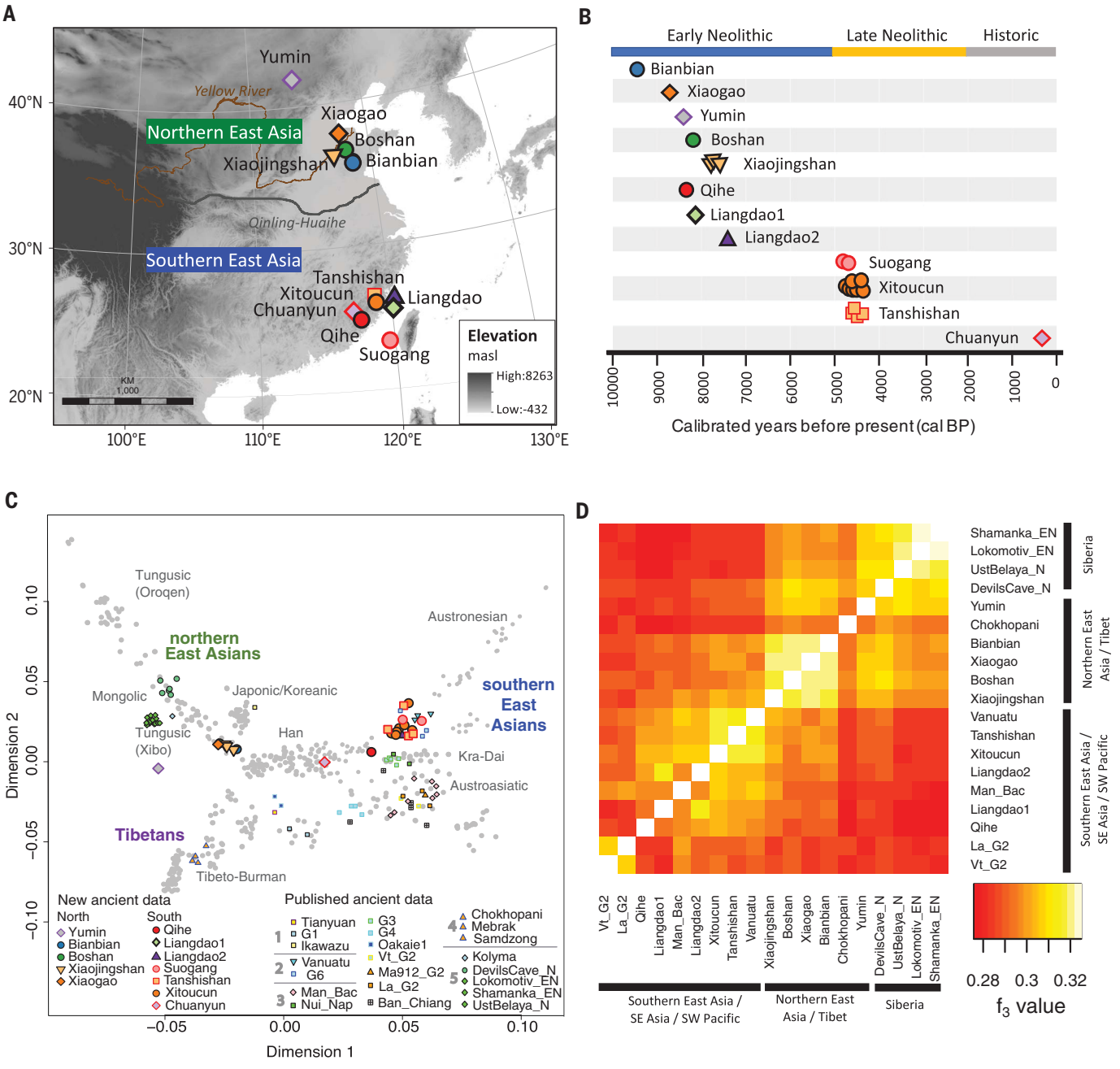


Fig. 1. Geographic, temporal, and genetic information for newly sampled East Asians. (**A**) Map showing the locations from which ancient individuals were sampled for this study. The Qinling-Huaihe line between the Huaihe River and the Qinling mountains is used to divide northern and southern East Asia. Gray shading indicates differences in elevation across the region. masl, meters above sea level. (**B**) Chronology of ancient individuals sampled in this study, where individuals date to 9500 to 7500 cal yr B.P. (Early Neolithic), 5000 to 4000 cal yr B.P. (Late Neolithic), or 300 cal yr B.P. (Historic). Individuals from northern East Asia all date to the Early Neolithic, whereas individuals from southern East Asia date to all three time periods. (**C**) Projection of ancient Asians onto a principal components analysis for present-day East Asians. Present-day East Asians (gray dots) form three main vertices—northern East Asians, southern East Asians, and Tibetans—that correspond to the major present-day East Asian populations. The gray text refers to language groups associated with present-day populations, and a more detailed listing of ethnic groups included can be found in table S1 or fig. S1. Ancient published samples are grouped by archaeological location and numbered in gray. 1 indicates Early Asians; 2 indicates ancient Austronesian-related islanders from Southeast Asia and the Southwest Pacific; 3 indicates Neolithic, Bronze Age, and historical Southeast Asians; 4 indicates Neolithic, Bronze Age, and historical Tibetans; and 5 indicates Paleosiberian (Kolyma) and Neolithic Siberians. (**D**) Pairwise outgroup f_3 statistics for ancient East Asians, where yellow indicates high genetic similarity between pairs. The outgroup f_3 test takes the form f_3(*Mbuti*; *X*, *Y*) where *X* and *Y* are Neolithic Asians listed in the rows and columns and Mbuti (a present-day population from Central Africa) is the outgroup. SE Asia refers to Southeast Asia, and SW Pacific refers to Southwest Pacific.

first layer. Consistent results were obtained in an outgroup f_3 analysis, where Neolithic East Asians were found to share more genetic similarity with Neolithic Siberians, Tibetans, and Southwest Pacific Islanders (f_3 = 0.28 to 0.32) than with early Asians (f_3 = 0.25 to 0.26; fig. S2).

A direct comparison of Neolithic East Asians with present-day East Asians and early Asians in a symmetry test shows that Neolithic East Asians tend to be more closely related to present-day East Asians than to any early Asian. In other words, most f_4(*Mbuti*, *X*; *present-day East Asian*, *early Asian*) values are significantly <0 (Han: $-22.1 < Z < -4.4$; table S2) and most f_4(*Mbuti*, *early Asian*; *present-day East Asian*, *X*) values are ~0 (Han: $-2.4 < Z < 2.6$; table S2). Thus, the newly sampled individuals, including those from Early Neolithic southern East Asia, are genetically closest to modern populations of East Asian ancestry.

We also investigated patterns of archaic admixture in the Neolithic East Asians and found them to be similar to patterns observed in present-day East Asians (table S3). Neolithic East Asians do not show consistent connections with a 35,000-year-old individual from Belgium, nor do they show connections to any South Americans—patterns that differ from those seen in Tianyuan (*14*, *20*) but are similar to those in present-day East Asians (fig. S3). Previous ancient DNA studies have found deeply divergent ancestries in pre-Neolithic Asia (*7*, *20*) associated with a potential first layer of populations in Asia (*13*). However, populations from the Early Neolithic in both northern and southern China show different patterns, which suggests that they do not belong to the first layer and already carry ancestry primarily associated with present-day East Asians.

Population division between northern and southern East Asians in the Early Neolithic

In the PCA, our newly sampled individuals separate along geographic lines. Coastal Neolithic southern East Asians cluster together and close to present-day southern East Asians, whereas coastal Neolithic northern East Asians cluster together and close to present-day northern East Asians (Fig. 1C). These results suggest that population structure between northern and southern East Asia has existed since the Early Neolithic.

To determine genetic relationships in Neolithic East Asia, we first used a pairwise comparison of outgroup f_3 values (*22*) to assess their relationship to other Neolithic populations with East Asian–related ancestry (*14*). Neolithic southern East Asians share high genetic similarity to each other, Neolithic Southeast Asians, and Austronesian-related islanders from the Southwest Pacific (Fig. 1D)—a pattern also observed in the PCA (Fig. 1C). In an f_4 analysis (fig. S4A), Neolithic southern East Asians and the 3000-year-old Austronesian-related islanders from the Southwest Pacific (*10*) consistently share a closer genetic relationship with each other than with coastal and inland Neolithic northern East Asians, Siberians (*8*, *9*), and Tibetans (*21*). Although some Late Neolithic southern East Asians share a connection to northern East Asians (fig. S4B), we suggest that this is related to admixture. In a maximum-likelihood phylogeny allowing migration events (*23*), Neolithic southern East Asians group together relative to Neolithic northern East Asians, Siberians, and Tibetans (Fig. 2A and figs. S5 and S6) in 99.3 to 100% of the bootstrap (bs) replicates. Thus, in mainland southern East Asia and Taiwan Strait islands, we find shared southern East Asian ancestry that differs from that observed in Neolithic northern East Asians—a pattern that persists even with the evidence of admixture that we describe below.

In northern East Asians dating to the Early Neolithic, we find shared ancestry that is not found in Early Neolithic southern East Asians but is present in Neolithic Siberians and Tibetans. For example, in a maximum-likelihood phylogeny, Neolithic northern East Asians, Siberians, and Tibetans group together relative to Neolithic southern East Asians and Austronesian-related islanders (bs = 95.3 to 99%; Fig. 2A and figs. S5 and S6), and they also share high genetic similarity with each other in an outgroup f_3 analysis (Fig. 1D). Particularly, coastal Neolithic northern East Asians clustered together, whereas Yumin, an inland Neolithic northern East Asian, clustered with Neolithic Siberians from the Eastern Steppe and the Primorye region of Far East Russia (*8*, *9*). Most Neolithic

Table 1. Ancient individuals sampled in this study. Red highlighting indicates high contamination, and yellow highlighting indicates that the sample was excluded from demographic analysis because of a low number of SNPs. Lat, latitude; Lon, longitude; C→T%, percentage of C→T substitutions; Covg, coverage; DS, double-stranded; SS, single-stranded; SS UDG, single-stranded with uracil-DNA-glycosylase treatment.

ID	# of libraries	Sample ID	Population	Name†	Skeletal element	Date cal yr B.P.	Region	Lat	Lon	Library style	Sex	C→T%	Contamination Chr X	Contamination MtDNA	Covg#	SNPs
Yumin	1	M1	inland nEastAsia_EN	Yumin	Teeth	8,415-8,335	Inner Mongolia	42	114.2	DS	F	17	-	0.028	7.25	876757
Bianbian	2	Bianbian	coastal nEastAsia_EN	Bianbian	Petrous	9,545-9,480	Shandong	36.1	118.5	SS,DS	M	11	0.01	0.014	2.21	701880
BS	11	BS11	coastal nEastAsia_EN	Boshan	Tibia	8,320-8,040&	Shandong	36.5	117.9	SS UDG	M	10	0.03	0.01	7.36	1045084
XJS1309_M7	3	M7	coastal nEastAsia_EN	Xiaojingshan	Petrous	7,872-7,721	Shandong	36.5	117.9	DS	F	31	-	0.018	0.22	191653
XJS1311_M16	3	M16	coastal nEastAsia_EN	Xiaojingshan	Petrous	7,935-7,786	Shandong	36.5	117	DS	F	24	-	0.029	0.64	387603
XJS1309_M4	4	M4	coastal nEastAsia_EN	Xiaojingshan	Petrous	7,877-7,735	Shandong	36.5	117	DS	M	21	-	0.019	0.58	386648
Xiaogao	4	M1	coastal nEastAsia_EN	Xiaogao	Petrous	8,777-8,591	Shandong	37.9	117.6	DS	F	30	-	0.018	7.6	948953
Qihe2	4	Qihe2	coastal sEastAsia_EN	Qihe	Petrous	8,428-8,359	Fujian	25.4	117.6	SS	F	34	-	0.13	0.45^	328913^
LD1	1	M1	island sEastAsia_EN	Liangdao1	Phalanx	8,320-8,060&	Liang Island	26.3	120.2	SS UDG	M	12	0.03	0.015	2.72	843051
LD2	1	M2	island sEastAsia_EN	Liangdao2	Phalanx	7,590-7,560&	Liang Island	26.3	120.2	SS UDG	F	19	-	0.005	1.68	744266
SuogangB1	7	B1	island sEastAsia_ LN	Suogang	Petrous	4,800-4,300*&	Penghu Island	23.5	119.6	SS	M	8	-0.02	0.12	0.04^	48386^
SuogangB3	7	B3	island sEastAsia_ LN	Suogang	Petrous	4,800-4,300*&	Penghu Island	23.5	119.6	SS	M	10	0.11	0.28	0.03^	30050^
L5705	1	M49	coastal sEastAsia_LN	Xitoucun	Petrous	4,419-4,246	Fujian	26.2	119.1	SS	F	33	-	0.026	0.66	434148
L5700	1	M32	coastal sEastAsia_LN	Xitoucun	Petrous	4,530-4,417	Fujian	26.2	119.1	SS	F	28	-	0.027	0.43	307394
L5692	1	M26	coastal sEastAsia_LN	Xitoucun	Petrous	4,530-4,417	Fujian	26.2	119.1	SS	M	24	-0.02	0.058	0.01^	16027^
L5706	1	M15	coastal sEastAsia_LN	Xitoucun	Petrous	4,527-4,406	Fujian	26.2	119.1	SS	M	25	0.02	0.043	0.04^	49141^
L5704	1	M13	coastal sEastAsia_LN	Xitoucun	Petrous	4,580-4,423	Fujian	26.2	119.1	SS	M	23	0.11	0.11	0.04^	47786^
L5703	1	M44	coastal sEastAsia_LN	Xitoucun	Petrous	4,644-4,500	Fujian	26.2	119.1	SS	M	22	0.12	0.074	0.02^	19114^
L5702	1	M46	coastal sEastAsia_LN	Xitoucun	Petrous	-	Fujian	26.2	119.1	SS	M	13	-	-	0.01	9407
L5701	1	M18 -2	coastal sEastAsia_LN	Xitoucun	Petrous	4,418-4,240	Fujian	26.2	119.1	SS	M	28	0.07	0.089	0.04^	47608^
L7415	1	M6	coastal sEastAsia_LN	Tanshishan	Petrous	4,419-4,246	Fujian	26.1	119.2	SS	M	22	0.03	0.03	0.41	338182
L7417	1	M20	coastal sEastAsia_LN	Tanshishan	Petrous	4,526-4,417	Fujian	26.1	119.2	SS	F	24	-	0.035	0.14^	144349^
L7416	1	M25	coastal sEastAsia_LN	Tanshishan	Petrous	-	Fujian	26.2	119.1	SS	M	11	-	-	0.01	8422
L5698	1	M26 -1	coastal sEastAsia_LN	Tanshishan	Petrous	4,410-4,225	Fujian	26.2	119.1	SS	F	25	-	0.06	0.02^	20920^
L5696	1	M12	coastal sEastAsia_LN	Tanshishan	Petrous	-	Fujian	26.2	119.1	SS	M	18	0.16	0.196	0.02^	20107^
L5694	1	Chuanyun	coastal sEastAsia_H	Chuanyun	Toe	334 281	Fujian	25.6	117.3	SS	M	11	0.01	0.005	0.38	322808

†Labels used in analyses, which is the same as the site name. If the same name, they were grouped together. #Based on a panel of 1.2 million SNPs (*14*–*16*). &Radiocarbon date published elsewhere. *The bone was not directly radiocarbon dated. For the Suogang site, shells were radiocarbon dated to 4633 to 4287 cal yr B.P. and 4793 to 4407 cal yr B.P. ^Alleles were determined using the damage-restricted library.

northern East Asians and Siberians share a closer genetic relationship to each other than to Neolithic southern East Asians (figs. S7 and S8). In a model developed using all f_2,f_3, and f_4 statistics for populations included in the model [admixture graph; details provided in (*14*)], we similarly observe that northern and southern East Asian–related ancestry separates into two lineages (Fig. 2B and figs. S16 to S18).

Though the Neolithic is characterized by distinct northern and southern East Asian–related ancestries, gene flow also affected populations during this time. In the PCA, the inland Early Neolithic northern East Asian Yumin does not cluster with other northern East Asians as Neolithic Siberians and coastal northern East Asians do; rather, Yumin falls in between northern and Tibetan present-day East Asians (Fig. 1C). Additionally, two patterns improve the fit of the maximum-likelihood phylogeny (Fig. 2A and figs. S5 and S6). First, Upper Paleolithic northern Siberians influenced the ancestry of Neolithic Siberians (i.e., Yana; Fig. 2A), likely mediated by populations in Siberia that are closely related to Native Americans (Paleosiberians; table S1) as has been previously observed (*9*) and which we confirmed in f_4 analyses showing Paleosiberian connections primarily to Neolithic Siberians (table S4). Second, some coastal Early Neolithic northern East Asians show affinities to coastal Late Neolithic southern East Asians when compared with more–inland Neolithic populations in Asia (northern East Asian, Siberian, and Tibetan; figs. S7 and S8). Coastal

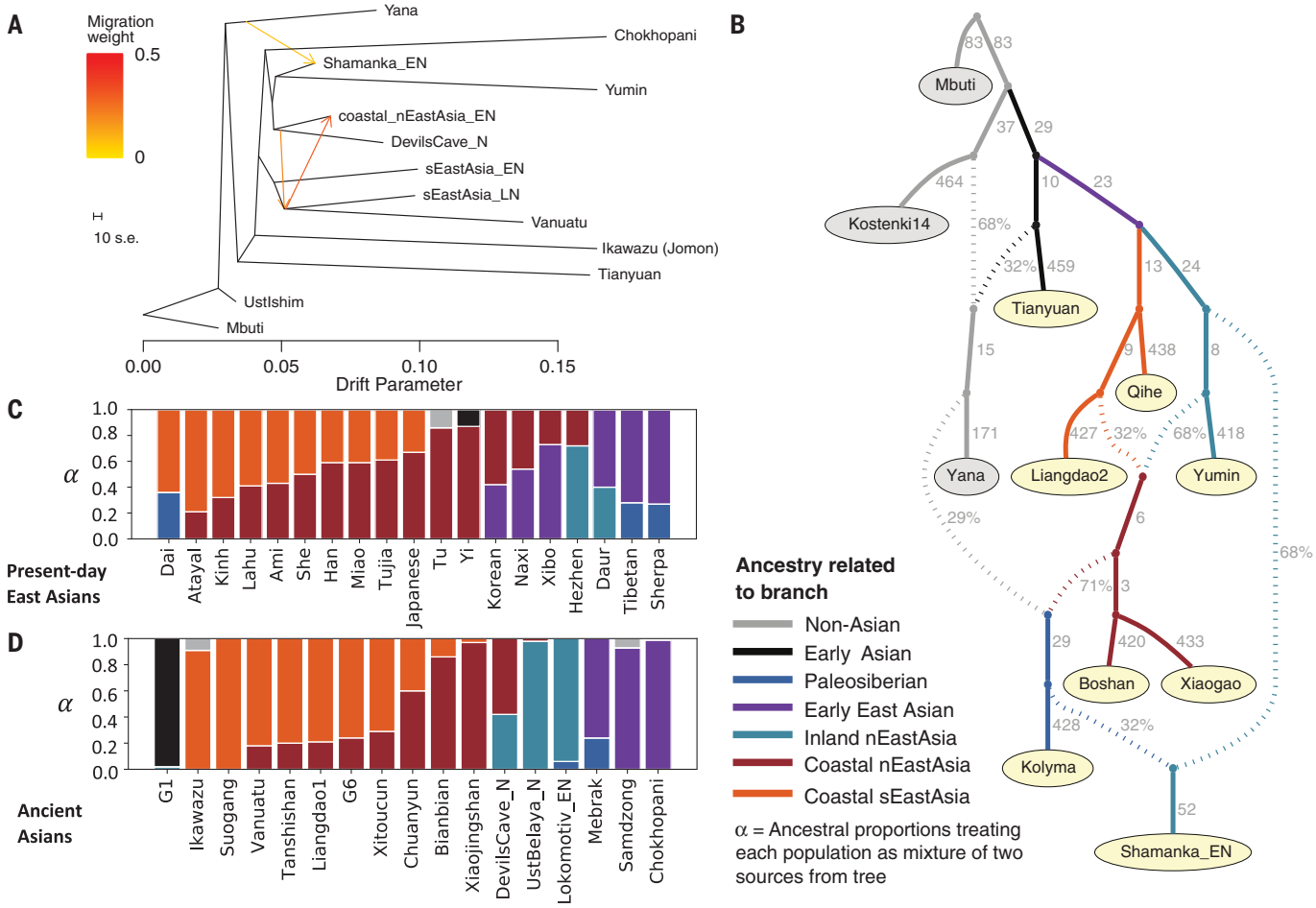


Fig. 2. Modeling tree relationships between ancient Asians. (**A**) A maximum-likelihood phylogeny allowing three migration events using the Treemix software (*23*). Aside from coastal nEastAsia_EN, sEastAsia_EN, and sEastAsia_LN (which refer to the labels in the Population columns in Table 1 or table S1 and group several ancient individuals from more than one site together), all other labels refer to the nomenclature used in the Name columns designated in Table 1 or table S1. s.e., average standard error of the entries in the sample covariance matrix. (**B**) An admixture graph that fits the data ($|\max Z| = 2.7$), where we include two Early Neolithic southern East Asians (sEastAsia: Qihe and Liangdao2), three Early Neolithic northern East Asians (inland and coastal nEastAsia: Yumin, Boshan, and Xiaogao), a Neolithic Siberian (Shamanka_EN), a Paleosiberian (Kolyma), an early Asian (Tianyuan), two individuals of non-Asian ancestry (Kostenki14 and Yana), and the Central African Mbuti, whom we used to root the tree. Branches are denoted by solid lines with branch lengths given in units of 1000 × the f_2 drift distance, and admixture events are denoted by dotted lines with mixture proportions as shown (all numbers rounded to the nearest integer). Other tree models that fit the data are shown in fig. S18. (**C** and **D**) The estimated mixture proportions observed when adding present-day (C) and ancient (D) East Asians to the graph in (B), allowing for each addition to be added as a single node or a mixture of two branches. We considered the tree with the best fit considering all f_2, f_3, and f_4 statistics comparing that individual or population with all tips in the tree. The color coding and associated key refers to the area of the tree where they attach and the estimated mixture proportions. In (A) and (B), we find a separation between Early Neolithic northern and southern East Asians with connections between coastal populations in the north and south.

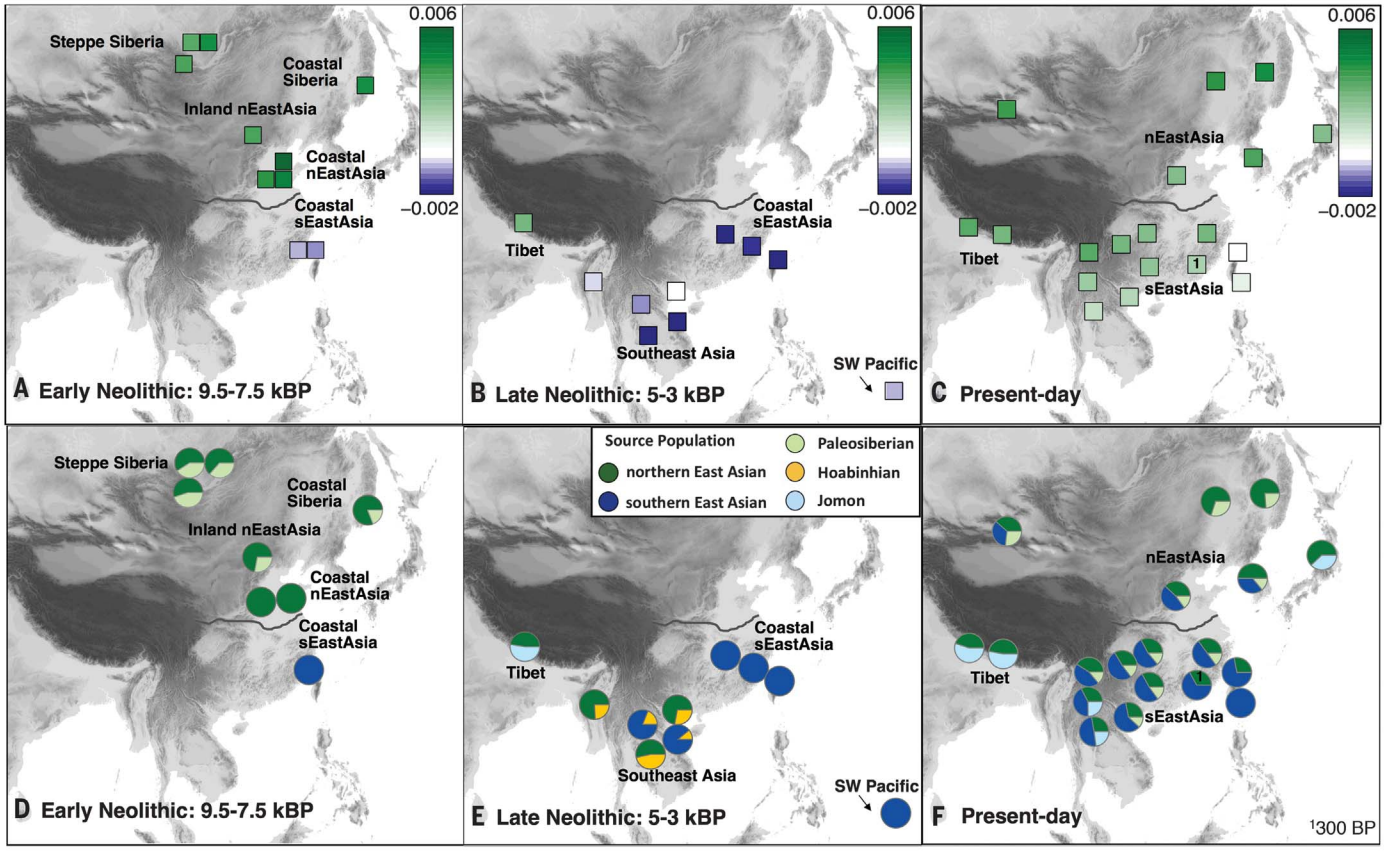


Fig. 3. Ancestry localized to the north and south in the Early Neolithic is found in admixed form across northern and southern East Asia today. (**A** to **C**) A heatmap showing f_4(*Mbuti*, *X*; *Qihe*, *Bianbian*), where Bianbian is the oldest northern East Asian sampled (~9500 cal yr B.P.) and Qihe is the oldest southern East Asian sampled (~8400 cal yr B.P.). *X* are East and Southeast Asians who date to the Early Neolithic (A), Late Neolithic (B), or present day (C). Green indicates more affinity to Early Neolithic northern East Asians, whereas blue indicates greater affinity to Early Neolithic southern East Asians. Numerical values can be found in table S5A. kBP, thousand calibrated years before the present. (**D** to **F**) Ancestry proportions estimated for Neolithic and present-day individuals. Possible ancestries are northern East Asian (green), southern East Asian (blue), Paleosiberian (light green), Hòabìnhian (orange), and Jōmon (light blue). Proportions were determined using *qpAdm* (*15*), with representative sources and outgroups described in (*14*). During the Neolithic, there was a division between East Asians north and south of the Qinling-Huaihe line (dark gray line on map) in ancestry; but in the present day, ancestries previously localized to only northern or southern East Asia can be found in appreciable frequencies in both regions. Results for present-day Han are in fig. S13 and point values are in table S6. Labels on the map indicate the populations listed in table S1.

Early Neolithic northern East Asians show these connections to Early Neolithic southern East Asians as well (fig. S8B and fig. S9). These connections demonstrate that the population relationships among coastal northern and southern Neolithic East Asians cannot be easily clarified without the inclusion of admixture (Fig. 2, B to D, and fig. S10).

Growing influence of northern East Asian ancestry in southern East Asia

In a spatiotemporal analysis, we next assessed differences in East Asian ancestry between the Neolithic and today. A major feature of Neolithic East Asian ancestry described above is a northern ancestry—represented by Neolithic northern East Asians, Siberians, and Tibetans—and a southern ancestry—represented by Neolithic southern East Asians and Austronesian-related islanders. A test of genetic differentiation [fixation index (F_{ST})] shows that coastal Neolithic northern East Asians and coastal Neolithic southern East Asians are more highly differentiated from each other [F_{ST} = 0.042, standard error (SE) = 0.004] than are present-day northern and southern East Asians (F_{ST} = 0.023, SE = 0.01; fig. S11). This difference indicates that East Asians today are more genetically homogeneous than Neolithic East Asians.

To determine what factors led to decreased genetic differentiation between present-day northern and southern East Asians relative to those from the Neolithic, we used an f_4 analysis to set up a symmetry test (Fig. 3, A to C), where we compared ancient and present-day individuals against the oldest northern (Bianbian) and southern (Qihe) East Asians sampled. That is, we tested f_4(*Mbuti*, *X*; *Bianbian*, *Qihe*) to determine northern and southern East Asian affinities across time and space. In the Early and Late Neolithic there is clear spatial separation, with northern populations sharing a closer relationship to Bianbian and southern populations sharing a closer relationship to Qihe (Fig. 3, A and B, and table S5A). This pattern persists when we substitute other Early Neolithic southern East Asians for Qihe and other Early Neolithic northern East Asians for Bianbian (fig. S12 and table S5, B and C). We then used *qpAdm* (*15*) to develop mixture models (Fig. 3, D to F) to estimate ancestry proportions in Neolithic populations (*14*), where Neolithic northern and southern East Asians each belong to a distinct ancestry associated with their location in northern

or southern East Asia (Fig. 3, D and E, and table S6A).

In contrast, patterns observed in present-day East Asians show that the main factor reducing genetic differentiation in present-day East Asians is increased northern East Asian–related ancestry in southern East Asia. In present-day East Asians, we observe a pronounced change where all East Asians (including mainland southern East Asians) share an affinity with Neolithic northern East Asians, as shown by f_4(*Mbuti*, *present-day East Asian*; *Qihe*, *Bianbian*)>0 ($3 < Z < 8.8$; Fig. 3C and table S5A), rather than demonstrating affinities on the basis of location. Estimates of ancestry proportions in our mixture models show that this is because of increased northern East Asian–related ancestry in southern mainland East Asia, with estimates of northern East Asian–related ancestry ranging from 21 to 55% (Fig. 3F, fig. S13, and table S6B). Southern East Asian ancestry also extends north, as it is found in Han populations from northern China (36 to 41%; fig. S13 and table S6B) as well as in some northern East Asians (e.g., Xibo and Korean populations; 35 to 36%; Fig. 3F and table S6B). Paleosiberian-related ancestry from Siberia also greatly affects recent East Asians, with the exception of a 300-year-old individual from coastal mainland southern East Asia, island populations, and Tibetans (Fig. 3F and table S6). This lack of Paleosiberian-related ancestry in the peripheral edges of East Asia suggests that different types of gene flow from north to south occurred in East Asia.

Although we are unable to estimate the timing of gene flow events, patterns in the f_4 comparisons suggest that population movement may have already started affecting East Asians by the Late Neolithic. In the test f_4(*Mbuti*, *Bianbian*; *Kolyma*, *Neolithic southern East Asians*), we would expect similar results for all Neolithic southern East Asians if none of them had northern East Asian ancestry. Instead, the Late Neolithic southern East Asians share a connection to the coastal northern East Asian Bianbian that Early Neolithic southern East Asians do not share (fig. S14). Other tests of admixture further support this finding: The admixture graph (Fig. 3, C and D) and ADMIXTURE (fig. S10) results both estimate northern East Asian–related ancestry in Late Neolithic southern East Asians that is reduced in Early Neolithic southern East Asians (*14*).

We also investigated whether the northern East Asian affinity is closer to inland (i.e., Yumin) or coastal (i.e., Bianbian) northern East Asians. In an f_4 test of symmetry, all Neolithic East Asians and coastal Neolithic Siberians share a closer relationship to Bianbian than to Yumin, whereas all inland Neolithic Siberians and Tibetans do not (fig. S15 and table S5D). This suggests that the northern East Asian ancestry found in all present-day mainland East Asians is primarily related to populations along the lower reaches of the Yellow River. These observations are consistent with archaeological and historical studies that argue for an origin of the Han ethnic group in northern China, along the Yellow River (*24*).

Proto-Austronesian origins in southern China and coastal connections

Austronesian-speaking populations extend from Taiwan to the Southwest Pacific and to Madagascar. Mainland southern China has been proposed to be the likely origin of the proto-Austronesian population that entered Taiwan, on the basis of geographic proximity across the Taiwan Strait and the occurrence of similar artifacts at archaeological sites from both regions (*11*, *25*, *26*). Furthermore, genetic modeling with mitochondrial DNA (mtDNA) genomes from present-day Southeast Asians (*27*) also suggests a southern Chinese origin. The mtDNA of the 8320- to 8060-year-old Liangdao1 individual belongs to haplogroup E1, which is common in Austronesian-speaking populations in Taiwan, the Philippines, and Indonesia today, and it is most similar to mtDNA found in the aboriginal Formosans from Taiwan (*28*).

Our data also support a proto-Austronesian origin in mainland southern China. Neolithic southern East Asians consistently fall on a cline leading to present-day Austronesians rather than to other present-day southern East Asian populations in a PCA (Fig. 1C), and they share a connection with present-day Austronesians (Ami) relative to other present-day southern East Asians, e.g., f_4(*Mbuti*, *Neolithic southern East Asians*; *Dai*, *Ami*)>0 ($5.2 < Z < 12.7$; table S7A). This connection to present-day Austronesians extending across all Neolithic samples tested from mainland southern China and Taiwan Strait islands supports a coastal southern China origin for proto-Austronesians (*29*).

We further investigated connections between the ancient samples and 3000-year-old Southwest Pacific Islanders from Vanuatu (*10*), who share a close relationship to present-day Austronesians. The Neolithic samples from mainland southern China and the Taiwan Strait islands group together with the ancient Vanuatu samples in a maximum-likelihood tree (bs = 82.3 to 100%; Fig. 2A and fig. S5) and share high genetic similarity with the Vanuatu in an outgroup f_3 analysis (Fig. 1D). Moreover, a direct f_4 comparison demonstrates that Late Neolithic southern East Asians share a close genetic relationship with these Austronesian-related Southwest Pacific Islanders (fig. S4A), with no excess connections from either group to Neolithic northern East Asians (fig. S4B). These results further support the proposed connection between Neolithic southern East Asians and proto-Austronesians (*29*).

A lack of genetic isolation between coastal populations can be observed all along the coast of East and Southeast Asia. Although most Neolithic Southeast Asians are a mixture of Hòabìnhian-related and southern East Asian–related ancestry, a 4000-year-old population from Vietnam shares an especially close relationship to coastal Late Neolithic southern East Asians (table S7B). Notably, these coastal connections extend further north. A 2700-year-old Jōmon individual from the Japanese archipelago that separated early from present-day East Asians was suggested to share genetic affinity to Hòabìnhians (*7*), but comparisons including Neolithic East Asians do not show this pattern (table S8). Instead, we find that the Jōmon individual shows affinities to several coastal Neolithic populations in Siberia as well as southern East Asia (table S9). The patterns demonstrated here indicate that coastal regions were areas of interconnectivity and gene flow rather than of isolation. These affinities between populations along the eastern coast of Asia (and island populations off the coast) that are not shared by more–inland Asian populations suggest that interactions along marine-related environments played an important role in the prehistory of coastal Asia.

Discussion

Our genetic survey of Neolithic northern and southern East Asians shows population differentiation during the Early Neolithic to an extent not observed in present-day East Asians. In a craniometric analysis of ancient and present-day Asians, it has been proposed that a second layer of northern East Asian–related populations spread across East Asia around the Early Neolithic, replacing, at least in part, a first layer of pre-Neolithic hunter-gatherers (*13*, *30*). Although we did not find evidence of a first layer population in coastal southern East Asia by 8400 years ago, we did observe increased northern influences in southern East Asia between the Early Neolithic and today. Thus, the argument for the spread of a second layer associated with northern East Asian ancestry is still an important model to explore in the context of East Asian prehistory.

However, the spread of northern East Asian ancestry led to increased admixture in both directions, such that most of today's East Asians are a mixture of northern and southern East Asian ancestries. Thus, not only was there spread of northern East Asian ancestry into southern East Asia, but southern East Asian–related ancestry can be found in some present-day northern East Asians. The fact that we do not

observe admixture to this extent in the Neolithic suggests that much of the human movement that contributed to present-day East Asian genetic patterns must have occurred after the Neolithic.

Shared ancestry amongst ancient individuals on the southeast coast of mainland East Asia, islands in the Taiwan Strait, and the Southwest Pacific island Vanuatu suggests that the Austronesian expansion derived from a population that came from southern China—a pattern supported by similarities in material culture on the southeast coast of China and in Austronesian-related material culture (*11*) as well as in studies of uniparental markers (*27*, *28*). Moreover, gene flow among coastal populations in East Asia is a common trend, as different coastal connections can be observed in Neolithic populations from as far north as coastal Siberia and the Japanese archipelago to as far south as coastal Vietnam.

In individuals from northern and southern China dating to 9500 to 300 cal yr B.P., we observe a close genetic relationship between East Asians, albeit with shifts in ancestry indicative of population movement and admixture during the Neolithic. Increased genetic sampling from the Paleolithic and from populations further inland in central China should help to further clarify the relationships among Paleolithic hunter-gatherers, Neolithic farmers, and present-day populations of East Asia.

REFERENCES AND NOTES

1. F. Zhang, B. Su, Y.-P. Zhang, L. Jin, *Phil. Trans. R. Soc. B* **362**, 987–996 (2007).
2. The HUGO Pan-Asian SNP Consortium, *Science* **326**, 1541–1545 (2009).
3. C. Tian *et al.*, *PLOS ONE* **3**, e3862 (2008).
4. J. Chen *et al.*, *Am. J. Hum. Genet.* **85**, 775–785 (2009).
5. A. Keinan, J. C. Mullikin, N. Patterson, D. Reich, *Nat. Genet.* **39**, 1251–1255 (2007).
6. M. Lipson *et al.*, *Science* **361**, 92–95 (2018).
7. H. McColl *et al.*, *Science* **361**, 88–92 (2018).
8. P. de Barros Damgaard *et al.*, *Science* **360**, eaar7711 (2018).
9. M. Sikora *et al.*, *Nature* **570**, 182–188 (2019).
10. P. Skoglund *et al.*, *Nature* **538**, 510–513 (2016).
11. L. Liu, X. Chen, *The Archaeology of China: From the Late Paleolithic to the Early Bronze Age* (Cambridge Univ. Press, 2012).
12. B. Sun, S. Cui, *Cultural Relics of Central China* **3**, 23–28 (2008).
13. H. Matsumura *et al.*, *Sci. Rep.* **9**, 1451 (2019).
14. See the supplementary materials for more details.
15. W. Haak *et al.*, *Nature* **522**, 207–211 (2015).
16. Q. Fu *et al.*, *Nature* **524**, 216–219 (2015).
17. A. W. Briggs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14616–14621 (2007).
18. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, *BMC Bioinformatics* **15**, 356 (2014).
19. Q. Fu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013).
20. M. A. Yang *et al.*, *Curr. Biol.* **27**, 3202–3208.e9 (2017).
21. C. Jeong *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7485–7490 (2016).
22. M. Raghavan *et al.*, *Nature* **505**, 87–91 (2014).
23. J. K. Pickrell, J. K. Pritchard, *PLOS Genet.* **8**, e1002967 (2012).
24. W. Yan, *Wen Wu* **3**, 38–50 (1987).
25. T. Jiao, in *A Companion to Chinese Archaeology*, A. P. Underhill, Ed. (Wiley, 2013), pp. 599–611.
26. Fujian Museum, Longyan Municipal Bureau of Culture and Publishing, *Kaogu.* **5**, 7–19 (2013).
27. W. Kutanan *et al.*, *Eur. J. Hum. Genet.* **26**, 898–911 (2018).
28. A. M.-S. Ko *et al.*, *Am. J. Hum. Genet.* **94**, 426–436 (2014).
29. P. Bellwood, *Asian Perspect.* **26**, 107–117 (1984).
30. H. Matsumura *et al.*, *PLOS ONE* **13**, e0198689 (2018).
31. Y. Wang *et al.*, *Genomics Proteomics Bioinformatics* **15**, 14–18 (2017).
32. BIG Data Center Members, *Nucleic Acids Res.* **46**, D14–D20 (2018).

ACKNOWLEDGMENTS

We thank J. Kelso, M. Slatkin, and J. Gardner for comments, as well as S. Pääbo and his laboratory from the Max Planck Institute of Evolutionary Anthropology for their help in setting up the laboratory at the Institute of Vertebrate Paleontology and Paleoanthropology, CAS. We also thank our reviewers for providing insightful suggestions on the manuscript. **Funding:** This work was supported by the Chinese Academy of Sciences (XDB26000000), the National Natural Science Foundation of China (91731303, 41925009, and 41630102), CAS (XDA1905010 and QYZDB-SSW-DQC003), Zhengzhou University through the "Research on the roots of Chinese civilization" grant (XKZDJC202006), the Tencent Foundation through the XPLORER PRIZE, and the Howard Hughes Medical Institute (grant no. 55008731). **Author contributions:** Q.F. conceived the idea for the study. X.Fa., B.S., C.C., J.L., C.T., Y.-C.K., H.C., Q.B., X.W., A.M.-S.K., W.Z., Y.Z., C.S., M.Z., X.G., and Q.F. assembled archaeological materials and performed dating. Q.F., P.C., R.Y., F.L., X.Fe., Q.D., C.N., and Y.C. performed or supervised wet laboratory work. M.A.Y. and Q.F. analyzed the data and wrote the manuscript, with critical input from all authors. M.A.Y., T.W., M.D., and Q.F. wrote and edited the supplementary materials. Y.-C.K., X.G., D.R., and M.S. helped revise the manuscript and supplementary materials. M.A.Y. performed analyses and writing at the Institute of Vertebrate Paleontology and Paleoanthropology, with additional revisions done at the University of Richmond. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** BAM files and genotype calls for the newly sequenced individuals are available at the Genome Sequence Archive (*31*) in the Beijing Institute of Genomics (BIG) Data Center (*32*) (https://bigd.big.ac.cn/gsa-human; accession no. HRA000123). All software used is freely available online and is referenced in (*14*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/282/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S18
Tables S1 to S9
References (*33*–*79*)
MDAR Reproducibility Checklist

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

4 November 2019; accepted 1 May 2020
Published online 14 May 2020
10.1126/science.aba0909

REPORT

CHEMICAL PHYSICS

Second critical point in two realistic models of water

Pablo G. Debenedetti[1]*†, Francesco Sciortino[2]†, Gül H. Zerze[1]†

The hypothesis that water has a second critical point at deeply supercooled conditions was formulated to provide a thermodynamically consistent interpretation of numerous experimental observations. A large body of work has been devoted to verifying or falsifying this hypothesis, but no unambiguous experimental proof has yet been found. Here, we use histogram reweighting and large-system scattering calculations to investigate computationally two molecular models of water, TIP4P/2005 and TIP4P/Ice, widely regarded to be among the most accurate classical force fields for this substance. We show that both models have a metastable liquid-liquid critical point at deeply supercooled conditions and that this critical point is consistent with the three-dimensional Ising universality class.

Water affects virtually every aspect of our lives and sculpts the physical environment that we inhabit (*1*); it is a key actor in the major physical and chemical processes that are essential to life as we know it (*2*). Water's ubiquity masks the fact that many of its physical properties are qualitatively different from those of most other substances: Liquid water at ambient pressure is denser than the solid into which it freezes; cold liquid water expands upon cooling and becomes less viscous upon compression; and water has at least 17 crystalline phases (*3*). These are just a few of the numerous examples of water's distinctiveness, which originates in its ability to form hydrogen bonds and the consequent tetrahedral coordination to which they, in conjunction with the water molecule's geometry, give rise (*4*). Tetrahedrally coordinated liquids such as silica and silicon share many of water's anomalies (*5*). The microscopic origin of these anomalies is the liquid's ability to form transient local arrangements characterized by low density, low energy, and low entropy (*6*, *7*).

In 1976, experiments by Speedy and Angell brought to light the fact that water's distinctive behavior becomes more pronounced upon supercooling (*8*). The sharp increases in liquid water's response functions (isothermal compressibility, isobaric heat capacity, and magnitude of the thermal expansion coefficient) upon cooling imply a corresponding enhancement of fluctuations (density, enthalpy, and entropy-volume covariance, respectively) (*6*, *7*). Providing an experimentally testable thermodynamically consistent explanation for these observations has since become a major goal for researchers, one that remains only partially fulfilled to date. A number of thermodynamic scenarios have been formulated that can explain experimental observations (*7*, *9*–*11*), but an unambiguous experimental discrimination among them has not yet been possible. The difficulty of performing measurements under conditions where the liquid is highly metastable with respect to the crystal, hence short-lived, remains a challenge to experimental ingenuity. Much progress has nevertheless been made in the characterization of cold, metastable forms of water in the supercooled (*12*, *13*), stretched (*14*), and glassy states (*15*).

An intriguing hypothesis was formulated in a 1992 computational study, according to which the enhanced fluctuations upon supercooling are indicative of an underlying first-order phase transition between two liquid phases, terminating at a liquid-liquid critical point (LLCP) located at higher-than-atmospheric pressure (*16*). The verification or falsification of the liquid-liquid phase transition (LLPT) hypothesis has been a major goal of experimental studies. Although no definitive proof of the existence of a metastable second critical point in water has been produced to date (*17*), the preponderance of evidence is consistent with this hypothesis (*18*, *19*).

Computational studies have played a central role in this field (*20*), because crystallization, which poses a major challenge in experiments, does not affect simulations. Indeed, in molecular models of water, the characteristic time scales for relaxation and crystal nucleation are well separated (*21*), enabling the accurate thermodynamic characterization of metastable liquid states without the interference of crystallization. These time scales are not as widely separated in atomic models of water (*22*), where avoiding crystallization at deeply supercooled conditions is a challenge.

Rigorous proof of the existence of a LLPT requires performing free energy calculations at subcritical temperatures, a computationally demanding undertaking. Although evidence consistent with a LLPT has been reported for several potentials (*20*, *23*, *24*), to date, rigorous proof of the existence of a LLPT has only been

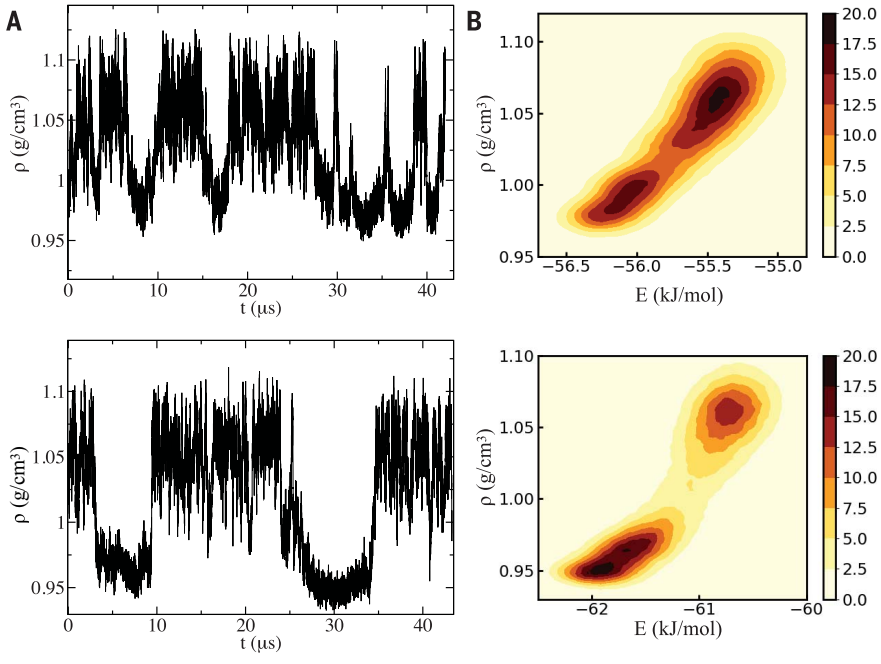


Fig. 1. Critical density and energy fluctuations. (**A**) Density fluctuations during isobaric molecular dynamics simulations in the vicinity of the critical point for N = 300 molecules of (top) TIP4P/2005 at 177 K and 1750 bar and (bottom) TIP4P/Ice at 190 K and 1725 bar. (**B**) Statistics of critical fluctuations displayed as two-dimensional (density and energy) probability densities for TIP4P/2005 (top) and TIP4P/Ice (bottom) at their respective critical points (see also Fig. 2).

[1]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA. [2]Department of Physics, Sapienza University of Rome, Rome, Italy.
*Corresponding author. Email: pdebene@princeton.edu
†These authors contributed equally to this work.

obtained for the ST2 model (*25*–*28*), a five-site model in which tetrahedrality is built into its constitutive geometry.

Here, we uncover critical behavior in two water models for which a LLPT has yet to be rigorously proven. We performed histogram reweighting (*29*) and large-system scattering calculations using two of the most accurate classical models of water, TIP4P/2005 (*30*) and TIP4P/Ice (*31*). Unlike in the ST2 model, the interacting sites in both TIP4P models lie on the same plane. We applied the theory of critical phenomena and statistical mechanical analysis to quantify density and energy fluctuations close to the critical point. We obtained evidence of critical behavior belonging to the three-dimensional (3D) Ising universality class, which demonstrates that, in addition to the regular vapor-liquid critical point, both TIP4P/2005 and TIP4P/Ice exhibit a second, metastable critical point at deeply supercooled conditions.

Figure 1A shows the temporal evolution of the density for TIP4P/2005 (top) and TIP4P/Ice (bottom), obtained from extremely long (>40 μs) isobaric-isothermal molecular dynamics simulations at slightly supercritical conditions (see supplementary text section of the supplementary materials and tables S1 and S2). Pronounced density fluctuations, characteristic of near-critical behavior, are evident. The statistics of these critical fluctuations are displayed in Fig. 1B for TIP4P/2005 (top) and TIP4P/Ice (bottom), in the form of two-dimensional probability densities, reweighted to the respective critical points using the multiple histogram method (Fig. 2 and fig. S1). Bimodal behavior, as well as a strong positive correlation between density and energy fluctuations, is evident. Both models experienced strong fluctuations between low density–low energy and high density–high energy states, and these states also differ in their structural properties (fig. S2). The time scales for relaxation and crystal nucleation remained well separated across the range of conditions probed in this work (fig. S3). A nucleation event, being irreversible, would prevent the density from oscillating between low and high values. The ergodicity of the simulated systems at the conditions explored in this work is documented in figs. S4 and S5.

Figure 2, A and B, shows the statistics of density fluctuations at selected temperatures for pressure values at which the variance of density fluctuations is maximized. The distributions were obtained using the multiple histogram method. The appearance of two peaks, characteristic of critical phenomena, signals the coexistence of regions of different density, correlated over distances comparable to the studied system size.

To ascertain the universality class of the metastable critical point, we compared the

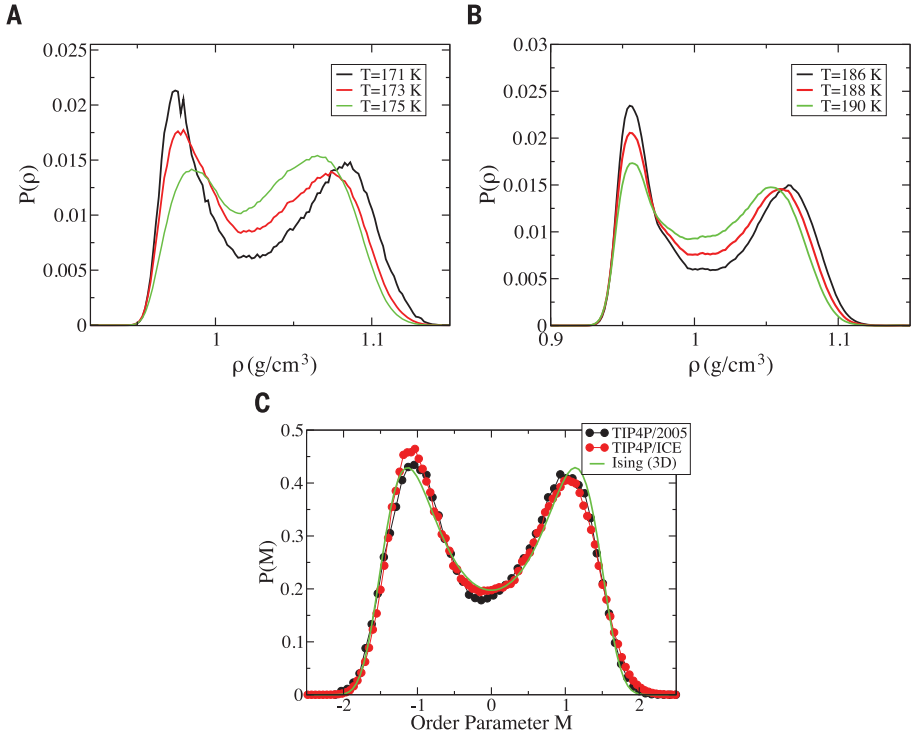


Fig. 2. Universal Ising-like behavior of order parameter fluctuations. Density probability distributions for (**A**) TIP4P/2005 and (**B**) TIP4P/Ice, obtained from density fluctuation data collected at several *T* and *P* and then reweighted to arbitrary state points using the histogram reweighting procedure (see supplementary materials). The pressure at each temperature was chosen so as to maximize the variance of the density distribution. (**C**) Simulation data (symbols) for the order parameter distribution ($M = \rho + sE$) matched to the corresponding 3D Ising universal distribution (green curve). The match yields optimal values for *s*, T_c (172 and 188.6 K for TIP4P/2005 and TIP4P/Ice, respectively), and P_c (1861 and 1725 bar for TIP4P/2005 and TIP4P/Ice, respectively) (table S3). *N* = 300 molecules.

distribution of order parameter fluctuations with the corresponding asymptotic form for the 3D Ising model. The relevant order parameter for single-component fluids is a linear combination of density and configurational energy ($\rho + sE$), where *s* is a field-mixing parameter (*32*). As described in the supplementary materials, we performed a nonlinear fit to regress the critical temperature T_c, critical pressure P_c, and field-mixing parameter such that the reweighted order parameter fluctuations best match the theoretical predictions. The quality of the fits shown in Fig. 2C for a system size of $N = 300$ water molecules (see figs. S6 and S7 for other system sizes) confirmed that the finite-size estimates of the metastable critical point for TIP4P/2005 and TIP4P/Ice are consistent with the 3D Ising universality class. Averaging over the three small-system estimates ($N = 300$, 500, and 1000 water molecules), our best estimate for the critical point is $T_c = 172 \pm 1$ K and $P_c = 1861 \pm 9$ bar for TIP4P/2005 and $T_c = 188 \pm 1$ K and $P_c = 1739 \pm 6$ bar for TIP4P/Ice (table S3).

To explore the anomalous scattering behavior expected in the vicinity of a LLCP (*33*), we computed the oxygen-oxygen static structure factor $S(k)$, focusing on the small–wave vector region (small k). The static structure factor is given by

$$S(k) = N^{-1}\sum_{j=1}^{N}\sum_{l=1}^{N}\left\langle e^{[i\mathbf{k}\cdot(\mathbf{r}_j-\mathbf{r}_l)]}\right\rangle \quad (1)$$

where the angle brackets denote thermal averaging and averaging over wave vectors $\mathbf{k}$ with the same magnitude k. The structure factor is related to the isothermal compressibility by

$$\lim_{k\to 0} S(k) = \rho k_B T K_T \quad (2)$$

In the above equations, N is the number of molecules, $\mathbf{r}_m$ is the position vector of the oxygen of molecule m, ρ is the number density, k_B is the Boltzmann constant, and K_T is the isothermal compressibility.

In the Ornstein-Zernike approximation (*34*), which is valid in the small-k limit, $S(k)$ close to a critical point can be written as

$$S(k) = S_N(k) + \frac{S_A(0)}{\xi^2 k^2 + 1} \quad (3)$$

where S_N is the normal component of the structure factor (not affected by critical fluctuations), S_A quantifies the contribution of critical (anomalous) fluctuations to the compressibility (see

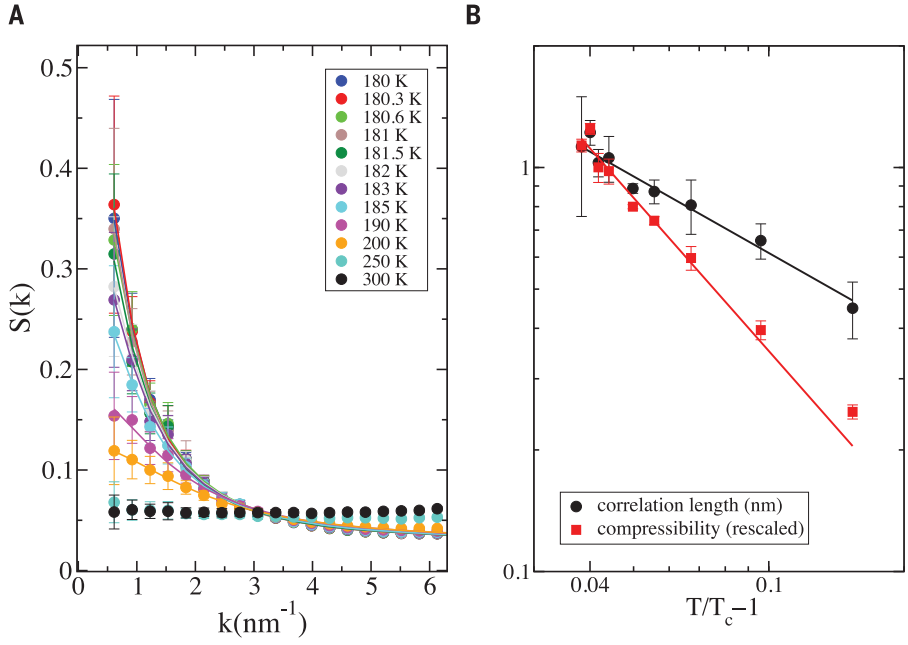


Fig. 3. Anomalous behavior of the structure factor near the critical point. (**A**) Low-*k* region of the structure factor computed for N = 36,424 TIP4P/2005 molecules at different temperatures and ρ = 1.012 g/cm^3, in proximity to the critical isochore (see supplementary materials). (**B**) Calculated values of the isothermal compressibility (Eq. 2) in units of inverse pascal, and scaled by a multiplication factor of 2 × 10^8, and the correlation length (Eq. 3). The slope of the lines through the data correspond to the 3D Ising critical exponents (ν = 0.63 and γ = 1.26; see Eqs. 4 and 5). Critical temperature obtained by a constrained least-squares fit to the data, yielding T_c = 173.3 K. Errors are standard deviations obtained using three equal, nonoverlapping data blocks.

Eq. 2), and ξ is the correlation length, i.e., the characteristic distance over which density fluctuations become uncorrelated. Thus, by computing $S(k)$ according to Eq. 1 over a range of wave vectors and fitting to Eq. 3 (fig. S8), one can obtain numerically, for a given T, the quantities $S_N(k)$, $S_A(0)$, and ξ. Asymptotically close to the critical point, the Lorentzian form in Eq. 3 should be replaced by $S(k) \sim (\xi^2 + k^2)^{\left(\frac{\eta}{2}-1\right)}$, with $\eta \approx 0.04$ (*35*, *36*); at the temperatures sampled in this work (table S4), the proportionality between $1/S(k)$ and k^2 was satisfied (fig. S9), indicating that the Ornstein-Zernike form (Eq. 3) is valid, and η corrections were negligible.

In the vicinity of the critical point, the correlation length and the isothermal compressibility are expected to diverge along the critical isochore according to (*36*)

$$K_T \sim \left(\frac{T}{T_c} - 1\right)^{-\gamma} \quad (4)$$

$$\xi \sim \left(\frac{T}{T_c} - 1\right)^{-\nu} \quad (5)$$

For the 3D Ising universality class, the critical exponents are $\nu \approx 0.63$ and $\gamma = 2\nu \approx 1.26$. Figure 3A shows the computed structure factor as a function of $k = |\mathbf{k}|$ over a range of temperatures, along the critical isochore (see supplementary materials), for the TIP4P/2005 model (see fig. S10 for TIP4P/Ice). A sufficiently large system size was used (N = 36,424 molecules) to ensure that the linear dimension of the cubic computational cell exceeded twice the largest correlation length. The sharp increase in the structure factor at low k values is evident. Because $S(k)$ is the Fourier transform of the pair correlation function (*36*), the low-k behavior shown in Fig. 3A corresponds to an increase in long-range density fluctuations, a distinctive signature of criticality. The regressed values of the correlation length and the critical component of the isothermal compressibility are shown in Fig. 3B, plotted according to Eqs. 4 and 5. The slope of each line has been fixed at its expected 3D Ising universality class value. It can be seen that the Ising slopes provide a very good representation of the data, further confirming that the critical behavior we have identified is consistent with the 3D Ising universality class. We report an approximately threefold increase in the correlation length and a corresponding approximately fivefold increase in the compressibility. An even closer approach to criticality for this large size (from the present $T/T_c - 1 \approx 0.04$ to, say, 0.001) is beyond current computational capabilities. Undertaking such an investigation, when it becomes possible, would provide a computational probing of criticality comparable in proximity to experimental studies of ordinary vapor-liquid criticality (*37*).

Our calculations, probing the boundaries of what is computationally possible in this area at present, provide clear evidence of the presence of a metastable critical point in the deeply supercooled region of the TIP4P/2005 and TIP4P/Ice models of water as well as consistency with the Ising universality class. TIP4P/2005 and TIP4P/Ice are among the most accurate classical models of water, having in particular been demonstrated to describe water's complex crystalline phase diagram with semiquantitative accuracy (*30*, *31*). Previously, rigorous free energy methods have been used to demonstrate the presence of a LLPT in the less accurate ST2 model (*20*, *25*–*28*), subsequently confirmed by anomalous scattering calculations (*33*).

The TIP4P/2005 and TIP4P/Ice models, although among the most realistic classical models of water, are not accurate enough to endow the present calculations with predictive value for the location of the critical point in water. The critical parameters of the two models differ by ~10% in temperature (but substantially less in pressure and density). The significance of this study is the demonstration that a LLCP is not a fortuitous effect in the less accurate ST2 model but is likely a considerably more general feature that can be observed, albeit with substantial numerical effort, in quite accurate classical models. The precise determination of the location of the critical point in these models opens the possibility of tackling several open questions connected with the existence of a LLCP. Specifically, we refer to the interplay between critical fluctuations and nucleation (*38*, *39*), the nucleation of a low-density liquid from within a high-density liquid and vice versa, and the possibility of highlighting specific features associated with the LLPT that may help design experiments capable of discriminating between different thermodynamic scenarios. It is also important to test whether our findings are confirmed when using even more accurate (but computationally far more expensive) potentials (*40*) and including nuclear quantum effects (*41*).

REFERENCES AND NOTES

1. F. Franks, *Water: A Matrix for Life* (RSC Paperbacks, Royal Society of Chemistry, ed. 2, 2000).
2. P. Ball, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13327–13335 (2017).
3. C. G. Salzmann, *J. Chem. Phys.* **150**, 060901 (2019).
4. F. H. Stillinger, *Science* **209**, 451–457 (1980).
5. J. Russo, K. Akahane, H. Tanaka, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3333–E3341 (2018).
6. H. E. Stanley, J. Teixeira, *J. Chem. Phys.* **73**, 3404–3422 (1980).
7. P. G. Debenedetti, *J. Phys. Condens. Matter* **15**, R1669–R1726 (2003).
8. R. J. Speedy, C. A. Angell, *J. Chem. Phys.* **65**, 851–858 (1976).
9. S. Sastry, P. G. Debenedetti, F. Sciortino, H. E. Stanley, *Phys. Rev. E* **53**, 6144–6154 (1996).
10. P. H. Handle, T. Loerting, F. Sciortino, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13336–13344 (2017).
11. N. J. Hestand, J. L. Skinner, *J. Chem. Phys.* **149**, 140901 (2018).
12. J. A. Sellberg *et al.*, *Nature* **510**, 381–384 (2014).

13. K. H. Kim *et al.*, *Science* **358**, 1589–1593 (2017).
14. G. Pallares *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7936–7941 (2014).
15. K. Amann-Winkel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17720–17725 (2013).
16. P. H. Poole, F. Sciortino, U. Essmann, H. E. Stanley, *Nature* **360**, 324–328 (1992).
17. A. K. Soper, *J. Chem. Phys.* **150**, 234503 (2019).
18. O. Mishima, H. E. Stanley, *Nature* **392**, 164–168 (1998).
19. S. Woutersen, B. Ensing, M. Hilbers, Z. Zhao, C. A. Angell, *Science* **359**, 1127–1131 (2018).
20. J. C. Palmer, P. H. Poole, F. Sciortino, P. G. Debenedetti, *Chem. Rev.* **118**, 9129–9151 (2018).
21. J. C. Palmer, R. S. Singh, R. Chen, F. Martelli, P. G. Debenedetti, *Mol. Phys.* **114**, 2580–2585 (2016).
22. E. B. Moore, V. Molinero, *Nature* **479**, 506–508 (2011).
23. Y. Li, J. Li, F. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12209–12212 (2013).
24. Y. Ni, J. L. Skinner, *J. Chem. Phys.* **144**, 214501 (2016).
25. J. C. Palmer *et al.*, *Nature* **510**, 385–388 (2014).
26. F. Sciortino, I. Saika-Voivod, P. H. Poole, *Phys. Chem. Chem. Phys.* **13**, 19759–19764 (2011).
27. T. A. Kesselring, G. Franzese, S. V. Buldyrev, H. J. Herrmann, H. E. Stanley, *Sci. Rep.* **2**, 474 (2012).
28. Y. Liu, J. C. Palmer, A. Z. Panagiotopoulos, P. G. Debenedetti, *J. Chem. Phys.* **137**, 214505 (2012).
29. A. M. Ferrenberg, R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195–1198 (1989).
30. J. L. F. Abascal, C. Vega, *J. Chem. Phys.* **123**, 234505 (2005).
31. J. L. F. Abascal, E. Sanz, R. García Fernández, C. Vega, *J. Chem. Phys.* **122**, 234511 (2005).
32. N. B. Wilding, A. D. Bruce, *J. Phys. Condens. Matter* **4**, 3087–3108 (1992).
33. J. Guo, R. S. Singh, J. C. Palmer, *Mol. Phys.* **116**, 1953–1964 (2018).
34. L. S. Ornstein, F. Zernike, *Proc. Sect. Sci. K Ned. Akad. Wet.* **17**, 793 (1914).
35. M. E. Fisher, *J. Math. Phys.* **5**, 944–962 (1964).
36. H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, 1971).
37. J. V. Sengers, J. M. H. L. Sengers, *Annu. Rev. Phys. Chem.* **37**, 189–222 (1986).
38. K. Binder, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9374–9375 (2014).
39. R. Kurita, H. Tanaka, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24949–24955 (2019).
40. G. A. Cisneros *et al.*, *Chem. Rev.* **116**, 7501–7528 (2016).
41. B. Cheng, E. A. Engel, J. Behler, C. Dellago, M. Ceriotti, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1110–1115 (2019).
42. P. G. Debenedetti, F. Sciortino, G. H. Zerze, Data for "Second critical point in two realistic models of water," Version 1, Zenodo (2020); https://doi.org/10.5281/zenodo.3836542.

ACKNOWLEDGMENTS

The simulations presented in this work were performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University, and on computational resources managed and supported by the Physics Department of Sapienza University Rome. **Funding:** P.G.D. acknowledges support from the National Science Foundation (grant CHE-1856704). F.S. acknowledges support from MIUR-PRIN 2018 (grant 2017Z55KCW). **Author contributions:** P.G.D. and F.S. conceived of the project. P.G.D., F.S., and G.H.Z. designed the calculations, and F.S. and G.H.Z. performed the calculations. All authors analyzed the data. P.G.D. wrote the paper, and all authors edited the paper. **Competing interests:** None declared. **Data and materials availability:** GROMACS codes used to perform the simulations in this work are publicly available (http://manual.gromacs.org/documentation/). The input files for simulations and energy, density, and structure factor data underlying the conclusions of this work are archived in Zenodo (*42*). All other data needed to evaluate the conclusions in this paper are present either in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/289/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S4
References (*43*–*58*)

30 March 2020; accepted 27 May 2020
10.1126/science.abb9796

PIEZOELECTRICS

Giant piezoelectricity in oxide thin films with nanopillar structure

Huajun Liu[1]*†, Haijun Wu[2]†, Khuong Phuong Ong[3], Tiannan Yang[4], Ping Yang[2,5], Pranab Kumar Das[5], Xiao Chi[6], Yang Zhang[2], Caozheng Diao[5], Wai Kong Alaric Wong[5], Eh Piew Chew[5], Yi Fan Chen[1], Chee Kiang Ivan Tan[1], Andrivo Rusydi[6], Mark B. H. Breese[5,6], David J. Singh[7], Long-Qing Chen[4], Stephen J. Pennycook[2], Kui Yao[1]*

High-performance piezoelectric materials are critical components for electromechanical sensors and actuators. For more than 60 years, the main strategy for obtaining large piezoelectric response has been to construct multiphase boundaries, where nanoscale domains with local structural and polar heterogeneity are formed, by tuning complex chemical compositions. We used a different strategy to emulate such local heterogeneity by forming nanopillar regions in perovskite oxide thin films. We obtained a giant effective piezoelectric coefficient $d^*_{33,f}$ of ~1098 picometers per volt with a high Curie temperature of ~450°C. Our lead-free composition of sodium-deficient sodium niobate contains only three elements (Na, Nb, and O). The formation of local heterogeneity with nanopillars in the perovskite structure could be the basis for a general approach to designing and optimizing various functional materials.

The well-established strategy for achieving high piezoelectric performance is to construct multiphase boundaries by tuning complex chemical compositions, such as the morphotropic phase boundary in lead zirconate titanate (PZT) (*1*) or the polymorphic phase boundary in potassium sodium niobate (KNN) (*2*). At these phase boundaries, energy differences among the distinct crystalline phases are adequately small with flattened thermodynamic energy profiles, giving rise to low domain wall energy and miniaturization of ferroelectric domains (*3*, *4*). The local structural and chemical heterogeneity at the nanoscale is recognized to play a key role in achieving ultrahigh piezoelectric coefficients (*5*–*7*). However, the strategy of constructing composition-controlled phase boundaries usually involves complex chemical compositions, and these often have poor temperature stability due to a substantially reduced Curie temperature (T_C) from extensive doping.

Defects, such as dislocations and dopants, are important for developing materials with certain properties, such as the mechanical properties of metals or the electronic properties of semiconductors. An out-of-phase boundary is a type of extended crystallographic defect characterized by a misregistry of a fraction of a unit cell dimension in neighboring regions of a crystal (*8*). When the offset of the misregistry is a half of a unit cell dimension, it is the special case known as the antiphase boundary (*8*, *9*). Although the origins of out-of-phase boundaries and their influences on properties have been well studied in metal alloys (*10*) and semiconductor thin films (*11*), their investigation in complex oxides is rare (*9*) and their link to functional properties remains elusive (*8*). We show that the out-of-phase boundaries between the nanopillar regions and regular perovskite matrix induce local structural and polar heterogeneity and emergent giant piezoelectric response in lead-free, Na-deficient $NaNbO_3$ thin films. We used atomic-scale images to reveal the formation of vertical nanopillars with size of a few to tens of nanometers. The structural distortions around nanopillar regions lower the crystal symmetry from tetragonal to monoclinic structure, facilitating polarization rotation and domain wall motion under electric fields to substantially enhance piezoelectric performance as supported by theoretical calculations. We discovered a giant effective piezoelectric coefficient $d^*_{33,f}$ of ~1098 pm/V at 1 kHz with an applied electric field of 125 kV/cm, which is more than twice that of the best PZT films (*12*) and four times that of the best KNN-based films with extensive doping (*13*). We maintained a simple composition of three elements (Na, Nb, and O) and a high T_C of ~450°C.

We grew epitaxial thin films of stoichiometric $NaNbO_3$ (NNO) with regular perovskite structure and Na-deficient NNO with nanopillar regions (NPR-NNO) on Nb-doped $SrTiO_3$ (001) substrates by sputter deposition (*14*). Our stoichiometric NNO film was homogeneous with a thickness of ~200 nm, which we determined using a low-magnification annular bright field (ABF) image taken by scanning transmission electron microscopy (STEM) (Fig. 1A). The Nb atoms are arranged in a typical perovskite lattice (Fig. 1, B and C). In contrast, we found vertically aligned nanopillars with a smooth surface in the cross-sectional view of NPR-NNO film (Fig. 1D). Those nanopillars are embedded in the matrix of regular perovskite structure. Inside the nanopillar regions, the Nb atoms occupy the original Na positions along both in-plane directions (Fig. 1H), and the Na atoms also occupy the original Nb positions along both in-plane directions (fig. S2B). In the cross-sectional view (Fig. 1E), we observed many extended columnar structures. By changing the microscope focus, we confirm that these columnar structures are nanoscale in in-plane dimensions (fig. S1), fully consistent with the nanoscale size of nanopillar regions ranging from a few to tens of nanometers in the plan-view image (Fig. 1G). Because the thickness of the sample greatly exceeds the thickness of the nanopillars, we actually see overlapping of the nanopillar regions (marked by arrows in Fig. 1E) and the matrix. The lattice in the nanopillar regions (yellow box in Fig. 1, E and H) is shifted slightly by δ_C along the out-of-plane direction and by approximately half of a unit cell along both in-plane directions from the lattice of the matrix (red box in Fig. 1, E and H). Hence, the misregistry between the nanopillar regions and matrix can be seen as normal perovskite phase shifted by $\langle \mathbf{a}/2, \mathbf{b}/2, \boldsymbol{\delta}_C \rangle$, where $\mathbf{a}$ and $\mathbf{b}$ are unit vectors along in-plane directions and $\boldsymbol{\delta}_C$ is along the out-of-plane direction with magnitude of a small fraction of the unit cell length. This boundary is an out-of-phase boundary, which is different from an antiphase boundary where the displacement vector would be $\langle \mathbf{a}/2, \mathbf{b}/2, \mathbf{c}/2 \rangle$.

On the basis of the above structure details, we constructed a schematic structural model (Fig. 1, F and I). At the out-of-phase boundary between the nanopillar regions and the matrix, the oxygen octahedra are edge-sharing instead of the corner-sharing octahedra normally found in perovskites. Analogous to the Ruddlesden-Popper stacking faults in perovskite structures of ABO_3 with A-site excess (*15*), these out-of-phase boundaries can accommodate more Nb atoms than the stoichiometric perovskite (fig. S2). The Na deficiency (fig. S3) in the NPR-NNO film drives the formation of out-of-phase boundaries with atoms sitting at in-plane antisite positions (Fig. 1, H and I). The interface atomic structure between the NPR-NNO film and the substrate (fig. S2G) shows that the first layer of film appears to be a Nb-occupied NaO layer, which largely determines the growth pattern of the nanopillar structures in the following layers. According to first-principles calculations (*14*), when Na is deficient and Nb is rich, the negative formation energy of antisite Nb_{Na} shows that it is energetically favorable for excess Nb atoms to occupy Na vacancies.

[1]Institute of Materials Research and Engineering, A*STAR (Agency for Science, Technology and Research), 138634 Singapore. [2]Department of Materials Science and Engineering, National University of Singapore, 117575 Singapore. [3]Institute of High Performance Computing, A*STAR (Agency for Science, Technology and Research), 138632 Singapore. [4]Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA 16802, USA. [5]Singapore Synchrotron Light Source (SSLS), National University of Singapore, 117603 Singapore. [6]Department of Physics, National University of Singapore, 117551 Singapore. [7]Department of Physics and Astronomy, University of Missouri, Columbia, MO 65211, USA.
*Corresponding author. Email: huajunliu87@gmail.com (H.L.); k-yao@imre.a-star.edu.sg (K.Y.)
†These authors contributed equally to this work.

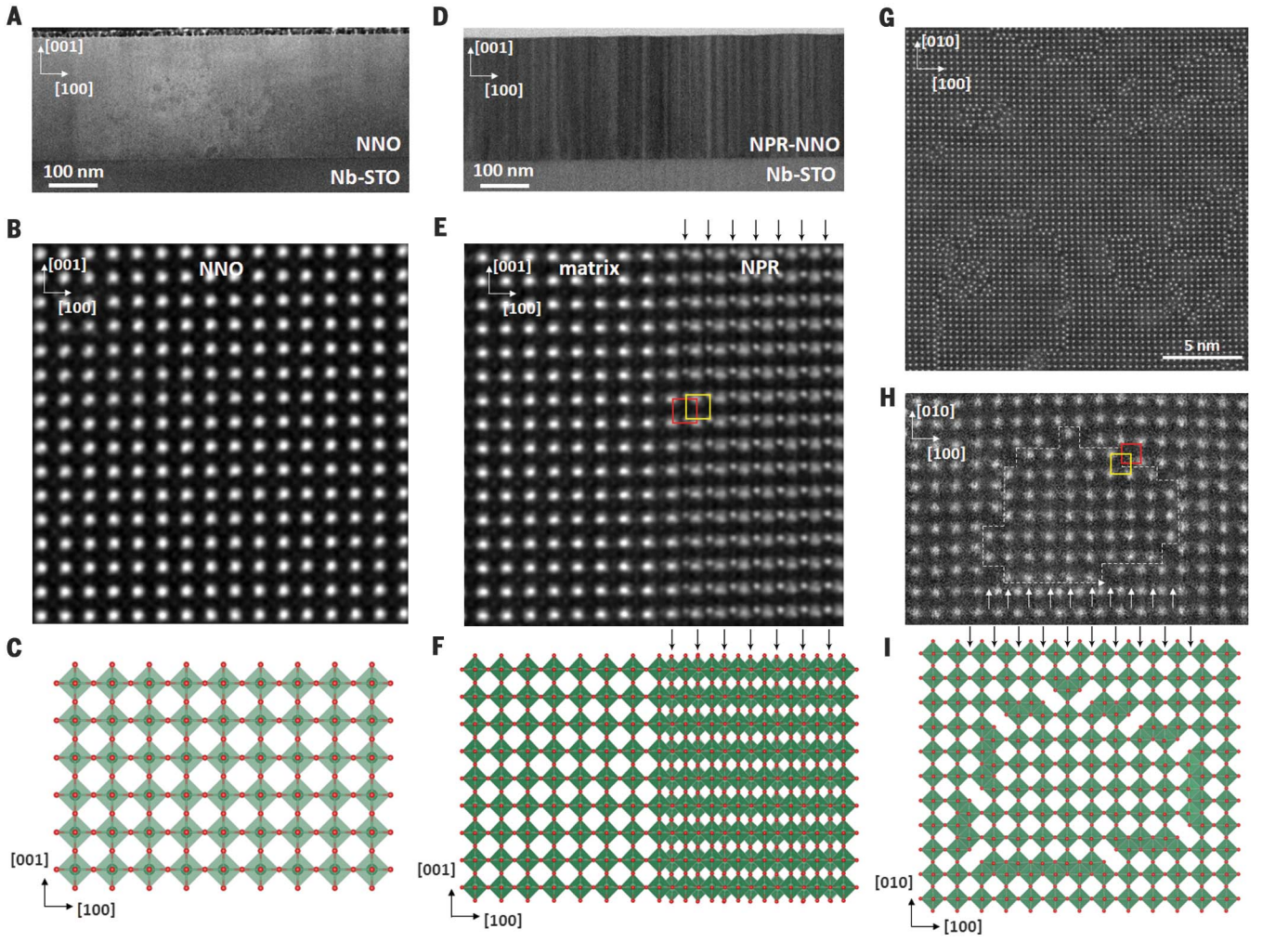


Fig. 1. Atomic structures of NNO and NPR-NNO thin films. (**A** to **C**) Cross-sectional low-magnification STEM ABF image (A), atomically resolved HAADF image (B), and schematic structural model (C) of NNO films on Nb-STO (001) substrates. (**D**, **E**, **G**, and **H**) Cross-sectional low-magnification ABF image (D), cross-sectional atomically resolved HAADF image (E), plan-view high-magnification image (G), and plan-view atomically resolved HAADF image (H) of NPR-NNO films on Nb-STO (001) substrates. (**F** and **I**) Schematic structural models of NPR-NNO films in cross-sectional view (F) and plan view (I). The arrows in (E), (F), (H), and (I) indicate the rows of Nb atoms in the nanopillar regions. They are much weaker than the Nb atom columns in the normal crystal from cross-sectional view [left side of (E)] because they are much shorter along the beam direction. The yellow and red boxes in (E) and (H) represent the lattice of the nanopillar region and the regular perovskite structure of the matrix, respectively. The dashed lines in (H) show the nanopillar region.

We used synchrotron x-ray diffraction to investigate the crystal structures and temperature-dependent phase transitions of the NNO and NPR-NNO films. The thickness fringes around the main diffraction (001) peak of the NPR-NNO film (Fig. 2A) reflect the high quality and smooth surface, whereas we observed no thickness fringes for NNO films. We found a single diffraction peak with the same in-plane lattice parameter as the Nb-STO substrate in the (103) reciprocal space mapping (RSM) for the NNO film (Fig. 2B). This indicates that the NNO film is coherently strained even at a thickness of 200 nm, and the observation is consistent with a previous report (*16*). This is probably due to the small lattice mismatch of ~0.47% between the NNO bulk structure and the Nb-STO substrate (fig. S4). Combined with other RSMs, we found that the NNO film has a tetragonal symmetry. However, the (103) RSM for the NPR-NNO film (Fig. 2C) shows peak splitting due to monoclinic symmetry (see fig. S4 for more RSMs and lattice parameters). Oxygen octahedral rotations in the perovskite structure give rise to doubling of the unit cell along certain directions, which can be probed by the half-order diffraction peaks (*17*) (Fig. 2D). The half-order (0 0.5 1.5) diffraction peak is equivalent to the (0.5 0 1.5) peak (fig. S4B) because of the isotropic nature of the in-plane directions of the Nb-STO substrate, thus suggesting that the in-phase rotation exists along both in-plane directions. According to Glazer's notation (*18*), the rotation patterns for both films are a mixture of $a^-b^+c^-$ and $a^+b^-c^-$. However, the NPR-NNO film shows substantially lower peak intensity, especially the (1.5 0.5 1.5) and (0.5 0.5 1.5) peaks corresponding to out-of-phase rotations. This is probably due to the distortion induced by the nanopillar regions and the edge-sharing octahedra at the out-of-phase boundary (Fig. 1, G to I) that breaks the long-range order of rotation.

We conducted temperature-dependent x-ray diffraction to study the phase transition (Fig. 2E). The out-of-plane lattice parameter c shows a change in slope at ~457°C for the NNO and NPR-NNO films. The changes in the in-plane lattice parameters a of both films are less obvious as a result of the constraint by the substrates. We also monitored the intensity of the half-order (0 0.5 1.5) diffraction peak (Fig. 2F). For both NNO and NPR-NNO films,

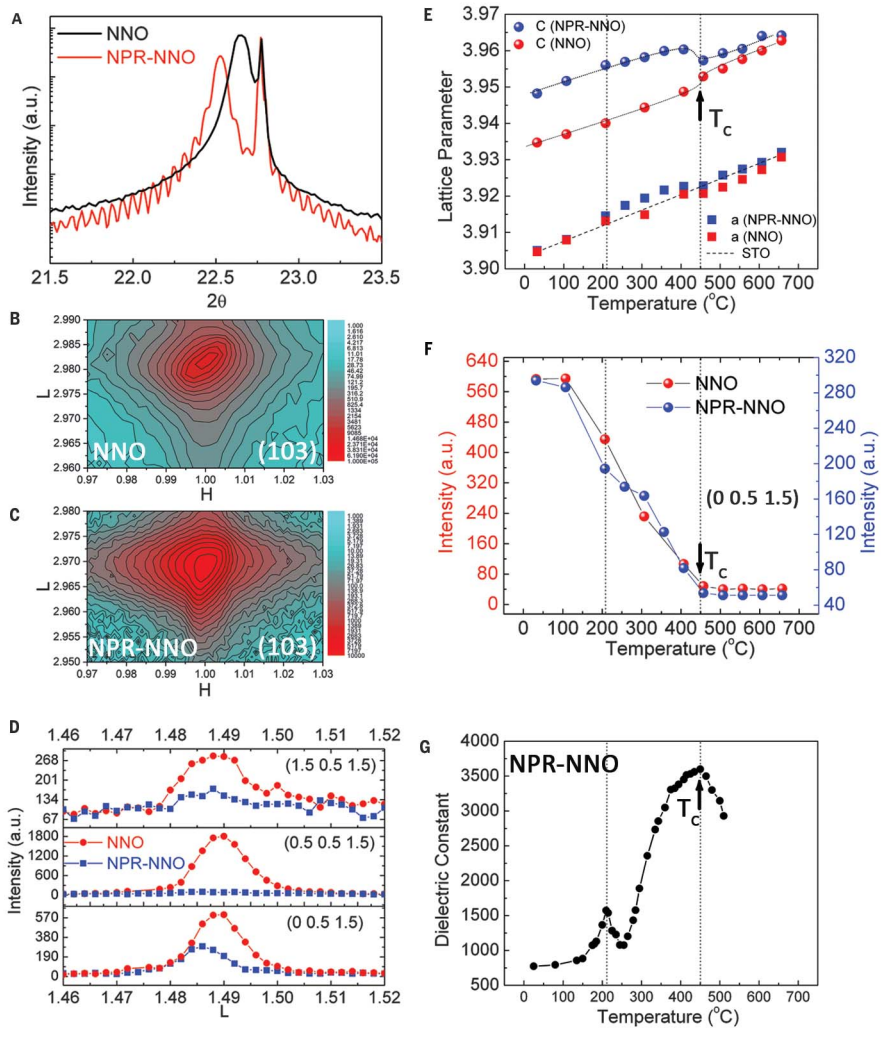


Fig. 2. Crystal structure and phase transition of NNO and NPR-NNO thin films. (**A**) X-ray diffraction patterns around the (001) peak for NNO and NPR-NNO films. (**B** and **C**) Reciprocal space mapping around the (103) peak for the NNO film (B) and the NPR-NNO film (C), plotted in reciprocal lattice units of the Nb-STO substrate. (**D**) Half-order diffraction peaks for the NNO and NPR-NNO films. (**E**) Lattice parameters of the NNO and NPR-NNO films as a function of temperature. (**F**) Intensity of the half-order (0 0.5 1.5) diffraction peaks as a function of temperature for the NNO and NPR-NNO films. (**G**) Dielectric constant measured at 1 kHz as a function of temperature for the NPR-NNO film.

the diffraction peaks disappear and the intensities decrease to the background level at ~457°C. This observation suggests a phase transition from the a^+ and b^+ tilt pattern to nontilted a^0 and b^0 of a higher symmetry phase, with the likely space group being *P4/mbm* above 457°C (*19*). Measuring the temperature dependence of the dielectric constant for the NPR-NNO film (Fig. 2G) shows a peak at ~450°C, confirming the high T_C. The smaller peak at ~210°C could be related to a ferroelectric-to-ferroelectric phase transition, as we observed no obvious structural change (Fig. 2, E and F). The hysteresis loops we measured demonstrate the ferroelectric nature of both films (fig. S5, A to C).

To investigate the macroscopic piezoelectric properties of the films, we used a laser scanning vibrometer (*20*–*22*) with an ac unipolar voltage of 2 V in magnitude at 1 kHz for the NNO film (Fig. 3A). We determined the effective piezoelectric coefficient $d^*_{33,f}$ to be ~22 pm/V based on the magnitude of the surface displacement. The electric field dependence of the piezoelectric coefficient and strain for the NNO film exhibits the linear dependence of strain on the electric field, as found in typical piezoelectric materials (Fig. 3B). We found no dependence of the piezoelectric coefficient on the driving frequency of the electric field for the NNO film (Fig. 3C), which agrees with the impedance spectroscopy results (fig. S5D). We measured the film displacement of the NPR-NNO film under the same conditions as the NNO film (Fig. 3D). We determined the effective piezoelectric coefficient $d^*_{33,f}$ to be ~1098 pm/V at 1 kHz with an applied electric field of 125 kV/cm (Fig. 4, E and F). The NPR-NNO film has nonlinear behavior and much-enhanced strain at a higher electric field (Fig. 3E). Decreasing the electric field leads to a very small strain hysteresis (Fig. 3E). This behavior departs from the strong hysteretic behavior of the point defect–induced large strain in aged $BaTiO_3$ single crystal (*23*). The effective piezoelectric strain of the NPR-NNO film increases nonlinearly with the electric field, saturating at ~1.4% above 125 kV/cm. We observed a strong dependence of the effective piezoelectric coefficient on the driving frequency of the electric field for the NPR-NNO film (Fig. 3F). At a low frequency of 100 Hz, the effective piezoelectric coefficient reaches a giant value of ~8114 pm/V, whereas it drops to ~112 pm/V at a high frequency of 10 kHz. The strain shows a similar strong dependence on frequency (see fig. S6). We believe that the strong amplitude and electric field frequency dependence of the piezoelectric strain of the NPR-NNO film indicates a substantial contribution from the motion of out-of-phase boundaries between nanopillar regions and the matrix perovskite (*24*). We observed a strong frequency dependence of dielectric constant and loss in NPR-NNO films in the impedance spectroscopy (fig. S5). X-ray absorption spectra (fig. S7) suggest a slightly occupied Nb 4d level in the NPR-NNO film as compared to the empty Nb 4d level in the NNO film, consistent with higher leakage in the NPR-NNO film (fig. S5F), which could come from the vertical boundaries of the nanopillar regions.

We performed first-principles calculations and phase-field simulations to better understand the mechanism of the giant piezoelectric response (*14*). We calculated the formation energy of the antisite Nb_{Na} defect as a function of chemical potential of Na under both Nb-deficient and Nb-rich conditions (Fig. 4A). The negative formation energy under Nb-rich and Na-deficient conditions shows that excess Nb atoms would occupy Na vacant positions to form antisite Nb_{Na} defects when NNO is Na-deficient. However, as a result of large differences between the ionic radii and valences of Na and Nb ions, antisite Nb atoms are not stable at the original antisite position but show substantial off-centering along the out-of-plane *c* direction. This behavior is similar to Sr-deficient $SrTiO_3$ thin films (*25*). First-principles calculations predict the energies of antisite Nb atoms at different positions (Fig. 4B), showing that large off-centering displacement is energetically favorable. The off-centering displacement induces a large ferroelectric polarization of ~16.2 μC/cm^2 along the out-of-plane *c* direction in the structure with antisite defects. This polarization is greatly enhanced from the polarization of ~1.2 μC/cm^2 in the regular

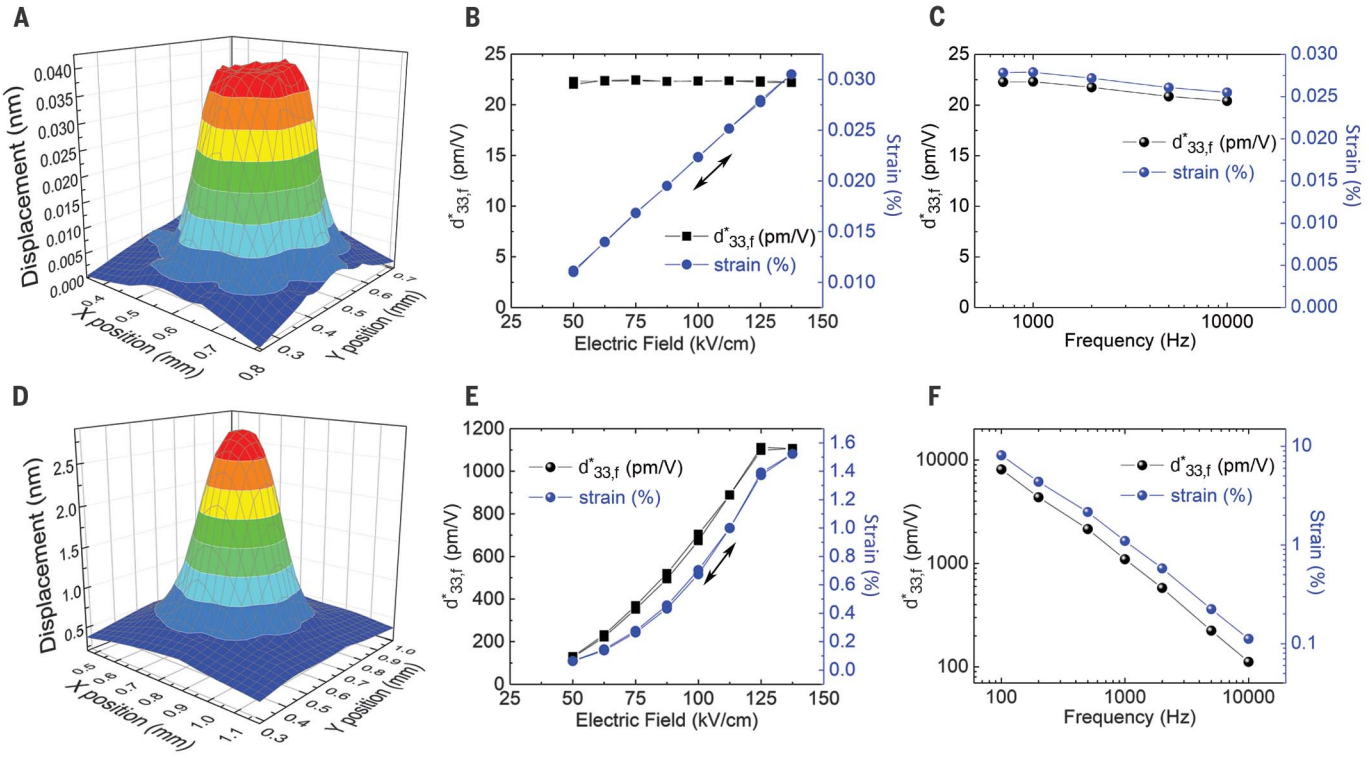


Fig. 3. Piezoelectric properties of NNO and NPR-NNO thin films. (**A** and **D**) 3D drawing of surface displacement of the NNO film (A) under ac unipolar voltage of 2 V in magnitude at 1 kHz and the NPR-NNO film (D) under ac unipolar voltage of 2.5 V in magnitude at 1 kHz. (**B** and **E**) Electric field dependence of effective piezoelectric coefficient and strain of the NNO film (B) and the NPR-NNO film (E) at frequency of 1 kHz. (**C** and **F**) Frequency dependence of effective piezoelectric coefficient and strain of the NNO film (C) and the NPR-NNO film (F).

perovskite structure of NNO, according to first-principles calculations. This large local polarization distorts the surrounding lattice to lower symmetry and gives rise to local polar heterogeneity at the nanoscale. However, the d_{33} we calculated by first principles from the piezoelectric stress tensor and elastic stiffness (*14*) is ~17 pm/V for the structural model with antisite defects and ~14 pm/V for the structural model of normal perovskite. The similar values indicate that the giant strain we observed in the experiments likely originates at a scale beyond the unit cells of the lattice.

We used phase-field simulations to study the contributions by nanopillar regions and their boundaries to the piezoelectric properties (*14*). We simulated the ferroelectric domain structure of the NPR-NNO film (fig. S9, C and D). The local stress developed around the boundaries of nanopillar regions (fig. S9E) causes the polarization directions to be rotated away from the out-of-plane direction, thus forming a multidomain structure with coexisting tetragonal *c* domains and nanoscale monoclinic domains with rich and mixed orientations of polarizations. We simulated the piezoelectric strain response to a static electric field applied along the out-of-plane direction (Fig. 4C). The tetragonal domains (e.g., region III in Fig. 4C) show a weak piezoelectric strain response, whereas the monoclinic domains (e.g., region I) show a much stronger response. The strongest piezoelectric strain response appears around the domain walls between tetragonal and monoclinic domains (e.g., region II). The motions of these domain walls under the electric field serve as substantial contributions to the piezoelectric response, giving rise to a factor of ~15 enhancement of the piezoelectric coefficient d_{33} ~ 290 pm/V for NPR-NNO film as compared with ~20 pm/V for pure NNO film from phase-field simulations. We simulated the frequency dependence of piezoelectric strain by applying an ac electric field at different frequencies (Fig. 4D). The domain wall motion is weakened upon increasing the frequency for the NPR-NNO film (fig. S9, F to H). The piezoelectric response of both the domain wall regions and the monoclinic domains decreases, while the piezoelectric response of the tetragonal domains is almost unchanged. For the NNO film with a single tetragonal domain, the piezoelectric strain is almost independent of frequency. Our simulation results suggest that the giant piezoelectric response is likely to originate from domain wall motion under electric fields. The substantially reduced domain wall response at high frequencies results in the strong frequency dependence of the piezoelectric property we observed.

Because the piezoelectric properties of ferroelectric materials are lost when the temperature exceeds T_C, a high T_C is preferred for applications. We compared the piezoelectric coefficient d_{33} of several major lead-free piezoelectric materials systems together with PZT ceramics as a function of T_C (Fig. 4E and fig. S10A). A common strategy to enhance the piezoelectric coefficient in bulk ceramics is to construct phase boundaries by chemically tuning the composition. However, the complex chemical composition at the phase boundaries often reduces T_C (Fig. 4E). Our design strategy of constructing nanoscale structural and polar heterogeneity intrinsic to the nanopillar structure in NPR-NNO thin films has the advantage of improving the piezoelectric coefficient while simultaneously maintaining a high T_C (Fig. 4E). An additional advantage of the nanopillar approach is the simple chemical composition relative to other thin films (Fig. 4F and fig. S10B). Complex chemical compositions make phase boundary construction challenging because the dopant levels are only a few percent and often involve highly volatile elements (Na, K, Pb, and Bi). This presents an impediment to obtaining reliable and reproducible synthesis of bulk ceramics. Thin films are even more problematic to reliably synthesize. The NPR-NNO thin film contains only three elements, so our design strategy shows great potential for improving composition control to obtain high-performance piezoelectric thin films.

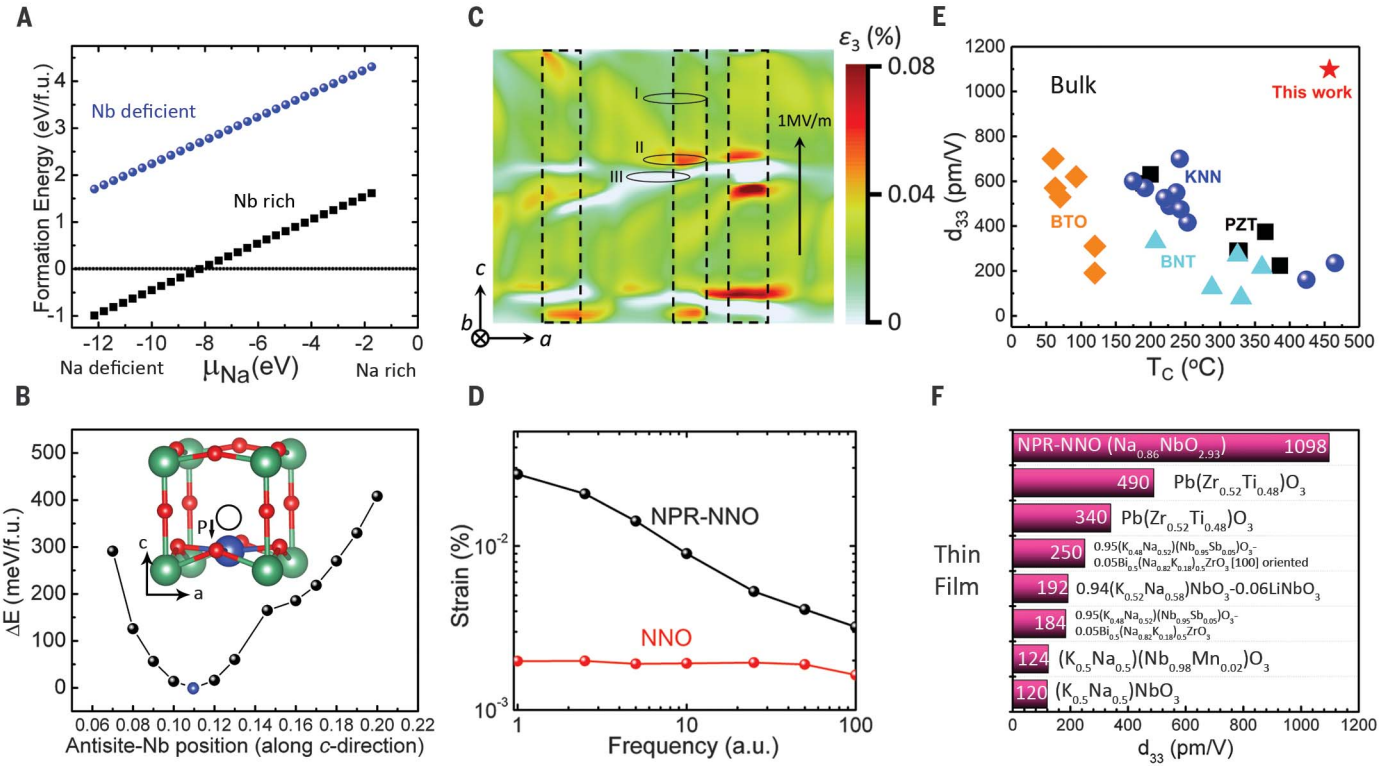


Fig. 4. Theoretical calculation and overview of piezoelectric properties. (**A**) Antisite formation energy as a function of the chemical potential of Na under Nb-rich and Nb-deficient conditions. (**B**) Energy curve of the antisite Nb atom at different positions along the out-of-plane *c* direction. The structural model in the inset illustrates the off-centering displacement of the antisite Nb atom (the blue sphere) in the lowest-energy position. Green spheres are regular Nb atoms; red spheres are oxygen atoms. (**C**) Mapping of the piezoelectric strain response of the NPR-NNO film under an applied electric field of 1 MV/m from phase-field simulation. Region I corresponds to the monoclinic domains, region II the domain wall regions, and region III the tetragonal domains. The dashed rectangles show the nanopillar regions. (**D**) Average piezoelectric strain response of the NPR-NNO and NNO films as a function of field frequency from phase-field simulation. (**E**) Comparison of piezoelectric coefficient d_{33} of bulk material systems with this work (thin film) as a function of Curie temperature T_C. PZT, $Pb(Zr_{0.52}Ti_{0.48})O_3$-based piezoelectric materials; BTO, $BaTiO_3$-based piezoelectric materials; BNT, $(Bi_{0.5}Na_{0.5})TiO_3$-based piezoelectric materials; KNN, $(K_{0.5}Na_{0.5})NbO_3$-based piezoelectric materials. (**F**) Piezoelectric coefficient of the NPR-NNO thin film with simple chemical composition compared to other piezoelectric thin films with much more complicated chemical compositions.

Furthermore, because the out-of-phase boundary can be coupled to magnetic (*26*) and superconducting properties (*27*), the formation of local heterogeneity through such nanopillar structures in perovskite oxides (ABO_3) with A-site deficiency could enable other emergent phenomena in functional materials.

REFERENCES AND NOTES

1. B. Jaffe, W. R. Cook, H. Jaffe, *Piezoelectric Ceramics* (Academic Press, 1971).
2. J. Wu, D. Xiao, J. Zhu, *Chem. Rev.* **115**, 2559–2595 (2015).
3. W. Liu, X. Ren, *Phys. Rev. Lett.* **103**, 257602 (2009).
4. J. Gao *et al.*, *Appl. Phys. Lett.* **99**, 092901 (2011).
5. F. Li *et al.*, *Nat. Mater.* **17**, 349–354 (2018).
6. Q. Liu *et al.*, *Energy Environ. Sci.* **11**, 3531–3539 (2018).
7. F. Li, S. Zhang, D. Damjanovic, L. Chen, T. R. Shrout, *Adv. Funct. Mater.* **28**, 1801504 (2018).
8. M. A. Zurbuchen *et al.*, *J. Mater. Res.* **22**, 1439–1471 (2007).
9. Z. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9485–9490 (2018).
10. M. J. Marcinkowski, N. Brown, R. M. Fisher, *Acta Metall.* **9**, 129–137 (1961).
11. D. B. Holt, *J. Phys. Chem. Solids* **30**, 1297–1308 (1969).
12. M. D. Nguyen, E. P. Houwman, G. Rijnders, *Sci. Rep.* **7**, 12915 (2017).
13. Y. Wang *et al.*, *Adv. Electron. Mater.* **5**, 1800691 (2019).
14. See supplementary materials.
15. R. J. D. Tilley, *J. Solid State Chem.* **21**, 293–301 (1977).
16. T. Saito, T. Wada, H. Adachi, I. Kanno, *Jpn. J. Appl. Phys.* **43**, 6627–6631 (2004).
17. A. M. Glazer, *Acta Crystallogr. A* **31**, 756–762 (1975).
18. A. M. Glazer, *Acta Crystallogr. B* **28**, 3384–3392 (1972).
19. S. K. Mishra, N. Choudhury, S. L. Chaplot, P. S. R. Krishna, R. Mittal, *Phys. Rev. B* **76**, 024110 (2007).
20. K. Yao, F. E. H. Tay, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **50**, 113–116 (2003).
21. L. Wang, K. Yao, W. Ren, *Appl. Phys. Lett.* **93**, 092903 (2008).
22. P. C. Goh, K. Yao, Z. Chen, *Appl. Phys. Lett.* **97**, 102901 (2010).
23. X. Ren, *Nat. Mater.* **3**, 91–94 (2004).
24. N. Bassiri-Gharb *et al.*, *J. Electroceram.* **19**, 47–65 (2007).
25. D. Lee *et al.*, *Science* **349**, 1314–1317 (2015).
26. T. Hibma *et al.*, *J. Appl. Phys.* **85**, 5291–5293 (1999).
27. J. Orenstein, A. J. Millis, *Science* **288**, 468–475 (2000).

ACKNOWLEDGMENTS

H.L. acknowledges helpful discussions with D. D. Fong and H. Zhou. **Funding:** Supported by the Institute of Materials Research and Engineering, A*STAR, RIE2020 AME Core Funds–SERC Strategic Funds: Seeding R&D Activities for Competitive Grants (grant A1718 g0056), National Research Foundation Competitive Research Programme grant NRF-CRP15-2015-04, and NSF grant DMR-1744213. Also supported by a Lee Kuan Yew Postdoctoral Fellowship through Singapore Ministry of Education Tier 1 grant R-284-000-212-114 (H.W.); Singapore Ministry of Education Tier 2 grant MOE2017-T2-1-129 (S.J.P. and H.W.); the Institute of High Performance Computing, Agency for Science, Technology And Research (IHPC, A*STAR) (K.P.O.); Singapore Ministry of Education AcRF Tier-2 grant MOE2019-T2-1-163 and Singapore National Research Foundation under its Competitive Research Funding R-398-000-087-281 (A.R.), and the Singapore Synchrotron Light Source (SSLS) via NUS Core Support Grant C-380-003-003-001, which is a National Research Infrastructure under the National Research Foundation, Singapore. Work at the University of Missouri (D.J.S.) is supported by the U.S. Department of Energy, Basic Energy Sciences, award DE-SC0019114. **Author contributions:** H.L. conceived the main idea, grew the films, and carried out the electric testing; H.L. and C.K.I.T. fabricated the ceramic targets; Y.F.C. deposited the electrodes; H.W. and Y.Z. performed aberration-corrected STEM and analyzed the data; H.L., P.Y., W.K.A.W., and E.P.C. performed x-ray diffraction and analyzed the data; P.K.D., C.D., and X.C. carried out x-ray spectroscopy and analyzed the data; K.P.O. performed DFT calculations and built theoretical models with input from H.L. and in discussion with D.J.S.; results of first-principles calculations were analyzed by K.P.O. and D.J.S.; T.Y. and L.-Q.C. performed and analyzed the phase-field simulations; K.Y. planned and guided the work, and analyzed the electric testing results; H.L., H.W., T.Y., D.J.S., S.J.P., A.R., M.B.H.B., and K.Y. wrote and modified the manuscript with input from others. All authors discussed the results and revised the manuscript. **Competing interests:** The authors have filed provisional Singapore Patent 10201908039P. **Data and materials availability:** All data are available in the main text and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/292/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*28*–*40*)

15 February 2020; accepted 19 May 2020
10.1126/science.abb3209

CORONAVIRUS

Introductions and early spread of SARS-CoV-2 in the New York City area

Ana S. Gonzalez-Reiche[1]*, Matthew M. Hernandez[2,3]*, Mitchell J. Sullivan[1], Brianne Ciferri[1], Hala Alshammary[2], Ajay Obla[1], Shelcie Fabre[4], Giulio Kleiner[2], Jose Polanco[1,2], Zenab Khan[1], Bremy Alburquerque[1,3], Adriana van de Guchte[1], Jayeeta Dutta[1], Nancy Francoeur[1], Betsaida Salom Melo[1,5], Irina Oussenko[1,5], Gintaras Deikus[1,5], Juan Soto[1,5], Shwetha Hara Sridhar[1,5], Ying-Chih Wang[1,5], Kathryn Twyman[6], Andrew Kasarskis[1,5,6], Deena R. Altman[1,7], Melissa Smith[1,5], Robert Sebra[1,5,8,9], Judith Aberg[7], Florian Krammer[2,10], Adolfo García-Sastre[2,7,10,11], Marta Luksza[1,12], Gopi Patel[5], Alberto Paniz-Mondolfi[3], Melissa Gitman[3], Emilia Mia Sordillo[3]†‡, Viviana Simon[2,5,6]†‡, Harm van Bakel[1,8]†‡

New York City (NYC) has emerged as one of the epicenters of the current severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. To identify the early transmission events underlying the rapid spread of the virus in the NYC metropolitan area, we sequenced the virus that causes coronavirus disease 2019 (COVID-19) in patients seeking care at the Mount Sinai Health System. Phylogenetic analysis of 84 distinct SARS-CoV-2 genomes indicates multiple, independent, but isolated introductions mainly from Europe and other parts of the United States. Moreover, we found evidence for community transmission of SARS-CoV-2 as suggested by clusters of related viruses found in patients living in different neighborhoods of the city.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; previously known as 2019-nCoV) is an emerging viral pathogen that was first reported to cause severe respiratory infections in Wuhan, China, in late December 2019. Over the past months, it rapidly spread across the globe, and the World Health Organization (WHO) declared a pandemic on 11 March 2020. Targeted screening of suspected coronavirus disease 2019 (COVID-19) cases, as well as a series of successive nationwide travel restrictions, were put in place to curtail SARS-CoV-2 introductions into the continental United States from outbreak hotspots in China (2 February 2020), Iran (2 March 2020), mainland European countries (13 March 2020), and the United Kingdom and Ireland (16 March 2020) (*1*–*4*). Despite these measures, the first COVID-19 case in New York State was identified in New York City (NYC) on 29 February 2020. During the first weeks of March, the number of detected cases rapidly increased because of expansion of screening capacity after implementation of automated platforms for detection of SARS-CoV-2 infections by local health system diagnostic laboratories and reference laboratories. As of 2 May 2020, there were 312,977 confirmed COVID-19 cases in New York state, including 172,354 (55%) in NYC [New York State Department of Health (NYS DOH); https://coronavirus.health.ny.gov]. With more than 13,300 fatalities in the metropolitan area, NYC has been one of the major epicenters of SARS-CoV-2 infections in the United States.

The Pathogen Surveillance Program (PSP) at the Icahn School of Medicine at Mount Sinai is a multidisciplinary, institutional infrastructure that seeks to generate high-resolution, near real-time genetic information on pathogens found to cause disease in the large and diverse patient population seeking care at the Mount Sinai Health System in NYC. After biospecimen coding, nucleic acid extraction, and polymerase chain reaction quantification, next-generation sequencing approaches based on Illumina and Pacific Biosciences technology provide information on the pathogen's genome. The process has been optimized for quick turnaround, optimized data assembly, and integration with deidentified demographic information.

We took advantage of the existing PSP infrastructure to investigate the origins of SARS-CoV-2 strains circulating in NYC and to dissect the spread of the virus in this metropolitan area with a high-density population. Here, we present the genomic diversity of 90 SARS-CoV-2 isolates obtained from 84 patients seeking care at the Mount Sinai Health System between 29 February 2020 and 18 March 2020. These genomes provide clear evidence for multiple, independent SARS-CoV-2 introductions into NYC during the first weeks of March 2020.

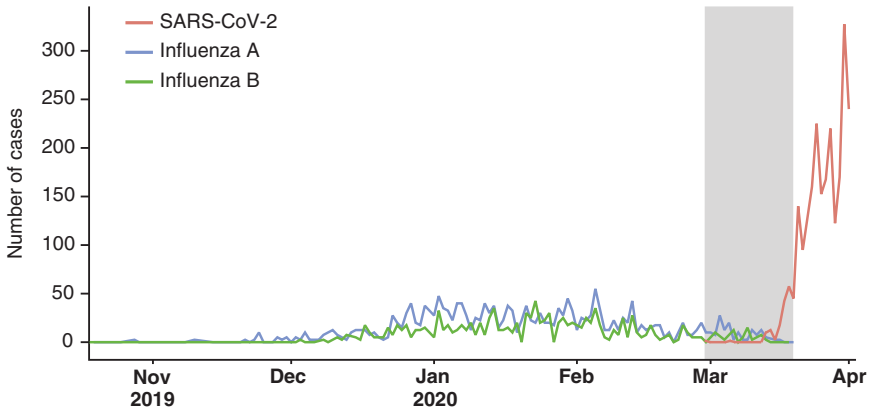


Fig. 1. SARS-CoV-2 confirmed cases. Number of patients with positive molecular tests for SARS-CoV-2 up to 31 March 2020, compared with the daily number of patients with influenza A and B tests in the 2019–2020 season. The shaded area indicates the period in which the samples sequenced in this study were obtained (29 February to 18 March). The number of positive tests per virus was not normalized for the number of tested samples.

[1]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [2]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [3]The Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [4]Clinical Microbiology Laboratory, Department of Pathology, Molecular, and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [5]Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [6]The Mount Sinai Data Office, Mount Sinai Health System, New York, NY 10029, USA. [7]Division of Infectious Diseases, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [8]Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [9]Sema4, Stamford, CT 06902, USA. [10]The Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [11]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [12]Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: harm.vanbakel@mssm.edu (H.v.B.); viviana.simon@mssm.edu (V.S.); emilia.sordillo@mssm.edu (E.M.S.) ‡These authors contributed equally to this work.

On the basis of genetic similarity and phylogenetic analysis of full-length viral genome sequences, most cases diagnosed during the 18 days after the first-reported COVID-19 case in New York state appear to be associated with untracked transmission and potential travel-related exposures. The majority of introductions appear to have been sourced from Europe and the United States. We also identified two clusters that total 21 closely related cases, suggesting community spread. These observations are also supported by the citywide distribution of these cases, which mapped to three of four represented NYC boroughs and five New York state counties. Our data point to the limited efficacy of travel restrictions into NYC for preventing spread in the metropolitan area once multiple introductions of SARS-CoV-2 and community-driven transmission had already occurred.

Initial diagnostic testing for SARS-CoV-2 in NYC was targeted and limited to individuals who fit a set of criteria outlined by the Centers for Disease Control and Prevention (CDC) and required preapproval by the local Department of Health. Between 9 March and 14 March 2020, screening capacity at the Mount Sinai Health System was greatly expanded, leading to a surge of newly diagnosed cases (Fig. 1). Within 1 week, the number of daily positive SARS-CoV-2 tests exceeded the normal volume of positive tests for influenza virus by a factor of 5. We sequenced 90 SARS-CoV-2 genomes from clinical isolates obtained from 84 of the more than 800 cases identified up to 18 March, yielding 72 complete and 18 partial (>95% coverage) genomes. These cases were drawn from 21 NYC neighborhoods across four boroughs (Manhattan, Bronx, Queens, and Brooklyn) as well as two towns in Westchester County. Sequenced isolates were obtained directly from nasopharyngeal swab samples collected from 44 females (52.4%) and 40 (47.7%) males ranging in age from 20 to 44 years (16%), 45 to 64 years (33%), 65 to 79 years (27.4%), and >80 years (16.7%). Of the 68 cases with available hospital visit information, 12 were discharged (17.6%), 53 were admitted (80%), and three were initially discharged and later admitted on a subsequent visit (2.4%).

We performed phylogenetic analysis of the 84 distinct patient isolates, together with 2363 sequences deposited in GISAID up to 1 April 2020 (Fig. 2A). NYC isolates were distributed throughout the phylogenetic tree, which is consistent with multiple independent introductions. We assigned each SARS-CoV-2 isolate derived from a patient seeking care at the Mount Sinai Health System to one of the main monophyletic clades on the basis of amino acid and nucleotide substitutions and statistical support from both maximum-likelihood (ML) and Bayesian methods. These mutations were used only for the purpose of classification because the functional impact of many of these substitutions remains unknown. As a reference, we adopted the clade nomenclature from the NextStrain tool (*5*, *6*). For each assigned clade, we identified different types of events according to the position of the NYC sequences and available epidemiological information (Table 1). The first isolate was obtained from a patient with documented exposure through travel to the Middle East (clade A3, node 5), and the second was obtained from a patient with documented travel to Europe (clade B, node 6). We therefore excluded these two cases from any inference made from the phylogenetic analyses. For the remaining isolates, the great majority (87%) clustered with clade A2a. This clade is largely composed of isolates obtained from patients with COVID-19 in Europe (72%) (Fig. 2B), suggesting that introductions from Europe account for the majority of cases found in NYC in the first weeks of March 2020.

Most of the NYC isolates within clade A2a were interspersed, and we did not observe grouping by country or geographical region (fig. S1). Despite their diverse origins, many sequences in this clade are highly similar or identical, which makes it impossible to resolve direct relationships or directionality between cases without additional epidemiological data. We next used ML and Bayesian phylodynamic analysis to infer the time of virus transmission events related to NYC (fig. S2A). For clade A2a, we estimated a period of untracked global transmission from late January to mid-February, which is consistent with epidemiological observations of the developing pandemic (Table 1, node 2, and fig. S2A). The earliest sequences at the base of clade A2a include isolates from Italy, Finland, Spain, France, the United Kingdom, and other European countries from late February, in addition to a few North American isolates (Canada and United States) from the first week of March 2020. Within this clade, we identified two mutations that distinguish clusters of sequences from NYC and elsewhere, suggesting at least two independent introductions that likely occurred in early to mid-February (Table 1, nodes 3 to 4, and fig. S2A). One of these two clusters contained 64 of the 72 NYC isolates positioned in clade A2a, suggesting local spread.

Table 1. Inferred SARS-CoV-2 virus transmission events related to NYC.

Clade	First NYC sequence	Mutations	Node	Cases	Event type	Inferred source	tMRCA 95% HPD*
A1a	4 March	ORF3a:G251V, C14805T	1	5	Putative introduction	Europe	25 January to 6 February
A2a	11 March	S:D624G, ORF1b:P314L	2	4	Untracked transmission	N/A	5 January to 29 January
		S:D624G, ORF1b:P314L, OFR3a:Q57H	3	5	Putative introduction	Europe/North America	31 January to 11 February
		S:D624G, ORF1b:P314L, OFR3a:Q57H, ORF1a:T265I	4	64	Putative introduction	Europe/North America	15 February to 19 February
A3	29 February	ORF1a:V378I	5	1	Travel exposure	Middle East (epi link)	29 February (Test date)
B	4 March	ORF8:L84S	6	1	Travel exposure	Europe (epi link)	4 March (Test date)
			7	1	Putative introduction	Europe/Asia	2 February to 22 February
			8	1	Putative introduction	North America/Domestic	28 January to 24 February
B1	17 March	C18060T, ORFb:P1427L	9	1	Putative introduction	Domestic	26 January to 7 February
B4	14 March	N:S202N, ORF14:V94I	10	1	Putative introduction	Asia/Oceania	31 January to 25 February

*Inferred 95% High Posterior Density (HPD) interval for the mean time of tMRCA from the root of the node containing New York–origin strains with ≥70% bootstrap support value in the ML tree and ≥0.9 posterior probability in the maximum clade credibility (MCC) trees inferred for major clades. The 95% HPD intervals shown correspond to the MCC trees inferred by using a strict clock model. Travel exposures and putative introductions are shown for the cases described in this study.

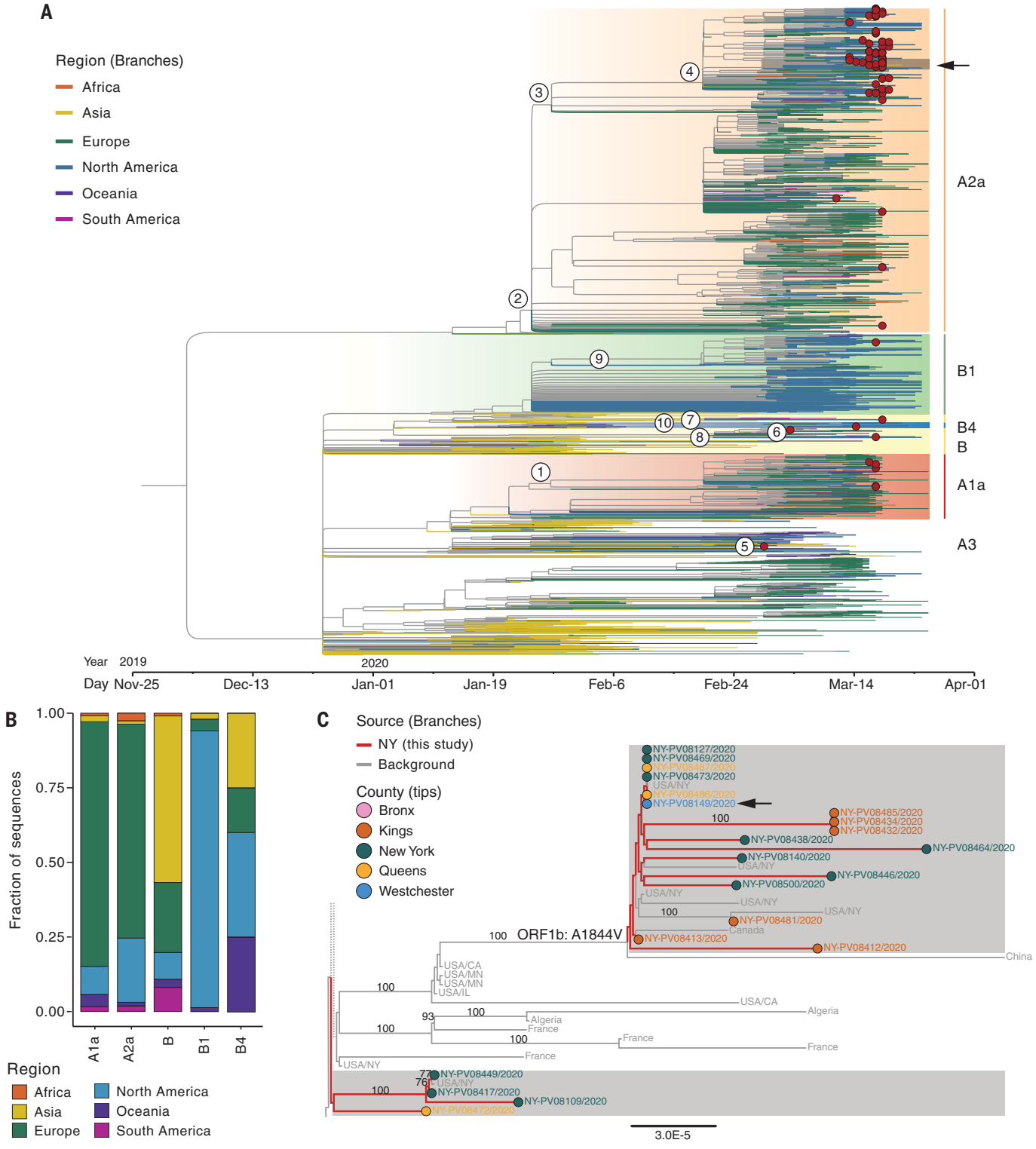


Fig. 2. Phylogenetic relationships of SARS-CoV-2 from New York and other global strains. (**A**) ML phylodynamic inference of 84 SARS-CoV-2 sequences from New York from this study in a global background of 2363 sequences available in the GISAID EpiCoV database as of 1 April 2020. Branches are colored according to the region of origin. Tip circles (red) indicate the position of the 84 New York sequences. Clades that contain New York sequences are highlighted, with names shown on the right; the local transmission clusters are indicated by the arrow. The node positions for the transmission events listed in Table 1 are marked by the numbers in white circles. The displayed time tree was inferred under a strict clock model with a fixed substitution rate of 0.8×10^{-3}. (**B**) Stacked barplot showing the fraction of sequences per region by clade. (**C**) Local transmission clusters on the ML tree showing the source of cases by county. Bootstrap support values ≥70% are shown; sibling clusters are collapsed for easier visualization. The mutation identified specific to the community transmission cluster is indicated. The scale bar at the bottom indicates the number of nucleotide substitutions per site.

Similar to isolates in clade A2a, SARS-CoV-2 isolates in our study positioned in clade A1a (6%) were interspersed among isolates from multiple regions with unknown directionality (Fig. 2A, node 1). This clade is also largely composed of European-origin isolates (82%) (Fig. 2B and fig. S1).

For the rest of the clades (B, B1, and B4), we identified four putative SARS-CoV-2 virus introductions to NYC sometime between February and early March (Table 1, nodes 7 to 10, and fig. S2A). Two of these introductions were inferred to be of domestic origin on the basis of their close relationship with U.S. isolates, including those from the main community transmission in Washington state (clade B1, node 9) (*7*). The introduction of this clade to the East Coast was recently reported (*8*). Although more than half of the sequences in clade B were of Asian origin (Fig. 2B), the closest relatives to the New York isolates were of European and North American origin. The isolate that belongs to clade B4 is within a node with an inferred date of early March and a prior period of untracked transmission in unknown location(s) during February (Table 1, node 10, and fig. S2A). Before this period, the closest viral isolates basal to this cluster are from Australia and China (fig. S1).

The sequenced isolates and assigned clades were spatially distributed throughout all NYC boroughs and 21 neighborhoods (Fig. 3). Isolates sequenced at other NYC hospitals collected within a comparable time window had a similar clade and spatial distribution (*9*).

Taken together, our results show that the NYC SARS-CoV-2 epidemic has been mainly sourced from untracked transmission between the United States and Europe, with limited evidence of direct introductions from China, where the virus originated.

Despite the relatively small number of SARS-CoV-2 sequences available in early April 2020, we identified two monophyletic clusters positioned within clade A2a that almost exclusively contained isolates from New York (Fig. 2B). One cluster included 23% (17) of the isolates contained in clade A2a and 20% of the total isolates sequenced. According to ZIP code information, the cases from this cluster were distributed across five counties, including one sample from New Rochelle, Westchester County, which is part of the metropolitan area directly north of NYC and reported the first documented cluster of community-acquired infections in New York state on 3 March 2020. This cluster was characterized by the amino acid substitution A1844V in the ORF1b gene [in which alanine (A) was replaced at position 1844 with valine (V)]. Basal to these clusters are isolates from the states of Minnesota, Washington,

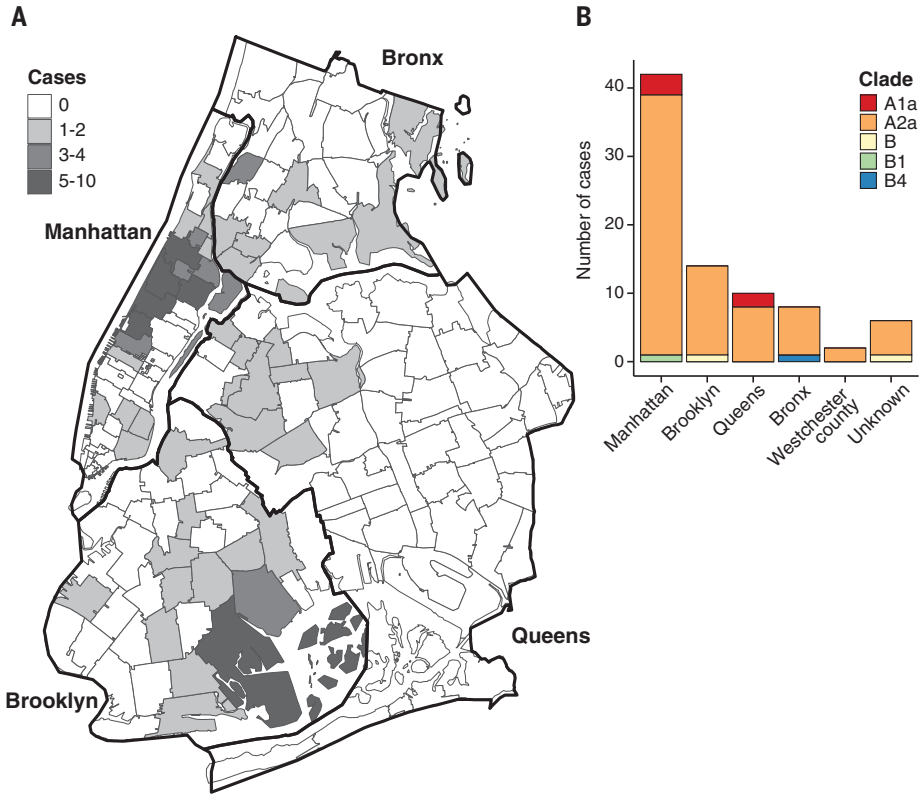


Fig. 3. Distribution of the geographic location of the patients with COVID-19 from which viral isolates were sequenced. (**A**) Distribution of 74 sequenced cases with available ZIP code information across NYC boroughs and neighborhoods. Neighborhoods are shaded according to the number of cases that were sampled. (**B**) Breakdown of sequenced cases according to phylogenetic clades across NYC boroughs and Westchester County. Cases without available ZIP code information are indicated as "Unknown."

and California. The relatedness of other U.S. and New York isolates suggests that viruses spreading locally could have been introduced to New York through a domestic route (fig. S1).

The second cluster was a smaller group that contained four isolates from Manhattan/NYC and one isolate from Queens/NYC. ZIP code information was available for three of the Manhattan cases, which were mapped to three different neighborhoods, further supporting community spread. Although most NYC cases are intermixed within this largely European clade, these results suggest that domestic introductions may have also been a source of early community spread within NYC.

SARS-CoV-2 is the cause of one of the largest noninfluenza pandemics of this century. Sequencing of SARS-CoV-2 isolates from 84 patients diagnosed with COVID-19 at one of the largest healthcare systems in NYC during the first weeks of March 2020 provides insights into the origin and diversity of this new viral pathogen in the region. We found clear evidence for multiple independent introductions into the NYC metropolitan area from different regions globally, as well as from other parts of the United States. Our data indicates that early introductions by cases that were identified on the basis of their known travel histories did not seed the larger community clusters, suggesting that their early quarantine and hospitalization were effective in curtailing further spread. With increased testing, we observed the emergence of community-acquired infections. Most of these were caused by viral isolates derived from clades that were circulating in Europe, likely reflecting local transmissions from undetected introductions.

A limitation of our analysis is the relatively small number of global isolates from cases identified in the first weeks of March 2020, which means that our model relies on inferences based on available background sequences at that time. Thus, some of our inferences may change as more complete and representative SARS-CoV-2 sequences become available. Moreover, because global sequencing efforts are disparate, the fraction of sequences available by region or country is not necessarily representative of the number of cases reported for each of these regions. As shown

by our estimates from different methods and molecular clock models, the time of the most recent common ancestor (tMRCA) estimates and precision can vary even with the same data. Last, in the absence of detailed epidemiological data on travel history and contacts, we were not able to associate periods of untracked transmissions with any specific regions or countries.

Taken together, we provide a first analysis of the SARS-CoV-2 viral genotypes collected from patients seeking medical care in the NYC metropolitan area. We found that NYC, as an international hub, provides not only a snapshot of the diversity of disease-causing SARS-CoV-2 at the global level but also informs on the dynamics of the pandemic at the local level. Future studies as well as additional sequencing are needed to define viral phenotypes and further explore the evolution of the SARS-CoV-2 epidemic in NYC, as well as to assess the impact of the public health measures on community transmission.

REFERENCES AND NOTES

1. Executive Office of the President, Proclamation No. 9984, 85 F.R. 6709 (2020).
2. Executive Office of the President, Proclamation No. 9992, 85 F.R. 12855 (2020).
3. Executive Office of the President, Proclamation No. 9993, 85 F.R. 15045 (2020).
4. Executive Office of the President, Proclamation No. 9996, 85 F.R. 15341 (2020).
5. J. Hadfield *et al*., *Bioinformatics* **34**, 4121–4123 (2018).
6. T. Bedford *et al*., *Nextstrain Build for Novel Coronavirus SARS-CoV-2* (Github, 2020); https://github.com/nextstrain/ncov.
7. T. Bedford *et al*., Cryptic transmission of SARS-CoV-2 in Washington State. medRxiv 2020.04.02.20051417 [Preprint] (16 April 2020); https://doi.org/10.1101/2020.04.02.20051417.
8. J. R. Fauver *et al*., *Cell* S0092-8674(20)30484-0 (2020).
9. M. T. Maurano *et al*., Sequencing identifies multiple, early introductions of SARS-CoV2 to New York City region. medRxiv 2020.04.15.20064931 [Preprint] (23 April 2020); https://doi.org/10.1101/2020.04.15.20064931.
10. mjsull, mjsull/COVID_pipe: initial release. Zenodo (2020); https://doi.org/10.5281/zenodo.3775031.

ACKNOWLEDGMENTS

We are greatly indebted to our clinical colleagues for their dedication and courage in providing continued high-quality medical care under very difficult conditions. We also gratefully acknowledge the authors and originating and submitting laboratories of sequences from GISAID's EpiFlu and EpiCoV (www.gisaid.org) that were used as background for our phylogenetic inferences. We thank D. Ruchnewitz, M. Lässig, M. Nelson, and N. Trovão for their input on the phylogenetic analyses. **Funding:** The Research reported in this paper was supported by the Office of Research Infrastructure of the National Institutes of Health (NIH) under awards S10OD018522 and S10OD026880 as well as institutional and philanthropic funds. Protocols established for this study were in part based on influenza virus sequencing protocols established with the support of CRIP (Center for Research on Influenza Pathogenesis), an NIH-funded Center of Excellence for Influenza Research and Surveillance (CEIRS, contract HHSN272201400008C). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. **Author contributions:** J.A., G.P., A.P.-M., M.G., and E.M.S. provided clinical samples for the study. D.R.A. and J.A. performed clinical evaluations. M.M.H., S.F., J.P., D.R.A., G.P., A.P.-M., M.G., E.M.S., V.S., and H.v.B. were involved in clinical sample accessioning. M.M.H., H.A., and G.K. performed RNA extraction and measured viral titers. A.S.G.-R., Z.K., B.A., A.v.d.G., J.D., N.F., and B.S.M. performed NGS experiments. N.F., B.S.M., I.O., G.D., J.S., S.H.S., Y.W., M.S., and R.S. provided NGS services. M.J.S. and A.O. developed the assembly pipeline and performed genome assembly. K.T. and A.K. performed mining of electronic medical records. A.S.G.-R., M.M.H., M.J.S., B.C., A.O., F.K., A.G.-S., M.L., E.M.S., V.S., and H.v.B. analyzed, interpreted, and/or discussed data. A.S.G.-R., E.M.S., V.S., and H.v.B. wrote the manuscript. E.M.S., V.S., and H.v.B. conceived the study. E.M.S., V.S., and H.v.B. supervised the study. A.G.-S., E.M.S., V.S., and H.v.B. raised financial support. **Competing interests:** R.S. is vice president of technology development and a stockholder at Sema4, a Mount Sinai venture. This work, however, was conducted solely at Icahn School of Medicine at Mount Sinai. **Data and materials availability:** SARS-CoV-2 genome consensus sequences for all study isolates were deposited in GISAID (accession nos. EPI_ISL_414476, EPI_ISL_415151, and EPI_ISL_421348 to EPI_ISL_421435) and GenBank (accession nos. MT370831 to MT370920). The code used to assemble the consensus genome sequences in this study is available online at https://github.com/mjsull/COVID_pipe and (*10*). This study was reviewed and approved by the Institutional Review Board of the Icahn School of Medicine at Mount Sinai (protocols: HS#13-00981 and HS#16-01215). 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/369/6501/297/suppl/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 and S2
References (*11*–*25*)
MDAR Reproducibility Checklist
Data Files S1 and S2

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

10 April 2020; resubmitted 29 May 2020
Accepted 27 May 2020
Published online 29 May 2020
10.1126/science.abc1917

ORGANIC CHEMISTRY

Boron-enabled geometric isomerization of alkenes via selective energy-transfer catalysis

John J. Molloy*†, Michael Schäfer†, Max Wienhold, Tobias Morack, Constantin G. Daniliuc, Ryan Gilmour*

Isomerization-based strategies to enable the stereodivergent construction of complex polyenes from geometrically defined alkene linchpins remain conspicuously underdeveloped. Mitigating the thermodynamic constraints inherent to isomerization is further frustrated by the considerations of atom efficiency in idealized low–molecular weight precursors. In this work, we report a general ambiphilic C_3 scaffold that can be isomerized and bidirectionally extended. Predicated on highly efficient triplet energy transfer, the selective isomerization of β-borylacrylates is contingent on the participation of the boron p orbital in the substrate chromophore. Rotation of the C(sp^2)–B bond by 90° in the product renders re-excitation inefficient and endows directionality. This subtle stereoelectronic gating mechanism enables the stereocontrolled syntheses of well-defined retinoic acid derivatives.

Complex polyenes provide an expansive structural framework for biological processes that occur in two-dimensional chemical space (*1*), yet synthetic paradigms that enable the stereocontrolled construction of these motifs remain conspicuous in their absence. In the biochemical context, information is stored in one of two stereoisomeric forms, such that the spatiotemporal control of constituent alkene units encodes for a specific biological function (*2*–*4*). In cross-conjugated alkenes such as retinal (vitamin A), this interplay facilitates the regulatory gating mechanism that underpins the mammalian visual cycle (*5*). Despite powerful manifestations of this structure-function interplay in human medicine and materials research, isomerization platforms that enable the stereodivergent construction of complex polyenes from common linchpins remain underdeveloped. Synthesis algorithms reported by Burke and co-workers have greatly facilitated the modular construction of polyenes from common building blocks in a manner that is reminiscent of biopolymer assembly (*6*, *7*). Despite these advances, the intractable challenges that are associated with inverting the geometry of small alkene fragments to enable stereodivergence from common precursors offer the opportunity for creative endeavor (*8*). This

Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Corrensstraße 36, 48149 Münster, Germany.
***Corresponding author. Email: molloy@uni-muenster.de (J.J.M.); ryan.gilmour@uni-muenster.de (R.G.)** †These authors contributed equally to this work.

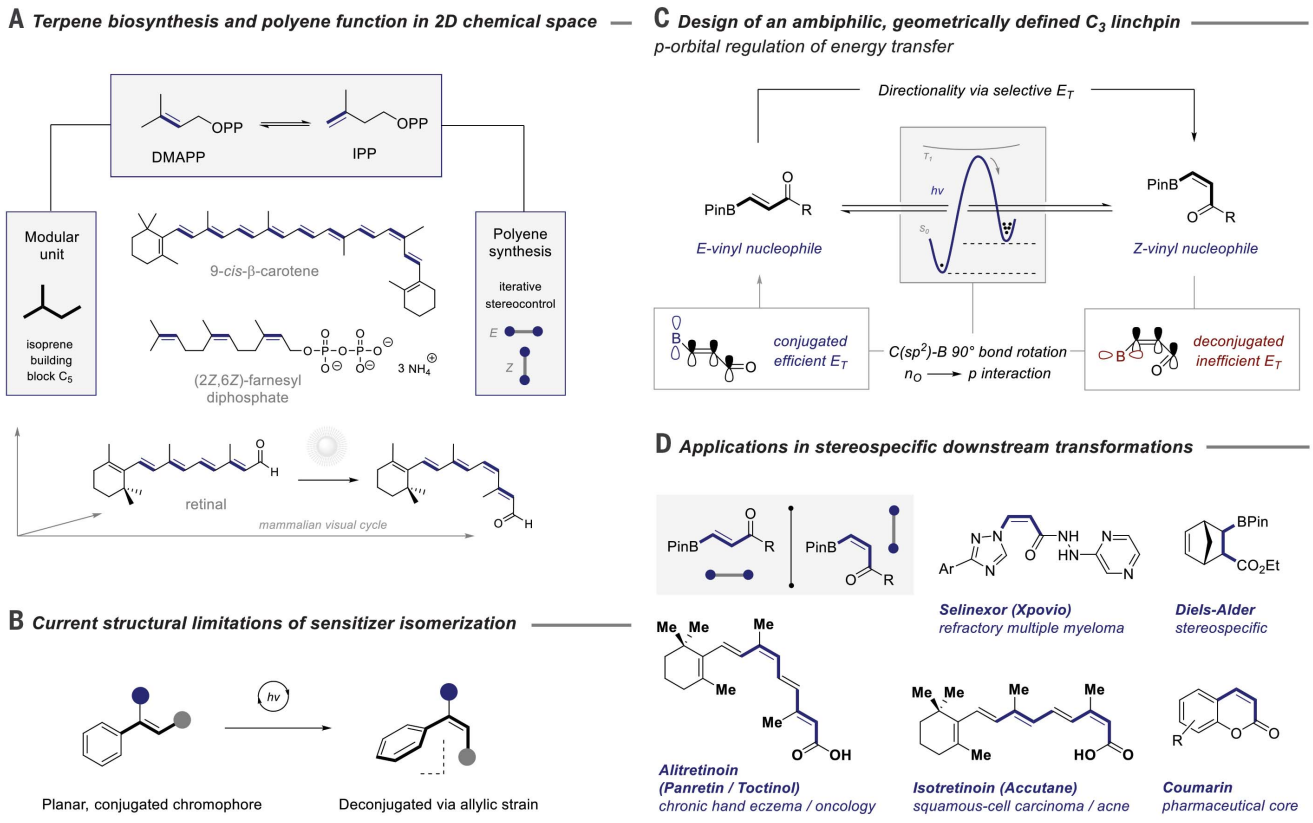


Fig. 1. Inspiration for the development of a photocatalytic isomerization. (**A**) The biosynthesis of complex terpenes via well-defined building blocks and the importance of alkene geometry in the visual cycle. 2D, two-dimensional; OPP, pyrophosphate; DMAPP, dimethylallyl pyrophosphate; IPP, isopentenyl pyrophosphate. (**B**) Photosensitized isomerization of alkenes and the precondition that the alkene is embedded in a larger chromophore (e.g., styrene). hν, photon energy. (**C**) Design of an ambiphilic C_3 linchpin. Selective energy-transfer (E_T) platform is based on the β-boryl-acrylate scaffold. (**D**) Stereospecific applications of the geometrically defined linchpins for the preparation of complex polyenes and pharmaceuticals. The stereochemical assignment of the alkene fragment follows International Union of Pure and Applied Chemistry (IUPAC) nomenclature and reflects the priority order C > B > H.

strategic deficiency in alkene isomerization technology has been highlighted by Baran and co-workers in the context of accessing dienophiles for cycloaddition chemistry (*9*). Terpene biosynthesis predicated on the sequential elongation of small (C5) alkene-containing linchpins such as DMAPP and IPP is a conceptually attractive blueprint in reconciling this disparity (Fig. 1A) (*10*–*12*). Under the auspices of enzyme catalysis, these fragments are unified with high fidelity, thereby enabling the synthesis of well-defined polyenes by either stereodivergent biosynthesis (e.g., farnesyl diphosphate) or postproduct isomerization routes (e.g., 9-*cis*-β-carotene) (*13*). Although challenging to implement in a synthetic context, these precedents highlight the need for an isomerization platform to enable stereodivergent access to both alkene geometries as a prerequisite for efficiency (*14*). Consideration of step and atom economy (*15*) necessitates the use of a common ambiphilic building block that contains the minimum number of carbon atoms needed for sequential manipulation to allow stereospecific downstream processing. The nondegeneracy of alkene stereoisomers and the associated contra-thermodynamic challenges render Dexter-type energy-transfer activation (*16*–*21*) ideally suited to mitigate microscopic reversibility that is intrinsic to ground state reactivity (*22*–*24*). Furthermore, the biocompatibility of this activation strategy has recently been validated by MacMillan and co-workers (*25*).

Photosensitized alkene isomerization has a venerable history, but efficiency often restricts

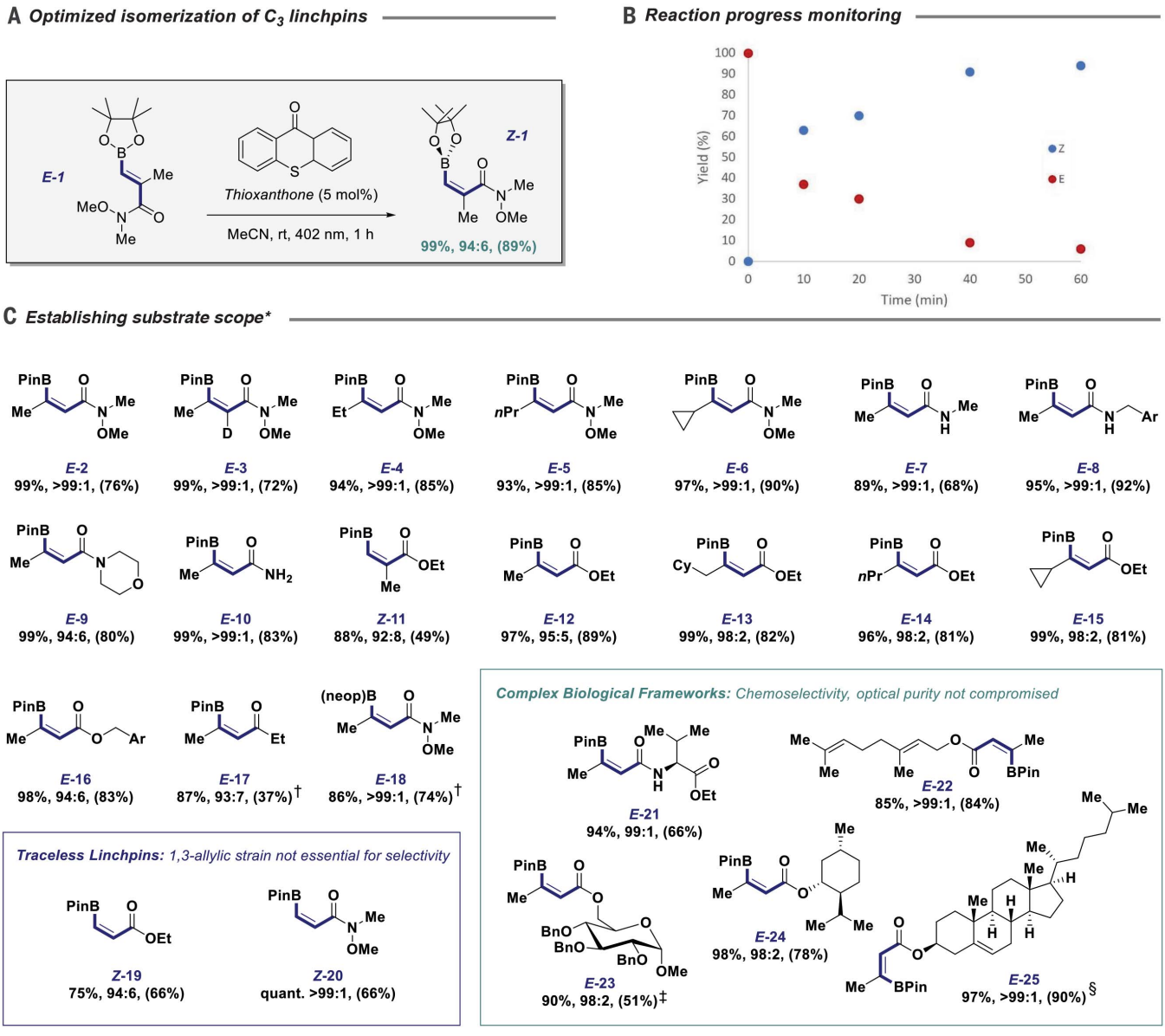


Fig. 2. The photocatalytic isomerization of small molecule ambiphilic C3 linchpins. (**A**) Optimized reaction conditions: substrate (0.3 mmol), thioxanthone (5 mol %), MeCN (9 mL), 402 nm light, 1 hour. rt, room temperature. (**B**) Reaction progress monitoring of ***E*-1** → ***Z*-1**. (**C**) Establishing substrate scope and demonstrating efficiency in simple 1,2-disubstituted systems, showing chemoselectivity and compatibility with biomolecules.

*The stereochemical assignment of the alkene fragment follows IUPAC nomenclature and reflects the priority C > B > H. †Yield over two steps after cross-coupling. ‡0.1 mmol scale. §16 hours. Yields and *E*:*Z* ratios were determined by ^{1}H NMR integration (using 1,3,5-trimethoxybenzene as internal standard). Numbers in parentheses refer to isolated yields of the major product isomer after column chromatography. quant., quantitative.

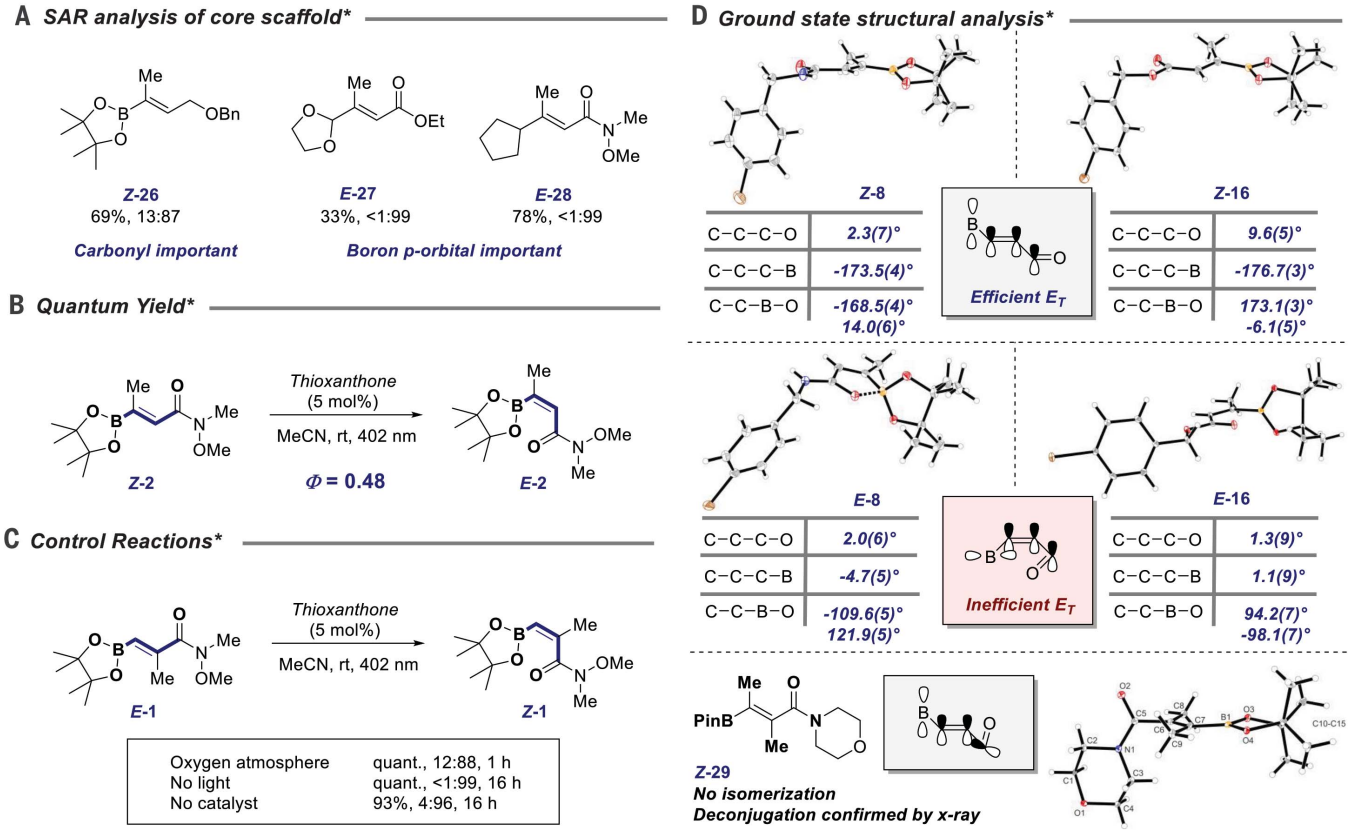


Fig. 3. Investigating the origin of stereoselectivity. (**A**) Structure activity relationship (SAR) of core scaffold. (**B**) Quantum yield studies. (**C**) Control reactions. (**D**) Ground state x-ray structural analysis of ***Z/E*-8**, ***Z/E*-16**, and ***Z*-29**. *The stereochemical assignment of the alkene fragment follows IUPAC nomenclature and reflects the priority C > B > H.

the scope to styrenyl chromophores (*3*). Seminal studies by Hammond *et al.* and Arai *et al.* have established benchmarks for stilbene and styrene isomerization through the use of a photosensitizer (*26*–*29*), and Walker and Radda have simulated triplet energy transfer–enabled selective isomerization of retinal in the visual cycle (*30*). More recently, the Weaver group (*31*) and our own research group (*32*–*35*) have disclosed operationally simple isomerization protocols that harness functionally diverse styrenes with various structural motifs in the β-position (Fig. 1B). However, the presence of an aryl alkene fragment is a precondition to act as a suitable chromophore. Thus, the identification of a traceless chromophore would represent a notable advance toward achieving operational parity with positional isomerization (*36*).

To circumvent this limitation and enable a modular assembly of polyenes (*6*–*8*, *37*) with sequential control of alkene geometry, we consolidated the importance of acetyl derivatives in biosynthesis and the need for an ambiphilic fragment in the β-borylacrylate scaffold (Fig. 1C). Specifically, we identified the boronic acid pinacol ester (β-BPin) as a group that might enable isomerization to generate both alkene isomers and provide a handle for traceless, stereospecific coupling. The conceptual framework of the alkene isomerization was predicated on rotation of the C(sp^2)–B bond as a stereoelectronic gating mechanism to enable selective energy transfer. The p orbital of trigonal planar boron systems plays a pivotal role in regulating the reactivity of organoboron compounds (*38*, *39*). We envisaged that the incorporation of a traceless boronic ester motif as an aryl surrogate would extend the acrylate chromophore and permit subsequent excitation of the planar *E* isomer. After excitation and subsequent isomerization, the C(sp^2)–B bond in the product *Z* isomer would be twisted by 90° and could potentially engage in a dative interaction with the carbonyl group ($n_O \rightarrow p$) (*40*, *41*). This subtle bond rotation would shorten the chromophore and provide a structural foundation to bias the photostationary composition.

In this work, we report a photocatalytic isomerization of β-borylacrylate derivatives that enables access to both geometric isomers of ambiphilic C_3 linchpins. Whereas hydroboration provides a convenient platform to generate a single alkene isomer, this approach harnesses Dexter-type selective energy transfer using an inexpensive small molecule photosensitizer (excited at 402 nm) to enable stereodivergence. Directionality ($E \rightarrow Z$) is a consequence of a subtle 90° C(sp^2)–B bond rotation, thereby contracting the π system of the *Z* isomer. This strategy circumvents the reliance of alkene isomerization on styrenyl chromophores by replacing inert aryl rings with a traceless BPin handle. Both isomers of the vinyl boron species are compatible with stereospecific cross-coupling, which provides the basis for an iterative, stereodivergent entry to complex polyenes. Mechanistic validation and the total syntheses of the therapeutics isotretinoin and alitretinoin are disclosed, and we showcase the synthetic utility of these small, densely functionalized units in stereospecific transformations (Fig. 1D).

Initially, we investigated the geometric isomerization of the ambiphilic Weinreb amide ***E*-1** to ***Z*-1** (Fig. 2A and supplementary materials). Reactions performed in acetonitrile under visible light irradiation (402 nm) using thioxanthone as a photocatalyst (5 mol %) proved to be operationally simple and highly

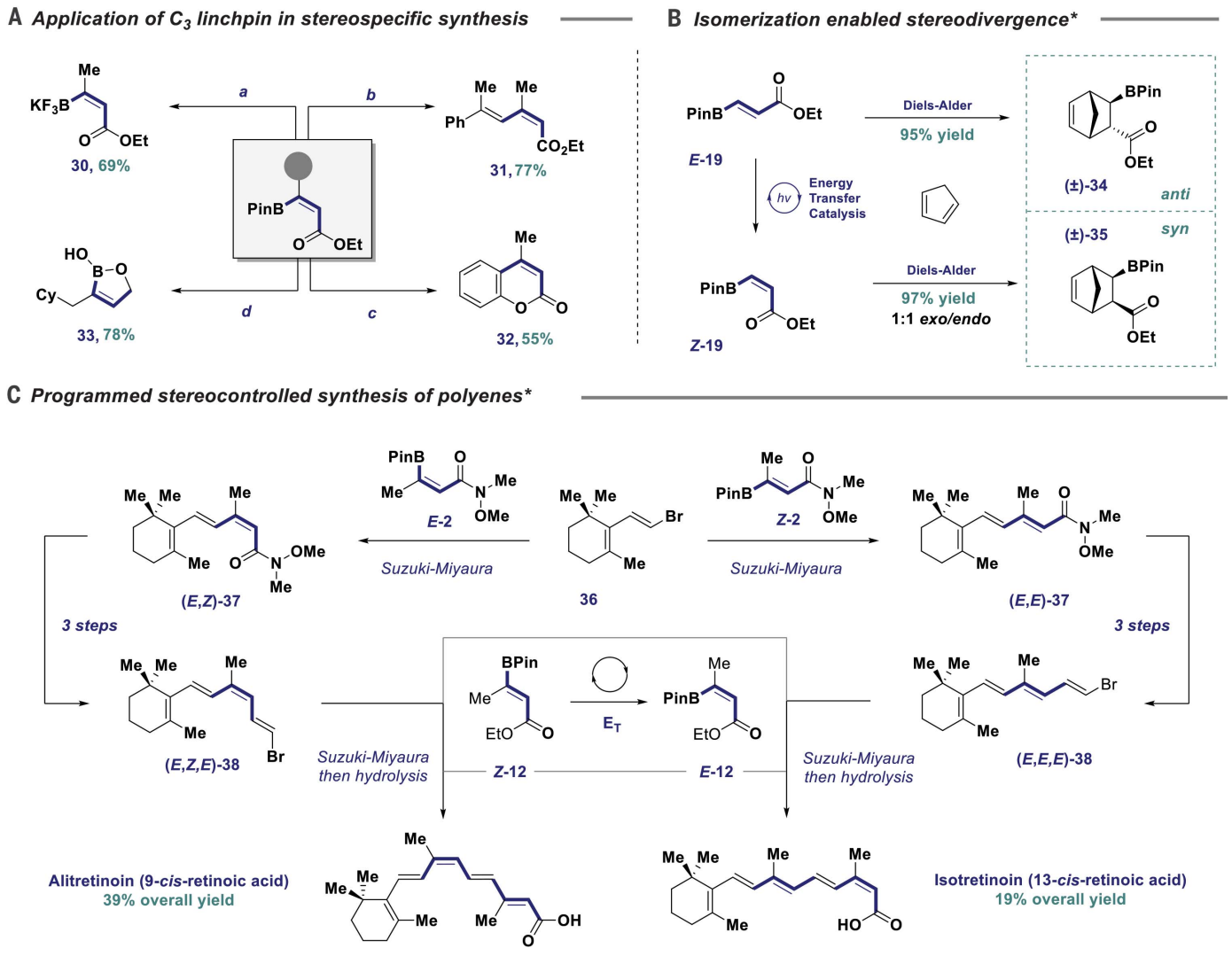


Fig. 4. Stereospecific and stereodivergent application of small molecule C_3 linchpins. (**A**) Application of product isomer in synthesis: (a) ***E*-12** (1 equiv), aq. KHF_2 (4.5 M, 5 equiv), MeOH, rt. (b) ***E*-12** (1 equiv), $Pd(OAc)_2$ (5 mol %), SPhos (10 mol %), K_3PO_4 (3 equiv), vinyl halide (1.2 equiv), H_2O (5 equiv), 1,4-dioxane, 80°C. (c) ***E*-12** (1 equiv), $Pd(OAc)_2$ (5 mol %), SPhos (10 mol %), K_3PO_4 (3 equiv), 2-bromophenol (1.2 equiv), H_2O (5 equiv), 1,4-dioxane, 80°C. (d) ***E*-13** (1 equiv), $NaBH_4$ (2 equiv), rt. (**B**) Stereodivergent Diels-Alder reaction: toluene, 16 hours. (**C**) Stereocontrolled synthesis of the polyene pharmaceuticals alitretinoin (9-*cis*-retinoic acid) and isotretinoin (13-*cis*-retinoic acid). Full details are provided in the supplementary materials. *The stereochemical assignment of the alkene fragment follows IUPAC nomenclature and reflects the priority C > B > H.

efficient [99% nuclear magnetic resonance (NMR) yield, *Z*:*E* = 94:6; Fig. 2, A and B]. With these optimized conditions, we explored the substrate scope and limitations of the method (Fig. 2C). Relocating the methyl substituent to the β-position of the Michael acceptor ***E*-2** and ***E*-3** led to notable selectivity (>99:1). Modification of the alkyl substituent was very well tolerated, furnishing the isomerized products in >99:1 selectivity and with high isolated yields (***E*-4**, ***E*-5**, and ***E*-6**).

Replacement of the Weinreb amide by secondary amides (***E*-7** and ***E*-8**), tertiary amides (***E*-9**), and primary carboxamides (***E*-10**) had no impact on the efficiency of the transformation (up to >99:1). Ester groups were also found to be compatible with the general isomerization conditions, but the α-substituted Michael acceptor (***Z*-11**) proved to be unstable toward silica gel chromatography. High levels of stereoselectivity and isolated yield were also observed with β-substituted analogs (***E*-12** to ***E*-16**, up to 98:2). Efficient isomerization was observed with ketones and neopentyl boronic esters (***E*-17** and ***E*-18**, 93:7 and >99:1, respectively), but the intrinsic instability of the products required a cross-coupling protocol after isomerization to determine stereoselectivity. To explore the limitations of the method in simple 1,2-disubstituted frameworks, we investigated acrylates **19** and **20**. Photochemical alkene isomerization typically requires a substituent to generate 1,3-allylic strain in the product, thereby causing deconjugation, reducing the efficiency of energy transfer, and enabling directionality. In contrast to styrenyl systems, the isomerization of the basic β-borylacrylate also occurs with high levels of stereoselectivity (***E*-19** and ***E*-20**, 94:6 and >99:1, respectively), which provides entry to simple 1,2-disubstituted ambiphilic linchpins. To ensure that the method does not compromise optical purity and is compatible with preexisting alkenes (thereby demonstrating chemoselectivity), we prepared **21** and **22**, respectively. Finally, extension of

the scope to common, biologically relevant frameworks included D-glucose derivative **23**, the (–)-menthol derivative **24**, and modified cholesterol **25** (up to >99:1).

The efficiency of isomerization using simple 1,2-disubstituted building blocks such as **19** and **20** indicates that stereoselectivity is not contingent on allylic strain as is observed with styrenyl systems. To discern the structural basis of the selectivity, we investigated the effects of single-site editing on the core substrate (Fig. 3A). These investigations revealed that both the boron and the carbonyl substituent are critical to achieve efficient directionality, as is clear from ***E*-26**, **27**, and **28**. Determination of the quantum yield excludes the possibility of an efficient radical chain process, and control experiments demonstrate the need for both light and a photocatalyst (Fig. 3, B and C). In establishing the structural basis of directionality, inspection of the ^{11}B NMR spectra of the products proved instructive. Accumulation of electron density in the boron p orbital is conspicuous in the product *E* isomer of the amide derivatives [17.19 ppm (parts per million) for ***E*-8**]. This is indicative of a stabilizing dative ($n_O \rightarrow p$) interaction that is only geometrically viable when the C(sp^2)–B bond rotates out of the plane of the π-system (fig. S6). This deconjugative contraction of the chromophore does not manifest itself in the ^{11}B NMR spectra of the ester substrates (30.76 ppm for ***E*-16**), which prompted further investigation.

It was possible to analyze the substrate and product antipodes of **8** and **16** by x-ray crystallography (Fig. 3D). Comparison of the starting isomers ***Z*-8** and ***Z*-16** confirmed the expected planar chromophore in which the boron group is conjugated (Fig. 3D, top). Consistent with the ^{11}B NMR studies, a clear dative interaction is evident in the amide product ***E*-8** resulting from a 90° rotation of the C(sp^2)–B bond (Fig. 3D, middle). This subtle contraction of the π-system would likely raise the triplet energy and render re-excitation inefficient. The ester derivative ***E*-16** shows no dative interaction, but, consistent with the amide scenario, the boron p orbital is perpendicular to the alkene π-system and no longer in conjugation. To further demonstrate the importance of conjugation for efficient sensitization, the *tetra*-substituted precursor ***Z*-29** was investigated. X-ray analysis revealed that the carbonyl group is out of plane, thereby rendering excitation inefficient. Although *tetra*-substituted alkenes are seldom found in polyenes, this substrate is a useful control.

To validate this isomerization platform in stereocontrolled polyene synthesis, we investigated the synthetic utility of simple ambiphilic linchpins in target synthesis and a series of downstream manipulations (Fig. 4). Boronic ester protection to generate the corresponding trifluoroborate was facile (**30**; Fig. 4A). Similarly, stereospecific cross-coupling to form diene **31** with alternating geometrical isomers proceeded in high yields, whereas the use of 2-bromo phenol allowed the rapid generation of the coumarin scaffold **32**. A known covalent binder in medicinal chemistry, the oxaborole scaffold (**33**) could be accessed by standard reduction conditions. Isomerization enabled stereodivergent Diels-Alder reactions to generate the *anti*-adduct **34** and the *syn*-adduct **35** (Fig. 4B). Our ultimate goal was to enable the programmed, stereocontrolled synthesis of polyenes by using simple, geometrically defined alkenes (Fig. 4C). Using easily accessible vinyl bromide **36**, stereospecific cross-coupling using (*Z*)- and (*E*)-**2** allowed expedient access to geometrical isomers (*E*, *E*)- and (*E*, *Z*)-**37** in high yield and selectivity. Subsequent reduction to the aldehyde—using the boron-Wittig reaction, as developed by Morken and coworkers (*42*)—followed by bromodeboronation furnished stereodefined vinyl bromide intermediate (*E*, *E*, *E*)- and (*E*, *Z*, *E*)-**38** (see supplementary materials for full details). Suzuki-Miyaura cross-coupling (*43*–*45*) with ***E*-12** proceeded smoothly to furnish 13-*cis*-retinoic acid (isotretinoin) after subsequent hydrolysis. The analogous sequence with ***Z*-12** enabled the synthesis of 9-*cis*-retinoic acid (alitretinoin) (*46*). This platform for the stereocontrolled generation of complex polyenes might prove to be expansive and may facilitate the exploration of these bioactive materials in drug discovery.

REFERENCES AND NOTES

1. C. Dugave, L. Demange, *Chem. Rev.* **103**, 2475–2532 (2003).
2. C. M. Pearson, T. N. Snaddon, *ACS Cent. Sci.* **3**, 922–924 (2017).
3. J. J. Molloy, T. Morack, R. Gilmour, *Angew. Chem. Int. Ed.* **58**, 13654–13664 (2019).
4. L. F. Maia, R. F. Fernandes, G. Lobo-Hajdu, L. F. C. de Oliveira, *Phil. Trans. R. Soc. A* **372**, 20140200 (2014).
5. P. D. Kiser, M. Golczak, K. Palczewski, *Chem. Rev.* **114**, 194–232 (2014).
6. E. M. Woerly, J. Roy, M. D. Burke, *Nat. Chem.* **6**, 484–491 (2014).
7. S. J. Lee, K. C. Gray, J. S. Paek, M. D. Burke, *J. Am. Chem. Soc.* **130**, 466–468 (2008).
8. J. W. Lehmann, D. J. Blair, M. D. Burke, *Nat. Rev. Chem.* **2**, 0115 (2018).
9. T.-G. Chen *et al*., *Nature* **560**, 350–354 (2018).
10. A. Eschenmoser, L. Ruzicka, O. Jeger, D. Arigoni, *Helv. Chim. Acta* **38**, 1890–1904 (1955).
11. A. Eschenmoser, D. Arigoni, *Helv. Chim. Acta* **88**, 3011–3050 (2005).
12. J. W. Cornforth, *Pure Appl. Chem.* **2**, 607–630 (1961).
13. G. W. Burton, K. U. Ingold, *Science* **224**, 569–573 (1984).
14. B. M. Trost, *Science* **219**, 245–250 (1983).
15. B. M. Trost, *Science* **254**, 1471–1477 (1991).
16. M. Kathan, S. Hecht, *Chem. Soc. Rev.* **46**, 5536–5550 (2017).
17. R. S. Stoll, S. Hecht, *Angew. Chem. Int. Ed.* **49**, 5054–5075 (2010).
18. M. Kathan *et al*., *Nat. Chem.* **10**, 1031–1036 (2018).
19. A. Hölzl-Hobmeier *et al*., *Nature* **564**, 240–243 (2018).
20. E. Ota, H. Wang, N. L. Frye, R. R. Knowles, *J. Am. Chem. Soc.* **141**, 1457–1462 (2019).
21. N. Y. Shin, J. M. Ryss, X. Zhang, S. J. Miller, R. R. Knowles, *Science* **366**, 364–369 (2019).
22. F. Strieth-Kalthoff, M. J. James, M. Teders, L. Pitzer, F. Glorius, *Chem. Soc. Rev.* **47**, 7190–7202 (2018).
23. L. Marzo, S. K. Pagire, O. Reiser, B. König, *Angew. Chem. Int. Ed.* **57**, 10034–10072 (2018).
24. Q.-Q. Zhou, Y.-Q. Zou, L.-Q. Lu, W.-J. Xiao, *Angew. Chem. Int. Ed.* **58**, 1586–1604 (2019).
25. J. B. Geri *et al*., *Science* **367**, 1091–1097 (2020).
26. G. S. Hammond, J. Saltiel, *J. Am. Chem. Soc.* **84**, 4983–4984 (1962).
27. G. S. Hammond *et al*., *J. Am. Chem. Soc.* **86**, 3197–3217 (1964).
28. T. Arai, H. Sakuragi, K. Tokumaru, *Chem. Lett.* **9**, 261–264 (1980).
29. T. Arai, H. Sakuragi, K. Tokumaru, *Bull. Chem. Soc. Jpn.* **55**, 2204–2207 (1982).
30. A. G. Walker, G. K. Radda, *Nature* **215**, 1483 (1967).
31. K. Singh, S. J. Staig, J. D. Weaver, *J. Am. Chem. Soc.* **136**, 5275–5278 (2014).
32. J. B. Metternich, R. Gilmour, *J. Am. Chem. Soc.* **137**, 11254–11257 (2015).
33. J. B. Metternich, R. Gilmour, *J. Am. Chem. Soc.* **138**, 1040–1045 (2016).
34. J. J. Molloy, J. B. Metternich, C. G. Daniliuc, A. J. B. Watson, R. Gilmour, *Angew. Chem. Int. Ed.* **57**, 3168–3172 (2018).
35. S. I. Faßbender, J. J. Molloy, C. Mück-Lichtenfeld, R. Gilmour, *Angew. Chem. Int. Ed.* **58**, 18619–18626 (2019).
36. A. Vasseur, J. Bruffaerts, I. Marek, *Nat. Chem.* **8**, 209–219 (2016).
37. M. Burns *et al*., *Nature* **513**, 183–188 (2014).
38. A. J. J. Lennox, G. C. Lloyd-Jones, *Chem. Soc. Rev.* **43**, 412–443 (2014).
39. D. G. Hall, Ed., *Boronic Acids* (Wiley, 2011).
40. R. J. Grams, R. G. Fritzemeier, C. Slebodnick, W. L. Santos, *Org. Lett.* **21**, 6795–6799 (2019).
41. K. Nagao, A. Yamazaki, H. Ohmiya, M. Sawamura, *Org. Lett.* **20**, 1861–1865 (2018).
42. J. R. Coombs, L. Zhang, J. P. Morken, *Org. Lett.* **17**, 1708–1711 (2015).
43. N. Schneider, D. M. Lowe, R. A. Sayle, M. A. Tarselli, G. A. Landrum, *J. Med. Chem.* **59**, 4385–4402 (2016).
44. J. W. B. Fyfe, A. J. B. Watson, *Chem* **3**, 31–55 (2017).
45. A. J. J. Lennox, G. C. Lloyd-Jones, *Angew. Chem. Int. Ed.* **52**, 7362–7370 (2013).
46. S. Khalil *et al*., *J. Dermatolog. Treat.* **28**, 684–696 (2017).

ACKNOWLEDGMENTS

This work is dedicated to the memory of Dr. Duilio Arigoni. We thank T. Neveselý for helpful discussions. **Funding:** We acknowledge financial support from the Westfälische Wilhelms-Universität Münster, the Deutsche Forschungsgemeinschaft (SFB 858), the Alexander von Humboldt Foundation (postdoctoral fellowship to J.J.M.), and the Verband der Chemischen Industrie (Kekulé Fellowship to T.M.). **Author contributions:** All authors contributed to the conceptualization of this study, the experimental design, and the interpretation of data. J.J.M. and R.G. directed the research. J.J.M. and R.G. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** X-ray structural data for ***Z*/*E*-8** (CCDC 1990167 and 1990168), ***Z*/*E*-16** (CCDC 1990169 and 1990170), and ***Z*-29** (CCDC 1990171) can be obtained free of charge from the Cambridge Crystallographic Data Centre (CCDC). All other experimental data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/302/suppl/DC1
Materials and Methods
Figs. S1 to S11
Table S1
NMR Spectra
References (*47*–*90*)

13 March 2020; accepted 28 May 2020
10.1126/science.abb7235

CHEMICAL PHYSICS

Molecular square dancing in CO-CO collisions

Zhong-Fa Sun[1,2], Marc C. van Hemert[3], Jérôme Loreau[4], Ad van der Avoird[5], Arthur G. Suits[6], David H. Parker[1]*

Knowledge of rotational energy transfer (RET) involving carbon monoxide (CO) molecules is crucial for the interpretation of astrophysical data. As of now, our nearly perfect understanding of atom-molecule scattering shows that RET usually occurs by only a simple "bump" between partners. To advance molecular dynamics to the next step in complexity, we studied molecule-molecule scattering in great detail for collision between two CO molecules. Using advanced imaging methods and quasi-classical and fully quantum theory, we found that a synchronous movement can occur during CO-CO collisions, whereby a bump is followed by a move similar to a "do-si-do" in square dancing. This resulted in little angular deflection but high RET to both partners, a very unusual combination. The associated conditions suggest that this process can occur in other molecule-molecule systems.

If art follows nature, then the square dance call "swing your partner, do-si-do" must have been inspired by "dance moves" laid out during the collision of two CO molecules, where despite a considerable "swing" (conversion of translational to rotational energy), both molecules leave the encounter moving forward. Up to now, the dance moves for rotational energy transfer (RET) in atom-molecule scattering have been described as being annoyingly predictable (*1*): Large swings are always accompanied by about-face or backward scattering. Advanced experimental methods, particularly velocity map imaging (VMI) (*2*), can now expose details of molecule-molecule scattering with a resolution sufficient to recognize new collision processes involving unexpected motion, as shown in this report. Although the likelihood of the unexpected process observed here depends strongly on the initial collision geometry, we show that it is responsible for production of up to 20% of high-energy rotational states.

Our current understanding of RET, which is the most probable energy-exchange process during a molecular collision, is nearly perfect for atom-diatom collisions, based in detail on accurate fully quantum calculations and in general on simple and precise models that treat the collision as a point mass colliding with a hard-shell ellipsoid scaled to the shape of the molecule. Along with predicting the directionality of scattering, these models also lead to the exponential energy gap law (EGL) (*3*), where the probability of collision-induced transfer is found to be proportional to $\exp(-C\Delta E)$, with ΔE being the energy difference between the ground and excited rotational states and C being a constant. The EGL ascribes a much lower probability of rotational excitation to higher j levels, where j is the rotational quantum number. A key question is whether this picture always applies to molecule-molecule scattering.

Because a single collision can cause rotational excitation of both collision partners in molecule-molecule interactions, the ideal RET experiment provides for each possible product-pair the directionality of scattering, also known as the pair-correlated differential scattering cross section (PC-DCS). This measurement is now possible when using crossed molecular beam scattering (*4*) combined with VMI, where each molecular beam has a narrow (*5*) or ultra-narrow and controlled (*6*) velocity distribution and contains only the lowest ($j = 0$) rotational state of the molecule of interest. Nascent product molecules in a chosen final (j') state are then converted to ions by resonant laser ionization, and their velocity distribution is mapped using VMI with high speed and angular resolution onto a two-dimensional image (*7*–*9*). Our study explores (nearly) identical–molecule scattering, which offers conceptual and computational advantages over mixed molecules. We characterize RET in the collision system $^{13}CO(j = 0) + {}^{12}CO(j = 0) \rightarrow {}^{13}CO(j') + {}^{12}CO(j'')$, where, for practical reasons, specific j' rotational quantum levels of ^{13}CO product molecules are detected. Besides being of fundamental interest, CO-CO scattering is relevant in the pure CO atmosphere of exotic exoplanets (*10*) and in cometary atmospheres, where CO is the second most abundant molecule after H_2O. CO has a small dipole moment and forms a dimer with two stable, end-on antiparallel geometries with relatively small binding energies of about 125 cm^{-1} (*11*). The CO-CO center-of-mass distance is substantially larger for the geometry when the C ends are closer than when the O ends are closer, because the C end is effectively larger than the O end. We show that the latter property also has interesting effects in CO-CO collisions.

We report here the scattering of $^{12}CO(j = 0)$ with $^{13}CO(j = 0)$ to form $^{13}CO(j' = 15)$, where the experiment yields the PC-DCS for each j'' level of $^{12}CO(j'' = 0, 1, 2, \ldots,$ up to $j''_{max})$ paired with $^{13}CO(j' = 15)$. However, in this so-called "heavy-heavy" collision system where the CO rotation energy spacing is much smaller than the collision energy, only partial (but sufficient) resolution of the product-pair DCSs was possible. Note that full resolution of PC-DCSs for collisions of NO with O_2 (*12*) and ND_3-D_2 (*13*) have been recently reported, both at a lower collision energy than that of the present system.

Crossed beam scattering with identical molecules poses severe overlap problems that were solved here using a beam of ^{13}CO and selective ^{13}CO ionization. Moreover, we detected final states that were too high in energy to be populated by collision of ^{13}CO with the carrier gas used for rotational-state cooling of CO in the partner beam, as described in the supplementary materials (SM). The conditions that avoided overlap allowed detection of $^{13}CO(j' > 10)$ final states from CO-CO scattering at a 1460-cm^{-1} collision energy, which, for a "heavy-heavy" system, has too many open channels at present for analysis by fully quantum theory. Quasi-classical trajectory (QCT) analysis on an accurate CO-CO potential energy surface (*14*, *15*), proven to be reliable for CO-He and CO-Ar collisions in previous publications (*16*, *17*), was used here to simulate the scattering process. The QCT results were confirmed at a lower collision energy by fully quantum close-coupling (CC) calculations and compared with those from a previous fully quantum theory study on CO-CO (*18*) described in the SM.

Details of the experiment and analysis methods are presented in the SM. Raw velocity map images of $^{13}CO(j' = 15)$ molecules formed by scattering with CO and with N_2 for comparison (fig. S1) showed an obvious difference of "extra" signal in the forward scattering region around high $j'' \approx j'$, hereafter labeled as forward-scattered symmetric excitation (FSSE), which was present for ^{13}CO-CO but missing for ^{13}CO-N_2. After correcting the raw images for ionization laser polarization sensitivity and the density-flux transform, the experimental data are plotted in Fig. 1 in (j'', θ) coordinates, where the image radius is converted to the j'' position. FSSE signal peaked at $j'' = 15$, which was seen only in the CO + CO data, as indicated by an arrow in both the experimental plot (Fig. 1, left) and the corresponding theoretical prediction (Fig. 1, right). An explanation for why FSSE is absent for the kinematically similar CO + N_2 system was provided by the QCT analysis, as described later in the text.

The dotted diagonal lines shown in Fig. 1 indicate regions of atom-molecule–type behavior for inelastic scattering, where conversion

[1]Department of Molecular and Laser Physics, Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, Netherlands. [2]Anhui Province Key Laboratory of Optoelectric Materials Science and Technology, Department of Physics, Anhui Normal University, Wuhu, Anhui 241000, China. [3]Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, Einsteinweg 55, 2333 CC Leiden, Netherlands. [4]Department of Chemistry, KU Leuven, B-3001 Leuven, Belgium. [5]Theoretical Chemistry, Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, Netherlands. [6]Department of Chemistry, University of Missouri, Columbia, MO 65211, USA.
*Corresponding author. Email: parker@science.ru.nl

of translational energy to an increasingly larger amount of rotational energy required an increasingly "head-on" collision with the target molecule, and thus a smaller impact parameter (the distance of closest approach that would have occurred without any interaction). Small–impact parameter collisions result in scattering toward increasingly larger (backward) angles. The longer dotted diagonal line follows the ridge of the primary rainbow [a build-up of trajectories that pass the threshold for rotational excitation (*1*)], whereas the shorter dotted line, at larger scattering angles, follows the expected secondary rainbow.

Figure 1 shows satisfactory agreement, including the FSSE signal, between experiment and QCT. A full description of the QCT results, including the FSSE versus EGL relative yields in CO-CO RET at 1460- and 500-cm^{-1} collision energies, is presented in the SM section. The (j',j'' = 15, 15) pair-correlated differential cross section with forward scattering angles of 30° or less originated predominantly from trajectories with an impact parameter between 3.0 and 3.5 Å. More than half of the trajectories leading to forward scattering were very similar to the prototypical trajectories shown in movie S1 and sketched in Fig. 2. For comparison, movie S2 shows non-FSSE collisions producing (15, 5) products.

Three unusual features in Fig. 2 characterized FSSE: (i) initial collision geometry, (ii) abruptness of RET, and (iii) the do-si-do movement. The collision geometry is shown in the middle region of Fig. 2 as a series of stills at time positions A to F, from a visual molecular dynamics (*19*) visualization of a prototypical (j', j'' = 15, 15) trajectory. The most effective initial geometry was where the two C ends collide, with an impact parameter of 3.2 Å. Forward scattering was most probable when both the molecular axes were in the same collision plane ($\phi \approx 0$), defined with respect to the internuclear separation vector ***R***. The interaction energy curve in Fig. 2 shows that collision at point B, 240 fs, was abrupt and accounted for the translation to RET, with both molecules rotating individually and in the same sense. Most notable in Fig. 2 was the relatively constant value of the center-of-mass distance between the two diatoms (R) from ~240 to ~600 fs, which was roughly half of a CO(j = 15) rotational period. The average speed here was ≥492 m/s, so near-zero recoil means that the molecules swing around each other with little repulsion or additional torque to reach the O-end facing or back-to-back (do-si-do) configuration. In square dancing, a feature of the do-si-do is that both pairs of feet point outward, avoiding tripping. In the case of CO-N_2, FSSE was not possible because of "tripping" by the ellipsoidal N_2 partner. Two ovoids, colliding first at their large ends, are needed for FSSE. As shown in fig. S5C, FSSE

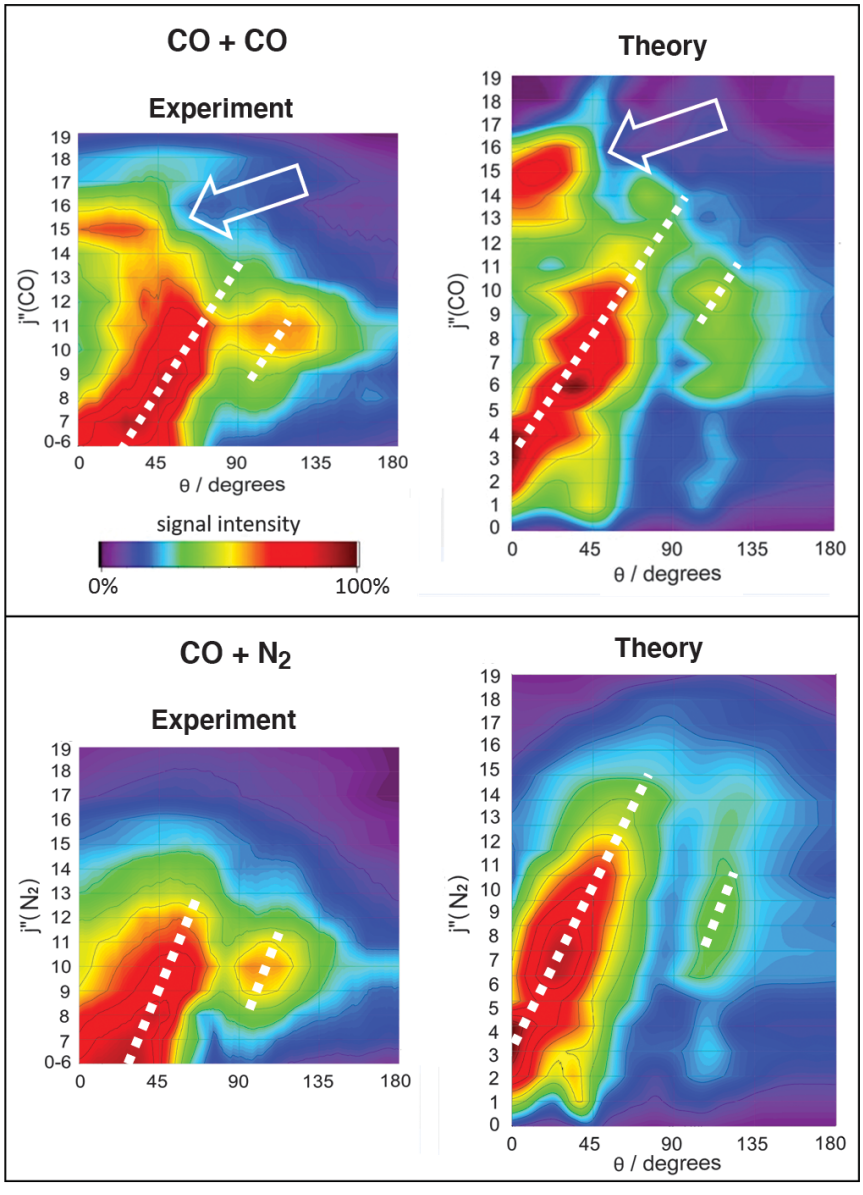


Fig. 1. Contour plot of experimental and theoretical PC-DCSs for ^{13}CO + CO and ^{13}CO + N_2 inelastic scattering to the j' (^{13}CO) = 15 final state. (Left) Partially resolved experimental PC-DCSs with scattering angle on the horizontal axis; the vertical axis indicates the final rotational state of the CO or N_2 coproduct, j''(CO or N_2), which ranges from the low states [j''(CO or N_2) = 0 to 6] at the bottom to the higher states at the top, up to j''_{max}(CO or N_2) = 19. The j''(CO or N_2) = 0 to 6 signals are not resolved. (Right) Theoretical PC-DCSs fully resolved in the individual final rotational state of coproduct CO or N_2, with j'' (CO or N_2) convoluted over the experimental uncertainty. The arrow in both panels for ^{13}CO + CO indicates the region of the unexpected forward-scattered high $j'' \approx j'$ signal, which is not observed for the kinematically similar ^{13}CO + N_2 scattering.

was not due to the displaced center of mass in CO but was rather a size (electronic volume) effect of the potential energy surface. Following the do-si-do, which allows RET far beyond EGL predictions, the rotationally excited molecules departed in the forward-scattering direction.

FSSE collisions showed two sequential interactions, where the first, C-C end, collision resulted in translation to RET, whereas the second interaction at the O-O end reoriented the recoil direction of the two molecules as they separated, without changing the rotational energy. Viewing the process in reverse, the O-O–end interaction first swings the molecules around closer to each other so that when they slam together on the C-C ends, all rotation energy is up-converted to faster linear motion. QCT yielded a final impact parameter (i.e., for the reverse reaction) of b_f = 4.0 ± 0.1 Å, a value that is 25% larger

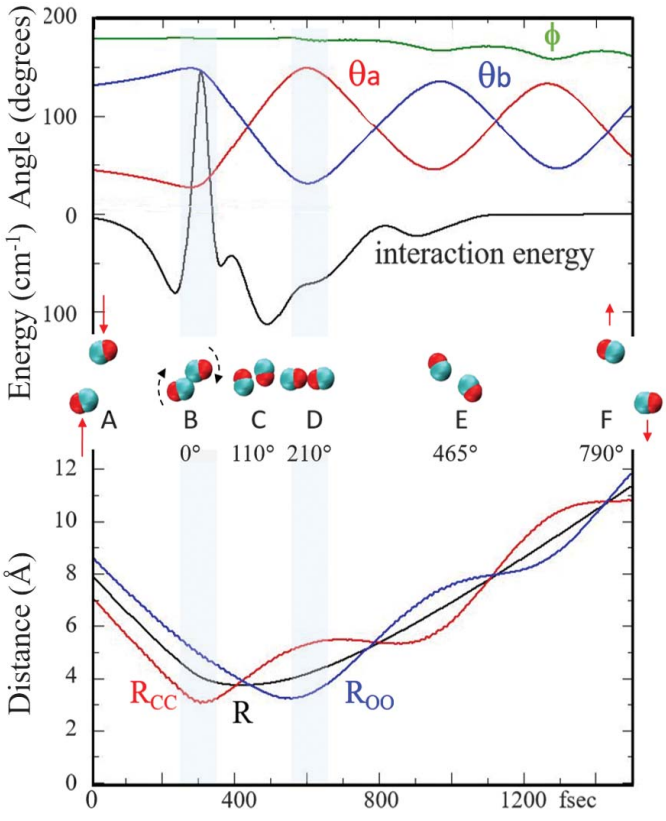


Fig. 2. Trajectory details for two colliding CO(j = 0) molecules, ending with both in j = 15. R_{CC} and R_{OO} are internuclear distances, and R is the center-of-mass distance between the two diatoms. The dihedral angle ϕ is 180° when the molecular axes are coplanar and antiparallel, and θ_a and θ_b are the angles between the ***R*** vector and the bond directions of molecules a (traveling up in the figure) and b (traveling down), respectively. In the top panel, energy (cm^{-1}) and angle (degrees) have the same vertical axis spacing. The collision geometry is shown in the middle of the figure for positions A to F, corresponding to times of 0, 275, 450, 600, 1000, and 1500 fs, respectively. Molecular rotation begins at position B and reaches the angle values shown under each position. Collision energy is 1460 cm^{-1}. See text and SM (movie S1) for more details.

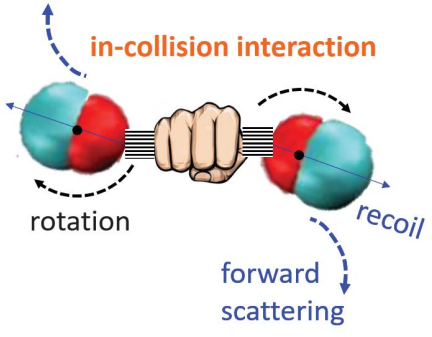


Fig. 3. Cartoon of scattering-angle reorientation. With the collision geometry at position A in Fig. 2, or its mirror image, both CO molecules rotate in the same sense, and when the intermolecular interaction during the collision couples with this rotational motion, the recoil direction will rotate toward the forward direction.

than the initial b_i = 3.2 Å, which was a simple result of conservation of energy and angular momentum when starting with two j = 0 molecules. As ***R*** expanded by 0.8 Å during the collision, it also rotated, causing the final recoil direction to turn from sideways toward forward, as illustrated in Fig. 3.

At a lower collision energy (500 cm^{-1}), QCT predicted the same FSSE process, peaking at (j', j'') = (8, 8), with forward scattering being most probable at b_i = 3.0 Å. FSSE occurred for both collision energies with roughly 60% conversion of translational to rotational energy. Fully quantum CC calculations at 500 cm^{-1} confirmed an amount of rotational excitation for near-symmetric excitation similar to that at 1460-cm^{-1} collision energy, peaking even closer to forward scattering than the QCT results (fig. S8). CC calculations also revealed a propensity for collisions resulting in even values of final j (fig. S7), reflecting the nearly homonuclear diatomic nature of CO.

Pair-correlated excitation cross sections observed previously for NO-O_2 collisions (*12*) also deviate from EGL behavior in that strong excitation in both the NO and O_2 molecules is preferred over excitation in one collision partner only, even though the total amount of energy transfer is higher. Opacity function analysis showed that the inelastic channels for high rotational-state excitation were governed by short-range head-on collisions and that only the long-range contributions follow "normal" behavior (*12*).

CO-CO is not a particularly unusual collision system. We suggest that FSSE-like behavior is also possible in mixed-molecule scattering systems that match the criteria found here: (i) large ends colliding first and (ii) an abrupt energy transfer coupled with a swing around toward the small-end configuration before departure. Potential energy surfaces with minor topological features compared with the collision energy could allow the pathway observed here for CO-CO scattering. As molecule-molecule scattering is explored in higher detail with new methods, more unusual molecular dance moves could be expected.

REFERENCES AND NOTES

1. R. J. Marsh, A. J. McCaffery, *Chem. Phys. Lett.* **335**, 134–139 (2001).
2. A. T. Eppink, D. H. Parker, *Rev. Sci. Instrum.* **68**, 3477–3484 (1997).
3. J. C. Polanyi, K. B. Woodall, *J. Chem. Phys.* **56**, 1563–1572 (1972).
4. Y. T. Lee, J. D. McDonald, P. R. LeBreton, D. R. Herschbach, *Rev. Sci. Instrum.* **40**, 1402–1408 (1969).
5. B. Yan *et al.*, *Rev. Sci. Instrum.* **84**, 023102 (2013).
6. S. Y. T. van de Meerakker, H. L. Bethlem, N. Vanhaecke, G. Meijer, *Chem. Rev.* **112**, 4828–4878 (2012).
7. G. Sarma, S. Marinakis, J. J. ter Meulen, D. H. Parker, K. G. McKendrick, *Nat. Chem.* **4**, 985–989 (2012).
8. S. N. Vogels *et al.*, *Phys. Rev. Lett.* **113**, 263202 (2014).
9. A. von Zastrow *et al.*, *Nat. Chem.* **6**, 216–221 (2014).
10. K. B. Sheppard *et al.*, *Astrophys. J.* **850**, L32 (2017).
11. G. W. M. Vissers, A. Hesselmann, G. Jansen, P. E. S. Wormer, A. van der Avoird, *J. Chem. Phys.* **122**, 54306 (2005).
12. Z. Gao *et al.*, *Nat. Chem.* **10**, 469–473 (2018).
13. Z. Gao, J. Loreau, A. van der Avoird, S. Y. T. van de Meerakker, *Phys. Chem. Chem. Phys.* **21**, 14033–14041 (2019).
14. G. W. M. Vissers, P. E. S. Wormer, A. van der Avoird, *Phys. Chem. Chem. Phys.* **5**, 4767–4771 (2003).
15. L. A. Surin *et al.*, *J. Phys. Chem. A* **111**, 12238–12247 (2007).
16. L. Song *et al.*, *J. Phys. Chem. A* **119**, 12526–12537 (2015).
17. Z. F. Sun *et al.*, *Phys. Chem. Chem. Phys.* **21**, 9200–9211 (2019).
18. S. A. Ndengué, R. Dawes, F. Gatti, *J. Phys. Chem. A* **119**, 7712–7723 (2015).
19. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **14**, 33–38 (1996).
20. Z. F. Sun, M. C. van Hemert, J. Loreau, A. van der Avoird, A. G. Suits, D. H. Parker, Molecular square dancing in CO-CO collisions dataset, Dryad (2020); https://doi.org/10.5061/dryad.80gb5mkp0.

ACKNOWLEDGMENTS

We thank B. v. d. Meerakker, G. C. Groenenboom, and R. J. A. Scheidsbach for discussion. R. J. A. Scheidsbach, C. K. Bishwakarma, G. Sarma, and L. Gerritsen developed the experimental methods used in this work. **Funding:** The experimental part of this work was funded by the Dutch National Science Foundation (NWO) under grant number TOP 715.013.002 and the Dutch Astrochemistry Network DANII under grant number 648.000.024. A.G.S. acknowledges support from the U.S. National Science Foundation (NSF) under award number CHE-1955239. **Author contributions:** The experimental work and data analysis were carried out by Z.F.S., A.G.S., and D.H.P. Theoretical calculations were done by M.C.v.H., J.L., and A.v.d.A. All authors contributed to interpreting the results and writing the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data are available in the main text or in the supplementary materials and at Dryad (*20*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/307/suppl/DC1
Experimental Methods
Theoretical Methods
Supplementary Text
Figs. S1 to S8
Table S1
References (*21*–*31*)
Data S1
Movies S1 and S2

19 March 2020; accepted 22 May 2020
10.1126/science.aan2729

ORGANIC SEPARATIONS

N-Aryl–linked spirocyclic polymers for membrane separations of complex hydrocarbon mixtures

Kirstie A. Thompson[1]*, Ronita Mathias[2]*, Daeok Kim[3]†, Jihoon Kim[3], Neel Rangnekar[4], J. R. Johnson[4], Scott J. Hoy[5], Irene Bechis[6], Andrew Tarzia[6], Kim E. Jelfs[6], Benjamin A. McCool[4], Andrew G. Livingston[3,7], Ryan P. Lively[2]‡, M. G. Finn[1]‡

The fractionation of crude-oil mixtures through distillation is a large-scale, energy-intensive process. Membrane materials can avoid phase changes in such mixtures and thereby reduce the energy intensity of these thermal separations. With this application in mind, we created spirocyclic polymers with *N*-aryl bonds that demonstrated noninterconnected microporosity in the absence of ladder linkages. The resulting glassy polymer membranes demonstrated nonthermal membrane fractionation of light crude oil through a combination of class- and size-based "sorting" of molecules. We observed an enrichment of molecules lighter than 170 daltons corresponding to a carbon number of 12 or a boiling point less than 200°C in the permeate. Such scalable, selective membranes offer potential for the hybridization of energy-efficient technology with conventional processes such as distillation.

Each day, ~100 million barrels of crude oil are fractionated in refineries around the globe, primarily by thermal distillation. Though distillation has the potential for high recoveries of molecules in discrete boiling-point ranges, it is energy intensive, requiring more than 1100 terawatt-hours per year—nearly 1% of global energy use (*1*). Distillation energy and carbon efficiency could be improved if it were paired with low-energy membrane-based separations that fractionate complex mixtures containing thousands of compounds into smaller groups of molecules. Membrane separations based on molecular differences in size, shape, and membrane-penetrant interactions have the potential for a 10-fold increase in energy efficiency over thermal processes (*2*). Seawater reverse-osmosis systems that currently process greater liquid flows than the largest refineries highlight the potential for membrane technology to match the scales required by the hydrocarbon-processing industry. Whereas membrane technologies are established for certain aqueous and gas separations, they have not been developed for the fractionation of organic solvent mixtures, because of the relative scarcity of effective materials that are easily processable, inexpensive, and stable (*3*). Polymeric membrane materials provide the best combination of these three metrics but to date have not been sufficiently effective for separations of small liquid organic molecules.

Polymers of intrinsic microporosity (PIMs), with their contorted ladder-like structures, have high guest permeability and have therefore been widely studied for gas separations and, to some extent, for organic solvent separations (*4*–*6*). For the prototypical polymer **PIM-1**, there is a substantial disparity between the apparent pore-size distributions in the "dry" and organic solvent "wet" states. Cryogenic physisorption experiments in solid **PIM-1** suggest an average pore size between 3.5 and 5.0 Å (*7*). By contrast, molecular weight cutoff measurements of styrene oligomers in toluene by **PIM-1** membranes were reported (and were confirmed in this study) to be greater than 600 g mol^{-1} (*8*), which corresponds to a pore size of 14 Å in a rigid membrane material such as a ceramic (*9*). This discrepancy derives from motion-enabled zones of activation in the semirigid polymer network that undergo swelling and plasticization in organic solvents (*5*). A PIM-like polymer that does not exhibit these phenomena should enable separations of small molecules in the organic solvent reverse osmosis (<200 g mol^{-1}) and "tight" organic solvent nanofiltration (200 to 300 g mol^{-1}) regimes. Although postpolymerization modifications can tune the properties of **PIM-1**–like materials (*10*), structures capable of performing difficult separations without postpolymerization or postfabrication modification would provide more chances to create solutions to these impactful, large-scale challenges.

We took note of previous gas separation work by McKeown and colleagues in which a spirobifluorene monomer was found to frustrate chain packing and rigidify ladder-type polymer structures beyond equivalent materials made by using a spirobisindane monomer (*6*), resulting in higher polymer free volume and permeability relative to **PIM-1**. Moreover, swelling and plasticization were reduced, resulting in no apparent loss of separation selectivity compared with **PIM-1**. We therefore employed a similar spirobifluorene building block but realized that this modification alone would likely be unable to address the challenge posed by swelling in organic solvent separation compared with gas-phase separations. Thus, we sought to enhance packing efficiency in certain segments of the backbone by using a modestly flexible linkage, choosing the aryl–*N*-aryl connection.

This design allowed the use of catalytic C–N bond formation as the polymerization process, rather than the nucleophilic aromatic substitution reaction that gives rise to the dibenzodioxin linkage of **PIM-1**. The palladium-catalyzed Buchwald-Hartwig amination reaction is a robust method for the coupling of aryl halides and amines that has been used extensively in medicinal, materials, and natural products chemistry (*11*); its application to polymer synthesis has been more limited. We chose a single spirobifluorene dibromide (**1**) (Fig. 1 and fig. S1) as the A–A monomer and a range of commercially available aromatic diamines as the B–B component (compounds **I** to **IV** in Fig. 1). Polymerization using the XantPhos Pd G4 palladacycle catalyst (*12*) was optimized (table S1) with respect to temperature and concentration, providing the spirobifluorene aryl diamine (SBAD) series (Fig. 1).

All SBAD polymers were readily soluble in volatile solvents (tetrahydrofuran, chloroform, CH_2Cl_2) (fig. S2) and were easily cast into films. Gel permeation chromatography analysis revealed a component of high apparent molecular weight, thought to correspond to a small amount of branched material derived from a second N–H insertion event at one or more amine sites along the polymer chain (fig. S3). Even a small amount of branching at secondary amine centers can substantially inflate observed weight-average molecular weights and dispersities (*13*). The SBAD polymers demonstrated high thermal stabilities, with no observable glass transition temperatures (figs. S4 and S5), as has often been reported for PIM systems (*14*). These materials also exhibited low uptakes of N_2 at 77 K (fig. S6A) but had only moderately lower CO_2 sorption than **PIM-1** at 273 K (Fig. 1B), suggesting the existence of isolated microporous free-volume elements.

To further probe the pore structure of these materials, computational models for the polymers in the dry state were generated by using the polymerization algorithm Polymatic (*15*). Three models were generated for each system to sample different structural arrangements and ensure the models were representative of the bulk material. The properties of the models were then characterized (figs. S7 and S8

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA. [2]School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. [3]Department of Chemical Engineering, Imperial College London, London SW7 2AZ, UK. [4]Corporate Strategic Research, ExxonMobil Research and Engineering, Annandale, NJ 08801, USA. [5]Analytical Sciences Laboratory, ExxonMobil Research and Engineering, Annandale, NJ 08801, USA. [6]Department of Chemistry, Imperial College London, London W12 0BZ, UK. [7]School of Engineering and Materials Science, Queen Mary University of London, London E1 4NS, UK.
*These authors contributed equally to this work.
†Present address: Environment and Safety Research Center, Samsung Electronics, Hwaseong-si, Gyeonggi-do, South Korea.
‡Corresponding author. Email: ryan.lively@chbe.gatech.edu (R.P.L.); mgfinn@gatech.edu (M.G.F.)

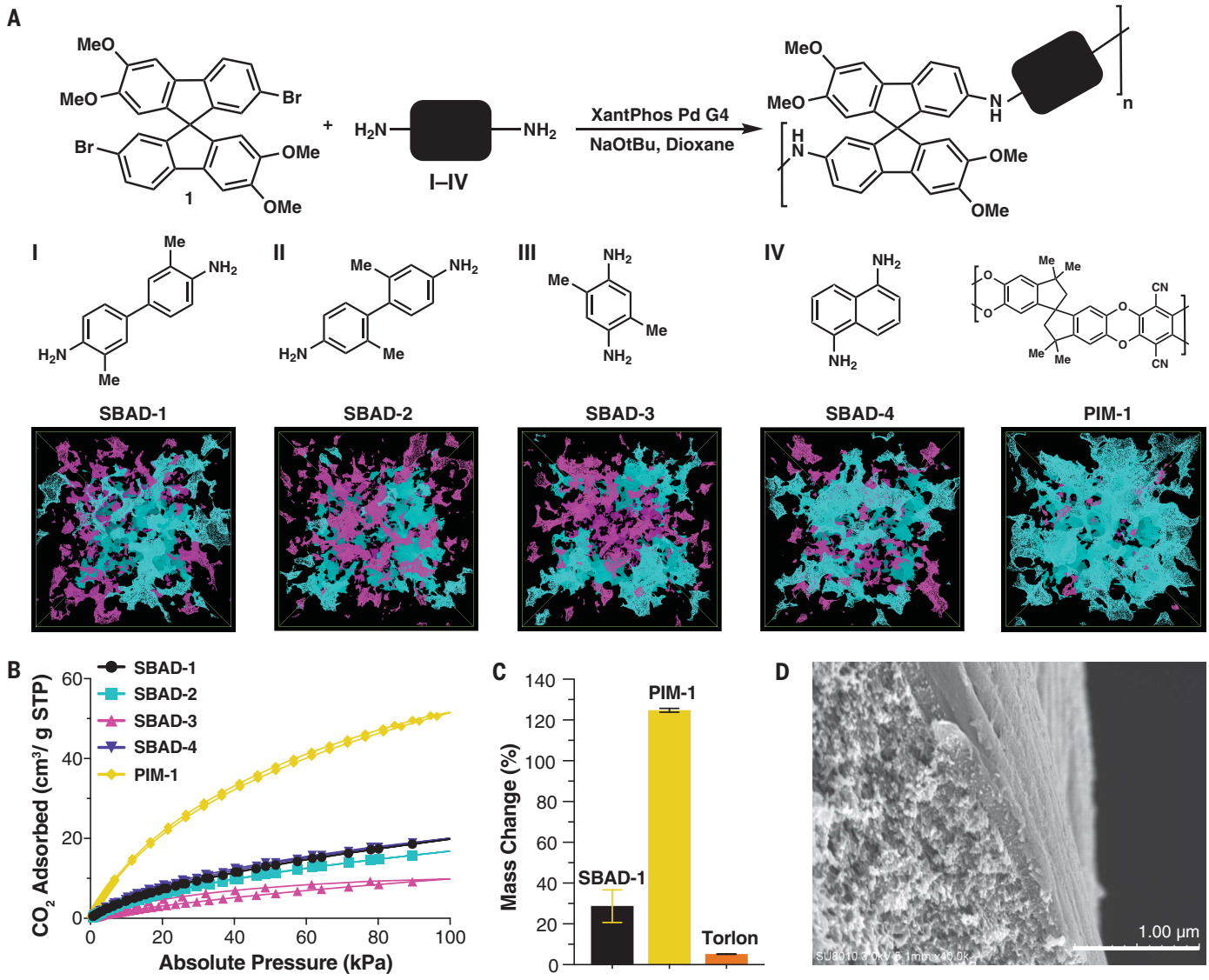


Fig. 1. Modular reaction scheme, computational simulation of porosity, powder physisorption, liquid sorption, and spectroscopic characterization. (**A**) Polymerization of 7,7′-dibromo-2,2′,3,3′-tetramethoxy-9,9′-spirobifluorene, **1**, with diamines I to IV giving the SBAD class of materials: **SBAD-1** to **SBAD-4**. Below the diamine structures are representative results of computational modeling of the pore surfaces for the SBAD series compared with **PIM-1** applying a 2.2 Å probe diameter, the largest probe that can pass through the SBAD materials; teal indicates accessible (interconnected) surface area, and magenta indicates nonaccessible surface area. (**B**) CO_2 uptake (at 273 K) by SBAD polymers compared with that by **PIM-1**. (**C**) Uptake of liquid toluene by dense films of **SBAD-1**, **PIM-1**, and **Torlon** (*22*), (N = 2) 22°C, 1 atm. Data are the mean of 2 films ± the range. (**D**) SEM image of **SBAD-1** thin-film composite showing an approximate barrier-layer thickness of 200 nm. tBu, tert-butyl; Me, methyl; STP, standard temperature and pressure.

and table S2) (*16*). **PIM-1** and the SBAD series models showed similar predicted densities (table S3) but notable differences in polymer voids, as seen in Fig. 1A, fig. S9, and table S3. The diameter of the largest probe that can pass through an interconnected pore network in the **PIM-1** models is 3.2 ± 0.3 Å, whereas the equivalent value for the SBAD series is 2.3 ± 0.1 Å. Only a small amount of flexible breathing motions or polymer swelling would be required for nitrogen [diameter, 3.64 Å] to diffuse through **PIM-1**, but much larger dilations would be required with the SBAD series (figs. S10 and S11). This is consistent with the observed nitrogen physisorption isotherms (fig. S6) showing high uptakes for **PIM-1** compared with the SBAD series at low relative pressures, which is the range that is typically used to calculate micropore distributions. By defining "interconnected pores" as those that are accessible by N_2 at 77 K within a reasonable time frame (2 days), it is apparent that "interconnected microporosity" is a characteristic feature of **PIM-1** that is suppressed in the SBAD polymers. This effect is likely due to the absence of the dibenzodioxin linkage and ladder-like morphology of **PIM-1** (figs. S12 and S13), replaced in the SBAD series with more-flexible single-bond C–N linkages, allowing for the formation of π–π stacking interactions between and within chains (fig. S14) and more-efficient chain packing.

The computational work also suggests the existence of narrow distributions of ultramicropores between 2 and 8 Å for **SBAD-1** and **-3** and 2 and 10 Å for **SBAD-2** and **-4**, whereas **PIM-1** has an apparently broader distribution between 2 and 12 Å (figs. S15 and S16). Even tighter distributions of ultramicropores are seen in traditional glassy polymers (*17*), but the pore sizes are often too small for meaningful organic solvent separations. Thus, the

presence of intermediate noninterconnected pores exhibited by the SBAD materials could offer penetrant-penetrant differentiation of small-molecule organic solvents at reasonable fluxes if diffusion from one pore to the next were reasonably fast. In the wet state, separating small organic molecules requires that the membrane material exhibit resistance to solvent-induced swelling (e.g., Torlon, a highly solvent-stable polyamide-imide) (Fig. 1C). Solvent-induced dilation experiments showed that **SBAD-1** experienced substantially less mass change than **PIM-1** at unit activity of toluene (Fig. 1C), suggesting that the SBAD materials retain their narrow distribution of ultramicropores even when exposed to solvents.

Mechanically robust thin-film composite membranes were fabricated on cross-linked polyetherimide supports with no observable interlayer delamination (Fig. 1D and fig. S17). The molecular separation performance of these materials was surveyed with binary mixture separations of 1,3,5-triisopropylbenzene (TIPB) in toluene (Fig. 2A). TIPB serves as a surrogate solute molecule for organic solvent separations, with a molecular weight (204.35 Da) at the lower end of the typical nanofiltration range (200 to 1000 g mol^{-1}) in which PIMs have often been tested (*9*). Membrane performance was evaluated through cross-flow filtration (fig. S18) after more than 48 hours of continuous permeation of the mixture.

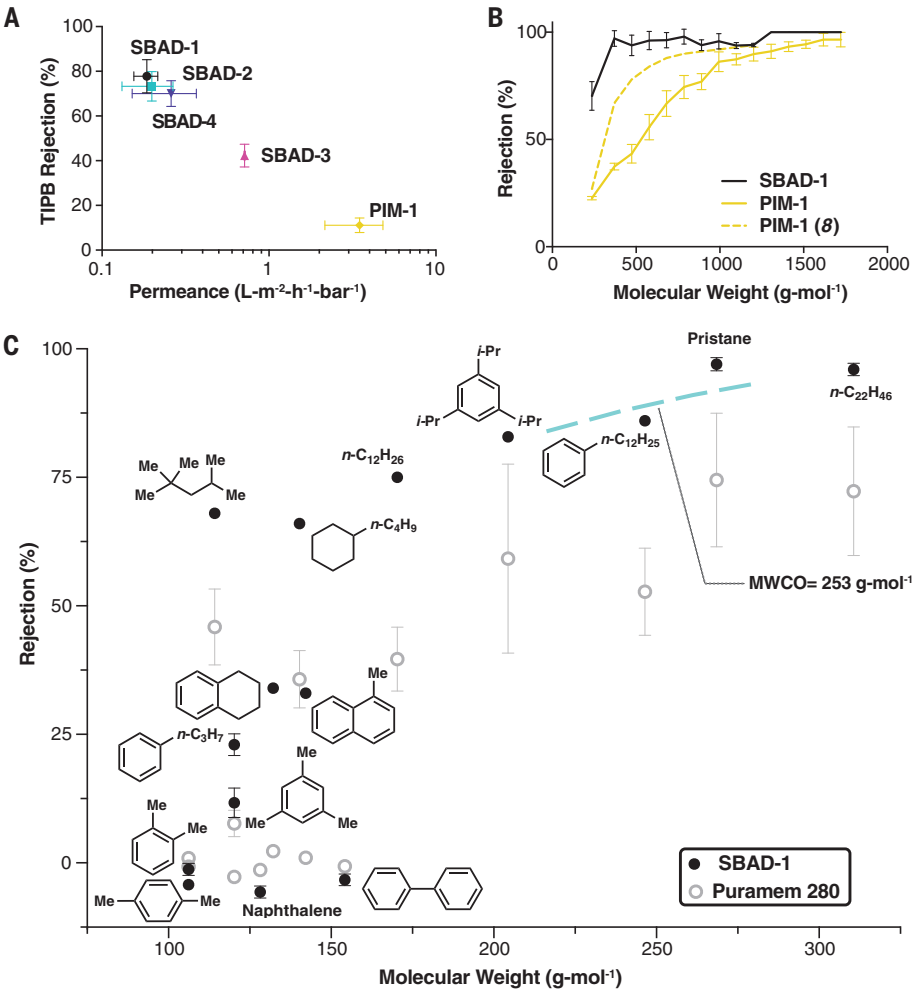


Fig. 2. Dilute hydrocarbon separation performance using polymer thin-film composites. (**A**) TIPB rejection versus toluene permeance for a 1/99 mol % feed of TIPB/toluene (N = 3), 22°C, 15 bar. (**B**) Molecular weight cutoff curves of **SBAD-1** (N = 3) at 30 bar (threshold = 335 g mol^{-1}), **PIM-1** at 5 bar (1220 g mol^{-1}), and **PIM-1** data reported at 30 bar (786 g mol^{-1}) (*8*) for polystyrene standards in toluene at 22°C. The high flux in **PIM-1** in our work limited the applied pressure difference to 5 bar to maintain <5% stage cuts. Data reported by Cook *et al.* at the same conditions as **SBAD-1** was shown for direct comparison. (**C**) Hydrocarbon molecular weight cutoff curve (threshold = 253 g mol^{-1}, fig. S21) for **SBAD-1** (N = 2) and Puramem 280 (N = 3) from solute mixtures listed in table S4. Data are the mean of N membranes ± standard deviation. *i*-Pr, isopropyl; MWCO, molecular weight cutoff.

As shown in Fig. 2A, all of our polymers exhibited significantly greater rejections at steady state than **PIM-1**, with permeances in the range of 0.1 to 0.7 liters m^{-2} $hour^{-1}$ bar^{-1}. **SBAD-1** showed the highest TIPB rejection of 80% throughout the duration of the experiment, and **SBAD-3** was the least effective, despite having similar predicted micropore size distributions as the other materials. The most relevant structural difference appeared to be the larger biphenyl bridge structure used in **SBAD-1**, **2**, and **4** versus the smaller phenyl bridge in **SBAD-3**. The improvement in TIPB rejection with the larger bridge suggests that aromatic chain-chain interactions, more available with biphenyl linkages (fig. S14), are advantageous. The rejection properties of **SBAD-2** and **SBAD-4** increased over 48 hours, reaching steady-state values only slightly lower than those of **SBAD-1** (fig. S19). Such time-dependent behavior is consistent with solvent-induced polymer relaxation toward a less swollen state (*18*). Only a handful of commercial organic solvent nanofiltration membranes and other polymeric membranes with complex fabrication procedures are capable of high rejections of ~200 g mol^{-1} solutes and have toluene permeances in the same range (*9*). As **SBAD-1** demonstrated the highest rejections, it was chosen for scale-up and further evaluation with more-complex mixtures.

SBAD-1 and **PIM-1** were compared in a standardized measurement of molecular weight cutoffs by using oligostyrene markers in toluene. **SBAD-1** was significantly better than **PIM-1** in the separation of small solutes (molecular-weight cutoff of 335 versus 1220 g mol^{-1}, respectively), albeit at a much lower permeance (0.2 versus 3.49 liters m^{-2} $hour^{-1}$ bar^{-1}). The molecular weight cutoff of **PIM-1** increased over the course of 3 days, whereas that of **SBAD-1** remained constant (fig. S20).

The separation performance of **SBAD-1** using a group of nonpolar hydrocarbon molecules in the organic solvent reverse-osmosis and nanofiltration ranges that are representative of hydrocarbons found in light crude oils is shown in Fig. 2C. A hydrocarbon molecular-weight cutoff of ~253 g mol^{-1} (fig. S21) was calculated. Below 150 g mol^{-1}, greater aromaticity (ratio of aromatic carbons to aliphatic carbons) of the solute was found to diminish rejection, perhaps because of greater sorption interactions with the aromatic rings present in the polymer backbone. Dual dependence on both size and aromatic interaction was supported by a linear correlation to a combination of aromaticity (sorption) and molecular weight (diffusion) factors, but to neither factor alone (fig. S22). When compared with Puramem 280, a commercial polyimide membrane, **SBAD-1** showed higher rejections, although this was accompanied by lower permeances, as shown

in table S4 at 22°C. As the thin-film thicknesses were not controlled in this work, the differences in permeance cannot be directly compared, because the **SBAD-1** films are considerably thicker.

The naphtha and kerosene fractions of crude oil make up the primary component of gasoline and jet fuel, respectively. These two fractions typically contain a range of alkanes, alkyl aromatics, and polyaromatics with boiling points from 90° to 300°C, representing a good candidate for membrane-based hydrocarbon fractionation. A thin-film composite membrane sheet of **SBAD-1** was produced through roll-to-roll coating of the polymer onto a prefabricated Ultem support (fig. S23, A and B). Membrane coupons from the sheet (diameter, 47 mm) were subjected to a complex feed of hydrocarbons that serve as a model naphtha-kerosene fraction. Several large spiral-wound modules (fig. S23C) (containing 1.8 m by 0.2 m of membrane) were also fabricated from the sheet to highlight the scalability of these materials.

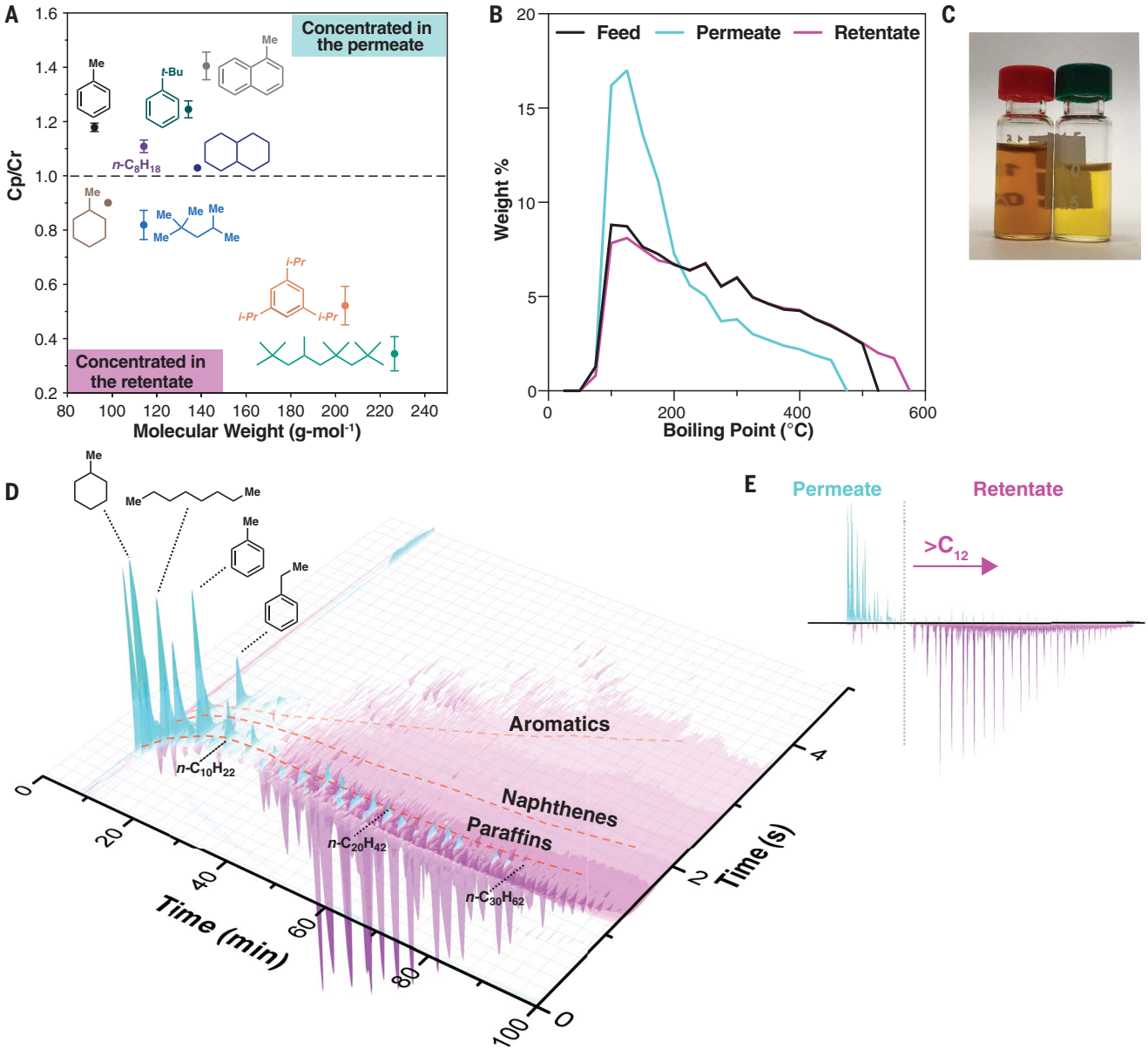


Fig. 3. Synthetic and real crude-oil separation performance using SBAD-1 thin-film composites. (**A**) Ratio of concentrations in permeate (Cp) versus retentate (Cr) of components in a complex, model crude oil fractionated by a thin-film composite membrane coupon of **SBAD-1** (N = 3); 40 bar, 22°C. Molecules above the reference line were more concentrated in the permeate compared with the feed, whereas the molecules below were more concentrated in the retentate. Data are the mean of three membranes ± standard deviation. (**B**) Boiling-point distribution of feed, permeate, and retentate from **SBAD-1** membrane fractionation of shale-based crude oil. (**C**) Picture of feed (left) and permeate (right) from **SBAD-1** membrane fractionation of real shale-based light whole crude oil. (**D**) GCxGC-Flame-ionized detection (FID) analysis of membrane fractionation of shale-based crude oil at 130°C and 55 bar showing the feed chromatogram subtracted from the permeate chromatogram. Teal elution peaks are concentrated in the permeate, whereas the magenta elution peaks are concentrated in the membrane retentate. General hydrocarbon classifications are highlighted with dashed red lines to guide the eye. (**E**) Side view of GCxGC-FID [shown in (D)] showing the carbon number partition of real crude oil obtained by **SBAD-1**.

Fig. 3A shows the membrane performance (feed composition shown in table S5) as the ratio of permeate to retentate concentrations plotted against the molecular weights of the components. A permeance of 0.022 ± 0.013 liters m^{-2} $hour^{-1}$ bar^{-1} for the coupon and 0.076 ± 0.003 liters m^{-2} $hour^{-1}$ bar^{-1} for the module (fig. S24) was obtained. This permeance is lower than that in the dilute-mixture experiment (Fig. 2) because of the high osmotic pressure of the complex, multicomponent mixture. However, good separation of molecules of different molecular weights, independent of molecular class, as well as for molecules from different classes without a molecular-weight advantage was obtained. Additionally, exposure to temperatures up to 75°C did not negatively affect membrane performance (fig. S25). Some notable separation factors (based on coupon data) are shown in table S6, illustrating the potential of **SBAD-1** for membrane-based concentration of solutes of both low molecular weights and high aromaticity. The values highlighted are comparable to separation factors obtained for liquid xylene isomers by using carbon molecular sieve membranes (*19*). Aromatic-rich permeates streams (observed in this work) are useful for the production of plastics and provide a rich source of industrial organic solvents. Fractionation of crude oil into groups of smaller and larger molecules also has direct value in generating products classified by the number of carbons present such as gasoline (C5-C12 alkanes and cycloalkanes) and jet fuel (C10-C18 alkanes and aromatics). The retentate stream shown in Fig. 3A contains both of these.

SBAD-1 was also challenged with a light shale crude oil to illustrate the potential of membrane materials as an alternative for the fractionation of crude oil. Figure 3B shows the boiling-point distribution of molecules in the feed, permeate, and retentate. Approximately 38% of the mixture boiled below 200°C in the feed tested in this work. After passing through a thin-film composite of **SBAD-1**, the permeate was found to be appreciably enriched in the lighter molecules present in the feed, such that more than 60% of the permeate boiled below 200°C. Figure 3C shows a picture of the crude oil and the lighter-colored permeate indicating removal of color bodies that are typically high in molecular weight. Figure 3D visualizes the differences in GCxGC-FID chromatograms of the feed and permeate, with the three axes representing boiling point (time scale in minutes), polarity (time scale in seconds), and response factor (*z* axis). The GCxGC-FID data indicates enrichment of molecules lighter than carbon numbers of 12 or around molecular weight of 170 Da (Fig. 3E). The membrane permeance is low (0.016 liters m^{-2} $hour^{-1}$ bar^{-1}), as the thin-film fabrication process has not yet been optimized to create more-productive composites. Overall, the **SBAD-1** membrane-based separation demonstrated selectivity for low–molecular-weight saturated molecules in the naphtha and early kerosene boiling range with slight permeation of higher–molecular-weight linear paraffins.

This result is similar to a boiling-point separation of a complex hydrocarbon mixture, but without expending the energy associated with vaporization. A clear "interclass" separation of the crude was observed with a preferential order of permeation: naphthenes, linear paraffins, aromatics, and branched paraffins (fig. S26 and table S7). Moreover, intraclass size-based separation also occurred for each of these individual classes of hydrocarbons, such that the permeate was enriched in the low boiling (or "light") compounds in the crude oil (fig. S27). Although we did not specifically explore fouling effects from the real crude feed, the permeance and separation factors changed minimally over a period of 2 months during synthetic crude testing at temperatures ranging from 25° to 75°C (fig. S25).

An interesting distinction emerges when comparing traditional binary separation experiments (i.e., toluene and TIPB) and dilute multicomponent hydrocarbon separations with the complex mixture fractionation experiments (i.e., fractionation of lab-made and real crude oil). Whereas molecular weight cutoff and rejection are important figures of merit for dilute separations such as water purification and nanofiltration, the utility of this metric is less clear in the case of complex mixture fractionation, in which there is no discrete solvent or solute: Every molecule in crude oil is a solvent. In this case, each molecule in the mixture experiences highly nonideal coupling effects (e.g., diffusion and sorption) during permeation. Moreover, the osmotic effects in this crude-oil system are complex and certainly contribute toward reducing the fluxes relative to the diluted mixtures shown in Fig. 2. The membrane-based crude-oil separation occurs through a combination of diffusion-based selection processes as well as solvent-membrane (and solvent-solvent) interactions, which further reinforces the importance of testing real mixtures. Beyond testing industrial feeds, it will be important to establish how these materials fit into an energy-efficient membrane-fractionation process (*20*). To produce distillation-like products through membrane fractionation, a cascade of membrane materials such as **SBAD-1**, capable of fractionating feeds across a specific range of hydrocarbon sizes, will be necessary (fig. S28).

The results described here highlight that the dibenzodioxin connection between spirocyclic monomers that characterizes **PIM-1**–like materials is not necessary to enable meaningful small-molecule membrane-based separations. We speculate that the advantageous properties of SBAD-based membranes derive from the nature of the aromatic nitrogen-carbon linkage, which has long been known to demonstrate hindered rotation (*21*). This connection between spirocyclic and aromatic building blocks gives rise to bulk rigidity and low swelling, isolated and large micropores, and enough segmental mobility to allow migration of molecules between micropores at rates influenced by both molecular size and molecular type.

Together, these properties make the SBAD family promising for the separation of small-sized molecules spanning the organic solvent reverse osmosis and the lower end of the organic solvent nanofiltration range (100 to 350 g mol^{-1}) in complex mixtures such as crude oil. Most other membranes that operate in similar regimes require postpolymerization or postfabrication modifications (e.g., cross-linking) that can be costly and difficult to scale up. By contrast, the polymers used in this study can be tailored with simple changes in composition and performance. The proposed dependence of SBAD performance on both pore structure and dynamic motion enables complex mixture fractionation that may be scalable and translatable to real crude-oil fractions.

REFERENCES AND NOTES

1. International Energy Agency, "World Balance IEA Sankey Diagram" (IEA, 2017).
2. National Academies of Sciences Engineering and Medicine, *A Research Agenda for Transforming Separation Science* (The National Academies Press, 2019).
3. R. P. Lively, D. S. Sholl, *Nat. Mater.* **16**, 276–279 (2017).
4. N. B. McKeown, P. M. Budd, *Chem. Soc. Rev.* **35**, 675–683 (2006).
5. W. J. Koros, C. Zhang, *Nat. Mater.* **16**, 289–297 (2017).
6. C. G. Bezzu *et al.*, *Adv. Mater.* **24**, 5930–5933 (2012).
7. R. Swaidan, B. S. Ghanem, E. Litwiller, I. Pinnau, *J. Membr. Sci.* **457**, 95–102 (2014).
8. M. Cook, P. R. J. Gaffney, L. G. Peeva, A. G. Livingston, *J. Membr. Sci.* **558**, 52–63 (2018).
9. P. Marchetti, M. F. Jimenez Solomon, G. Szekely, A. G. Livingston, *Chem. Rev.* **114**, 10735–10806 (2014).
10. D. Fritsch, P. Merten, K. Heinrich, M. Lazar, M. Priske, *J. Membr. Sci.* **401-402**, 222–231 (2012).
11. P. Ruiz-Castillo, S. L. Buchwald, *Chem. Rev.* **116**, 12564–12649 (2016).
12. N. C. Bruno, N. Niljianskul, S. L. Buchwald, *J. Org. Chem.* **79**, 4161–4166 (2014).
13. A. B. Foster *et al.*, *Macromolecules* **53**, 569–583 (2020).
14. M. Heuchel, D. Fritsch, P. M. Budd, N. B. McKeown, D. Hofmann, *J. Membr. Sci.* **318**, 84–99 (2008).
15. L. J. Abbott, K. E. Hart, C. M. Colina, *Theor. Chem. Acc.* **132**, 1334 (2013).
16. T. F. Willems, C. Rycroft, M. Kazi, J. C. Meza, M. Haranczyk, *Microporous Mesoporous Mater.* **149**, 134–141 (2012).
17. C. Nagel *et al.*, *Macromolecules* **33**, 2242–2248 (2000).
18. J. L. Duda, I. Hadi Romdhane, R. P. Danner, *J. Non-Cryst. Solids* **172-174**, 715–720 (1994).
19. D. Y. Koh, B. A. McCool, H. W. Deckman, R. P. Lively, *Science* **353**, 804–807 (2016).
20. B. A. McCool, D. A. Bhandari, Y. V. Joshi, U.S. Patent Application no. 20190367820A1 (2019).
21. A. Mannschreck, H. Muensch, *Tetrahedron Lett.* **9**, 3227–3230 (1968).
22. H. Y. Jang *et al.*, *AIChE J.* **65**, e16757 (2019).

ACKNOWLEDGMENTS

We thank N. C. Bruno (Georgia Institute of Technology) for Pd-catalyzed cross-coupling guidance, G. Zhu (Georgia Institute of Technology) for assistance with nitrogen and carbon dioxide

physisorption experiments, Y. Ma (Georgia Institute of Technology) for assistance with the dense film fabrication, D. A. Bhandari (ExxonMobil Research and Engineering) for suggesting the synthetic kerosene mixture composition, and M. L. Jue (Georgia Institute of Technology) for the synthesis of PIM-1. **Funding:** This work was supported by ExxonMobil Research and Engineering. K.A.T. acknowledges support from the Department of Education Graduate Assistance in Areas of National Need (GAANN) program at Georgia Institute of Technology (award no. P200A180075). K.E.J. acknowledges a Royal Society University Research Fellowship. A.T. acknowledges a Royal Society Enhancement Grant (2018) and the European Research Council (ERC) under FP7 (CoMMaD, ERC grant no. 758370). **Author contributions:** Monomer and polymer synthesis, gel permeation chromatography, and elemental characterization were conducted by K.A.T. Polymer model generation and analyses were carried out by I.B., A.T., and K.E.J. Matrimid support fabrication, blade coating, spectroscopic characterization, coupon separation performance, thermal analysis of powders, and physisorption model fitting were done by R.M. Ultem support fabrication and roll-to-roll coating were performed by D.K. Spiral-wound–module fabrication and spiral-wound–module performance analysis were conducted by D.K. and J.K. Whole crude separation was conducted by N.R. and J.R.J. GCxGC analysis was conducted by S.J.H. Complex mixture separation results were confirmed by R.M., D.K., J.K., N.R., and J.R.J. at separate facilities. Temperature resistance of membranes was investigated by N.R., and J.R.J., B.A.M., A.G.L., R.P.L., and M.G.F. conceived the research. **Competing interests:** A.G.L. serves as a non-Executive Director of Evonik Membrane Extraction Technology UK Limited. **Data and materials availability:** Detailed descriptions of materials and methods, monomer synthesis, polymer synthesis, polymer model generation and analysis, powder characterization, membrane fabrication, spectroscopic characterization, and separation performance are in the supplementary materials. K.A.T., R.M., B.A.M., R.P.L., and M.G.F. are inventors on patent applications (US20190275469, US20190390126, and 62/640,253) submitted by Georgia Tech Research Corporation and ExxonMobil Research and Engineering that covers SBAD polymers, membrane fabrications, and crude-oil fractionation applications. All data needed to evaluate the conclusions in the paper present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/310/suppl/DC1
Materials and Methods
Figs. S1 to S28
Tables S1 to S7
NMR Spectra
References (*23*–*39*)

30 January 2020; accepted 22 May 2020
10.1126/science.aba9806

CORONAVIRUS

Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic

Rachel E. Baker[1,2]*, Wenchang Yang[3], Gabriel A. Vecchi[1,3], C. Jessica E. Metcalf[2,4], Bryan T. Grenfell[2,4,5]

Preliminary evidence suggests that climate may modulate the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Yet it remains unclear whether seasonal and geographic variations in climate can substantially alter the pandemic trajectory, given that high susceptibility is a core driver. Here, we use a climate-dependent epidemic model to simulate the SARS-CoV-2 pandemic by probing different scenarios based on known coronavirus biology. We find that although variations in weather may be important for endemic infections, during the pandemic stage of an emerging pathogen, the climate drives only modest changes to pandemic size. A preliminary analysis of nonpharmaceutical control measures indicates that they may moderate the pandemic-climate interaction through susceptible depletion. Our findings suggest that without effective control measures, strong outbreaks are likely in more humid climates and summer weather will not substantially limit pandemic growth.

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic represents an unprecedented public health, social, and economic challenge. Sustained local transmission is present in multiple countries and in all continents, and the implications in terms of morbidity and mortality are expected to be severe (*1*, *2*). The role of seasonal and geographic climate variations in modulating the transmission of the virus has received increasing attention. Studies using a regression framework have found a role for temperature and relative and specific humidity in the transmission of SARS-CoV-2 (*3*–*7*), suggesting that cold, dry conditions increase the transmission of the virus. However, with limited data on the current epidemic, these early-stage results are inevitably inconclusive. Furthermore, the relative importance of climate drivers when compared with high population susceptibility during the pandemic stage of an emerging infection such as SARS-CoV-2 has not been fully characterized.

Climate affects the transmission of several directly transmitted pathogens (*8*). Specific humidity (the mass of water vapor in a unit mass of moist air) has been shown to be important for influenza transmission in both laboratory settings (*9*–*11*) and population-level studies (*12*). Respiratory syncytial virus (RSV), a childhood pathogen, has also been found to be dependent on specific humidity (*13*) and exhibits latitudinal correlations with climate (*14*). For both influenza and RSV, low specific humidity increases transmission, and epidemics tend to peak in the wintertime in northern latitudes. However, other directly transmitted infections exhibit different patterns (*15*), with enteroviruses, for instance, often peaking in the summer months (*16*).

Prior work on climate and directly transmitted diseases has typically considered endemic infections, such as seasonal influenza or RSV. Emerging pathogens, by contrast, have distinct dynamics driven by high population susceptibility (*17*). A key question is the extent to which seasonal and geographic climate variations are relevant in the pandemic phase of an emerging infection. Here, we build on known features of endemic human coronaviruses and other directly transmitted infections to probe this question. Although we do not yet know the climate sensitivity of SARS-CoV-2 transmission directly, data exists on four other coronaviruses that currently circulate in human populations. Two of these coronaviruses, human coronavirus HKU1 (HCoV-HKU1) and HCoV-OC43, are of the same betacoronavirus genus as SARS-CoV-2 (*18*).

We use data on HCoV-HKU1 and HCoV-OC43 from U.S. census regions to understand the potential climate dependence of betacoronavirus transmission (*19*). We fit a susceptible-infected-recovered-susceptible (SIRS) model to case data of HCoV-HKU1 and HCoV-OC43 where the fitted parameters include the climate dependence of transmission and the length of immunity after infection. All other parameters are fixed, based on values from Kissler *et al.* (*18*). Motivated by the climate dependence of influenza and RSV, we posit that transmission depends on specific humidity: We use population-weighted average climatology of specific humidity over 2014–2020 taken from the ERA5 reanalysis dataset (*20*), with population data from (*21*). We note that specific humidity is dependent on temperature through the Clausius-Clapeyron relation, and results using both variables have been found to be equivalent for other diseases (*13*). After fitting the model parameters, we run simulations of the SARS-CoV-2 pandemic under three scenarios. In the first scenario, we assume that SARS-CoV-2 has the same sensitivity to climate as influenza, based on a prior model from laboratory studies (*9*, *12*). In the second and third scenarios, we assume that SARS-CoV-2 has the same climate dependence and length of immunity as HCoV-OC43 and HCoV-HKU1, respectively. Although we assume that the climate dependence is the same as these three infections, our simulations use the basic reproduction number (R_0) based on current estimates of SARS-CoV-2 (*18*, *22*).

We first consider the seasonality of the endemic betacoronaviruses. Figure 1 shows the average seasonal pattern of endemic betacoronaviruses HCoV-OC43 and HCoV-HKU1 (hereafter OC43 and HKU1) for different regions in the United States. Cases of both diseases increase as specific humidity declines (fig. S1). We therefore assume that, to some extent, transmission will decline with specific humidity; however, the extent of the decline is yet to be determined. We characterize the link between specific humidity and the transmission of SARS-CoV-2 using plausible estimates derived from the two endemic betacoronaviruses as well as influenza. Figure 1A shows different potential functional forms for the climate-transmission relationship. Changes to specific humidity modulate R_0 between a maximum wintertime value and a hypothesized lower bound, taken from prior studies (*18*, *22*). In the extreme cases, transmission (R_0) either rapidly declines as specific humidity increases or has no relationship with specific humidity. The highlighted influenza relationship is based on laboratory studies using the guinea pig animal model (*9*–*11*) and later used to predict influenza epidemics in human populations (*12*). In this case R_0 values correspond to SARS-CoV-2 estimates.

The other two scenarios in Fig. 1A correspond to the relationship between climate and OC43 and HKU1 transmission. We evaluate the functional form of this relationship by fitting our climate-driven SIRS model to U.S. case data for the two infections (Fig. 1, B and C, and figs. S2 and S3). Our results (Fig. 1, D and E) suggest a somewhat wide range of climate dependency for the two coronaviruses, with HKU1 having a much steeper response to specific humidity than OC43. Strong seasonal forcing has been linked to biennial outbreaks, as observed for HKU1 (fig. S3), in other respiratory pathogens (*13*) and implies endemic dynamics driven by herd immunity; however, this inference may be complicated by cross-protection

[1]Princeton Environmental Institute, Princeton University, Princeton, NJ, USA. [2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. [3]Department of Geosciences, Princeton University, Princeton, NJ, USA. [4]Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ, USA. [5]Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA.
*Corresponding author. Email: racheleb@princeton.edu

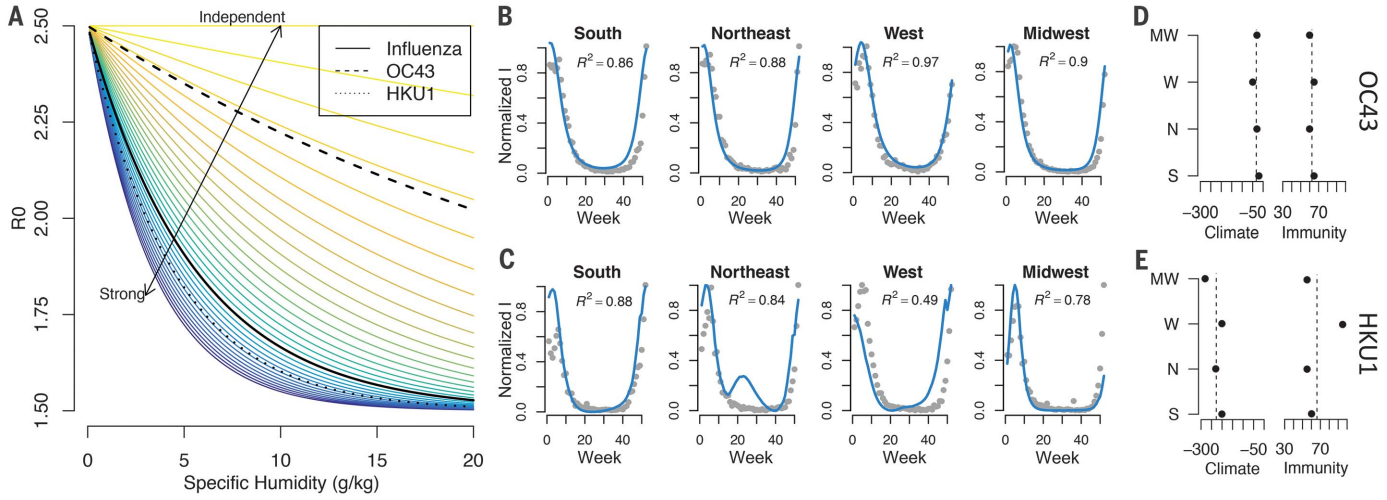


Fig. 1. Specific humidity and transmission. (**A**) Colored lines represent different hypotheses for the relationship between climate and transmission for SARS-CoV-2. Values of R_0 reflect SARS-CoV-2 estimates. The functional climate-dependence of influenza transmission, OC43 transmission, and HKU1 transmission is shown with solid, dashed, and dotted black lines, respectively. (**B** and **C**) A summary of seasonal model fits (blue line) for scaled average weekly cases (gray) of (B) OC43 and (C) HKU1. (Our model captures the biennial cycles of HKU1, shown in fig. S3, and detailed model fits for OC43, shown in fig. S2.) Coefficient of determination (R^2) values are shown. I, number infected. (**D** and **E**) Fit results in terms of climate dependence and immunity length (weeks) for (D) OC43 and (E) HKU1, where mean fits are shown with dashed lines. MW, Midwest; W, West; N, North; S, South.

from other circulating strains (*18*). Although there is some uncertainty in our estimates, simulating a pandemic outbreak using a range of climate-transmission dependencies allows us to explore a wide plausible range of potential climate effects.

We simulate a pandemic invasion for all locations (Fig. 2, A and B) and focus on the results for nine exemplar cities (Fig. 2, C to E), each with a very distinct mean and seasonal cycle of specific humidity (fig. S4). We stress that these initial simulations explore only the interaction of the epidemic (SIRS) model clockwork and seasonality in transmission; they do not address complexities of demography, control, and other environmental factors. In Fig. 2, C to E, we show the evolution of the simulated pandemic, holding population constant, for Northern Hemisphere, Southern Hemisphere, and tropical locations. The model assumes that the outbreak starts at the same time and that no control measures are in place, revealing only the effect of climate on pandemic size and duration. For the Northern Hemisphere locations, we do not see any substantial difference in pandemic size across all three scenarios, despite very different climates in New York, London, and Delhi. In the influenza and HKU1 scenarios, tropical locations experience a more sustained, lower-intensity pandemic than those in the Northern Hemisphere. These scenarios represent a stronger dependence on climate than the OC43 scenario, such that the lack of really dry conditions (low specific humidity) in tropical regions means that these locations do not experience the high transmission rates of the higher latitudes. However, the outbreak in the tropical cities remains substantial, and factors we do not explore here, such as population density, could further exacerbate the size of the epidemic.

We also simulate the pandemic in a range of Southern Hemisphere locations (Fig. 2D). We see only minor delays in the peak of Southern Hemisphere locations relative to those in the Northern Hemisphere (Fig. 2B), despite the 6-month shift in specific humidity seasonality between the two hemispheres (fig. S4). For the OC43 scenario, pandemics are temporally aligned across all locations and of similar magnitude. A stronger climate response for influenza and HKU1 parameters leads to slight regional differences. It is worth noting that our different scenarios also reflect a range of immunity lengths. The size of the pandemic peak is not affected by changes in immunity length (fig. S11), but the timing of later-stage outbreaks is partially dictated by this parameter. The differential timing of secondary peaks in the influenza and HKU1 scenarios, which have a similar climate dependence, is driven by this variability.

During the pandemic stage of an emerging pathogen, the lack of population immunity—that is, high susceptibility—is a crucial driver. To illustrate this in the general case, we run our simulation model for different climates (represented by the seasonal range of humidity values a location experiences) and different levels of population susceptibility, using the mean specific humidity and seasonality of New York. Figure 3, A to C, shows the results in terms of the size of the pandemic peak. Although humidity range does modulate pandemic size, population susceptibility exhibits a much steeper gradient. For novel pathogens, such as SARS-CoV-2, the proportion of the population susceptible to infection may be close to 1. To illustrate the potential longer-term behavior of the pandemic, we plot a typical pandemic trajectory on the susceptible-infected (SI) phase plane (Fig. 3D). The initial pandemic trajectory (red) is relatively independent of seasonal forcing. This then gives way to the endemic attractor (blue), which oscillates around the equilibrium of the unforced model (green). These longer-term dynamics show a much stronger signature of seasonal forcing than the initial pandemic phase (figs. S9 and S10).

Figures 3E and 4 show a preliminary exploration of the impact of nonpharmaceutical control on the epidemic trajectory. In Fig. 3E, we show the SI phase plane where the HKU1 parameters of R_0 are artificially controlled to R_0 = 1.1 for a 6-month period. In this scenario, the control measures result in a moderate reduction in peak incidence as the outbreak is shifted to the summer months; however, high susceptibility still results in a substantial number of cases. In Fig. 4, we explore the interaction between the climate and control measures in more detail. We consider four scenarios: climate-dependencies based on OC43 and HKU1 as well as control measures where R_0 = 1.1 or 1.3, representing limited transmission. For each scenario, we vary the length of the control measure and the location; however, for simplicity, we assume that all control measures start at two times: 4 and 6 weeks after the disease is introduced. We note that these control measures are simplified test cases and

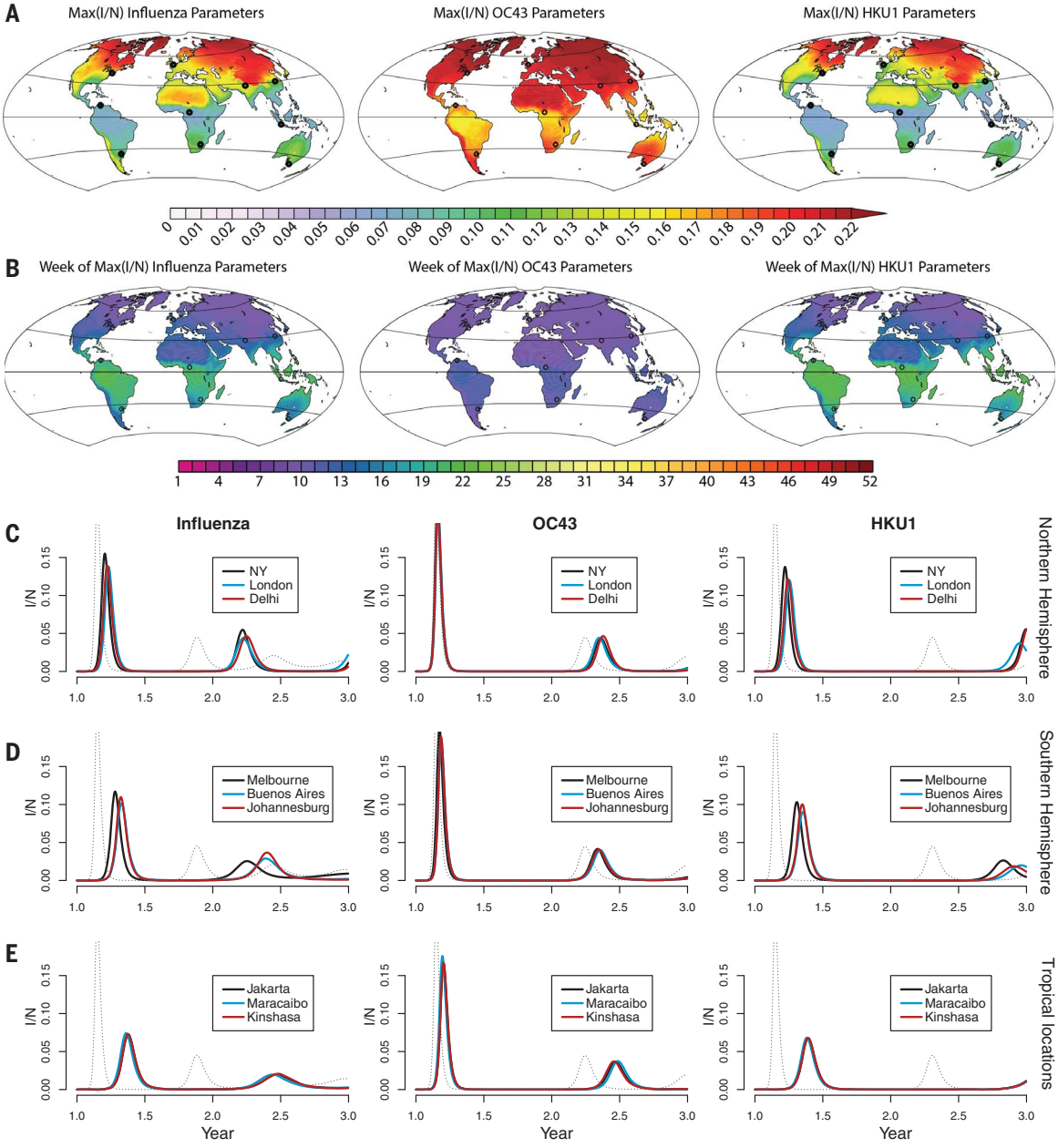


Fig. 2. Global model results and nine example trajectories. (**A** and **B**) The (A) maximum number of infections per capita (I/N) and (B) timing of peak I/N for global locations. Black circles show locations where trajectories are explicitly shown. The color scale shows maximum I/N in (A) and week of peak I/N in (B). (**C** to **E**) Simulated pandemics are shown for cities in (C) the northern hemisphere, (D) the southern hemisphere, and (E) tropical locations. The dotted line represents a pandemic with no climate dependence. NY, New York.

do not represent the local heterogeneity and efficacy of current controls, which are yet to be determined. These results show changes to peak incidence; changes to the number infected are shown in fig. S5.

For all control scenarios, we assume a degree of transmission during the control period, such that $R_0 > 1$, resulting in an increase in population immunity over time. In the scenarios where $R_0 = 1.3$, immunity builds faster and control measures work to reduce the pandemic peak after several months. In the $R_0 = 1.1$ scenario, more time is required for population immunity to build such that the pandemic peak is reduced across all locations. In this scenario, as susceptibility declines, the climate plays a more substantial role in determining pandemic peak size (Fig. 4, A and B). When $R_0 = 1.1$ in both the HKU1 and OC43 scenarios, releasing control measures close to the month of maximum transmission may result in a larger pandemic peak compared with the no-control scenario, particularly in the higher latitudes where transmission likely increases in the winter (fig. S6).

More broadly, our simulated control measures imply that the key determinant of reduced peak incidence is the extent to which population immunity builds over the control period, as demonstrated by the higher efficacy of the $R_0 = 1.3$ control scenario in mitigating peak incidence. The climate plays a complex role in tuning the efficacy of potential control efforts, resulting in differential outcomes depending on location; however, population susceptibility remains a fundamental driver. Further exploration of these complexities may be necessary when evaluating location-specific controls. The timing of introduction and the efficacy of local control measures, as well as factors such as population density and contact patterns, could also shape future outcomes. Serological surveys, at the local level, will be important for tracking the build-up of immunity over time. Moreover, implementing control measures buys crucial time while vaccines and other treatments are developed.

There are several caveats to interpreting these results. Primarily, these simulations do not address location-specific factors such as spatial and social mixing patterns, contact

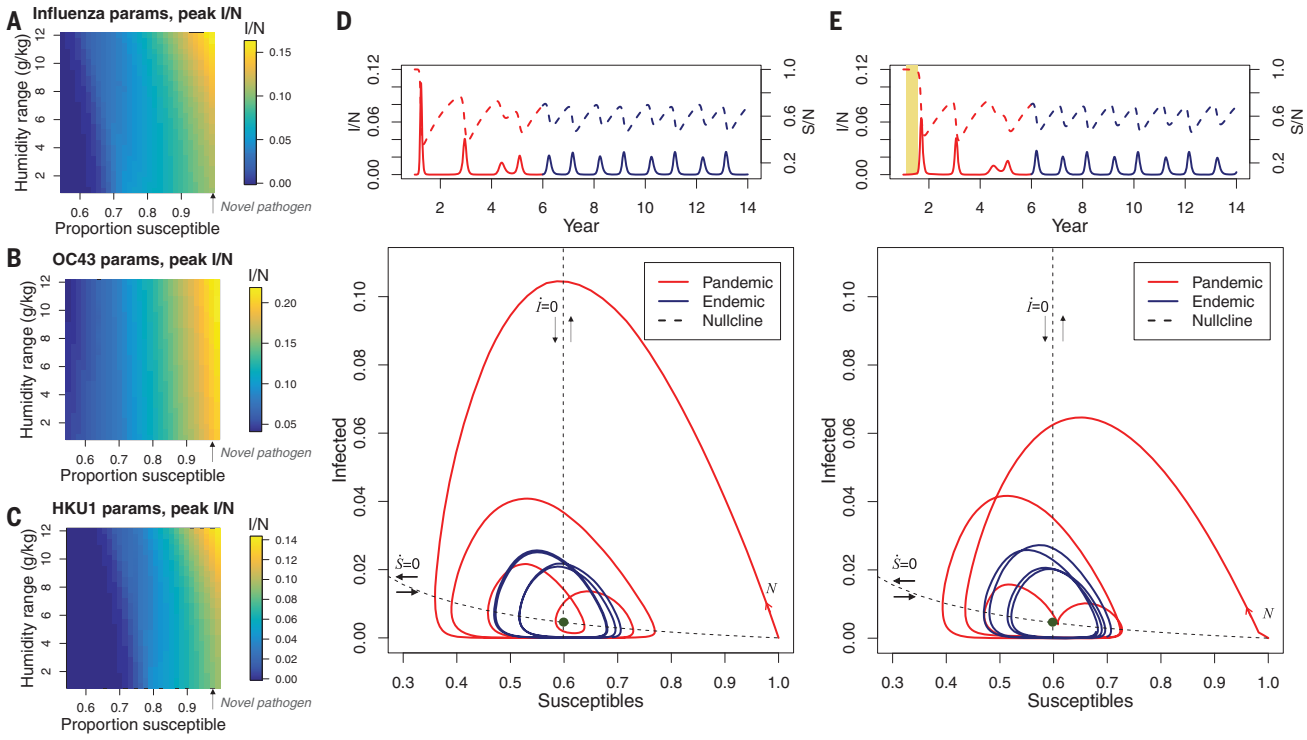


Fig. 3. Pandemic peak size depends on the proportion of the population that is susceptible. (**A** to **C**) For the three scenarios, (A) influenza, (B) OC43, and (C) HKU1, the surface plot shows the dependence of maximum pandemic incidence per capita on the seasonal range of humidity and the proportion of the population that is susceptible, assuming mean humidity of New York. params, parameters. (**D**) The time series from pandemic to endemic outbreaks for an example location (Wuhan with HKU1 params) (top) and the equivalent SI phase plane of pandemic and epidemic trajectories (bottom). The two nullclines are from the unforced SIRS using mean R_0. The green circle represents the equilibrium of the unforced model. S/N, proportion susceptible; $\dot{I}$, equilibrium infected; $\dot{S}$, equilibrium susceptible; N, total population. (**E**) The same trajectory but with a 6-month control period (reducing R_0 to 1.1). The yellow shading indicates the timing of the control period.

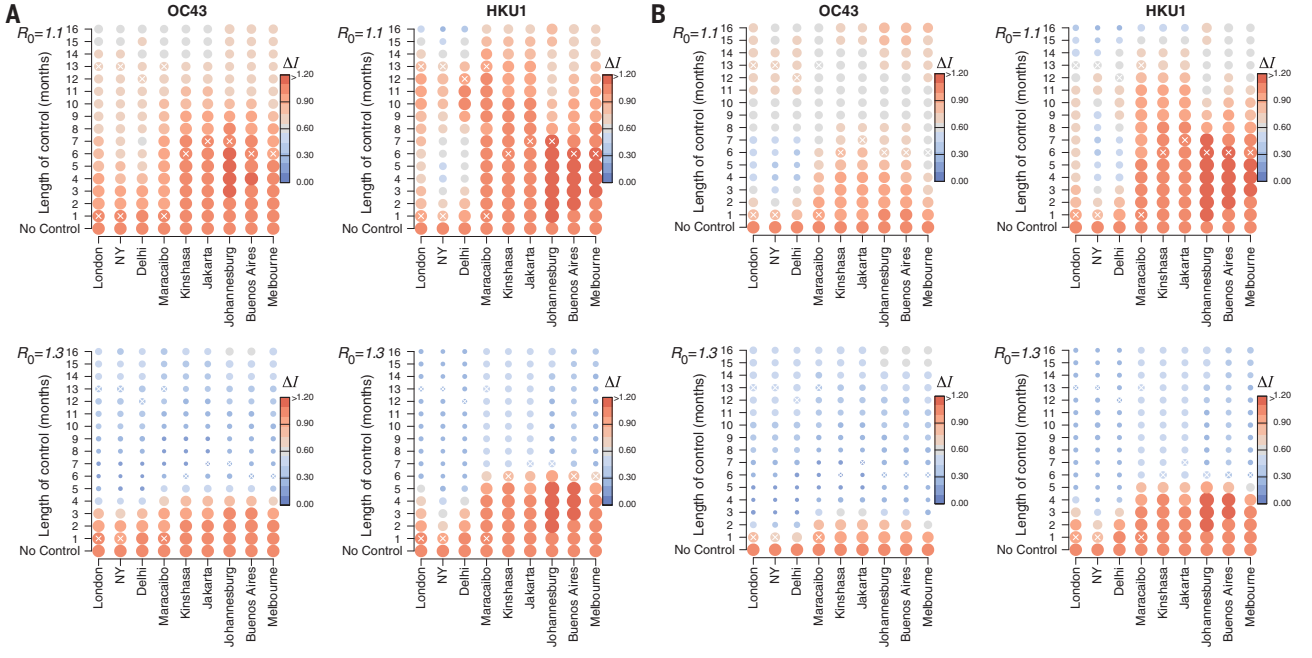


Fig. 4. Interaction between control measures and the climate. (**A** and **B**) Four scenarios representing the interaction of different climate dependencies (OC43 and HKU1 params) with two potential control measures [R_0 =1.1 and 1.3 in the control period, occurring (A) 1 month and (B) 6 weeks after pandemic start]. The size and color of the circles represent the size of peak incidence (within 2 years of pandemic start) relative to the no-control scenario. White crosses show the month of maximum climate-driven transmission, i.e., lowest specific humidity.

networks, population density, and the specifics of control timing and efficacy. In particular, our results apply most closely to relatively well-mixed epidemics in large cities. Rural areas (with potentially lower R_0; see figs. S7 and S8) may have a more delayed initial epidemic with complex consequences for ensuing interactions with climate drivers. Our model also does not account for potential cross-protection from other coronavirus infections (*18*). Cross-immunity may contribute to the seasonality of endemic coronaviruses, meaning that estimated climate drivers could be even weaker than we suggest and that our main findings are conservative (*16*, *23*). Finally, results from influenza and RSV suggest that high precipitation may play a role in driving transmission (*13*, *24*), particularly in tropical locations. Owing to limited data on betacoronaviruses from tropical locations, we have not been able to confirm whether a rainfall signal exists. Precipitation effects and other drivers such as schooling may also affect the epidemic trajectory, particularly after the pandemic. We further test for the sensitivity of our results to changes in core parameter values (materials and methods and figs. S7 to S11). These analyses suggest that our results are qualitatively robust to variations in climate dependency and weather fluctuations.

Our results suggest that although climate may play a role in modulating detailed aspects of the size and time scales of a pandemic outbreak within a particular location, population immunity is a much more fundamental driver of pandemic invasion dynamics. Although our HKU1 scenario presents a modest role for climate in terms of shifting the timing and intensity of the pandemic, a scenario with OC43 parameters is equally likely. In terms of the SARS-CoV-2 pandemic, our results imply that both tropical and temperate locations should prepare for severe outbreaks of the disease and that summertime temperatures will not effectively limit the spread of the infection. However, this does not mean that the climate is not important in the longer term. Endemic cycles of the disease will likely be tied to climate factors, and seasonal peaks may vary with latitude (figs. S3, S9, and S10). A more detailed understanding of climate drivers as well as immunity length will be crucial for understanding the implications of control measures. Furthermore, weather and near-term climate forecasts could be helpful for predicting secondary outbreaks after the initial pandemic phase has passed.

REFERENCES AND NOTES

1. R. Verity *et al*., medRxiv 2020.03.09.20033357 [Preprint]. 13 March 2020; https://doi.org/10.1101/2020.03.09.20033357.
2. R. Li *et al*., medRxiv 2020.03.09.20033241 [Preprint]. 16 March 2020; https://doi.org/10.1101/2020.03.09.20033241.
3. Q. Bukhari, Y. Jameel, SSRN 3556998 [Preprint]. 17 March 2020; https://dx.doi.org/10.2139/ssrn.3556998.
4. J. Wang, K. Tang, K. Feng, W. Lv, SSRN 3551767 [Preprint]. 9 March 2020; https://dx.doi.org/10.2139/ssrn.3551767.
5. M. M. Sajadi *et al*., *JAMA Netw. Open* **3**, e2011834 (2020).
6. W. Luo *et al*., medRxiv 2020.02.12.20022467 [Preprint]. 17 February 2020; https://doi.org/10.1101/2020.02.12.20022467.
7. Y. Ma *et al*., medRxiv 2020.03.15.20036426 [Preprint]. 18 March 2020; https://doi.org/10.1101/2020.03.15.20036426.
8. R. E. Baker, A. S. Mahmud, C. J. E. Metcalf, *Clim. Change* **148**, 547–560 (2018).
9. J. Shaman, M. Kohn, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3243–3248 (2009).
10. A. C. Lowen, J. Steel, *J. Virol.* **88**, 7692–7695 (2014).
11. A. C. Lowen, S. Mubareka, J. Steel, P. Palese, *PLOS Pathog.* **3**, 1470–1476 (2007).
12. J. Shaman, V. E. Pitzer, C. Viboud, B. T. Grenfell, M. Lipsitch, *PLOS Biol.* **8**, e1000316 (2010).
13. R. E. Baker *et al*., *Nat. Commun.* **10**, 5512 (2019).
14. V. E. Pitzer *et al*., *PLOS Pathog.* **11**, e1004591 (2015).
15. M. E. Martinez, *PLOS Pathog.* **14**, e1007327 (2018).
16. S. Takahashi *et al*., *PLOS Med.* **13**, e1001958 (2016).
17. J. R. Gog *et al*., *PLOS Comput. Biol.* **10**, e1003635 (2014).
18. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* eabb5793 (2020).
19. Centers for Disease Control and Prevention, Surveillance for common human coronaviruses (2020); www.cdc.gov/surveillance/nrevss/coronavirus/index.html.
20. L. Hoffmann *et al*., *Atmos. Chem. Phys.* **19**, 3097–3124 (2019).
21. NASA Socioeconomic Data and Applications Center (SEDAC), Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. Palisades, NY; https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11 [accessed March 2020].
22. S. W. Park *et al*., medRxiv 2020.01.30.20019877 [Preprint]. 28 February 2020; https://doi.org/10.1101/2020.01.30.20019877.
23. M. Kamo, A. Sasaki, *Physica D* **165**, 228–241 (2002).
24. J. D. Tamerius *et al*., *PLOS Pathog.* **9**, e1003194 (2013).
25. R. E. Baker, rebaker64/climate-cov: R1. Zenodo (2020); http://doi.org/10.5281/zenodo.3787766.

ACKNOWLEDGMENTS

Funding: R.E.B. is supported by the Cooperative Institute for Modeling the Earth System (CIMES). W.Y. is supported by the Princeton Environmental Institute and the Princeton Institute for International and Regional Studies. **Author contributions:** Conceptualization: R.E.B., W.Y., G.A.V., C.J.E.M., B.T.G.; Data curation: R.E.B., W.Y.; Formal analysis: R.E.B.; Methodology: R.E.B., W.Y., G.A.V., C.J.E.M., B.T.G.; Software: R.E.B., W.Y., G.A.V.; Visualization: R.E.B., G.A.V.; Writing – original draft: R.E.B.; Writing – reviewing and editing: R.E.B., W.Y., G.A.V., C.J.E.M., B.T.G. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** U.S. coronavirus data are available from the National Respiratory and Enteric Virus Surveillance System (NREVSS) on request at www.cdc.gov/surveillance/nrevss/coronavirus/index.html. Climate data are publicly available from ERA5 and NASA: ERA5 data at www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5 and NASA MERRA data at https://gmao.gsfc.nasa.gov/reanalysis/MERRA/. Code for running the model is available at Zenodo (*25*). 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/369/6501/315/suppl/DC1
Materials and Methods
Figs. S1 to S11
References (*26*–*29*)
MDAR Reproducibility Checklist

15 April 2020; accepted 14 May 2020
Published online 18 May 2020
10.1126/science.abc2535

IMMUNOLOGY

Autoreactivity in naïve human fetal B cells is associated with commensal bacteria recognition

Jeff W. Chen[1]*, Tyler A. Rice[1], Jason M. Bannock[1], Agata A. Bielecka[1]†, Juliet D. Strauss[1]‡, Jason R. Catanzaro[1], Haowei Wang[1], Laurence C. Menard[1]§, Jennifer H. Anolik[2], Noah W. Palm[1]¶, Eric Meffre[1,3]¶

Restricted V(D)J recombination during fetal development was postulated to limit antibody repertoire breadth and prevent autoimmunity. However, newborn serum contains abundant autoantibodies, suggesting that B cell tolerance during gestation is not yet fully established. To investigate this apparent paradox, we evaluated the reactivities of more than 450 antibodies cloned from single B cells from human fetal liver, bone marrow, and spleen. We found that incomplete B cell tolerance in early human fetal life favored the accumulation of polyreactive B cells that bound both apoptotic cells and commensal bacteria from healthy adults. Thus, the restricted fetal preimmune repertoire contains potentially beneficial self-reactive innate-like B cell specificities that may facilitate the removal of apoptotic cells during development and shape gut microbiota assembly after birth.

V(D)J recombination generates a diverse adult antibody repertoire, but immunoglobulin heavy chain (IgH) rearrangements are restricted during prenatal development (*1*). Restriction of the fetal antibody repertoire is thought to limit autoimmunity by preventing the generation of self-reactive B cell clones (*1*). However, fetal B cell receptors are often polyreactive (*2*), and human newborns display abundant serum IgM autoantibodies (*3*), which suggests incomplete B cell tolerance during gestation. We sought to determine the extent to which B cell tolerance is established in early human life and whether differences between fetal and adult B cell receptor repertoires result in distinct antibody binding specificities (*4*). We generated more than 450 recombinant antibodies cloned from single B cells isolated from liver, bone marrow (BM), and spleens from four second-trimester human fetuses and observed an enrichment in the downstream variable regions of immunoglobulin heavy chain (V_H), D7-27, and upstream heavy-chain joining region (J_H) gene segment usage, as previously reported (tables S1 to S3 and figs. S1 and S2) (*1*, *5*).

Central B cell tolerance in adults is established via elimination of developing autoreactive clones between the early immature and immature B cell stages in the BM (*4*). We observed a significant purging of self-reactive and antinuclear clones in immature B cells from both fetal liver and BM, suggesting that central B cell tolerance is largely functional in early fetal life (Fig. 1, A to D, and fig. S3, A to D). By contrast, polyreactive clones were only modestly reduced during early fetal B cell selection but were purged from adult immature B cells (Fig. 1E). We next evaluated whether central B cell tolerance was mediated by secondary recombination of the immunoglobulin kappa (Igκ) and then lambda (Igλ) loci (i.e., receptor editing) (*6*). Diverse V–J recombination events were observed on both the Igκ and Igλ loci in fetal B cells, but fetal immature B cells were enriched in in-frame Igλ transcripts as compared with their adult counterparts (Fig. 1, F and G, and figs. S5, A to C, and S6). Increased secondary recombination events in fetal immature B cells were also evidenced by greater usage of distal variable regions of immunoglobulin κ chain (Vκs) and downstream kappa chain joining region 2 (Jκ2) and Jκ3 segments, decreased usage of the most-upstream Jκ1 segment, and increased recombination with the κ-deleting element, which enables Igλ recombination, as compared with adult counterparts (Fig. 1, H to J). Thus, central B cell tolerance in fetal primary lymphoid organs silences developing autoreactive clones via enhanced secondary recombination (*6*) but is not as efficient at eliminating polyreactive clones.

In adults, a peripheral B cell tolerance checkpoint counterselects autoreactive new emigrant or transitional B cells before they enter the mature naïve B cell compartment (Fig. 2, A and B) (*4*). Spleens from second-trimester human fetuses are immature and contain naïve B cells that may resemble those from fetal peripheral blood (figs. S7 and S8). Thus, we assessed peripheral B cell tolerance by evaluating autoreactivity in new emigrant B cells and mature naïve B cells from fetal spleens versus adult peripheral blood (Fig. 2, A and B, and fig. S7). Unlike in adults, human epithelial type 2 (HEp-2) reactive clones in fetuses remained stable between the new emigrant or transitional B cell and mature naïve B cell stages (Fig. 2A and fig. S9A). Fetal mature naïve B cells also contained an elevated proportion of polyreactive clones (Fig. 2B and fig. S9B). These data suggest that the peripheral B cell tolerance checkpoint is not yet functional during second-trimester gestation.

Because peripheral B cell tolerance is controlled by T cells, including regulatory T cells (T_{regs}) (*7*, *8*), we assessed the proportions of T cells and T_{regs} versus B cells in fetal spleens (Fig. 2, C to F). Although the proportion of T cells in adult blood far exceeded that of B cells, B cells were more prevalent than T cells in fetal spleens (Fig. 2, C and D) (*5*). In addition, the ratios of T_{regs} to new emigrant or transitional B cells were significantly reduced in fetal spleen, liver, and BM compared with adult blood and spleen, despite T_{regs} being more prevalent in fetuses than adults (Fig. 2, E and F, and fig. S8) (*9*). The limited numbers of T cells and T_{regs} present during early fetal life may be insufficient to establish robust peripheral B cell tolerance and thus enable the accumulation of autoreactive and polyreactive innate-like B cells in fetuses. These cells may resemble murine autoreactive B-1 cells produced by the fetal liver and positively selected to secrete natural antibodies (*10*, *11*). Thus, human fetal liver-derived B cells, which can self-renew and persist long-term in the peripheral B cell pool (*12*), may be an important source of natural antibodies into adulthood.

To explore the hypothesis that fetal self-reactive B cells encode potentially beneficial natural antibodies that may aid in clearance of dead cells (*13*), we examined the binding of recombinant antibodies cloned from fetal and adult B cells to live, early apoptotic, or dead cells (Fig. 3A). Although antibodies from a few early immature fetal B cells bound live cells, live-cell reactivity was purged from more mature B cell compartments (Fig. 3B). By contrast, early apoptotic and dead-cell-reactive clones remained present throughout B cell development in both fetuses and adults and were significantly more prevalent in fetal than in adult B cells (Fig. 3B and table S3). Thus, tolerance does not appear to counterselect early apoptotic and dead-cell-reactive clones, and increased autoreactivity among fetal B cells may favor the clearance of cellular debris generated by developmental apoptosis.

Polyreactivity has been associated with cross-reactivity to commensal bacteria in IgA^+ B cells in both mice and humans (*14*, *15*). We therefore assessed the ability of recombinant antibodies cloned from fetal (n = 95) or adult (n = 131) mature naïve B cells to bind human gut

[1]Department of Immunobiology, Yale University, New Haven, CT 06511, USA. [2]Division of Allergy, Immunology and Rheumatology, University of Rochester Medical Center, Rochester, NY 14642, USA. [3]Section of Rheumatology, Allergy and Clinical Immunology, Yale School of Medicine, New Haven, CT 06511, USA.
*Present address: ClearView Healthcare Partners, Newton, MA 02458, USA.
†Present address: Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany.
‡Present address: Tulane School of Medicine, New Orleans, LA 70112, USA.
§Present address: Bristol-Myers Squibb, Princeton, NJ 08540, USA.
¶Corresponding author. Email: eric.meffre@yale.edu (E.M.); noah.palm@yale.edu (N.W.P.)

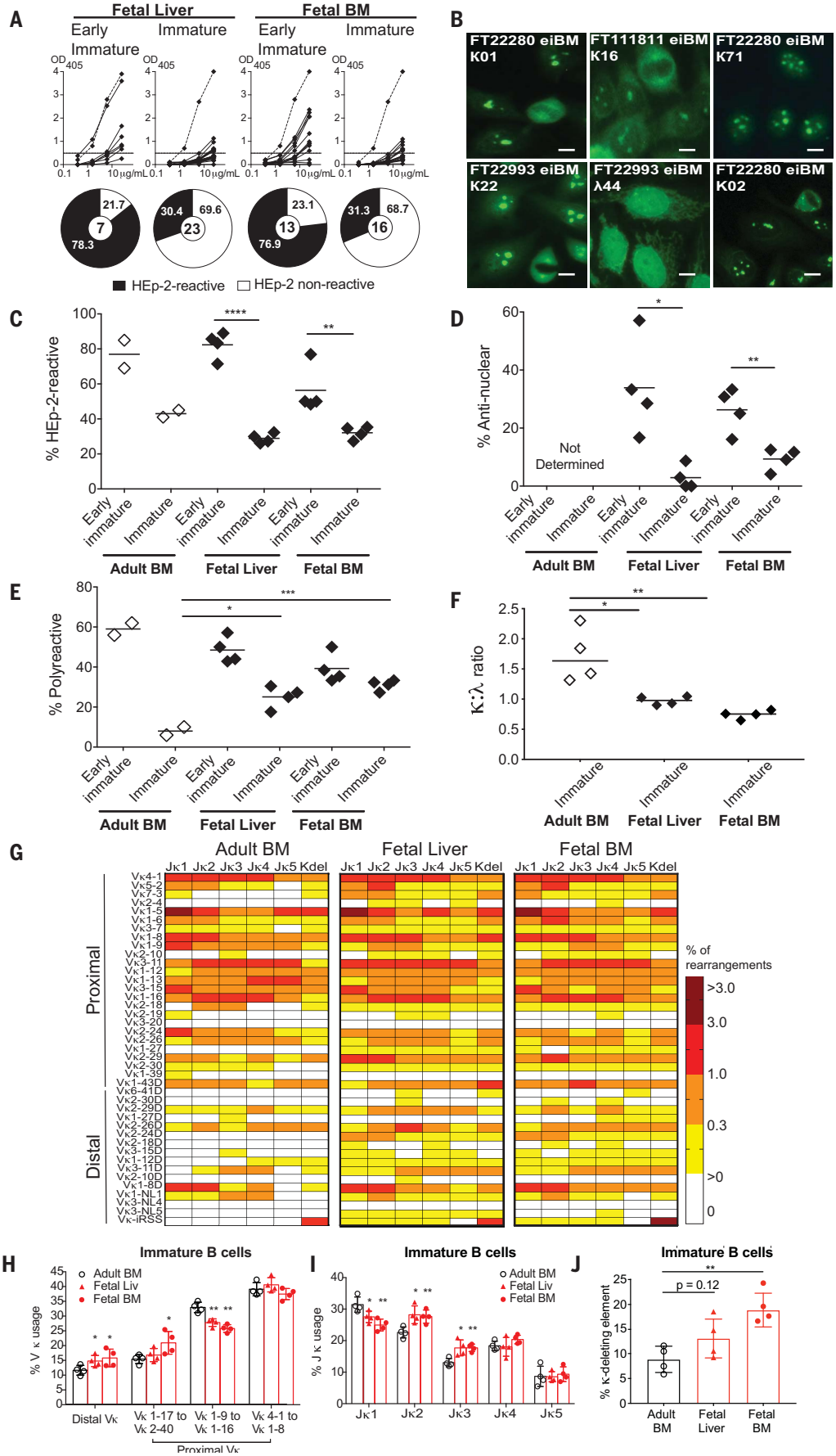


Fig. 1. Fetal B cells undergo active central tolerance through Igκ secondary recombination. (**A**) Representative HEp-2 reactivities of antibodies from individual fetal B cells. Dotted lines show the positive control. Pie charts represent the frequencies of reactive (solid) and nonreactive (open) clones, with the number of clones tested (*n*) indicated in the center. (**B**) Representative nuclear staining patterns for antibodies cloned from fetal early immature B cells. Scale bars, 25 μm. (**C** to **E**) Frequencies of HEp-2 reactive (C), antinuclear reactive (D), and polyreactive (E) clones in fetal and adult B cell fractions. Each data point summarizes the reactivity data from an average of *n* = 17 cloned recombinant antibodies from an individual fetal (*n* = 4) or adult (*n* = 2) B cell fraction. (**F**) Summary of κ:λ light chain ratios in fetal and adult immature B cells (*n* = 4 samples per group). More than 40,000 in-frame κ and λ sequences were analyzed per sample. (**G**) Vκ and Jκ gene pairing in immature B cells. NL, not localized in the locus; iRSS, intronic heptamer recombination signal sequence; Kdel, κ-deleting element. (**H** to **J**) Vκ gene usage (H), Jκ gene usage (I), and frequency of κ-deleting element (J) in fetal and adult immature B cells. Error bars represent SEM. Mann-Whitney *U* test with Bonferroni correction; **P* < 0.05; ***P* < 0.01; *****P* < 0.0001.

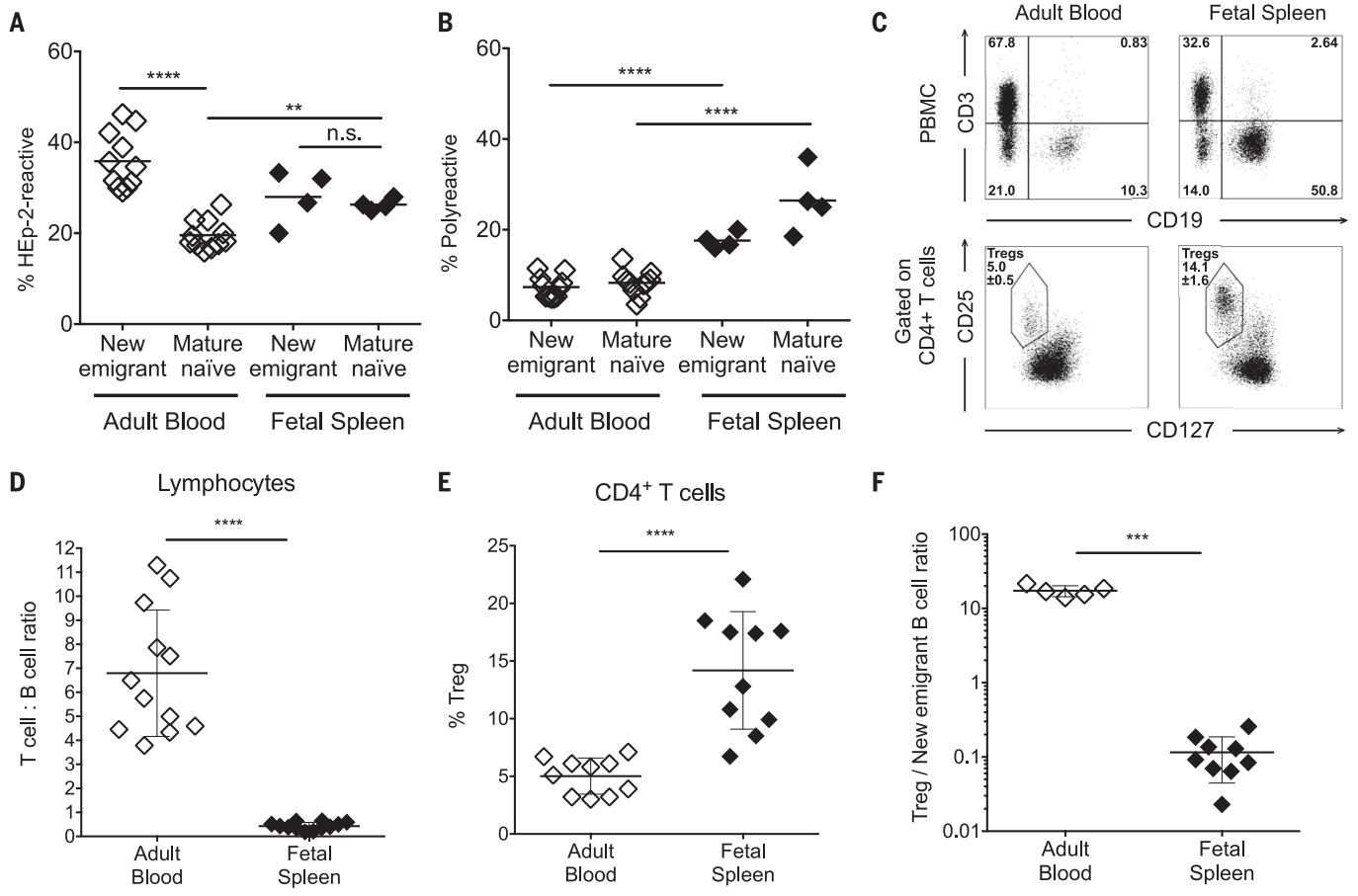


Fig. 2. Impaired peripheral B cell tolerance checkpoint in fetuses is associated with scarcity of T cells and T_{regs}. (**A** and **B**) Frequency of HEp-2 reactivity (A) and polyreactivity (B) in adult and fetal new emigrant and mature naïve B cells. Each data point summarizes the reactivity of n = 16 to 30 recombinant antibodies cloned from individual fetal (n = 4) or adult (n = 12) samples. (**C**) Representative dot plots showing the relative proportions of peripheral T ($CD3^+$) versus B ($CD19^+$) cells (top row) and $CD3^+CD4^+$ $CD25^{hi}CD127^{lo}$ T_{regs} (bottom row) in adults and fetuses. (**D** to **F**) Summaries of T:B cell ratios (D), the proportion of T_{regs} out of $CD4^+$ T cells (E), and T_{reg}: transitional B cell ratios (F) in adult blood (n = 12) and fetal spleen (n = 10) samples. Error bars represent SEM. Mann-Whitney U test with Bonferroni correction; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; n.s., not significant.

commensals. About 65% of these antibodies showed no detectable commensal reactivity (fig. S10A). However, some clones stained up to 10 to 15% of fecal bacteria, roughly half of which were also bound by secretory IgA (Fig. 3C). Fetal naïve B cell–derived clones were significantly enriched in commensal reactivity as compared to adult-derived clones (Fig. 3D). Thus, commensal reactivity is prevalent in early prenatal B cells despite the largely sterile womb environment (*16*). Reactivity was consistent across 22 antibodies and four fecal samples from unrelated donors, suggesting that these antibodies recognize prevalent bacterial taxa in the healthy human microbiome (fig. S10B).

To determine which commensal bacteria were bound by stool-reactive recombinant antibodies, we stained human fecal bacteria with 10 stool-reactive antibodies cloned from fetal B cells and 12 stool-reactive antibodies cloned from adult B cells, separated the antibody-bound and -unbound bacteria, and performed 16*S* rRNA gene sequencing (*17*). Stool-reactive antibodies enriched bacteria from the genera *Blautia* and *Dorea* by threefold and more than 100-fold, respectively, but did not enrich phylogenetically related taxa from the genera *Faecalibacterium* and *Ruminococcus* (Fig. 3, E and F, and fig. S10, C and D). *Dorea* species were similarly enriched across four stool samples from independent donors, whereas other bacterial binding patterns were specific to each individual microbiome (fig. S10E). *Blautia* and *Dorea* were not enriched by six non-stool-reactive antibodies (fig. S11). Stool-reactive antibodies also bound pure *Dorea* and *Blautia* isolates but did not bind *Faecalibacterium* or *Brevundimonas* strains (figs. S12 and S13). Mature naïve B cells with intrinsic commensal bacteria reactivity may serve as precursors for IgA-secreting plasma cells in the gut. The resulting antibodies may favor the selection or expansion of *Blautia* and *Dorea* species, which are highly coated with secretory IgA in healthy humans and reduced in IgA-deficient subjects (*17*, *18*).

Fetal mature naïve B cells were four times more likely to be polyreactive, recognize dead cells, and/or bind commensals than were their adult counterparts (Fig. 4, A and B). Hence, self-reactivity of the fetal antibody repertoire is associated with binding to commensal bacteria. V_H3 family genes—the most highly conserved across evolution—were specifically associated with commensal binding in both fetuses and adults (Fig. 4C). Furthermore, nearly all commensal-binding antibodies were encoded by V_H3 segments, which may also modestly favor dead-cell binding in fetuses but not polyreactivity (Fig. 4D). However, commensal binding by V_H3-encoded antibodies could not be explained by known V_H3-binding proteins made by bacteria, including *Staphylococcus* protein A (*19*) (fig. S14) or IbpA and IbpB superantigens from *Ruminococcus gnavus* (fig. S15) (*20*).

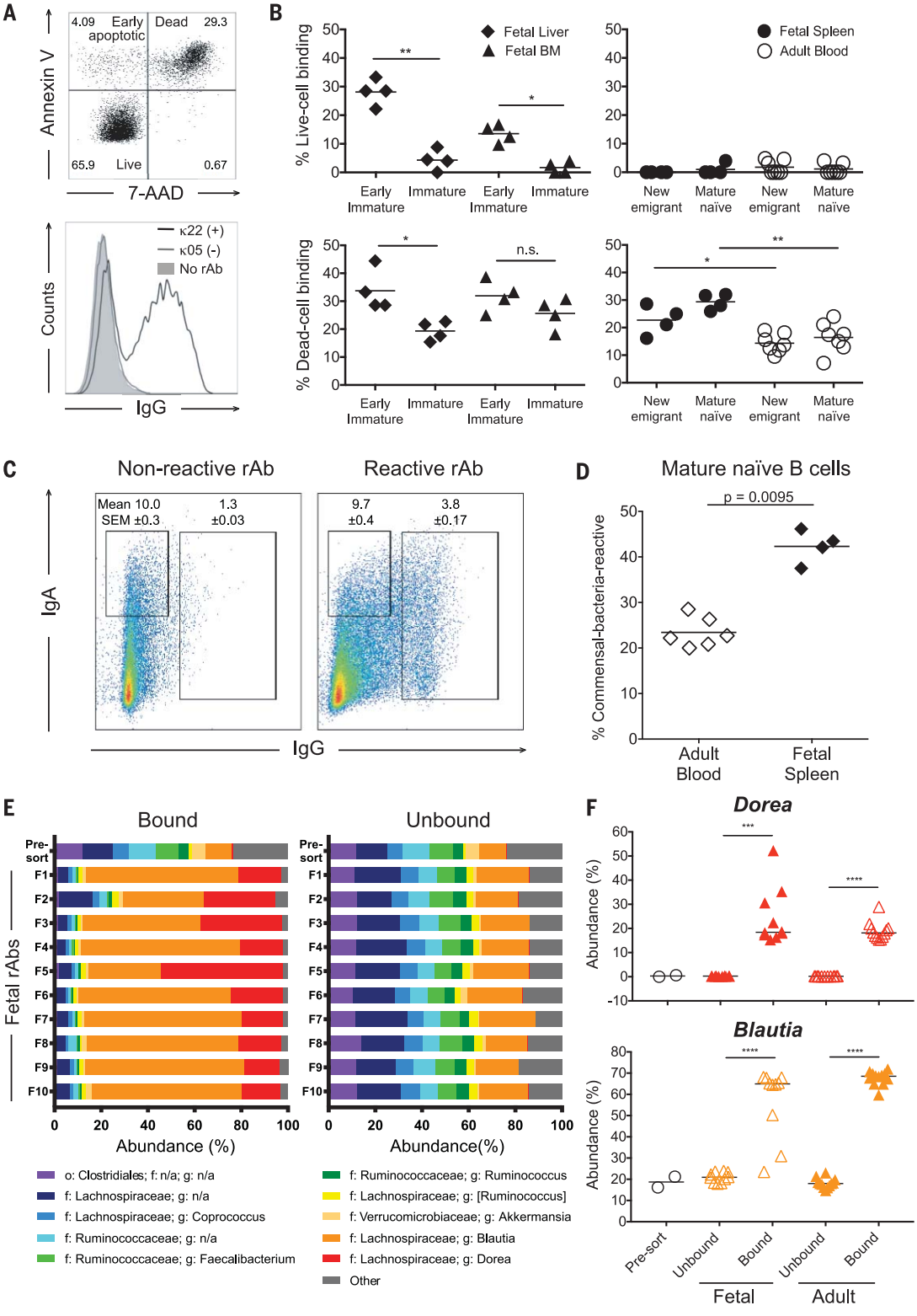


Fig. 3. Fetal mature naïve B cells are enriched in dead-cell- and commensal-bacteria-reactive clones. (**A**) Gating of live and dead Jurkat cells (top) and representative dead-cell binding by a reactive (κ22) versus nonreactive (κ05) fetal antibody (bottom). (**B**) Summary of live- and dead-cell binding by B cell clones across developmental stages. Each data point summarizes the reactivity of n = 16 to 30 recombinant antibodies cloned from individual fetal (n = 4) or adult (n = 7) samples. (**C**) Coating of human stool bacteria with endogenous secretory IgA (IgA) and commensal bacteria-reactive (n = 77) or nonreactive (n = 135) recombinant antibodies (rAb). Mean and SEM for gated fractions are indicated. (**D**) Frequencies of commensal bacteria-reactive clones in mature naïve B cells isolated from six adults (total of 121 tested clones) versus four fetuses (total of 91 tested clones). Each data point summarizes n = 16 to 30 clones per individual. (**E** and **F**) Relative abundances of all bacterial genera >1% (E) or *Dorea* and *Blautia* (F) in presort, rAb-bound, and rAb-unbound gut microbiome fractions (n = 10 different fetal rAbs; F1 to F10). o, order; f, family; g, genus. Mann-Whitney U test with Bonferroni correction; $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

Swapping of IgH complementarity-determining region 3 sequences (CDR3s) from stool-reactive V_H3-expressing clones to nonreactive V_H1 or V_H4 antibodies also was insufficient to confer commensal reactivity, underscoring the importance of germline-encoded V_H3 sequences (fig. S16). Attempts to identify amino acid residues favoring commensal reactivity associated with specific V_H3 family members were also unsuccessful (figs. S17 and S18). Long and positively charged IgH CDR3s favored polyreactivity (*4*) but did not correlate with commensal reactivity or dead-cell binding (fig. S19). Elevated D7-27 gene segment usage and increased Igλ expression in fetal antibodies also were not associated with any specific pattern of reactivity (fig. S20). J_H2 gene segments, which are common in fetal but not adult B cells

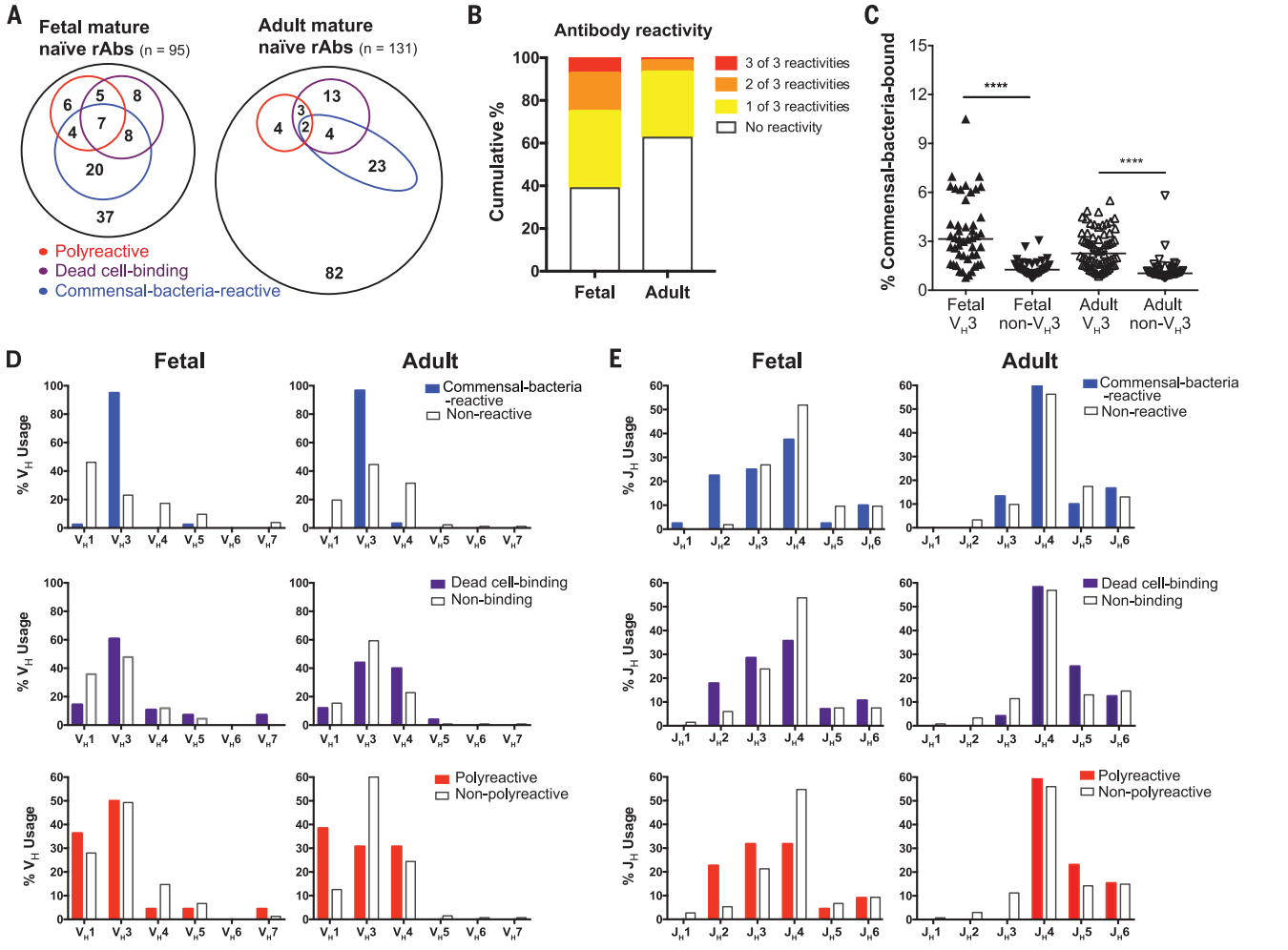


Fig. 4. Commensal bacteria reactivity in fetuses is associated with self-reactivity. (**A** and **B**) Fetal and adult mature naïve B cell reactivity patterns represented as Venn diagrams (A) and stacked histograms (B). (**C**) Proportion of human gut commensal bacteria bound by V_H3- or non-V_H3-encoded antibodies cloned from fetal or adult mature naïve B cells. Mann-Whitney *U* test with Bonferroni correction; $^{****}P < 0.0001$. (**D** and **E**) Distribution of V_H family (D) and J_H gene (E) usage in antibody clones from fetal ($n = 95$) and adult ($n = 131$) B cells that bind or do not bind commensal bacteria or dead cells, or display polyreactivity.

(figs. S1 and S2), were enriched among fetal commensal-reactive, dead-cell-binding, and polyreactive clones, but this trend was absent among adult B cells, in which J_H2 gene usage was rare (Fig. 4E). Thus, commensal reactivity in fetal B cells likely lies within the germline-encoded V_H3 gene segments and potentially the J_H2 gene segment. Nonetheless, although our CDR3-swapping and mutagenesis studies strongly suggest canonical binding, we cannot rule out CDR-independent commensal binding or interactions with undiscovered superantigens.

These studies illustrate that incomplete B cell tolerance mechanisms in early human fetal life favor the accumulation of autoreactive fetal innate-like B cells that also bind to specific commensal bacteria. The prevalence of anticipatory commensal reactivities before microbial exposure may promote colonization of the gut by beneficial commensal taxa and facilitate the development of advantageous host-microbiota interactions after birth (*21*, *22*).

REFERENCES AND NOTES

1. H. W. Schroeder Jr., J. L. Hillson, R. M. Perlmutter, *Science* **238**, 791–793 (1987).
2. J. L. Hillson, I. R. Oppliger, E. H. Sasso, E. C. Milner, M. H. Wener, *J. Immunol.* **149**, 3741–3752 (1992).
3. Y. Merbl, M. Zucker-Toledano, F. J. Quintana, I. R. Cohen, *J. Clin. Invest.* **117**, 712–718 (2007).
4. H. Wardemann *et al.*, *Science* **301**, 1374–1377 (2003).
5. E. Rechavi *et al.*, *Sci. Transl. Med.* **7**, 276ra25 (2015).
6. D. Nemazee, M. Weigert, *J. Exp. Med.* **191**, 1813–1818 (2000).
7. J. Sng *et al.*, *Sci. Immunol.* **4**, eaav6778 (2019).
8. T. Kinnunen *et al.*, *Blood* **121**, 1595–1603 (2013).
9. J. E. Mold *et al.*, *Science* **322**, 1562–1565 (2008).
10. K. Hayakawa *et al.*, *Science* **285**, 113–116 (1999).
11. J. F. Kearney, P. Patel, E. K. Stefanov, R. G. King, *Annu. Rev. Immunol.* **33**, 475–504 (2015).
12. A. Roy *et al.*, *Clin. Immunol.* **183**, 8–16 (2017).
13. Y. Chen, Y. B. Park, E. Patel, G. J. Silverman, *J. Immunol.* **182**, 6031–6043 (2009).
14. M. A. Berkowska *et al.*, *J. Immunol.* **195**, 1417–1426 (2015).
15. J. J. Bunker *et al.*, *Science* **358**, eaan6619 (2017).
16. M. C. de Goffau *et al.*, *Nature* **572**, 329–334 (2019).
17. N. W. Palm *et al.*, *Cell* **158**, 1000–1010 (2014).
18. J. Fadlallah *et al.*, *Sci. Transl. Med.* **10**, eaan1217 (2018).
19. A. J. Ulloa-Morales, C. S. Goodyear, G. J. Silverman, *Front. Immunol.* **9**, 2011 (2018).
20. J. J. Bunker *et al.*, *Sci. Transl. Med.* **11**, eaau9356 (2019).
21. J. L. Kubinak, J. L. Round, *Nat. Rev. Immunol.* **16**, 767–774 (2016).
22. D. B. Sutherland, K. Suzuki, S. Fagarasan, *Immunol. Rev.* **270**, 20–31 (2016).

ACKNOWLEDGMENTS

We thank L. Devine and C. Wang for cell sorting and W. Schofield for advice on 16S library preparation and sequencing. We also thank M. Kriegel for insightful discussions. **Funding:** This work was supported by grant funding from the NIH AI061093 and AI118855 (to E.M.). J.W.C. is supported by the Gruber Science Foundation and the Richard K. Gershon fellowship from the Department of Immunobiology at Yale University. N.W.P. also acknowledges funding from the NIH (DK125119, AI137935, AI123477), Helmsley and Pew Charitable Trusts, and the Kenneth Rainin and Smith Family Foundations. **Author contributions:** E.M. initiated the collaboration with N.W.P. for testing antibody reactivity to commensal bacteria. J.W.C., T.A.R., J.M.B., A.A.B., H.W., J.D.S., and L.C.M. performed the experiments. A.A.B. and J.R.C. provided stool samples and monocultured bacteria, and J.H.A. provided

adult BM samples. J.W.C., T.A.R., N.W.P., and E.M. interpreted and discussed the data. J.W.C., N.W.P., and E.M. wrote the manuscript. All authors reviewed the manuscript and provided scientific input. **Competing interests:** E.M. is an adviser for and receives funding from AbbVie, Inc., and N.W.P. is a cofounder of and receives funding from Artizan Biosciences, Inc. Other authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main paper or the supplementary materials. Sequencing data are available through the immuneACCESS portal (https://clients.adaptivebiotech.com/pub/chen-2020-science) and the Sequence Read Archive (accession PRJNA631283). The authors agree to provide materials upon request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/320/suppl/DC1
Materials and Methods
Figs. S1 to S21
Tables S1 to S5
References (*23*–*26*)
MDAR Reproducibility Checklist

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

3 August 2019; resubmitted 15 January 2020
Accepted 4 June 2020
10.1126/science.aay9733

AGING

A programmable fate decision landscape underlies single-cell aging in yeast

Yang Li[1], Yanfei Jiang[1], Julie Paxman[1], Richard O'Laughlin[2], Stephen Klepin[1], Yuelian Zhu[1], Lorraine Pillus[1,3], Lev S. Tsimring[4], Jeff Hasty[1,2,4], Nan Hao[1,4]*

Chromatin instability and mitochondrial decline are conserved processes that contribute to cellular aging. Although both processes have been explored individually in the context of their distinct signaling pathways, the mechanism that determines which process dominates during aging of individual cells is unknown. We show that interactions between the chromatin silencing and mitochondrial pathways lead to an epigenetic landscape of yeast replicative aging with multiple equilibrium states that represent different types of terminal states of aging. The structure of the landscape drives single-cell differentiation toward one of these states during aging, whereby the fate is determined quite early and is insensitive to intracellular noise. Guided by a quantitative model of the aging landscape, we genetically engineered a long-lived equilibrium state characterized by an extended life span.

Many damage factors, including chromatin instability, mitochondrial dysfunction, and reactive oxygen species (*1*, *2*), contribute to cellular aging. In each individual cell, how these factors combine to drive the aging process remains unclear. For instance, aging could be driven by independent damage mechanisms that accumulate at varying rates, resulting in different aged phenotypes in individual cells, or, alternatively, by the deterioration of overall cellular condition, leading to a common aging pathway in all cells. Single-cell analysis can reveal which scenario actually underlies the aging process. We investigated replicative aging of the yeast *Saccharomyces cerevisiae*, a genetically tractable model for the aging of mitotic cell types such as stem cells (*3*, *4*). Yeast aging research has focused on life span, as measured by the number of cell divisions before death (*5*). We integrated microfluidics (fig. S1) with time-lapse microscopy to measure life span along with dynamic changes of gene expression, chromatin state, and organelle morphology, enabling us to reveal a long-lived mode of the single-cell aging process in yeast.

We found that isogenic wild-type (WT) cells exhibited two types of phenotypic changes during aging (*6*, *7*). About half of cells produced daughters with an elongated morphology during later stages of their life span. In contrast, the other half continuously produced small, round daughter cells until death. We designate these phenotypic changes as "mode 1" and "mode 2" aging, respectively (Fig. 1A and figs. S2 and S3). Aberrant structural changes of the nucleolus and mitochondria are hallmarks of aging in many organisms, indicating age-associated organellar dysfunction (*8*, *9*). Nucleoli were enlarged and fragmented in all mode 1 aged cells but not in mode 2 cells (Fig. 1B, top, and fig. S4). Mitochondria in mode 1 cells, however, retained a normal tubular morphology throughout life, whereas those in mode 2 became aggregated before cell death (Fig. 1B, bottom), in accord with mitochondrial decline observed previously in a fraction of cells (*10*). Thus, the two aging modes were associated with distinct organellar failures: mode 1 with nucleolar decline and mode 2 with mitochondrial decline.

Age-dependent nucleolar enlargement is related to instability of ribosomal DNA (rDNA) (*11*, *12*). The conserved lysine deacetylase Sir2, encoded by a well-studied longevity gene, maintains rDNA stability by mediating rDNA silencing (*13*). To track rDNA silencing, we used a green fluorescent protein (GFP) reporter inserted at the nontranscribed spacer region of rDNA (rDNA-GFP) (*6*). Expression of the reporter is repressed by silencing, so increased fluorescence indicates a loss of silencing. Cells exhibited sporadic loss of silencing during early phases of aging (*6*). Late in aging, mode 1 cells underwent sustained loss of silencing, but mode 2 cells did not (Fig. 1C), in accord with nucleolar decline only in mode 1 cells. Exposure to nicotinamide (NAM), a Sir2 inhibitor that disrupts rDNA silencing, induced mode 1 aging in most cells (fig. S5A), suggesting a role of Sir2 and rDNA silencing in driving mode 1 aging. NAM treatment can also lead to elongated cell morphology, which may be a consequence of loss of Sir2 activity (fig. S5B).

For mode 2 aging, we monitored heme, an iron-containing compound proposed to be a key factor for mitochondrial decay during human aging (*14*). To determine how heme abundance changed during aging, we used two independent reporters: a fluorescent HS1 heme sensor (*15*) and a nuclear anchored infrared fluorescent protein reporter (nuc. iRFP) (*16*). For HS1, because binding of heme to its receptor quenches GFP fluorescence, increased GFP signal indicates decreased heme. For nuc. iRFP, fluorescence depends on a product of heme catabolism, biliverdin, and was found to correlate with the amount of cellular heme (fig. S6). These two reporters, expressed in the same cells, showed consistently that the amount of heme decreased in mode 2 aging but not mode 1 aging (Fig. 1D). Heme activates the heme activator protein (HAP) transcriptional complex to maintain mitochondrial biogenesis and function (*17*). Consistently, expression of HAP-regulated genes *COX5a* and *CYC1*, both of which encode key mitochondrial components, decreased specifically in mode 2 cells (figs. S7 and S8). Depletion of heme by succinylacetone induced mode 2 aging in most cells (fig. S5C), indicating that the decrease in heme abundance and HAP activity may drive mode 2 aging. Heme depletion also coincided with increased abundance of the plasma membrane proton pump Pma1 (fig. S9), which causes decreased vacuolar acidity and mitochondrial dysfunction during aging (*9*, *18*, *19*). Moreover, mode 2 aging featured a shorter life span and a greater extension of cell cycle length than mode 1 aging (Fig. 1, E and F).

To track rDNA silencing and heme abundance simultaneously, we introduced rDNA-GFP and nuc. iRFP reporters into the same cells. Newborn cells began with relatively uniform amounts of the two reporters and diverged in two opposite directions during the early half of their lifetimes. Mode 1 cells (47.3% of the population) ended life with high GFP and high iRFP signals, corresponding to a state with low rDNA silencing and a high abundance of heme. In contrast, mode 2 cells (52.7% of the population) ended in a state with high rDNA silencing and a low abundance of heme (Fig. 1G, fig. S10, and movie S1). Thus, isogenic cells age toward two discrete terminal states with anticorrelated rDNA silencing and heme abundance, suggesting interactions between these pathways.

To test for such an interaction, we deleted *SIR2* to disrupt rDNA silencing in the dual reporter strain. The majority (83.1%) of *sir2*Δ cells showed increased iRFP fluorescence during aging, consistent with increased heme abundance and HAP activity in the absence of Sir2 (Fig. 2A). The expression of HAP-regulated genes *COX5a* and *CYC1* was also increased in *sir2*Δ cells (fig. S11). The majority (83.1%) of *sir2*Δ cells aged with mode 1 phenotypes. Whereas WT mode 1 cells lost silencing at a late phase of aging, *sir2*Δ cells showed sustained loss of rDNA silencing throughout their lives, resulting in accelerated progression toward death (fig. S12) (*6*). When *HAP4*, which encodes

[1]Section of Molecular Biology, Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA. [2]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA. [3]UCSD Moores Cancer Center, University of California San Diego, La Jolla, CA 92093, USA. [4]BioCircuits Institute, University of California San Diego, La Jolla, CA 92093, USA.
*Corresponding author. Email: nhao@ucsd.edu

a major component of the HAP complex, was deleted, 89.8% of cells showed decreased rDNA-GFP fluorescence during aging, consistent with increased rDNA silencing in the absence of HAP (Fig. 2B). The majority (89.8%) of *hap4*Δ cells aged with mode 2 phenotypes and quickly approached a state with high rDNA silencing and a low abundance of heme before death.

When Hap4 expression was increased, 95.9% of cells showed decreased rDNA silencing during aging and aged with mode 1 phenotypes, consistent with an inhibition of Sir2 by HAP during aging (Fig. 2C and fig. S13). Caloric restriction (CR), which induces Hap4 expression (*20*, *21*), also increased the proportion of mode 1 cells (fig. S14). In contrast, twofold overexpression of Sir2 (fig. S15) increased the proportion of mode 2 cells with decreased heme abundance, in agreement with Sir2-mediated inhibition of HAP. Additionally, Sir2

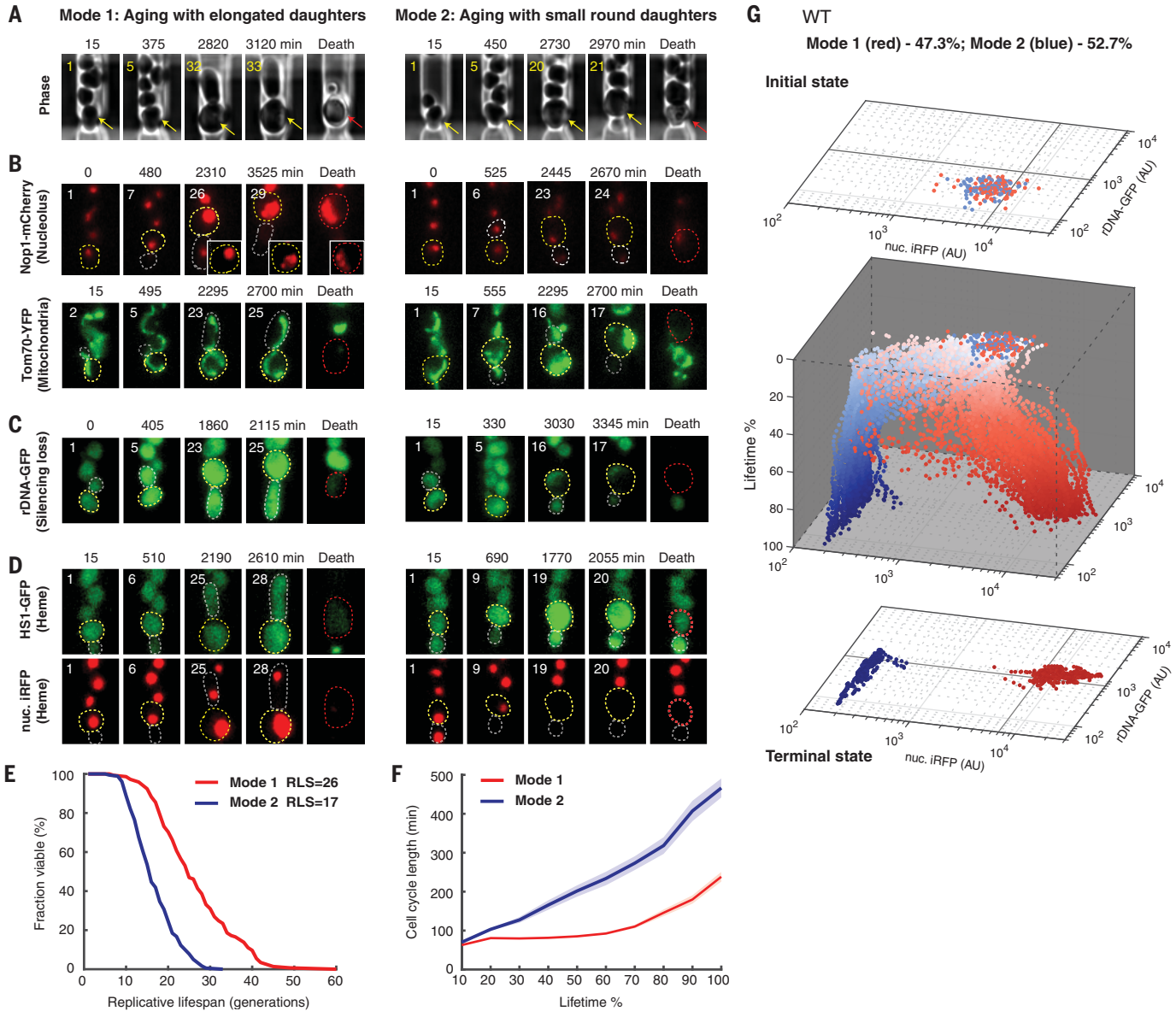


Fig. 1. Sir2 and HAP mediate divergent aging of isogenic cells. Representative time-lapse images of mode 1 and mode 2 aging processes, for (**A**) phase (n = 187 cells); (**B**) Nop1-mCherry (nucleolar marker; n = 116) and Tom70-YFP (mitochondrial marker; n = 142); (**C**) rDNA-GFP (n = 256); and (**D**) HS1-GFP and nuc. iRFP (heme reporters in the same cells; n = 230). Time-lapse images are representative of all mode 1 and mode 2 cells measured in this study. Replicative age of the mother cell is shown at the top left corner of each image. For phase images, aging and dead mother cells are marked by yellow and red arrows, respectively. In fluorescence images, aging mother cells, newborn daughter cells, and dead mother cells are circled in yellow, gray, and red, respectively. Images in Nop1-mCherry insets use different contrast to illustrate nucleolar fragmentation. (**E**) Replicative life spans (RLSs) of mode 1 and mode 2 cells (P < 0.0001, log-rank test; P < 0.0001, Gehan-Breslow-Wilcoxon test). (**F**) Changes of cell cycle length during mode 1 (red) and mode 2 (blue) aging (mode 1: n = 89; mode 2: n = 99). Shaded areas represent standard errors of the mean (SEM). (**G**) Aging trajectories of a population of WT cells (n = 187). Initial states of all the cells are projected onto a rDNA-GFP versus nuc. iRFP plane. The three-dimensional space shows the aging processes of all cells, in which the z axis represents the percentage of lifetime. Each dot represents quantified rDNA-GFP and nuc. iRFP fluorescence in a single cell at a given moment (mode 1, red; mode 2, blue). Terminal states of all the cells are projected onto another rDNA-GFP versus nuc. iRFP plane. Experiments were independently performed at least three times. AU, arbitrary units.

overexpression generated a third mode of aging, one that ended with high rDNA silencing and high heme abundance (Fig. 2D). Cells undergoing this mode of aging (mode 3) were long-lived (Fig. 2E), and they differed from mode 1 cells in that, although they also produced elongated daughters, their cell cycle lengths were unchanged throughout their life spans (Fig. 2F, fig. S16, and movie S2).

To analyze the mechanisms underlying divergent aging, we devised a model composed of two ordinary differential equations incorporating the proposed mutual inhibition between Sir2 and HAP as well as positive autoregulation of each (Fig. 3A). The mutual inhibition may act through transcriptional regulation, as the transcription of many HAP components and HAP-regulated genes is

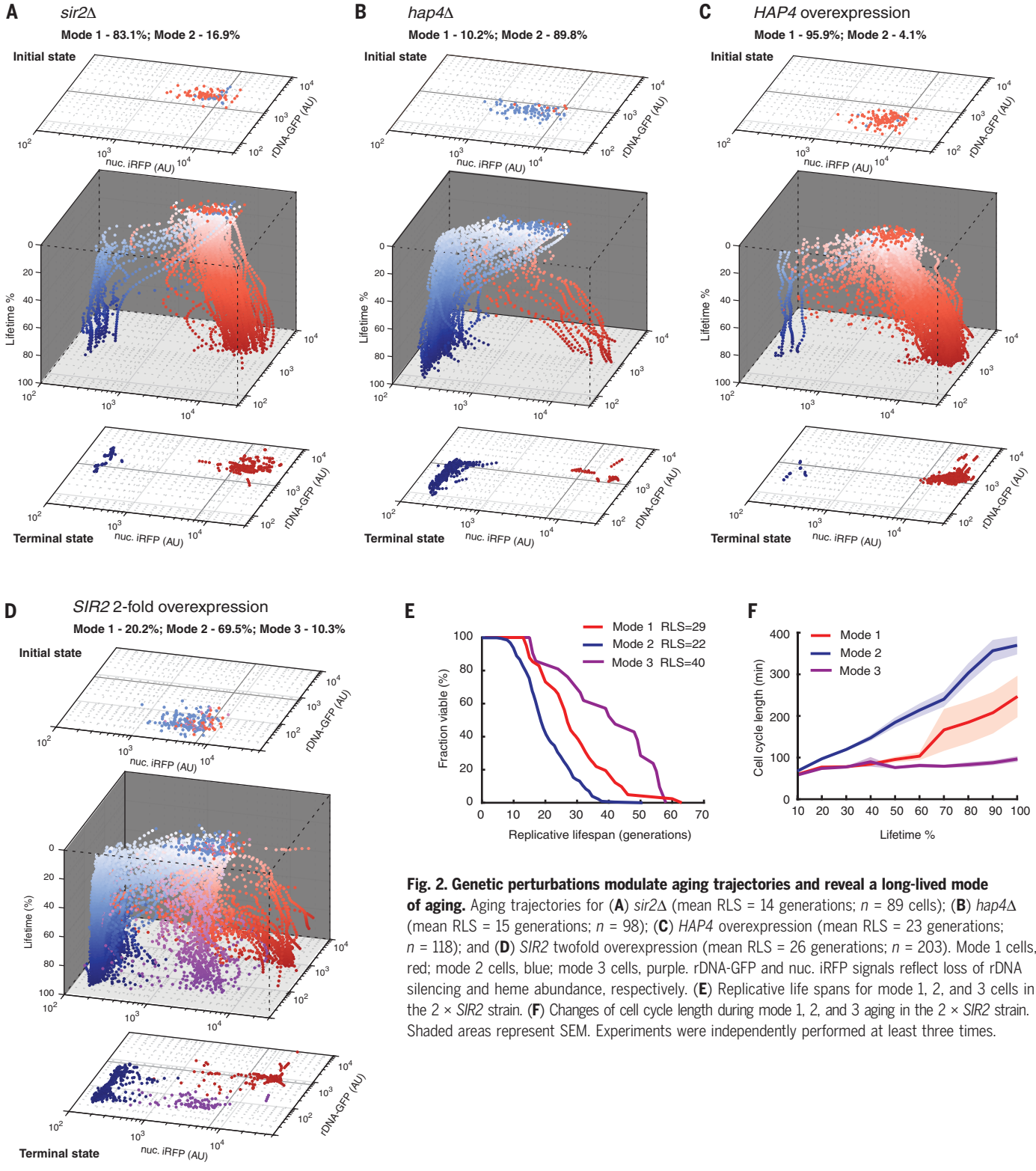


Fig. 2. Genetic perturbations modulate aging trajectories and reveal a long-lived mode of aging. Aging trajectories for (**A**) *sir2*Δ (mean RLS = 14 generations; *n* = 89 cells); (**B**) *hap4*Δ (mean RLS = 15 generations; *n* = 98); (**C**) *HAP4* overexpression (mean RLS = 23 generations; *n* = 118); and (**D**) *SIR2* twofold overexpression (mean RLS = 26 generations; *n* = 203). Mode 1 cells, red; mode 2 cells, blue; mode 3 cells, purple. rDNA-GFP and nuc. iRFP signals reflect loss of rDNA silencing and heme abundance, respectively. (**E**) Replicative life spans for mode 1, 2, and 3 cells in the 2 × *SIR2* strain. (**F**) Changes of cell cycle length during mode 1, 2, and 3 aging in the 2 × *SIR2* strain. Shaded areas represent SEM. Experiments were independently performed at least three times.

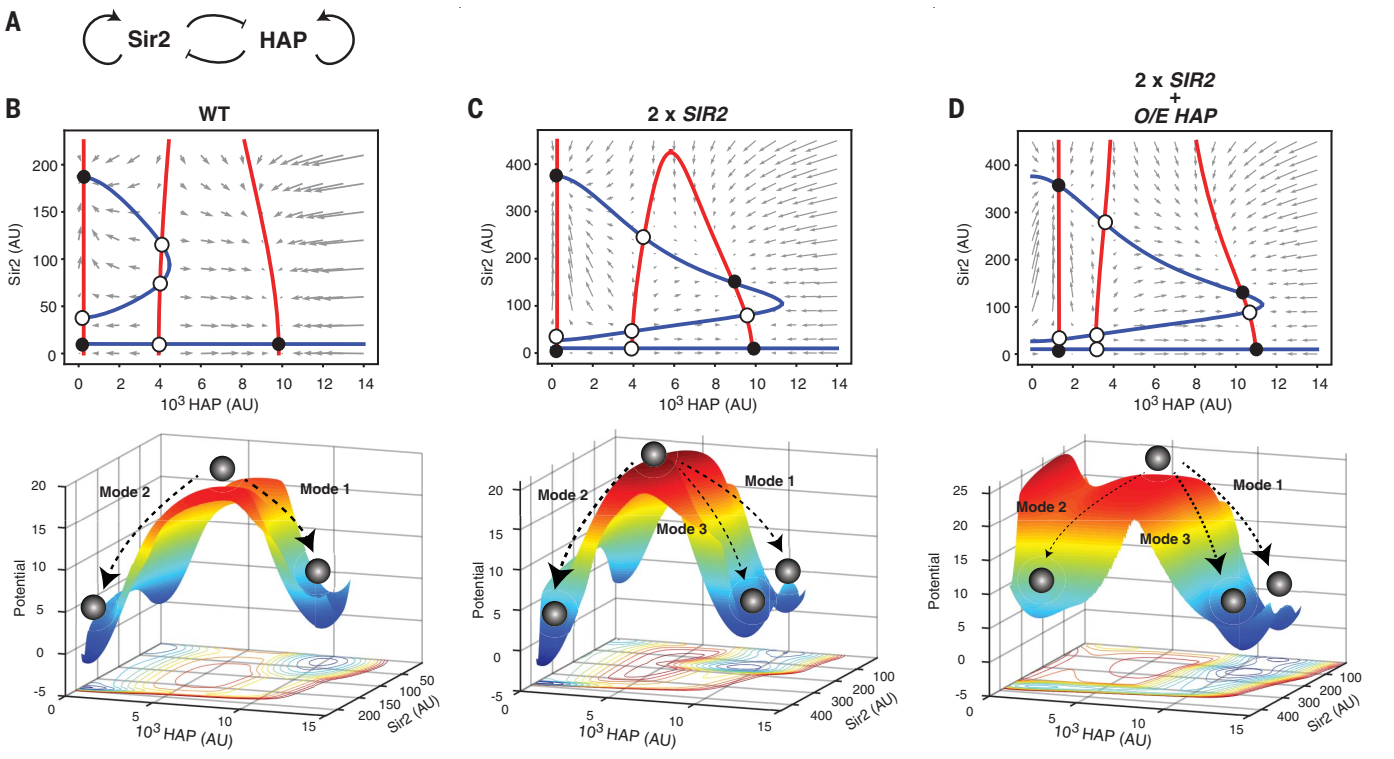


Fig. 3. Computational modeling unravels multistability of the Sir2-HAP circuit that may enable distinct modes of aging. (**A**) The diagram of the circuit topology. Graphic analyses of the model were conducted for (**B**) WT, (**C**) 2 × *SIR2*, and (**D**) 2 × *SIR2* + O/E *HAP* (overexpression). Top panels show phase plane diagrams. The nullclines for Sir2 and HAP are shown in blue and red, respectively. Gray arrows represent the vector field of the system, in which quivers show the rate and direction of movement. Fixed points are indicated with open (unstable) and filled (stable) circles. Bottom panels show computed potential landscapes. Well depths represent the probabilities of cells attracted to them. Single-cell aging processes can be depicted as a ball rolling into one of the wells on the surface.

enhanced in *sir2*Δ (*22*), and, conversely, HAP functions in transcriptional regulation of specific cofactors in Sir2 silencing complexes (*23*–*25*). Positive autoregulation could occur by Sir2 deacetylation of H4-Lys16, creating high-affinity nucleosome binding sites to promote further Sir2 recruitment (*26*–*28*). For positive feedback of HAP, heme activates the HAP complex, leading to increased expression of tricarboxylic acid cycle genes, which, in turn, increases the biosynthesis of heme (*29*). Collectively, this network topology has the potential to generate multistability (*30*).

The dynamical behavior of the system without noise can be analyzed graphically by plotting the nullclines and vector field in a Sir2-HAP phase plane. With appropriate parameters, there are seven fixed points for WT, of which three are stable (Fig. 3B, top). To address the probabilistic nature of single-cell aging, we considered a stochastic version of our model and computed an aging landscape by numerically solving the corresponding two-dimensional Fokker-Planck equation (Fig. 3B, bottom). Each of the stable fixed points of the deterministic system corresponds to a local well on the landscape, with the depth of the well reflecting its fate decision probability. Of the three wells, the deep wells with low Sir2, high HAP or high Sir2, low HAP correspond, respectively, to the experimentally observed mode 1 and mode 2 aging. The low Sir2, low HAP well is shallower (lower probability) and corresponds to a portion of mode 2 cells with low rDNA silencing (Fig. 1G) (see supplementary text section of the supplementary materials; figs. S17 to S21; and tables S1 and S2).

In the model, when the amount of Sir2 was increased twofold, its increased activity partially counteracted the inhibition from HAP, leading to the emergence of a fourth fixed stable point with high Sir2 and high HAP on the phase plane and a new well on the landscape (Fig. 3C), corresponding to experimentally observed mode 3 aging. However, Sir2 overexpression also biases the fate decision toward the high Sir2, low HAP state, consistent with the data indicating that many cells (69.5%) underwent the short-lived mode 2 aging and that the average life span of the whole population was only slightly extended (Fig. 2D). We predicted that an increase in basal HAP abundance under this condition can reshape the landscape to bias toward the high HAP states (Fig. 3D), directing most cells to mode 1 and mode 3 aging and a longer life span.

To test this prediction, we overexpressed both Sir2 and Hap4. Consistent with the model, very few cells (5.8%) underwent mode 2 aging and, instead, many (44.7%) experienced long-lived mode 3 aging (Fig. 4A and fig. S22). The combined perturbation extended the life span more than overexpression of either factor alone (Fig. 4B and table S3) and enabled maintenance of a relatively normal cell cycle length during aging (Fig. 4C). From the network perspective, the combined perturbation can partially neutralize the mutual inhibition between Sir2 and HAP, potentiating a newly observed, distinct long-lived mode with high Sir2 and HAP activities (mode 3) and resulting in a synergistic, rather than additive, effect on lifespan extension (Fig. 4B). A similar effect was observed when we combined the *fob1*Δ long-lived mutant with Hap4 overexpression (fig. S23). This mutant, similar to Sir2 overexpression, enhances rDNA stability. Our model could also help explain the synergy between CR (promoting HAP) and Sir2 (*31*), two seemingly independent longevity factors.

Cellular aging can thus be considered a fate-decision process, in which single cells age toward either silencing loss and nucleolar decline or heme depletion and mitochondrial decline. This process can be viewed as a divergent progression on a Sir2-HAP landscape, which can be reshaped by model-guided genetic

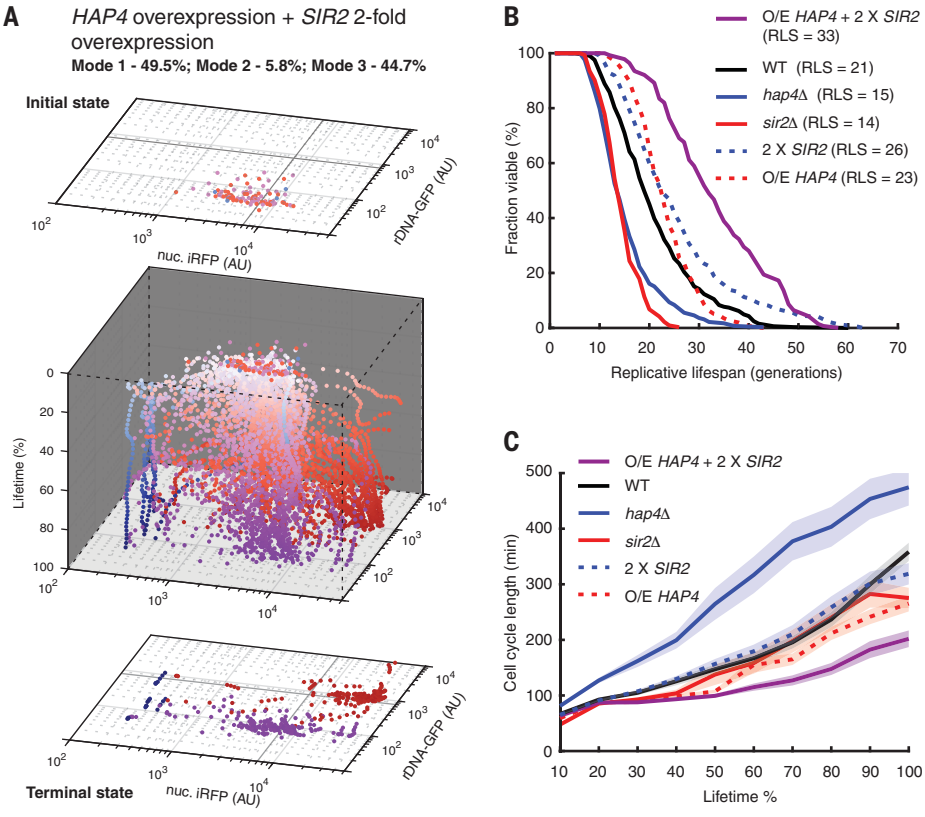


Fig. 4. Combined perturbations of Sir2 and HAP reshape the computed aging landscape and synergistically promote longevity. (**A**) Aging trajectories for the strain with combined *HAP4* overexpression and *SIR2* twofold overexpression (n = 103). Mode 1 cells, red; mode 2 cells, blue; mode 3 cells, purple. (**B**) Replicative life spans for WT, *hap4*Δ, *sir2*Δ, *SIR2* twofold overexpression (2 × *SIR2*), *HAP4* overexpression (O/E *HAP4*), and combined *HAP4* overexpression and *SIR2* twofold overexpression (O/E *HAP4* + 2 × *SIR2*). Significant numbers of life-span changes are included in table S3. (**C**) Changes of cell cycle length during aging for WT, *hap4*Δ, *sir2*Δ, 2 × *SIR2*, O/E *HAP4*, and O/E *HAP4* + 2 × *SIR2*. Shaded areas represent SEM. Experiments were independently performed at least three times.

perturbations, thereby enriching a long-lived mode of aging.

REFERENCES AND NOTES

1. B. K. Kennedy *et al*., *Cell* **159**, 709–713 (2014).
2. M. M. Crane, M. Kaeberlein, *Curr. Opin. Syst. Biol.* **8**, 25–31 (2018).
3. C. He, C. Zhou, B. K. Kennedy, *Biochim. Biophys. Acta Mol. Basis Dis.* **1864**, 2690–2696 (2018).
4. M. Kaeberlein, *Nature* **464**, 513–519 (2010).
5. K. K. Steffen, B. K. Kennedy, M. Kaeberlein, *J. Vis. Exp.* **2009** 1209 (2009).
6. Y. Li *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11253–11258 (2017).
7. M. Jin *et al*., *Cell Syst.* **8**, 242–253.e3 (2019).
8. D. A. Sinclair, K. Mills, L. Guarente, *Science* **277**, 1313–1316 (1997).
9. A. L. Hughes, D. E. Gottschling, *Nature* **492**, 261–265 (2012).
10. S. Fehrmann *et al*., *Cell Rep.* **5**, 1589–1599 (2013).
11. D. A. Sinclair, L. Guarente, *Cell* **91**, 1033–1042 (1997).
12. S. Morlot *et al*., *Cell Rep.* **28**, 408–422.e4 (2019).
13. M. R. Gartenberg, J. S. Smith, *Genetics* **203**, 1563–1599 (2016).
14. H. Atamna, D. W. Killilea, A. N. Killilea, B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14807–14812 (2002).
15. D. A. Hanna *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7539–7544 (2016).
16. G. S. Filonov *et al*., *Nat. Biotechnol.* **29**, 757–761 (2011).
17. S. Buschlen *et al*., *Comp. Funct. Genomics* **4**, 37–46 (2003).
18. K. A. Henderson, A. L. Hughes, D. E. Gottschling, *eLife* **3**, e03504 (2014).
19. J. R. Veatch, M. A. McMurray, Z. W. Nelson, D. E. Gottschling, *Cell* **137**, 1247–1258 (2009).
20. S. J. Lin *et al*., *Nature* **418**, 344–348 (2002).
21. J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **278**, 680–686 (1997).
22. A. Ellahi, D. M. Thurtle, J. Rine, *Genetics* **200**, 505–521 (2015).
23. Z. Hu, P. J. Killion, V. R. Iyer, *Nat. Genet.* **39**, 683–687 (2007).
24. J. Reimand, J. M. Vaquerizas, A. E. Todd, J. Vilo, N. M. Luscombe, *Nucleic Acids Res.* **38**, 4768–4777 (2010).
25. M. Becerra *et al*., *Mol. Microbiol.* **43**, 545–555 (2002).
26. L. Pillus, J. Rine, *Cell* **59**, 637–647 (1989).
27. D. Moazed, *Cell* **146**, 510–518 (2011).
28. K. Sneppen, I. B. Dodd, *Epigenetics* **10**, 293–302 (2015).
29. T. Zhang, P. Bu, J. Zeng, A. Vancura, *J. Biol. Chem.* **292**, 16942–16954 (2017).
30. F. Wu, R. Q. Su, Y. C. Lai, X. Wang, *eLife* **6**, e23702 (2017).
31. M. Kaeberlein, K. T. Kirkland, S. Fields, B. K. Kennedy, *PLOS Biol.* **2**, e296 (2004).
32. yaj030, yaj030/aging_science2020 v1.1, Version 1.1,Zenodo (2020); http://doi.org/10.5281/zenodo.3770529.

ACKNOWLEDGMENTS

We thank G. Zhu for help with image analysis, A. R. Reddi (Georgia Tech) and D. E. Gottschling (Calico) for generously providing strains and reagents, and E. L. Petty for guidance with CNV analysis. **Funding:** This work was supported by the National Institutes of Health, National Institute on Aging, AG056440 (to N.H., J.H., L.P., and L.S.T.); NSF MCB-1616127 (N.H.); and NSF MCB-1716841 (L.P.). **Author contributions:** Conceptualization, Y.L., Y.J., L.P., L.S.T., J.H., and N.H.; Methodology, Y.L., Y.J., R.O., and S.K.; Investigation, Y.L., J.P., and S.K.; Formal Analysis, Y.L., Y.J., Y.Z., and L.S.T.; Writing – Original Draft, Y.L., Y.J., and N.H.; Writing – Review & Editing, Y.L., Y.J., J.P., R.O., L.P., L.S.T., J.H., and N.H.; Resources, L.P., L.S.T., J.H., and N.H.; Funding Acquisition, L.P., L.S.T., J.H., and N.H.; Supervision, N.H.; Project Administration, N.H. **Competing interests:** J.H. is a founder of GenCirq and Quantitative BioSciences, which focus on cancer therapeutics and agricultural synthetic biology, respectively; there is no competing interest with this work. The other authors also declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. The code from this work is available at https://github.com/yaj030/aging_science2020 and Zenodo (*32*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/325/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 to S6
References (*33*–*37*)
MDAR Reproducibility Checklist
Movies S1 and S2

7 May 2019; resubmitted 24 January 2020
Accepted 13 May 2020
10.1126/science.aax9552

CORONAVIRUS

Site-specific glycan analysis of the SARS-CoV-2 spike

Yasunori Watanabe[1,2,3]*, Joel D. Allen[1]*, Daniel Wrapp[4], Jason S. McLellan[4], Max Crispin[1]†

The emergence of the betacoronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus 2019 (COVID-19), represents a considerable threat to global human health. Vaccine development is focused on the principal target of the humoral immune response, the spike (S) glycoprotein, which mediates cell entry and membrane fusion. The SARS-CoV-2 S gene encodes 22 N-linked glycan sequons per protomer, which likely play a role in protein folding and immune evasion. Here, using a site-specific mass spectrometric approach, we reveal the glycan structures on a recombinant SARS-CoV-2 S immunogen. This analysis enables mapping of the glycan-processing states across the trimeric viral spike. We show how SARS-CoV-2 S glycans differ from typical host glycan processing, which may have implications in viral pathobiology and vaccine design.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative pathogen of coronavirus 2019 (COVID-19) (*1*, *2*), induces fever, severe respiratory illness, and pneumonia. SARS-CoV-2 uses an extensively glycosylated spike (S) protein that protrudes from the viral surface to bind to angiotensin-converting enzyme 2 (ACE2) to mediate host-cell entry (*3*). The S protein is a trimeric class I fusion protein, composed of two functional subunits, responsible for receptor binding (S1 subunit) and membrane fusion (S2 subunit) (*4*, *5*). The surface of the envelope spike is dominated by host-derived glycans, with each trimer displaying 66 N-linked glycosylation sites. The S protein is a key target in vaccine design efforts (*6*), and understanding the glycosylation of recombinant viral spikes can reveal fundamental features of viral biology and guide vaccine design strategies (*7*, *8*).

Viral glycosylation has wide-ranging roles in viral pathobiology, including mediating protein folding and stability and shaping viral tropism (*9*). Glycosylation sites are under selective pressure as they facilitate immune evasion by shielding specific epitopes from antibody neutralization. However, we note the low mutation rate of SARS-CoV-2 and that as yet, there have been no observed mutations to N-linked glycosylation sites (*10*). Surfaces with an unusually high density of glycans can also enable immune recognition (*9*, *11*, *12*). The role of glycosylation in camouflaging immunogenic protein epitopes has been studied for other coronaviruses (*10*, *13*, *14*). Coronaviruses form virions by budding into the lumen of endoplasmic reticulum–Golgi intermediate compartments (*15*, *16*). However, observations of complex-type glycans on virally derived material suggests that the viral glycoproteins are subjected to Golgi-resident processing enzymes (*13*, *17*).

High viral glycan density and local protein architecture can sterically impair the glycan maturation pathway. Impaired glycan maturation resulting in the presence of oligomannose-type glycans can be a sensitive reporter of native-like protein architecture (*8*), and site-specific glycan analysis can be used to compare different immunogens and monitor manufacturing processes (*18*). Additionally, glycosylation can influence the trafficking of recombinant immunogen to germinal centers (*19*).

To resolve the site-specific glycosylation of the SARS-CoV-2 S protein and visualize the distribution of glycoforms across the protein surface, we expressed and purified three biological replicates of recombinant soluble material in an identical manner to that which was used to obtain the high-resolution cryo–electron microscopy (cryo-EM) structure, albeit without a glycan-processing blockade using kifunensine (*4*). This variant of the S protein contains all 22 glycans on the SARS-CoV-2 S protein (Fig. 1A). Stabilization of the trimeric prefusion structure was achieved by using the 2P stabilizing mutations (*20*) at residues 986 and 987, a GSAS (Gly-Ser-Ala-Ser) substitution at the furin cleavage site (residues 682 to 685), and a C-terminal trimerization motif. This helps to maintain quaternary architecture during glycan processing. Before analysis, supernatant containing the recombinant SARS-CoV-2 S was purified by size exclusion chromatography to ensure

[1]School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK. [2]Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. [3]Division of Structural Biology, University of Oxford, Wellcome Centre for Human Genetics, Oxford OX3 7BN, UK. [4]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: max.crispin@soton.ac.uk

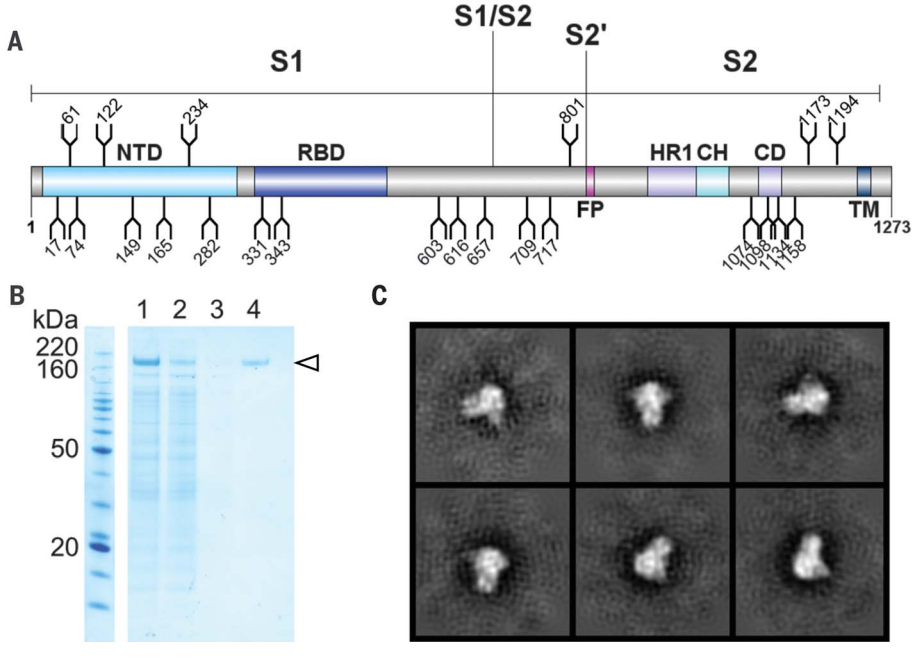


Fig. 1. Expression and validation of the SARS-CoV-2 S glycoprotein. (**A**) Schematic representation of the SARS-CoV-2 S glycoprotein. The positions of N-linked glycosylation sequons (N-X-S/T, where X ≠ P) are shown as branches (N, Asn; X, any residue; S, Ser; T, Thr; P, Pro). Protein domains are illustrated: N-terminal domain (NTD), receptor binding domain (RBD), fusion peptide (FP), heptad repeat 1 (HR1), central helix (CH), connector domain (CD), and transmembrane domain (TM). (**B**) SDS–polyacrylamide gel electrophoresis analysis of the SARS-CoV-2 S protein (indicated by the arrowhead) expressed in human embryonic kidney (HEK) 293F cells. Lane 1: filtered supernatant from transfected cells; lane 2: flow-through from StrepTactin resin; lane 3: wash from StrepTactin resin; lane 4: elution from StrepTactin resin. (**C**) Negative-stain EM 2D class averages of the SARS-CoV-2 S protein. 2D class averages of the SARS-CoV-2 S protein are shown, confirming that the protein adopts the trimeric prefusion conformation matching the material used to determine the structure (*4*).

that only native-like trimeric protein was analyzed (Fig. 1B and fig. S1). The trimeric conformation of the purified material was validated by using negative-stain EM (Fig. 1C).

To determine the site-specific glycosylation of SARS-CoV-2 S, we used trypsin, chymotrypsin, and α-lytic protease to generate three glycopeptide samples. These proteases were selected to generate glycopeptides that contain a single N-linked glycan sequon. The glycopeptides were analyzed by liquid chromatography–mass spectrometry, and the glycan compositions were determined for all 22 N-linked glycan sites (Fig. 2). To convey the main processing features at each site, the abundances of each glycan are summed into oligomannose-type, hybrid-type, and categories of complex-type glycosylation based on branching and fucosylation. The detailed, expanded graphs showing the diverse range of glycan compositions are presented in table S1 and fig. S2.

Two sites on SARS-CoV-2 S are principally oligomannose-type: N234 and N709. The predominant oligomannose-type glycan structure observed across the protein, with the exception of N234, is $Man_5GlcNAc_2$ (Man, mannose; GlcNAc, N-acetylglucosamine), which demonstrates that these sites are largely accessible to α-1,2-mannosidases but are poor substrates for GlcNAcT-I, which is the gateway enzyme in the formation of hybrid- and complex-type glycans in the Golgi apparatus. The stage at which processing is impeded is a signature related to the density and presentation of glycans on the viral spike. For example, the more densely glycosylated spikes of HIV-1 Env and Lassa virus (LASV) GPC exhibit numerous sites dominated by $Man_9GlcNAc_2$ (*21*–*24*).

A mixture of oligomannose- and complex-type glycans can be found at sites N61, N122, N603, N717, N801, and N1074 (Fig. 2). Of the 22 sites on the S protein, 8 contain substantial populations of oligomannose-type glycans, highlighting how the processing of the SARS-CoV-2 S glycans is divergent from host glycoproteins (*25*). The remaining 14 sites are dominated by processed, complex-type glycans.

Although unoccupied glycosylation sites were detected on SARS-CoV-2 S, when quantified they were revealed to form a very minor component of the total peptide pool (table S2). In HIV-1 immunogen research, the holes generated by unoccupied glycan sites have been shown to be immunogenic and potentially give rise to distracting epitopes (*26*). The high occupancy of N-linked glycan sequons of SARS-CoV-2 S indicates that recombinant immunogens will not require further optimization to enhance site occupancy.

Using the cryo-EM structure of the trimeric SARS-CoV-2 S protein [Protein Data Bank (PDB) ID 6VSB] (*4*), we mapped the glycosylation status of the coronavirus spike mimetic onto the experimentally determined three-dimensional (3D) structure (Fig. 3). This combined mass spectrometric and cryo-EM analysis reveals how the N-linked glycans occlude distinct regions across the surface of the SARS-CoV-2 spike.

Shielding of the receptor binding sites on the SARS-CoV-2 spike by proximal glycosylation sites (N165, N234, N343) can be observed, especially when the receptor binding domain is in the "down" conformation. The shielding of receptor binding sites by glycans is a common feature of viral glycoproteins, as observed on SARS-CoV-1 S (*10*, *13*), HIV-1 Env

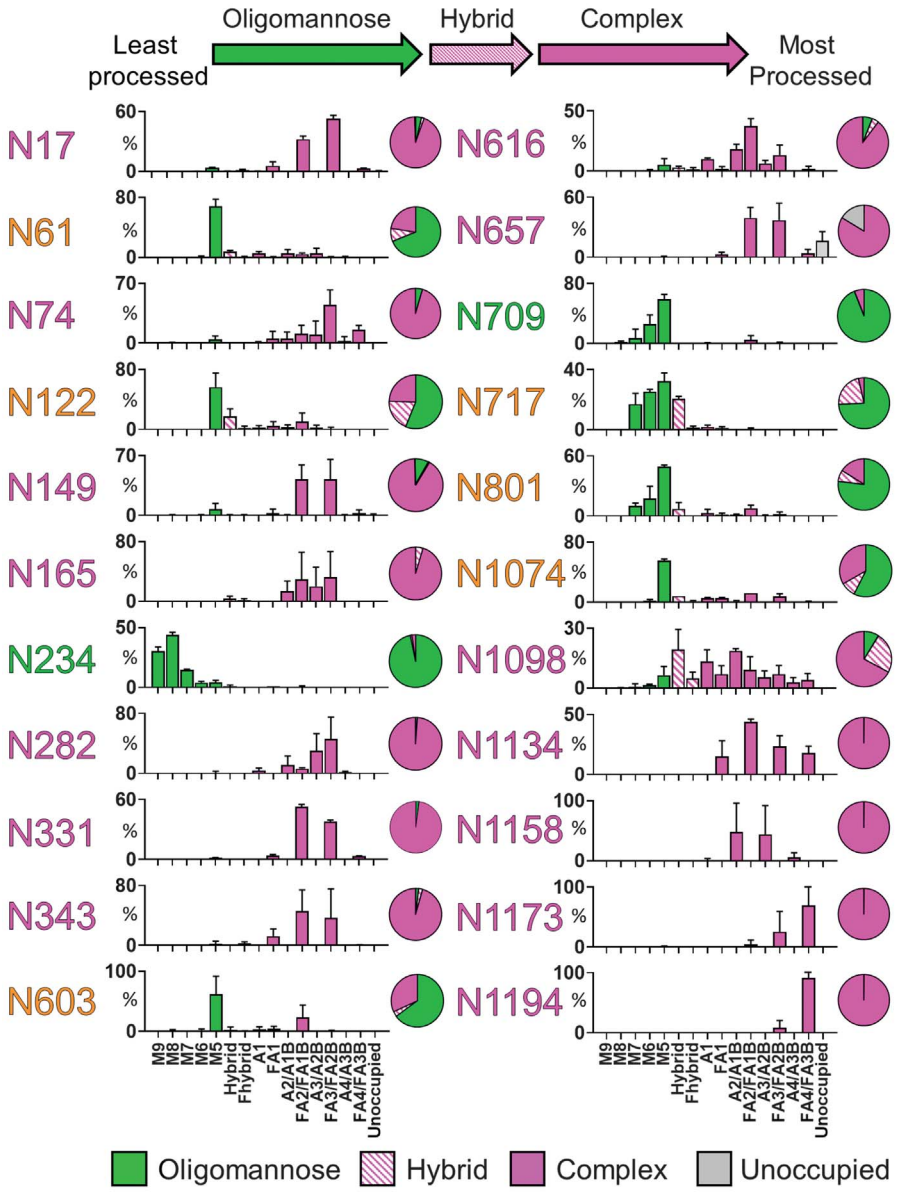


Fig. 2. Site-specific N-linked glycosylation of the SARS-CoV-2 S glycoprotein. The schematic illustrates the color code for the principal glycan types that can arise along the maturation pathway from oligomannose- to hybrid- to complex-type glycans. The graphs summarize quantitative mass spectrometric analysis of the glycan population present at individual N-linked glycosylation sites simplified into categories of glycans. The oligomannose-type glycan series (M9 to M5; $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$) is colored green, afucosylated and fucosylated hybrid-type glycans (hybrid and F hybrid) are dashed pink, and complex glycans are grouped according to the number of antennae and presence of core fucosylation (A1 to FA4) and are colored pink. Unoccupancy of an N-linked glycan site is represented in gray. The pie charts summarize the quantification of these glycans. Glycan sites are colored according to oligomannose-type glycan content, with the glycan sites labeled in green (80 to 100%), orange (30 to 79%), and pink (0 to 29%). An extended version of the site-specific analysis showing the heterogeneity within each category can be found in table S1 and fig. S2. The bar graphs represent the mean quantities of three biological replicates, with error bars representing the standard error of the mean.

(*27*), influenza hemagglutinin (*28*, *29*), and LASV GPC (*24*). Given the functional constraints of receptor binding sites and the resulting low mutation rates of these residues, there is likely selective pressure to use N-linked glycans to camouflage one of the most conserved and potentially vulnerable areas of their respective glycoproteins (*30*, *31*).

We note the dispersion of oligomannose-type glycans across both the S1 and S2 subunits. This is in contrast to other viral glycoproteins; for example, the dense glycan clusters in several strains of HIV-1 Env induce oligomannose-type glycans that are recognized by antibodies (*32*, *33*). In SARS-CoV-2 S, the oligomannose-type structures are likely protected by the protein component, as exemplified by the N234 glycan, which is partially sandwiched between the N-terminal and receptor binding domains (Fig. 3).

We characterized the N-linked glycans on extended flexible loop structures (N74 and N149) and at the membrane-proximal C terminus (N1158, N1173, N1194) that were not resolved in the cryo-EM maps (*4*). These were determined to be complex-type glycans, consistent with steric accessibility of these residues.

Whereas the oligomannose-type glycan content (28%) (table S2) is above that observed on typical host glycoproteins, it is lower than other viral glycoproteins. For example, one of the most densely glycosylated viral spike proteins is HIV-1 Env, which exhibits ~60% oligomannose-type glycans (*21*, *34*). This suggests that the SARS-CoV-2 S protein is less densely glycosylated and that the glycans form less of a shield compared with other viral glycoproteins, including HIV-1 Env and LASV GPC, which may be beneficial for the elicitation of neutralizing antibodies.

Additionally, the processing of complex-type glycans is an important consideration in immunogen engineering, especially considering

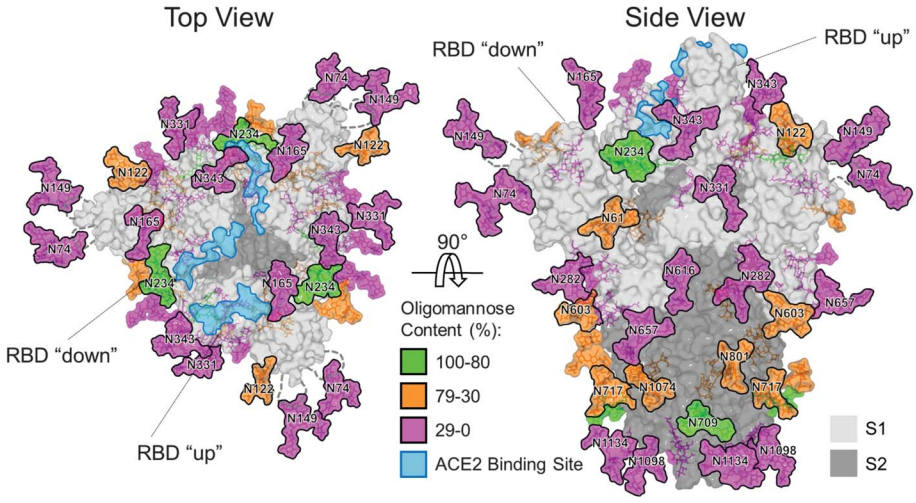


Fig. 3. Structure-based mapping of SARS-CoV-2 S N-linked glycans. Representative glycans are modeled onto the prefusion structure of the trimeric SARS-CoV-2 S glycoprotein (PDB ID 6VSB) (*4*), with one RBD in the "up" conformation and the other two RBDs in the "down" conformation. The glycans are colored according to oligomannose content as defined by the key. ACE2 receptor binding sites are highlighted in light blue. The S1 and S2 subunits are rendered with translucent surface representation, colored light and dark gray, respectively. The flexible loops on which the N74 and N149 glycan sites reside are represented as gray dashed lines, with glycan sites on the loops mapped at their approximate regions.

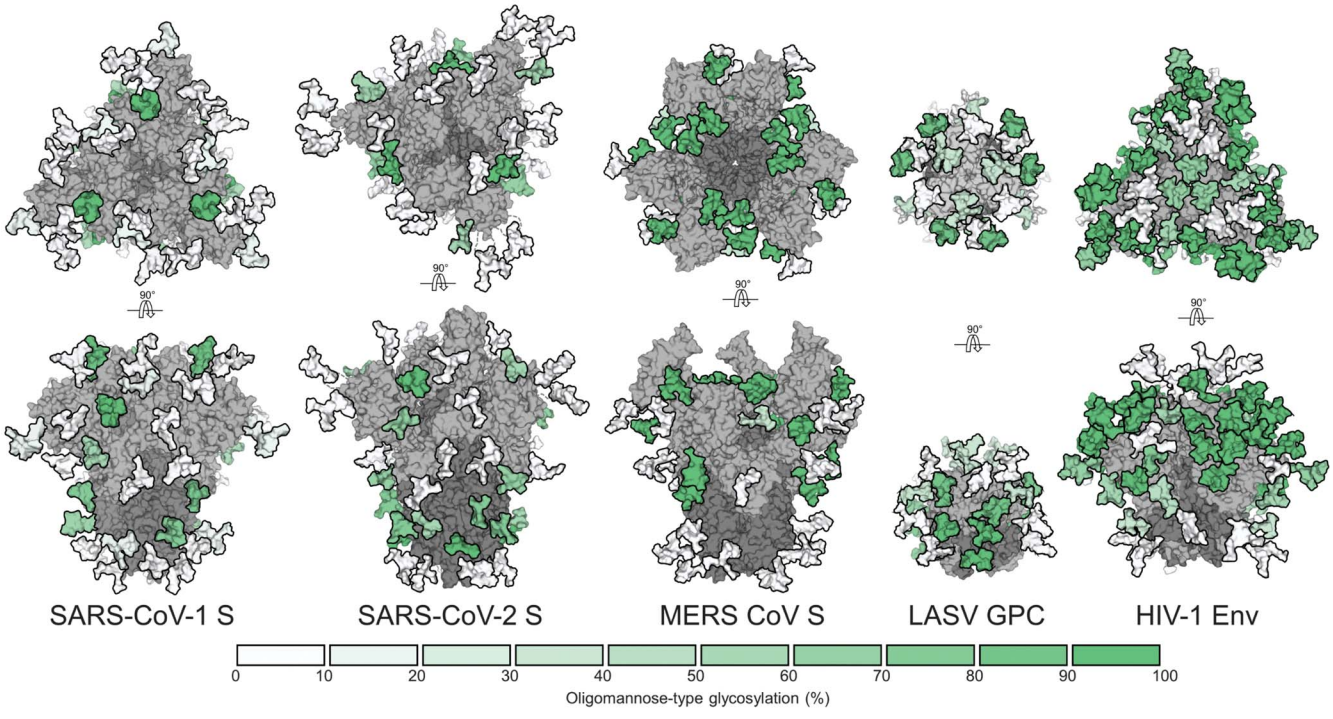


Fig. 4. Underprocessing of viral glycan shields. From left to right, MERS-CoV S (*10*), SARS-CoV-1 S (*10*), SARS-CoV-2 S, LASV GPC (*24*), and HIV-1 Env (*8*, *21*). Site-specific N-linked glycan oligomannose quantifications are colored according to the key. All glycoproteins were expressed as soluble trimers in HEK 293F cells apart from LASV GPC, which was derived from virus-like particles from Madin-Darby canine kidney II cells.

that epitopes of neutralizing antibodies against SARS-CoV-2 S can contain fucosylated glycans at N343 (*35*). Across the 22 N-linked glycosylation sites, 52% are fucosylated and 15% of the glycans contain at least one sialic acid residue (table S2 and fig. S3). Our analysis reveals that N343 is highly fucosylated with 98% of detected glycans bearing fucose residues. Glycan modifications can be heavily influenced by the cellular expression system used. We have previously demonstrated for HIV-1 Env glycosylation that the processing of complex-type glycans is driven by the producer cell but that the levels of oligomannose-type glycans were largely independent of the expression system and are much more closely related to the protein structure and glycan density (*36*).

Highly dense glycan shields, such as those observed on LASV GPC and HIV-1 Env, feature so-called mannose clusters (*22*, *24*) on the protein surface (Fig. 4). Whereas small mannose-type clusters have been characterized on the S1 subunit of Middle East respiratory syndrome (MERS)–CoV S (*10*), no such phenomenon has been observed for the SARS-CoV-1 or SARS-CoV-2 S proteins. The site-specific glycosylation analysis reported here suggests that the glycan shield of SARS-CoV-2 S is consistent with other coronaviruses and similarly exhibits numerous vulnerabilities throughout the glycan shield (*10*). Last, we detected trace levels of O-linked glycosylation at Thr^{323}/Ser^{325} (T323/S325), with over 99% of these sites unmodified (fig. S4), suggesting that O-linked glycosylation of this region is minimal when the structure is native-like.

Our glycosylation analysis of SARS-CoV-2 offers a detailed benchmark of site-specific glycan signatures characteristic of a natively folded trimeric spike. As an increasing number of glycoprotein-based vaccine candidates are being developed, their detailed glycan analysis offers a route for comparing immunogen integrity and will also be important to monitor as manufacturing processes are scaled for clinical use. Glycan profiling will therefore also be an important measure of antigen quality in the manufacture of serological testing kits. Last, with the advent of nucleotide-based vaccines, it will be important to understand how those delivery mechanisms affect immunogen processing and presentation.

REFERENCES AND NOTES

1. C. Huang *et al.*, *Lancet* **395**, 497–506 (2020).
2. X. Yang *et al.*, *Lancet Respir. Med.* 10.1016/S2213-2600(20)30079-5 (2020).
3. M. Letko, A. Marzi, V. Munster, *Nat. Microbiol.* **5**, 562–569 (2020).
4. D. Wrapp *et al.*, *Science* **367**, 1260–1263 (2020).
5. A. C. Walls *et al.*, *Cell* **181**, 281–292.e6 (2020).
6. F. Amanat, F. Krammer, *Immunity* **52**, 583–589 (2020).
7. L. Cao *et al.*, *Nat. Commun.* **8**, 14954 (2017).
8. A.-J. Behrens *et al.*, *J. Virol.* **91**, e01894–e16 (2017).
9. Y. Watanabe, T. A. Bowden, I. A. Wilson, M. Crispin, *Biochim. Biophys. Acta* **1863**, 1480–1497 (2019).
10. Y. Watanabe *et al.*, *Nat. Commun.* **11**, 2688 (2020).
11. M. Dalziel, M. Crispin, C. N. Scanlan, N. Zitzmann, R. A. Dwek, *Science* **343**, 1235681 (2014).
12. C. N. Scanlan, J. Offer, N. Zitzmann, R. A. Dwek, *Nature* **446**, 1038–1045 (2007).
13. A. C. Walls *et al.*, *Cell* **176**, 1026–1039.e15 (2019).
14. T. J. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1438–1446 (2020).
15. S. Stertz *et al.*, *Virology* **361**, 304–315 (2007).
16. P. Venkatagopalan, S. M. Daskalova, L. A. Lopez, K. A. Dolezal, B. G. Hogue, *Virology* **478**, 75–85 (2015).
17. G. Ritchie *et al.*, *Virology* **399**, 257–269 (2010).
18. A. A. Hargett, M. B. Renfrow, *Curr. Opin. Virol.* **36**, 56–66 (2019).
19. T. Tokatlian *et al.*, *Science* **363**, 649–654 (2019).
20. J. Pallesen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7348–E7357 (2017).
21. W. B. Struwe *et al.*, *Cell Rep.* **24**, 1958–1966.e5 (2018).
22. A.-J. Behrens *et al.*, *Cell Rep.* **14**, 2695–2706 (2016).
23. M. Panico *et al.*, *Sci. Rep.* **6**, 32956 (2016).
24. Y. Watanabe *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7320–7325 (2018).
25. I. Loke, D. Kolarich, N. H. Packer, M. Thaysen-Andersen, *Mol. Aspects Med.* **51**, 31–55 (2016).
26. M. Bianchi *et al.*, *Immunity* **49**, 288–300.e8 (2018).
27. J. Jardine *et al.*, *Science* **340**, 711–716 (2013).
28. C.-J. Wei *et al.*, *Sci. Transl. Med.* **2**, 24ra21 (2010).
29. R. Xu *et al.*, *Science* **328**, 357–360 (2010).
30. X. Wei *et al.*, *Nature* **422**, 307–312 (2003).
31. M. Zhang *et al.*, *Glycobiology* **14**, 1229–1246 (2004).
32. G. B. E. Stewart-Jones *et al.*, *Cell* **165**, 813–826 (2016).
33. D. Sok *et al.*, *Sci. Transl. Med.* **6**, 236ra63 (2014).
34. L. Cao *et al.*, *Nat. Commun.* **9**, 3693 (2018).
35. D. Pinto *et al.*, *Nature* 10.1038/s41586-020-2349-y (2020).
36. L. K. Pritchard, D. J. Harvey, C. Bonomelli, M. Crispin, K. J. Doores, *J. Virol.* **89**, 8932–8944 (2015).
37. M. Crispin, SARS-CoV-2 spike site-specific N-linked glycan analysis. MassIVE Database (2020); https://doi.org/10.25345/C54X4K.

ACKNOWLEDGMENTS

We thank M. Dixon and M. Gowland-Pryde for supporting our work on this project during the difficulties arising from the pandemic and G. Ould for critical reading of the manuscript. **Funding:** This work was funded by the International AIDS Vaccine Initiative, Bill and Melinda Gates Foundation through the Collaboration for AIDS Vaccine Discovery (OPP1084519 and 1196345 to M.C.), the NIAID (R01-AI127521 to J.S.M.), and the Scripps Consortium for HIV Vaccine Development (CHAVD) (AI144462 to M.C.). M.C. is a Supernumerary Fellow of Oriel College, Oxford, and professor adjunct at Scripps Research, CA. **Author contributions:** Y.W. and J.D.A. performed mass spectrometry experiments and analyzed data. Y.W. built glycosylated models. J.S.M. and M.C. supervised experiments. Y.W., J.D.A., and D.W. expressed and purified proteins. Y.W., J.D.A., and M.C. wrote the manuscript with input from all authors. **Competing interests:** J.S.M. is an inventor on U.S. patent application no. 62/412,703 ("Prefusion Coronavirus Spike Proteins and Their Use"), and D.W. and J.S.M. are inventors on U.S. patent application no. 62/972,886 ("2019-nCoV Vaccine"). **Data and materials availability:** Mass spectrometry raw files have been deposited in the MassIVE proteomics database (*37*). The plasmid is available from J.S.M. under a material transfer agreement with The University of Texas at Austin. 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/369/6501/330/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 and S2
References (*38*, *39*)
MDAR Reproducibility Checklist
Data S1

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

31 March 2020; accepted 29 April 2020
Published online 4 May 2020
10.1126/science.abb9983

GENOME EDITING

CRISPR-CasΦ from huge phages is a hypercompact genome editor

Patrick Pausch[1,2]*, Basem Al-Shayeb[1,3]*, Ezra Bisom-Rapp[4], Connor A. Tsuchida[1,5], Zheng Li[6], Brady F. Cress[1,2], Gavin J. Knott[1,2,7], Steven E. Jacobsen[6,8], Jillian F. Banfield[1,9], Jennifer A. Doudna[1,2,8,10,11,12]†

CRISPR-Cas systems are found widely in prokaryotes, where they provide adaptive immunity against virus infection and plasmid transformation. We describe a minimal functional CRISPR-Cas system, comprising a single ~70-kilodalton protein, CasΦ, and a CRISPR array, encoded exclusively in the genomes of huge bacteriophages. CasΦ uses a single active site for both CRISPR RNA (crRNA) processing and crRNA-guided DNA cutting to target foreign nucleic acids. This hypercompact system is active in vitro and in human and plant cells with expanded target recognition capabilities relative to other CRISPR-Cas proteins. Useful for genome editing and DNA detection but with a molecular weight half that of Cas9 and Cas12a genome-editing enzymes, CasΦ offers advantages for cellular delivery that expand the genome editing toolbox.

Competition between viruses and their host microbes fostered the evolution of CRISPR-Cas systems that employ nucleases and noncoding CRISPR RNAs (crRNAs) to target foreign nucleic acids by complementary base pairing (*1*). Processing of CRISPR array transcripts, which consist of repeats and spacer sequences acquired from viruses or other mobile genetic elements (MGEs) (*2*), generates mature crRNAs that guide Cas proteins to detect and destroy previously encountered viruses (*3*). Although found almost exclusively in microbial genomes, the recent discovery of ubiquitous huge bacteriophages (viruses of bacteria) revealed the surprising prevalence of CRISPR-Cas systems encoded in their genomes (*4*). These systems notably lack CRISPR spacer acquisition machinery (Cas1, Cas2, and Cas4 proteins) and generally harbor compact CRISPR arrays (median of five spacers per array), some of which target the genes of competing phages or phage hosts. CasΦ (Cas12j) is a family of Cas proteins encoded in the Biggiephage clade (*4*). CasΦ contains a C-terminal RuvC domain with remote homology to that of the TnpB nuclease superfamily from which type V CRISPR-Cas proteins are thought to have evolved (*4*, *5*) (fig. S1). However, CasΦ shares <7% amino acid identity with other type V CRISPR-Cas proteins and is most closely related to a TnpB group distinct from miniature type V (Cas14) proteins (Fig. 1A).

CasΦ's unusually small size of ~70 to 80 kDa, about half the size of Cas9 and Cas12a (Fig. 1B), and its lack of co-occurring genes raised the question of whether CasΦ functions as a bona fide CRISPR-Cas system. We investigated three divergent CasΦ orthologs from metagenomic assemblies (fig. S2), here referred to as CasΦ-1, CasΦ-2, and CasΦ-3. To examine CasΦ's ability to recognize and target DNA in bacterial cells, we tested whether CasΦ could protect *Escherichia coli* from plasmid transformation. CRISPR-Cas systems target DNA sequences that follow or precede a 2– to 5–base-pair (bp) protospacer-adjacent motif (PAM) for self versus nonself discrimination (*6*). To determine whether CasΦ uses a PAM, we transformed a library of plasmids containing randomized regions adjacent to crRNA-complementary target sites, thereby depleting plasmids harboring functional PAMs. This revealed the crRNA-guided double-stranded DNA (dsDNA)–targeting capability of CasΦ and minimal T-rich PAM sequences, including 5′-TBN-3′ PAMs (where B is G, T, or C) depleted for CasΦ-2 (Fig. 1C).

We next used the *E. coli* expression system and plasmid interference assay to determine the components required for CRISPR-CasΦ system function. RNA sequencing (RNA-seq) analysis revealed transcription of the *cas*Φ gene and the reduced CRISPR array but no evidence of other noncoding RNA, such as a trans-activating CRISPR RNA (tracrRNA) within the locus (Fig. 1D). In addition, CasΦ activity could be readily reprogrammed to target other plasmid sequences by altering the guide RNA (fig. S3). These findings suggest that in its native environment, CasΦ is a functional phage protein and bona fide CRISPR-Cas effector capable of cleaving crRNA-complementary DNA, such as other phages (Fig. 1E). Furthermore, these results demonstrate that this single-RNA system is much more compact than other active CRISPR-Cas systems (Fig. 1F).

We next investigated the DNA recognition and cleavage requirements of CasΦ in vitro. RNA-seq revealed that the crRNA spacer, which is complementary to DNA targets, is 14 to 20 nucleotides (nt) long (Fig. 1D). Incubation of purified CasΦ (fig. S4) with crRNAs of different spacer sizes along with supercoiled plasmid or linear dsDNA revealed that DNA cleavage requires the presence of a cognate PAM and a spacer of ≥14 nt (Fig. 2A; fig. S5A). Analysis of the cleavage products showed that CasΦ generated staggered 5′-overhangs of 8 to 12 nt (Fig. 2, B and C, and fig. S5, B and C), similar to the staggered DNA cuts observed for other type V CRISPR-Cas enzymes, including Cas12a and CasX (*7*, *8*). We also observed that CasΦ-2 and CasΦ-3 were more active in vitro than CasΦ-1, and the non–target strand (NTS) was cleaved faster than the target strand (TS) within the RuvC active site (Fig. 2D; figs. S6A and S7; supplementary text). Furthermore, CasΦ was found to cleave single-stranded DNA (ssDNA) but not ssRNA in cis or in trans (figs. S6B and S8), suggesting that CasΦ may also target ssDNA MGEs or ssDNA intermediates. The trans-cleavage activity of CasΦ, observed only upon DNA recognition in cis (fig. S8), coupled with a minimal PAM requirement (Fig. 1C), may be useful for broader nucleic acid detection as previously demonstrated for type V and type VI Cas proteins (*9*–*11*).

CRISPR-CasΦ systems must produce mature crRNA to guide foreign DNA cleavage. Other type V CRISPR-Cas proteins process pre-crRNAs by using an internal active site distinct from the RuvC domain (*12*) or by recruiting ribonuclease III to cleave a pre-crRNA–tracrRNA duplex (*13*–*16*). The absence of a detectable tracrRNA for CasΦ hinted that CasΦ may catalyze crRNA maturation on its own. To test this possibility, we incubated purified CasΦ with substrates designed to mimic the pre-crRNA structure (Fig. 3A). Reaction products corresponding to a 26- to 29-nt-long repeat and 20-nt spacer sequence of the crRNA were observed only in the presence of wild-type CasΦ; this was corroborated by RNA-seq analysis of native loci (Figs. 1D and 3, A and C; fig. S9). In control experiments, we found that pre-crRNA processing was strictly magnesium dependent (Fig. 3B; fig. S9), which is different from other CRISPR-Cas RNA processing reaction conditions and suggests a distinct cleavage mechanism. Notably, the RuvC domain requires magnesium to cleave DNA (*17*), and some RuvC domains have been reported to have endoribonucleolytic activity (*15*). Based on these observations, we tested CasΦ containing a

[1]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA. [3]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, USA. [4]College of Natural Resources, University of California, Berkeley, Berkeley, CA, USA. [5]University of California, Berkeley—University of California, San Francisco Graduate Program in Bioengineering, Berkeley, CA, USA. [6]Department of Molecular, Cellular and Developmental Biology, University of California, Los Angeles, CA, USA. [7]Monash Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Victoria, Australia. [8]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA, USA. [9]Department of Earth and Planetary Sciences, University of California, Berkeley, Berkeley, CA, USA. [10]Department of Chemistry, University of California, Berkeley, Berkeley, CA, USA. [11]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [12]Gladstone Institute of Data Science and Biotechnology, Gladstone Institutes, San Francisco, CA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: doudna@berkeley.edu

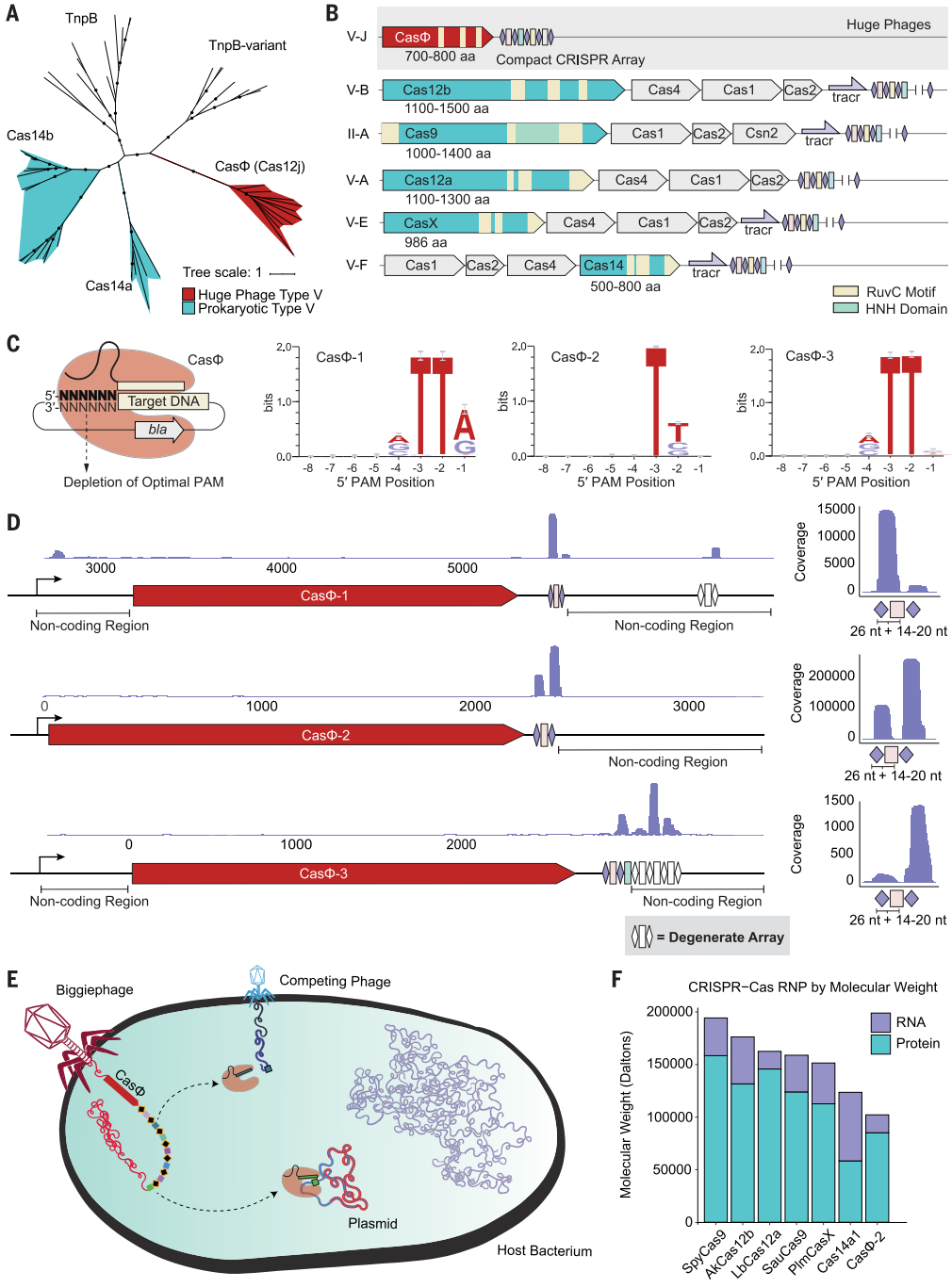


Fig. 1. CasΦ is a bona fide CRISPR-Cas system from huge phages. (**A**) Maximum likelihood phylogenetic tree of type V effector proteins and respective predicted ancestral TnpB nucleases. Bootstrap and approximate likelihood ratio test values of ≥90 are denoted on the branches with black circles. (**B**) Illustrations of genomic CRISPR-Cas loci of CasΦ, Cas14, and systems previously employed in genome-editing applications. (**C**) Graphical representation of the PAM depletion assay and the resulting PAMs for three CasΦ orthologs. (**D**) (Left) RNA-seq results mapped onto the native genomic loci of CasΦ orthologs and their upstream and downstream noncoding regions cloned with reduced CRISPR arrays into expression plasmids. (Right) Enlarged view of RNA mapped onto the first repeat (diamond)-spacer (rectangle) pair. (**E**) Schematic of the hypothesized function of Biggiephage-encoded CasΦ in an instance of superinfection of its host. CasΦ may be used by the huge phage to eliminate competing MGEs. (**F**) Predicted molecular weights of the RNP complexes of small CRISPR-Cas effectors and those functional in editing of mammalian cells.

RuvC-inactivating mutation and found it to be incapable of processing pre-crRNAs (Fig. 3B; fig. S9, A and B). Both wild-type and catalytically inactivated CasΦ proteins bound crRNA, and their reconstituted complexes with pre-crRNA had similar elution profiles from a size exclusion column, suggesting no pre-crRNA binding or protein stability defect resulting from the RuvC mutation (fig. S10).

We hypothesized that if the RuvC domain is responsible for pre-crRNA processing, the products should contain 5′-phosphate and 2′- and 3′-hydroxyl moieties as observed in RNAs generated by the RuvC-related RNase HI enzymes (*17*). By contrast, other type V CRISPR-Cas enzymes process pre-crRNA by metal-independent acid-base catalysis in an active site distinct from the RuvC, generating 2′-3′-cyclic phosphate crRNA termini, as observed for Cas12a (*18*). Phosphatase treatment of CasΦ-generated crRNA and then denaturing acrylamide gel analysis showed no change in the crRNA migration, distinct from the change in mobility detected for crRNA generated by Cas12a (Fig. 3C; fig. S9C). This result implies that no 2′-3′-cyclic phosphate is formed during the reaction catalyzed by CasΦ, in contrast to the acid-base– catalyzed processing reaction for Cas12a (Fig. 3, C and D). Together, these data demonstrate that CasΦ uses a single RuvC active site for both pre-crRNA processing and DNA cleavage.

The versatility and programmability of CRISPR-Cas systems for genome editing in virtually any organism have sparked a revolution in biotechnology and fundamental research (*19*). To investigate whether CasΦ can be harnessed for human genome editing, we performed a gene disruption assay (*8*) using CasΦ coexpressed

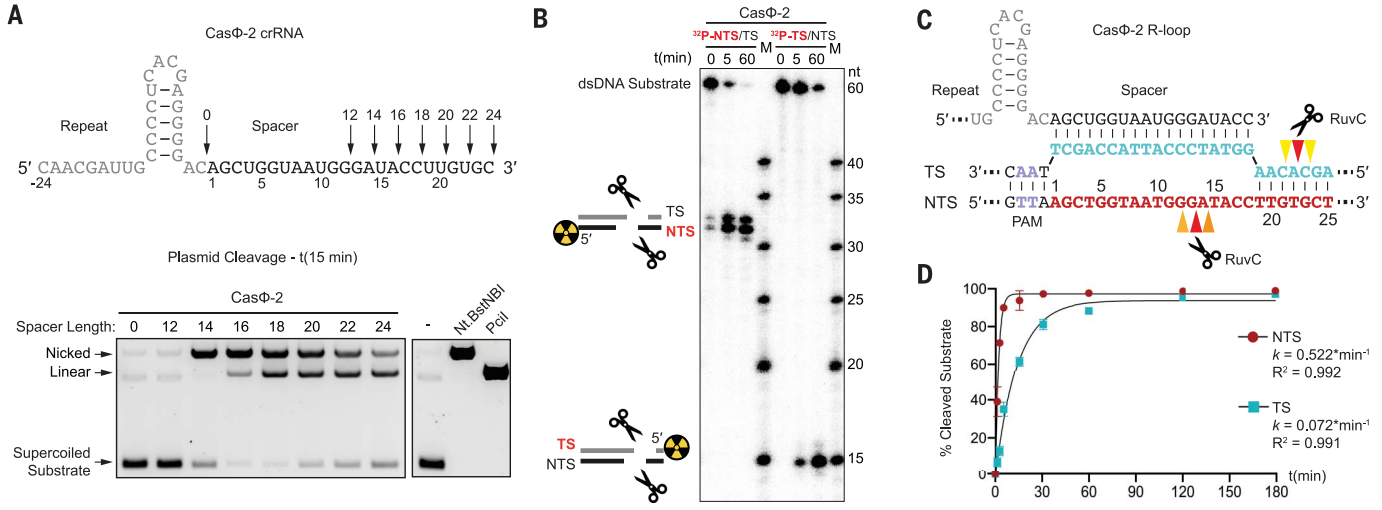
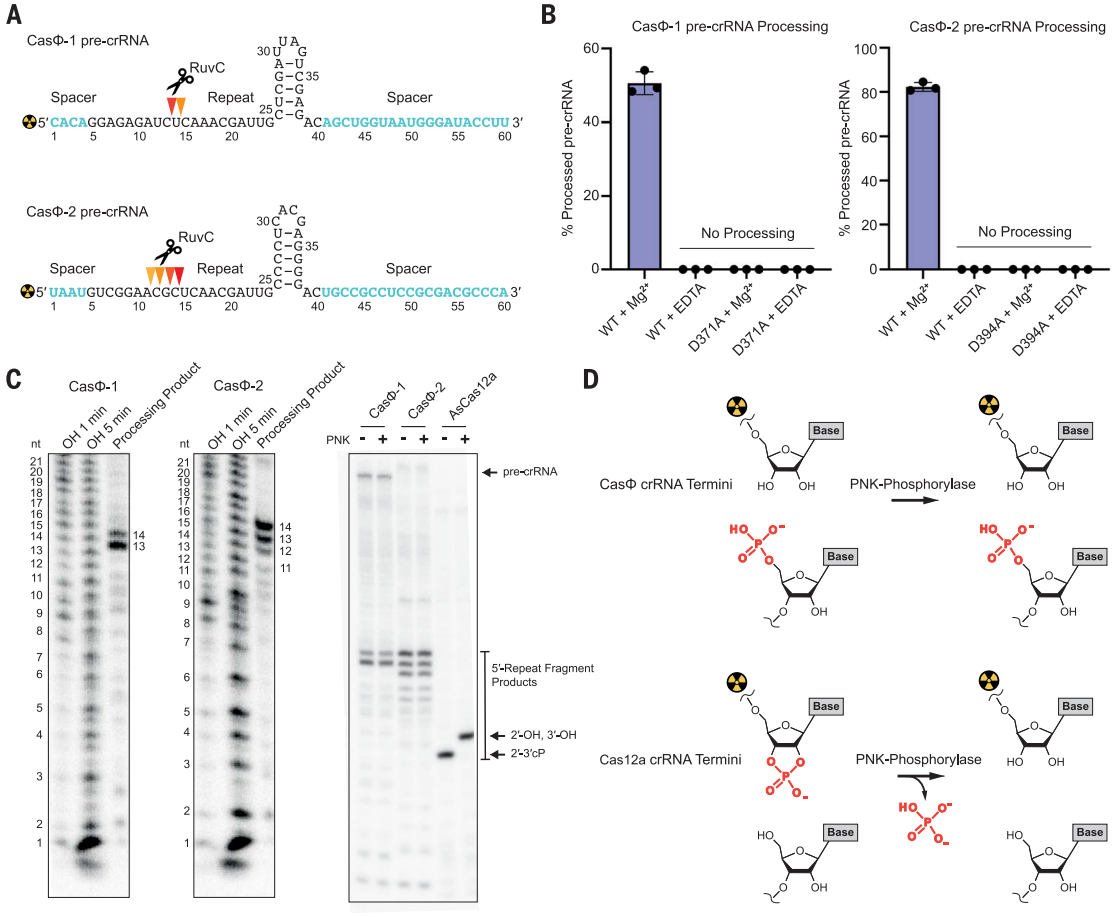


Fig. 2. CasΦ cleaves DNA. (**A**) Supercoiled plasmid cleavage assay testing CasΦ RNPs reconstituted with crRNAs of different spacer lengths. (**B**) Cleavage assay targeting dsDNA oligo-duplices for mapping of the cleavage structure. (**C**) Scheme illustrating the cleavage pattern. (**D**) NTS and TS DNA cleavage efficiency (n = 3 each; means ± SD). Data are shown in fig. S7B.

Fig. 3. CasΦ processes pre-crRNA within the RuvC active site. (**A**) Pre-crRNA substrates and processing sites (red triangles) derived from the OH ladder in panel C. (**B**) Pre-crRNA processing assay for CasΦ-1 and CasΦ-2 in dependence of Mg^{2+} and RuvC active site residue variation [Asp^{371}→Ala (D371A) and D394A] (n = 3 each; means ± SD; time = 60 min). Data are shown in fig. S9B. (**C**) (Left and middle) Alkaline hydrolysis ladder (OH) of the pre-crRNA substrate. (Right) T4 polynucleotide kinase (PNK)–phosphatase treatment of the CasΦ and *Acidaminococcus* sp. Cas12a cleavage products. (**D**) Graphical representation of the mature crRNA termini chemistry of CasΦ and Cas12a and PNK-phosphorylase treatment outcomes.



with a crRNA in HEK293 cells (Fig. 4A). We found that CasΦ-2 and CasΦ-3 induced targeted disruption of a genomically integrated enhanced green fluorescent protein (EGFP) gene (Fig. 4A; fig. S11). In one case, CasΦ-2 with an individual guide RNA was able to edit up to 33% of cells (Fig. 4A), which is comparable to levels initially reported for CRISPR-Cas9, CRISPR-Cas12a, and CRISPR-CasX (*7*, *8*, *20*). We next tested if CasΦ-2 could be delivered as ribonucleoproteins (RNPs) into plant protoplasts to edit the endogenous *Arabidopsis thaliana PDS3* gene (Fig. 4B; fig. S12). Next-generation sequencing revealed that CasΦ-2 introduced primarily 8- to 10-bp deletions (Fig. 4B), consistent with the cleavage pattern observed in vitro (Fig. 2C). The small size of CasΦ in combination with its minimal PAM

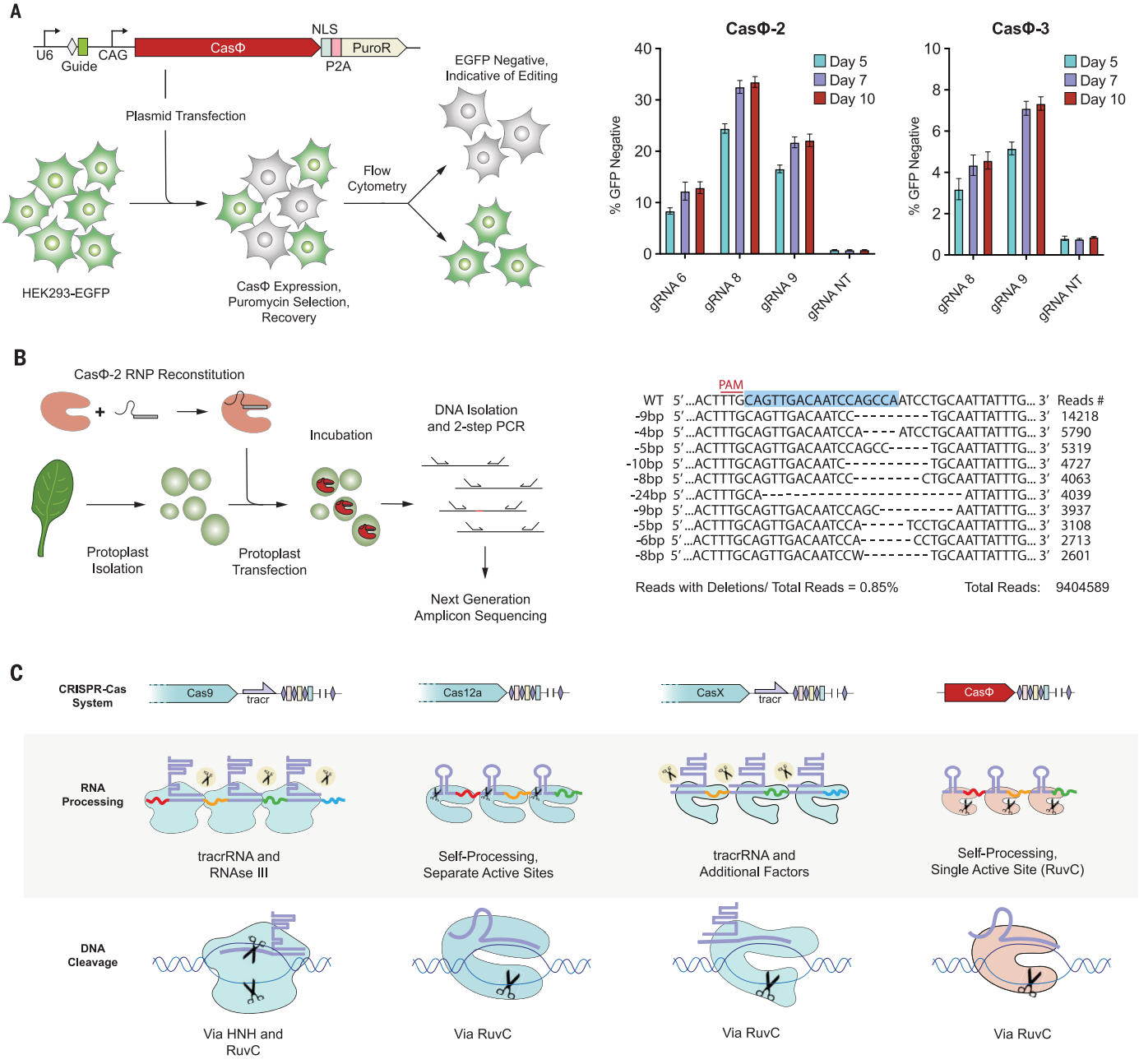


Fig. 4. CasΦ is functional for genome editing. (**A**) Experimental workflow of the GFP disruption assay (left) and GFP disruption using CasΦ-2 and CasΦ-3 and a nontargeting (NT) guide as a negative control (n = 3 each; means ± SD). (**B**) (Left) Experimental workflow of CasΦ2 RNP-mediated genome editing in *A. thaliana* mesophyll protoplasts. (Right) Amplicon sequencing data showing the most frequent deletions for gRNA33 in the targeted region (blue) within the *AtPDS3* gene. (**C**) Scheme illustrating the differences in RNA processing and DNA cutting for Cas9, Cas12a, CasX, and CasΦ.

requirement will be particularly advantageous for both vector-based delivery into cells and a wider range of targetable genomic sequences, and as such it will provide a powerful addition to the CRISPR-Cas toolbox.

Three other well-characterized Cas enzymes, Cas9, Cas12a, and CasX, use one (Cas12a and CasX) or two (Cas9) active sites for DNA cutting and rely on a separate active site (Cas12a) or additional factors (CasX and Cas9) for crRNA processing (Fig. 4C). The finding that a single RuvC active site in CasΦ is capable of crRNA processing and DNA cutting suggests that size limitations of phage genomes, possibly in combination with large population sizes and higher mutation rates in phages than in prokaryotes (*21*–*23*), led to a consolidation of chemistries within one catalytic center. Such compact proteins may be particularly amenable to engineering and laboratory evolution to create new functionalities for genome manipulation, and they highlight huge phages as an exciting forefront for discovery and biotechnological applications for human health.

REFERENCES AND NOTES

1. R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
2. J. McGinn, L. A. Marraffini, *Nat. Rev. Microbiol.* **17**, 7–12 (2019).
3. F. Hille *et al.*, *Cell* **172**, 1239–1259 (2018).
4. B. Al-Shayeb *et al.*, *Nature* **578**, 425–431 (2020).
5. S. Shmakov *et al.*, *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
6. D. Gleditzsch *et al.*, *RNA Biol.* **16**, 504–517 (2019).
7. B. Zetsche *et al.*, *Cell* **163**, 759–771 (2015).
8. J.-J. Liu *et al.*, *Nature* **566**, 218–223 (2019).
9. J. S. Chen *et al.*, *Science* **360**, 436–439 (2018).
10. A. East-Seletsky *et al.*, *Nature* **538**, 270–273 (2016).
11. J. S. Gootenberg *et al.*, *Science* **356**, 438–442 (2017).
12. I. Fonfara, H. Richter, M. Bratovič, A. Le Rhun, E. Charpentier, *Nature* **532**, 517–521 (2016).

13. D. Burstein *et al*., *Nature* **542**, 237–241 (2017).
14. L. B. Harrington *et al*., *Science* **362**, 839–842 (2018).
15. W. X. Yan *et al*., *Science* **363**, 88–91 (2019).
16. S. Shmakov *et al*., *Mol. Cell* **60**, 385–397 (2015).
17. M. Nowotny, *EMBO Rep.* **10**, 144–151 (2009).
18. D. C. Swarts, J. van der Oost, M. Jinek, *Mol. Cell* **66**, 221–233.e4 (2017).
19. G. J. Knott, J. A. Doudna, *Science* **361**, 866–869 (2018).
20. P. Mali *et al*., *Science* **339**, 823–826 (2013).
21. S. Duffy, L. A. Shackelton, E. C. Holmes, *Nat. Rev. Genet.* **9**, 267–276 (2008).
22. M.-C. Lee, C. J. Marx, *PLOS Genet.* **8**, e1002651 (2012).
23. M. Lynch, *Annu. Rev. Microbiol.* **60**, 327–349 (2006).

ACKNOWLEDGMENTS

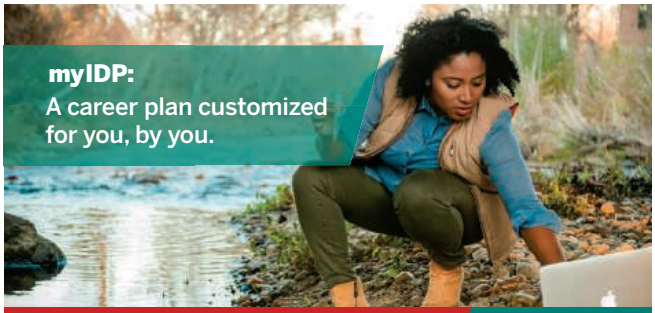
We thank Anabel Schweitzer and Hannah Spinner for providing plasmids (pBFC545, pBFC546, pBFC547, and the PAM library plasmid) and Juan Hurtado for providing the EGFP-HEK293 cell line. We thank Doudna Lab members as well as Alexander Crits-Christoph and Jacob West-Roberts for helpful discussions and Rohan Sachdeva for bioinformatic support. **Funding:** P.P. was supported by the Paul G. Allen Frontiers Group. P.P. and B.A.-S. are supported by the Somatic Cell Genome Editing Program of the Common Fund of the NIH (NIH U01AI142817-02). B.A.-S. is supported by an NSF Graduate Research Fellowship (no. DGE 1752814). G.J.K. is supported by an NHMRC Investigator Grant (EL1, APP1175568) and previously an American Australian Association Fellowship. J.A.D. receives funding from the NSF (1817593). J.A.D. and S.E.J. are Investigators of the Howard Hughes Medical Institute. **Author contributions:** B.A.-S. and J.F.B. conceived the initial study, which was further developed with P.P. and J.A.D. P.P. and B.A.-S. designed experiments and analyzed data with input from J.A.D. and J.F.B. P.P., E.B-.R., and B.A.-S. cloned constructs with guidance from P.P. B.A.-S. conducted metagenomic assemblies, genome curation, protein and crRNA sequence analyses, active residue identification, and computational analyses for PAM depletion assays and RNA-seq. P.P. conducted PAM depletion assays, efficiency of transformation assays, performed RNA-seq experiments, purified proteins, and performed biochemical experiments. P.P., B.A.-S., C.A.T., and Z.L. designed genome editing experiments with input from S.E.J. C.A.T. performed tissue culture and flow cytometry, and Z.L. performed editing experiments in *A. thaliana*. B.F.C. and G.J.K. provided materials and experimental advice. B.A.-S. and P.P. wrote the manuscript with input from J.A.D. and J.F.B. P.P. and B.A.-S. prepared figures. The manuscript was reviewed and approved by all coauthors. **Competing interests:** The Regents of the University of California have patents pending for CRISPR technologies on which the authors are inventors. J.A.D. is a cofounder of Caribou Biosciences, Editas Medicine, Intellia Therapeutics, Scribe Therapeutics, and Mammoth Biosciences. J.A.D. is a scientific advisory board member of Caribou Biosciences, Intellia Therapeutics, eFFECTOR Therapeutics, Scribe Therapeutics, Synthego, Algen Biotechnologies, Felix Biosciences, and Mammoth Biosciences. J.A.D. is a director at Johnson & Johnson and has sponsored research projects by Pfizer, Roche, AppleTree Partners, and Biogen. J.F.B. is a founder of Metagenomi. S.E.J. is a scientific cofounder of Inari Agriculture and S.E.J. and J.A.D. are members of its scientific strategy board. **Data and materials availability:** All data are available in the manuscript or the supplementary material. Reagents are available through Addgene and upon request from J.A.D.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6501/333/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S4
References (*24–29*)
MDAR Reproducibility Checklist
Protein, repeat, and locus sequences for CRISPR-CasΦ

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

5 February 2020; accepted 6 May 2020
10.1126/science.abb1400



For your career in science, 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:

FASEB · University of Massachusetts Medical School UMASS · UCSF University of California San Francisco · UIC · Burroughs Wellcome Fund

EGL Charitable Foundation

The EGL Charitable Foundation invites you to apply to the

Gruss Lipper Post-Doctoral Fellowship Program

Eligibility

- Israeli citizenship
- Candidates must have completed PhD and/or MD/PhD degrees in the Biomedical Sciences at an accredited Israeli University/Medical School or be in their final year of study
- Candidates must have been awarded a postdoctoral position in the U.S. host research institution.

Details regarding the fellowship are available at www.eglcf.org

Application Deadline is October 4, 2020

CAREER FEATURE:

Postdoctoral Careers

Issue date: August 21
Reserve ad space by August 6
Ads accepted until August 14 if space allows



To book your ad, contact:
advertise@sciencecareers.org

The Americas
202 326 6577
Europe
+44 (0) 1223 326527
Japan
+81 3 6459 4174
Greater China, South Korea, Singapore, Thailand
+86 131 4114 0012

Produced by the *Science*/AAAS Custom Publishing Office.

Why choose this feature for your advertisement?

- Relevant ads lead off the career section with a special postdoctoral banner.

Expand your exposure by posting your print ad online:

- Additional marketing driving relevant job seekers to the job board
- *Science* online job postings are now being served to thousands of passive job seekers in the Wiley Online Library.

Science Careers AAAS
SCIENCECAREERS.ORG



FOR RECRUITMENT IN SCIENCE, THERE'S ONLY ONE *SCIENCE*.

By **Lala Tanmoy Das**

When the jokes aren't funny

At our first lab meeting since the COVID-19 lockdowns were lifted, I sensed unease. After taking a seat at a large table, I noticed that no one seemed to want to sit next to me. As more people shuffled in, most of my labmates huddled on the opposite side of the table. Someone mentioned the "kung flu." Another labmate "joked" that Asians really ought to be wearing masks, especially in small gatherings such as lab meetings. It took me a few moments to realize that—as the only Asian person in the room—they were likely talking about me. I know I should have been wearing a mask, but I wasn't the only person in the room without one.

A few days before that, I was sitting in the building's atrium with a Black colleague discussing an experimental method when a scientist from another lab joined us. Upon noticing my colleague's Black Lives Matter pin, the scientist said, "You really ought to keep your politics at home." We were both taken aback. Our visitor went on to declare that looting is not the right solution to George Floyd's killing, that people need to get jobs instead of protesting all day, and that Black people shouldn't be so angry all the time. We sat there, appalled, but didn't say anything.

On another occasion, a former co-worker asked me—as an immigrant—how I felt about President Donald Trump's immigration bans. I said I wished immigrants weren't constantly viewed as a threat to the economy because we have so much to offer, particularly in niche work sectors. The person's response was along the lines of, "At least Silicon Valley will be less of a 'brown town' and give others a chance." I laughed but felt seriously uncomfortable.

Why didn't I speak up? Because I feared repercussions—and I'm sure I'm not alone in feeling that way. On the rare occasions when I have voiced my concern about racist comments, I've been told I am "too much of a social justice warrior." I'm at a precarious point in my training, because I'm currently doing lab rotations, and I will settle on a Ph.D. adviser next year. Strained relationships with my colleagues may make it harder to find a permanent lab, or they may lead to poor letters of recommendation or lost authorship opportunities.

Yet racist comments, no matter how funny some people may think they are, should not be normalized and tolerated. I feel deeply unsettled that in the current political environment, some people—including some scientists—think it's OK to say these things. And I wish early-career scientists like me felt more comfortable questioning them.



"Racist comments ... should not be normalized and tolerated."

It will take action at all rungs of scientific institutions—from deans' offices down to individual labs—to change the climate. We need more than a few hours of mandated online bias training to really mobilize a shift. Universities need to create safe spaces to discuss racial issues and microaggressions head-on—for example, by developing seminar series or journal clubs to educate the community about problems and to generate solutions. And science itself needs a culture change. We should commit to holding one another accountable for problematic behavior. When issues arise, institutions should have a clear procedure for filing grievances.

And, more broadly, there should be open and honest dialogue about the culture of intimidation and hierarchy in science. Universities and labs should seek actionable ways to flatten research team structures so that early-career scientists can feel empowered to speak up against problematic behavior.

Change is hard and takes time. Talking about racism with your colleagues is even harder. Mopping up centuries of racism that have percolated into everyday parlance will take a long time. However, science thrives when there are alternate perspectives. Now more than ever, we must acknowledge that and champion a culture of positive change. ■

Lala Tanmoy Das is an M.D.-Ph.D. student in a tri-institutional program offered by Weill Cornell Medicine, Rockefeller University, and Memorial Sloan Kettering Cancer Center in New York City.
Do you have an interesting career story that you would like to share? Send it to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS