Does a record-setting 2020 suggest global warming is accelerating? p. 334

Clock genes and mosquito pheromones pp. 340 & 411

Explaining the Solar System's two populations p. 365

Science

$15
22 JANUARY 2021
sciencemag.org

AAAS



SPITTING IMAGE

Convergent changes in cobra venom p. 386



The heart of the matter.

The NEBNext® Ultra™ II workflow lies at the heart of NEB's portfolio for next gen sequencing library preparation. With specially formulated master mixes and simplified workflows, high quality libraries can be generated with low inputs and reduced hands-on time.

As sequencing technologies improve and applications expand, the need for compatibility with ever-decreasing input amounts and sub-optimal sample quality grows. Scientists must balance reliability and performance with faster turnaround, higher throughput and automation compatibility.

NEBNext Ultra II modules and kits for Illumina® are the perfect combination of reagents, optimized formulations and simplified workflows, enabling you to create DNA or RNA libraries of highest quality and yield, even when starting from extremely low input amounts.

The Ultra II workflow is central to many of our NEBNext products, including:

- Ultra II DNA & FS DNA Library Prep
- Enzymatic Methyl-seq
- Ultra II RNA & Directional RNA Library Prep
- Single Cell/Low Input RNA Library Prep
- Module products for each step in the workflow

The Ultra II workflow is available in convenient kit formats or as separate modules – it is easily scalable and automated on a range of liquid handling instruments.

End Repair/ dA-Tailing	Adaptor Ligation	Clean Up/ Size Selection	PCR Enrichment	Clean Up
Module NEB #E7546	Module NEB #E7595	Component of NEB #E7103	Module NEB #M0544	Component of NEB #E7103
NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB #E7645)				
NEBNext Ultra II Library Prep with Sample Purification Beads (NEB #E7103)				

The NEBNext Ultra II workflow has been cited in thousands of publications, as well as a growing number of preprints and protocols related to COVID-19. Citation information and extensive performance data for each product is available on neb.com.

To learn more about why NEBNext is the choice for you, visit **NEBNext.com**.

One or more of these products are covered by patents, trademarks and/or copyrights owned or controlled by New England Biolabs, Inc. For more information, please email us at gbd@neb.com. The use of these products may require you to obtain additional third party intellectual property rights for certain applications. Illumina® is a registered trademark of Illumina, Inc.
© Copyright 2020, New England Biolabs, Inc.; all rights reserved.

NEW ENGLAND BioLabs® Inc.

be INSPIRED
drive DISCOVERY
stay GENUINE

CONTENTS

22 JANUARY 2021 • VOL 371 ISSUE 6527



NEWS

IN BRIEF

326 News at a glance

IN DEPTH

329 New mutations raise specter of 'immune escape'
SARS-CoV-2 variants found in Brazil and South Africa may evade human antibodies
By K. Kupferschmidt

331 Disgraced COVID-19 studies are still routinely cited
Journal mentions of studies based on disputed data from Surgisphere often fail to clearly flag retractions *By C. Piller*

332 Controversial study says U.S. labs use 111 million mice, rats
Figure is more than seven times some estimates, but critics say analysis is flawed
By D. Grimm
PODCAST

333 What causes IBS pain? Maybe immune flare-ups
New study suggests certain foods trigger a localized reaction in the gut *By J. Couzin-Frankel*
PODCAST

334 Global temperatures in 2020 tied record highs
Defying cooling from La Niña, warming planet speeds toward breach of climate targets
By P. Voosen

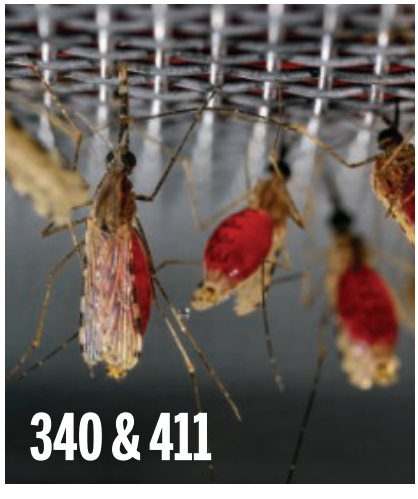
FEATURES

336 Separation anxiety
Phase separation, an idea about how cells organize their contents and functions into dropletlike compartments, has divided biologists
By M. Leslie

INSIGHTS

PERSPECTIVES

340 Drivers of mosquito mating
Light cycle genes and environmental and chemical cues modulate male mosquito swarming *By N. C. Manoukis*
REPORT p. 411



341 Snapshots of a light-triggered transition
A new method captures real-space movies of ordered states on ultrafast time scales
By A. Kogar
REPORT p. 371

343 Thermopower and harvesting heat
A metric to compare low-grade heat-to-electricity conversion can be misleading *By X. Shi and J. He*

344 A host metabolite promotes *Salmonella* survival
Succinate is used by *Salmonella* in different ways when it is in macrophages or the gut
By J. P. Lynch and C. F. Lesser
REPORT p. 400

345 Racing against unwanted isomerization
Control of relative reaction rates paves the way for (*Z*)-alkene stereochemical retention
By S. J. Malcolmson
REPORT p. 380

347 Sex differences in immune responses
Biological sex differences in immunity potentially underlie male bias for severe COVID-19
By T. Takahashi and A. Iwasaki

349 Retrospective: Masatoshi Koshiba (1926–2020)
Innovative founder of neutrino astronomy
By M. Nakahata and A. Suzuki

POLICY FORUM

350 Rethinking immigration policies for STEM doctorates
Many Ph.D.'s follow an inefficient path to green cards through visas aimed at entry-level workers *By M. Roach and J. Skrentny*

BOOKS *ET AL.*

353 Darwin and human evolution
An anniversary appraisal of *The Descent of Man* probes the naturalist's prescience and prejudices *By E. L. Milam*

354 A barrier breaker's complicated legacy
Certain of her own exceptional nature, Elizabeth Blackwell dismissed those who aided her *By L. Stark*

LETTERS

355 Shortfin mako sharks speeding to the brink
By D. W. Sims et al.

355 Mediterranean sharks and rays need action
By M. Milazzo et al.

CREDITS: (ILLUSTRATION) V. ALTOUNIAN/*SCIENCE*; (PHOTO) SOLVIN ZANKL/NPL/MINDEN PICTURES



Apply for our exciting research Prize!

SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

$25, 000 Grand Prize!
Get published in *Science*!

The *Science*-PINS Prize is a highly competitive international prize that honors scientists for their excellent contributions to neuromodulation research. For purposes of the Prize, neuromodulation is any form of alteration of nerve activity through the delivery of physical (electrical, magnetic, or optical) stimulation to targeted sites of the nervous system with impications for translational medicine.

For full details, judging criteria and eligibility requirements, visit:

www.sciencemag.org/prizes/pins

Submission Deadline: 15 March 2021

Science AAAS

PINS 品驰

Science Translational Medicine AAAS

356 Tilapia farming threatens Brazil's waters
By P. Charvet et al.

RESEARCH

IN BRIEF

358 From *Science* and other journals

REVIEW

361 Neurodevelopment
Mapping the molecular and cellular complexity of cortical malformations
E. Klingler et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABA4517

RESEARCH ARTICLES

Chromatin

362 JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications
V. Kasinath et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABC3393

363 Regulation of the Dot1 histone H3K79 methyltransferase by histone H4K16 acetylation
M. I. Valencia-Sánchez et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABC6663

364 Single-cell genomics
Developmental cell programs are co-opted in inflammatory skin disease
G. Reynolds et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABA6500

365 Planet formation
Bifurcation of planetary building blocks during Solar System formation
T. Lichtenberg et al.

REPORTS

371 Ultrafast microscopy
Ultrafast nanoimaging of the order parameter in a structural phase transition
T. Danz et al.
PERSPECTIVE p. 341

375 Surface chemistry
Unraveling CO adsorption on model single-atom catalysts *J. Hulva* et al.

380 Organic chemistry
Iridium-catalyzed *Z*-retentive asymmetric allylic substitution reactions
R. Jiang et al.
PERSPECTIVE p. 345

386 Convergent evolution
Convergent evolution of pain-inducing defensive venom components in spitting cobras *T. D. Kazandjian* et al.
PODCAST

390 Solar cells
Nanoscale localized contacts for high fill factors in polymer-passivated perovskite solar cells *J. Peng* et al.

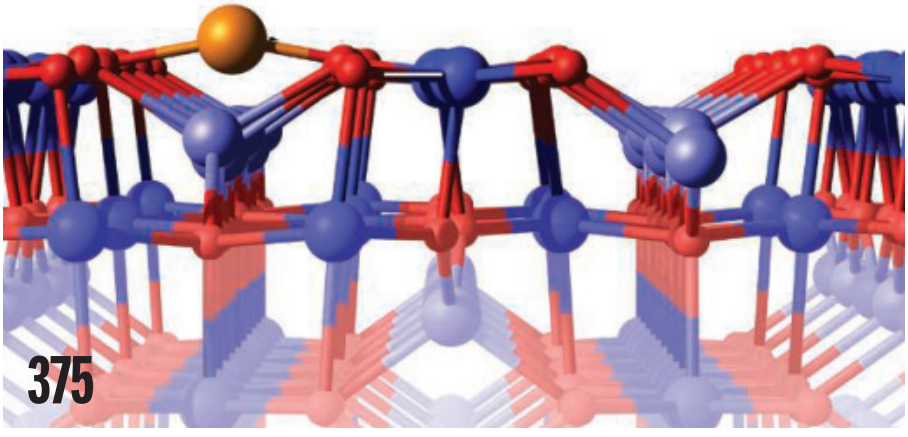
396 Sexual dimorphism
Transvection regulates the sex-biased expression of a fly X-linked gene
C. C. Galouzis and B. Prud'homme

400 Host defense
Host succinate is an activation signal for *Salmonella* virulence during intracellular infection *G. Rosenberg* et al.
PERSPECTIVE p. 344

405 Immunometabolism
Glycolysis fuels phosphoinositide 3-kinase signaling to bolster T cell immunity
K. Xu et al.

411 Mosquito biology
Clock genes and environmental cues coordinate *Anopheles* pheromone synthesis, swarming, and mating *G. Wang* et al.
PERSPECTIVE p. 340

415 Metabolic evolution
Ancient balancing selection maintains incompatible versions of the galactose pathway in yeast *J. Boocock* et al.





DEPARTMENTS

325 Editorial
Target student mental well-being
By Alan I. Leshner

434 Working Life
A department terminated *By Paul Bierman*

ON THE COVER

Some cobras, such as this Mozambique spitting cobra (*Naja mossambica*), use venom spitting as an effective long-range defense mechanism. The three origins of venom spitting in snakes coincide with convergent changes to venom composition, allowing venom to serve a defensive function by producing enhanced pain in the recipient. These findings emphasize that evolution can funnel complex adaptations down repeatable pathways. See page 386. *Photo: Guy Edwardes/NPL/Minden Pictures*



Science Staff 322
AAAS Meeting Program 421
New Products 430
Science Careers 431

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 © 2021 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.

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

Science AAAS

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

Executive Editor Monica M. Bradford

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

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

News Editor Tim Appenzeller

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

Creative Director Beth Rakouskas

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

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

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

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

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

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

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

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

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

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

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

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

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

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

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

CLASSIFIED ADVERTISING
advertising.sciencemag.org/science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

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

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

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

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

Flexible microplate reader solutions for virology research

Omega series

- Reporter gene assays
- Immunoassays – ELISA
- DNA/RNA quantification

SPECTROstar® *Nano*

- A260 DNA/RNA quantification
- Protein/cytokine quantification
- MTT, MTS, WST assays



CLARIOstar® *Plus*

- Live cell assays
- CPE and TCiD50 determination
- Virus detection with LAMP assay

PHERAstar® *FSX*

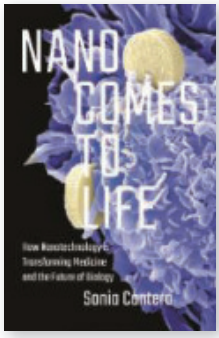
- Antiviral drug screening
- Protein-protein interactions
- Antigen binding and affinity

www.bmglabtech.com

BMG LABTECH

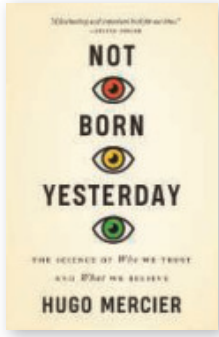
The Microplate Reader Company

© 2021 All rights reserved. All logos and trademarks are the property of BMG LABTECH.



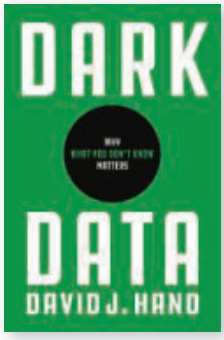
"[A] succinct study.... Contera frames this near-future transmaterial science, with its focus on human well-being, as an effort allied to social justice even as it probes existential questions of what it means to be human."

—Barbara Kiser, *Nature*



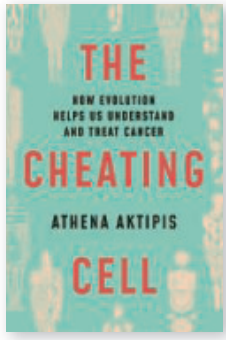
"At the risk of being seen as credulous, I'd say [Mercier] makes a strong case for gullibility being a far less prevalent and important trait than we thought."

—*New Scientist*



"A topical and relevant book.... Hand knows his subject well and takes us on a tour de force of the many, varied reasons why data, and the elements that we don't have – the dark data – is such a complex area to get right."

—Simon Cocking, *Irish Tech News*



"Aktipis takes an evolutionary approach to cancer, tracing the ways cells 'cheat' natural selection and showing how the human body evolved to outsmart many of those threats."

—Erin Blakemore, *Washington Post*

PRINCETON UNIVERSITY PRESS

Science Webinars
AAAS

Science Webinars help you keep pace with emerging scientific fields!

Stay informed about scientific breakthroughs and discoveries.
Gain insights into current research from top scientists.
Take the opportunity to ask questions during live broadcasts.
Get alerts about upcoming free webinars.

Sign up at: webinar.sciencemag.org/stayinformed

EDITORIAL

Target student mental well-being

Mental health and well-being are critical to student success in higher education and beyond but rarely get the attention they need. Colleges and universities in the United States have been seeing increases in mental illness and emotional distress among students for decades. Unfortunately, the COVID-19 pandemic and its social and economic consequences are accelerating these trends. Furthermore, renewed focus on systemic racism in many spheres of American life is highlighting the stresses and mental health consequences faced by students from traditionally underrepresented groups and by women who continue to suffer the lingering effects of sexism. The damaging consequences of these problems for students, their institutions, and a modern world that depends on a well-educated, healthy workforce demand a much better response. Promoting student mental health and emotional well-being on campus requires more than a well-functioning counseling center. It calls for a campus-wide commitment and coordinated action plan. For many institutions, accomplishing this will involve a major rethinking of their responsibilities to students.

> "Promoting student mental health...requires more than a well-functioning counseling center."

American undergraduate students have been reporting increasing rates of anxiety, depression, suicidal thoughts, trauma, and substance use for decades. In the 2018–2019 Healthy Minds Study (HMS) survey, conducted before the pandemic, 40% of about 300,000 students at some 300 colleges and universities reported experiencing a mental health problem, and 60% said they were having difficulty accessing mental health care on campus or in the community. A spring 2020 survey conducted in the midst of the COVID-19 pandemic by the HMS and American College Health Association showed that self-reported prevalence of depression among undergraduates had increased by 15% compared to the fall of 2019, and more students reported that mental health problems negatively affected their academic performance. Other studies estimate that the dropout rate for U.S. undergraduate students with diagnosed mental health problems is over 40%, and that the prevalence of mental health problems among graduate students is six times that of the general population. These statistics reflect a major problem that has clear implications not only for students but also for their institutions that depend on, if nothing else, tuition payments.

A December 2020 survey by the American Council on Education reported that 68% of U.S. college and university presidents ranked mental health issues among the five most pressing concerns facing their institutions. Most academic institutions provide some mental health services, often through a dedicated counseling center. But no matter how good that center is, it alone cannot solve the problems. Last week, a U.S. National Academies of Sciences, Engineering, and Medicine committee (which I chaired) laid out recommendations for how institutional leadership, campus mental health professionals, faculty, staff, and students can together address this mental health crisis. An important early step is for each institution to assess the extent of the crisis on its campus and whether there are sufficient treatment services either at the institution or in the local community. Another important question is whether there are elements of the institution's own environment or culture that contribute to the problems that students are experiencing, like unreasonable workloads or deadlines for assignments. In addition, student orientation should emphasize the value of pursuing one's own wellness while providing information about campus and community services available to help when needed. The stigma that inevitably accompanies mental health and substance use problems and inhibits students from seeking help must also be tackled. Faculty should be trained to recognize students in distress and refer them to professional services.

Many institutions may need to undergo a change in culture and perception of their mission—from seeing their role as dealing only with severe or acute mental health episodes to promoting the mental health and well-being of all students. Every college and university should implement a comprehensive action plan to achieve this goal. This may require additional resources during a trying time for institutional budgets. But supporting the whole student is important enough that financial priorities should be rearranged, if that is what it takes.

The mission of higher education should be to develop the whole student. After all, physically and mentally healthy, well-educated individuals are what society really needs from academia.

–Alan I. Leshner



Alan I. Leshner is chief executive officer emeritus of the American Association for the Advancement of Science, former executive publisher of *Science*, former deputy director and acting director of the U.S. National Institute of Mental Health, and former director of the U.S. National Institute on Drug Abuse of the National Institutes of Health. alan.i.leshner@gmail.com

PHOTO: PROFESSIONAL IMAGES PHOTOGRAPHY

10.1126/science.abg5770

NEWS



Former President Donald Trump visited the headquarters of the U.S. Centers for Disease Control and Prevention in March 2020.

IN BRIEF

Edited by **Jeffrey Brainard** *and* **Jocelyn Kaiser**

Trump's presidency ends with flurry of controversial moves on science

In its waning days, former President Donald Trump's administration moved to finalize an array of controversial new regulations, policies, and decisions involving science. But many of the measures might not last long, because they could be rewritten or revoked by President Joe Biden, Congress, or the courts. Trump's last-gasp efforts include:

SCIENTISTS' TERM LIMITS The Trump administration moved to impose 5-year term limits on top scientists at the U.S. Department of Health and Human Services (HHS). The rule, released on 15 January, requires that directors of seven centers at the Food and Drug Administration, as well as 17 positions at the Centers for Disease Control and Prevention, undergo a performance review that could lead to a new 5-year appointment, or to the staffer's transfer. A 2016 law mandates such 5-year reviews for institute and center directors at the National Institutes of Health. But some current and former officials worry the term limits will subject such positions to political interference from the White House, and they could face legal challenges, *Politico* reported.

GLOBAL WARMING A surprise Environmental Protection Agency (EPA) rule finalized last week would effectively ban the government from regulating greenhouse gas emissions from heavy industries other than power plants. The agency substantially rewrote a draft rule originally focused on regulating carbon emissions from new power plants, expanding it to exempt other "stationary" sources, such as refineries and oil and gas wells. The exemption covers an entire class of sources if its collective emissions are less than 3% of the U.S. total. Only power plants, which produce 27% of U.S. carbon emissions, exceed that bar. Analysts say the rule is vulnerable to a court challenge, and the Biden administration can likely suspend its implementation.

SPOTTED OWLS The U.S. Fish and Wildlife Service slashed protections for the endangered northern spotted owl on 13 January, declaring more than 1.4 million hectares of Pacific Northwest forests would no longer be considered critical habitat for the bird. The decision comes despite findings by agency scientists that the owl's population is declining and that it warrants stricter protection. The habitat reduction is part of a move by the outgoing Trump administration to settle a lawsuit by the timber industry and counties that earn revenue from logging. The land area is 17 times the amount that the agency initially proposed to remove from protections in August 2020.

TRANSGENIC ANIMALS A push by the White House would essentially eliminate the Food and Drug Administration's (FDA's) authority to regulate genetically modified animals and put the U.S. Department of Agriculture (USDA) in charge. The two agencies had been negotiating on dividing the task, and critics of the White House move say it would put USDA in the problematic position of both promoting and regulating genetically modified animals. FDA opposes the shift, *Politico* reported.

ARCTIC OIL DRILLING The first-ever auction of oil drilling rights inside the Arctic National Wildlife Refuge in Alaska, a policy priority for the Trump administration, met with a tepid response this month.

PHOTO: T. J. KIRKPATRICK/REDUX

Just three bidders paid $14.4 million to claim 11 parcels covering 220,000 hectares—about half of the land up for auction. In an unusual move, a state agency, the Alaska Industrial Development and Export Authority, won bids for nine of the parcels. It joined the auction in part because the agency fears the Biden administration will slow or block further leasing, reducing the state's potential economic gains. The leases must still be finalized by the Bureau of Land Management. Wildlife scientists have warned that the drilling could harm caribou herds and other parts of the ecosystem.

FETAL TISSUE RESEARCH Scientists who use fetal tissue obtained from elective abortions would need to comply with new rules under a proposal released by the Trump administration on 13 January. Among other changes, the policy would add new requirements to forms used to obtain informed consent from women who donate tissue for research. It would also limit the source of fetal tissue, which often comes from nonprofit clinics, to federally or state funded hospitals or academic medical centers. In 2019, Trump's administration banned fetal tissue research by federal researchers and required a new ethics review for studies by scientists receiving federal grants; research groups have urged the Biden administration to reverse that policy. The new proposal, which is open for comment for 30 days, is not expected to move forward.

CENSUS FIGHT The Trump administration last week abandoned a 2-year effort to prod the Census Bureau to provide a separate tally of undocumented U.S. residents as part of the 2020 census. In 2019, the president had ordered that the separate tally, and a rushed compilation of the decennial head count used for apportioning the 435 seats of the U.S. House of Representatives, be delivered before he left office. Most demographers said it could not be done and called the directive political interference. On 11 January, U.S. government lawyers told a federal judge that the apportionment numbers would not be ready until 6 March. Civil rights organizations want Census Director Steven Dillingham to resign before his term ends in December, saying he has failed to uphold the agency's high standards for data quality.

PHOTO: SAN DIEGO ZOO GLOBAL ARCHIVES

LEADERSHIP

Geneticist named as Biden's science adviser

President Joe Biden on 15 January named Eric Lander to be his science adviser and director of the White House Office of Science and Technology Policy. A mathematician turned molecular biologist, the 63-year-old Lander will take leave from his post as president and founder of the Broad Institute, jointly run by Harvard University and the Massachusetts Institute of Technology (MIT). The first biologist to hold the job, Lander spent 8 years as co-chair of the nation's top science advisory panel under former President Barack Obama. He also co-led the public Human Genome Project, which completed its first draft in 2001. Biden has picked chemistry Nobel laureate Frances Arnold and MIT's Maria Zuber to lead his science advisory panel. He named David Kessler, a former commissioner of the U.S. Food and Drug Administration, to direct Operation Warp Speed, the federal effort to speed development of COVID-19 vaccines. And Biden said Francis Collins has agreed to remain as director of the National Institutes of Health.

Biden unveils big COVID-19 plan

POLICY | President Joe Biden announced a sweeping, $400 billion plan last week to tackle the "dismal failure" of the COVID-19 vaccine rollout, safely reopen schools by March, and ramp up testing people for the pandemic coronavirus. The measures are part of an ambitious, $1.9 billion "American Rescue Plan" unveiled by Biden ahead of his inauguration to help people who are struggling financially because of the pandemic—a proposal that depends on Congress providing the money. The federal government would pay for 100,000 new public health workers to assist states in vaccination and other pandemic response efforts. Biden promised to invoke the Defense Production Act to provide vaccine-makers with whatever they need to increase production. Biden's administration would also work more closely with pharmacies to move vaccines from freezers into arms. "The more people we vaccinate and the faster we do it, the sooner we can put this pandemic behind us," Biden said.

Captive gorillas test positive

INFECTIOUS DISEASES | All eight gorillas at the San Diego Zoo Safari Park were exposed to the coronavirus that causes COVID-19, and at least two have begun to cough, zoo officials said last week. Tests of fecal samples showed that two were infected, marking the first known cases in nonhuman apes. The officials suspect the western lowland gorillas caught the virus from an asymptomatic staff member who tested positive for SARS-CoV-2; the zoo has been closed to the public for weeks because of the pandemic. The news confirmed fears that the virus can infect endangered great apes. Human respiratory viruses are already a leading cause of death for chimpanzees in the wild.



Gorillas at the San Diego Zoo Safari Park will be monitored for COVID-19 symptoms.

Saliva signals worst outcomes

COVID-19 | High virus levels in saliva are correlated with later hospitalization, serious illness, or death from COVID-19, raising the prospect that testing saliva for the coronavirus that causes the disease will help identify patients most at risk, a study has found. The standard test to detect the SARS-CoV-2 virus analyzes samples of nasal mucus taken with nasopharyngeal (NP) swabs. But patients with the worst outcomes were more likely to have high virus loads in their saliva, but not in their NP swabs, report Akiko Iwasaki of Yale University and colleagues in a 10 January preprint. That may reflect that nasal mucus comes from the upper respiratory tract,

whereas severe disease is associated with damage deep in the lungs; coughing regularly brings up viral particles to the throat, where they can pervade saliva. If the results are confirmed, saliva tests could help doctors prioritize which patients in the early stages of the disease should receive medicines that drive down levels of the virus.

'Mole' heat probe lost on Mars

PLANETARY SCIENCE | The Red Planet has claimed another robot. Scientists at NASA and the German Aerospace Center last week called off a 2-year effort to rescue the failed rod-shaped heat probe, or "mole," of the InSight lander. The mole was designed to burrow 5 meters into the martian soil and tease out how quickly heat escapes from Mars—a clue to how the planet formed. But soil compacted instead of crumbling as the rod tried to dig in, leaving it stuck at the surface. Even after engineers used InSight's robotic arm to push the probe down and scraped dirt on top, the probe failed a final attempt this month to dig on its own, leaving the mole buried in a shallow grave. InSight's other primary instrument, a seismometer, continues to function normally.

Scientist with China ties charged

FOREIGN INFLUENCES | The U.S. Department of Justice (DOJ) has added another prominent scientist to its crackdown on U.S.-based academics with allegedly undisclosed ties to China. On 14 January, police arrested nanotechnologist Gang Chen of the Massachusetts Institute of Technology (MIT) at his Cambridge home, charging him with violating federal wire fraud, banking, and tax laws. DOJ alleges Chen held various appointments with Chinese institutions and provided technical advice, "often in exchange for financial compensation and awards." He allegedly failed to disclose these affiliations as required when applying for U.S. Department of Energy grants, and did not tell tax authorities about a bank account in China. MIT said, "We take seriously concerns about improper influence in U.S. research." Chen was born in China and is a naturalized U.S. citizen. He is a member of the National Academy of Engineering and fellow of AAAS (which publishes *Science*).



A researcher attempts to catch a rare Raetzer's ringlet butterfly on a mountain slope in Italy.

PHOTO: ANDREA BATTISTI/HYBRIDWILDLIFE.COM

CONSERVATION

Rock climbing reveals good news on rare butterfly

To study one of Europe's rarest butterflies, researchers pioneered a new method of observation: rappelling down vertiginous mountainsides along the border of Italy and Switzerland. Scientists first described the orange-and-brown Raetzer's ringlet (*Erebia christi*) more than 100 years ago, but its dangerous, inaccessible habitat complicated population surveys. Drawing on decades of climbing experience, independent biologists Andrea Battisti and Matteo Gabaglio slid down ropes to count butterflies in several areas during the past 6 years. It was "like being an explorer … going where nobody has ever [gone]," Gabaglio says. Researchers sighted the ringlet 177 times at two key sites in Italy, they reported this month in the *Journal of Insect Conservation*. That's good news, they add: The ringlet appears to be more abundant that previous studies suggested. But because of climate change and other threats, they recommend reclassifying the species as endangered rather than vulnerable.

Drive honors paleontology pioneer

HISTORY OF SCIENCE | Fans of Mary Anning are hoping to raise £33,000 by next month to fund a statue honoring the paleontological pioneer, who discovered and interpreted key fossils along England's Jurassic Coast. Anning, who lived in Lyme Regis in the early 1800s, was the first to correctly identify an ichthyosaur, and discovered England's first pterosaur. Her discoveries were profoundly influential, but as a self-taught, working-class woman she was excluded from meetings of the Geological Society of London. Its members discussed and built on her discoveries, but often failed to acknowledge her. The £100,000 statue project was inspired by 13-year-old local Evie Swire. Organizers hope the cause will be helped by the film *Ammonite*, starring Kate Winslet as Anning, which opened in U.S. theaters in November 2020.

AAAS offers open-access route

PUBLISHING | AAAS, which publishes the *Science* family of journals, said last week it will offer its authors a free way to comply with funder requirements that their papers be open access on publication. Under a new policy, authors may deposit near-final, peer-reviewed versions of papers accepted by paywalled *Science* titles in public repositories where they are free to read. This "green open-access" route will apply for now only to authors of papers funded by Coalition S, a group of mostly European funders and foundations behind a mandate for immediate open access that takes effect this month. AAAS said it will pilot the new policy for 1 year.



Relatives attend a COVID-19 victim's burial in Manaus, Brazil, on 13 January.

PHOTO: MICHAEL DANTAS/AFP/GETTY IMAGES

COVID-19

New mutations raise specter of 'immune escape'

SARS-CoV-2 variants found in Brazil and South Africa may evade human antibodies

By **Kai Kupferschmidt**

When the number of COVID-19 cases began to rise again in Manaus, Brazil, in December 2020, Nuno Faria was stunned. The virologist at Imperial College London had just co-authored a paper in *Science* estimating that three-quarters of the city's inhabitants had already been infected with SARS-CoV-2, the pandemic coronavirus—more than enough, it seemed, for herd immunity to develop. The virus should be done with Manaus. Yet hospitals were filling up again. "It was hard to reconcile these two things," Faria says. He started to hunt for samples he could sequence to find out whether changes in the virus could explain the resurgence.

On 12 January, Faria and his colleagues posted their initial conclusions on the website virological.org. Thirteen of 31 samples collected in mid-December in Manaus turned out to be part of a new viral lineage they called P.1. Much more research is needed, but they say one possibility is that in some people, P.1 eludes the human immune response triggered by the lineage that ravaged the city earlier in 2020.

Emerging variants of the coronavirus have been in the news ever since scientists raised the alarm over B.1.1.7, a SARS-CoV-2 variant that first caught scientists' attention in England in December and that is more transmissible than previously circulating viruses (*Science*, 8 January, p. 108). But now, they're also focusing on a potential new threat: variants that could do an end run around the human immune response. Such "immune escapes" could mean more people who have had COVID-19 remain susceptible to reinfection, and that proven vaccines may, at some point, need an update.

At a World Health Organization (WHO) meeting on 12 January, hundreds of researchers discussed the most important scientific questions raised by the wave of new mutations. WHO also convened its COVID-19 Emergency Committee on 14 January to discuss the impact of the new variants and the travel restrictions that many countries are imposing to contain them. The committee called for a global effort to sequence more SARS-CoV-2 genomes to help track mutations.

The more transmissible variant, B.1.1.7, is already spreading rapidly in the United Kingdom, Ireland, and Denmark, and probably in many other countries. But scientists are just as worried about 501Y.V2, a variant detected in South Africa. Some of the mutations it carries, including ones named E484K and K417N, change its surface protein, spike, and have been shown in the lab to reduce how well monoclonal antibodies combat the virus. In a preprint published earlier this month, Jesse Bloom, an evolutionary biologist at the Fred Hutchinson Cancer Research Center, showed that E484K also reduced the potency of convalescent sera from some donors 10-fold—although he is quick to add this does not necessarily mean the mutation would cause people's immunity to the new strain to drop 10-fold.

P.1 adds to the concerns because it appears to have hit on a similar constellation of muta-

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

tions and has emerged in a place with a high level of immunity. "Anytime you see the same mutations arising and starting to spread multiple times, in different viral strains across the world, that's really strong evidence that there's some evolutionary advantage to those mutations," Bloom says.

Like B.1.1.7, the Brazilian variant is already on the move. Just as Faria was finishing his analysis of the Brazilian genomes, a report was published of a variant detected in travelers arriving in Japan from Brazil—and it turned out to be P.1. (As *Science* went to press, U.S. researchers also reported several new variants, but their importance remained unclear.)

HOW THESE NEW variants are affecting the course of the pandemic is unclear. In Manaus, for example, P.1 might have nothing to do with the new surge in infections; people's immunity might simply be waning, says University of Oxford epidemiologist Oliver Pybus. Or it might be driving the boost because it is transmitted more easily, like B.1.1.7, not because it can evade the immune response. "Of course it could be a combination of these factors, too," Pybus says.

Similarly, in a recent modeling study, researchers at the London School of Hygiene & Tropical Medicine calculated that South Africa's 501Y.V2 variant could be 50% more transmissible but no better at evading immunity, or just as transmissible as previous variants but able to evade immunity in one in five people previously infected. "Reality may lie between these extremes," the authors wrote.

Ester Sabino, a molecular biologist at the University of São Paulo, São Paulo, has launched a study to find reinfections in Manaus that could help decide between these hypotheses for P.1. Lab studies investigating the variants are also underway. The United Kingdom on 15 January launched a new consortium, G2P-UK (for "genotype to phenotype-UK"), headed by Wendy Barclay of Imperial College London, to study the effects of emerging mutations in SARS-CoV-2. One idea discussed at the 12 January WHO meeting is to set up a biobank that would aid studies by housing virus samples, as well as plasma from vaccine recipients and recovered patients.

Interactions between the new mutations may make it harder to tease out their effects. The variants from the United Kingdom, South Africa, and Manaus all share a mutation named N501Y, for instance, or Nelly, as some researchers call it. But the mutation, which affects the spike protein, also occurs in some variants that do not spread faster, suggesting N501Y does not operate alone, says Kristian Andersen of Scripps Research: "Nelly might be innocent, except maybe when she's hanging with her bad friends."

Bloom thinks none of the changes is likely to let the virus escape the immune response entirely. "But I would expect that those viruses have some advantage when a lot of the population has immunity"—which might help explain the surge in Manaus.

SO FAR THE VIRUS does not appear to have become resistant to COVID-19 vaccines, says vaccinologist Philip Krause, who chairs a WHO working group on COVID-19 vaccines. "The not-so-good news is that the rapid evolution of these variants suggests that if it is possible for the virus to evolve into a vaccine-resistant phenotype, this may happen sooner than we like," he adds. That possibility adds to the urgency of putting good surveillance in place to detect such escape variants early on, says biostatistician Natalie Dean of the University of Florida.



People line up to receive a COVID-19 vaccine in Birmingham, U.K.

PHOTO: JACOB KING/PA WIRE/BLOOMBERG/GETTY IMAGES

Some scientists worry that proposed changes in vaccine dosing regimens could hasten the evolution of such strains. Desperate to tame a massive surge in cases, the United Kingdom on 30 December decided to allow up to 12 weeks between the first and second dose of two authorized vaccines, rather than the 3 or 4 weeks used in the vaccines' clinical trials, so more people can get their first dose quickly and have at least some immunity. And the Trump administration decided to ship all available doses immediately, rather than holding back 50% to guarantee that people receive their second doses on time. That policy, which the Biden administration has said it will follow, could inadvertently extend the dosing interval if future vaccine deliveries don't arrive or aren't administered on time.

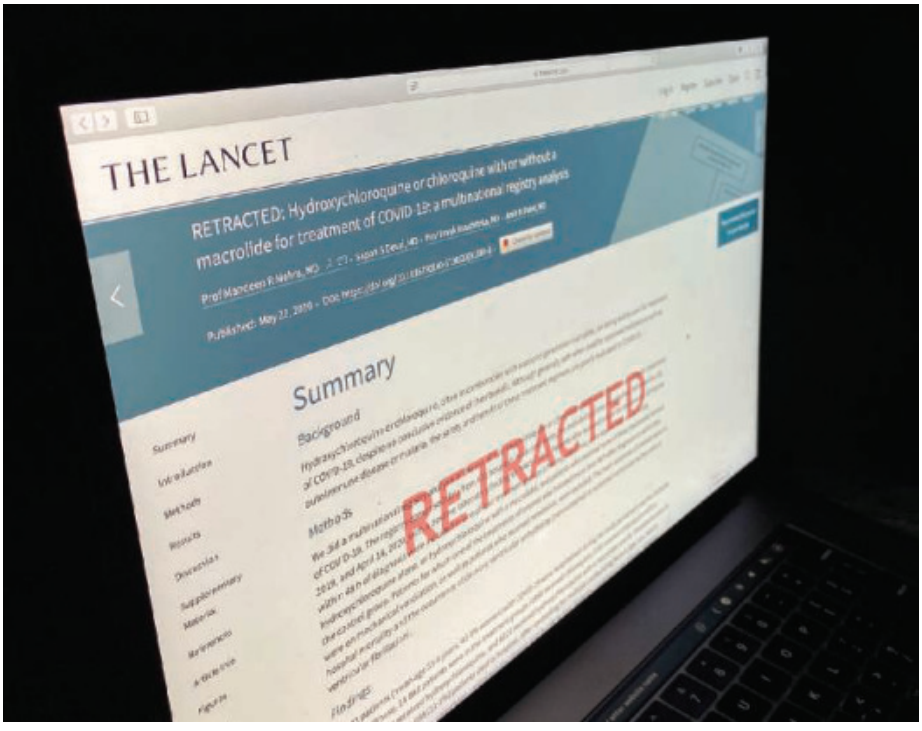
Widespread delays of the second dose might create a pool of millions of people with enough antibodies to slow the virus and avoid getting sick, but not enough to wipe it out. That could well be the perfect recipe for creating vaccine-resistant strains, says virologist Florian Krammer of the Icahn School of Medicine at Mount Sinai: "If we end up with everybody just getting one dose with no doses available for a timely boost, that would in my opinion, be a problem."

But others say unchecked spread of the virus poses greater risks. "It's carnage out there," says evolutionary microbiologist Andrew Read of Pennsylvania State University, University Park. "Twice as many people with partial immunity has got to be better than full immunity in half of them." Historically, few viruses have managed to evolve resistance to vaccines, with the notable exception of seasonal influenza, which evolves so rapidly on its own—without vaccine pressure—that it requires a newly designed vaccine every year.

If vaccine-resistant SARS-CoV-2 strains emerge, vaccines might need to be updated. Several vaccines could be easily changed to reflect the latest changes, but regulators might balk at authorizing them without seeing updated safety and efficacy data, Krause says. If new variants circulate alongside older strains, multivalent vaccines, effective against several lineages, might even be needed. "To be clear: These are downstream considerations," Krause says. "The public should not think that this is imminent, and that new vaccines will be needed." But Ravindra Gupta, a researcher at the University of Cambridge, says manufacturers should start to produce vaccines designed to generate immunity to mutated versions of the spike protein, because they keep cropping up. "It tells us that we should have these mutations in our vaccines, so that you shut off one of the avenues for the virus to go down."

For now, increased transmissibility is the biggest worry, says virologist Angela Rasmussen of Georgetown University. "I'm puzzled why [that] isn't a bigger part of the conversation," she says. The U.S. hospital system, she says, "is at capacity in many places and further increases in transmission can tip us over the edge where the system collapses. Then we'll start seeing potentially huge increases in mortality." ■

With reporting by Meredith Wadman.



On 4 June 2020, *The Lancet* and *The New England Journal of Medicine* each retracted a study on COVID-19 patients.

COVID-19

Disgraced COVID-19 studies are still routinely cited

Journal mentions of studies based on disputed data from Surgisphere often fail to clearly flag retractions

By **Charles Piller**

In June 2020, in the biggest research scandal of the pandemic so far, two of the most important medical journals each retracted a high-profile study of COVID-19 patients. Thousands of news articles, tweets, and scholarly commentaries highlighted the scandal, yet many researchers apparently failed to notice. In an examination of the most recent 200 academic articles published in 2020 that cite those papers, *Science* found that more than half—including many in leading journals—used the disgraced papers to support scientific findings and failed to note the retractions.

52.5%
Share of recent articles citing two retracted COVID-19 papers that did not mention their status.

COVID-19 "is such a hot topic that publishers are willing to publish without proper vetting," even in the face of retractions that made global headlines, says Elizabeth Suelzer, a reference librarian at the Medical College of Wisconsin who has written about problematic citations to a retracted 1998 study in *The Lancet* falsely linking vaccination to autism.

Both of the retracted COVID-19 papers, one in *The New England Journal of Medicine* (*NEJM*) and the other in *The Lancet*, were based on what appeared to be a huge database of patient records compiled from hospitals worldwide by Surgisphere, a small company operated by vascular surgeon Sapan Desai, who was a co-author on each article. The 22 May 2020 *Lancet* paper ostensibly showed that hydroxychloroquine, an antimalarial drug promoted by former President Donald Trump and others, could harm rather than help COVID-19 patients. Its publication led to a temporary halt in a major clinical trial and inflamed an already-divisive debate over the drug, which has proved to be useless against COVID-19. The 1 May *NEJM* article corroborated other evidence that people already taking certain blood pressure medicines did not face a greater risk of death if they developed COVID-19.

Questions soon arose about the validity, and even existence, of the Surgisphere database, however, and the retractions followed on 4 June (*Science*, 12 June 2020, p. 1167). But of the 200 papers examined by *Science*—all published after the retractions—105 inappropriately cited one of the disgraced studies. In several cases it was a primary source for a meta-analysis combining multiple studies to draw overarching conclusions. In most, the studies were cited as scientific support or context. *Science* also found a handful of articles that uncritically cited an influential April preprint based on the same Surgisphere data set, which described the antiparasitic drug ivermectin as beneficial in critical COVID-19 cases. (There is no standard way to retract preprints, however.)

Ivan Oransky, co-founder of the website Retraction Watch, says such blunders occur because "people are either willfully or negligently not checking references." Many authors copy and paste lists of apparently relevant citations from similar papers without actually reading them, he says. "It's frightening. It's terrible, but common."

Many of the tainted citations appeared in papers published by little-known journals, but at least a dozen found their way into major publications. For example, three articles in *PLOS ONE*, the prominent open-access journal, cited the retracted papers in discussions of pandemic conditions in Europe. A 28 December paper in the *Proceedings of the National Academy of Sciences* (*PNAS*)—one of the most influential journals—discussed the risks and benefits of drugs to treat COVID-19, including the *Lancet*'s hydroxychloroquine findings. It called them "controversial" and only noted the paper was retracted in its citation, without citing the retraction notice directly.

Editors at those two publications said they would correct the references and take steps to prevent such problems in the future. Renee Hoch, a *PLOS ONE* editor and publication ethics manager, wrote in an email that the publication relies on authors and its outside, volunteer editors to check citations, and she was caught by surprise when contacted by *Science*. "We are currently following up on this issue with high priority in light of the implications for public health and ongoing COVID-19 research," she wrote.

Hoch added that reliance on retracted work, "either directly or in the form of sup-

PHOTO: E. PETERSEN/*SCIENCE*

porting references," can be harmful. "[W]here the retracted work has clinical implications, this can result in direct risks to patients."

In a written response to questions about the citation of the *Lancet* paper, May Berenbaum, editor-in-chief of *PNAS*, said, "The authors really should have either removed the citation, added more text about why they included it, or cited the retraction notice itself." Given that no editor or reviewer caught the problem, she said, "I plan to discuss with the staff incorporating such screening into manuscript processing." A co-author of the paper, biostatistician Clelia Di Serio of Vita-Salute San Raffaele University, later said the reference to the *Lancet* paper would be expunged.

Some journal editors noted extenuating factors. In *Stroke*, a leading medical journal, a December article about ischemic stroke in COVID-19 patients referenced the *NEJM* paper without mentioning the retraction. *Stroke* Editor-in-Chief Ralph Sacco wrote in an email that the retraction occurred after the article was initially received. However, a revision of the article was resubmitted months after the event, according to the paper itself. Sacco said he would issue no correction because the retraction "is not material to the findings."

A 16 December paper on SARS-CoV-2 genetics in *Nature Communications*, another high-profile journal, also cited the *NEJM* article without reference to the retraction. Elisa De Ranieri, the journal's editor-in-chief, told *Science* her journal does not routinely check for "retractions or other postpublication updates." A lead author of the paper, biomathematician Maik Pietzner at the University of Cambridge, said that although the paper was submitted after the retractions occurred, it was written beforehand, and "the current pandemic requires immediate response." However, the paper was published 4 months after its submission.

Suelzer says inappropriate citations of retracted articles are hard to excuse. Retraction Watch publishes a free retraction database that has been integrated into a number of automated services to check citations, including scite.ai, Zotero, and RedacTek. The failure to use such tools "is a disservice to the readers and researchers," Suelzer says. "They are pretty low bars."

Yet Oransky estimates that in biomedicine, up to 90% of citations to retracted papers don't mention their fall from grace. "Half the time [as seen with the Surgisphere papers] is an improvement. That's what's shocking about it." ■

This story was supported by the *Science* Fund for Investigative Reporting.



Lab mice in their cages at an unnamed research institution.

PHOTO: UNOL/ISTOCK

ANIMAL RESEARCH

Controversial study says U.S. labs use 111 million mice, rats

Figure is more than seven times some estimates, but critics say analysis is flawed

By **David Grimm**

The most numerous mammals in U.S. research are also the most invisible. Mice and rats comprise the vast majority of lab mammals, yet no one knows exactly—or even approximately—how many are used in scientific experiments every year. Now, for the first time, someone has attempted to calculate this number using data from U.S. labs, and it's big: Nearly 111 million mice and rats are used annually in U.S. biomedical research, according to a new study. That represents more than 99% of all lab animals.

"It's a very thoughtful and reasonable analysis," says Sue Leary, president of the Alternatives Research & Development Foundation (ARDF), which seeks to reduce the number of animals in labs. The figures are troubling, she says, because mice and rats are not covered by the federal Animal Welfare Act (AWA), which enforces the humane treatment of research animals. "If the numbers are anywhere near correct, the amount of pain and suffering that's occurring in these animals is completely unacceptable."

But many in the biomedical community say the figures, published last week in *Scientific Reports*, are a gross overestimation, and that the study itself is deeply flawed. "It's a really disappointing analysis," says Allyson Bennett, senior editor at Speaking of Research, which advocates for the use of lab animals. The organization's own estimates—extrapolated from European labs, which do count mice and rats—put the number of rodents used annually in U.S. facilities at 10 million to 25 million. That would represent 93% to 97% of all U.S. research mammals. Those numbers are in line with other estimates from ARDF and the National Association for Biomedical Research (NABR).

Because the AWA doesn't cover most mice or rats, the federal government doesn't report their numbers, as it does for animals like dogs, rabbits, and monkeys. So Larry Carbone, a veterinarian who worked for 4 decades in laboratory animal care at Cornell University and the University of California, San Francisco, tried to figure them out for himself. He did something he had long been the target of when he ran animal facilities: File Freedom of Information Act requests. Animal rights groups often use such requests to dig up animal welfare

infractions. But Carbone's efforts were much more focused.

He made records requests of large public scientific institutions, asking only for the number of mice and rats they reported to the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International in 2017 and 2018. The organization inspects many U.S. animal facilities, where it counts rodents and other animals, but it keeps the figures confidential. Carbone's expertise helped. "I knew exactly which appendix of their annual report to point them to," he says.

Carbone also emailed people he knew at private institutions, which are not subject to public records requests, for their rodent records.

He was able to collect data from 11 public and five private facilities, representing 16 of the top 30 funded by the U.S. National Institutes of Health. The institutions reported using a total of about 39,000 animals covered by the AWA, and more than 5.5 million mice and rats.

That means rats and mice made up 99.3% of all mammals used at those places. If all U.S. institutions use the same proportion of these rodents versus other mammals, then the federal numbers include just 0.7% of all mammals used. Given that the government reported about 780,000 animals at all biomedical facilities in 2018, Carbone concluded that approximately 111 million mice and rats are being used in U.S. research every year. His data also indicate that mice make up the vast majority of these rodents, more than 97%.

Carbone says the figures argue for including mice and rats in the AWA. "We keep telling the public, 'Don't worry about research animals because there is rigorous government oversight,'" he says. "But that oversight covers less than 1% of the animals in labs."

Leary says the numbers are also important if scientists are serious about one of the cornerstones of biomedical research: the three Rs. The biomedical community has been saying for decades that it is dedicated to reducing and replacing animals in labs, and to finding ways to minimize their suffering, she notes. "How can you measure your progress …if you're not even counting the animals?"

Peter Smith, associate director of Yale University's Animal Resources Center, thinks Carbone's estimates are close to the mark. "It's a well–thought-out approach," he says, "and about as good as anyone can do."

Others strongly disagree. The new analysis is "fundamentally flawed," and the numbers are "inaccurate and vastly overstated," says Nadia Rosenthal, scientific director of the Jackson Laboratory, which supplies mouse strains to universities and other institutions. Spending and shipping numbers from Charles River Laboratories, the world's largest commercial distributor of rodents, suggest commercial breeders supply, at most, 15 million mice and rats to U.S. labs every year, Rosenthal says. Combined with what labs themselves likely breed, she says the total number of those animals is likely closer to 20 million.

Bennett faults the study for extrapolating from a handful of top-funded academic institutions to more than 900 other biomedical facilities, including less well-funded academic centers, pharmaceutical companies, and contract research organizations. "If the same ratios of mice and rats don't hold at these places," she says, "the entire exercise is invalid."

> **"Somebody has to be keeping track of these animals. You can't track progress if you can't measure progress."**
> **Larry Carbone**, retired from the University of California, San Francisco

Bennett also takes issue with the idea that omitting mice and rats from the AWA means they are unprotected. She notes that all labs that receive federal funding must answer to institutional animal care and use committees, which follow the three Rs and oversee the welfare of all lab animals. Those that don't receive federal funds typically submit to AAALAC inspections, she says, which also enforce humane standards.

Still, Cathy Liss, president of the Animal Welfare Institute, which has also fought to include mice and rats in the AWA, says only this federal law gives someone the ability to file a complaint that must be investigated. "These animals merit the same level of protection that other animals get."

Matthew Bailey says that's not realistic. The president of NABR, which has repelled efforts to include rats and mice in the AWA, he says the government doesn't have the time or money to track so many animals. "Now is not the time to be seeking additional restrictions on biomedical research or endeavoring to make it more difficult and more expensive."

Regardless, Carbone says he hopes his study will start an important conversation about whether to count mice and rats. "If we're serious about reducing the number of lab animals we use and curbing the number of painful experiments, somebody has to be keeping track of these animals," he says. "You can't track progress if you can't measure progress." ■

BIOMEDICINE

What causes IBS pain? Maybe immune flare-ups

New study suggests certain foods trigger a localized reaction in the gut

By **Jennifer Couzin-Frankel**

For millions of people with irritable bowel syndrome (IBS), the condition's cramping, diarrhea, and constipation are frustrating and distressing. And it's made worse when doctors sometimes dismiss the symptoms as "all in the mind or between the ears," says Guy Boeckxstaens, a gastroenterologist at KU Leuven.

Boeckxstaens rejects that view. For years he's sought to understand one hallmark of IBS, abdominal pain brought on by eating. Last week, he and more than 40 colleagues presented data to support a new hypothesis: that this pain is caused by a kind of localized allergic reaction to food. "The idea that you can have a specific allergic response going on in the gut" is "a really new concept," says Robin Spiller, a gastroenterologist at the University of Nottingham who was not involved in the study and finds it persuasive.

An early clue to IBS pain emerged about 15 years ago. Researchers found that in IBS patients, "the immune system was different," says Giovanni Barbara, a gastroenterologist at the University of Bologna who led the work. In biopsies of intestinal tissue, immune cells called mast cells were activated. Normally, mast cells act as an alert system, spewing out chemicals such as histamine when threatened with infection or other dangers, like parasites. But these IBS patients had no apparent infections, yet their mast cells were activated and caused nearby nerve cells to fire excessively. Barbara recalls that "nobody believed me." Doubts that IBS pain was rooted in gut biology ran deep.

Boeckxstaens, however, was intrigued and went looking for more evidence. Intestinal infections, ranging from acute food poisoning to mild illness, are followed

by IBS in about 10% of previously healthy people, and some researchers thought low-grade inflammation persisting in the gut after infection might be the trigger. Yet when Boeckxstaens examined intestinal biopsies of IBS patients, he didn't find inflammation.

Instead, he thought a gut infection may disrupt how the organ tolerates protein fragments called antigens, which many foods contain. When battling an infection, the intestine's immune cells may mistake food antigens for enemy forces. A gut reaction to food antigens that persists after infection could explain the pain and cramping that often accompany a meal.

To test the idea, Boeckxstaens and colleagues infected mice with harmful gut bacteria, and at the same time fed them antigens from egg whites. After the gut infection cleared, the mice ingested the antigens again—and this time, they seemed to experience abdominal pain, as measured by stomach muscle contractions. Mice that didn't get egg white protein while they were infected had no trouble.

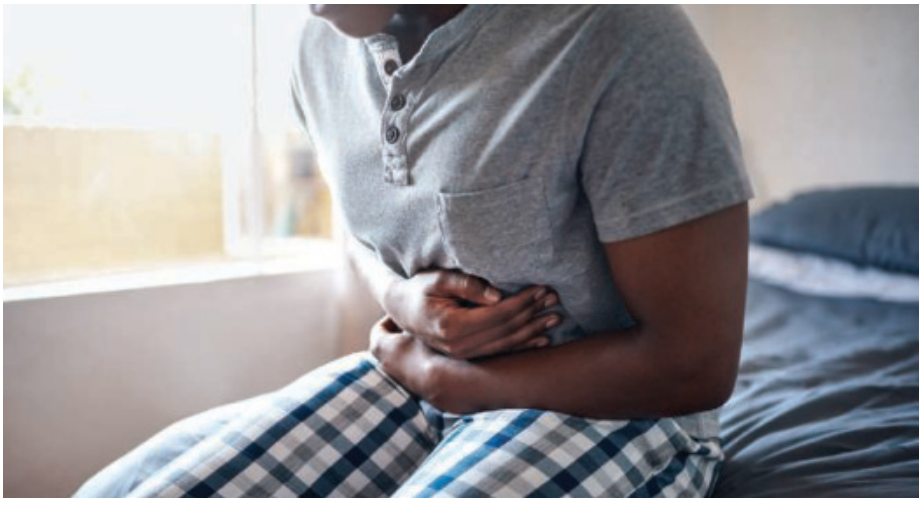
Probing further, the researchers found that after an infection, the egg white protein set off a chain reaction similar to what happens in food allergies. The protein anchored itself to antibodies called immunoglobulin E (IgE), which are bound to mast cells. In the mice—as in people with allergies—mast cells then became activated and released their chemicals. The reaction to egg protein endured for at least 4 weeks after the infection. The persistent reaction underscored what Barbara had observed years earlier in people: If mast cells were shooting out chemicals, neurons could become hypersensitive and more excitable, which the body would interpret as pain.

These new findings could "open people's minds" to the idea that IBS pain is rooted in biology, says Bana Jabri, a pediatric gastroenterologist and mucosal immunologist at the University of Chicago. Jabri has pursued a similar idea with celiac disease, in which patients can't tolerate gluten, reporting with colleagues in 2019 that ingesting gluten causes a strong immune reaction specific to that substance; the reaction in turn drives abdominal symptoms such as pain and nausea.

To find out whether the mouse results applied to patients, Boeckxstaens's group tested 12 people with IBS for four common food allergies: to cow's milk, gluten, wheat, and soy. All were negative. That didn't surprise him: What his mice experienced isn't a classic food allergy, because the animals' immune reaction was localized to the gut. But when the researchers injected these potential allergens rectally, every volunteer had a localized reaction to at least one of the antigens. In a control group of eight healthy people, only two had a borderline gut reaction to either soy or gluten, the researchers reported last week in *Nature*. (Boeckxstaens suspects some unaffected people may have mild reactions, which their gut can tolerate.)

"It's pretty novel and thought-provoking," says Dan Mucida, an immunologist at Rockefeller University. If the mice's localized IgE reaction holds up in people with IBS, he wonders whether a range of foods can trigger it, or only a few. Another line of inquiry involves treatment. Currently, IBS treatments focus on managing symptoms, but if mast cells and IgE are driving them, at least in some cases, immune therapies could prove useful.

Finally, Mucida wonders whether this immune reaction plays a role in cases of IBS induced by triggers other than infection, such as stress. Boeckxstaens is exploring that question now, studying whether in mice, stress alone can induce a similar immune cascade in the gut. ■



The pain of irritable bowel syndrome may be caused by an immune reaction to food in the gut.

PHOTO: PEOPLEIMAGES/ISTOCK

CLIMATE CHANGE

Global temperatures in 2020 tied record highs

Defying cooling from La Niña, warming planet speeds toward breach of climate targets

By **Paul Voosen**

Housebound by a pandemic, humanity slowed its emissions of greenhouse gases in 2020. But Earth paid little heed: Temperatures last year tied the modern record, climate scientists reported last week. Overall, the planet was about 1.25°C warmer than in preindustrial times, a trend that puts climate targets in jeopardy, according to jointly reported assessments from NASA, Berkeley Earth, the U.K. Met Office, and the National Oceanic and Atmospheric Administration.

The annual update of global surface temperatures—an average of readings from thousands of weather stations and ocean probes—shows 2020 essentially tied records set in 2016. But the years were nothing alike. Temperatures in 2016 were boosted by a strong El Niño, a weather pattern that warms the globe by blocking the rise of cold deep waters in the eastern Pacific Ocean. Last year, however, the Pacific entered La Niña, which has a cooling effect. That La Niña didn't provide more relief is an unwelcome surprise, says Nerilie Abram, a climate scientist at Australian National University. "It makes me worried about how quickly the global warming trend is growing."

The past 6 years are the six warmest on record, but the warming of the atmosphere is unsteady because of its chaotic nature. The ocean, which absorbs more than 90% of the heat from global warming, displays a steadier trend, and here, too, 2020 was a record year. The upper levels of the ocean contained 20 zettajoules (10^{21} joules) more heat than in 2019, and the rise was double the typical annual increase, scientists reported last week in *Advances in Atmospheric Sciences*. The subtropical Atlantic



In 2020, wildfires raged in Siberia's vast forests, one of many effects of record high temperatures.

Ocean was particularly hot, fueling a record outbreak of hurricanes, says Lijing Cheng, a climate scientist at the Chinese Academy of Sciences's Institute of Atmospheric Physics who led the work.

This heat, monitored down to 2000 meters by a fleet of 4000 robotic probes, is spreading deeper into the ocean while also migrating toward the poles. An extreme heat wave struck the northern Pacific, killing marine life. For the first time, warm Atlantic waters were seen penetrating into the Arctic Ocean, melting sea ice from below and reducing its extent nearly to a record low (*Science*, 28 August 2020, p. 1043). The warming ocean and melting ice sheets are raising sea levels by 4.8 millimeters per year, and the rate is accelerating (*Science*, 20 November 2020, p. 901).

On land, 2020 was even more relentless, with temperatures rising 1.96°C above preindustrial levels, a clear record, Berkeley Earth reported. It was the warmest year ever in Asia and Europe and tied for the warmest in South America. Russia was particularly hot, breaking its previous record by 1.2°C, while swaths of Siberia were 7°C warmer than in preindustrial times, leading to large-scale fires and thawing permafrost that caused buildings to founder and set off oil spills (*Science*, 7 August 2020, p. 612). "Siberia was crazy," says Zeke Hausfather, a climate scientist at the Breakthrough Institute and coauthor of the Berkeley Earth analysis. "That heat would effectively be impossible without the warming we've seen."

In Australia, record-setting heat and drought fueled catastrophic bushfires at the start of 2020. Fires torched nearly one-quarter of southeastern Australia's forests and destroyed 3000 homes. Climate change was to blame for the country's "Black Summer," Abram and co-authors concluded in a study published this month in *Communications Earth & Environment*.

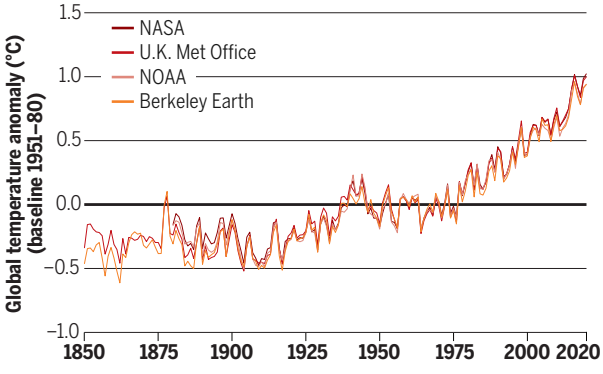
Meanwhile, in the United States, unprecedented heat came to the desert Southwest, which is already warming faster than the rest of the country. Phoenix wilted under its hottest summer ever, averaging 36°C. Arizona's Maricopa county, home to Phoenix, is a leader in addressing heat exposure, yet its heat deaths have hit a new record each year since 2016. In 2020, the number approached 300, a jump of some 50% over the previous year, says David Hondula, a climatologist who studies heat mortality at Arizona State University, Tempe. "It was just off the charts in terms of heat."

Although the global economic slowdown of the COVID-19 pandemic cut carbon dioxide (CO_2) emissions by some 7%, atmospheric CO_2 is long-lived, and warming from previous emissions is preordained. In any case, the drop in emissions is unlikely to last. Later this year, in May, before photosynthesis in the Northern Hemisphere draws down CO_2, the U.K. Met Office predicts that levels of atmospheric CO_2 will pass 417 parts per million for several weeks, 50% higher than preindustrial levels. Only dramatic action by the world's countries, far beyond existing efforts, can begin to halt this build up, Cheng says.

Should the current rate of warming continue, the world will breach the targets set in the Paris climate agreement—limiting warming to 1.5°C or 2°C—by 2035 and 2065, respectively. But Hausfather says it's quite possible that warming, which has largely held steady for the past few decades at 0.19°C per decade, will actually speed up. The rate of warming over the past 14 years is well above the long-term trend. The debate now, he says, is whether that is an omen of an even darker future. ■

Turning up the heat

Temperatures in 2020 tied 2016's record levels. They were about 1°C above a 1951–80 average, or 1.25°C hotter than preindustrial levels.



CREDITS: (PHOTO) DENIS BUSHKOVSKY/TASS/NEWSCOM; (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) MET OFFICE; NASA; BERKELEY EARTH; NOAA

FEATURES



SEPARATION

Phase separation, an idea about how cells organize their contents and functions into dropletlike compartments, has divided biologists

By **Mitch Leslie**

For 7 years as president of the Howard Hughes Medical Institute, Robert Tjian helped steer hundreds of millions of dollars to scientists probing provocative ideas that might transform biology and biomedicine. So the biochemist was intrigued a couple of years ago when his graduate student David McSwiggen uncovered data likely to fuel excitement about a process called phase separation, already one of the hottest concepts in cell biology.

Phase separation advocates hold that proteins and other molecules self-organize into denser structures inside cells, like oil drops forming in water. That spontaneous sorting, proponents assert, serves as a previously unrecognized mechanism for arranging the cell's contents and mustering the molecules necessary to trigger key cellular events. McSwiggen had found hints that phase separation helps herpesviruses replicate inside infected cells, adding to claims that the process plays a role in functions as diverse as switching on genes, anchoring the cytoskeleton, and repairing damaged DNA. "It's pretty clear this process is at play throughout the cell," says biophysicist Clifford Brangwynne of Princeton University.

The pharmaceutical industry is as excited as some academic researchers, given studies linking phase separation to cancer, amyotrophic lateral sclerosis (ALS), diabetes, and other diseases. Dewpoint Therapeutics, a startup pursuing medical treatments targeting cellular droplets, recently signed development deals worth more than $400 million with pharma giants Merck and Bayer. And three other companies looking to exploit the process opened their doors late last year. Reflecting that enthusiasm, *Science* picked phase separation as a runner-up in its 2018 Breakthrough of the Year issue.

Tjian says he was agnostic at first about the importance of the process. But after McSwiggen's findings inspired him and colleagues to look more closely at the range of claims, the researchers began to have doubts. Tjian and a camp of similarly skeptical biologists now argue that the evidence that liquidlike condensates arise naturally in cells is largely qualitative and obtained with techniques that yield equivocal results—in short, they believe much of the research is shoddy.

Moreover, the contention that those intracellular droplets perform important roles "has gone from hypothetical to established dogma with no data," says Tjian, who stepped down as president of Howard Hughes in 2016 and now co-directs a lab at the University of California (UC), Berkeley. "That to me is so perverse and destructive to the scientific discovery process."

Although proponents of phase separation bridle at some of those criticisms, many scientists agree that the research requires a jolt of rigor. "I don't think the whole field is bunk," says biophysicist Stephanie Weber of McGill University. "But we do need to be more careful" in identifying instances of phase separation in cells and ascribing functions to them.

The process may be less important than many scientists now assert, adds quantitative cell biologist Amy Gladfelter of the University of North Carolina, Chapel Hill. Some researchers, she says, have tried to make it "the answer to everything."

PHASE SEPARATION COULD ANSWER a fundamental question that has nagged biologists for more than 100 years: How do cells arrange their contents so that the molecules necessary to carry out a particular job are in the right place at the right time? One obvious way is with internal membranes, such as those fencing off the Golgi bodies and mitochondria. Yet many other well-known cellular structures, including the nucleolus—an organelle within the nucleus—and the RNA-processing Cajal bodies, lack membranes.

Phase separation is an appealing answer. Many proteins sport sticky patches that attract other proteins of the same or a different type. Test tube studies have shown that under certain conditions, such as when protein concentration climbs above a certain level, the molecules may begin to huddle, forming dropletlike condensates. Researchers understand the mechanics best for proteins, but nucleic acids such as RNA could also aggregate with proteins. If the process happens in the cell, it could generate and maintain organelles and permit unique functions. "It's a principle that could explain how many things in the cell and nucleus are organized," says biophysicist Mustafa Mir of the University of Pennsylvania, who as a postdoc once worked with Tjian.

Although biologists mooted a role for intracellular droplets as far back as the 1890s, evidence that they are vital began to coalesce a little over 10 years ago. Brangwynne, then a postdoc at the Max Planck Institute of Molecular Cell Biology and Genetics, was tracing P granules, flecks of protein and RNA that, in nematode embryos, mark the cells that go on to produce sperm and eggs. To

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*



ANXIETY

observe the granules' movements, Brangwynne squeezed worm gonads that harbor the structures between two microscope cover slips. Under pressure, P granules responded not like solids but like liquids, flowing along the surface of the nucleus and dripping off, he and colleagues reported in *Science* in 2009. The granules' watery behavior "was mind-blowing. It was so different than anything in cells," says Weber, a former postdoc of Brangwynne's.

In 2012, Brangwynne and colleagues saw similar fluid features in the nucleolus, a dense mix of proteins, RNA, and DNA that manufactures ribosomes, the cell's protein factories. The same year, biophysicist Michael Rosen of the University of Texas Southwestern Medical Center and colleagues showed that three proteins that collaborate to organize part of the cytoskeleton form liquid droplets in a test tube solution. They found that the process speeds the assembly of one type of skeletal fiber in vitro—and might do the same in the cell. Scientists have since reported dozens of examples of cellular structures that are round, prone to fuse, and tend to bead on or flow across surfaces—hallmarks of droplets formed by phase separation (see graphic, p. 338).

To confirm that a molecular gathering in a cell is a liquid and not something more solid, scientists often deploy a technique called fluorescence recovery after photobleaching (FRAP). Using a cell that contains fluorescent proteins, researchers zap the region in question with a laser to darken the molecules and then trace how long the fluorescence takes to diffuse back in from other parts of the cell. A liquid, which the fluorescent proteins easily penetrate, should light up more quickly than a solid. Another test involves applying 1,6-hexanediol, a compound that fractures some of the molecular interactions that hold droplets together, to determine whether the structure dissolves.

Rosen notes that three papers published last year in *Cell* offer some of the strongest evidence for phase separation in cells. One, from Brangwynne's lab, showed a particular protein had to reach a threshold concentration in cells to allow formation of stress granules—organelles that pop up during hard times and have been attributed to phase separation. The other two studies also identified threshold conditions for phase separation. Because a threshold is an attribute of the process, the studies provide "good but not perfect data that these structures are going through phase separation," Rosen says.

Many researchers are now convinced that phase separation explains many aspects of cell organization and function. Several research groups have reported that the mechanism helps convene the hundreds of proteins that carry out transcription, the process of reading DNA to produce the RNA instructions for making proteins. Similar molecular corralling may underlie functions including memory in fruit flies, immune cells' responses to pathogens, DNA silencing, transmission of nerve impulses across synapses, and reproduction of SARS-CoV-2, the pandemic coronavirus.

Conversely, phase separation may cause disease when it goes awry. In 2018, for example, biophysicist Tanja Mittag of St. Jude Children's Research Hospital and colleagues revealed that mutations that promote several kinds of tumors disrupt the ability of the protein SPOP, which helps eliminate proteins that spur growth of cancer cells, to form droplets in test tube solutions. The researchers proposed that phase separation is key to SPOP's cleanup function in cells, and thwarting it allows cancer-promoting proteins to accumulate.

Faulty phase separation could also spur damage by aiding the formation of the toxic intracellular inclusions, or protein globs, that amass in neurodegenerative illnesses such as ALS, Alzheimer's disease, and Parkinson's disease. For example, in some ALS patients the protein FUS is mutated and forms inclusions in their neurons. In the test tube, the mutated protein condenses into droplets that then morph into furry knots of fibers resembling the inclusions. In 2018, biochemist Dorothee Dormann of the Ludwig Maximilian University of Munich and colleagues discovered a possible reason: The mutated version of FUS shrugs off a protein bodyguard that prevents the normal variety from undergoing phase separation and clumping in the test tube.

YET THAT SATISFYING PICTURE may be growing murky as more researchers have raised doubts about phase separation. In 2019, for instance, scientists organized a debate at Wiston House, a posh 16th century manor south of London, in part to mull whether the process helped control gene activity. About 30 participants hashed over the evidence that the process occurs in cells with

the help of "free-flowing champagne," recalls Mir, one of the presenters. The group's conclusion, he says, was that the support for many putative cases of phase separation in cells is shaky.

Tjian, who was not at the meeting, came around to a similar conclusion because of new data from McSwiggen. McSwiggen's early evidence showed that in herpesvirus-infected cells, the replication compartments—clusters of protein and DNA that help produce new copies of the pathogen—are round and merge with each other, suggesting they result from phase separation.

After tracking individual proteins within cells, though, McSwiggen and colleagues determined the molecules diffuse just as fast through the compartments as through the rest of the nucleus. In a true condensate, molecular crowding should have hindered diffusion. Other researchers found the negative evidence compelling when it was published later in 2019, soon after the Wiston House debate. The study is "a really important cautionary tale," Weber says.

The results spurred Tjian, McSwiggen, Mir, and Xavier Darzacq, a cell biologist who co-directs the UC Berkeley lab with Tjian, to scrutinize the phase separation literature. Later that year, in a December 2019 issue of *Genes and Development*, they published a scathing review of 33 studies that claimed to detect the process in cells. Tjian says he was "really disappointed by the quality of the papers." The evidence, he and his co-authors wrote, was "often phenomenological and inadequate to discriminate between phase separation and other possible mechanisms."

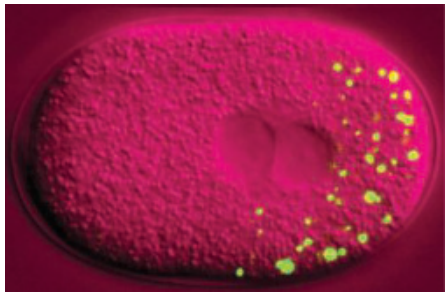
Too often, he and the other review authors asserted, researchers looking for phase separation rely on qualitative indicators—shape, for example—rather than quantitative data. Moreover, because many intracellular structures possibly formed by phase separation are so small, they are near what's known as the diffraction limit of traditional light microscopes. As a result, the structures may look like fuzzy orbs, but their real shape isn't discernible.

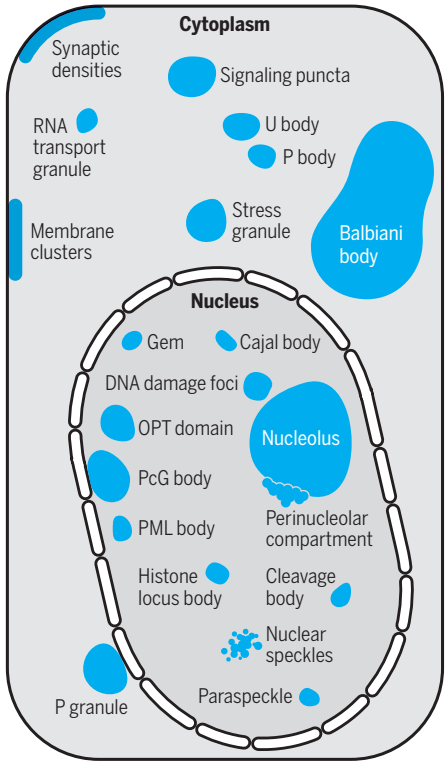
Tjian and colleagues also chastised researchers for often assuming the protein concentration in a cell is high enough to trigger phase separation, instead of actually measuring it. Overinterpretation "is rampant" in this type of research, Tjian says.

The scientists questioned the FRAP measurements that underpin many claims of phase separation. In the hands of different scientists, the group noted, FRAP recovery rates for the same molecule can range from less than 1 second to several minutes, indicating the technique is too variable to confirm phase separation. Darzacq adds that

Dropping in

P granules (top, green), pockets of protein and RNA in early worm embryos that mark where sperm or egg cells will arise, have become the classic example of phase-separated regions in cytoplasm. But researchers propose that many other cellular features (bottom, blue), including some in the nucleus, form in the same way.





FRAP "only shows you have a liquid. You have liquid everywhere in the cell." Many of the congregations that researchers have identified with FRAP or other techniques are probably transient collections of molecules that only last a few seconds, Darzacq and Tjian say.

The review was "an invitation for all of us to proceed with a more careful and thoughtful in-depth analysis of cellular condensates," says molecular biophysicist Sua Myong of Johns Hopkins University. Although some scientists have been meticulous, "it has not been true of the field," Rosen adds.

Brangwynne says he, too, sees value in the critique. "I agree that we need quantitative approaches." For example, he concurs that researchers need to be more rigorous when interpreting imaging results so that "every diffraction-limited blob" isn't declared an example of phase separation.

Other recent papers have also raised doubts about cases of phase separation. In 2019 in *Non-Coding RNA*, Weber and a coauthor weighed the support for phase separation in the cell nucleus and concluded that solid data back its role in forming three structures, including the nucleolus, but not two other structures commonly attributed to the process.

And in April 2020 in *Molecular Cell*, biophysicist Fabian Erdel of the Center for Integrative Biology in Toulouse, France, and colleagues published a new investigation of heterochromatin—silenced regions of the genome in which DNA coils tightly with various proteins. Previous work suggested phase separation of the intracellular protein HP1 helped stretches of heterochromatin bunch up. But Erdel's team discovered that HP1 didn't form stable liquid droplets in mouse cells and that the size of the densely packed DNA regions didn't depend on the amount of the protein.

Brangwynne and other researchers argue that even if some individual findings cited by Tjian and colleagues remain in dispute, the field is making progress toward more solid results. To provide some of the rigor of test tube studies, he and his team have developed a technique for seeding cells with what they call corelets, combinations of molecular fragments that cluster when exposed to light. The corelets trigger droplet formation in cells, allowing the researchers to more precisely probe what protein concentrations are necessary for phase separation and which parts of the molecule are required for the behavior. Even Tjian and colleagues give the approach high marks.

Mir, who has been skeptical of much of the evidence for phase separation, agrees that the field seems to be moving away from the "everything is phase separation" stage to a more nuanced discussion of the formation and functions of condensates. "It's like any supertrendy thing in science. The noise subsides, and you are left with the truth."

To get to that truth, however, researchers "desperately need" new tools and a better understanding of the basic rules for how condensates form in cells, Gladfelter says. Scientists also need patience, she says, noting the field "tried to grow up and answer everything really fast." But she's confident researchers will eventually sort out the real importance of phase separation in cells. "Give us time. We'll get there." ■

CREDITS: (GRAPHIC) V. ALTOUNIAN/*SCIENCE*; (IMAGE) C. BRANGWYNNE *ET AL.*, *SCIENCE*, 324, 5935, 1729 (2020)



HOW ARE LSU SCIENTISTS HELPING FIGHT COVID-19?

LSU researchers' foresight may have saved countless lives in the fight against the novel coronavirus.

To alleviate some of the burden placed on Louisiana hospitals, Stephania Cormier (pictured), a respiratory immunologist and the Wiener Chair Professor in the LSU Department of Biological Sciences, along with LSU Veterinary Medicine's Rebecca Christofferson, an infectious disease and emerging viruses expert, worked to provide much-needed support to the local community as Louisiana faced a first surge in March 2020. Through an expedited process, they received the necessary authorizations and the LSU River Road Testing Lab was born.

The LSU River Road Testing Lab began providing the most critical patients their results within hours, rather than days.

The lab conducted PCR-testing, which detects genetic components specific to SARS-Cov-2. Within the first two weeks, the researchers created and processed more than 1,200 sample kits and diagnostic tests for SARS-CoV2, the cause of COVID-19.

"This is what science should be," said Christofferson, who is also a LSU College of Science alumna. "Science as scholarship is wonderful, and we come into academic research with a romantic idea of hallowed halls and ideation. But the privilege of exploring every day comes with a responsibility: When there is an emergency that you have the ability and capacity to respond to, get in the trenches and go."

Join LSU's community of fearless explorers.
lsu.edu/science

LSU | College of Science

YOUR QUESTION NEXT.



INFECTIOUS DISEASES

Drivers of mosquito mating

Light cycle genes and environmental and chemical cues modulate male mosquito swarming

By **Nicholas C. Manoukis**

At first glance, the sex lives of mosquitoes may seem an esoteric topic. Yet, elucidating the details of mosquito mating may affect hundreds of millions of human lives each year. *Anopheles* mosquitoes are the principal vectors of the parasites that cause malaria in Africa, where in 2018, 93% of the world's estimated 228 million cases and 94% of its 405,000 malaria deaths occurred (*1*). On page 411 of this issue, Wang *et al.* (*2*) link clock gene expression, light, and temperature to the formation of male swarms and mating of *Anopheles* mosquitoes. They also suggest a role for the desaturase 1 (*desat1*) gene in the production of the cuticular hydrocarbon (surface chemicals that serve multiple roles for insects) heptacosane, which they found stimulates mating. These findings add to our knowledge of molecular factors and their interaction with the environment that together drive mating behavior in these mosquitoes.

A critical step in the life history of anopheline mosquitoes is the mating swarm (*3*). These are crepuscular airborne aggregations (leks) composed almost entirely of males, numbering from a few individuals to thousands. Females enter these swarms, and coupling occurs within their bounds, with insemination happening in flight or on the ground near the swarm. Females generally mate once in their lives. Factors thought to be involved in swarm formation include basic environmental variables such as light (*3*), but mating might depend on complex behaviors such as speed matching and harmonization of wingbeat frequencies (*4*, *5*). The details are important to the mating success of any male *Anopheles* mosquito.

Male success in mating swarms is relevant to the most technologically advanced and promising interventions to reduce the burden of mosquito-borne diseases—such as those built on gene drive systems (*6*). Gene drive systems are based on the release of organisms whose genomes have

Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Agricultural Research Service, U.S. Department of Agriculture, Hilo, HI 96720, USA.
Email: nicholas.manoukis@usda.gov

PHOTO: SOLVIN ZANKL/NPL/MINDEN PICTURES

Mosquitoes (*Anopheles stephensi*) used for malaria research in laboratories are fed on blood through a thin membrane.

been modified or engineered to spread a desired allele or trait (such as resistance to the parasites that cause malaria) through a population. Success will depend on the release of genetically modified males that will be able to mate with wild females. Beyond gene drive strategies, in mosquitoes it is understood that only males can be released as part of any genetic pest control (GPC) program (*7*); females feed on blood to lay eggs, and releasing insects that will feed on humans is widely unacceptable.

The grandparent of all GPC insect control technologies is the sterile insect technique (SIT), which consists of releasing mass-produced, irradiation-sterilized males in sufficient numbers to "flood" the wild male population and thereby compete for wild female insects. This leads to the production of inviable offspring, suppression of the population, and potentially the eradication of the targeted species (*8*). Although this method has been successful—for example, eradicating the screwworm (a serious pest of livestock) from the southern United States, Mexico, and most of Central America—the difference between feasible and successful programs and those that falter rests heavily on the competitiveness of released sterile males in securing wild female mates. "Traditional" SIT has not succeeded against mosquitoes, and the complexities of mating behavior might be a contributing reason.

Wang *et al*. found in swarming males upregulation of two clock genes, *tim* (*timeless*) and *per* (*period*), that they showed are important to swarming and successful mating. They also measured the effect of two basic environmental factors (temperature and light). Integrating environmental variables into the experiment was the key to determining how gene expression operates in a realistic context and led to major insights, such as finding that photoperiod (the daily duration of light and dark) affects both clock genes and *desat1*. Previous studies have tended to focus on one type of factor (environmental, genetic, or molecular), but clearly these separate factors do not operate independently.

A further integrative feature of the study of Wang *et al*. is that experiments were conducted under semi–field conditions in addition to the laboratory. The cross-disciplinary collaboration that made this possible is an important model for developing our understanding of the whole phenomenon of mating in mosquitoes and for applying the results of experiments to new vector control programs.

Wang *et al*. open the door a little wider to a future in which mass-produced lines of male mosquitoes are better adapted and more competitive with wild males. Achieving this goal may be as simple and direct as enhancing expression of *desat1* to improve the mating success of released males or may result from a more subtle manipulation of swarming behavior through environmental and gene expression factors (for example, optimizing mass rearing or engineering superior swarm competitors). A related possibility is that the swarming ability of the males reared for release might be assessed by means of genetic screen. This may prove critical for determining when to switch out or refresh mass-reared lines, all of which suffer from genetic inbreeding. Enhancing the probability of success of GPCs would be a welcome instance of the integration of molecular and genetic information with physical and environmental factors, delivering on decades of promises in medical entomology (*9*).

By integrating molecular and environmental factors, Wang *et al*. make a crucial contribution to the growing body of research on mosquito mating, which now spans chemical cues to motion coordination (*10*, *11*). Future work should expand on their efforts, particularly by integrating more sophisticated physical variables with gene expression data (for example, oscillating temperature profiles rather than a single fixed overnight temperature for mass rearing). In the end, the more that is known about mosquito mating generally and *Anopheles* species specifically, the better our ability to intrude in just the right way to reduce disease transmission around the world. ■

REFERENCES AND NOTES

1. World Health Organization (WHO), World malaria report 2019 (WHO, 2019), p. 232; https://www.who.int/publications/i/item/9789241565721.
2. G. Wang *et al*., *Science* **371**, 411 (2021).
3. P. I. Howell, B. G. Knols, *Malar. J.* **8** (suppl. 2), S8 (2009).
4. S. Butail *et al*., *J. R. Soc. Interface* **9**, 2624 (2012).
5. C. Pennetier, B. Warren, K. R. Dabiré, I. J. Russell, G. Gibson, *Curr. Biol*. **20**, 131 (2010).
6. J. Champer, A. Buchman, O. S. Akbari, *Nat. Rev. Genet*. **17**, 146 (2016).
7. C. F. Curtis, *Biol. J. Linn. Soc. Lond*. **26**, 359 (1985).
8. E. F. Knipling, *J. Econ. Entomol*. **48**, 459 (1955).
9. W. K. Reisen, *Genet. Evol* **28**, 573 (2014).
10. R. Mozūraitis *et al*., *Nat. Ecol. Evol*. **4**, 1395 (2020).
11. D. Shishika, N. C. Manoukis, S. Butail, D. A. Paley, *Sci. Rep*. **4**, 6318 (2014).

ACKNOWLEDGMENTS

The author is employed by the U.S. Department of Agriculture (USDA), Agricultural Research Service. Opinions in this publication are those of the author and should not be construed to represent any official USDA or U.S. government determination or policy. USDA is an equal opportunity provider and employer.

10.1126/science.abf7917

PHASE TRANSITIONS

Snapshots of a light-triggered transition

A new method captures real-space movies of ordered states on ultrafast time scales

By **Anshul Kogar**

In an iconic 1964 photo of a bullet piercing an apple, Harold "Doc" Edgerton captured a snapshot that took a mere millionth of a second (*1*). The same working principle behind his high-speed images have been adopted today to study the evolution of ordered states of matter after excitation by extremely short light pulses. To trace the time evolution of these ordered states after excitation, in which the atoms move on the femto- to picosecond time scale (10^{-12} to 10^{-15} s), images need to be taken with both extraordinary temporal and spatial resolution. On page 371 of this issue, Danz *et al*. (*2*) show that this goal can be accomplished using time-resolved dark-field electron microscopy by obtaining real-space snapshots of a charge density wave phase transition triggered with pulses of light.

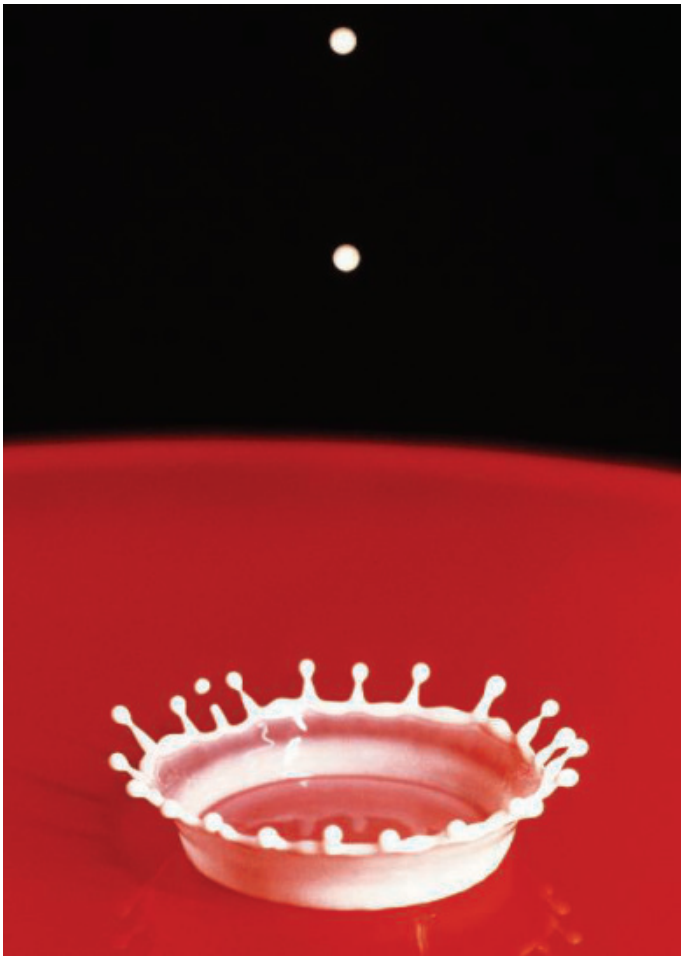
The famous high-speed photos taken by Doc Edgerton help us to understand how these experiments were carried out (see the figure). This image is captured after a drop lands in an otherwise quiescent pool of milk (*3*). Taking these kinds of photos, which Edgerton started pursuing in the 1930s, required ingenuity, because shutter speeds even on most modern cameras are only in the millisecond range. Edgerton, an early innovator in electronic flashtube technology, simply turned off the lights in the room and used a microsecond flash to capture the images at a particular instant. Despite the long camera exposure, light is only captured at the instant of the flash. As long as the perturbing drop (the "pump") is appropriately synchronized with the flash (the "probe"), the result is a perfectly timed snapshot.

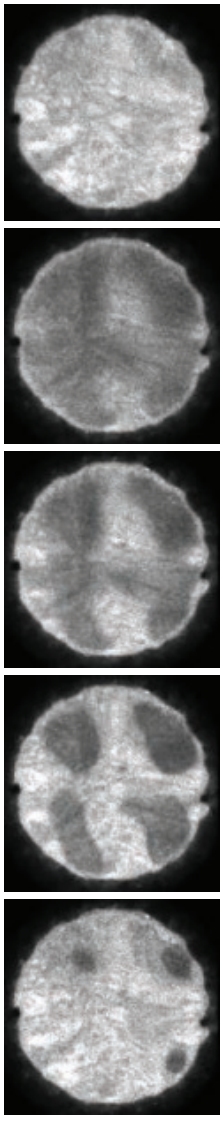
To synchronize the events, the drop falling from a pipette triggers an electronic circuit that is used to delay the flash precisely to

Department of Physics, University of California, Los Angeles, Los Angeles, CA, USA. Email: anshulkogar@physics.ucla.edu

Capturing fast physics in a flash

This famous image from Harold Edgerton (below) was captured on the microsecond time scale using a synchronized electronic flashtube. Analogously, the evolution of a charge density wave transition is captured on the picosecond scale (right) with time-resolved dark-field imaging in an electron microscope.





the moment that the drop hits the pool. By electronically delaying the time between the pump and probe events, a sequence of images can be captured to illustrate the splash evolving in time. Using this method, each photo in a sequence is actually captured with different milk drops. Thus, a sequence of images patched together is not a "movie" in the truest sense, but so long as the event is repeatable, a movie can be meaningfully constructed. Capturing the evolution of the milk drop disturbing the pool away from its quiescent equilibrium state inspires those today who use the pump-probe technique.

Before the advent of pump-probe spectroscopy, states of matter were studied at, or very close to, thermal equilibrium. Studies of glass provide a counterpoint, but the time evolution for glasses is extremely slow. In the past two decades, building on advances in laser technology, the first glimpses have begun into how ordered states evolve on femtosecond time scales after being driven far from equilibrium with intense light pulses. Major goals of this enterprise are to use the light pulses to manipulate and control existing states of matter and to discover different ones that do not exist under equilibrium conditions. For example, prior studies have hinted that light pulses can substantially raise transition temperatures in molecular superconductors (*4*, *5*) and that pulses can create quantum mechanical hybrid photon-electron states, so-called Floquet-Bloch states, in insulating materials (*6*, *7*).

Because of the extreme conditions to which matter is subjected in these experiments, some of the most basic and prized experimental probes are unusable. Even measurements of the electrical resistivity, perhaps the most run-of-the-mill among probes of near-equilibrium matter, become monumentally difficult to perform on the femtosecond time scale. Thus, attention has shifted to expanding the experimental repertoire to better capture the far-from-equilibrium dynamics of ordered matter in pursuit of what happens on extremely short time scales. Only in the past 10 to 15 years have methods such as time- and angle-resolved photoemission spectroscopy (*7*, *8*) and ultrafast electron scattering (*9*–*11*) matured to present us with meaningful insights into condensed phases of matter.

Danz *et al*. substantially boost our experimental arsenal by capturing real-space images of the spatiotemporal order parameter dynamics during a light-triggered phase transition. They take snapshots, in the manner of Doc Edgerton, of charge density wave material 1*T*-TaS_2 (see the figure). A charge density wave is an ordered state that gives rise, among other phenomena, to a periodic lattice distortion. In the authors' experiment, a femtosecond light pulse melts this density wave locally, and over time, the sample returns to its distorted state. Images in their scheme are captured by using electron pulses in a conventional transmission electron microscope, but the authors spearhead two advances. They use an inhomogeneous light excitation to melt the density wave in specific regions within their field of view. And instead of exposing the camera to all scattered electrons from the sample, they design a sample-specific, tailor-made mask to collect only diffraction peaks from the density wave. Imaging just these diffraction peaks—so-called dark-field imaging—allows the authors to watch how the vaporized regions interact with each other, shrink, and eventually recondense into the charge density wave state.

These never-before-seen images will inspire others to come up with schemes that do not require the use of the tailor-made mask, which currently prohibits the technique from being adopted more widely. The future will no doubt see strategies for masking "dynamically" and applying the method more systematically across materials. Had Doc Edgerton been alive to watch the kinds of movies directed today by those pursuing studies of far-from-equilibrium physics, we imagine that they would evoke in him both awe and a nostalgic flashback. ■

REFERENCES AND NOTES

1. H. E. Edgerton, *Bullet Through Apple* (MIT Museum, 1964).
2. T. Danz, T. Domröse, C. Ropers, *Science* **371**, 371 (2021).
3. H. E. Edgerton, *Birth of the Milk Drop* (National Gallery of Australia, 1934).
4. M. Mitrano *et al*., *Nature* **530**, 461 (2016).
5. M. Buzzi *et al*., *Phys. Rev. X* **10**, 031028 (2020).
6. Y. H. Wang *et al*., *Science* **342**, 453 (2013).
7. F. Mahmood *et al*., *Nat. Phys*. **12**, 306 (2016).
8. F. Schmitt *et al*., *Science* **321**, 1649 (2008).
9. M. Eichberger *et al*., *Nature* **468**, 799 (2010).
10. M. J. Stern *et al*., *Phys. Rev. B* **97**, 165416 (2018).
11. A. Kogar *et al*., *Nat. Phys*. **16**, 159 (2020).

10.1126/science.abf1406

PHOTO (LEFT): H. EDGERTON/SCIENCE SOURCE; (SERIES ON RIGHT) DANZ *ET AL*./*SCIENCE*

Thermopower and harvesting heat

A metric to compare low-grade heat-to-electricity conversion can be misleading

By **Xun Shi**[1] *and* **Jian He**[2]

Harvesting and directly converting the enormous, yet largely untapped, reservoir of low-grade heat into electricity has the potential for immense economic and environmental impacts. To this end, thermoelectrics (*1*), thermionic capacitors (*2*), and thermocells (*3*) are among the state-of-the-art options. Despite distinct working principles of the internal circuit, the trio are heat engines in a general sense because they convert a temperature difference input ΔT into a device voltage output ΔV_D that empowers the electrons in the external circuit to do work. The unit of the $\Delta V_D/\Delta T$ ratio, a numeric performance index of the energy conversion device, is the same as the thermopower. However, it is misleading to equate or compare this ratio to the thermopower, an intrinsic physical property of material described by a rank-2 tensor (*2*–*7*).

The concept of thermopower originates in the Seebeck effect. When a ΔT is applied across an electronic conductor, charge carriers diffuse from the hot end to the cold end until the material's resultant open-circuit voltage difference (ΔV_M) counterbalances further diffusion. The ΔV_M measures the electrochemical potential difference due to the ΔT-induced carrier concentration gradient. The thermopower of an electronic conductor is encoded in the material-specific $\Delta V_M/\Delta T$ ratio with the sign denoting the type of the majority carriers and the magnitude relating to the average entropy per carrier charge (on the order of μV/K) (*8*). Likewise, the concept of ionic thermopower can be defined through the Soret effect of an ionic conductor (on the order of mV/K) (*9*).

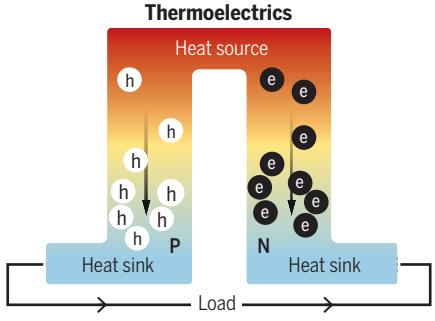
Although the thermoelectric device and thermionic capacitor build on the material's Seebeck effect and ionic Soret effect, respectively, a device is far more than the constituent active materials. Not only is the ΔV_M of the constituent active materials and the electrodes the physical cause of the ΔV_D, chemical reactions can also contribute to the voltage difference. The ΔV_D of the thermocell, a ΔT-driven battery with a redox couple–containing electrolyte (*3*), includes the thermodiffusive and thermogalvanic voltages. The $\Delta V_D/\Delta T$ ratio of the thermocell hardly reflects the constituent active materials' thermopower.

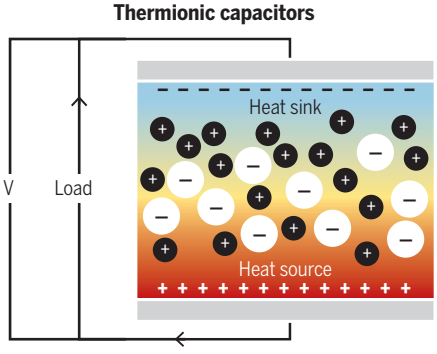
The $\Delta V_D/\Delta T$ ratio is therefore distinct from the thermopower in concept and in practice. It is faulty to reappropriate the formula of thermoelectric figure of merit of a material to assess the performance of a device, or to compare the $\Delta V_D/\Delta T$ ratios at face value. With the ΔT acting as the common driving force of the energy conversion process, the specific cause of the ΔV_D dictates the device's working mode. The thermoelectric device and thermocell work in a continuous mode, and the power density is the primary performance metric. Meanwhile, the working mode of the thermionic capacitor is intermittent, and its performance ought to be assessed by the power density and energy density, like the performance of regular capacitors.

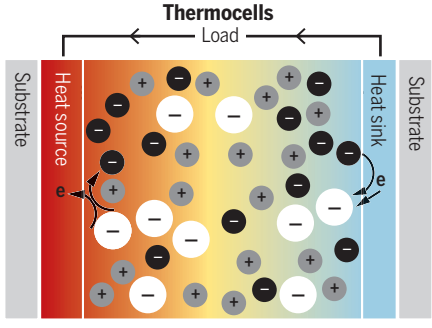
The high $\Delta V_D/\Delta T$ ratios, on the order of mV/K, that have been reported for thermionic capacitors (*2*, *5*) and thermocells (*3*, *6*, *7*)

Making heat work for us

Differences in temperature can be used to drive current with a variety of different devices. Thermoelectrics rely on a p-type and an n-type material that drives electrons (e) or holes (h) from source to sink. Thermionic capacitors rely on separation of positive from negative ions. Thermocells require temperature difference–specific ion and electron cycling. The relative effectiveness of these devices requires care when comparing performance metrics.

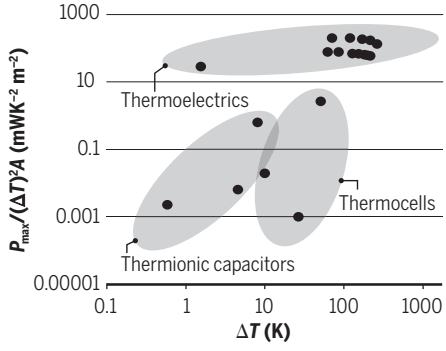


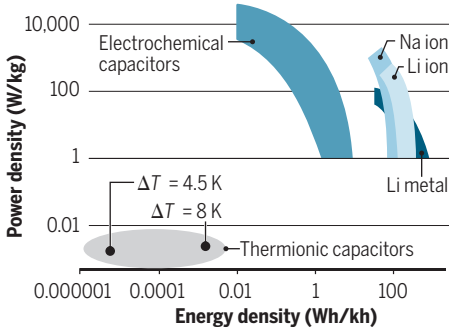




Comparing devices

The normalized power density as a function of the temperature difference (left panel) shows the relative difference in performance between the device types. The power density versus energy density plot (right panel) shows the difference between thermionic capacitors and other methods of storing energy. Large black dots: experimental measurements.





[1]State Key Laboratory of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, China. [2]Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA. Email: xshi@mail.sic.sc.cn; jianhe@g.clemson.edu

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

are vital to the application of wearable electronics, but with certain trade-offs. For thermionic capacitors and thermocells (see the figure), the normalized maximum power densities, $PA^{-1}\Delta T^{-2}$, where P is the maximum power output and A is the module's cross-sectional area perpendicular to the heat flow direction, are at least one order of magnitude lower than for thermoelectric devices (*10*–*12*). This difference is due to large internal resistances. Shortening the interelectrode spacing to reduce the internal resistance falls short as it will diminish the driving force, ΔT. A ΔT = 20 K typical in low-grade heat harvesting energetically equals 1.72 meV, compared to eV-level driving forces in regular capacitors and batteries. The power density and energy density of thermionic capacitors (*2*) are orders of magnitude lower than those of regular capacitors and batteries (*13*) because of weak driving forces.

The multifold nature of the energy conversion process in the device further explains these performance discrepancies, where the challenges and opportunities lie. The efficiency of a multifold energy conversion process is subject to the efficiency of each subprocess and the coupling of the subprocesses. Thermoelectricity is a single-fold process with the simplest working principle. The working media in the internal and external circuits are both electrons. The energy conversion process of the thermionic capacitor is twofold. The ion diffusion is driven by ΔT and the interactions between the accumulated ions and the electrodes (*2*). The energy conversion process of the thermocell includes the ionic diffusion driven by ΔT and by the chemical reaction–induced ionic concentration gradient, the redox reactions at the electrodes, and the interactions between the electrolyte and the electrodes (*3*). Understanding these differences is critical to the integrative development of low-grade heat harvesting technologies in the short and long term. ■

REFERENCES AND NOTES

1. J. He, T. M. Tritt, *Science* **357**, 6358 (2017).
2. D. Zhao *et al.*, *Energy Environ. Sci.* **9**, 1450 (2016).
3. B. Yu *et al.*, *Science* **370**, 342 (2020).
4. C.-G. Han *et al.*, *Science* **368**, 1091 (2020).
5. H. L. Cheng, X. He, Z. Fan, J. Ouyang, *Adv. Energy Mater.* **9**, 1901085 (2019).
6. P. Yang *et al.*, *Angew. Chem. Int. Ed.* **55**, 12050 (2016).
7. A. Taheri *et al.*, *Sustainable Energy Fuels* **2**, 1806 (2018).
8. P. Chaikin, in *Organic Superconductivity*, V. Z. Kresin, W. A. Little, Eds. (Springer, 1990), pp. 101–115.
9. J. Janek, C. Corte, A. B. Lidiard, in *Fluid Mixtures*, W. Köhler, S. Wiegand, Eds. (Springer, 2002), pp. 146–183.
10. T. Kuroki *et al.*, *J. Electron. Mater.* **43**, 2405 (2014).
11. F. Hao *et al.*, *Energy Environ. Sci.* **9**, 3120 (2016).
12. F. Suarez *et al.*, *Appl. Energy* **202**, 736 (2017).
13. P. Simon, Y. Gogotsi, *Nat. Mater.* **19**, 1151 (2020).

ACKNOWLEDGMENTS

We acknowledge P. Qiu, S. Qu, L. Chen, and D. Liebenberg for helpful discussions. X.S. is supported by the National Natural Science Foundation of China (no. 51625205).

10.1126/science.abf3342

IMMUNOLOGY

A host metabolite promotes *Salmonella* survival

Succinate is used by *Salmonella* in different ways when it is in macrophages or the gut

By **Jason P. Lynch**[1,2] *and* **Cammie F. Lesser**[1,2]

The Krebs cycle is a key metabolic pathway that is common to bacteria and mitochondria, often referred to as the central hub of aerobic metabolism. However, on recognition of invading bacteria and other proinflammatory stimuli, the Krebs cycle of macrophages is reprogrammed, shifting from oxidative phosphorylation to glycolysis, resulting in accumulation of intermediate metabolites. These metabolites, which include succinate, act to induce expression of genes involved in promoting inflammation. On page 400 of this issue, Rosenberg *et al.* (*1*) provide evidence that the bacterium *Salmonella enterica* serovar Typhimurium (*S.* Tm), in response to increases in succinate, induces the expression of virulence determinants that promote survival within reprogrammed inflammatory macrophages.

S. Tm is a highly versatile enteric pathogen and one of the most common bacterial causes of gastroenteritis. *S.* Tm is considered an intracellular pathogen that establishes a replicative niche in phagosomes of either intestinal epithelial cells or macrophages. In addition, *S.* Tm colonizes the lumen of the large intestines, ensuring its successful transmission by a fecal-oral route. The pathogenesis of *S.* Tm is dependent on the coordinated regulation of a variety of virulence determinants controlled by complicated regulatory networks that integrate numerous environmental cues, including temperature, oxygenation, osmolarity, and nutrient availability (*2*).

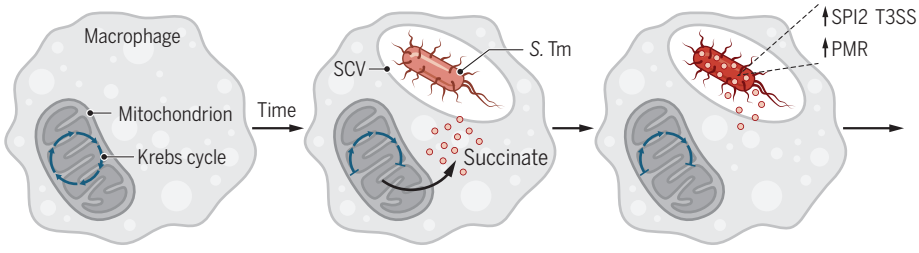
Macrophages are critical mediators of the host innate immune response. They recognize, engulf, and kill invading pathogens, as well as produce proinflammatory cytokines and chemokines to initiate inflammation. Part of this response involves metabolic reprogramming whereby the Krebs cycle in macrophages is broken at multiple steps (*3*). An important consequence of this is the accumulation of Krebs cycle metabolites, including succinate, α-ketoglutarate, citrate, and itaconate, each of which can act as a signaling molecule to modulate the inflammatory response.

To investigate whether *S.* Tm within infected macrophages adapt to host cell metabolic reprogramming, Rosenberg *et al.* used dual RNA sequencing to assess gene expression of both infected macrophages and intracellular bacteria. Analyses of bacterial gene expression changes led to the discovery of enhanced expression of two regulons that promote the survival of

[1]Center for Bacterial Pathogenesis, Division of Infectious Diseases, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA. [2]Department of Microbiology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA. Email: clesser@mgh.harvard.edu

Succinate promotes expression of virulence determinants

The invasion of *Salmonella enterica* serovar Typhimurium (*S.* Tm) into macrophages triggers reprogramming of the Krebs cycle. The subsequent accumulation of succinate is sensed by *S.* Tm within *Salmonella*-containing vacuoles (SCVs), resulting in increased expression of the *Salmonella* pathogenicity island 2 type III secretion system (SPI2 T3SS) and polymixin resistance (PMR) regulons that promote intracellular bacterial replication.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

intracellular *S.* Tm. The first induced regulon, the *Salmonella* pathogenicity island 2 (SPI2) type III secretion system (T3SS), promotes the expression of a protein delivery vehicle that injects proteins into host cells that promote the remodeling of a *S.* Tm–containing phagosome. The second regulon, called the polymixin resistance (PMR) regulon, mediates the expression of genes that promote *S.* Tm resistance to antimicrobial peptides by promoting lipopolysaccharide modifications.

Rosenberg *et al.* found that succinate alone was sufficient to induce expression of both regulons in *S.* Tm and that *S.* Tm cells that lack succinate transporters were impaired in their ability to replicate within infected macrophages. Furthermore, the succinate-mediated effects were unaltered in the absence of an intact bacterial Krebs cycle, suggesting that succinate directly modulates the expression of *S.* Tm regulons akin to how it regulates macrophage gene expression, as opposed to utilizing succinate for metabolic purposes (see the figure).

Prior to invading macrophages and intestinal epithelial cells, *S.* Tm establishes a replicative niche within the large intestines. This is an inhospitable environment to *S.* Tm because this pathogen prefers to grow in the presence of oxygen, and the large intestines are anoxic and densely colonized by obligate anaerobes primarily of the classes *Clostridia* and *Bacteroidia*. However, the recognition of *S.* Tm by intestinal epithelial cells leads to the initiation of innate immune responses and the development of an inflammatory milieu in the gut that promotes *S.* Tm replication.

Neutrophils recruited to the gut release elastase and reactive oxygen species (ROS), reagents to which *S.* Tm is inherently resistant, but to which the *Clostridia* and *Bacteroides* are sensitive (*4*–*6*). The depletion of these anaerobes by neutrophils leads to loss of a source of bacterial short-chain fatty acids, nutrients that promote the viability of intestinal epithelial cells (*7*). Starved of these metabolites, intestinal epithelial cells become dysfunctional, shifting from oxidative phosphorylation to glycolysis, which diminishes their consumption of oxygen. Consequently, oxygen passively diffuses across the epithelium from the surrounding vasculature into the gut lumen. The increased oxygen tension results in expression of all of the enzymes of the *S.* Tm Krebs cycle (*8*). This enables *S.* Tm to perform a complete Krebs cycle, which is fed by microbiota-derived metabolites, including succinate. This promotes aerobic metabolism and results in a burst of *S.* Tm replication within the intestinal lumen.

These studies, together with the findings of Rosenberg *et al.*, demonstrate that *S.* Tm can respond to succinate in two distinct ways. Within the gut lumen, *S.* Tm utilizes microbiota-derived succinate as a nutrient source, whereas within metabolically reprogrammed macrophages, succinate is used as an activation signal. These findings also provide examples of how the Krebs cycle of both host cells and bacteria can be reprogrammed in the context of an infection.

There is growing evidence for microbiota-driven metabolites influencing the growth of invading pathogens. For example, microbiota-derived succinate can also promote blooms of the pathogen *Clostridium difficile* (*9*), and the growth of multiple intracellular pathogens, including *Legionella*, *Francisella*, *Shigella*, and *Anaplasma* species, is driven by host-derived amino acids or metabolites (*10*). Given the intricate ways that successful bacterial pathogens have evolved to combat host defense pathways, it seems likely that succinate and/or other Krebs cycle metabolites are sensed by other pathogens. In this regard, succinate has been reported to induce expression of at least one component of the T3SS of *Citrobacter rodentium*, an extracellular intestinal pathogen, through the regulation of Cra, a global transcriptional regulator (*11*). It will also be interesting to discover how a single metabolite, succinate, is able to regulate the expression of two distinct *S.* Tm regulons. Is it through the modulation of a yet-to-be-identified common regulator? Is it due to the direct binding to one or more transcription factors or regulatory small RNAs? Perhaps succinate acts within the invading pathogen to promote post-translational succinylation of regulatory proteins and thereby alters their function (*12*). More research is needed to understand succinate-mediated gene regulation. ■

REFERENCES AND NOTES

1. G. Rosenberg, D. Yehezkel, D. Hoffman, *Science* **371**, 400 (2021).
2. A. Fàbrega, J. Vila, *Clin. Microbiol. Rev.* **26**, 308 (2013).
3. D. G. Ryan, L. A. J. O'Neill, *Annu. Rev. Immunol.* **38**, 289 (2020).
4. N. Gill *et al.*, *PLOS ONE* **7**, e49646 (2012).
5. I. Sekiro *et al.*, *Gut Microbes* **1**, 30 (2010).
6. C. C. Gillis *et al.*, *Cell Host Microbe* **23**, 54 (2018).
7. F. Rivera-Chávez *et al.*, *Cell Host Microbe* **19**, 443 (2016).
8. L. Spiga *et al.*, *Cell Host Microbe* **22**, 291 (2017).
9. J. A. Ferreyra *et al.*, *Cell Host Microbe* **16**, 770 (2014).
10. A. Best, Y. Abu Kwaik, *Trends Microbiol.* **27**, 550 (2019).
11. M. M. Curtis *et al.*, *Cell Host Microbe* **16**, 759 (2014).
12. J. Liu, C. Qian, X. Cao, *Immunity* **45**, 15 (2016).

ACKNOWLEDGMENTS

The authors are supported by the National Institutes of Health (R01DK113599, R01AI128360, R21AI144716) and the Kenneth Rainin Foundation. C.F.L. is also the Brit d'Arbeloff MGH research scholar and J.P.L. a Crohn's and Colitis Foundation Association Postdoctoral Scholar. C.F.L. is on the scientific advisory board of Synlogic Therapeutics.

10.1126/science.abf8414

ORGANIC CHEMISTRY

Racing against unwanted isomerization

Control of relative reaction rates paves the way for (*Z*)-alkene stereochemical retention

By **Steven J. Malcolmson**

Carbon-carbon double bonds are found in a diverse array of natural products as well as synthetic pharmaceuticals and agrochemicals. A common substitution pattern is for each carbon to bear one proton and one other functional group. Because the double bond lacks free rotation, the coplanar geometry creates two possible stereochemical isomers. The lower-energy trans (*E*) isomer has protons on opposite sides of the double bond, and the higher-energy cis (*Z*) isomer has the protons on the same side of the double bond. Often the *Z* isomer is desired, but its higher energy can hamper its chemical synthesis, and when present in a molecule, given the chance, it can (and usually will) isomerize to the lower-energy *E* isomer (see the figure, top). On page 380 of this issue, Jiang *et al.* (*1*) report the catalytic substitution of (*Z*)-alkene–containing allylic acetates wherein the (*Z*)-alkene is retained in the products.

Allylic substitution is a hallmark reaction in modern organic synthesis (*2*–*4*). The transformation can be promoted by transition metal catalysts containing either early transition metals such as molybdenum or late transition metals such as palladium and iridium (see the figure, bottom). This reaction has been used to form a large variety of chemical bonds, including the carbon-carbon bonds that provide the skeletal backbone of organic molecules. Single-enantiomer catalysts (one of two possible mirror-image isomers) have been used in forming single-enantiomer products composed of stereogenic carbons (those with four different groups attached at one carbon). Much less attention has been paid to controlling the alkene stereochemistry of the products.

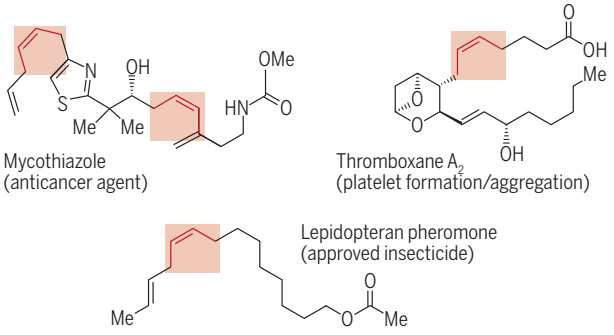
Department of Chemistry, Duke University, Durham, NC 27708, USA. Email: steven.malcolmson@duke.edu

How to catch some (Z)-alkenes

Cis or (Z)-alkenes are the disfavored higher-energy isomer, but a new catalytic route can retain this isomer.

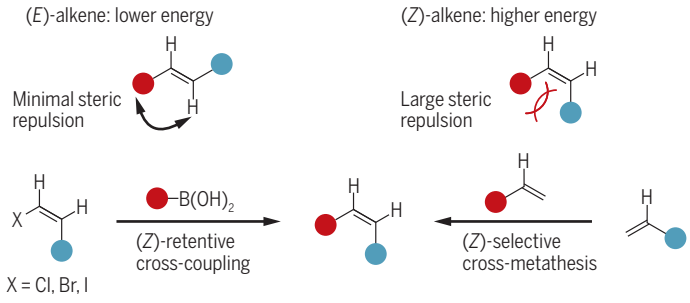
Synthetic relevance

Routes to (*Z*)-alkenes are needed, as they are found in many natural products.



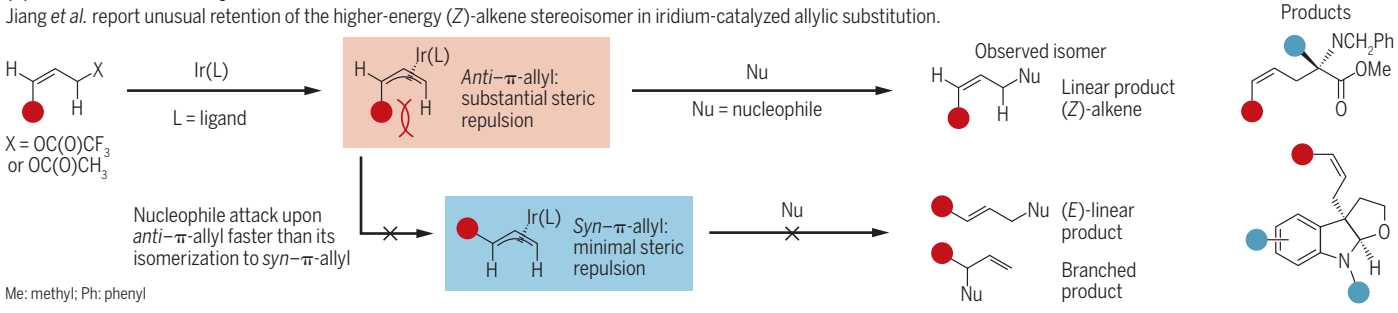
Synthetic challenges and solutions

The energetic difference between (*E*) and (*Z*) isomers (top) and established approaches to (*Z*) isomers (bottom) are shown.



(Z)-isomer retention in allylic substitution

Jiang *et al.* report unusual retention of the higher-energy (*Z*)-alkene stereoisomer in iridium-catalyzed allylic substitution.



Me: methyl; Ph: phenyl

Many allylic substitutions generate cyclic alkenes, which impart a geometric constraint for (*Z*)-alkene formation. However, in acyclic systems, when stereoisomers might be formed, it is the (*E*)-alkene that is typically observed. The favoring of the (*E*)-alkene can be understood in terms of two key mechanistic steps: the displacement of the leaving group (ionization) by the transition metal catalyst to form a metal–π-allyl complex, and the subsequent attack of a nucleophile upon this electrophilic intermediate to furnish the product. Unlike the alkenes of the starting materials and products, the metal–π-allyl intermediate itself is not geometrically static but can exist as two stereoisomers that may readily equilibrate. Thus, the lower-energy isomer, called the *syn*–π-allyl, if not formed kinetically by the metal catalyst in the leaving-group ionization step, may be garnered by isomerization of the higher-energy *anti*–π-allyl complex.

The unimolecular isomerization of the π-allyl complex might be expected to be faster on an entropic basis than its bimolecular coupling with the nucleophile. However, Jiang *et al.* illustrate that two related classes of single-enantiomer iridium-based catalysts belie this logic. From (*Z*)-alkene–containing substrates, these researchers show that catalyst displacement of either a trifluoroacetate or acetate leaving group kinetically delivers the expected *anti*–π-allyl complex, but this intermediate is then trapped intermolecularly by nucleophiles faster than its isomerization to the *syn*–π-allyl can occur, leading to (*Z*)- rather than (*E*)-alkene products (*5*). In a series of experiments wherein both π-allyl isomers could be observed spectroscopically, they convincingly demonstrate that it is this difference in relative reaction rates in these two critical steps that is responsible for (*Z*)-retention.

An important factor in the observed retention of (*Z*)-stereochemistry is that the nucleophiles studied undergo attack at the carbon bearing the greater number of protons within the *anti*–π-allyl to afford the linear product. This finding is astounding in that these iridium catalysts typically lead to nucleophilic attack at the carbon with fewer protons within the *syn*–π-allyl complex and deliver the branched product with an alkene that lacks stereochemistry (*4*, *6*). The reason for this uncommon positional selectivity is unclear at this time (*7*). Two different classes of nucleophiles are shown to generate (*Z*)-alkene products with accompanying high enantiomer ratios: Tryptamines and tryptophols afford tricyclic products, and α–amino acid derivatives furnish valuable α,α-disubstituted amino acids.

The (*Z*)-retentive allylic substitution strategy disclosed by Jiang *et al.* does rely on pre-installation of the higher-energy (*Z*)-alkene, and so it does not constitute de novo synthesis of this functionality, as does (*Z*)-selective cross-metathesis (*8*, *9*). Still, unlike metal-catalyzed cross-coupling reactions, which are far less prone to alkene isomerization, the resistance toward alkene stereochemical isomerization discovered by Jiang *et al.* in iridium-catalyzed allylic substitution is remarkable. That two disparate classes of nucleophiles show the same (*Z*)-retentive reactivity suggests the exciting possibility that this phenomenon is widely applicable. Given the overall worth of allylic substitution in chemical synthesis, this finding is perhaps just the first step toward new pathways for chemists to access highly valuable (*Z*)-alkenes. ■

REFERENCES AND NOTES

1. R. Jiang, L. Ding, C. Zheng, S.-L. You, *Science* **371**, 380 (2021).
2. B. M. Trost, M. R. Machacek, A. Aponick, *Acc. Chem. Res.* **39**, 747 (2006).
3. Z. Lu, S. Ma, *Angew. Chem. Int. Ed.* **47**, 258 (2008).
4. Q. Cheng *et al.*, *Chem. Rev.* **119**, 1855 (2019).
5. S. T. Madrahimov, J. F. Hartwig, *J. Am. Chem. Soc.* **134**, 8136 (2012).
6. S. T. Madrahimov, Q. Li, A. Sharma, J. F. Hartwig, *J. Am. Chem. Soc.* **137**, 14968 (2015).
7. X. Zhang, L. Han, S.-L. You, *Chem. Sci.* **5**, 1059 (2014).
8. S. J. Meek, R. V. O'Brien, J. Llaveria, R. R. Schrock, A. H. Hoveyda, *Nature* **471**, 461 (2011).
9. M. J. Koh et al., *Nature* **517**, 181 (2015).

10.1126/science.abf8851

GRAPHIC: JOSHUA BIRD/*SCIENCE*

VIEWPOINT: COVID-19

Sex differences in immune responses

Biological sex differences in immunity potentially underlie male bias for severe COVID-19

By **Takehiro Takahashi**[1] *and* **Akiko Iwasaki**[1,2]

Evidence increasingly indicates that male sex is a risk factor for more severe disease and death from COVID-19. Male bias in COVID-19 mortality is observed in nearly all countries with available sex-disaggregated data, and the risk of death in males is ~1.7 times higher than in females (*1*). Aging is strongly associated with higher risk of death in both sexes, but at all ages above 30 years, males have a significantly higher mortality risk, rendering older males the most vulnerable group (*1*). Sex differences are intertwined with differences in gender roles socially and with behavioral factors, which also influence COVID-19 incidence and outcomes. However, there are also possible biological mechanisms of male sex bias that affect the severity of COVID-19, particularly with respect to immune responses.

Sex differences beyond sex organs are present across species and extend to physiological systems, including the immune system. Infection by different pathogens results in differential immune responses and disease outcomes by sex, and although the pattern depends on age and other host factors, male sex is more often associated with lower immune responses and higher susceptibility and/or vulnerability to infections in animals. This is generally also the case in humans (*2*): Male patients have higher viral loads for hepatitis B virus (HBV) and HIV (*2*). Conversely, females generally mount a more robust immune response to vaccines, such as influenza vaccines. However, the heightened immune responses in females can also lead to detrimental immunopathology in infections (*2*).

The physiological response to virus infection is initiated when virus replication is detected by pattern recognition receptors. This leads to two antiviral programs by the infected cells: (i) cellular antiviral defense programs mediated by type I and type III interferons (IFNs) to limit viral replication and spread, and (ii) cytokine and chemokine production to recruit and coordinate immune cells, such as monocytes and neutrophils that can phagocytose and eliminate infected cells. Notably, COVID-19 is characterized by strong innate immune cytokine and chemokine responses despite the disproportionately low antiviral defense signature mediated by IFNs (*3*). Patients with severe COVID-19 exhibit high serum concentrations of proinflammatory cytokines and chemokines, with particularly high concentrations of interleukin-6 (IL-6) and the inflammasome-associated cytokines IL-1β and IL-18 (*4*). Exacerbated systemic inflammation is associated with extensive pulmonary pathology, including massive infiltration of monocytes and neutrophils. Increased neutrophil count is associated with poor clinical outcome (*5*). The robust induction of these inflammatory cytokines and cells may be a compensatory response to the ability of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) to evade IFN responses, which necessitates the engagement of IFN-independent defense mechanisms (*6*).

The plasma concentrations of several key innate immune cytokines and chemokines, such as IL-8 and IL-18, are increased in male patients compared with female patients in the early phase of COVID-19. Conversely, female patients have higher plasma concentrations of type I IFN (IFNα) over the disease course (*7*). Notably, autoantibodies that inhibit type I IFN signaling have been reported in a subset of severely ill patients, the majority (94%) of whom were older males (*8*). By contrast, T cell activation at the early phase of SARS-CoV-2 infection is robust even in older female patients, whereas male patients have a significant decline with age. Male patients with poor T cell activation at the early phase of disease onset have worse COVID-19 outcomes, whereas no such difference is observed in female patients (*7*).

What could be the potential underlying mechanism of this sex dimorphism in immune response? One culprit is the sex chromosomes: A substantial number of important immune-related genes are encoded on the X chromosome. Although one of the two copies of the X chromosome in females is usually epigenetically silenced [X chromosome inactivation (XCI)], some critical immune-related genes, including Toll-like receptor 7 [*TLR7*, whose viral RNA sensing elicits strong type I IFN production in its major producer, plasmacytoid dendritic cells (pDCs)], can escape from XCI in some proportion of cells. This makes the population "mosaic" for biallelic expression, which leads to higher gross expression of some immune-related genes in females (*9*). Human pDCs have also been reported to have higher expression of interferon regulatory factor 5 (IRF5) in females. Higher expression of these genes leads to more robust type I IFN responses in females, and this is one of the potential mechanisms involved in females' enhanced protection against viral infections, including COVID-19 (*1*, *2*).

Sex has a major impact on immune cell transcriptomes; immune cells or even the immune system are differentially affected by aging, depending on sex. Aging induces a decline in the proportion of naïve T cells that is more prominent in males, and B cells decline after age 65 only in males (*10*). Males have abrupt and drastic changes in the epigenetic landscape of their immune cells between ages 62 and 64, and subsequently males exhibit an accelerated immunosenescence phenotype that is characterized by enhanced innate proinflammatory gene expression and lower gene expression related to adaptive immunity, which could potentially predispose older males to hyperinflammation and poor adaptive immune responses.

By contrast, major changes in the epigenetic landscape of immune cells occur in females 5 to 6 years later than in males, with this gap largely corresponding to the life span differences between sexes (*10*). It is notable that females generally mount more pronounced cytokine responses in viral infections, although this is not the case with COVID-19. Instead, males have higher plasma concentrations of innate proinflammatory cytokines such as IL-8 and IL-18 (*7*). This could be because patients who have severe disease are generally of older age, and the background transcriptomic and epigenetic sex differences in immune cells in these older individuals are possibly amplified and more explicitly manifested in the context of SARS-CoV-2 infection (see the figure).

The other important biological factors are sex hormones. In a mouse model of SARS-CoV infection, higher mortality in male mice was observed and was attributed to the protective roles of the female sex hormone estrogen (*1*). Studies using various cell types and animal models have shown that the expression of angiotensin-converting enzyme 2 (ACE2), the host cell entry receptor for SARS-CoV-2, is modulated by estrogen (*1*). This likely occurs through transcriptional regulation by estrogen receptor signaling, although the detailed molecular mechanisms

[1]Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA. [2]Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT, USA. Email: akiko.iwasaki@yale.edu

are yet to be identified and the consequences of the type of ACE2 modulation (up- or down-regulation) could be cell type– and context-dependent. Estrogen receptors are widely expressed by many cell types, including immune cells, and estrogen is a critical regulator of gene expression and functions in innate immune cells, including monocytes, macrophages, and dendritic cells, as well as in lymphocytes such as T helper 1/2 ($T_H1/2$) cells, regulatory T cells (T_{regs}), and B cells. One of the major forms of estrogen, estradiol, has been shown to dampen the production of excessive innate inflammatory cytokines by monocytes and macrophages (*1*). Menopause is an independent risk factor for female patients, and concentrations of estradiol and anti-Müllerian hormone, which is produced by cells in developing follicles and is an endocrine marker for oocytes remaining in the ovaries (ovarian reserve), are inversely correlated with COVID-19 severity (*11*). Conversely, androgen-deprivation therapies for prostate cancer have been reported to reduce the risk for SARS-CoV-2 infection (*12*). Like estrogen receptors, androgen receptors are expressed widely among immune cells, regulating the transcription of various genes, and the outcome of their signaling has been shown to be generally immunosuppressive (*1*).

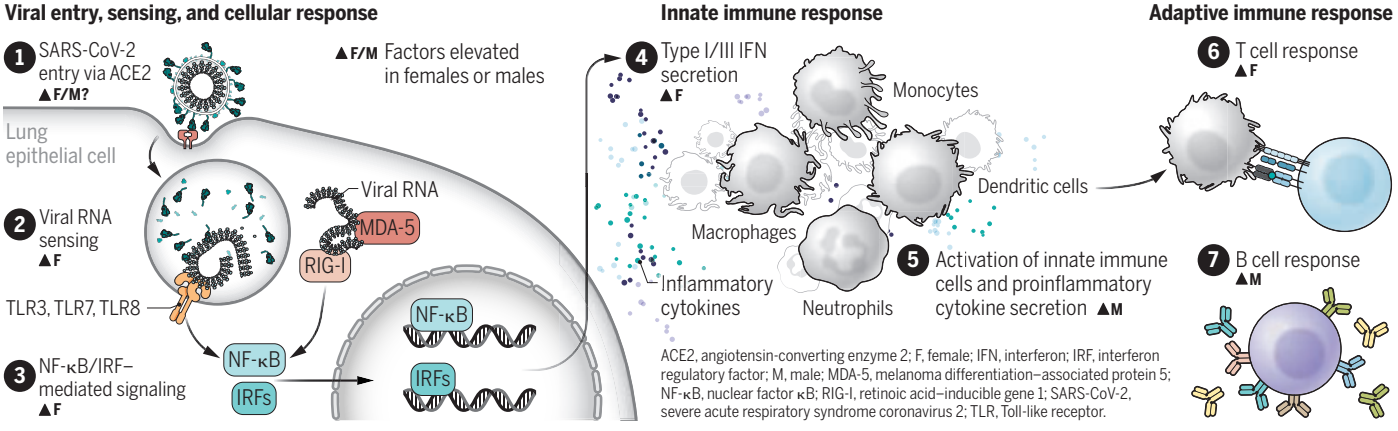
Despite the emerging understanding of sex differences of immune responses in COVID-19, many questions remain. Sex is not binary, and little is known regarding immune responses to viral infections, including COVID-19, in individuals with disorders of sex development (DSD) or transgender individuals. DSD describes congenital conditions in which the development of chromosomal, gonadal, and anatomical sex is atypical. For example, Klinefelter syndrome (also known as 47, XXY) results in phenotypes such as gynecomastia (enlarged breast tissue in males) and small testes with hypogonadism. This condition accompanies a series of comorbidities, in particular a high frequency of autoimmune diseases, such as systemic lupus erythematosus (SLE). The prevalence of SLE among Klinefelter syndrome patients is higher by a factor of 14 than among 46, XY males; this is comparable to its prevalence among 46, XX females, which suggests a gene dosage effect of the X chromosome (*13*). Additionally, transgender is a collective term encompassing those whose gender identities or gender roles differ from those typically associated with the sex they were assigned at birth. Little is known about the immune responses in these individuals, some of whom are undergoing gender reassignment hormone therapies. It is possible that DSD and transgender individuals mount distinct immune responses to viral infection in general and also in COVID-19.

Sex differences in immunity also have implications for responses to SARS-CoV-2 vaccination or reinfection. Analysis of convalescent plasma showed that male sex, older age, and hospitalization for COVID-19 were associated with greater SARS-CoV-2 antibody titers (*14*). This could be related to increased disease severity in this demographic of patients, with increased viral load driving more B cell activation and antibody production. Alternatively, higher amounts of antiviral antibodies could be due to compensatory increases in antibody production due to suboptimal qualities of antibodies generated in older males, which are unable to neutralize the virus effectively. Antibodies with suboptimal neutralizing ability have the capacity to promote viral invasion into host cells such as macrophages [called antibody-dependent enhancement (ADE)]. However, there has been no clear evidence of ADE in COVID-19. With other viral infections, fatal diseases in dengue virus infection (dengue hemorrhagic fever and dengue shock syndrome) have been associated with ADE. Upon secondary infection with a virus of a serotype different from that of the primary infection, the fatalities from these severe diseases were primarily seen in adult females despite the male dominance in dengue hemorrhagic fever prevalence (*15*). The mechanism remains unknown but may be due to heightened inflammation in females, leading to capillary permeability.

It is important that studies of COVID-19 patients report results in a sex-disaggregated manner, not only to elucidate differential disease pathogenesis, but also to enable a deeper understanding of this disease and the eventual development of better treatment and preventive strategies. It should be standard practice to collect and report sex-disaggregated data for this and for all infectious disease and vaccine studies in the future. ■

Sex differences in factors that affect infection and immunity in COVID-19

SARS-CoV-2 binds to ACE2 to initiate host cell entry. This activates the viral RNA sensors TLR3/7/8 and RIG-I–MDA-5, which induce secretion of IFNs and other inflammatory cytokines, leading to innate and adaptive immune responses. In each of these steps, sex differences may shape the antiviral immune response.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. E. P. Scully *et al.*, *Nat. Rev. Immunol.* **20**, 442 (2020).
2. S. L. Klein, K. L. Flanagan, *Nat. Rev. Immunol.* **16**, 626 (2016).
3. D. Blanco-Melo *et al.*, *Cell* **181**, 1036 (2020).
4. C. Lucas *et al.*, *Nature* **584**, 463 (2020).
5. J. Wang *et al.*, *Front. Immunol.* **11**, 2063 (2020).
6. S. Nish, R. Medzhitov, *Immunity* **34**, 629 (2011).
7. T. Takahashi *et al.*, *Nature* **588**, 315 (2020).
8. P. Bastard *et al.*, *Science* **370**, eabd4585 (2020).
9. M. Souyris *et al.*, *Sci. Immunol.* **3**, eaap8855 (2018).
10. E. J. Márquez *et al.*, *Nat. Commun.* **11**, 751 (2020).
11. T. Ding *et al.*, *Clin. Infect. Dis.* 10.1093/cid/ciaa1022 (2020).
12. M. Montopoli *et al.*, *Ann. Oncol.* **31**, 1040 (2020).
13. R. H. Scofield *et al.*, *Arthritis Rheum.* **58**, 2511 (2008).
14. S. L. Klein *et al.*, *J. Clin. Invest.* **130**, 6141 (2020).
15. S. S. Sam *et al.*, *PLOS Negl. Trop. Dis.* **7**, e2194 (2013).

10.1126/science.abe7199

RETROSPECTIVE

Masatoshi Koshiba (1926–2020)

Innovative founder of neutrino astronomy

By **Masayuki Nakahata**[1] *and* **Atsuto Suzuki**[2]

Masatoshi Koshiba, eminent experimental particle physicist, passed away on 12 November 2020. He was 94. By conducting electron–positron (e^-–e^+) collider experiments, Koshiba used his creativity to advance the field of particle physics. He also adapted equipment to make ground-shifting discoveries, leading to the new fields of underground neutrino physics and neutrino astronomy.

Koshiba was born in Toyohashi, Japan, on 19 September 1926. He earned his B.S. in physics from the University of Tokyo in 1951 and completed his master's thesis detecting cosmic rays with a nuclear emulsions technique. In 1953, he enrolled in the Graduate School of Physics at Rochester University in New York. After obtaining his Ph.D. in cosmic ray physics within 2 years, he was appointed as a research fellow at the University of Chicago in 1955. In 1958, he returned to the University of Tokyo, where he remained until his retirement in 1987.

After contributing to international cosmic ray collaborations, Koshiba began his exploration of accelerator particle physics in 1969, when he joined the construction of the e^-–e^+ collider at Budker Institute of Nuclear Physics in Russia. Although the project was terminated in 1972, Koshiba pushed on with e^-–e^+ collider experiments and later joined the Double Arm Spectrometer (DASP) experiment at Deutsches Elektronen-Synchrotron (DESY) in Germany. This project evolved into the JADE (Japan, Deutschland, and England) experiment at DESY. Through these experiments, Koshiba established the Standard Model of particle physics and discovered the gluon, which mediates strong interactions between quarks. In 1985, in recognition of these efforts, he received the Great Cross of Merit, the Order of Merit of the Federal Republic of Germany. Koshiba dreamed of constructing the next-generation e^-–e^+ linear collider in Japan, a vision that will be realized with the completion of the e^-–e^+ International Linear Collider (ILC).

In 1980, detecting evidence of a so-called Grand Unified Theory became a major challenge for experimentalists. This theory predicted that protons would decay, and Koshiba proposed a way to detect the process. We both joined Koshiba as he prepared to execute his experiment, known as the Kamioka Nucleon Decay Experiment (Kamiokande). He had constructed a 3000-ton water Cherenkov detector located 1000 m underground in the Kamioka mine. The distinguishing feature of this experiment was the use of 1000 photomultiplier tubes with a diameter of 51 cm, the largest in the world. Kamiokande started to collect data in 1983. Unfortunately, the lifetime of a proton was found to be much longer than the sensitivity of Kamiokande, and we could not find any positive signals. However, Koshiba—with characteristic ingenuity—noticed that the Kamiokande detector was sensitive to astrophysical neutrinos, including solar neutrinos (generated inside the Sun). He proposed an update to the detector to observe neutrinos instead.



PHOTO: YAMAGUCHI HARUYOSHI/GETTY IMAGES

The results of the first solar neutrino observation experiment, led by physicist Raymond Davis Jr. in 1968, had shown a deficit of neutrinos relative to the solar model prediction. Within 2 months of the start of Kamiokande's solar neutrino data collection in 1987, billions of extragalactic messengers—the supernova (SN) 1987A neutrinos—swept through Earth. SN 1987A occurred in the Large Magellanic Cloud, 170,000 light-years away from the Solar System. On 23 February 1987, 11 neutrino events were observed in Kamiokande within 13 seconds. The energies and number of observed events in the Kamiokande data showed that the origin of a supernova was gravitational collapse and that 99% of the released gravitational energy is carried out by neutrinos, which confirmed the basic idea of the star core collapse. Davis and Koshiba shared the 2002 Nobel Prize in Physics "for pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos," and the research field of neutrino astronomy was born.

Kamiokande succeeded in observing solar neutrinos, and the data—collected through 1990—clearly confirmed the solar neutrino deficit, the explanation for which remained a puzzle. In addition to the solar neutrino deficit, Kamiokande first observed the atmospheric neutrino anomaly that shows a deficit of muon neutrino to electron neutrino ratio of atmospheric neutrinos. Both the solar neutrino deficit and the atmospheric neutrino anomaly were predicted to be solved by neutrino oscillations induced by finite neutrino masses that had been predicted by the Grand Unified Theory. These anomalies were finally resolved in 1998 by the detection of evidence of neutrino oscillations in Super-Kamiokande, a 50,000-ton water Cherenkov detector originally proposed by Koshiba. For this work, physicist Takaaki Kajita, one of Koshiba's Ph.D. students, was awarded the 2015 Nobel Prize in Physics.

Koshiba nurtured many talented physicists and educated them using his signature didactic messages. For example, Koshiba suggested that a researcher should always hold three or four new "eggs" (research ideas) at a time but should regularly ask whether each egg can still become a bird. If not, then the idea is no longer worth holding. He also recommended that experimental equipment incorporate as many distinctive devices as possible. That way, even if you fail to catch the prey you are aiming for, you will have a chance to catch other prey.

Koshiba's incredible enthusiasm and outstanding ability to get things done were embodied by Kamiokande, which led to both neutrino astronomy and the discovery of finite neutrino masses. Koshiba always worked on his research with keen intuition, passion, foresight, and leadership. His innovative ideas inspired younger researchers, who built on his work to realize his dreams. Those young scientists will no doubt continue to advance the field of physics on the basis of the foundations that Koshiba put in place. ■

10.1126/science.abg1561

[1]Kamioka Observatory, University of Tokyo, Tokyo, Japan. [2]Iwate Prefectural University, Iwate, Japan. Email: nakahata@suketto.icrr.u-tokyo.ac.jp; atsuto.suzuki@kek.jp

POLICY FORUM

STEM WORKFORCE

Rethinking immigration policies for STEM doctorates

Many Ph.D.'s follow an inefficient path to green cards through visas aimed at entry-level workers

By **Michael Roach**[1] *and* **John Skrentny**[2]

Despite broad consensus that high-skilled immigration can contribute to innovation and economic growth (*1*–*3*), there is considerable controversy regarding how to reform immigration policies, including for workers in science, technology, engineering, and mathematics (STEM) professions. Much attention has centered on entry-level IT workers (*4*, *5*), but less consideration has been given to the visa pathways of STEM doctorates (*6*, *7*) who, unlike entry-level IT workers, can bring firms advanced training at the frontiers of science and technology and contribute disproportionately to innovation and entrepreneurship relative to workers with bachelor's or master's degrees (*3*). We provide new insights on the visa progression, qualifications, and starting salaries of STEM doctorates in the U.S. context from a survey that follows a cohort of 1597 U.S. citizen (69.7%) and foreign-born (30.3%) science and engineering Ph.D.'s from U.S. research universities into their first-time industry R&D employment [see details in the supplementary materials (SM)]. We show that the H-1B visa has become the predominant first step for STEM Ph.D.'s employed in industrial R&D, not because it is legally required or the most suitable visa but because of inefficiencies and delays on the path to permanent residency. Our findings show that the H-1B—a highly contentious visa used primarily for entry-level workers—may be an inefficient pathway for U.S.-trained STEM doctorates and suggest the need to rethink visa policies to retain these highly specialized workers.

Lawmakers of both major U.S. political parties have supported policies that would facilitate retention of U.S.-trained STEM Ph.D.'s. In recent years, however, attention has centered on the H-1B, which is problematic for U.S.-trained Ph.D.'s for two reasons. First, there are widespread concerns that firms use the H-1B to import cheap labor that displaces U.S. workers, especially in entry-level IT jobs (*5*). As a result, some reformers seek a reduction in the number of H-1Bs issued, and recently the Trump Administration tried to suspend issuance of new H-1B visas. However, there is little evidence on whether these concerns also apply to Ph.D.'s. To the extent that the H-1B is used to hire STEM Ph.D.'s, such sweeping reforms could drastically restrict U.S. companies' access to talented Ph.D.'s (*8*).

Second, STEM Ph.D.'s compete with entry-level workers for a limited number of H-1B visas allocated through a lottery system regardless of employer demand, level of occupation, or pay. This has heightened uncertainty for both workers and employers, leading U.S. firms—including Amazon, Google, and Microsoft—to open R&D centers in countries such as Canada where visas favor high-skilled STEM talent (*9*). Visa uncertainties also deter foreign Ph.D.'s from working in startups (*10*), placing young technology companies at a disadvantage in their ability to hire recent doctorates. Congress did not design the H-1B to retain elite workers with advanced degrees from U.S. universities, and visa reforms tailored to foreign-born Ph.D.'s may be warranted.

VISA PROGRESSION

Upon graduation, foreign Ph.D.'s have different visa pathways to industry employment. First, STEM Ph.D.'s may work on their F-1 student visa for up to 3 years without an employment-based visa through the Optional Practical Training (OPT) program. However, to remain in the U.S. long-term, foreign Ph.D.'s need to transition to an employment-based visa. There are two visa paths—one authorizing permanent residency and the other authorizing temporary work.

The Immigration Act of 1990 created a complex series of merit-based permanent resident ("green card") visa categories to attract and retain highly skilled workers. These include the highest category (EB-1) for outstanding researchers and workers with "extraordinary ability." The second category (EB-2) is for workers with advanced degrees or "exceptional ability." The EB-2 also includes a special "National Interest Waiver" (NIW) category to expedite permanent residence for workers with skills deemed especially valuable to the country. STEM doctorates from U.S. universities are qualified upon graduation for EB-2 for advanced degrees, and they may also be qualified for the EB-1 under extraordinary ability, the EB-2 under exceptional ability, or NIW if they can meet specific criteria such as evidence of scientific publications, patents, awards, and commanding a high salary. Although workers must be sponsored by their employers for the EB-2 advanced degree or exceptional ability, they may self-sponsor through either the EB-1 or NIW.

Doctorates are eligible to apply for permanent residency while on OPT and are not required to first transition to a temporary work visa such as the H-1B. The time to secure permanent residency varies depending upon the EB visa type and the worker's nationality, with processing times as short as 6 to 12 months for many Ph.D.'s on EB-1 to as long as 5 to 10 years for Ph.D.'s on EB-2 visas from countries where the high number of applicants has resulted in a wait list, such as India and China.

Another pathway for foreign Ph.D.'s is through the employer-sponsored H-1B, a "non-immigrant" temporary visa that Congress designed to fill short-term labor shortages for workers in specialty occupations with at least a bachelor's degree. Each year, the U.S. Citizenship and Immigration Services (USCIS) allocates 65,000 H-1B visas, plus an additional 20,000 for workers with a master's degree or higher from U.S. universities, for employees of for-profit firms (nonprofits such as universities are exempt from these caps). Unlike EB visas, the H-1B gives little priority to worker qualifications and no special recognition to Ph.D.'s. Given that each year the number of H-1B applications far exceeds the number of visas available, the USCIS randomly selects workers until the quotas are met. Doctorates not selected in a given year may reapply the following year, so long as they have not exhausted their OPT eligibility. The H-1B is valid for 3 years and may be renewed for an additional 3 years. Although the H-1B is not required for permanent residence eligibility, in practice Ph.D.'s may use the H-1B as a bridge between OPT and a green card.

The array of different pathways that doctorates may take through the visa system is complex, and there is little empirical un-

[1]Charles H. Dyson School of Applied Economics and Management, SC Johnson College of Business, Cornell University, Ithaca, NY, USA. [2]Department of Sociology, University of California-San Diego, La Jolla, CA, USA. Email: michael.roach@cornell.edu

derstanding of which visas recent doctorates use in their first industry job, nor the sequencing and timing of visa progression. We provide detailed insights on the visa progression of foreign STEM doctorates from 39 leading U.S. research universities who graduated between 2010 and 2016 and transitioned to industry employment (see SM for details of sample by nationality and degree field). The survey provides previously unavailable microdata on STEM doctorates' visa progression, employment qualifications, starting salary and work benefits, and employer characteristics. We complement our survey with administrative data from the Department of Labor (DOL) on 2461 Labor Certification applications for foreign Ph.D.'s sponsored for an EB-2 visa. Our sample represents young, early-career Ph.D.'s who have many years of productivity ahead of them and are the focus of governments around the world seeking to attract and retain leading STEM researchers.

We asked foreign Ph.D.'s to report their first work visa, as well as their current visa after at least 3 years of industry employment. In their first job, two-thirds of Ph.D.'s were sponsored for an H-1B, the same visa used to hire entry-level workers with bachelor's degrees and that is at risk of restriction by Congress (see the first figure). Approximately 10% of foreign doctorates were either employer- or self-sponsored for permanent residency, and 16% remained on OPT without a sponsored work visa.

After at least 3 years of employment, 68% of doctorates have either received or been sponsored for permanent residency (see the first figure). Among Ph.D.'s who were first sponsored on an H-1B, 76% have transitioned to or are being sponsored for permanent residency (table S2). Although our sample does not reflect Ph.D.'s who may have left the United States as a result of visa delays, these patterns illustrate that the majority of Ph.D.'s obtain permanent residency early in their careers and that most do so by passing through the H-1B.

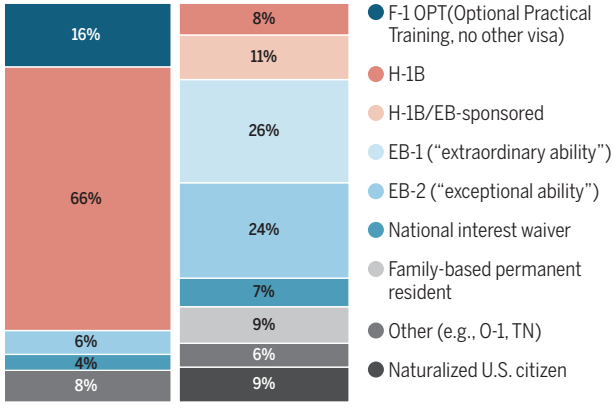
To obtain deeper insights into pathways for Ph.D.'s to gain permanent residency, we use DOL administrative data on EB-2 Labor Certification applications. Among foreign STEM Ph.D.'s employed in industrial R&D and sponsored by their employer for EB-2 visas, 20% progressed directly from OPT, whereas nearly 80% progressed through an H-1B (table S4). The average time from graduation to filing an EB-2 application is 1.1 years under OPT and 2.8 years under H-1B (fig. S2). Although Ph.D.'s on H-1B wait ~18 months longer to be sponsored for permanent residency than Ph.D.'s on OPT, this wait time is much shorter than the full 6-year term of the H-1B visa. This suggests that employers are not using the H-1B as a trial period before sponsoring for a green card.

At the same time, the short duration of OPT work eligibility makes applying for permanent residency without first obtaining an H-1B an often impossible path for doctorates from India and China, who may wait several years to receive their green cards because of per-country quotas. Our survey shows that Indian (78%) and Chinese (67%) doctorates are especially likely to be first sponsored for an H-1B compared to doctorates from other countries (55%) (table S4). Although the longest wait times of 5 to 10 years are for the EB-2, wait times for the EB-1 are just over 2 years (*11*). To examine whether Ph.D.'s differ in their use of EB-1 and EB-2 visas, we asked survey respondents sponsored for permanent residency their EB preference category: 46% reported that they were sponsored for an EB-2, 41% were sponsored for EB-1, and 10% for NIW (table S3). Moreover, a disproportionate share of doctorates from India and China are on EB-1 (50.8% and 54.2%, respectively) compared to doctorates from the rest of the world (20.0%, table S4), likely because of the shorter wait time relative to the EB-2.

Together these findings demonstrate that a meaningful share of doctorates transition directly to permanent residency while on OPT. At the same time, the majority first pass through the H-1B, suggesting that the H-1B may be pressed into service to overcome gaps resulting from an inefficient immigration system.

Visa progression of foreign-born STEM Ph.D.'s

Employment visa in first job is shown at left, current visa after 3 to 9 years of employment is at right. Sample is foreign-born Ph.D.'s who were temporary residents at graduation and required a work visa in their first industrial job (N = 305).



This raises the question of why more doctorates don't transition directly to permanent residency. One possibility is that doctorates sponsored for permanent residence in their first job may be more qualified than those sponsored for H-1B (see SM). In regression analyses that control for demographics and degree field (table S6), we find no significant difference between Ph.D.'s sponsored for EB or H-1B visas in terms of the number of publications or patents prior to industry employment, two evidentiary requirements of a worker's qualifications used in permanent residency applications. We also find no significant difference in starting salary as reported in our survey, nor in the offered wage as reported in the DOL EB-2 applications (see the second figure). Although we lack data to explore why firms sponsor some Ph.D.'s for EB and others for H-1B, these results show that Ph.D.'s sponsored for an H-1B may be qualified for permanent residency when first hired.

COMPARISON TO U.S. PH.D.'S

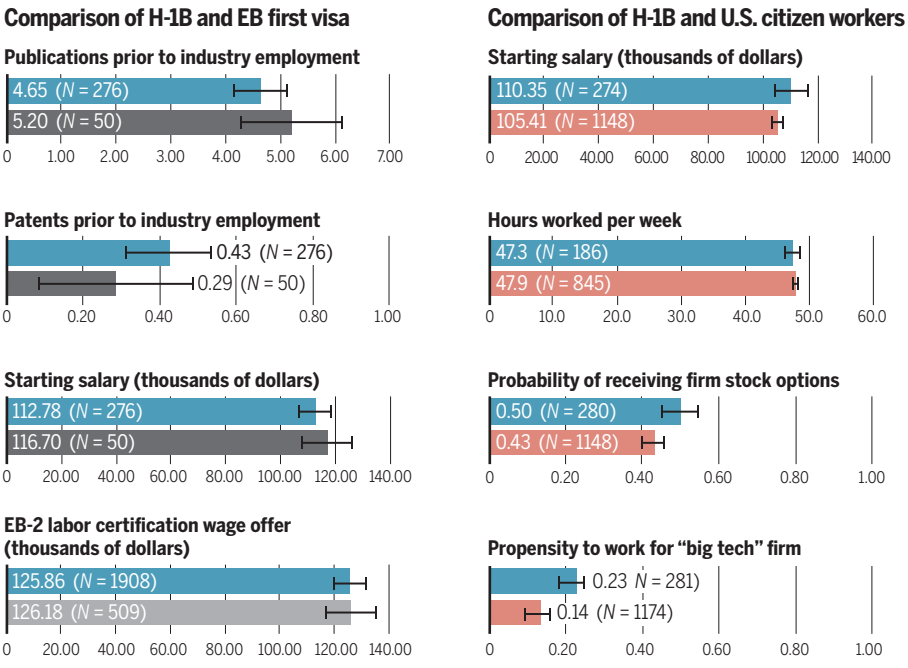
Arguably the greatest concern when considering changes in visa policies is potential adverse impacts on U.S. citizens. Though the preponderance of prior research has compared U.S. citizens to H-1B entry-level IT workers (*5*, *12*), there is little existing evidence comparing U.S. and foreign doctorates first-time employees in similar industry occupations. Data from the U.S. Bureau of Labor Statistics show that doctorates have the lowest unemployment and highest salaries of any level of educational attainment, with pre–COVID-19 unemployment of 1.1% and median annual salaries of approximately $100,000 (*13*). Although these numbers do not distinguish between U.S. and foreign doctorates, the tight labor market for highly skilled Ph.D.'s does not suggest crowding out.

To examine possible differences in pay, we use our survey to compare the starting salary of U.S. citizen and foreign Ph.D.'s on temporary resident visas in their first industrial R&D job (see SM). In regression analyses that control for worker ability, demographics, work benefits, job start year, and employer type, we find no significant difference in pay between U.S. citizen and foreign doctorates (see the second figure). Instead, starting salary is driven primarily by degree field and proxies for worker ability (table S7). Though these comparisons do not allow for careful identification of the

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

Comparing qualifications

Graphs show predicted values from regression analyses (see tables S6, S7, and S9). (Left) Foreign doctorates first sponsored for H-1B (blue) are compared with those first sponsored for permanent residency (EB)(dark gray); EB-2 Labor Certification application wage offer is compared between applicants on H-1B and on OPT (Optional Practical Training, light gray). (Right) Foreign doctorates on H-1B (blue) are compared with U.S. citizen doctorates (red).



causal effects of immigration policies on wages, nor do they rule out the possibility that a greater number of foreign Ph.D.'s in the workforce could drive down wages for native Ph.D.'s (*14*), they provide suggestive evidence that often-cited concerns of foreign entry-level STEM workers being paid less than their native peers do not apply to STEM Ph.D.'s (*5*).

Another frequent concern is that employers may use visa sponsorship as leverage to exploit foreign workers in ways other than salary. For example, employers may require sponsored employees to work longer hours relative to U.S. citizens, or they may offer fewer benefits to offset the high costs of visa sponsorship. In regression analyses (table S9), we find no significant difference between U.S. and H-1B Ph.D.'s in hours worked, with both reporting an average of 47 hours per week (see the second figure). We also find that H-1B Ph.D.'s have a higher probability than U.S. Ph.D.'s of receiving firm stock options (see the second figure), a highly coveted financial benefit. This appears to be due to foreign doctorates' higher propensity to work in "big tech" firms like Google and Amazon that are more likely to give new employees stock options (see SM). H-1B Ph.D.'s are also nearly twice as likely to work for a big tech firm compared to other firms (see the second figure).

IMPLICATIONS

The role of the H-1B as the modal start of a complex path to permanent residency for elite, high-skilled STEM doctorates from U.S. universities is inconsistent with the U.S. government's classification of the H-1B as a guest worker visa for entry-level jobs. However, delays in the processing and wait times for a green card have made the H-1B the de facto first visa for the majority of doctorates, especially individuals from India and China. This inefficient visa path through the H-1B increases not only uncertainty for workers but also the costs for many employers who spend thousands of dollars per employee to first sponsor them for an H-1B and then several thousand more for permanent residency. These inefficiencies suggest that a streamlining of the path from doctorate to permanent residency or changes to the H-1B program may be warranted (*15*).

Visa reforms that specifically target STEM Ph.D.'s from U.S. universities—with oversight to avoid fraudulent dissertations and job offers—could have broad benefits by facilitating and easing doctorates' transitions to permanent residence. A simplified path toward permanent residency could be especially beneficial in leveling the playing field for early-stage technology startups that are at a disadvantage in hiring workers who require visa sponsorship (*10*). In addition, exempting STEM Ph.D.'s from national quotas for green cards would not only facilitate their pathway to permanent residency but could also help U.S. firms to retain top scientists and engineers where they can contribute to innovation and economic growth. Our findings align with the proposed immigration policies of President-elect Biden, who within his first 100 days intends to implement a program that would provide recent STEM doctorates from U.S. universities with a green card that is exempt from national quotas. ■

REFERENCES AND NOTES

1. W. R. Kerr, W. F. Lincoln, *J. Labor Econ.* **28**, 473 (2010).
2. J. Hunt, M. Gauthier-Loiselle, *Am. Econ. J. Macroecon.* **2**, 31 (2010).
3. J. Hunt, *J. Labor Econ.* **29**, 417 (2011).
4. R. Hira, in *U.S. Engineering in a Global Economy*, R. B. Freeman, H. Salzman, Eds. (Univ. of Chicago Press, 2018), pp. 263–283.
5. D. Costa, R. Hira, "H-1B visas and prevailing wage levels" (Economic Policy Institute, Washington, D.C., 2020).
6. S. Kahn, M. MacGarvie, *Res. Policy* **49**, 103879 (2020).
7. M. G. Finn, L. A. Pennington, "Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2013" (Oak Ridge Institute for Science and Education, 2018).
8. National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, *The Arc of the Academic Research Career* (National Academies Press, 2014).
9. B. Glennon, "How Do Restrictions on High-Skilled Immigration Affect Offshoring? Evidence from the H-1B Program," University of Pennsylvania Working Paper (2019); https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2019/07/FP0450_WP_2020Mar.pdf.
10. M. Roach, J. Skrentny, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16805 (2019).
11. https://travel.state.gov/content/travel/en/legal/visa-law0/visa-bulletin.html
12. J. Bound, G. Khanna, N. Morales, in *High-Skilled Migration to the United States and Its Economic Consequences*, W. R. K. G. H. Hanson, S. Turner, Eds. (Univ. of Chicago Press, 2018), pp. 109–175.
13. U.S. Bureau of Labor Statistics, www.bls.gov/emp/chart-unemployment-earnings-education.htm.
14. G. J. Borjas, in *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, R. B. Freeman, D. L. Goroff, Eds. (Univ. of Chicago Press, 2009).
15. C. Gafner, S. Yale-Loehr, *Fordham Urban Law J.* **38**, 183 (2019).
16. M. Roach, "Replication Data for: Rethinking immigration policies for STEM doctorates," Harvard Dataverse, V1 (2020); https://doi.org/10.7910/DVN/8WCCSB.

ACKNOWLEDGMENTS

Special appreciation to S. Yale-Loehr for sharing his invaluable insights and expertise on immigration law and practice. We also thank B. Rissing, R. Hira, W. Kandel, H. Salzman, H. Sauermann, K. Wenger, and the editor and two referees for valuable comments and suggestions. We also thank the doctorates who have participated in the Science & Engineering PhD Panel Survey over the past decade who made this research possible. M.R. recognizes support from the Ewing Marion Kauffman Foundation Junior Faculty Fellowship. J.S. recognizes support from the Alfred P. Sloan Foundation (Award B2012-51) and the National Science Foundation (NCSE Award 1322945). Survey data used in the manuscript and supplementary materials will be made available on Harvard Dataverse (*16*). U.S. Department of Labor data are publicly available at www.foreignlaborcert.doleta.gov/.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/350/suppl/DC1

10.1126/science.abe7151

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*



BOOKS *et al.*

Darwin's hypotheses were made before the discovery of hominin species such as *Australopithecus afarensis*.

HISTORY OF SCIENCE

Darwin and human evolution

An anniversary appraisal of *The Descent of Man* probes the naturalist's prescience and prejudices

By **Erika Lorraine Milam**

Charles Darwin has for many years been cast as the hero of the life sciences. As a young man, he embarked on the adventure of a lifetime, joining the HMS *Beagle* on its voyage around the world. He returned older and wiser (and chronically ill), slowly transforming into a recognizably bearded sage. He published his most celebrated book, *On the Origin of Species*, in 1859 at the age of 50 (*1*).

If you read just that one book, it can be easy to interpret Darwin as being primarily interested in the evolution of animals and plants and to find him remarkable for all that he got right in elucidating the process of evolution by natural selection. Although he devoted considerable space in this work to discussing humans as agents of morphological change through domestication and "artificial selection," Darwin included only a single sentence on humanity being shaped by evolution: "Light will be thrown on the origin of man and his history" (*1*).

More than a decade later, he developed this promissory note in both *The Descent of Man, and Selection in Relation to Sex* and *The Expression of the Emotions in Man and Animals*, published in quick succession (*2*, *3*). To mark the 150th anniversary of *The Descent of Man*, paleoanthropologist Jeremy DeSilva has gathered a team of experts, mostly scientists, to pen reflections and update Darwin's analysis: one essay each on *The Descent of Man*'s introduction and its seven chapters about human evolution, one summarizing the 11 chapters Darwin devoted to sexual selection, one on sexual selection in humans, and a conclusion. The essays in *A Most Interesting Problem* collectively present an image of Darwin as both "remarkably prophetic" with regard to some predictions and "flat out wrong" with others.



A Most Interesting Problem
Jeremy M. DeSilva, editor
Princeton University Press, 2021. 288 pp.

As several authors point out, scientists now have far better evidence to work with than Darwin did, including thousands of fossilized hominin specimens spanning the past 7 million years, as well as robust genetic analyses of contemporary and historic populations. Darwin made the case for human evolution without any of this information, relying instead on comparative data from skeletons, embryos, brains, and behavior.

He concluded that the major factors defining humanity—large brains, complex moral codes, culture, and an appreciation for beauty—exist in other primates too, although to a lesser extent. The differences between humans and other primates were thus a matter of degree, not of kind. This reasoned assertion of Darwin's proved correct. So, too, did his intuition that human evolution had taken place largely in Africa.

Unsurprisingly, the most controversial of Darwin's claims, in his time and ours, regard race and sex. It is in the essays on racialized differences and sexual selection in humans—written by Agustín Fuentes and Holly Dunsworth, respectively—that his ideas come under direct fire. Fuentes calls out Darwin's "ethnocentric, Eurocentric, and anti-African biases." Dunsworth suggests that for Darwin, "women were wives, but men were so much more than husbands."

Still, Fuentes countenances Darwin such a "good scientist" that he speculates the Victorian would agree that the division of humans into biological races is scientifically unsound, if only he were presented with our current data on biological diversity. Dunsworth, by contrast, argues that scientists owe it to the human species to finally break the link between "Darwin's inchoate offerings" on gender and "their perceived social implications."

Had Darwin's chapters in *The Descent of Man* on sexual selection in animals received individualized attention, the volume would have looked rather different. For Darwin, questions of human evolution were partly constituted by his understanding of variations between "the races" and "the sexes"—sexual selection provided an evolutionary mechanism explaining the origins of human diversity (*2*, *4*). Greater scrutiny of sexual selection would thus have highlighted an additional dimension of how much has changed in the past century and a half, that scientists today treat race and sex as distinct subjects.

DeSilva's volume provides a welcome opportunity to reflect on the history of evolutionary theory as a legacy complicated by Darwin's prescience as well as prejudice. Indeed, *The Descent of Man*, writes noted historian Janet Browne in the volume's introduction, "shows Darwin at his most Victorian." The essays develop this insight as both a compliment and a criticism, revealing Darwin as far more human than the heroic role he is often assigned. ■

The reviewer is at the Department of History, Princeton University, Princeton, NJ 08544, USA. Email: emilam@princeton.edu

REFERENCES AND NOTES

1. C. Darwin, *On the Origin of Species* (John Murray, 1859).
2. C. Darwin, *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
3. C. Darwin, *The Expression of the Emotions in Man and Animals* (John Murray, 1872).
4. E. Richards, *Darwin and the Making of Sexual Selection* (Univ. of Chicago Press, 2017).

10.1126/science.abf6947

PHOTO: DAVE EINSEL/STRINGER/GETTY IMAGES

HISTORY OF MEDICINE

A barrier breaker's complicated legacy

Certain of her own exceptional nature, Elizabeth Blackwell dismissed those who aided her

By **Laura Stark**

In 1849 Elizabeth Blackwell became the first woman to receive a medical degree from an American college. She was a resourceful, dedicated, and imaginative thinker. She was also a greedy, racist, and rivalrous snob. In *The Doctors Blackwell*, Janice Nimura shifts between Blackwell's own account of her singular greatness and the story of her relationship with her medical colleague, occasional roommate, and little sister, Emily. Nimura's smart and skillful collective biography layers an account of an exceptional individual onto a narrative of the interdependence and political structures that made the myth of Elizabeth Blackwell possible.

The Blackwell sisters were immigrants from a large, loyal family that financially deteriorated as they grew up. The family business was sugar, and it profited directly from slavery, imbuing the sisters with the values of racist capitalism: competition, individualism, and hierarchical organization. Before medical school, Elizabeth apprenticed in the U.S. South, where she witnessed the cruelty of slavery firsthand. Yet she clung to the notion of racial hierarchy convenient to white political power. She spared respect for only her blood relations and clipped any threads of friendship and ties of political solidarity when such alliances were no longer of use. It takes a brazen tongue to bad-mouth Florence Nightingale, and Elizabeth Blackwell had one.

Elizabeth benefited from the early women's movement, although she refused to support its political efforts. She jeered women's rights advocates, labor organizers, and abolitionists who might have been natural allies in her projects. She spurned political organizing aimed at opening opportunities to more people and promoted access to medical care and education for women to build her own renown.

The reviewer is at the Department of History and the Department of Medicine, Health, and Society, Vanderbilt University, Nashville, TN 37235, USA.
Email: laura.stark@vanderbilt.edu

Throughout her life, Elizabeth had an unintuitive, self-serving view of women's rights. She neither wanted to be seen as a man (no bloomers for her) nor did she think that women and men were equal. It was important to her to be seen as a woman—as a member of a group believed to be generally inferior—in order to demonstrate her own exceptionalism.

White women in science are allowed to be ambitious and successful, so long as they are tenderhearted or at least tactful. Elizabeth Blackwell was neither. But the energy it takes to judge Elizabeth harshly is better spent, Nimura tacitly suggests, exploring the gendered standards of judgment behind ideas of good doctoring.



Emily (left) did much of the work while Elizabeth (right) took much of the credit.

Elizabeth showed a vexing combination of dependence and public disavowal in her relationship with her sister Emily. A few years after Emily graduated from Cleveland Medical College in 1854, the sisters started a clinic for poor women and children in New York City. The eventual success of the clinic was largely Emily's doing. Elizabeth had only one eye—the other was removed after it was infected with gonorrhea at the hospital in Paris where she apprenticed after medical school—so Emily did most of the doctoring in the clinic for which the elder Blackwell is remembered.

Elizabeth also dreamed of starting a medical school with her sister for the exceptional few women whom she saw fit to train in her image. Shortly after the school opened, however, Elizabeth decamped for England, their mother country, where she remained for the rest of her long life, while Emily remained and ran the school.

The Doctors Blackwell: How Two Pioneering Sisters Brought Medicine to Women—and Women to Medicine
Janice P. Nimura
Norton, 2021. 336 pp.



In 1910, both Blackwell sisters died. Although not mentioned in Nimura's insightful biography, that year also saw the publication of the infamous Flexner Report, which laid out uniform criteria for medical training that ultimately shuttered most medical schools in the United States. The schools that remained open were primarily in urban areas of the U.S. Northeast. Most admitted only white men and typically came with a steep price tag—ensuring that only elites became doctors.

In the book's last lines, Nimura writes that "when the Blackwell sisters died, there were more than nine thousand women doctors in the United States, about six percent of all physicians." "Today," Nimura continues, "thirty five percent of physicians—and slightly more than half of all medical students—are female." There are two ways to read these final words. The comparison might suggest the success, over time, of efforts to expand access to medical training, of which the Blackwell sisters were important drivers. Yet it is far from clear whether either Elizabeth or Emily were interested in democratizing medical education.

Nimura's closing comparison, interpreted differently, shows how unequal the U.S. medical profession remains more than 100 years after the Blackwell sisters' intrepid, institution-building work: The preponderance of medical students are women, and yet women make up a mere third of all American doctors. Nimura's closing lines are brilliant because they allow the implications of the Blackwell sisters' work to remain ultimately ambiguous. ■

10.1126/science.abf7416

PHOTOS: SCHLESINGER LIBRARY, RADCLIFFE INSTITUTE, HARVARD UNIVERSITY



Overfishing threatens the endangered shortfin mako shark (*Isurus oxyrinchus*).

PHOTO: CHRIS & MONIQUE FALLOWS/NATUREPL/SCIENCE SOURCE

Edited by **Jennifer Sills**

Shortfin mako sharks speeding to the brink

Overfishing has reduced abundances of sharks worldwide, with functional extinction of some species in former habitat strongholds (*1*, *2*). The world's fastest shark, the shortfin mako (*Isurus oxyrinchus*), is globally endangered (*3*) but remains overfished in the North Atlantic and in high demand commercially for its meat and fins (*3*, *4*). Despite these threats to the species, influential fishing nations continue to block policies that would follow scientific recommendations for rebuilding shortfin mako populations by reducing catches in areas beyond national jurisdictions (the "High Seas"). At the November 2020 Meeting of the International Commission for the Conservation of Atlantic Tunas (ICCAT), the European Union and United States ignored ICCAT's own scientific advice and blocked a catch retention ban for North Atlantic mako proposed by Canada, Senegal, the United Kingdom, Chinese Taipei, and Gabon (*5*). Blocking a retention ban delays hopes of ending overfishing.

The latest shortfin mako stock assessment estimates that North Atlantic stock recovery can be achieved by 2070 with a 60% chance if the total allowable catch—including dead discards—is less than 300 tons (*6*). However, current catches—excluding discards—are already more than 6 times this number (*4*). Even with zero retention, reducing capture and post-release mortalities will be necessary. An annual mortality of 1100 tons, already greatly exceeded in 2019 (*4*), gives only an 8% chance of rebuilding by 2070 (*6*).

In 2019, shortfin makos were included in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora, regulating trade for this species (*7*), but additional measures are needed. In addition to reversing course to support a retention ban, the European Union and United States should show appropriate leadership in shark conservation by decreasing fishing effort and implementing area-based management tools (such as marine reserves) (*8*) that are capable of protecting key shark hotspots (*9*). Achieving these as part of a new United Nations High Seas treaty (*10*) may fast become the last management refuge for rebuilding oceanic shark populations.

David W. Sims[1,2]*, Gonzalo Mucientes[3,4], Nuno Queiroz[1,4]

[1]Marine Biological Association of the UK, Plymouth PL1 2PB, UK. [2]University of Southampton, Southampton SO14 3ZH, UK. [3]Instituto de Investigaciones Marinas–Spanish Research Council (IIM-CSIC), 36208 Vigo, Spain. [4]Centro de Investigação em Biodiversidade e Recursos Genéticos/Rede de Investigação em Biodiversidade e Biologia Evolutiva, Universidade do Porto, 4485-661 Vairão, Portugal.
*Corresponding author. Email: dws@mba.ac.uk

REFERENCES AND NOTES

1. D. J. McCauley *et al.*, *Science* **347**, 1255641 (2015).
2. M. A. MacNeil *et al.*, *Nature* **583**, 801 (2019).
3. C. L. Rigby *et al.*, *Isurus oxyrinchus* (The IUCN Red List of Threatened Species, 2019).
4. International Commission for the Conservation of Atlantic Tunas, "2020 Standing Committee on Research and Statistics (SCRS) advice to the commission" (Madrid, Spain, 2020).
5. ICCAT, "2020 Commission Documents: Summary Report by the Chair of Panel 4" (Madrid, Spain, 2020).
6. ICCAT, "Report of the 2019 Shortfin Mako Shark Stock Assessment Update Meeting" (Madrid, Spain, 2019); www.iccat.int/Documents/SCRS/DetRep/SMA_SA_ENG.pdf
7. Convention on International Trade in Endangered Species of Wild Fauna and Flora, in *Resolution Conf. 12.6 (Rev. CoP18)* (2019); https://cites.org/sites/default/files/document/E-Res-12-06-R18.pdf.
8. G. Ortuño Crespo *et al.*, *Mar. Pol.* **122**, 104102 (2020).
9. N. Queiroz *et al.*, *Nature* **572**, 461 (2019).
10. G. Ortuño Crespo *et al.*, *Nat. Ecol. Evol.* **3**, 1273 (2019).

10.1126/science.abg2355

Mediterranean sharks and rays need action

After news spread that endangered sharks were being openly sold off in chunks in marketplaces in Tunisia, more than 30 nongovernmental organizations appealed to Tunisian authorities to put an end to the practice (*1*). In response, the Tunisian government circulated a note to fishery officials at ports to enforce the ban on catching, landing, and marketing of 21 out of 24 species of sharks and rays listed on Annex II of the Barcelona Convention

(*2*), pending the inclusion of these species in the national legislation. This is a positive step in the protection of these species, given that Tunisia has the largest North African fishing fleet (*3*), second only to Libya for shark fishing in the Mediterranean, and that North African countries landed 70% of Mediterranean elasmobranch catches in 2018 (*4*).

More than half of Mediterranean sharks and rays that are caught deliberately or as bycatch are Critically Endangered, Endangered, or Vulnerable (*5*). An additional 20% of species lack sufficient data to classify (*5*). The list of species landed at ports includes many that have already disappeared in several parts of the Mediterranean, such as the spiny butterfly ray, angel shark, guitarfish, sandbar shark, and blue shark, as well as the majestic shortfin mako and iconic great white shark (*6*).

The recent action by Tunisian authorities will not be enough to curtail fishing of elasmobranchs that are in danger of extinction. The country must enact a new national fishery law that without legal uncertainty will protect threatened sharks and rays along its coast. Similarly, the governments of all Mediterranean countries, which have for too long turned a blind eye to mislabeled products and food fraud (*7*, *8*) and illegal, unreported, and unregulated fishing of sharks and rays (*6*), should more strictly enforce their existing conservation and food security policies by increasing controls in harbors, markets, and processing industries. Further protective measures should include international agreements to create sanctuaries offshore and in the birthing and nursery grounds where many of the endangered sharks and rays congregate in the Mediterranean (*9*). These should be key priorities for EU policymakers tasked with the success of the European Green Deal and the international community's efforts to meet sustainable targets on biodiversity and healthy oceans by 2030.

Marco Milazzo[1]*, Carlo Cattano[2], Sara A. A. Al Mabruk[3], Ioannis Giovos[4,5]
[1]Department of Earth and Marine Sciences, University of Palermo, I-90123 Palermo, Italy. [2]Department of Integrative Marine Ecology, Stazione Zoologica Anton Dohrn (SZN), Lungomare Cristoforo Colombo, I-90149 Palermo, Italy. [3]Omar Almukhtar University, Al Beida, Green Mountain, Libya. [4]iSea, GR-54645 Thessaloniki, Greece. [5]Department of Animal Production, Fisheries, & Aquaculture, University of Patras, Mesolonghi, Greece.
*Corresponding author.
Email: marco.milazzo@unipa.it

REFERENCES AND NOTES

1. Letter to Minister Akissa Bahri (2020); https://isea.com.gr/wp-content/uploads/2020/11/Tunisia-WS-24-11-20.pdf.
2. United Nations Environment Programme, Mediterranean Action Plan, Mediterranean Biodiversity Centre, "SPA/BD protocol—Annex II: List of endangered or threatened species" (2018); www.rac-spa.org/sites/default/files/annex/annex_2_en_20182.pdf.
3. Food and Agriculture Organization of the United Nations, General Fisheries Commission for the Mediterranean, "The State of Mediterranean and Black Sea Fisheries 2020" (Rome, 2020); www.fao.org/documents/card/en/c/cb2429en.
4. Food and Agriculture Organization of the United Nations, Fisheries Division, Statistics and Information Branch, "GFCM (Mediterranean and Black Sea) capture production" (2020); www.fao.org/fishery/statistics/gfcm-capture-production/en.
5. F. Serena *et al.*, *Eur. Zool. J.* **87**, 497 (2020).
6. World Wildlife Fund Mediterranean Marine Initiative, "Sharks and rays: A deadly harvest" (2020); https://wwfeu.awsassets.panda.org/downloads/wwf_iuu_sharks_and_rays_briefing_2020_v6__1_.pdf.
7. V. Ferrito *et al.*, *Foods* **8**, 537 (2019).
8. I. Giovos *et al.*, *Mar. Pol.* **112**, 103730 (2020).
9. United Nations Environment Programme, Mediterranean Action Plan, Mediterranean Biodiversity Centre, "Action plan for the conservation of cartilaginous fishes (Chondrichtyans) in the Mediterranean Sea," M. N. Bradai, Ed. (Tunis, 2020); www.rac-spa.org/sites/default/files/action_plans/elasmo.pdf.

COMPETING INTERESTS

M.M. receives funding from the University of Palermo Fondo Finalizzato alla Ricerca project. Elasmobranch conservation efforts of iSea (I.G.'s employer) in the Mediterranean sea include support from in-kind contributions to OceanCare, the Shark Conservation Fund, the Green Fund, and the Save Our Seas Foundation.

10.1126/science.abg1943

Tilapia farming threatens Brazil's waters

Once again, the Brazilian government is threatening native biodiversity by facilitating invasive alien species farming (*1*). The Brazilian Aquaculture and Fisheries Secretary recently announced a Federal Decree (10576/2020) to boost aquaculture by implementing farming of Nile tilapia in 60 of Brazil's 73 reservoirs (*2*). The Decree also transfers responsibilities, such as permits, licensing, and enforcement—which until now have rightly been the domain of the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA)—to the Brazilian Aquaculture and Fisheries Secretary. The Nile tilapia (*Oreochromis niloticus*), a cichlid fish native to Africa, is an invasive species that has spread worldwide (*3*). The ecological impacts associated with its farming are extensive (*3*, *4*), and these policy changes should be halted before they cause further environmental damage to Brazil's fragile aquatic ecosystems.

Politicians are relying on the misplaced idea that if the Nile tilapia escape, wild fishes would prey upon them, and they would not thrive in the reservoirs' depths (*2*, *5*, *6*). However, there is extensive evidence showing that once the Nile tilapia escapes, it survives, reproduces, establishes populations, and causes environmental and fisheries impacts (*7*, *8*). Moreover, cage farming brings negative impacts beyond escapes. The practice can alter reservoirs' abiotic conditions, feed leftovers and fish excrement increase nitrogen (i.e., cause eutrophication), and escaped fish alter the ecosystems' trophic interactions (*5*).

There could also be economic fallout: Local fishermen already face damages caused by habitat changes such as artificial lakes and dams (*9*). The addition of cage aquaculture would increase competition for space, decrease the availability of native fish (*10*), and shift or even destroy traditional fisheries. Such losses could be catastrophic, with long-term consequences for the communities (*6*, *11*) who live around the reservoirs and adjacent rivers, natural lakes, and streams.

In areas of Brazil where invasive alien species were introduced decades ago, regional and national policies have been promoting them to "native" status, "naturalizing" them by decree (*6*). This supposed solution hides the issue rather than offering protection and violates international treaties (such as the Convention on Biological Diversity) as well as invasion science consensus definitions. Given that cage farming of the invasive Nile tilapia threatens native biodiversity (*12*), Brazil must consult the scientific community before proceeding with its plans.

Patricia Charvet[1]*, Thiago Vinicius Trento Occhi[1], Larissa Faria[1], Barbara Carvalho[1], Clemerson Richard Pedroso[1], Laís Carneiro[1], Matheus Freitas[1], Miguel Petrere-Junior[2], Jean Ricardo Simões Vitule[1]
[1]Laboratório de Ecologia e Conservação, Setor de Tecnologia, Departamento de Engenharia Ambiental, Universidade Federal do Paraná, Curitiba, PR, 81531-970, Brazil. [2]Programa de Pós-raduação em Sustentabilidade de Ecossistemas Costeiros e Marinhos–Universidade Santa Cecília, Boqueirão, 11045-040, Santos, Brazil.
*Corresponding author. Email:pchalm@gmail.com

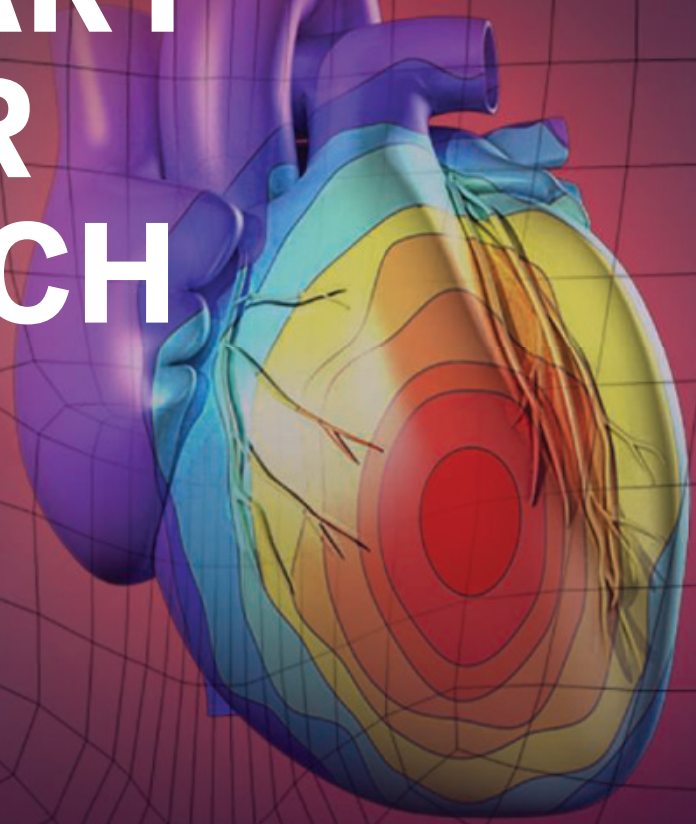
REFERENCES AND NOTES

1. M. F. G. Brito *et al.*, *Science* **361**, 139 (2018).
2. E. Conteúdo, "Peixe que Bolsonaro quer criar em Itaipú pode acabar com espécies nativas" (2020); www.gp1.com.br/noticias/peixe-que-bolsonaro-quer-criar-em-itaipu-pode-acabar-com-especies-nativas-491911.html [in Portuguese].
3. G. C. Canonico, A. Arthington, J. K. McCrary, M. L. Thieme, *Aquat. Conserv.* **15**, 463 (2005).
4. F. A. S. Cassemiro, D. Bailly, W. J. da Graça, A. A. Agostinho, *Hydrobiologia* **817**, 133 (2018).
5. C. W. Martin, M. M. Valentine, J. F. Valentine, *PLOS One*, **5**, e14395 (2010).
6. F. M. Pelicice *et al.*, *Fish Fish.* **18**, 1119 (2017).
7. S. F. Forneck, F. M. Dutra, M. P. Camargo, J. R. S. Vitule, A. M. Cunico, *Hydrobiologia*, 10.1007/s10750-020-04430-8 (2020).
8. V. M. Azevedo-Santos, O. Rigolin-Sá, F. M. Pelicice, *Neotrop. Icthyol.* **9**, 915 (2011).
9. A. A. Agostinho, F. M. Pelicice, L. C. Gomes, *Braz. J. Biol.* **68**, 1119 (2008).
10. J. L. Attayde, J. Brasil, R. A. Menescal, *Fish. Manag. Ecol.* **18**, 437 (2011).
11. G. M. N. Silva, R. M. C. M O. Carvalho, A. C. A. El-Deir, M. C. M. Sobral, M. Siegmund-Schultze, *Reg. Environ. Change*, **18**, 1889 (2018).
12. J. R. S. Vitule, C. A. F. Freire, D. Simberloff, *Fish Fish.* **10**, 98 (2009).

10.1126/science.abg1346

ScienceTranslationalMedicine.org

PUT HUMAN HEALTH AT THE HEART OF YOUR RESEARCH



Submit your research:
cts.ScienceMag.org

Science Translational Medicine
AAAS

Twitter: @ScienceTM
Facebook: @ScienceTranslationalMedicine

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

SEXUAL DIMORPHISM

Enhancer-gene interactions drive split

In many species across the animal kingdom, male and female phenotypes differ. One such example is the wing spot seen in male *Drosophila biarmipes* flies but not the female flies. Galouzis and Prud'homme investigate the X-linked *yellow* gene and its enhancer. Investigating the genetics of the trait, they found evidence that the male-specific phenotype is caused by a trans interaction between the enhancer and gene that silences the gene when it is present in only one copy in the male (which only has a single X chromosome) versus the two copies found in the female. This gene thus appears to be regulated by differing genomic interactions in male and female flies and is an example of the phenomenon known as transvection. —LMZ *Science*, this issue p. 396



The wing spot on the male side of a bilateral gynandromorph *Drosophila biarmipes* fly is the result of a complex, sex-linked expression pattern.

IMMUNOMETABOLISM

A metabolic circuit in T cell immunity

Naïve T cells are metabolically reprogrammed when they differentiate into T effector (T_{eff}) cells, transitioning from a reliance on mitochondrial oxidative phosphorylation to aerobic glycolysis. Xu *et al.* found that lactate dehydrogenase A (LDHA), a glycolytic enzyme that converts pyruvate to lactate, is a key player in this process. T_{eff} cells that differentiate in mice infected with the bacterium *Listeria monocytogenes* turned on LDHA through phosphoinositide 3-kinase (PI3K) signaling. By promoting adenosine triphosphate (ATP) production, LDHA in turn facilitated PI3K-dependent inactivation of the transcription factor Foxo1 needed for effective T_{eff} cell responses. Thus, glycolytic ATP acts like a rheostat that both gauges and regulates PI3K-dependent signaling. This type of positive feedback circuit may also provide a mechanistic explanation for the Warburg effect observed in cancer cells. —STS *Science*, this issue p. 405

METABOLIC EVOLUTION

Yeast switches for glucose and galactose

Some organisms can switch metabolic pathways depending on their environment. One such example is yeast, which can transition between the sugars glucose and galactose as carbon sources. Boocock *et al.* show that this ability has undergone selection, resulting in the maintenance of two incompatible metabolic pathways in a select set of yeast strains within a single species. A phylogenetic analysis supports that these different pathways are mediated by three genes that differ between strains within and among yeast species and likely have been maintained over 10 million to 20 million years. —LMZ *Science*, this issue p. 415

COSMOCHEMISTRY

Two-part formation of the Solar System

Measurements of meteorites have shown that the inner and outer Solar System formed from two distinct reservoirs of material. Existing models have proposed that these were split



False-color transmission electron microscope image of Brewer's yeast

IMAGES (FROM TOP): D. DHARMADHIKARI/N. GOMPEL ©2019; EYE OF SCIENCE/SCIENCE SOURCE

by Jupiter forming first, which would open a gap in the protoplanetary disc. Lichtenberg *et al.* instead argue that the snow line, the boundary between regions containing water vapor and solid ice, migrated first outward and then inward, forming two separate populations of planetesimals. Those planetesimals then grew through collisions to form the planets. Their simulation of this model explains the meteorite data and is consistent with astronomical observations of protoplanetary disks around other stars. —KTS

Science, this issue p. 365

SOLAR CELLS

Opening charge transport pathways

In perovskite solar cells, the insulating nature of passivation layers needed to boost open-circuit voltage also increases the series resistance of the cell and limits the fill factor. Most improvements in power conversion efficiency have come from higher open-circuit voltage, with most fill factor improvements reported for very small-area cells. Peng *et al.* used a nanostructured titanium oxide electron transport layer to boost the fill factor of larger-area cells (1 square centimeter) to 0.84 by creating local regions with high conductivity. —PDS

Science, this issue p. 390

CONVERGENT EVOLUTION

From offense to defense

Venom in snakes is largely used to subdue and/or kill prey, and most venoms have clear actions that facilitate death or paralysis. In one group of snakes, however, venom has evolved and shifted from predation to protection. Specifically, in three different lineages of "spitting" snakes, venom is used to deter predators. Kazandjian *et al.* show that similar adaptions have occurred within these lineages that transform cytotoxic components into a mixture that acts on mammalian sensory neurons and causes pain. The authors argue that increased predation on these lineages led to similar shifts in venom function. —SNV

Science, this issue p. 386

CANCER

An innate response to ovarian cancer

Favorable outcomes for patients with epithelial ovarian cancer have been associated with infiltration of immune cells into tumors. Foord *et al.* investigated the contribution of a subset of immune cells, γδ T cells, to clinical outcomes in these patients. The authors demonstrate that tumor-infiltrating γδ T cells isolated from patients with ovarian cancer have distinct T cell receptors and exhibit more innate immune cell–like functions compared with blood- or ascites-derived γδ T cells. Furthermore, γδ T cell infiltration into tumors was associated with increased survival in patients with ovarian cancer. These findings suggest that promoting γδ T cell responses may be a therapeutic option for ovarian cancer. —CSM

Sci. Transl. Med. **13**, eabb0192 (2021).

ANIMAL ROBOTS

Resonance on the pulse

Jellyfish and squid efficiently propel themselves through water by pulse-jetting, a cyclic process of expanding and contracting a hollow cavity in their bodies that in turn drives the intake and expulsion of fluid. Inspired by these creatures, Bujard *et al.* developed a squid-like robot that mimics their propulsive behavior. The flexible, self-propelled robot can achieve a swimming efficiency that is comparable to its biological counterparts by leveraging resonance to enhance its jetting propulsion. —MML

Sci. Robot. **6**, abd2971 (2021).

IN OTHER JOURNALS

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



FRAGRANCE CHEMISTRY

Unveiling the secret of vetiver's scent

You may not have heard of vetiver oil, but your nose is almost certainly familiar with the characteristic woody fragrance that it lends to thousands of perfumes and colognes. Given the widespread use of this natural extract, it is surprising that the specific molecular component most responsible for the odor has remained ambiguous. List *et al.* report a total synthesis of the tricyclic vetiver constituent (+)-2-*epi*-ziza-6(13)en-3-one, which olfactometry supports as the most potent source of the woody-amber scent. An enantioselective Mukaiyama-Michael addition was key to setting the stereochemistry. —JSY

Angew. Chem. Int. Ed. 10.1002/anie.202014609 (2020).

The major active olfactory component of vetiver (pictured), the basis of numerous perfumes and colognes, has been synthesized.

NEUROIMMUNOLOGY

Letting the nodes know

The idea that the immune and nervous systems can take a coordinated approach to potential harms is beginning to win traction. The lymph nodes constantly monitor lymph draining from peripheral tissues for potential pathogens. Using a variety of techniques, Huang *et al.* mapped sensory neuronal, immune, and stromal components of skin-draining lymph nodes in mice. They found sensory neurons at the lymph node periphery, poised to monitor inflammation-induced mechanical, chemical, and cellular changes. The sensory neurons were predominantly found to represent peptidergic nociceptors. The lymph node–innervating neurons were able to respond to inflammatory signals, potentially through communication with other lymph node–resident cell types. Lymph node stromal cells exhibited the highest potential for interaction with local sensory

PHOTO: DINODIA PHOTOS/ALAMY STOCK PHOTO



At the boundary between forest and savanna, trees create a fire break by shading out highly flammable grasses.

FIRE ECOLOGY

Forest-savanna fire guard

Natural fire regimes play an important role in controlling the boundaries between tropical forests and savannas. Cardoso *et al.* show that the boundaries themselves play a role in controlling fire. In an analysis of forest-savanna boundaries in Africa, they found that the boundary vegetation communities acted as a fire buffer. The boundary trees shaded out the more flammable grasses characteristic of savanna, thus shielding the forest interior from fire. Under future climate change, boundary tree communities may be expected to encroach into savanna areas, extending the spread of woody vegetation and reducing the extent of savannas. —AMS

J. Ecol. 10.1111/1365-2745.13549 (2020).

fibers. Stimulating the local sensory neurons elicited responses from lymphatic endothelial cells, blood endothelial cells, nonendothelial stroma, and certain populations of innate leukocytes. The identification of sensory neurons that innervate lymph nodes reveals a sensory neuron-stroma axis likely to play a role in regulating responses to infection and inflammation. —SMH

Cell 10.1016/j.cell.2020.11.028 (2020).

IMMUNOMETABOLISM

Neutrophil bioenergetics are specialized

Neutrophils are short-lived and require rapid access to adenosine triphosphate (ATP) through glycolysis for energy-intensive responses such as migration, pathogen control, and apoptosis. Sadiku *et al.* describe the metabolic processes that enable this specialization. The authors found that neutrophils cycle between gluconeogenesis and glycogenesis to maintain a consistent supply of ATP even when glucose is limited, such as at sites of infection. This process is dynamically regulated; for example, altitude-induced hypoxia in humans leads to increased glycogen storage in neutrophils. Furthermore, neutrophils from patients with chronic obstructive pulmonary disease show defective glycogen cycling, a phenomenon that is associated with impaired bacterial killing of *Streptococcus pneumoniae* and accelerated apoptosis. —STS

Cell Metab. 10.1016/j.cmet.2020.11.016 (2020).

EXTINCTION

Coevolution in India

Large-bodied vertebrates are widely threatened in today's world, but the loss of such megafauna began many thousands of years ago at the end of the Pleistocene. The main explanations for this phenomenon have converged on climate change and human impact. That Africa, the center of human evolution, experienced fewer extinctions than other regions supported the idea of coevolution with humans. Hominids were present in the Indian subcontinent for more than a million years, and many Indian megafauna should also have had the opportunity to coevolve. Jukar *et al.* found a rate of megafaunal loss in India more like that in Africa than in the heavily hit North America. Although the number of species lost in India was smaller, the extinctions were among the largest vertebrates present at the time. These results suggest that animal species in India may have had some opportunity to adapt to hominid presence, at least until highly lethal *Homo sapiens* came along. —SNV

Palaeogeogr. Palaeoclimatol, Palaeoecol. **562**, 110137 (2020).

IMPACT CRATERS

Oceans suppress crater formation

Asteroid impacts leave craters on land, but much of Earth's surface is ocean. Impacts on sufficiently shallow water should still produce some craters, though fewer than expected have been found. Nishizawa *et al.* performed laboratory hypervelocity impact experiments onto targets submerged in water. Extrapolating these to asteroid-scale impacts shows that the energy dissipation is more efficient than previous computer simulations had suggested, owing to deformation and fragmentation of the impactor in the water. Crater formation is suppressed when the water depth is only a few times the diameter of the impacting object, potentially explaining why so few underwater craters have been found on Earth. —KTS

J. Geophys. Res. E Planets **125**, e2019JE006291 (2020).

BIOINSPIRED MATERIALS

Variations enhance sensitivity

Gradient and hierarchical materials allow for a gradual change in composition, behavior that can be seen in the mechanical or optical properties of many biological organisms, even for those components composed of very few elements. Lee *et al.* considered gradient materials that enhance the touch sensitivity of electronic skin materials. The conducting sensors are composed of two multilayer films, each with a domed surface pattern that allows for gradual contact and compression on squeezing. On the top layer, the PDMS domes are coated with a dispersion of PEDOT:PSS and polyurethane to create layers with varying mechanical and electrical properties. This combination gives a far greater sensitivity, linear response range, and lowered response time when a pressure is applied. —MSL

ACS Nano 10.1021/acsnano.0c09581 (2020).

PHOTO: ARTERRA PICTURE LIBRARY/ALAMY STOCK PHOTO

REVIEW SUMMARY

NEURODEVELOPMENT

Mapping the molecular and cellular complexity of cortical malformations

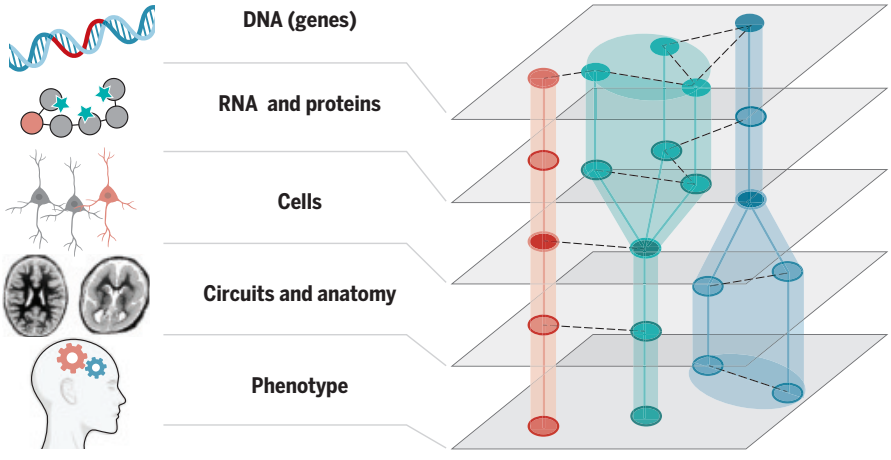
Esther Klingler, Fiona Francis, Denis Jabaudon*, Silvia Cappello*

INTRODUCTION: The cerebral cortex, or neocortex, is critical to key mammalian skills such as language, sociability, and sensorimotor control. This structure consists of dozens of specialized types of neurons organized across layers and areas, which are generated during development by diverse types of progenitors. Newborn neurons then undergo sequential molecular programs that drive their specific local and long-range circuit connectivity and adult function. Although they are necessary for proper cortical function to emerge, these myriads of molecular and cellular developmental processes provide multiple points of vulnerability for "cortical malformations," which cause various combinations of intellectual and/or motor disabilities that are often associated with seizures. These disabilities include microcephaly (decreased brain size), lissencephaly (loss of cortical folding), polymicrogyria (numerous small cortical folds), dysplasia (abnormal cortical lamination, which can be focal), and heterotopias (abnormally positioned cells in periventricular or subcortical regions). Despite the toll on patients and their caregivers, only limited treatments exist and although some causal genes have been identified, the sequence of events linking molecular disruption with clinical expression mostly remains obscure.

RATIONALE: To better understand cortical malformations and to highlight potential points of intervention, we first present basic principles of neocortical development and point out vulnerable cellular compartments and processes. Second, we dissect different "levels" of organization, from genes to cells, circuits, and clinical expression, and illustrate how complex interactions within and across these levels may account for variable disease patterns in cortical malformations. We finally propose a framework integrating these different levels of organization to assist in better understanding and treating such diseases.

RESULTS: We first present basic principles of neocortical development that result from billions of cells undergoing four key sequential and partially overlapping processes: (i) progenitor division and neurogenesis, (ii) neuron migration, (iii) extension of axon and dendrites, and (iv) synaptogenesis. We point out vulnerable cellular compartments and processes, with particular focus on neurogenesis and neuron migration, and highlight potential sources of variability that have precluded the establishment of clear causal relationships across genes and molecules, cell types, circuits, and clinical expression. Starting with genetic and molecular dysfunction, we examine monogenic versus polygenic causes of disease and their convergent (i.e., mutations in distinct genes leading to the same phenotype), divergent (i.e., mutations in a single gene leading to distinct phenotypes), or mixed relationships with disease phenotype(s). The contribution of redundant molecular mechanisms and versatility in protein function to the variability of disease processes is discussed and illustrated by examples. Disrupted spatiotemporal expression of genes, cell type–specific defects, and relationships between cell position and circuit wiring are also covered. Finally, we argue that comparison of gene expression across brain development in different animal models (including mouse and monkey), in humans, and in human-derived brain organoids is particularly important to identify affected processes. As a step in this direction, we provide an online resource (http://www.humous.org) compiling transcriptional maps across embryonic development and neuron differentiation for mouse embryos, human embryos, and human brain organoids.

CONCLUSION: Using select examples, we highlight several levels (i.e., genetic, molecular, cellular, circuit, and behavioral) within and across which combinatorial interactions occur during cortical malformations and which hamper a causal understanding of the disease process. Discerning the processes involved at each of these levels for individual patients is key to providing them and their families with prognostic indicators and therapeutic perspectives. Integrative approaches including electrophysiological, imaging, clinical, and biological data in patients using state-of-the art artificial intelligence algorithms may allow bridging DNA mutation(s) to molecular, cellular, anatomical, and circuit features. This will be instrumental for the physiopathogenic classification of diseases, which is an essential step in patient stratification and in the design of personalized diagnostic and therapeutic tools. ■



Emergence of complexity in cortical malformations. Shown are levels of organization during corticogenesis, from genes to gene expression (RNA and proteins), cells, circuits and anatomy, and phenotype. Each circle represents a given feature of that level. Interactions within levels are linked through complex relationships to states at other levels. Examples of abnormal linear (brandy red), convergent (cyan), and divergent (blue) feature relationships across levels in disease are highlighted.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: denis.jabaudon@unige.ch (D.J.); silvia_cappello@psych.mpg.de (S.C.)
Cite this article as: E. Klingler *et al.*, *Science* 371, eaba4517 (2021). DOI: 10.1126/science.aba4517

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba4517

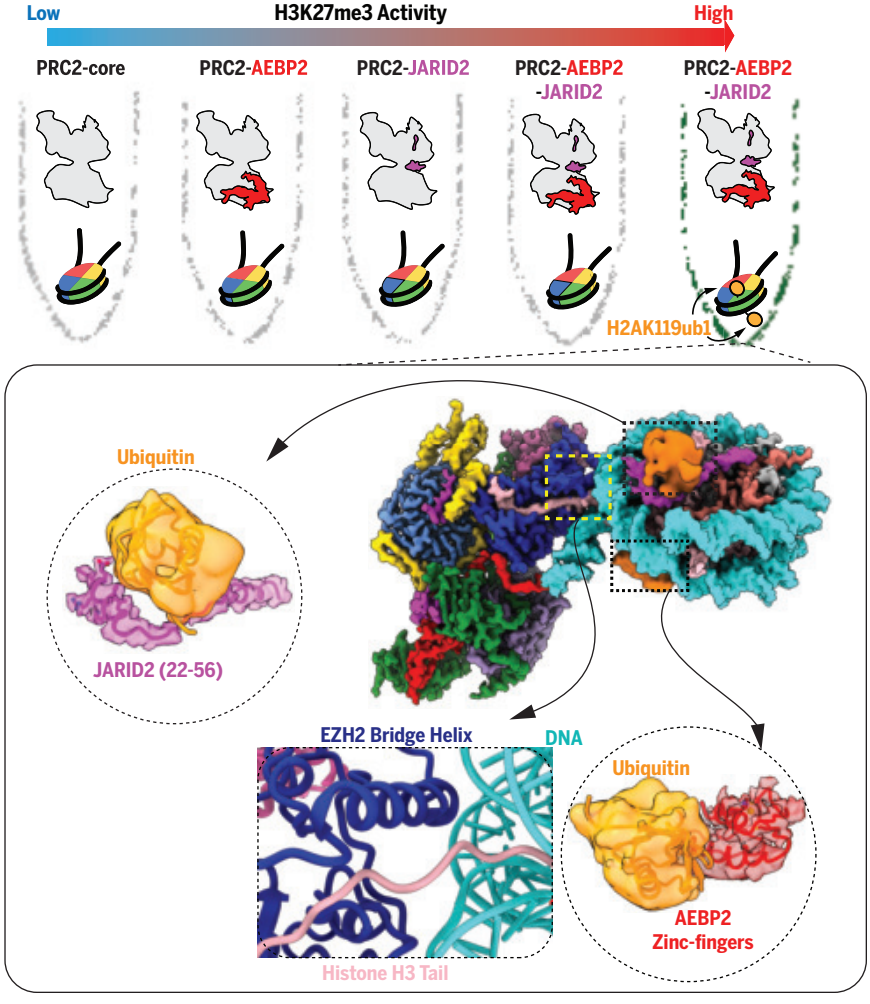
RESEARCH ARTICLE SUMMARY

CHROMATIN

JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications

Vignesh Kasinath*, Curtis Beck, Paul Sauer, Simon Poepsel, Jennifer Kosmatka, Marco Faini, Daniel Toso, Ruedi Aebersold, Eva Nogales*

INTRODUCTION: Histone modification activity of the polycomb repressive complexes 1 and 2 (PRC1 and PRC2) is critical for the establishment and maintenance of gene expression patterns and, thus, to the maintenance of cell identity. Distinct classes of cofactor proteins are known to regulate the functional activity and interplay of these two complexes, but we presently lack a comprehensive, mechanistic understanding of this process. Furthermore, PRC2 cofactors like AEBP2 and JARID2 also play a role in mediating the cross-talk between different histone posttranslational modifications and PRC2 recruitment and activity—a function that is important for the regulated control of gene expression.



Regulation of PRC2 methyltransferase activity by cofactors JARID2 and AEBP2 and by histone posttranslational modification. Although core PRC2 is a weak enzyme, it is allosterically activated by JARID2 and AEBP2. The presence of monoubiquitinated histone H2A, the product of PRC1 activity, is recognized by both JARID2 and AEBP2 through interactions that likely mediate recruitment of PRC2 to polycomb sites in the genome and further activate the methyltransferase activity of PRC2 on K27 of histone H3.

RATIONALE: PRC1 is an E3 ubiquitin ligase responsible for the monoubiquitination of histone H2A (H2AK119ub1), a histone mark recognized by PRC2 and linked to its genomic recruitment. We used cryo–electron microscopy (cryo-EM) and biochemical activity assays to probe the role played by PRC2 cofactors JARID2 and AEBP2 in the recognition of H2AK119ub1 and the regulation of PRC2 activity. We extended our cryo-EM and biochemical activity analysis to examine the possible role played by JARID2 and AEBP2 in the cross-talk between the histone H3K4me3, H3K36me3 modifications linked to transcriptionally active regions, and PRC2 activity.

RESULTS: We find that JARID2 recognizes both the ubiquitin moiety in H2AK119ub1 and the conserved histone H2A-H2B acidic patch. We also observe that the tandem zinc fingers of AEBP2 interact with ubiquitin and the histone H2A-H2B surface on the other side of the nucleosome. Biochemical assays show a secondary activation of PRC2 by JARID2 and AEBP2 on H2AK119ub1-containing nucleosomes besides the primary EED-mediated allosteric activation of PRC2 by methylated JARID2. Furthermore, we also find that the joint presence of JARID2 and AEBP2 partially reduces the inhibition of PRC2 methyltransferase activity by the transcriptionally active histone posttranslational modifications H3K4me3 and H3K36me3. Cryo-EM visualization of PRC2 that contains JARID2 and AEBP2 interacting with a H3K4me3-containing nucleosome shows the coexistence of states in which the histone H3 tail is either absent or engaged and reaching the catalytic site in PRC2, which provides a physical basis for the partial activity of the complex on H3K4me3-containing nucleosomes.

CONCLUSION: Our studies indicate that cofactors JARID2 and AEBP2 play a crucial role in both the recruitment and activation of PRC2 through their recognition of H2AK119ub1, which is generated by PRC1. Additionally, our work suggests that JARID2 and AEBP2 are likely to play a key role in regulating PRC2 activity on genomic regions with active transcription marks. The examination of the genomic distribution in embryonic stem cells of PRC2 core proteins together with JARID2 and AEBP2 will be important to further define their role in the tight regulation of PRC2 activity. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: enogales@lbl.gov (E.N.); vigneshk@berkeley.edu (V.K.)
Cite this article as V. Kasinath *et al.*, *Science* 371, eabc3393 (2021). DOI: 10.1126/science.abc3393

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc3393

RESEARCH ARTICLE SUMMARY

CHROMATIN

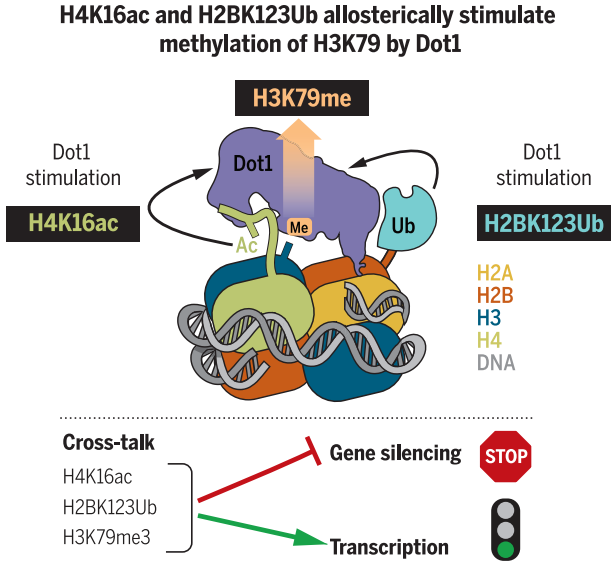
Regulation of the Dot1 histone H3K79 methyltransferase by histone H4K16 acetylation

Marco Igor Valencia-Sánchez*, Pablo De Ioannes*, Miao Wang*, David M. Truong, Rachel Lee, Jean-Paul Armache, Jef D. Boeke, Karim-Jean Armache†

INTRODUCTION: Nucleosomes—the primary, repeating unit of chromatin—package and protect the genome while also transmitting various regulatory signals through, in part, posttranslational modifications of histones. These histone modifications (for example, acetylation, ubiquitination, or methylation) affect critical processes such as transcription, replication, recombination, and repair, often forming complicated networks to ensure finely tuned signaling for chromatin enzymes. One such enzyme—evolutionarily conserved disruptor of telomeric silencing (Dot1)—catalyzes mono-, di-, and trimethylation of histone H3 lysine 79 (H3K79). H3K79 methylation by Dot1 is a prominent example of trans-histone cross-talk, a process in which one histone and its modification affects the modification of another histone. In mammals, the human homolog Dot1L plays critical roles in embryogenesis and hematopoiesis. Although recent advances have provided insights into Dot1L stimulation through histone H2B ubiquitination, how other modifications mechanistically regulate Dot1 activity is not known. We were particularly interested in how histone lysine acetylation contributes to the regulation of chromatin enzymes. Lysine acetylation plays pivotal roles in chromatin decondensation, transcriptional activation, and maintenance of euchromatin, serving as a general antisilencing mark in eukaryotes. Here, we present mechanistic studies that show how histone acetylation regulates the activity of Dot1.

RATIONALE: Histone H4 acetylation, histone H2B ubiquitination, and H3 methylation are conserved cotranscriptional histone modifications that work together to ensure the appropriate regulation of chromatin structure during transcription. The mechanisms of cross-talk between these modifications and enzymes that deposit them are crucial for understanding transcription. Advances in cryo–electron microscopy (cryo-EM), the ability to make specifically acetylated and ubiquitinated nucleosomes, and established in vitro and in vivo assays allowed us to determine the detailed mechanisms of yeast Dot1 regulation through histone acetylation and ubiquitination.

RESULTS: We tested nucleosomes with different acetylation states of histone H4 in vitro, and we show that Dot1 is allosterically stimulated by acetylation of H4 and that this effect is specific to lysine 16 (H4K16ac). The other known acetylation targets on histone H4 (H4K5ac, H4K8ac, and H4K12ac) do not stimulate the activity of Dot1, which highlights the distinctive role of H4K16 acetylation in regulating chromatin structure. We also show that the effect of H4K16 acetylation is direct and further enhanced by H2B ubiquitination (H2BUb), resulting in an optimal catalytic rate for Dot1. To gain mechanistic insights into stimulation by H4K16ac and its coordination with H2BUb, we determined two cryo-EM structures: one of Dot1 in complex, with nucleosomes bearing both H4K16ac and H2Bub, and the second of Dot1, with a nucleosome bearing only H2BUb. Upon examining our cryo-EM dataset of Dot1 bound to the nucleosome containing unacetylated H4, the particles classified into two main three-dimensional (3D) classes. In the first class, Dot1 is bound to the nucleosome in a catalytic conformation. In the second class, it is bound in a noncatalytic conformation. This is different from the dataset in which Dot1 is bound to an H4K16 acetylated nucleosome. When H4K16ac is present, the cryo-EM data is more homogeneous, and Dot1 is bound to the nucleosome predominantly in a catalytic conformation. This suggests a model in which acetylation of the H4 tail restricts the sampling space of Dot1, resulting in an active conformation leading to increased activity. We therefore propose that H2BUb partially restricts the conformation of yeast Dot1 on the nucleosome and that H4K16ac further restricts and stabilizes the active conformation. Comparing both of these cryo-EM structures allowed us to identify residues that are critical for Dot1 stimulation by H4K16ac and H2BUb. Site-directed mutagenesis of Dot1 coupled with enzymatic assays on nucleosomes revealed the details of these interfaces. These results show that the allosteric stimulation of Dot1 by H4K16ac and H2BUb plays a crucial role in H3K79 di- and trimethylation.



H3K79 methylation by Dot1 is allosterically stimulated by H4K16 acetylation and by H2BK123 ubiquitination. Cross-talk between different histone posttranslational modifications can orchestrate distinct chromatin states to regulate transcription and gene silencing.

CONCLUSION: This work demonstrates how Dot1 is regulated by histone acetylation and how H4K16ac coordinates with H2BUb to regulate Dot1. H4K16ac plays a critical role in opening chromatin structure by counteracting the binding of silencing proteins, while simultaneously stimulating an enzyme that is important for transcription. We provide an example in which the activity of the fundamental methyltransferase Dot1 is modulated through cross-talk between distinct histone modifications to ensure optimal maintenance and propagation of an epigenetic state. Cross-talk such as this may represent a general property of chromatin enzymes. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: karim-jean.armache@nyulangone.org
Cite this article as M. I. Valencia-Sánchez *et al.*, *Science* 371, eabc6663 (2021). DOI: 10.1126/science.abc6663

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc6663

RESEARCH ARTICLE SUMMARY

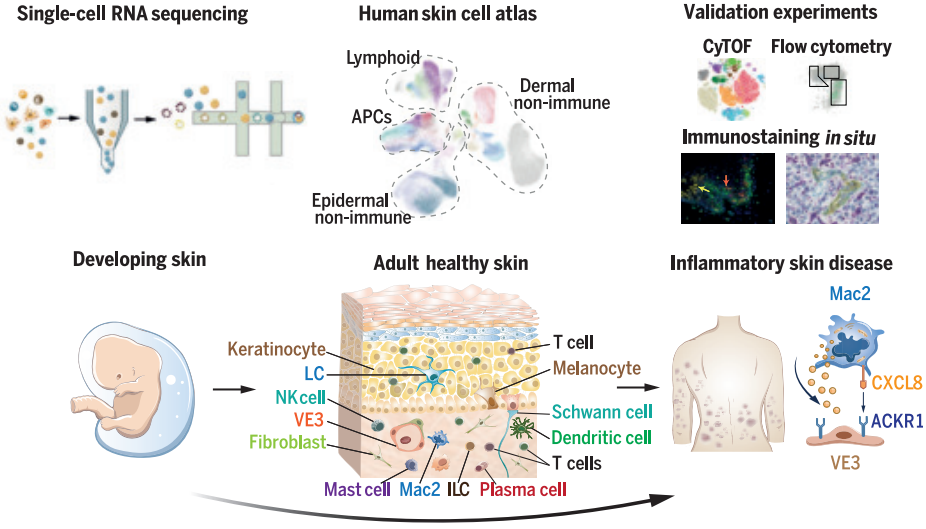
SINGLE-CELL GENOMICS

Developmental cell programs are co-opted in inflammatory skin disease

Gary Reynolds*, Peter Vegh*, James Fletcher*, Elizabeth F. M. Poyner*, Emily Stephenson, Issac Goh, Rachel A. Botting, Ni Huang, Bayanne Olabi, Anna Dubois, David Dixon, Kile Green, Daniel Maunder, Justin Engelbert, Mirjana Efremova, Krzysztof Polański, Laura Jardine, Claire Jones, Thomas Ness, Dave Horsfall, Jim McGrath, Christopher Carey, Dorin-Mirel Popescu, Simone Webb, Xiao-nong Wang, Ben Sayer, Jong-Eun Park, Victor A. Negri, Daria Belokhvostova, Magnus D. Lynch, David McDonald, Andrew Filby, Tzachi Hagai, Kerstin B. Meyer, Akhtar Husain, Jonathan Coxhead, Roser Vento-Tormo, Sam Behjati, Steven Lisgo, Alexandra-Chloé Villani, Jaume Bacardit, Philip H. Jones, Edel A. O'Toole, Graham S. Ogg, Neil Rajan, Nick J. Reynolds, Sarah A. Teichmann†, Fiona M. Watt†, Muzlifah Haniffa†

INTRODUCTION: Human skin provides vital protection from water loss and from external insult through structural adaptations and interplay with the innate and adaptive immune systems. The skin develops and functions in an aquatic environment in utero but rapidly adapts to a contrasting set of physical and pathogenic challenges after birth. The changes that take place across this complex multicellular system during development and upon perturbation by immune-mediated inflammatory diseases are poorly understood. A detailed study will facilitate the development of therapeutic interventions for inflammatory skin disease.

RATIONALE: We generated a comprehensive atlas of human skin in early prenatal life, in adulthood, and during inflammatory skin disease by profiling the transcriptomes of more than 500,000 single cells. We analyzed human embryonic skin between 7 and 10 post-conception weeks, healthy adult skin surplus from mammoplasty surgery, and skin biopsies from patients affected by atopic dermatitis (AD) and psoriasis, two common inflammatory skin diseases. Additionally, we performed single-cell T cell receptor analysis to assess T cell clonality in disease. Validation experiments were conducted at the protein level and used mass cytometry, flow cytometry, and immunostaining in situ of skin biopsies from healthy skin and patients with AD and psoriasis, including a cohort of AD patients before and during treatment with oral methotrexate.



The human skin cell atlas. We used single-cell RNA sequencing to build human skin cell atlases across development, homeostasis, and disease; the analyses were supported by various validation modalities. By delineating cell states in embryonic skin, healthy adult skin, and inflammatory skin diseases (atopic dermatitis and psoriasis), we identified a reemergence of developmental programs in disease mediated by interactions between Mac2 and VE3. APCs, antigen-presenting cells; CyTOF, cytometry by time-of-flight; LC, Langerhans cell; NK, natural killer cell; VE, vascular endothelial cell; Mac, macrophage; ILC, innate lymphoid cell.

RESULTS: Thirty-four cell states were identified in healthy human skin across the collective dataset, with dynamic changes in the nature and abundance of single-cell gene expression profiles identified across embryonic and adult life and upon perturbation during inflammatory skin disease. This resource can be accessed via an interactive browsable web portal, https://developmentcellatlas.ncl.ac.uk/datasets/hca_skin_portal. Analyses revealed that the immune system of first-trimester embryonic skin consists mainly of innate lymphocytes and macrophages. In adult skin, we defined two inferred trajectories for keratinocyte differentiation and the presence of endothelial cells that formed dilated postcapillary venules. We revealed a migratory dendritic cell (DC) signature in healthy adult skin that is conserved in murine DCs. The migratory DC signature was also evident in the developing human thymus and additional disease states. We identified clonally expanded disease-associated cytotoxic T cells (Tc IL13/IL22 cells) in lesional AD and Tc17/T helper 17 (T_H17) cells in lesional psoriasis. We demonstrated the reemergence of prenatal cellular programs mediated by Mac2 macrophages via the chemokine CXCL8 interacting with the venular capillary marker ACKR1 on VE3 vascular endothelial cells in diseased skin. This interaction is implicated in lymphocyte recruitment and angiogenesis. We identified and validated in situ the expansion of Mac2 and VE3 in lesional AD and lesional psoriasis skin, their close apposition in AD and psoriasis tissue, and their reduction in AD skin during methotrexate treatment, which aligned with an improvement in the clinical severity of disease in this patient cohort.

CONCLUSION: Our single-cell atlas of human skin from prenatal life, healthy adults, and AD and psoriasis patients highlights the dynamic nature of cutaneous homeostasis and immunity. Our study provides insights into perturbed and co-opted developmental cellular programs in inflammatory skin disease. These results may provide potential future translational targets to improve the diagnosis and molecular classification of these diseases and to guide treatment strategies. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: st9@sanger.ac.uk (S.A.T.); fiona.watt@kcl.ac.uk (F.M.W.); m.a.haniffa@newcastle.ac.uk (M.H.)
Cite this article as G. Reynolds *et al.*, *Science* 371, eaba6500 (2021). DOI: 10.1126/science.aba6500

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba6500

RESEARCH ARTICLE

PLANET FORMATION

Bifurcation of planetary building blocks during Solar System formation

Tim Lichtenberg[1]*, Joanna Drążkowska[2], Maria Schönbächler[3], Gregor J. Golabek[4], Thomas O. Hands[5]

Geochemical and astronomical evidence demonstrates that planet formation occurred in two spatially and temporally separated reservoirs. The origin of this dichotomy is unknown. We use numerical models to investigate how the evolution of the solar protoplanetary disk influenced the timing of protoplanet formation and their internal evolution. Migration of the water snow line can generate two distinct bursts of planetesimal formation that sample different source regions. These reservoirs evolve in divergent geophysical modes and develop distinct volatile contents, consistent with constraints from accretion chronology, thermochemistry, and the mass divergence of inner and outer Solar System. Our simulations suggest that the compositional fractionation and isotopic dichotomy of the Solar System was initiated by the interplay between disk dynamics, heterogeneous accretion, and internal evolution of forming protoplanets.

Planetary systems, including the Solar System, form by accretion from a protoplanetary disk of gas and dust. Astronomical observations of these disks provide evidence for rapid dust coagulation (*1*), show ringed substructure (*2*), and indicate a decrease in total dust mass with disk age (*3*), to below the total masses of fully assembled exoplanetary systems (*4*). This suggests that planet formation starts early. At the onset of planetary accretion, small dust grains coagulate to form larger aggregates (pebbles). These pebbles drift inward under aerodynamic drag and gravitationally collapse when they reach a sufficiently high local overdensity (*5*), which preferentially produces birth planetesimals of diameter ≈100 km (*6*). These bodies provide the seeds for the accretion process, but their direct growth by pebble accretion only becomes efficient after $\gtrsim 10^5$ to 10^6 years, once they have grown to larger embryos through mutual collisions (*7*).

Meteorites record planet formation in the Solar System and constrain the astronomical time scales: Radiometric dating of meteorites indicates that iron core formation (differentiation) in planetesimals occurred within ≈1 million years (Myr) after the formation of Ca-Al–rich inclusions (CAIs) (*8*), the oldest known solids that formed together with the proto-Sun (*9*). Core formation in first-generation, birth planetesimals is driven by internal radiogenic heating by ^{26}Al (half-life of ~7×10^5 years). Therefore, the time interval between planetesimal formation and differentiation ($\gtrsim 10^5$ to 10^6 years) increases with later formation time, so the earliest planetesimals in the inner Solar System formed ≲0.3 Myr after CAI formation (*8*). Combined, evidence from astronomy and geochemistry suggests that planet formation began during the earliest phases of the solar protoplanetary disk.

Complementing astronomical evidence for disk substructure, meteorite data indicate spatial heterogeneity in the isotopic composition of planetary materials (*10*): Isotopic studies of Ti, Cr, Mo, and other elements show variabilities between individual planetary bodies from different orbits (*11*–*13*). Combined with temporal constraints on the accretion process, these divide the outer and inner Solar System reservoirs into chronologically and spatially distinct populations, also known as the carbonaceous chondrite (CC) and noncarbonaceous (NC) reservoirs, after their representative meteorite classes. The isotopic signatures record the heterogeneous distribution of presolar dust from multiple sites of stellar nucleosynthesis and trace transport and mixing processes during planet formation. By contrast, the elemental abundances and internal structure of planetary bodies were modified through geophysical evolution (*14*). Earth's nucleosynthetic isotope composition is NC-like and differs from that of outer Solar System materials (*10*), but its depletion pattern of moderately volatile elements is close to that of CCs (*15*). This has been attributed to impact delivery (*16*) or addition of pebbles (*17*). The inner Solar System planets are depleted in highly volatile elements, such as hydrogen, carbon, and nitrogen, which influenced the availability of surface water and the composition of their atmospheres. This has been linked with the accretion of the terrestrial planets inside the snow line—the boundary between regions where solid water ice is and is not stable during the disk phase—and later delivery of volatile-rich, outer Solar System materials (*18*, *19*).

The observed CC/NC dichotomy has been suggested to arise from an early formation of proto-Jupiter ≲1 Myr after CAI formation, which would inhibit the aerodynamic drift of dust grains and prevent initially heterogeneous disk materials from mixing (*20*, *21*). This scenario constrains the chronology and physical mechanism of Jupiter's growth (*22*), but potentially contradicts astronomical and geochemical evidence on the accretion process (*10*).

Formation of two distinct planetesimal populations

Multiple planetesimal populations can be formed during the build-up and evolution of the solar protoplanetary disk (*23*). We test whether such multistage planetesimal formation is consistent with the accretion chronology as inferred from meteorites. We use the output of previous simulations (*10*, *23*) to explore the CC/NC dichotomy. The simulations cover the initial infall and build-up of gas and dust, during which the disk itself accretes from the surrounding molecular cloud (Class I stage), and include the drift, coagulation, and compositional evolution of the dust and viscous evolution of the gas during later evolution of the disk (Class II stage). In these simulations, the initial material accretes onto the innermost regions and continuously feeds more distant orbits through viscous expansion of the disk. The snow line moves outward during the Class I stage because a dense, compact disk forms, which heats up viscously. With the onset of the Class II stage, the gas density decreases, the temperature structure becomes dominated by stellar irradiation, and the snow line moves inward again.

The snow line location affects the redistribution of dust grains. These are initially well-mixed with the disk gas but radially drift because of coagulation and aerodynamic drag from the ambient gas. Inward-drifting, icy pebbles undergo rapid dehydration and size reduction at the snow line, which reduces their drift velocity and causes a pile-up of solid material. Outward diffusion of water vapor additionally leads to its recondensation onto icy grains just beyond the snow line (the cold-finger effect), locally increasing the density of solids. When the conditions for the formation of dense, gravitationally unstable dust filaments (the streaming instability) are met in the simulation, planetesimals are formed. Planetesimal formation in the simulation therefore preferentially occurs around the

[1]Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK. [2]University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität München, Munich, Germany. [3]Institute for Geochemistry and Petrology, Department of Earth Sciences, ETH Zurich, Zurich, Switzerland. [4]Bayerisches Geoinstitut, University of Bayreuth, Bayreuth, Germany. [5]Institute for Computational Science, University of Zurich, Zurich, Switzerland.
*Corresponding author. Email: tim.lichtenberg@physics.ox.ac.uk

snow line. At 5 Myr into the simulation, which is the approximate upper lifetime of the solar nebula (*10*), we assume that the gas disk dissipates and planetesimal formation halts.

In wind-driven disks with low levels of midplane turbulence, the global angular momentum transport is dominated by near-surface layers (*10*). In this scenario, the outward-then-inward tacking of the snow line generates two distinct episodes of planetesimal formation in different orbital regions and time intervals (Fig. 1). The first, early-formed planetesimal population (hereafter Reservoir I) is triggered by the cold-finger effect between ≈0.2 to 0.35 Myr at orbital locations between ≈1.3 and 7.5 astronomical units (au), proceeding from the inside-out. The second planetesimal population (Reservoir II) mainly arises from large-scale inward-drift and pebble pile-up at the snow line in the Class II stage. This second population starts forming from ≈0.7 Myr between ≈17 and 3 au and proceeds from the outside-in. Because of differences in mechanism efficiency and local dust availability, the two reservoirs differ substantially in their total formed mass: Reservoir I forms about an Earth mass of planetesimals, and Reservoir II about a Jupiter mass (*10*, *23*). The build-up of both reservoirs depends on the local evolution of the pebble flux but in turn influences the coagulation and drift of pebbles toward the inner disk (figs. S1 to S3). We therefore assess the feedback between reservoir formation and continuing planetary accretion.

Heterogeneous accretion and reservoir separation

We aim to determine the dominant mode of solid mass transfer in this disk build-up scenario and to evaluate its consistency with the time scales of planetary accretion derived from radiometric ages (*8*, *24*, *25*). To relate the numerical simulations to geochemical chronology, we equate time zero in the disk model with the time of CAI formation (*9*, *10*). For fixed orbits at 2 and 15 au, approximately representative of Reservoir I and II in the inner and outer disk, we evaluate the anticipated growth time scales for two scenarios of protoplanet accretion (*10*). We first consider collisional growth due to planetesimal-planetesimal interactions, during which the largest planetesimals grow by accreting smaller bodies. Second, we investigate the efficacy of pebble accretion, which is driven by the abundance of dust grains that cross the protoplanet orbit. Consistent with dynamical evidence from the asteroid main belt (*6*) and numerical simulations of the streaming instability (*5*), we assume a maximum planetesimal radius of 300 km and that collisional growth is dominated by planetesimals of 50 km in radius. The orbit- and time-dependent evolution of the pebble flux (i.e., the transfer of solid mass in the disk), pebble size, and the gas disk characteristics are self-consistently derived from the disk and coagulation simulation described above (*10*, *23*).

Under these conditions, the pebble flux is dominated by outward-moving dust grains (Fig. 2A) between ≈0.2 and 0.35 Myr after CAI formation, because the disk accretes and viscously expands during the infall stage. The associated outward-directed pebble flux that passes Reservoir I is not efficient enough to drive substantial growth by pebble accretion, but planetesimal-planetesimal interactions lead to substantial growth within the time scale of disk evolution (Fig. 2B). Between ≈0.35 and 1 Myr, inward-drifting pebbles reach the inner disk, producing approximately equal contributions to accretion from both pebbles and collisions once the disk transitions into the Class II stage (Fig. 1). Subsequently, pebbles substantially contribute to planetesimal growth for ≈1 Myr.

A final disk phase occurs from ≈2 Myr after CAI formation until gas disk dispersal (*10*),

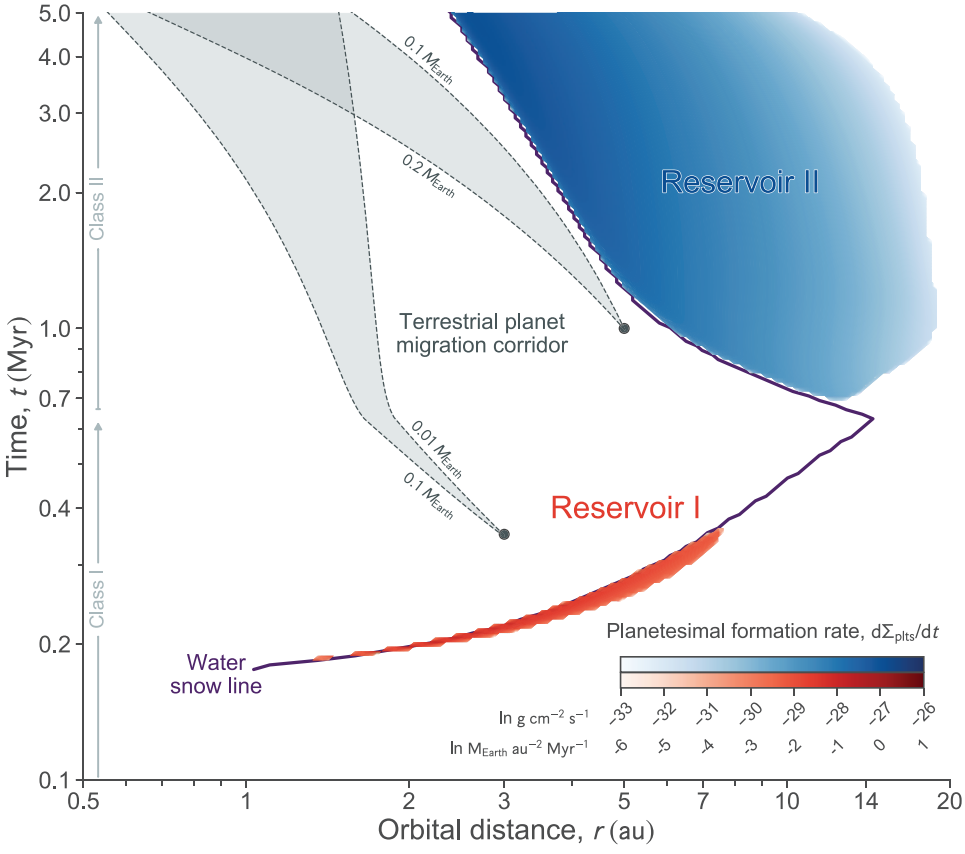


Fig. 1. Formation of two distinct planetesimal populations in the disk simulation. The rate of planetesimal formation (change in surface density per timestep) is shown on logarithmic scale in centimeter–gram-second and natural units. The solid purple line indicates the orbital migration of the snow line; red and blue areas highlight the formation regions of the two planetesimal populations (Reservoir I and Reservoir II). The dashed gray lines (labeled "Terrestrial planet migration corridor"; compare to fig. S9) bound the range of possible planet migration scenarios consistent with the present-day orbits of the terrestrial planets. Although smaller embryos move slowly inward, more rapidly accreting (and thus more massive) embryos would migrate faster toward the proto-Sun. The light gray arrows on the left side indicate the transition from the Class I to Class II disk stage.

during which planetesimal growth by pebbles in Reservoir I stalls (*10*). This is because a substantial fraction of solids, which drift from the outer toward the inner disk, are trapped in Reservoir II, where the pebbles gravitationally collapse to form planetesimals starting from ≈1 Myr after CAI formation onward (Fig. 2A). The build-up and onset of planetary accretion in Reservoir II shields Reservoir I from drifting grains that originate at larger orbital distances. The difference in pebble flux between the inner and outer Solar System reaches more than an order of magnitude. The pebble flux in the inner Solar System decreases to less than one Earth mass per million years after 3 Myr after CAI formation (Fig. 2A).

Divergent geophysical evolution of planetesimal populations

These simulation results suggest that both planetesimals and pebbles contribute to planetary accretion. Rocky protoplanets that form from the planetesimal populations in both reservoirs inherit geochemical and geophysical properties from their precursor bodies (*14*). The two planetesimal reservoirs differ substantially in accretion time and thus in radiogenic, internal heating from the decay of short-lived ^{26}Al. Meteorite data have shown that ^{26}Al heating in early-formed Solar System planetesimals drove their internal evolution and thus altered their structure and bulk composition (*10*, *26*, *27*). We determine the internal and compositional evolution of the Reservoir I and II planetesimal populations that formed in the disk simulation. We constrain the timing of iron core formation and hydrothermal activity (aqueous alteration and degassing) using geodynamic simulations of the thermochemical evolution of planetesimals. Our numerical setup assumes isolated planetesimals that form instantaneously and evolve according to a fluid mechanical model (*10*, *14*, *28*).

We perform 700 single-planetesimal simulations, spanning the parameter range of planetesimal formation time, $t_{form} \in [0.1, 3.0]$ Myr, and planetesimal radius, $R_P \in [1, 300]$ km (Fig. 3A). To quantify their compositional evolution, we use four thermochemical criteria (Fig. 3B) to determine the onset and cessation of core formation and hydrothermal activity (*10*). Core formation in planetesimals initiates with the percolation of Fe,Ni–S liquids before silicate melting occurs and terminates once the interior is sufficiently melted to form an internal magma ocean. At this stage, metal droplets rain out from the convecting magma flow, producing complete differentiation between the metal core and liquid silicate mantle (*28*). Quantitatively, we assume that core formation occurs in regions with temperatures T higher than the threshold for metal-sulfide percolation, $T_{perc} \equiv 1273$ K, and completes for silicate melt fractions φ above the rheological transition, $\varphi_{rain} \equiv 0.4$, where rocks start to behave like a liquid rather than a solid. As temperature thresholds for hydrothermal (water-rock) activity, we assume that water ice melts above $T_{hydr} \equiv 273$ K, and hydrous rock phases fully decompose above $T_{decomp} \equiv 1223$ K.

Averaged over the planetesimal volume, Fig. 3A indicates that larger and earlier-formed planetesimals experience higher temperatures during their evolution, leading to a greater degree of internal processing. Compared to Reservoir II, Reservoir I planetesimals reach systematically higher temperatures and undergo body-wide metal-silicate segregation. Reservoir I bodies larger than about 30 km in radius also reach near-complete dehydration. Reservoir II planetesimals only experience limited degrees of heating and therefore less core formation and dehydration. The peak temperatures (Fig. 3) only reveal approximate trends in the simulations, as heating and compositional evolution are highly variable on time scales of 10^5 to 10^6 years (figs. S4 to S7).

To compare the simulations with the meteorite record, we explore the time-dependent evolution of the planetesimal populations in both reservoirs. We assume the planetesimal number distributions follow that of the streaming instability (*10*) in the interval $R_P \in [1, 300]$ km, and take into account the generation of new planetesimals in the simulation. We then evolve the birth planetesimal populations in Reservoir I and II and derive the fraction of each reservoir that falls within the threshold criteria, normalized to the final reservoir population after 5 Myr in the disk simulation.

Figure 4 compares the radiometric ages of metal-silicate separation in meteoritic measurements with the thermochemical evolution in our simulations. Reservoir I planetesimals undergo core formation between ≈0.4 and 3.0 Myr after CAI formation, with a peak between ≈0.6 and 1.3 Myr. On average, core formation via percolation occurs ≈0.4 Myr earlier than through rain-out in internal magma oceans. Multistage core formation scenarios, in which cores form by multiple stages of percolation and rain-out, are bracketed by these time intervals. Reservoir II planetesimals undergo core formation at later stages, starting at ≈1.1 Myr after CAI formation for percolation, and at ≈1.8 Myr after CAI formation for rain-out, and peaking between ≈1.8 and 3.1 and ≈2.3 and 4.0 Myr, respectively. The peak interval and the fraction of material

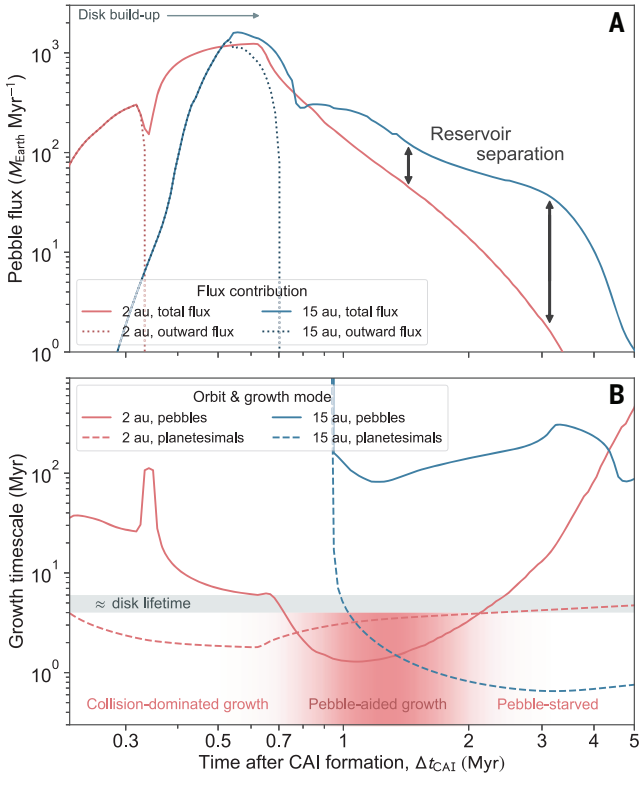


Fig. 2. Pebble flux and planetesimal growth time scales in the disk simulation. (**A**) Pebble flux at 2 (red) and 15 (blue) au over time. During the disk build-up stage, the pebble flux is dominated by outward-moving dust (dotted lines), after which the pebbles start drifting inward. Reservoir I and II in the simulation progressively diverge in pebble flux by more than one order of magnitude owing to accretion of Reservoir II (black arrows, labeled "Reservoir separation"). (**B**) Comparison of growth time scales for birth-sized planetesimals of 300-km radius by either pebble or planetesimal (collisional) accretion (*10*) within the approximate lifetime of the solar protoplanetary disk (gray horizontal band, labeled "≈ disk lifetime"). The red shaded area in the bottom indicates the time interval during which pebble accretion is more effective than collisional growth (labeled "Pebble-aided growth"; compare to fig. S11).

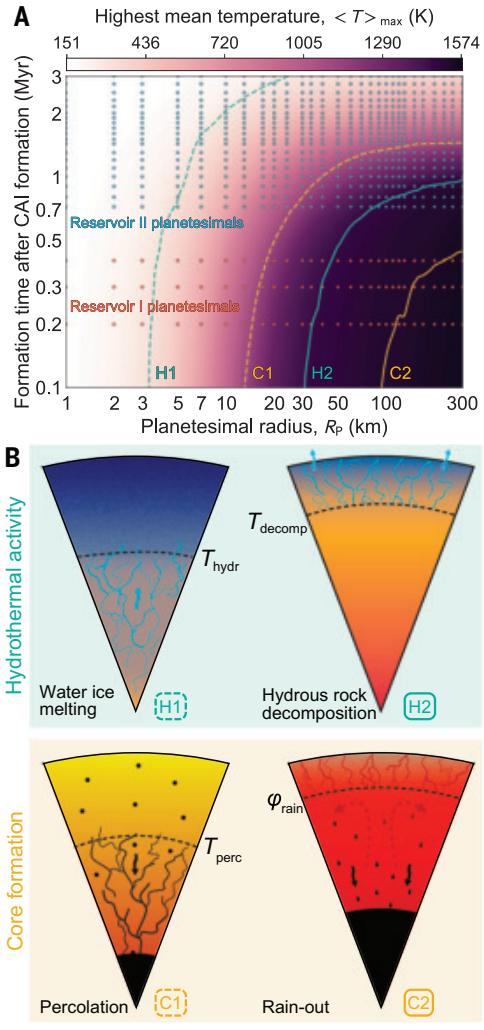


Fig. 3. Simulated thermochemical evolution of planetesimals. (**A**) Highest mean internal temperature (color bar), interpolated from simulated planetesimal radii and formation times (dots). Reservoir I and II in the disk model are indicated by red and blue dots, respectively (compare to Fig. 1). Green and yellow lines (labeled H1 to C2) indicate bodies for which 50 vol % (dashed lines) or 90 vol % (solid lines) of their interiors undergo hydrothermal activity (H1, H2) and iron core formation (C1, C2). (**B**) Schematic illustration of the scenarios indicated by the lines in (A) at peak heating and the threshold parameters [see text and (*10*)] to distinguish between regimes. H1, H2: Dark and light blue represent water ice and liquid water, respectively. Silicates are colored yellow to orange, indicating increasing temperature. Dashed black lines indicate the thresholds for water ice melting (T_{hydr}) and hydrous rock decomposition (T_{decomp}). The light blue arrows indicate hydrothermal circulation (H1) or degassing of volatiles (H2). C1, C2: Dark gray represents Fe,Ni–S metals. Silicates are colored yellow to red, indicating increasing temperature (yellow to orange) and progressive melting (red). Dashed black lines indicate the thresholds for metal percolation (T_{perc}) and rain-out (φ_{rain}). Black arrows indicate core formation by percolation (C1) or rain-out (C2). Dark red dashed arrows (C2) indicate convection in the internal magma ocean.

experiencing differentiation in Reservoir I is earlier and higher than in Reservoir II owing to the differing formation times and thus varying initial abundance of ^{26}Al. The timing of core formation in both simulated populations agrees with the meteorite data within the uncertainties of the individual ages (*8*, *20*, *29*). The spread in early-formed NCs is reproduced within the uncertainties, with the exception of the IAB meteorite group (*29*). The fractional volumes in both simulated reservoirs differ, with ≈2 to 10 vol % of Reservoir II undergoing core formation, whereas Reservoir I planetesimals differentiate to ≳90 vol %.

Figure 5 compares aqueous alteration ages from meteorites with our simulations. Simulated planetesimals in Reservoir I experience a brief phase of hydrothermal activity between ≈0.25 and 0.7 Myr after CAI formation and then dehydrate rapidly at ≈0.7 Myr. After that initial peak of water ice melting and hydrothermal activity, ≈10 vol % of rock contains hydrous silicate phases, and ≲1 vol % of water ice is retained. Reservoir II planetesimals, which form later with less ^{26}Al, experience protracted hydrothermal activity lasting for several million years with a peak at ≈5 Myr after CAI formation. At ≈3.5 Myr after CAI formation, more than 50 vol % of the Reservoir II planetesimals have undergone hydrothermal activity. At ≈5 Myr after CAI formation, the simulation reaches a steady state with ≈15 vol % water ice and ≈75 vol % hydrous rock. The peak for hydrothermal activity in Reservoir II reproduces the clustering of aqueous alteration ages in the CC meteorite record. The single available age for the NC population (from an ordinary chondrite, OC) does not coincide with the peak in the Reservoir I population, which we discuss below.

Compositional dichotomy between inner and outer Solar System

Our simulations indicate that the formation of spatially and temporally distinct planetesimal populations resulted in differing evolutionary pathways for planet formation in the inner and outer Solar System. The initial seed population of the inner Solar System planets formed during the infall stage, whereas the outer Solar System planetesimal population started to form later, at the beginning of the Class II disk stage. Figure 6 shows our interpretation of the isotopic evolution during early Solar System formation. Measurements of CAIs indicate that the earliest Solar System infall material was initially dominated by dust enriched in supernovae-derived isotopes, then transitioned to a more depleted isotopic composition (*21*, *30*, *31*). The distribution of isotopes evolved with time because of viscous disk spreading and increasing specific angular momentum of infalling matter. As a result, the outer disk became progressively dominated by early-infalling, enriched material, whereas the inner disk became dominated by late-infalling, depleted material (*21*, *30*, *31*). We suggest that the distinct NC and CC reservoirs measured in the meteorite record result from the combined influence of early planetesimal formation in the inner Solar System and the subsequent dust pile-up at the snow line. The latter effect suppressed the rapid inward drift of enriched material and prevented isotopic homogenization of the disk. It also led to the onset of planetesimal formation and planetary accretion in the outer Solar System, which further limited mixing by dust drift from the outer to the inner reservoir.

Figure 6 also illustrates our proposed timeline of compositional evolution and accretion, based on our simulations. The first planetesimal population formed as a result of the cold-finger effect. These bodies incorporated a substantial amount of ^{26}Al and thus dehydrated and differentiated rapidly. Initial growth in the inner Solar System proceeded from mutual collisions among planetesimals to an intermediate phase of pebble accretion. During the final phase before disk dissipation, the pebble flux in the inner Solar System diminished, and pebble growth may have stalled (see below). The outer Solar System population began to form later, during the Class II disk stage, and incorporated more total dust mass into its planetesimals. The initial phase of collisional growth was then succeeded by pebble accretion. This series of growth stages proceeded faster than in the inner Solar System because of the higher local pebble flux.

This chronology is consistent with meteoritic and astronomical evidence and makes several potentially testable predictions. The rapid decrease in the pebble flux in the inner Solar System (Reservoir I), to less than one Earth mass per million years, reduces the growth by pebble accretion during later disk stages. Parameter space exploration (*10*) and previous

work (*32*) indicate that under such conditions, pebble accretion onto inner Solar System planetary embryos is inefficient. This implies that the intermediate, pebble-driven accretion phase stalled, and subsequent growth of the terrestrial protoplanets was driven by collisional interactions during the final disk stages and afterwards. Such protracted accretion time scales for the terrestrial planets are supported by radiometric evidence for both proto-Earth and proto-Mars (*24*, *25*), and the absence of rocky planets larger than Earth (super-Earths) in the inner Solar System. The formation of super-Earth exoplanets has been attributed to a high pebble flux and rapid inward migration with increasing planet mass (*7*, *10*, *18*, *33*). The secular transition of an early, low pebble flux to a brief period of high pebble flux could amplify preceding mass differences between accreting protoplanets and thus explain why Earth is the inner Solar System body with the closest composition to that of CI-chondrites (*12*, *15*–*17*).

The secular transition between growth regimes, the internal evolution of planetesimals, and the temporal variation of the local pebble composition would lead to a heterogeneous delivery of water and other highly volatile compounds to the inner Solar System (*10*): The birth planetesimals in Reservoir I form an initial seed population of water-rich bodies that subsequently dehydrate from ^{26}Al heating and accrete by means of collisions. With progressing time and planetary growth, the protoplanets in the inner Solar System experience an influx of dry pebbles between ≈1 and 2 Myr after CAI formation. Forming such initially icy and subsequently dehydrating planetesimals in the inner Solar System is consistent with measurements of hydrogen content in enstatite (*34*), carbonaceous chondrite (*35*), eucrite and angrite (*36*) meteorites, Earth's deep mantle, and bulk measurements of the terrestrial planets and asteroid families (*19*, *27*). The timing of hydrothermal activity in Reservoir I planetesimals in our simulations, between ≈0.3 to 0.7 Myr after CAI formation, is offset from the Reservoir II peak at ≳4 Myr after CAI formation (Fig. 5). Most of the hydrous phases that are formed in the Reservoir I peak, however, are subsequently destroyed by high-temperature internal processing in our simulations, consistent with evidence from NC meteorites (*37*). Such intense (planetary) thermal processing in Reservoir I may thus lead to fractionation in mass-dependent isotopes (*38*), consistent with evidence from terrestrial planetary bodies (*39*, *40*). The disk simulation does not produce direct analogs of ordinary or enstatite chondrites from gravitational collapse, but the predicted growth stages (Figs. 2 and 6) would lead to recycling of collisionally disrupted planetesimals (*28*) and secondary layers accreted onto growing

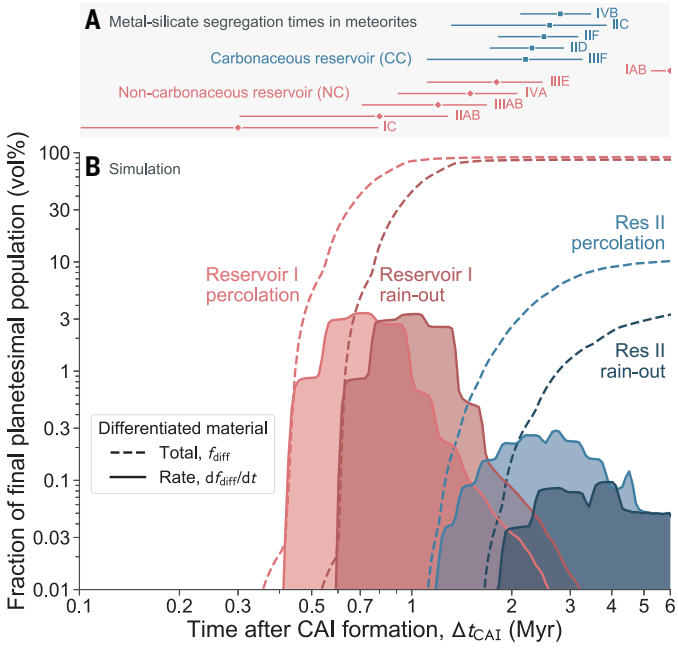


Fig. 4. Comparison of metal-silicate separation times in simulated planetesimal populations with the meteorite record. (**A**) Metal-silicate separation times in NC and CC meteorite classes (table S1). (**B**) Timing of core formation in Reservoir I (red) or Reservoir II (blue). Solid lines with colored shading represent the fraction of material undergoing metal-silicate separation [compare Fig. 2 and (*10*)]. Dashed lines show the total volumetric fraction for each scenario over time. Light and dark red or blue indicate percolation or rain-out, respectively.

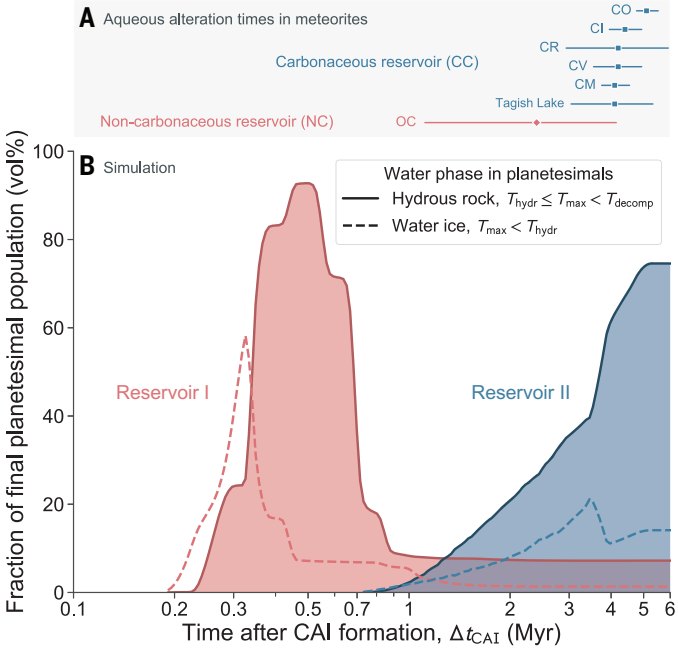


Fig. 5. Comparison of hydrothermal activity in simulated planetesimal populations with aqueous alteration times in the meteorite record. (**A**) Aqueous alteration times in NC and CC meteorite classes (table S1). (**B**) Fraction of planetary materials that retain primordial water ice (dashed lines) or undergo hydrothermal activity and retain water in hydrous silicate phases (solid lines with colored shading) in the simulations.

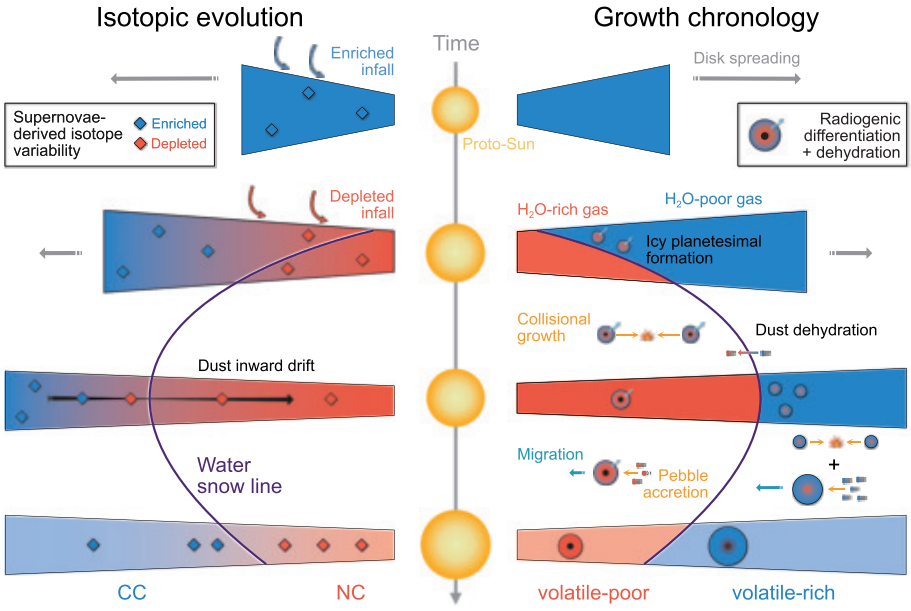


Fig. 6. Schematic illustration of our proposed chronology of early Solar System accretion. Nucleosynthetic isotope variability (left) across the disk due to varying composition of infall material is retained by the pile-up of inward-drifting dust grains at the snow line. The formation of two distinct planetesimal populations initiates divergent evolutionary pathways of inner and outer Solar System (right) due to the secular variation of local material composition, internal radiogenic heating, and dominant mode of planetary growth.

bodies (*41*), which could be potential origin locations.

The formation of Reservoir II planetesimals with initially higher water abundances than exist today is consistent with evidence for extensive iron oxidation and deuterium enrichment during aqueous alteration of CC meteorites (*26*). Simulated Reservoir II planetesimals undergo a prolonged million-year-long phase of hydrothermal activity with a peak at ≈4 to 5 Myr after CAI formation, which reproduces the aqueous alteration ages of CC meteorites (*27*). However, later-formed Reservoir II bodies remain almost unmodified by hydrothermal alteration (Fig. 3), so surviving bodies today should display varying internal processing, ranging from extensive aqueous alteration to fully pristine for small and late-formed bodies. Reservoir II planetesimals sample material originating from a wide range of orbital distances and form over an extended time interval. Therefore, compared to Reservoir I planetary materials, Reservoir II bodies may exhibit larger intrareservoir isotope variability. This is consistent with the nucleosynthetic isotope variability (*11*) and abundance of presolar interstellar materials (*42*) in carbonaceous chondrites, and the scatter in cometary D/H ratios (*27*).

The chronology of Solar System formation (Fig. 6) that we infer from our simulations links several characteristics found by geochemical laboratory analyses and astronomical observations. We interpret the chemical (volatile-poor versus volatile-rich) and isotopic (NC versus CC) dichotomy as causally linked by the build-up of distinct planetesimal populations in the inner and outer Solar System. Mixing by dust drift between them is limited to the earliest disk phases and declines with time owing to the progressive accretion of the outer Solar System planetary population. The temporal and spatial variation in the main volatile reservoirs and heterogeneous planet growth are intrinsically coupled to the disk structure, the redistribution and fractionation of volatile ices, the abundance of short-lived radionuclides, and the geophysical evolution of protoplanets during planetary formation.

REFERENCES AND NOTES

1. D. M. Segura-Cox *et al.*, *Nature* **586**, 228–231 (2020).
2. S. M. Andrews *et al.*, *Astrophys. J. Lett.* **869**, L41 (2018).
3. M. Ansdell *et al.*, *Astrophys. J.* **828**, 46 (2016).
4. J. R. Najita, S. J. Kenyon, *Mon. Not. R. Astron. Soc.* **445**, 3315–3329 (2014).
5. A. Johansen, M.-M. Mac Low, P. Lacerda, M. Bizzarro, *Sci. Adv.* **1**, e1500109 (2015).
6. M. Delbo', K. Walsh, B. Bolin, C. Avdellidou, A. Morbidelli, *Science* **357**, 1026–1029 (2017).
7. B. Liu, C. W. Ormel, A. Johansen, *Astron. Astrophys.* **624**, A114 (2019).
8. T. S. Kruijer *et al.*, *Science* **344**, 1150–1154 (2014).
9. J. N. Connelly *et al.*, *Science* **338**, 651–655 (2012).
10. Materials and methods and supplementary text are available as supplementary materials.
11. I. Leya, M. Schönbächler, U. Wiechert, U. Krähenbühl, A. N. Halliday, *Earth Planet. Sci. Lett.* **266**, 233–244 (2008).
12. A. Trinquier *et al.*, *Science* **324**, 374–376 (2009).
13. P. H. Warren, *Earth Planet. Sci. Lett.* **311**, 93–100 (2011).
14. T. Lichtenberg *et al.*, *Nat. Astron.* **3**, 307–313 (2019).
15. N. Braukmüller, F. Wombacher, C. Funk, C. Münker, *Nat. Geosci.* **12**, 564–568 (2019).
16. M. Schönbächler, R. W. Carlson, M. F. Horan, T. D. Mock, E. H. Hauri, *Science* **328**, 884–887 (2010).
17. M. Schiller, M. Bizzarro, V. A. Fernandes, *Nature* **555**, 507–510 (2018).
18. S. N. Raymond, A. Izidoro, A. Morbidelli, in *Planetary Astrobiology*, V. S. Meadows, G. N. Arnay, B. E. Schmidt, D. J. Des Marais, Eds. (Univ. of Arizona Press, 2020), pp. 287–324.
19. A. Peslier, M. Schönbächler, H. Busemann, S.-I. Karato, *Space Sci. Rev.* **212**, 743–810 (2018).
20. T. S. Kruijer, C. Burkhardt, G. Budde, T. Kleine, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6712–6716 (2017).
21. S. J. Desch, A. Kalyaan, C. M. O'D. Alexander, *Astrophys. J. Suppl. Ser.* **238**, 11 (2018).
22. Y. Alibert *et al.*, *Nat. Astron.* **2**, 873–877 (2018).
23. J. Drążkowska, C. P. Dullemond, *Astron. Astrophys.* **614**, A62 (2018).
24. J. F. Rudge, T. Kleine, B. Bourdon, *Nat. Geosci.* **3**, 439–443 (2010).
25. N. Dauphas, A. Pourmand, *Nature* **473**, 489–492 (2011).
26. S. Sutton *et al.*, *Geochim. Cosmochim. Acta* **211**, 115–132 (2017).
27. C. M. O'D Alexander, K. D. McKeegan, K. Altwegg, *Space Sci. Rev.* **214**, 36 (2018).
28. T. Lichtenberg *et al.*, *Icarus* **302**, 27–43 (2018).
29. A. C. Hunt *et al.*, *Earth Planet. Sci. Lett.* **482**, 490–500 (2018).
30. J. A. M. Nanne, F. Nimmo, J. N. Cuzzi, T. Kleine, *Earth Planet. Sci. Lett.* **511**, 44–54 (2019).
31. E. Jacquet, F. C. Pignatale, M. Chaussidon, S. Charnoz, *Astrophys. J.* **884**, 32 (2019).
32. H. F. Levison, K. A. Kretke, K. J. Walsh, W. F. Bottke, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14180–14185 (2015).
33. B. Bitsch, S. N. Raymond, A. Izidoro, *Astron. Astrophys.* **624**, A109 (2019).
34. L. Piani *et al.*, *Science* **369**, 1110–1113 (2020).
35. L. Piani, H. Yurimoto, L. Remusat, *Nat. Astron.* **2**, 317–323 (2018).
36. A. R. Sarafian *et al.*, *Philos. Trans. A Math. Phys. Eng. Sci.* **375**, 20160209 (2017).
37. J. A. Lewis, R. H. Jones, *Meteorit. Planet. Sci.* **51**, 1886–1913 (2016).
38. M. R. Benedikt *et al.*, *Icarus* **347**, 113772 (2020).
39. R. C. Hin *et al.*, *Nature* **549**, 511–515 (2017).
40. C. A. Norris, B. J. Wood, *Nature* **549**, 507–510 (2017).
41. C. Maurel *et al.*, *Sci. Adv.* **6**, eaba1303 (2020).
42. J. Davidson *et al.*, *Geochim. Cosmochim. Acta* **139**, 248–266 (2014).

ACKNOWLEDGMENTS

We thank C. P. Dullemond, A. C. Hunt, I. Pascucci, S. Ida, J. Wade, S.-J. Paardekooper, T. Birnstiel, S. M. Stammler, J. J. Barnes, A. Morbidelli, W. Kley, and members of the ERC EXOCONDENSE project at Oxford for discussions; B. Liu, J. F. J. Bryson, M. Ek, R. D. Alexander, and S. Charnoz for comments on earlier draft versions; T. V. Gerya for usage of the I2ELVIS code family; and C. P. Dullemond for usage of the disk evolution code. **Funding:** T.L. received funding from the Simons Foundation (SCOL award no. 611576) and the Swiss National Science Foundation (grant no. P2EZP2-178621). J.D. received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant no. 714769). T.O.H. was supported by the University of Zurich Forschungskredit Postdoc. Parts of this work were carried out within the framework of the National Centre for Competence in Research PlanetS (grant no. 51NF40-141881) supported by the Swiss National Science Foundation. **Author contributions:** Conceptualization, T.L., J.D., M.S., G.J.G.; Methodology, T.L., J.D., T.O.H., G.J.G.; Software, T.L., J.D., T.O.H., G.J.G.; Validation, all authors; Formal Analysis, T.L., J.D., T.O.H.; Investigation, T.L., J.D., M.S.; Resources, T.L., J.D., M.S.; Writing, T.L., J.D., T.O.H., M.S.; Review & Editing, all authors; Visualization, T.L. **Competing interests:** We declare no competing interests. **Data and materials availability:** The simulation codes, output data, and plotting scripts are available at https://osf.io/e2kfv.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/365/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Table S1
References (*43–204*)

2 March 2020; accepted 10 December 2020
10.1126/science.abb3091

REPORTS

ULTRAFAST MICROSCOPY

Ultrafast nanoimaging of the order parameter in a structural phase transition

Thomas Danz[1], Till Domröse[1], Claus Ropers[1,2]*

Understanding microscopic processes in materials and devices that can be switched by light requires experimental access to dynamics on nanometer length and femtosecond time scales. Here, we introduce ultrafast dark-field electron microscopy to map the order parameter across a structural phase transition. We use ultrashort laser pulses to locally excite a 1*T*-TaS_2 (1*T*-polytype of tantalum disulfide) thin film and image the transient state of the specimen by ultrashort electron pulses. A tailored dark-field aperture array allows us to track the evolution of charge-density wave domains in the material with simultaneous femtosecond temporal and 5-nanometer spatial resolution, elucidating relaxation pathways and domain wall dynamics. The distinctive benefits of selective contrast enhancement will inspire future beam-shaping technology in ultrafast transmission electron microscopy.

Optical control over physical and chemical properties of materials is a recurring motif in research areas from femtochemistry to ultrafast condensed matter physics. This broad interest is based on the impact of optical control strategies in current and future technology, such as data storage devices (*1*, *2*), neuromorphic computing (*3*), photonic circuits (*4*), and energy conversion (*5*). In parallel, fundamental scientific discoveries reveal specific light-induced features and correlation effects, involving coupled order parameters (*6*), metastable or hidden states (*7*–*10*), superconductivity (*11*), changes in topology (*12*), and metal-insulator transitions (*13*, *14*).

Typically, microscopic correlations evolve on femto- to picosecond time scales, a regime that is accessible by ultrafast measurement methodology. This approach enables spatially averaged probing of properties such as electronic gaps (*15*), optical conductivity (*16*), magnetization (*17*), and instabilities against periodic lattice and charge-density modulations (*18*, *19*). Functionality of devices, however, usually arises from nanoscale structuring or from interfacing different materials. This calls for experimental approaches that capture the dynamics in terms of a spatially dependent order parameter.

Recent experiments accomplished time-resolved mapping of the local free-carrier response in a correlated metal-insulator phase transition using near-field probing (*20*–*22*). However, direct sensitivity to the spatiotemporal structural modifications governing the transition on the atomic scale remains a challenge. Ultrafast transmission electron microscopy (UTEM) has proven a valuable tool to study lattice dynamics (*23*–*29*) but does not provide direct contrast of the structural order parameter.

In this work, we demonstrate real-space imaging of charge-density wave (CDW) phases with nanometer spatial and femtosecond temporal resolution. By introducing a tailored ultrafast dark-field (DF) electron microscopy scheme, we obtain maps of the order parameter in a prototypical CDW system. We observe the formation, stabilization, and relaxation of CDW domains after optical excitation and identify drastic differences between the emerging domain patterns for continuous-wave and pulsed illumination. Using time-dependent Ginzburg-Landau simulations to corroborate our findings, we discuss the nonequilibrium evolution of the order parameter near domain walls.

Specifically, we investigate the transition metal dichalcogenide 1*T*-TaS_2 (1*T*-polytype of tantalum disulfide). Owing to its low-dimensional character and strong correlation effects, this system is a promising target for nontrivial order parameter dynamics. It displays several CDW phases with both Mott and Peierls contributions and particularly complex orbital textures (*30*, *31*). The CDW states and the transitions between them have been the subject of various ultrafast spectroscopy and electron diffraction experiments in the past (*18*, *32*–*37*).

Our study focuses on the room- and high-temperature phases of the material (*30*, *38*–*40*). In all experiments, the sample is initially held at room temperature. Above a temperature of $T^* = 353$ K, the material exhibits an incommensurate superstructure aligned with the underlying hexagonal lattice (IC CDW phase; "IC phase" hereafter). Below the phase transition temperature, the CDW transforms into a nearly commensurate superstructure with an angle of ~12° between the modulation wave vectors and the lattice directions, effectively reducing the symmetry of the material (NC CDW phase; "NC phase" hereafter). Each of the CDW phases is accompanied by a pronounced periodic lattice distortion (PLD). In electron diffraction images, the PLD of the NC phase is evident from satellite diffraction peaks arranged around the bright structural

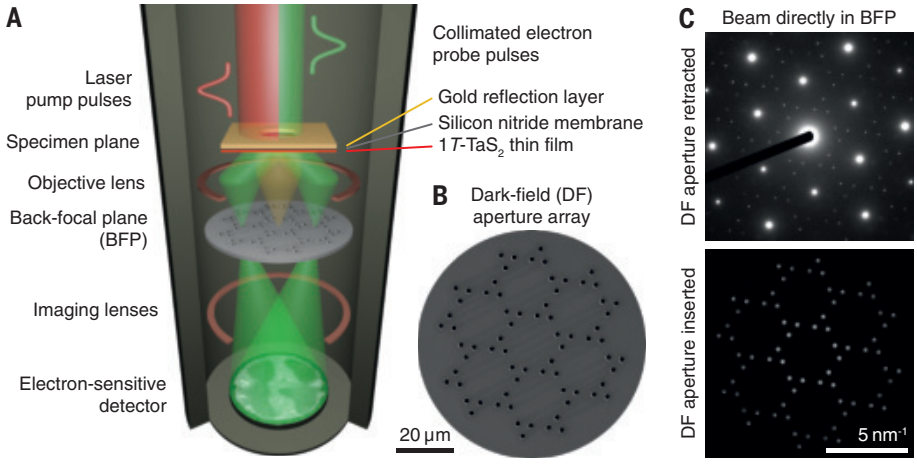


Fig. 1. Dark-field imaging in the ultrafast transmission electron microscope. (**A**) Sectional drawing of the experimental setup. Electron (green) and optical (red) pulses are incident close to perpendicular on the specimen. (**B**) Scanning electron micrograph of the DF aperture array placed in the back-focal plane (BFP) in (A). (**C**) Electron diffractograms of the 1*T*-TaS_2 thin film at room temperature, with the DF aperture array shown in (B) retracted from (top) and inserted into (bottom) the electron beam path. Only superstructure reflections are transmitted through the aperture array.

[1]4th Physical Institute – Solids and Nanostructures, University of Göttingen, 37077 Göttingen, Germany. [2]Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany.
*Corresponding author. Email: cropers@gwdg.de

reflections (Fig. 1C, top). Owing to a threefold out-of-plane stacking periodicity in the NC phase, mainly second-order satellites are visible (*40*).

In real-space images, NC CDW contrast is obtained by means of DF imaging—that is, by inserting a DF aperture into the back-focal plane of the microscope's objective lens. Early work has indicated that DF imaging is, in principle, suitable to map the spatial distribution of PLDs (*38*). However, standard circular DF apertures are too large for this purpose (*38*), and a single satellite diffraction peak is too weak because of brightness limitations of pulsed electron beams (<1 % intensity compared with structural reflections) (*41*). Thus, the present work uses focused ion beam etching to manufacture a tailored array of 72 small apertures, selecting the brightest second-order satellite reflections and blocking all other diffraction peaks (Fig. 1B) (*42*).

When the DF aperture array is inserted into the column of the electron microscope (Fig. 1C, bottom), only electrons scattered into NC satellite diffraction peaks can reach the detector and contribute to the image contrast, whereas electrons elastically scattered into main lattice peaks and IC diffraction peaks are blocked (fig. S1). Accordingly, the DF images directly reflect the local order parameter ϕ of the NC CDW with a spatial resolution below 5 nm (*42*). Kinematical diffraction simulations based on the NC structure (*40*) indicate that the dependency between local image intensity and order parameter is close to quadratic with the specific set of reflections considered (*42*). Inelastic scattering contributions account for some background in the image, even in the IC phase (*39*).

This approach now allows us to directly image CDW dynamics in the material (movie S1), with a contrast unattainable in conventional bright-field imaging (movie S2). In our time-resolved experiments, we use femtosecond laser pulses [800-nm center wavelength, 80-fs duration, 420-kHz repetition rate] to excite a freestanding 70-nm 1*T*-TaS_2 film and probe the transient state of the specimen using ultrashort electron pulses (120-keV electron energy, 530-fs duration). After filtering with the DF aperture array, a spatial image is formed in the detector plane (Fig. 1A; see fig. S2 for more details). Spatially inhomogeneous excitation is required to trigger spatiotemporal dynamics in this experiment. Therefore, we use the circular gold aperture that supports the film as a means to structure the excitation profile by interference between the main beam and edge reflections (see supplementary text for further information). Additionally, the gold layer acts as a heat bath and prevents thermal load on the specimen outside the region of interest (*42*). This very local excitation and effective thermal coupling ensures reversibility of the observed dynamics at a previously unattainable repetition rate for a structural phase transition.

Ultrafast DF images are shown in Fig. 2A as a function of the optical pump–to–electron probe delay Δt (*42*). Initially, the image within the circular aperture is flat with minor diffraction features and spatial variations of the NC phase PLD amplitude at room temperature. Directly after the temporal overlap of electron and laser pulses ("time zero"), an inhomogeneous suppression of image intensity is evident, with its shape governed by the excitation profile (Fig. 2E). Within 2 ps, well-defined domains have emerged, and the boundaries separating bright from dark regions have become visibly sharper (Fig. 2D). On a 100-ps time scale, a slight growth of the darkened regions and a further contrast sharpening is observed. Finally, these domains diminish in area, and a homogeneous NC contrast is reestablished after a few nanoseconds.

The image series contains considerably more information than spatially averaged diffraction data alone. Specifically, we can individually analyze the PLD dynamics in weakly and strongly pumped regions, which show either a transient quench and recovery or a persistent phase transformation with a full suppression within the temporal resolution (green and orange curves in Fig. 2C, bottom) (*18*). The practically binary contrast is a clear sign of a completed phase separation into distinct NC and IC regions no later than a few picoseconds after the pump. This allows us to carry out image segmentation to identify the temporal evolution of the respective area fractions (Fig. 2B and fig. S4). In particular, this analysis shows that the delayed additional suppression of the spatially averaged contrast (black curve in Fig. 2C), starting after a few picoseconds and lasting up to 1 ns, is caused by the growth of the IC domains (see the moderate increase of the IC fraction in Fig. 2C, top).

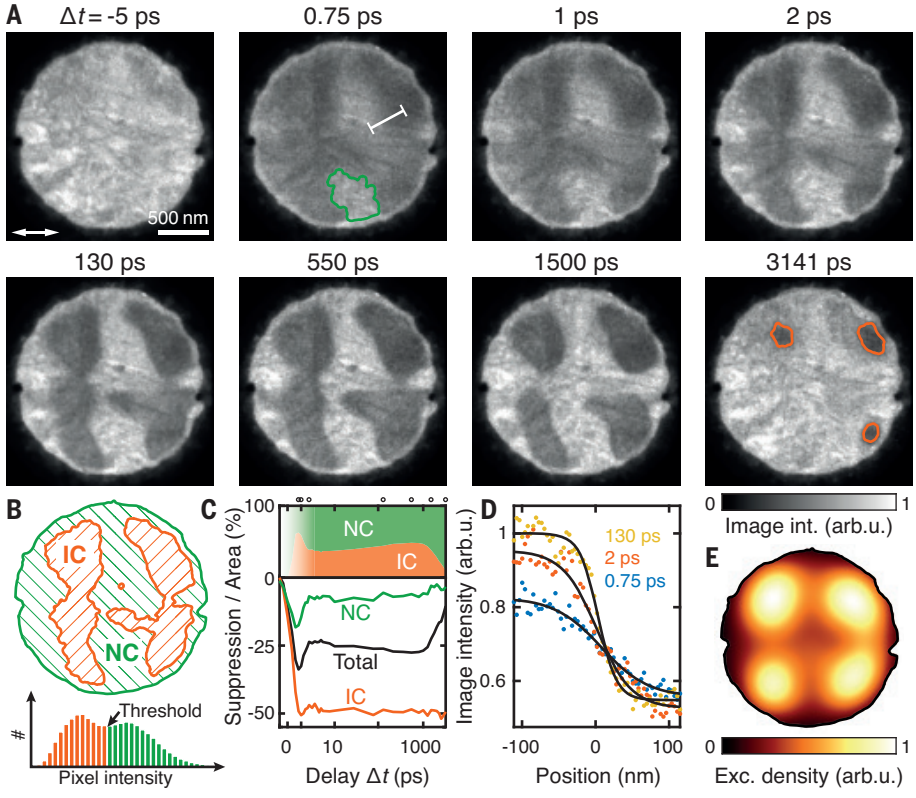


Fig. 2. Ultrafast dark-field domain imaging of charge-density wave dynamics. (**A**) Ultrafast DF micrographs of transient domain configurations in the 1*T*-TaS_2 film obtained in the laser pump–electron probe scheme (2.6-mJ/cm^2 pump fluence, linear pump polarization indicated by white arrow). Pump-probe delay steps were chosen so as to capture all major stages of the dynamics [see black circles above (C)]. (**B**) (Top) Image segmentation at 130-ps delay time. (Bottom) The segmentation threshold is determined from the intensity histogram of the full image series within the circular aperture. (**C**) (Top) Area fractions of NC and IC regions after completed phase separation, as determined from the segmented images. (Bottom) Average intensity of the image series within the entire aperture (black curve), and average intensity in weakly and strongly pumped regions [green and orange curves, respectively; evaluated regions are indicated in (A) with corresponding colors]. (**D**) Exemplary profiles of NC/IC phase boundaries taken on the white line indicated in (A). (**E**) Spatial profile of the excitation density giving rise to the initial suppression pattern [see (*42*) and fig. S3].

The CDW phase pattern in the time-resolved experiment is a direct consequence of the spatiotemporal nonequilibrium and strongly deviates from the situation observed under constant heating of the structure using a 532-nm continuous-wave laser (Fig. 3A and movie S3): Upon increasing the laser heating intensity, an IC domain nucleates close to the center of the thin film and grows until it almost fills the circular aperture. Decreasing the laser intensity results in a reduction of the switched area.

The steady-state phase pattern arises from the dynamical equilibrium between deposited and dissipated energy. Owing to efficient equilibration within the film compared with the thermal coupling to the support and heat bath, the optical excitation profile is not evident in this experiment. This provides us with a detailed picture of the sample's thermal state. We use a heat-transfer simulation to predict the temperature distribution for each laser intensity (Fig. 3, B to D) and quantify the thermal boundary resistance between the film and the substrate. The value obtained falls well in line with those of related interfaces (*42*).

A slight hysteresis of <4 K is observed in the heating cycle (orange dots in Fig. 3B), a signature of the first-order character of the phase transition (*38*). On microscopic length scales, we identify two different mechanisms affecting the hysteretic behavior: The up-sweep reveals a preferential formation of NC/IC phase boundaries along specific crystallographic directions (indicated by orange arrows in Fig. 3A;

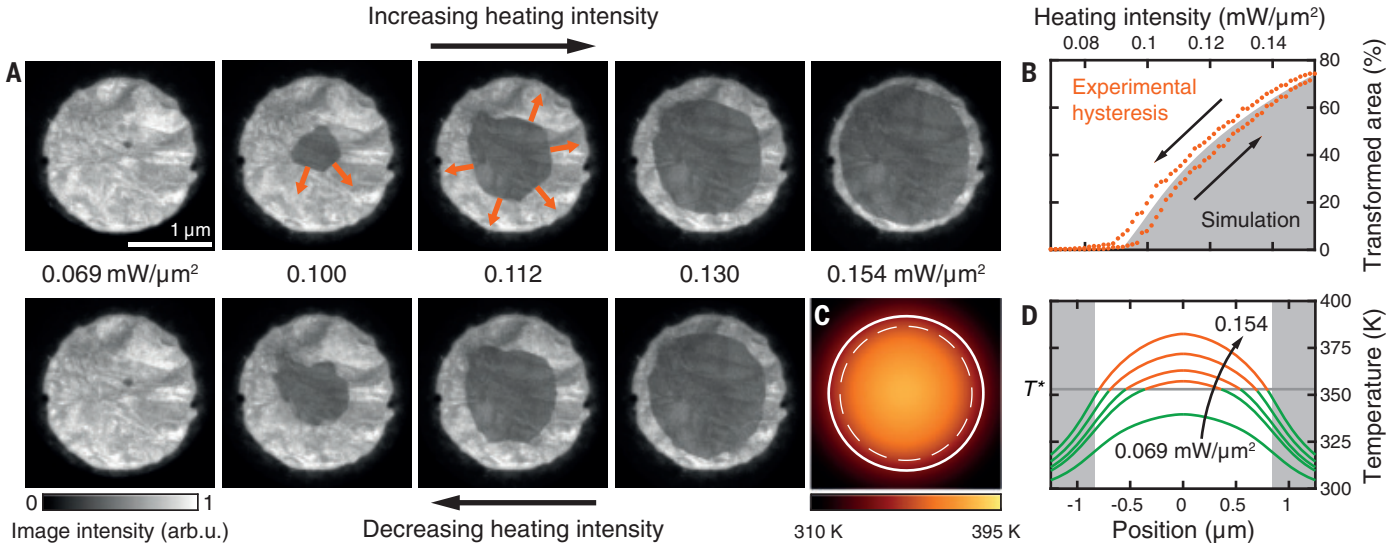


Fig. 3. Dark-field domain imaging with continuous-wave laser excitation. (**A**) DF micrographs of laser intensity up-sweep and down-sweep. Orange arrows indicate phase boundaries oriented along preferential lattice directions. (**B**) Hysteretic behavior as extracted from the micrographs (orange dots). The characteristic intensity dependency is reproduced by heat-transfer simulations (gray). (**C**) Temperature distribution simulated for a laser heating intensity of 0.154(9) mW/µm^2 for comparison with the experimental result (solid line, edge of circular aperture; dashed line, edge of transformed area). (**D**) Radial profiles of the simulated temperature distribution in the film for the experimental laser intensities presented in (A). The green and orange segments of the curves indicate temperatures below and above, respectively, the phase transition temperature T^*. Shaded regions are outside the aperture. See (*42*) for simulation details.

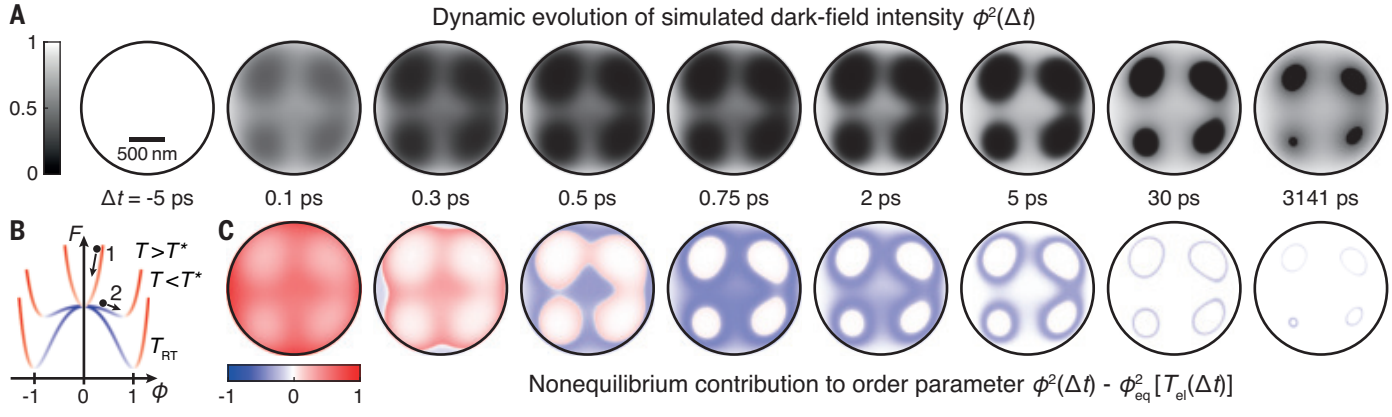


Fig. 4. Time-resolved Ginzburg-Landau simulations of optically induced order parameter dynamics. (**A**) Series of simulated DF micrographs (the squared order parameter ϕ^2 is projected along the specimen thickness to mimic DF image contrast). (**B**) Sketch of free-energy landscapes for different electron temperatures [color scale as indicated in (C)]. *F*, free-energy functional. (**C**) Deviation of the squared order parameter from its local equilibrium value ϕ^2_{eq} (red and blue regions denote order parameters that are larger and smaller, respectively, than at the free-energy minimum). See (*42*) for simulation details.

see supplementary text for details). On the other hand, local pinning sites visibly affect the down-sweep, owing to some degree of structural inhomogeneity (see fig. S5, A to C).

The appearance of these effects only under continuous excitation is a result of very different energy and time scales governing the dynamics. The steady-state experiment is rather sensitive to local defects and structural anisotropy, as there is sufficient time for small spatial variations in the CDW free-energy landscape to determine the phase pattern. On the other hand, the time-resolved experiment is characterized by electron-lattice nonequilibrium involving strong thermal gradients.

To obtain a better understanding of the spatiotemporal CDW evolution and energy redistribution, we simulate the three-dimensional dynamics of the order parameter as well as the electron and lattice temperatures (*42*). Employing a Ginzburg-Landau model for a nonconserved order parameter, we capture the symmetry-breaking nature of the first-order NC/IC phase transition [see (*43*), model A]. As in the experiments, we do not separately analyze the individual wave vector components of the CDW. Accordingly, we consider the spatially varying PLD amplitude as a single, real-valued order parameter ϕ with a generic free-energy functional for a first-order phase transition, parametrized by the local electron temperature T_{el}. The evolution of the electron and lattice temperatures is treated with coupled heat diffusion equations using the thermal parameters extracted from the steady-state experiment (*42*).

For values of T_{el} below the phase transition temperature T^*, the free-energy landscape exhibits a double-well shape (Fig. 4B). Initially, the simulation volume is in thermodynamical equilibrium with $\phi = 1$, corresponding to one of the potential minima at room temperature. The optical excitation at time zero is modeled as a jump in electronic temperature, with the spatial distribution given by the excitation profile (Fig. 2E). In this highly nonequilibrium state with $T_{el} > T^*$ in large areas of the sample, the reshaped potential is close to parabolic, with a single IC CDW minimum at $\phi = 0$. The resulting motion of the order parameter toward the high-symmetry state causes a global order parameter suppression, followed by a rapid electron-lattice energy transfer on a time scale of 350 fs (*18*). The simulation in Fig. 4A and movie S4 closely reproduces the transient suppression, phase separation, and cooling observed in the experimental dataset (see Fig. 2A).

Local trajectories of the order parameter can be classified according to the magnitude of its nonequilibrium contribution (Fig. 4C and movie S5). In regions with sufficient excitation density, the order parameter relaxes into the IC potential minimum within the electron-lattice coupling time (point 1 in Fig. 4B). The spatiotemporal evolution of the corresponding four white patches adiabatically follows the reshaping of the potential on heat diffusion time scales. On the other hand, and as a consequence of the nearly flat potential just below the phase transition temperature, more weakly excited regions undergo a transient quench of the NC phase, relaxing on a few-picosecond time scale (point 2) (see also fig. S6).

Close to the phase boundaries, the suppression of NC charge-ordering lasts until 30 ps after excitation (blue contours forming around the white patches after 2 ps in Fig. 4C), and a minor slowdown of the IC formation is observed (light red regions at 0.5 ps). These processes are a direct consequence of the delicate interplay between fast electron-lattice coupling and order parameter dynamics in a reshaped free-energy landscape, and they ultimately result in the sharpening of domain walls observed in simulation and experiment (fig. S7).

Our study illustrates how contrast enhancements by postspecimen (*44*–*47*) or even prespecimen beam shaping (*48*) allow for sensitivity to further degrees of freedom in complex materials, paving the way for other types of investigation. These possibilities include imaging of transient phonon populations (*49*) and, particularly, visualizing amplitude and phase modes in CDW materials (*41*). Furthermore, spatially resolved investigations of topological defects in the charge-ordering of these materials may lead to a deeper understanding of phase formation kinetics (*19*, *33*–*35*). Lastly, careful design of the beam-shaping parameters may facilitate multiphase imaging or the simultaneous mapping of chiral mirror domains (*36*). Building on these opportunities, ultrafast transmission electron microscopy will offer new perspectives for observing nanoscale dynamics in correlated materials.

REFERENCES AND NOTES

1. D. Loke *et al.*, *Science* **336**, 1566–1569 (2012).
2. M. Salinga *et al.*, *Nat. Commun.* **4**, 2371 (2013).
3. J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, W. H. P. Pernice, *Nature* **569**, 208–214 (2019).
4. H. Yang *et al.*, *Sci. Adv.* **4**, eaar7954 (2018).
5. K. Sivula, R. van de Krol, *Nat. Rev. Mater.* **1**, 15010 (2016).
6. N. A. Spaldin, M. Fiebig, *Science* **309**, 391–392 (2005).
7. L. Stojchevska *et al.*, *Science* **344**, 177–180 (2014).
8. M. Yoshida, R. Suzuki, Y. Zhang, M. Nakano, Y. Iwasa, *Sci. Adv.* **1**, e1500606 (2015).
9. D. Svetin, I. Vaskivskyi, S. Brazovskii, D. Mihailovic, *Sci. Rep.* **7**, 46048 (2017).
10. A. Kogar *et al.*, *Nat. Phys.* **16**, 159–163 (2020).
11. D. Fausti *et al.*, *Science* **331**, 189–191 (2011).
12. E. J. Sie *et al.*, *Nature* **565**, 61–66 (2019).
13. V. R. Morrison *et al.*, *Science* **346**, 445–448 (2014).
14. S. Wall *et al.*, *Science* **362**, 572–576 (2018).
15. T. Rohwer *et al.*, *Nature* **471**, 490–493 (2011).
16. S. M. Falke *et al.*, *Science* **344**, 1001–1005 (2014).
17. C. Stamm *et al.*, *Nat. Mater.* **6**, 740–743 (2007).
18. M. Eichberger *et al.*, *Nature* **468**, 799–802 (2010).
19. A. Zong *et al.*, *Nat. Phys.* **15**, 27–31 (2019).
20. M. A. Huber *et al.*, *Nano Lett.* **16**, 1421–1427 (2016).
21. S. A. Dönges *et al.*, *Nano Lett.* **16**, 3029–3035 (2016).
22. X. Fu *et al.*, *Nat. Commun.* **11**, 5770 (2020).
23. M. S. Grinolds, V. A. Lobastov, J. Weissenrieder, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18427–18431 (2006).
24. P. Baum, D.-S. Yang, A. H. Zewail, *Science* **318**, 788–792 (2007).
25. R. M. van der Veen, O.-H. Kwon, A. Tissot, A. Hauser, A. H. Zewail, *Nat. Chem.* **5**, 395–402 (2013).
26. L. Piazza *et al.*, *Struct. Dyn.* **1**, 014501 (2014).
27. A. Feist *et al.*, *Ultramicroscopy* **176**, 63–73 (2017).
28. A. Feist, N. Rubiano da Silva, W. Liang, C. Ropers, S. Schäfer, *Struct. Dyn.* **5**, 014302 (2018).
29. Y. Zhang, D. J. Flannigan, *Nano Lett.* **19**, 8216–8224 (2019).
30. K. Rossnagel, *J. Phys. Condens. Matter* **23**, 213001 (2011).
31. T. Ritschel *et al.*, *Nat. Phys.* **11**, 328–331 (2015).
32. C. Sohrt, A. Stange, M. Bauer, K. Rossnagel, *Faraday Discuss.* **171**, 243–257 (2014).
33. K. Haupt *et al.*, *Phys. Rev. Lett.* **116**, 016402 (2016).
34. C. Laulhé *et al.*, *Phys. Rev. Lett.* **118**, 247401 (2017).
35. S. Vogelgesang *et al.*, *Nat. Phys.* **14**, 184–190 (2018).
36. A. Zong *et al.*, *Sci. Adv.* **4**, eaau5501 (2018).
37. G. Storeck *et al.*, *Struct. Dyn.* **7**, 034304 (2020).
38. J. A. Wilson, F. J. Di Salvo, S. Mahajan, *Adv. Phys.* **24**, 117–201 (1975).
39. C. B. Scruby, P. M. Williams, G. S. Parry, *Philos. Mag.* **31**, 255–274 (1975).
40. A. Spijkerman, J. L. de Boer, A. Meetsma, G. A. Wiegers, S. van Smaalen, *Phys. Rev. B* **56**, 13757–13767 (1997).
41. A. W. Overhauser, *Phys. Rev. B* **3**, 3173–3182 (1971).
42. Materials and methods are available as supplementary materials.
43. P. C. Hohenberg, B. I. Halperin, *Rev. Mod. Phys.* **49**, 435–479 (1977).
44. R. Danev, K. Nagayama, *Methods Enzymol.* **481**, 343–369 (2010).
45. J. Verbeeck, H. Tian, P. Schattschneider, *Nature* **467**, 301–304 (2010).
46. V. Grillo *et al.*, *Nat. Commun.* **8**, 15536 (2017).
47. O. Schwartz *et al.*, *Nat. Methods* **16**, 1016–1020 (2019).
48. C. Ophus *et al.*, *Nat. Commun.* **7**, 10719 (2016).
49. M. J. Stern *et al.*, *Phys. Rev. B* **97**, 165416 (2018).
50. T. Danz, T. Domröse, C. Ropers, Data for: Ultrafast nanoimaging of the order parameter in a structural phase transition. GRO.data (2021); doi:10.25625/6GFN9K.

ACKNOWLEDGMENTS

We thank M. Sivis for technical support in focused ion beam milling of specimen and DF aperture array, as well as K. Rossnagel (University of Kiel) for supplying high-quality 1*T*-TaS_2 crystals. We acknowledge assistance from the Göttingen UTEM Team, especially T. Harvey, who designed the excitation beam path. Furthermore, we are grateful to C. Wichmann for providing the in-plane rotation specimen holder. **Funding:** This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the Collaborative Research Center "Atomic scale control of energy conversion" (217133147/SFB 1073, project A05) and via resources from the Gottfried Wilhelm Leibniz Prize (RO 3936/4-1). Th.D. gratefully acknowledges a scholarship by the German Academic Scholarship Foundation. **Author contributions:** Th.D. and Ti.D. conducted the experiments and analyzed the data. Th.D. designed and manufactured the DF aperture array. Th.D. and Ti.D. prepared the specimen and implemented the time-resolved simulations. Th.D. carried out the steady-state simulations. C.R. conceived and directed the study. All authors discussed the results and their interpretation and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data in the manuscript are available from GRO.data (*50*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/371/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Table S1
References (*51*–*71*)
Movies S1 to S5

11 June 2020; accepted 11 December 2020
10.1126/science.abd2774

SURFACE CHEMISTRY

Unraveling CO adsorption on model single-atom catalysts

Jan Hulva[1]*, Matthias Meier[1,2]*, Roland Bliem[1]†, Zdenek Jakub[1], Florian Kraushofer[1], Michael Schmid[1], Ulrike Diebold[1], Cesare Franchini[2,3], Gareth S. Parkinson[1]‡

Understanding how the local environment of a "single-atom" catalyst affects stability and reactivity remains a challenge. We present an in-depth study of copper$_1$, silver$_1$, gold$_1$, nickel$_1$, palladium$_1$, platinum$_1$, rhodium$_1$, and iridium$_1$ species on Fe_3O_4(001), a model support in which all metals occupy the same twofold-coordinated adsorption site upon deposition at room temperature. Surface science techniques revealed that CO adsorption strength at single metal sites differs from the respective metal surfaces and supported clusters. Charge transfer into the support modifies the d-states of the metal atom and the strength of the metal–CO bond. These effects could strengthen the bond (as for Ag_1–CO) or weaken it (as for Ni_1–CO), but CO-induced structural distortions reduce adsorption energies from those expected on the basis of electronic structure alone. The extent of the relaxations depends on the local geometry and could be predicted by analogy to coordination chemistry.

The reactivity of oxide-supported metal nanoparticle catalysts is traditionally understood by using the d-band model, which was developed for extended metal surfaces (*1*). Such a picture cannot be applied to so-called "single-atom" catalysts (*2*–*8*) because the isolated metal atoms are stabilized by chemical bonds to the support and are often charged. In this regard, single-atom catalysis systems resemble coordination complexes, and there is much excitement at the prospect that single-atom catalysts can be used to "heterogenize" problematic reactions currently performed in solution (*9*–*11*). Although there are similarities, there are also practical differences. Homogeneous catalysts are designed for purpose on the basis of a fundamental understanding of the structure-function relationship, and the ligands play an important role in activating reactants and stabilizing intermediates. Complexes containing O^{2-} ligands are rare (*5*), and the binding environment of the metal adatom on the metal oxide is difficult to ascertain and control. Because robust, inexpensive metal oxides are set to continue as the support of choice in single-atom catalysis, it is vital to learn how the coordination of the metal site on an oxide surface affects its adsorption properties and ultimately catalytic activity.

Most studies of oxide-supported single-atom catalysts feature transmission electron microscopy images (*12*) of powder samples showing that the metal adatoms align with the cationic sublattice. These structures are in agreement with x-ray absorption near-edge structure spectra, which typically suggest coordination to oxygen. Density functional theory (DFT) calculations (*13*–*16*) also predict such sites to be most stable (albeit on simplified models of the support surface). Often, a positive charge state can be inferred from x-ray photoelectron spectroscopy, infrared vibrational spectroscopy, or both of adsorbed CO (*8*, *13*). More definitive interpretation requires simulated or experimental reference spectra and, thus, a knowledge of the atomic configuration around the active site (*17*). Nevertheless, the charge state of the metal is frequently invoked to explain reactivity, although this explanation is not without controversy. The cationic nature of Pt_1 has been linked to both stronger (*8*) and weaker (*15*) CO adsorption, with diametrically opposed conclusions regarding CO oxidation activity.

In this paper, we combine temperature-programmed desorption, x-ray photoelectron spectroscopy, scanning tunneling microscopy, and DFT calculations to study CO adsorption on a series of model single-atom catalysts: Cu_1, Ag_1, Au_1, Ni_1, Pd_1, Pt_1, Rh_1, and Ir_1 atoms on Fe_3O_4(001). Upon deposition at room temperature, all of the metals assume the same twofold coordination to surface oxygen, which allows for direct comparison. We selected CO as the probe molecule because of the abundance of experimental and computational results on metal surfaces and nanoparticles and because many single-atom catalysis studies focus on CO oxidation and the water gas shift reaction, in which CO is a reactant. Moreover, the stretching frequency of adsorbed CO is often used to probe the charge state of the metal in single-atom catalysis systems (*8*, *13*, *17*, *18*), which makes a detailed understanding of the metal-CO interaction particularly important.

Our results show that the CO binding strength broadly follows the trends established for the corresponding low-index metal surfaces, i.e., group 9 > 10 > 11, but differences emerge within the groups. Our DFT-based calculations reproduce the experimental data well and show that the adsorption energy is linked to both the electronic structure of the adatom and adsorbate-induced structural relaxations of the system. The combination of the two effects differs from metal to metal because each responds differently to the twofold-coordination environment and because each has a different relative affinity for CO and O. Where the agreement between experiment and theory is imperfect, we show that the catalyst structure evolves during the experiment. Ultimately, our results corroborate that the behavior of single-atom catalysts is better rationalized by analogy to coordination complexes than to metal nanoparticles and that the reactivity of the metal atoms in single-atom catalysis can be tuned if the coordination environment can be controlled.

The experiments described here were performed on several natural Fe_3O_4(001) single crystals over a period of 4 years. The samples were prepared in ultrahigh vacuum by cycles of Ar^+ or Ne^+ sputtering and 900-K annealing. The resulting surfaces exhibited the $(\sqrt{2} \times \sqrt{2})$R45° low-energy electron diffraction pattern and scanning tunneling microscopy signature characteristic of the so-called subsurface cation vacancy reconstruction (*19*, *20*), which is known to stabilize dense arrays of metal atoms to temperatures as high as 700 K (*20*, *21*). Metal was evaporated directly onto the as-prepared Fe_3O_4(001) support in ultrahigh vacuum (i.e., no additional ligands are present). We define coverages in monolayers [ML, which corresponds to 1 atom per $(\sqrt{2} \times \sqrt{2})$R45° unit cell, or 1.42×10^{14} per square centimeter]. In Fig. 1A, scanning tunneling microscopy shows that Cu, Ag, Au, Ni, Pd, Pt, Rh, and Ir atoms all adsorb at the same location between the underlying rows of surface Fe atoms. The surface oxygen atoms are not imaged in scanning tunneling microscopy because they have no electronic states in the vicinity of the Fermi level, but their position is well known from quantitative electron (*19*) and x-ray diffraction (*22*) measurements and DFT-based calculations. The as-deposited structure and thermal stability of the adatom systems shown in Fig. 1A have all been thoroughly characterized previously (*21*, *23*–*30*), which makes this an ideal model system to systematically compare the adatom properties. The adsorption site is twofold coordinated to the surface oxygen atoms that do not have a subsurface tetrahedral Fe neighbor (Cu is shown as an example in Fig. 1B). This is a site where the next Fe cation would reside if the bulk structure continued outward, and the nearest neighbor distance of these sites is the periodicity of the surface reconstruction (8.4 Å). The height of the adatoms above the surface (z) varies

[1]Institute of Applied Physics, TU Wien, Vienna, Austria. [2]Computational Materials Physics, University of Vienna, Vienna, Austria. [3]Alma Mater Studiorum–Università di Bologna, Bologna, Italy.
*These authors contributed equally to this work. †Present address: Advanced Research Center for Nanolithography, Science Park 106, 1098XG Amsterdam, Netherlands.
‡Corresponding author. Email: parkinson@iap.tuwien.ac.at

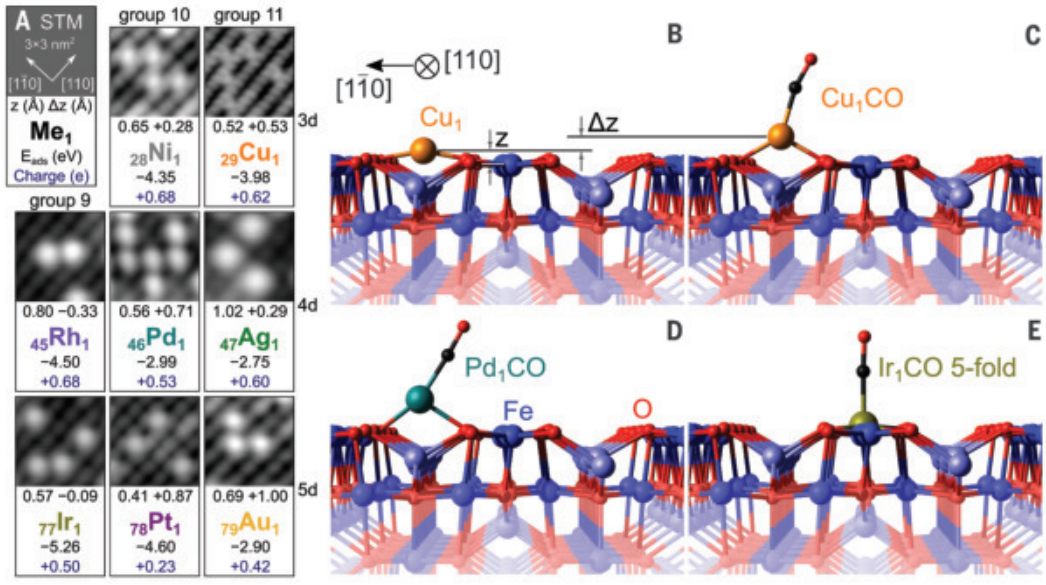


Fig. 1. Characterization of the Fe_3O_4(001)-based single-atom catalysis model systems. (**A**) Representative scanning tunneling microscopy images (V_{sample} = +1 to 1.5 V, I_{tunnel} = 0.1 to 0.3 nA) showing metal adatoms adsorbed midway between the bright Fe rows of the Fe_3O_4(001) support. This corresponds to the twofold adsorption geometry. The nearest neighbor distance is the periodicity of the surface reconstruction (8.4 Å). Alongside each image are the computationally-derived adsorption energies, Bader charges, and heights of the Me_1 adatom (z) above the surface Fe atoms in the twofold adsorption geometry, as well as the CO-induced vertical displacement (Δz). (**B** to **D**) Computationally-derived minimum-energy structure for the twofold-coordinated Cu_1/Fe_3O_4(001) adatom before (B) and after (C) adsorption of CO, as well as the Pd_1CO carbonyl (D). Both are lifted from the surface upon CO adsorption. (**E**) IrCO replaces a fivefold-coordinated surface Fe atom during the temperature-programmed desorption ramp, meaning that CO desorption ultimately occurs from the depicted fivefold Ir_1 geometry (*29*).

greatly, and we recently measured this parameter for Cu, Ag, and Ni (*27*, *28*).

These data were used to benchmark our theoretical approach, and Fig. 1A shows results from calculations using the optimal density functional theory (DFT+U) parameters [effective Hubbard U parameter (U_{eff}) = 3.61 eV, optB88-DF functional] (*27*, *28*). All of the metal adatoms are cationic with Bader charges in the range 0.23*e* to 0.68*e*. Although the Bader-derived charges generally underestimate the formal oxidation states, they are a useful scheme to correlate effective and ionic charges (*17*, *31*). For example, the Ni adatom exhibited a Bader charge (+0.68*e*) midway between the metal (+0.01*e*) and a fivefold-coordinated cation at the NiO(100) surface (+1.19*e*). Because the cations in NiO are nominally 2+, we assign a 1+ oxidation state for the Ni_1/Fe_3O_4(001) adatom. The charge transferred from the metal atom is donated to the undercoordinated surface oxygen atoms, which have a small magnetic moment on the clean surface (*19*). These revert to zero magnetic moment and an O^{2-} state when their fourfold environment is recovered. Ag, Au, and Pd are rather weakly bound (E_{ads} = −2.75, −2.90, and −2.99 eV, respectively), whereas Ir shows the strongest interaction (E_{ads} = −5.26 eV). Cu, Ni, Rh, and Pt lie in between −3.98 and −4.60 eV. Thus, the binding energy scales somewhat with the oxophilicity of the metal, which is in line with previous reports for oxide-supported metal clusters (*32*) and single-atom catalysts (*33*). Ni (*25*), Rh (*30*), and Ir (*29*) adatoms were ultimately unstable against incorporation into the surface lattice, which is important in interpreting our results.

To determine how strongly CO binds to the different metal adatoms, we conducted a series of temperature-programmed desorption experiments. The interaction of CO with the Fe_3O_4(001) support is weak (*34*), and the molecule desorbs from surface Fe^{3+} sites in two peaks between 60 and 100 K (fig. S1). Additional small desorption peaks between 100 and 220 K arise from CO desorbing from Fe^{2+}-containing defects in the surface, such as antiphase domain boundaries and step edges (*34*). The gray curves in Figs. 2, A to G, show the "clean-surface" CO temperature–programmed desorption data acquired before adsorption of the metal adatoms (omitting the much larger desorption peaks from the regular Fe sites, fig. S1). The small differences in the relative intensity of the different defect peaks from experiment to experiment were the result of these data being acquired with several different Fe_3O_4(001) samples.

The colored curves in Fig. 2 show selected CO temperature–programmed desorption data obtained after deposition of metal adatoms at 300 K. In each case, the sample was cooled to 100 K, CO was adsorbed, and the sample was heated with a ramp of 1 K s^{-1}. An arrow marks the temperature-programmed desorption peak corresponding to desorption of CO from the metal adatom in each case. The peak assignments were the result of a series of scanning tunneling microscopy, x-ray photoelectron spectroscopy, and temperature-programmed desorption experiments for different adatom coverages and experimental conditions. This assignment was necessary because the Me_1/Fe_3O_4(001) systems were not static and evolved differently upon CO exposure and heating. We briefly summarize the basis of the assignment for each metal. Further details can be found in the supplementary materials, and an exhaustive account of these experiments is contained within the Ph.D. thesis by Hulva (*35*).

Ag and Cu were straightforward to interpret because scanning tunneling microscopy measurements showed that the adatom phase was stable for densities up to almost 0.5 ML (*27*). Moreover, the adatom arrays were thermally stable to temperatures as high as 700 K (*24*) as well as after exposure to CO. For all adatoms studied in this work, CO adsorption induced a positive core-level shift in the photoemission spectra associated with the adatom (fig. S2) that disappeared when CO desorbed from the adatom, along with the peak in carbon 1s from the adsorbed CO (fig. S3).

For both Cu and Ag, a small shoulder was present on the high-temperature side of the main temperature-programmed desorption peak. This shoulder resulted from a fraction of adatoms occupying a metastable geometry after room-temperature deposition and decreased in intensity after annealing the system before CO adsorption (figs. S4 and S5). For Cu, an additional sharp peak was observed at 120 K that could arise from $Cu(CO)_2$ dicarbonyls, which can be stabilized at low temperature. Details regarding dicarbonyl formation, including how the structures mimic common coordination complexes, can be found in the supplementary materials (tables S3 and S4).

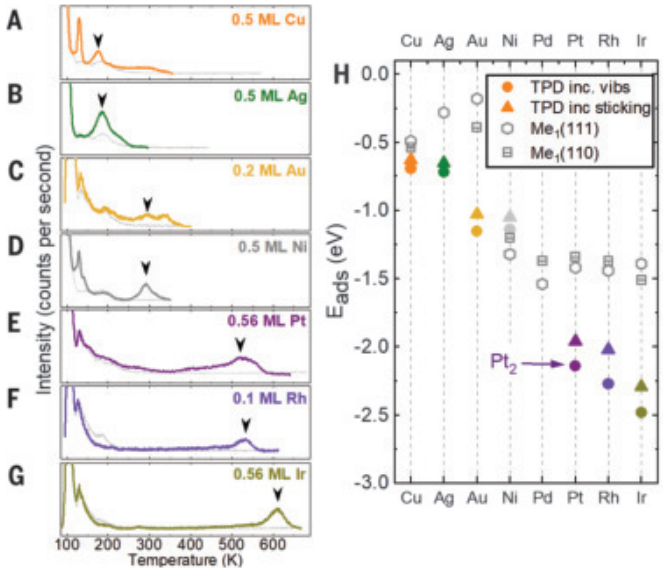


Fig. 2. Determining the CO adsorption on $Me_1/Fe_3O_4(001)$ model catalysts with temperature-programmed desorption. (**A** to **G**) CO temperature–programmed desorption (TPD) curves for various adatoms (1 ML corresponds to 1 metal atom per surface unit cell, or 1.42×10^{14} atoms per square centimeter). Desorption peaks associated with the metal adatoms are labeled using an arrow. Reasoning behind their position is given in the main text. The light-gray curves show CO temperature–programmed desorption data acquired for the ultrahigh vacuum–prepared $Fe_3O_4(001)$ surface before deposition of the metal. (**H**) Plot of experimental and calculated CO adsorption and desorption energies (E_d), alongside experimental values for respective metal (111) and (110) surfaces taken from the literature (table S6).

Au was the most complicated system studied in this work because clusters coexisted with adatoms even at very low coverages [<0.15 ML (*21*)] and because two CO temperature–programmed desorption peaks grew together at 335 K and 285 to 300 K as the Au coverage was increased (fig. S6). These results suggest the presence of two inequivalent sites for adsorption. We assigned the peak at 285 to 300 K to CO desorbing from regular Au adatoms because pre-annealing the system before CO adsorption increased the intensity of the 300-K peak relative to the 335-K peak and because the 335-K peak intensity increased after exposure to water (fig. S7). The latter observation suggested that the 335-K peak was probably related to Au adatoms interacting with surface hydroxyl groups, as has been observed previously by scanning tunneling microscopy for Pd adatoms on this surface (*23*).

Turning to the group 10 metals, the CO peak desorbing from Ni_1 at 300 K was straightforward to assign on the basis of the CO-induced core-level shift. The additional peak at 200 K may be related to Ni incorporated in the surface, as Ni can move to subsurface sites at room temperature (*28*). Temperature-programmed desorption data for Pd are not shown because CO adsorption destabilized the adatoms and led to rapid agglomeration (*23*). A similar process occurred for Pt, but the temperature-programmed desorption data show a main peak that resulted from the decomposition of $Pt_2(CO)_2$ species (*26*).

The CO desorption peaks from the group 9 Rh and Ir adatoms are again straightforward to assign, although these metals incorporated into the oxide lattice when CO desorbed (*29*). Thus, the postdesorption x-ray photoelectron spectroscopy peaks of Rh and Ir exhibited a higher binding energy than the initial twofold adatom (*29*) rather than shifting back to the as-deposited position. The data acquired at 0.1-ML Rh coverage are shown because this metal exhibited a lower-temperature desorption peak for coverages above 0.2 ML (*35*).

In Fig. 2H, we convert the temperature-programmed desorption peak temperatures into desorption energies and plot these alongside values for the corresponding (111) and (110) metal surfaces obtained from table S6 and ref (*36*). Assuming the adsorption-desorption process is reversible and there is no barrier for adsorption, the desorption energy is equivalent to E_{ads}. The basic trend is somewhat similar to the adatoms studied in this work (i.e., group 9 > 10 > 11), but there are large differences in the absolute magnitudes as well as in the tendencies within groups. Most studies on extended metal surfaces typically use the Redhead (*37*) or Polanyi-Wigner equation with a preexponential factor of ~10^{13} s^{-1}, which corresponds to a molecule adsorbed in a corrugation-free surface potential. In our case, however, scanning tunneling microscopy experiments before and after CO desorption reveal no cluster formation, which suggests the CO molecule is constrained at an immobile adatom even at the desorption temperature. We thus treat the system as an ideal two-dimensional lattice gas (*38*). The low density of adsorption sites results in much higher preexponential factors, and assuming the adsorbed CO has three vibrational modes with a frequency of 400 cm^{-1}, we obtain values between 1.27×10^{19} s^{-1} for Cu and 3.25×10^{19} s^{-1} for Ir. The upper and lower bounds for the experimental energies include a temperature uncertainty of ±10 K (±20 K for Au).

Figure 2H also includes the adsorption energies determined from DFT+U for a single CO molecule adsorbed on the metal adatoms (Fig. 2H, solid circles, and table S3). The calculations track the experimental trend with an offset of approximately −0.2 eV for all metals (other than Pt, which formed dimers), which shows that the optB88-DF functional systematically overbinds the CO molecule (as suggested previously for water adsorption on Fe_3O_4) (*39*). Nevertheless, the agreement is excellent, which gave us confidence to delve into the details of the calculations and understand how the "single-atom" nature of the systems affects CO binding.

The d-band center, d-band center of mass, and d-band filling are well-known descriptors affecting the CO adsorption energy on transition metal surfaces (*1*, *36*, *40*). When the outermost d-shell is full, as for group 11 metals, CO binding is dominated by d→2π* back-donation. When the d-shell is partially empty, as for group 9 metals, 5σ→d donation plays a more important role and even dominates in the case of Ir (*41*). Placing a metal atom on a metal-oxide support leads to the formation of chemical bonds, which affects the symmetry of the d-states as well as the filling. On this basis, one would expect the CO adsorption energy to differ from site to site and metal to metal and that there might be parallels with metal-oxide surfaces.

To understand how the nature of the single-atom catalysis site affects the CO adsorption energy, an extensive set of calculations was performed in which CO adsorption on each $Me_1/Fe_3O_4(001)$ system was compared with CO (i) in an on-top site of the corresponding (111) metal surface and (ii) on surface cation sites at the most stable facets of the respective metal oxide. The full dataset, all acquired using the same computational setup, is shown and discussed in the supplementary materials. In what follows, we discuss three representative metals, Ni, Ag, and Ir, which allow us to illustrate the most important factors involved. We found that the reactivity of metals was governed by electronic structure but that CO-induced structural distortions modulate the CO adsorption energy, particularly when the coordination is low.

It would be reasonable to expect that the oxide-supported metal adatoms should have properties somewhat in between a pure metal and a metal oxide. This case was exemplified by the Ni adatom, which exhibited a Bader charge (+0.68*e*) midway between the metal surface (+0.01*e*) and a fivefold-coordinated cation at the NiO(100) surface (+1.19*e*). Because the cations in NiO are nominally 2+,

we assigned a 1+ oxidation state for the $Ni_1/Fe_3O_4(001)$ adatom. Ni^{1+} is not common in nature, which immediately shows that unusual properties can arise when metal atoms are stabilized in undercoordinated geometries. In Fig. 3A, we plot the CO adsorption energy [E_{ads}(CO)] as a function of d-band center of mass for the three different environments mentioned above. As the metal atom becomes more oxidized (Bader charge increased), the d-band center of mass moves further below the Fermi energy E_F, and the CO binding weakens. If the Ni atom replaces a fivefold-coordinated Fe at the $Fe_3O_4(001)$ surface, as has been observed in experiment (*42*), the system resembles NiO (100) because the local environment is almost identical. In this configuration, the d-band center of mass is far away from E_F and d→2π* back-donation cannot occur effectively, so the CO–Ni bond is primarily electrostatic (*43*).

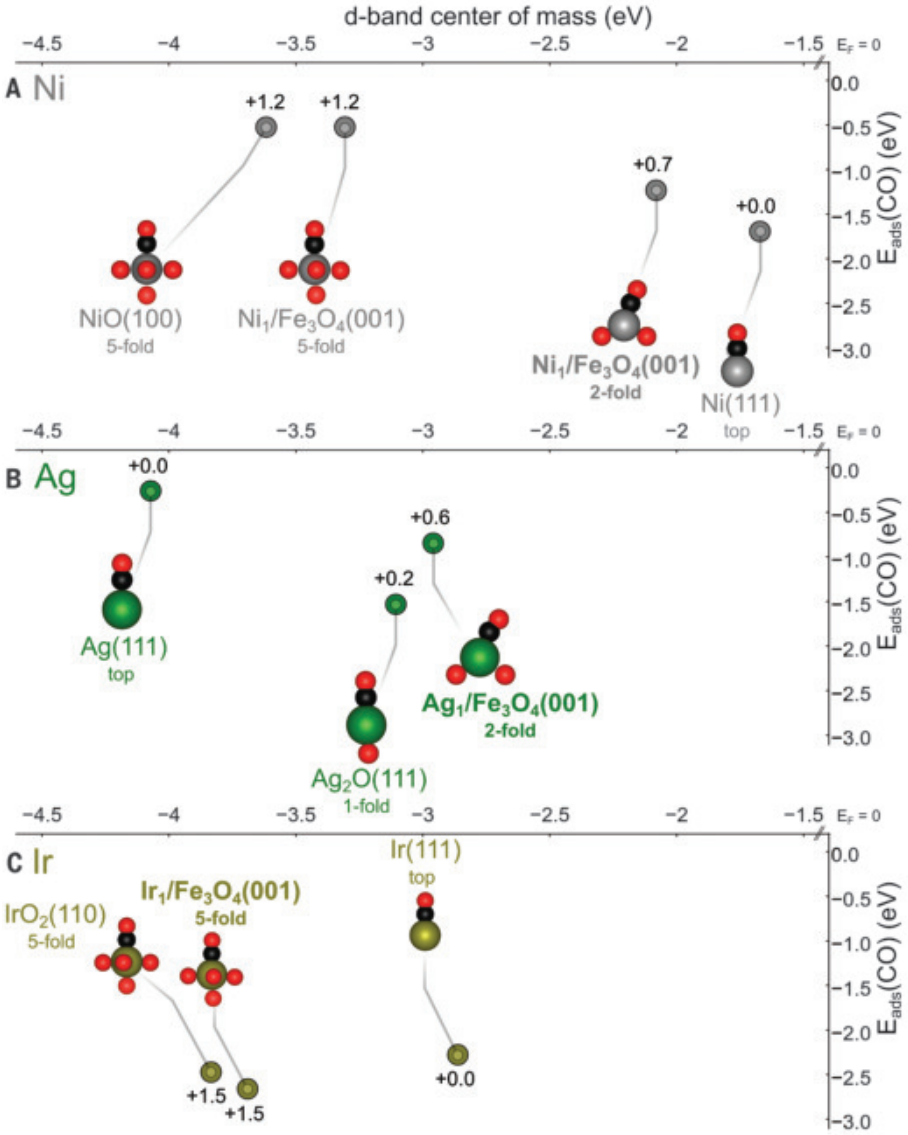


Fig. 3. CO adsorption energy for $Me_1/Fe_3O_4(001)$ single-atom catalysts compared with Me(111) and MeO_x oxide surfaces, plotted against the d-band center of mass. The first coordination sphere of each configuration is illustrated with oxygen (red) and carbon (black), and each point is labeled with the corresponding Bader charge. (**A**) Nickel: As the oxidation state increases, the d-band center of mass moves away from E_F and the CO binding energy decreases. Ni incorporated into a fivefold site on $Fe_3O_4(001)$ is close to NiO(100) because of the similar bonding environment. (**B**) Silver: Oxidized Ag has its d-band center of mass shifted toward E_F, leading to a stronger CO binding energy compared with the metal. The onefold-coordinated Ag atom on $Ag_2O(111)$ exhibits a strong E_{ads}(CO) due to a linear $O_{surface}$–Ag–CO geometry. (**C**) Iridium: The E_{ads}(CO) at fivefold $Ir_1/Fe_3O_4(001)$ is enhanced over Ir(111) because of the strong oxidation of the cation, which enhances 5σ→d donation. In this site, the Ir cation is similar to a cation in the $IrO_2(110)$ surface.

In the case of $Ag_1/Fe_3O_4(001)$ (Fig. 3B), CO is bound much more strongly than for the corresponding metal surface because the d-band center of mass shifts by 1.1 eV toward E_F in the twofold-coordinated adatom site. Because Ag retains a filled d-shell in the 1+ state, d→2π* back-donation continues to dominate, and the CO bond is substantially strengthened. However, the interaction is so strong that the Me_1-support bonds are weakened to accommodate the molecule in a more favorable geometry. Our reference calculations for the $Ag_2O(111)$ surface suggest that the CO adsorption energy is optimal when the O–Me_1–CO bond angle is 180° (supplementary materials). Such an ideal geometry cannot be achieved at the $Fe_3O_4(001)$ lattice, and the system comes to equilibrium when the energetic gain achieved by strengthening the Me_1–CO bond balances the cost of weakening the Me_1–O bonds. Thus, the adsorption energy is weaker than what would be expected on the basis of electronic structure of the adatom alone.

A similar distortion of the local structure was observed for all group 10 and 11 metals. In the $Cu_1/Fe_3O_4(001)$ system, the adsorption energy was almost exactly the same as calculated for the metal surface. This situation is a coincidence, however, because the stronger bond caused by the upward shift in the d-band center of mass was almost exactly compensated by the large CO-induced distortion (Fig. 1, B and C). The most extreme consequence of CO-induced distortion is the sintering observed for Pd and Pt, where the formation of a stable carbonyl weakens the support interaction to such an extent that diffusion becomes facile even at room temperature.

Finally, Ir exhibited the strongest CO binding of all the systems considered in this work. Upon CO adsorption, the twofold-coordinated Ir adatom formed a bond to a subsurface O atom and created a pseudo–square-planar environment for the Ir atom (*29*). Binding a second CO molecule created a highly stable $Ir(CO)_2$ dicarbonyl with two bonds to surface oxygen (*29*), again highlighting the preference for the square-planar coordination. Dicarbonyls can be formed on almost all metals studied in this work, and the structure follows that expected on the basis of Me(I) coordination complexes (table S3 and associated discussion).

Regardless, the twofold Ir atom is ultimately unstable against incorporation into the $Fe_3O_4(001)$ lattice and replaces a fivefold-coordinated Fe cation at ~450 K. Thus, CO desorption occurs from a fivefold $Ir_1/Fe_3O_4(001)$ site in the temperature-programmed desorption experiment. Because the Ir–CO bonding is dominated by 5σ→d donation (*41*), the high

oxidation state of the Ir atom (Bader charge = 1.5*e*) in this coordination enhances the CO adsorption energy. The downward shift in d-band center of mass weakens d→2π* back-donation, but this is counteracted by the better alignment of the 5σ and d levels. Little distortion occurs when CO completes the favored sixfold-coordination environment (Fig. 2D), so the adsorption energy is purely based on the electronic structure. The fivefold $Ir_1/Fe_3O_4(001)$ site closely resembles a fivefold Ir cation at the rutile $IrO_2(110)$ surface in terms of Bader charge, location of the d-band center of mass, and E_{ads}(CO). This result suggests that fivefold Ir single-atom catalysts could be cost-effective replacements for IrO_2 electrocatalysts for water splitting, which are highly effective but prohibitively expensive.

Whether a metal incorporates into the Fe_3O_4 support is governed by the oxygen affinity and preferential O coordination. Table S7 lists the overall energetic gains moving from a two- to a fivefold site. The group 11 elements will not incorporate because they prefer the twofold coordination in their bulk oxide and would be oversaturated in a fivefold geometry. The group 10 elements Pd and Pt prefer to be fourfold coordinated to oxygen and have similar energies for two- and fivefold geometries studied in this work. In any case, neither is competitive against sintering into clusters, so incorporation does not occur. Last, the group 9 elements as well as Ni prefer incorporation into the lattice because they have a high oxygen affinity and have octahedral coordination in their most stable oxide.

The implications of our results for single-atom catalysis are twofold. The choice of metal for a specific reaction is generally governed by the Sabatier principle, which states that the optimum catalyst should be reactive enough to adsorb and activate reactants but not so reactive that the rate is hampered by sluggish desorption of products. Single-atom catalysis was originally developed because expensive metals such as Pt exhibit the optimal balance and sit at the top of a "volcano plot." Our study shows that binding the metal atoms to an oxide alters their electronic structure. For most metals, the twofold binding site strongly shifted the d-band center of mass toward E_F and increased the metal–CO bond strength. This resulted in shorter Me–C bond lengths compared with the metal surface (table S4). Moreover, because adsorption energies are a good proxy for activation barriers (*44*), we conclude that the metal adatoms become more reactive, so the metals previously limited by poor reactivity will shift closer to the top of a volcano plot. We refer to bond strength here, not adsorption energy, because the structural distortions render these two quantities inequivalent for the twofold-coordinated structures.

This result also has consequences for the catalysts limited by product desorption because destabilization by structural distortions will lead to product desorption at lower temperatures, that is, avoid poisoning by adsorbed products. Stabilizing low-coordination "single-atom" sites could provide ways to tailor the activity of specific metals and allow non–Pt-group metals to become competitive. This conclusion is in line with two recent theoretical screening studies (*14*, *16*) suggesting that FeO_x-supported Pd_1, Rh_1, Ru_1, and even Ti_1 or Co_1 species can have similar activity as Pt_1 for CO oxidation.

Knowledge of the local geometry is thus essential to understand the adsorption properties of supported metal adatoms. The observed properties stem directly from how the particular metal reacts to a particular bonding environment, and we observed a preference for the systems to adopt the coordination and bond angles seen in corresponding coordination complexes when CO is adsorbed. This, coupled to the observed ability of the single-atom catalysis systems to distort the structure to accommodate reactants, suggests that adsorbed single atoms should be viewed as undercoordinated complexes and not as the smallest possible nanoparticle. The results seen in this work will be transferrable to many oxide supports, including zeolites, which also offer a twofold-coordination environment. Particularly stable configurations are possible if the metal-oxide system forms solid solution in the bulk [e.g., the fivefold $Ni/Fe_3O_4(001)$], in which case the metal site is straightforward to predict. Although we observed a clear driving force for higher coordination (incorporation) for many of the metals studied in this work, low coordination to the oxide could be stabilized by the adsorption of additional ligands (reactants, OH, water) in the reactive environment, provided that these create a stable geometry for the metal atom. If control of the active site can be achieved in industrial systems, single-atom catalysts really do have the potential to combine the best of heterogeneous and homogeneous catalysis.

REFERENCES AND NOTES

1. B. Hammer, Y. Morikawa, J. K. Norskov, *Phys. Rev. Lett.* **76**, 2141–2144 (1996).
2. X. F. Yang *et al.*, *Acc. Chem. Res.* **46**, 1740–1748 (2013).
3. S. Liang, C. Hao, Y. Shi, *ChemCatChem* **7**, 2559–2567 (2015).
4. J. Liu, *ACS Catal.* **7**, 34–59 (2016).
5. B. C. Gates, M. Flytzani-Stephanopoulos, D. A. Dixon, A. Katz, *Catal. Sci. Technol.* **7**, 4259–4275 (2017).
6. A. Wang, J. Li, T. Zhang, *Nat. Rev. Chem.* **2**, 65–81 (2018).
7. H. Zhang, G. Liu, L. Shi, J. Ye, *Adv. Energy Mater.* **8**, 1701343 (2018).
8. K. Ding *et al.*, *Science* **350**, 189–192 (2015).
9. F. Chen, X. Jiang, L. Zhang, R. Lang, B. Qiao, *Chin. J. Catal.* **39**, 893–898 (2018).
10. X. Cui, W. Li, P. Ryabchuk, K. Junge, M. Beller, *Nat. Catal.* **1**, 385–397 (2018).
11. R. Lang *et al.*, *Angew. Chem. Int. Ed.* **55**, 16054–16058 (2016).
12. J. Liu, *Chin. J. Catal.* **38**, 1460–1472 (2017).
13. H. V. Thang, G. Pacchioni, L. DeRita, P. Christopher, *J. Catal.* **367**, 104–114 (2018).
14. J. Liang, Q. Yu, X. Yang, T. Zhang, J. Li, *Nano Res.* **11**, 1599–1611 (2018).
15. B. Qiao *et al.*, *Nat. Chem.* **3**, 634–641 (2011).
16. F. Li, Y. Li, X. C. Zeng, Z. Chen, *ACS Catal.* **5**, 544–552 (2014).
17. L. DeRita *et al.*, *Nat. Mater.* **18**, 746–751 (2019).
18. J. Resasco, S. Dai, G. Graham, X. Pan, P. Christopher, *J. Phys. Chem. C* **122**, 25143–25157 (2018).
19. R. Bliem *et al.*, *Science* **346**, 1215–1218 (2014).
20. G. S. Parkinson, *Surf. Sci. Rep.* **71**, 272–365 (2016).
21. Z. Novotný *et al.*, *Phys. Rev. Lett.* **108**, 216103 (2012).
22. B. Arndt *et al.*, *Surf. Sci.* **653**, 76–81 (2016).
23. G. S. Parkinson *et al.*, *Nat. Mater.* **12**, 724–728 (2013).
24. R. Bliem *et al.*, *ACS Nano* **8**, 7531–7537 (2014).
25. R. Bliem *et al.*, *Phys. Rev. B* **92**, 075440 (2015).
26. R. Bliem *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8921–8926 (2016).
27. M. Meier *et al.*, *Nanoscale* **10**, 2226–2230 (2018).
28. P. T. P. Ryan *et al.*, *Phys. Chem. Chem. Phys.* **20**, 16469–16476 (2018).
29. Z. Jakub *et al.*, *Angew. Chem. Int. Ed.* **58**, 13961–13968 (2019).
30. Z. Jakub *et al.*, *Nanoscale* **12**, 5866–5875 (2020).
31. A. Walsh, A. A. Sokol, J. Buckeridge, D. O. Scanlon, C. R. A. Catlow, *Nat. Mater.* **17**, 958–964 (2018).
32. S. L. Hemmingson, C. T. Campbell, *ACS Nano* **11**, 1196–1203 (2017).
33. N. J. O'Connor, A. S. M. Jonayat, M. J. Janik, T. P. Senftle, *Nat. Catal.* **1**, 531–539 (2018).
34. J. Hulva *et al.*, *J. Phys. Chem. B* **122**, 721–729 (2018).
35. J. Hulva, T. U. Wien, thesis, TU Wien (2019).
36. M. Gajdoš, A. Eichler, J. Hafner, *J. Phys. Condens. Matter* **16**, 1141–1164 (2004).
37. P. A. Redhead, *Vacuum* **12**, 203–211 (1962).
38. C. T. Campbell, L. H. Sprowl, L. Árnadóttir, *J. Phys. Chem. C* **120**, 10283–10297 (2016).
39. M. J. Gillan, D. Alfè, A. Michaelides, *J. Chem. Phys.* **144**, 130901 (2016).
40. A. Stroppa, G. Kresse, *New J. Phys.* **10**, 063020 (2008).
41. M. T. M. Koper, R. A. van Santen, S. A. Wasileski, M. J. Weaver, *J. Chem. Phys.* **113**, 4392–4407 (2000).
42. Z. Jakub *et al.*, *J. Phys. Chem. C* **123**, 15038–15045 (2019).
43. G. Pacchioni *et al.*, *J. Phys. Condens. Matter* **16**, S2497–S2507 (2004).
44. L. C. Grabow, B. Hvolbæk, J. K. Nørskov, *Top. Catal.* **53**, 298–310 (2010).

ACKNOWLEDGMENTS

We thank P. Blaha (TU Wien) and P. T. P. Ryan (Imperial College London) for useful discussions. **Funding:** G.S.P., J.H., M.M., Z.J., and R.B. acknowledge funding from the Austrian Science Foundation (FWF) Start Prize (Y847-N20), and U.D. and F.K. acknowledge the Austrian Science Fund FWF (Project Wittgenstein Prize, Z250-N27). G.S.P. and M.M. acknowledge funding from the European Research Council (ERC) under the European Union's HORIZON2020 Research and Innovation program (ERC grant agreement 864628). Z.J. also acknowledges support from the TU Wien Doctoral College TU-D. The computational results were achieved in part by using the Vienna Scientific Cluster (VSC 3 and VSC 4). **Author contributions:** J.H. and R.B. performed the experiments under the supervision of G.S.P., who conceptualized the research and acquired research funding for the project. M.M. performed the theoretical calculations under the supervision of C.F. J.H., M.M., and G.S.P. wrote the paper with substantial input and revision from U.D., M.S., F.K., Z.J., and C.F. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/375/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S7
References (*45*–*95*)

31 August 2020; accepted 9 December 2020
10.1126/science.abe5757

ORGANIC CHEMISTRY

Iridium-catalyzed *Z*-retentive asymmetric allylic substitution reactions

Ru Jiang, Lu Ding, Chao Zheng, Shu-Li You*

Z-Olefins are challenging synthetic targets owing to their relative thermodynamic instability. Transition metal–catalyzed asymmetric allylic substitution reactions are well known for installing stereocenters adjacent to branched or *E*-linear olefins. However, analogous reactions for the synthesis of optically active *Z*-olefin products are rare. Here we report iridium-catalyzed asymmetric allylic substitution reactions that retain *Z*-olefin geometries while establishing an adjacent quaternary stereocenter. The formation of transient *anti*-π-allyl-iridium intermediates and their capture by external nucleophiles before isomerization to the thermodynamically more stable *syn*-π-allyl-iridium counterparts have been observed. These results provide a promising method for preparing chiral *Z*-olefinic compounds.

The *Z*-type geometric isomers of olefins (i.e., *Z*-olefins), in which the large substituents on each carbon reside on the same side of the central bond, are widely observed as basic structural motifs in nature. However, *Z*-olefins are thermodynamically less stable than their *E*-type counterparts, rendering their efficient and highly selective synthesis a challenging task in organic chemistry (*1*). Presently, common methods to access *Z*-olefins include the photo-induced isomerization of *E*-olefins (*2*–*4*), Wittig olefination with reactive ylides (*5*, *6*), *Z*-selective cross-metathesis using molybdenum (Mo) or ruthenium (Ru) catalysts (*7*–*10*), the semi-hydrogenation of alkynes (*11*, *12*), *Z*-retentive cross-coupling (*6*, *13*–*15*), and the asymmetric addition of *Z*-allylboron nucleophiles to ketones and imines (*16*–*18*). *Z*-Olefins with a homoallylic stereocenter are frequently embedded in diverse natural products and bioactive molecules (*19*, *23*) and have therefore generated considerable interest (Fig. 1A). In this regard, asymmetric allylic substitution reactions catalyzed by transition metals (*24*–*26*), such as Mo, palladium (Pd), iridium (Ir), and rhodium (Rh), have been useful for constructing stereochemically defined carbon-carbon and carbon-heteroatom bonds at the homoallylic position (Fig. 1B). The catalyst initially reacts with an *E*-linear olefin, branched olefin, or *Z*-linear olefin substrate, all of which include a leaving group (LG) at the allylic position, leading to the formation of π-allyl-metal complexes in either the syn configuration (from the *E*-linear or branched substrates) or the anti configuration (from the *Z*-linear substrate). The relative thermodynamic stabilities of the complexes result in the *syn*-π-allyl-metal complexes dominating the equilibrium between these two types of species via the π-σ-π interconversion process (*27*–*30*), except in the case of unusual outliers (*31*). Subsequently, the well-known regioselective attacks of prochiral nucleophiles on the *syn*-π-allyl-metal complexes at each terminus provide either *E*-linear or branched olefin products enantioselectively. However, direct outer-sphere attack of prochiral nucleophiles on the thermodynamically unfavorable, and therefore transient, *anti*-π-allyl-metal complexes, which would enantioselectively generate *Z*-olefinic products bearing a homoallylic stereocenter, has been elusive. Consequently, the efficient syntheses of *Z*-olefinic products by transition metal–catalyzed asymmetric allylic substitution reactions are uncommon, and only sporadic examples are known in which *Z*-olefinic products bearing homoallylic stereocenters are furnished in enantioenriched form via the inner-sphere reductive elimination mechanism from allyl-Pd or allyl-copper (Cu) species (*32*–*34*).

We posited that the potential for establishing a direct enantioselective nucleophilic attack on the thermodynamically unfavorable *anti*-π-allyl-metal species may be offered by the relatively slow π-σ-π interconversion process observed for π-allyl-Ir complexes (*35*–*38*). The *anti*-π-allyl-Ir complexes generated from the oxidative addition of a judiciously chosen chiral Ir catalyst with *Z*-linear allylic substrates might have a sufficiently long lifetime to be trapped by external prochiral nucleophiles and thereby provide a distinctive *Z*-retentive reaction pathway to install a stereocenter at the homoallylic position of *Z*-olefins. Bearing this hypothesis in mind, we found that the *Z*-retentive asymmetric allylic substitution reactions of *Z*-allyl trifluoroacetates acetates could be coupled with catalytic asymmetric dearomatization (*39*, *40*). Here we report reactions of *Z*-linear allylic substrates with tryptamines and tryptophols, providing convenient access to intriguing privileged polycyclic molecules, such as furoindolines and pyrroloindolines, which are well-known scaffolds in natural products and bioactive molecules (*41*), with a *Z*-olefin moiety installed at the ring junction with high enantiopurity and without the loss of the geometric specificity of the C=C double bond. Comprehensive analysis of the reaction mechanism provides evidence for the formation of key *anti*-π-allyl-Ir complexes and their capture by external nucleophiles before the occurrence of isomerization leading to the thermodynamically more stable *syn*-π-allyl-Ir counterparts. Under modified reaction conditions, the *Z*-retentive asymmetric allylic substitution reactions of *Z*-linear allyl acetates with various α-amino acid–derived aldimine esters were also realized. The reactions incorporated a *Z*-olefin moiety into the α-amino acid skeleton with the concomitant formation of a quaternary stereocenter in high enantioselectivity.

We first investigated the asymmetric allylic dearomatization reaction involving *N*-methyl tryptophol **1a** with *Z*-cinnamyl derivatives [Fig. 2; see tables S1 to S8 for more details]. When *Z*-cinnamyl acetate **2A'** was used, the desired reaction could be achieved with the catalyst derived from the iridium cyclooctadiene (COD) chloride dimer $[Ir(COD)Cl]_2$ (3 mol %) and Carreira-type chiral (phosphoramidite, olefin) ligand (*42*, *43*) (*S*)-**L1** (6 mol %) in the presence of $Zn(OTf)_2$ [1.0 equivalent (equiv), OTf = OSO_2CF_3, triflate] as a promoter in $CHCl_3$ at 0°C. Target product *Z*-**3aA** was obtained in moderate yield (28%), enantioselectivity [78% enantiomeric excess (ee)], and *Z*-type to *E*-type geometric ratio (*Z*/*E*) of 2:1. The reaction was initiated with the nucleophilic attack of the C3 position of the indole ring of **1a** to a proposed *anti*-π-allyl-Ir intermediate. Subsequently, the cyclization of the pendant nucleophile to the iminium moiety formed in the initial step afforded *Z*-**3aA**. Owing to the ring stain of the [3.3.0] bicyclic structure of *Z*-**3aA**, the facial selectivity of the cyclization step should be directed by the stereocenter established during the asymmetric allylic substitution step. Therefore, *Z*-**3aA** was generated as a single cis diastereoisomer (>20:1 diastereoselectivity; see fig. S1 for detailed discussion). The examination of Lewis acid additives suggested that $In(OTf)_3$ was the optimal choice, allowing the exclusive formation of *Z*-**3aA** in 84% yield (>20:1 *Z*/*E*) with 83% ee. The use of a Brønsted acid additive such as triflic acid (HOTf) did not give satisfactory results. The reaction became faster at room temperature (25°C, 1 day) but resulted in slightly decreased yield (77%) and enantioselectivity (78% ee). Switching **2A'** to *Z*-cinnamyl alcohol (**2A"**) or *Z*-cinnamyl trifluoroacetate acetate (**2A**) provided comparable yields (76 and 88%, respectively) and enantiopurities (84 and 87% ee, respectively) of *Z*-**3aA**. With **2A** as the reaction partner, the molar ratio of Ir atoms to (*S*)-**L1** in the catalyst was next considered, and the optimal value was 1:1. With a newly synthesized ligand

State Key Laboratory of Organometallic Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, 345 Lingling Lu, Shanghai 200032, China.
*Corresponding author. Email: slyou@sioc.ac.cn

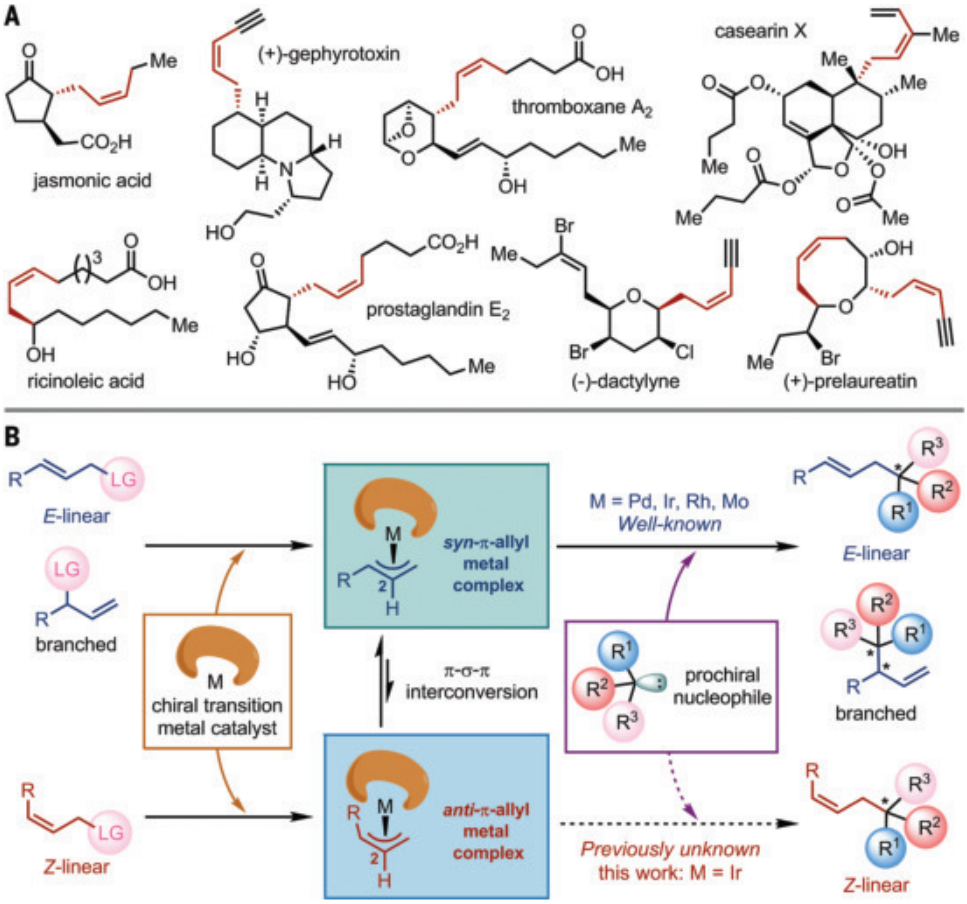
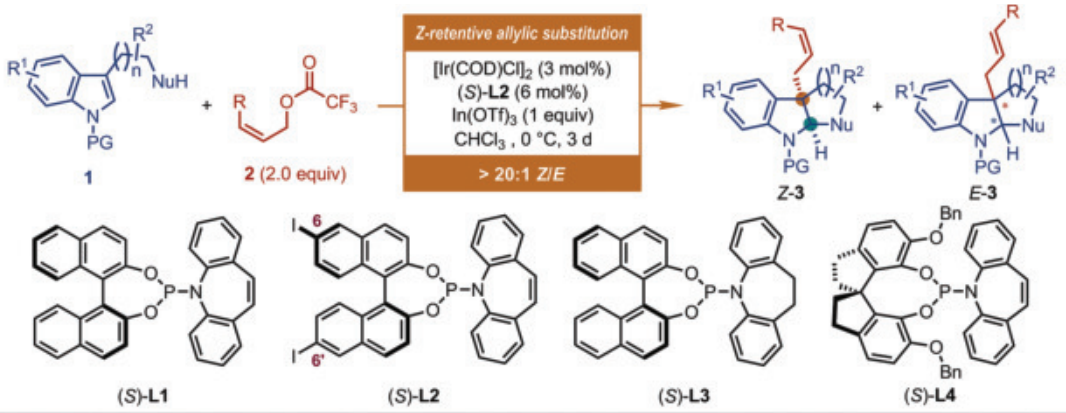


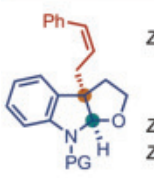
Fig. 1. Strategy for the construction of *Z*-olefins via asymmetric allylic substitution reactions catalyzed by transition metal catalysts. (**A**) Selected natural products containing *Z*-olefins with homoallylic stereocenters. Me, methyl. (**B**) Proposed general mechanism of transition metal–catalyzed asymmetric allylic substitution reactions involving *E*-linear, branched, or *Z*-linear olefin substrates, all of which include an atom or group of atoms denoted as a leaving group (LG).

(*S*)-**L2**, bearing two iodine atoms incorporated at the 6,6' positions of the 1,1'-bi-2-naphthol (BINOL) skeleton of the ligand, the optimal reaction outcomes in terms of the yield (84%) and ee (93%) of *Z*-**3aA** were achieved. In contrast, the chiral phosphoramidite (*S*)-**L3** with the original olefin moiety hydrogenated failed to promote the desired reaction. The above results highlight the key role of the bidentate coordination mode of the chiral (phosphoramidite, olefin) ligand, which was supported by extensive mechanistic studies (see below). The Ir catalyst consisting of a 1,1'-spirobiindane-7,7'-diol (SPINOL)–derived (phosphoramidite, olefin) ligand (*S*)-**L4** was also effective in this reaction, providing *Z*-**3aA** in 75% yield with 90% ee.

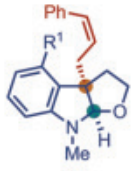
The proposed *Z*-retentive asymmetric allylic substitution reaction exhibited diverse functional group compatibility and tolerated a broad substrate scope, where more than 50 examples involving various (homo)tryptophols, tryptamines, and indole substrates bearing a pendant carbon-based nucleophile were successfully used with **2A** as the coupling partner. Nearly all reactions afforded *Z*/*E* ratios greater than 20:1 (the exceptions for **3j'A**, **3m'A**, and **3aP** are specified in the Fig. 2 caption). The corresponding *Z*-cinnamylated furoindolines (*Z*-**3aA** to *Z*-**3pA**), tetrahydropyranindoline (*Z*-**3qA**), pyrroloindolines (*Z*-**3a'A** to *Z*-**3o'A**), and hexahydrocyclopenta[*b*]indoles (*Z*-**3a''A** to *Z*-**3c''A**) were obtained in good yields of up to 93% and high ee values of up to 94%. In general, substituents of varied electronic and steric properties at all available positions of the indole ring (except for C2) had no substantially deleterious impact on either the efficiency or the enantioselectivity of the reaction. The structures and absolute configurations of *Z*-**3hA** and *Z*-**3b'A** were established by x-ray crystallographic analysis. The performance of D- and L-tryptophan derivatives was also examined. Both *exo*- and *endo*-*Z*-**3p'A** were obtained efficiently with 75 and 71% yields and high diastereoselectivity of 13:1 and 10:1, respectively. The use of (*S*)-**L2** as the chiral ligand indicates that the newly installed stereocenters in these reaction products were mainly controlled by the chiral catalyst. Reactions of **1a** with *Z*-cinnamyl trifluoroacetates possessing substituents in the meta or para positions proceeded smoothly and gave rise to the enantioenriched products *Z*-**3aB** to *Z*-**3aQ** with ee values of up to 93% in reasonable yields of up to 92%. In contrast, the reaction of **1a** with a *Z*-cinnamyl trifluoroacetate possessing an ortho substituent was not accommodated, probably because of unfavorable steric congestion during the reaction. Electron-rich allylic substrates were found to be incompatible with the reaction. The major reaction pathway observed for reaction of **1a** with *p*-methoxy-*Z*-cinnamyl trifluoroacetate was a Friedel–Crafts-type allylation at the C2 position of the indole ring (see page S76 of the supplementary materials for more details). The reaction could be further expanded to alkyl-substituted *Z*-allyl acetates, although only moderate enantioselectivity was obtained for *Z*-**3aR** (60% ee). In addition, *Z*-cinnamyl acetates participated well in the reaction. Chiral furoindolines bearing a *Z*-vinyl iodide (*Z*-**3aS**; 87% yield and 81% ee) or a *Z*-vinyl arene moiety (*Z*-**3aT**; 57% yield and 90% ee; or *Z*-**3aU**; 63% yield and 90% ee) were obtained under slightly modified reaction conditions, as outlined in the Fig. 2 caption. The reaction of **1a** with **2A** could also be conducted on a gram-scale without loss of the enantiopurity of *Z*-**3aA** (1.37 g, 78% yield, 93% ee).

Figure 3 illustrates how a number of products of the *Z*-retentive asymmetric allylic substitution reaction can be readily elaborated in diverse downstream transformations. For example, the *Z*-vinyl iodide of *Z*-**3aS** was reactive towards aryl boronic acids and organozinc

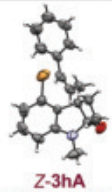




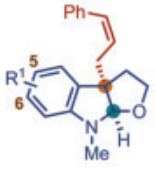
Z-3aA (PG = Me) 84% yield, 93% ee
[Gram-scale reaction] 78% yield, 93% ee 1.37 g
Z-3bA (PG = Bn) 70% yield, 89% ee
Z-3cA (PG = PMB) 84% yield, 89% ee



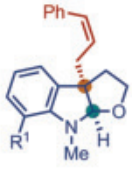
Z-3dA (R^1 = Me) 93% yield, 94% ee
Z-3eA (R^1 = cyclopropyl) 81% yield, 93% ee
Z-3fA (R^1 = Ph) 80% yield, 94% ee
Z-3gA (R^1 = Cl) 83% yield, 92% ee
Z-3hA (R^1 = Br) 87% yield, 93% ee



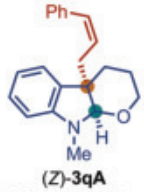
Z-3hA
(CCDC 2007144)



Z-3iA (R^1 = 5-Br) 42% yield, 78% ee
Z-3jA (R^1 = 6-Me) 60% yield, 90% ee
Z-3kA (R^1 = 6-Cl) 78% yield, 93% ee
Z-3lA (R^1 = 6-Br) 67% yield, 93% ee



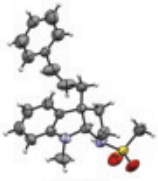
Z-3mA (R^1 = Me) 85% yield, 93% ee
Z-3nA (R^1 = cyclopropyl) 80% yield, 94% ee
Z-3oA (R^1 = Cl) 51% yield, 90% ee
Z-3pA (R^1 = Br) 80% yield, 89% ee



(*Z*)-**3qA**
76% yield, 91% ee



Z-3a'A (PG = Ts) 78% yield, 88% ee
Z-3b'A (PG = Ms) 76% yield, 92% ee
Z-3c'A (PG = CO_2Me) 81% yield, 88% ee
Z-3d'A (PG = CO_2Et) 76% yield, 89% ee
Z-3e'A (PG = CO_2*t*Bu) 69% yield, 85% ee



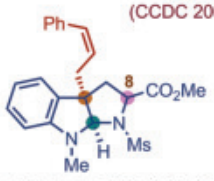
Z-3b'A
(CCDC 2007085)



Z-3f'A (R^1 = 4-Me) 86% yield, 94% ee
Z-3g'A (R^1 = 4-Cl) 74% yield, 94% ee
Z-3h'A (R^1 = 4-Br) 88% yield, 94% ee
Z-3i'A (R^1 = 4-Ph) 84% yield, 92% ee
Z-3j'A (R^1 = 5-Me)* 41% yield, 59% ee
Z-3k'A (R^1 = 5-Cl) 66% yield, 75% ee



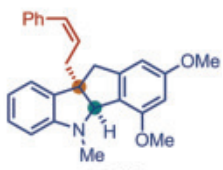
Z-3l'A (R^1 = 6-Me)
73% yield, 88% ee
Z-3m'A (R^1 = 6-Cl)†
83% yield, 92% ee
Z-3n'A (R^1 = 7-Me)
83% yield, 93% ee
Z-3o'A (R^1 = 7-Cl)
89% yield, 90% ee



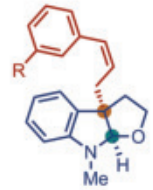
exo-**Z-3p'A** (8*R*) 75% yield, 13:1 dr
endo-**Z-3p'A** (8*S*) 71% yield, 10:1 dr



Z-3a"A (R = Me) 88% yield, 84% ee
Z-3b"A (R = Et) 93% yield, 82% ee



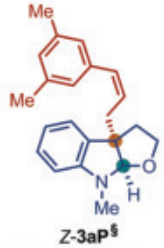
Z-3c"A
78% yield, 85% ee



Z-3aB (R = OMe) 82% yield, 91% ee
Z-3aC (R = Me) 77% yield, 93% ee
Z-3aD (R = OBn) 88% yield, 85% ee
Z-3aE (R = F) 87% yield, 88% ee
Z-3aF (R = NO_2)‡ 70% yield, 82% ee
Z-3aG (R = CF_3)‡ 91% yield, 86% ee
Z-3aH (R = CN) 71% yield, 82% ee



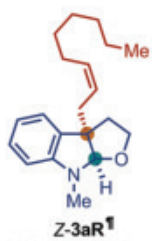
Z-3aI (R^1 = F, R^2 = R^3 = H) 83% yield, 90% ee
Z-3aJ (R^1 = Cl, R^2 = R^3 = H) 89% yield, 90% ee
Z-3aK (R^1 = Br, R^2 = R^3 = H) 87% yield, 87% ee
Z-3aL (R^1 = CO_2Me, R^2 = R^3 = H) 92% yield, 90% ee
Z-3aM (R^1 = OCF_3, R^2 = R^3 = H)‡ 74% yield, 89% ee
Z-3aN (R^1 = Ph, R^2 = R^3 = H) 80% yield, 92% ee
Z-3aO (R^1 = F, R^2 = Me, R^3 = H) 73% yield, 91% ee



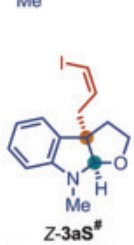
Z-3aP§
54% yield, 92% ee



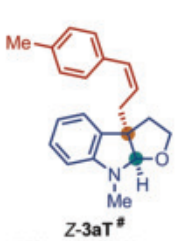
Z-3aQ
82% yield, 79% ee



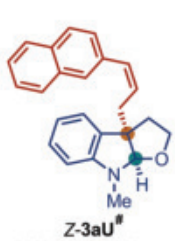
Z-3aR¶
89% yield, 60% ee



Z-3aS#
87% yield, 81% ee



Z-3aT#
57% yield, 90% ee



Z-3aU#
63% yield, 90% ee

Fig. 2. Substrate scope of Ir-catalyzed *Z*-retentive asymmetric allylic substitution reactions of indole derivatives. All reactions were conducted according to the conditions given here, unless otherwise indicated. The *Z*-type to *E*-type geometric ratio (*Z*/*E*) and diastereomeric ratio (dr) were determined by ^{1}H NMR analysis of the crude mixtures. The isolated yields of *Z*-**3** are reported. The ee values were determined by high-performance liquid chromatography (HPLC) analysis with a chiral stationary phase. Asterisk indicates a *Z*/*E* ratio of 7:1. Single-dagger symbol indicates a *Z*/*E* ratio of 8:1. Double-dagger symbol indicates that the reactions were conducted for 1 day. Section symbol indicates a *Z*/*E* ratio of 11:1. Paragraph symbol indicates that the reactions were conducted using (*Z*)-non-2-en-1-yl acetate, $[Ir(COD)Cl]_2$ (3 mol %), (*S*)-**L1** (6 mol %), $InBr_3$ (1.0 equiv), and tetrabutylammonium bromide (0.5 equiv) in $CHCl_3$ at 0°C. Pound symbol indicates that the reactions were conducted using the corresponding *Z*-allylic acetates, $[Ir(COD)Cl]_2$ (4.5 mol %), (*S*)-**L2** (9 mol %), $In(OTf)_3$ (0.5 equiv for *Z*-**3aS** and *Z*-**3aT**; 1 equiv for *Z*-**3aU**) in $CHCl_3$ at −10°C for 2 days. Bn, benzyl; Bu, butyl; Et, ethyl; Ms, mesyl; Ph, phenyl; PMB, *para*-methoxybenzyl; *p*-Tol, *para*-tolyl; Ts, tosyl.

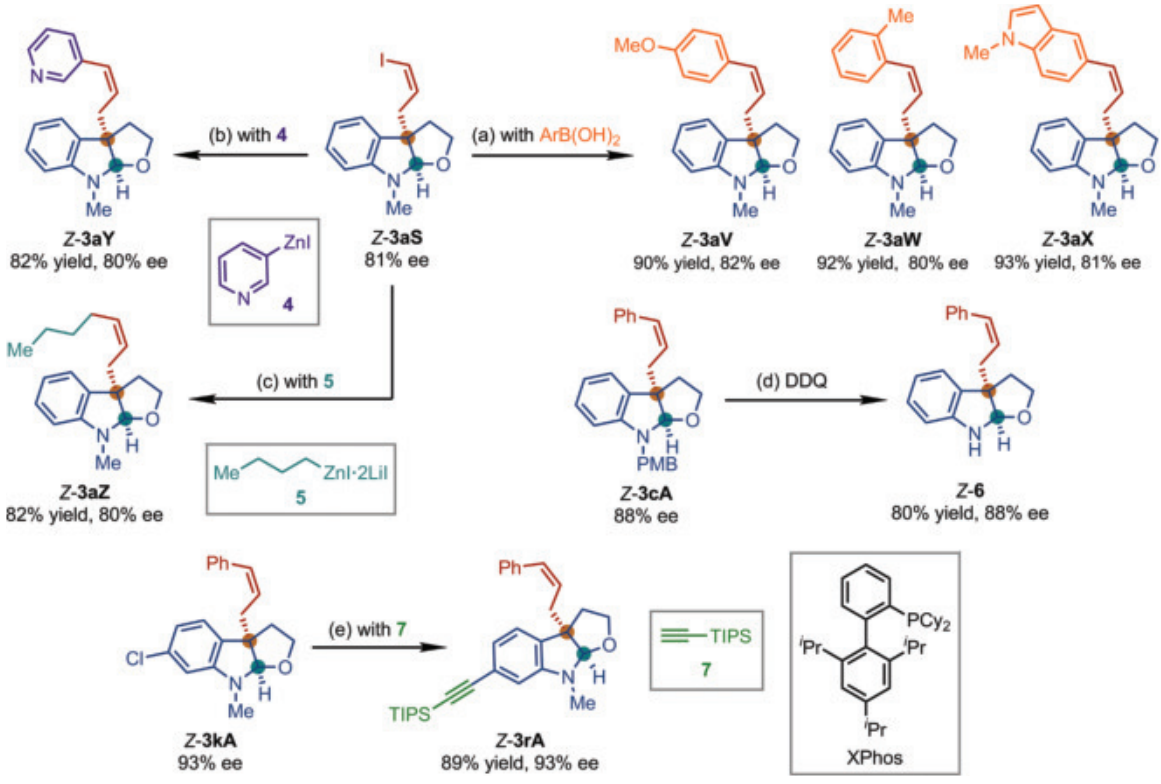


Fig. 3. Representative products obtained in subsequent reactions with the products of Ir-catalyzed *Z*-retentive asymmetric allylic substitution reaction. Reaction conditions: (a) $Pd(PPh_3)_4$ (0.5 mol %), Cs_2CO_3 (2 equiv), $ArB(OH)_2$ (1.5 equiv), acetonitrile, 50°C; (b) **4** (2 equiv), THF/diglyme (1:1), 110°C, 24 hours; (c) **5** (2 equiv), LiCl (4 equiv), toluene/THF/TMU, 110°C, 24 hours; (d) DDQ (1.5 equiv), CH_2Cl_2, NaH_2PO_4 (aqueous 1 M); (e) **7** (2 equiv), $PdCl_2(CH_3CN)_2$ (5 mol %), XPhos (15 mol %), Cs_2CO_3 (2.5 equiv), acetonitrile, 80°C. Ar, aryl; THF, tetrahydrofuran; TIPS, triisopropylsilyl; TMU, 1,1,3,3-tetramethylurea.

reagents **4** and **5**. This provides an alternative route to products *Z*-**3aV** to *Z*-**3aZ** with high yields of 82 to 93%, which could not be obtained via direct asymmetric *Z*-retentive allylic substitution reactions. In addition, the chlorine atom of *Z*-**3kA** facilitated Sonogashira coupling with silyl-protected acetylene **7** to afford *Z*-**3rA** in 89% yield. Removal of the N-PMB group of *Z*-**3cA** by 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) oxidation produced *Z*-**6** in 80% yield.

The specific reaction mechanisms involved in the proposed *Z*-retentive asymmetric allylic substitution were evaluated by a series of analyses. As illustrated in entries 7 to 9 of table S1, the *Z*-**3aA** product of the reaction was obtained with optimal results (88% yield and 87% ee) when the molar ratio of Ir atoms to (*S*)-**L1** in the catalyst was 1:1. In addition, the ee values of (*S*)-**L1** and *Z*-**3aA** are highly correlated linearly with a coefficient of determination $R^2 = 0.9996$ [see the mechanistic studies section (table S10 and fig. S3) in the supplementary materials for details]. These features strongly suggest that the active catalytic species consists of only a single chiral ligand coordinated to Ir. Further characterization of the key intermediates is summarized in the catalytic cycle in Fig. 4A. Mixing (*S*)-**L1** with $[Ir(COD)Cl]_2$ in an Ir to (*S*)-**L1** ratio of 1:1 in $CHCl_3$ produced the neutral Ir complex **K1**. The chloride anion of **K1** was readily abstracted by treating with AgOTf or $In(OTf)_3$ (1 equiv), leading to the cationic Ir complex **K2**. The structures of these Ir complexes were fully elucidated by ^{1}H nuclear magnetic resonance (NMR) spectroscopy, while x-ray crystallographic analysis was applied for **K2**•H_2O as well. Only a single chiral ligand was evident in each complex. We found that (*S*)-**L1** acted as a monodentate ligand in **K1**, while the bidentate coordination of (*S*)-**L1** to the Ir(I) center was confirmed in **K2**, as expected. The **K2** complex was also catalytically active when HOTf was used as the promoter of the allylic substrates [see the mechanistic studies section in the supplementary materials (table S11) for details]. The reactions of **K2** with *Z*- or *E*-cinnamyl carbonate **2** in the presence of HOTf were monitored by electrospray ionization mass spectrometry (ESI-MS) and ^{31}P NMR [see the

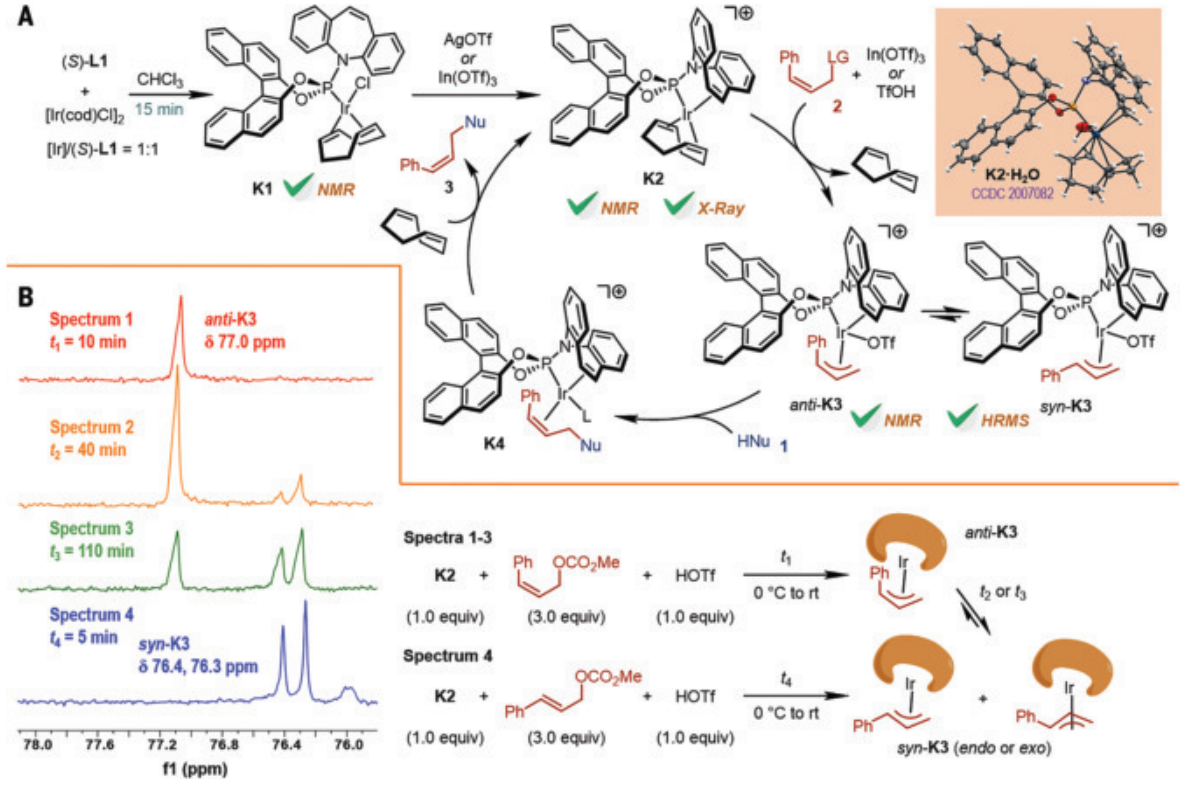


Fig. 4. Reaction mechanism studies. (**A**) A proposed catalytic cycle. The key intermediates **K1** and **K2** were characterized by NMR spectroscopy, while **K2**•H_2O was characterized by x-ray crystallographic analysis. HRMS, high-resolution mass spectrometry. (**B**) The reactions of **K2** with *Z*- or *E*-cinnamyl methyl carbonate demonstrated isomerization from *anti*-**K3** to *syn*-**K3** according to the results of monitoring by electrospray ionization mass spectrometry (ESI-MS) and ^{31}P NMR spectroscopy, the results of which are shown at left, with the details given by the reaction process diagrams to the right.

mechanistic studies section (figs. S4 to S11) in the supplementary materials for details]. The ^{31}P NMR results are presented in Fig. 4B, which demonstrate the isomerization of the complex **K3** from the *anti*- to the *syn*-π-allyl-Ir complex. In spectrum 1, a peak is observed at a chemical shift (δ) of 77.0 parts per million (ppm) when mixing **K2** (1 equiv), *Z*-cinnamyl carbonate (3 equiv), and HOTf (1 equiv, 1 M in ether) in $CDCl_3$ at 0°C, which was assigned to the cationic *anti*-π-allyl-Ir complex *anti*-**K3** [$[M]^+$ mass-to-charge ratio (*m*/*z*) = 966.1232]. This species was quenched immediately when adding *N*-methyl tryptophol (3 equiv), with the formation of *Z*-**3aA** and the regeneration of **K2**. When *anti*-**K3** was allowed to stand at room temperature without the addition of the nucleophile, two peaks gradually emerged (δ = 76.4 and 76.3 ppm, spectra 2 and 3, respectively), which were assigned to the cationic *syn*-π-allyl-Ir complex *syn*-**K3** in either endo or exo configuration. These two species were also observed in the results of ^{31}P NMR and ESI-MS ($[M]^+$ *m*/*z* = 966.1240) analyses of *syn*-**K3** prepared independently starting with *E*-cinnamyl carbonate, as shown in spectrum 4. Collectively, these analyses demonstrate that the success of the proposed *Z*-retentive asymmetric allylic substitution reaction originates from the rapid and enantioselective capture of the transient *anti*-π-allyl-Ir complex by external nucleophiles before its conversion to any other more thermodynamically stable species.

We next explored extension of the Ir-catalyzed *Z*-retentive asymmetric allylic substitution reaction to other nucleophiles. Besides the Ir catalysts consisting of Carreira-type chiral (phosphoramidite, olefin) ligands, which typically promote the reactions under acidic conditions, Ir catalysts derived from a Feringa-type chiral phosphoramidite ligand (*44*), often applied under basic conditions, also proved compatible with this reaction mode. Therefore, diverse nucleophiles beyond tryptophols and tryptamines could be used. Figure 5 summarizes the results for the Ir-catalyzed *Z*-retentive asymmetric allylic substitution reactions of alkyl *Z*-allylic acetates **9**. Herein, the prochiral nucleophiles were formed in situ from a Cu source, $Cu(MeCN)_4OTf$ (5 mol %), a chiral phosphinooxazoline (PHOX)–type ligand (S,S_p)-**L5**, and α-amino acid–derived aldimine esters **8** (*45*, *46*). The independently prepared chiral cyclometallated Ir complex **K5** (2 mol %) promoted the generation of the transient *anti*-π-allyl-Ir complexes from **9** (1.2 equiv) as well as their regio- and enantioselective captures of the aforementioned nucleophiles. The corresponding α-amino acid derivatives *Z*-**10** possessing a *Z*-allyl moiety at the α- stereocenter were delivered after the *Z*-retentive asymmetric allylic substitution and the subsequent reduction of the imine moiety by $NaBH_4$. Systematic evaluation of the reaction conditions revealed that although two chiral components, the chiral Ir complex and PHOX ligand, were involved in the reactions, the chiral discrimination process was dominated by the latter, which was embedded in the prochiral nucleophiles. A slight match-mismatch effect between these two chiral components was observed, and switching **K5** to an achiral Ir complex also led to comparable results (see table S9 for details). The scope of this reaction was quite broad. In more than 30 examples, the desired products *Z*-**10aA** to *Z*-**10tA** and *Z*-**10aB** to *Z*-**10aO** were synthesized exclusively (>20:1 *Z*/*E*, the exception for *Z*-**10aD** is noted in the Fig. 5 caption) in good yields (up to 97%) with excellent enantioselectivity (up to >99% ee). Various functionalities including halogen atoms, ethers, amines, thioethers, silylethers, esters, and sulfonates could be well tolerated. The structure and the absolute configuration of *Z*-**10kA** were established by x-ray crystallographic analysis. When *Z*-cinnamyl acetate was used, a mixture of the desired *Z*-cinnamylated product and its branched isomer (1:1) was obtained. The results presented herein suggest that the proposed protocol has the potential to

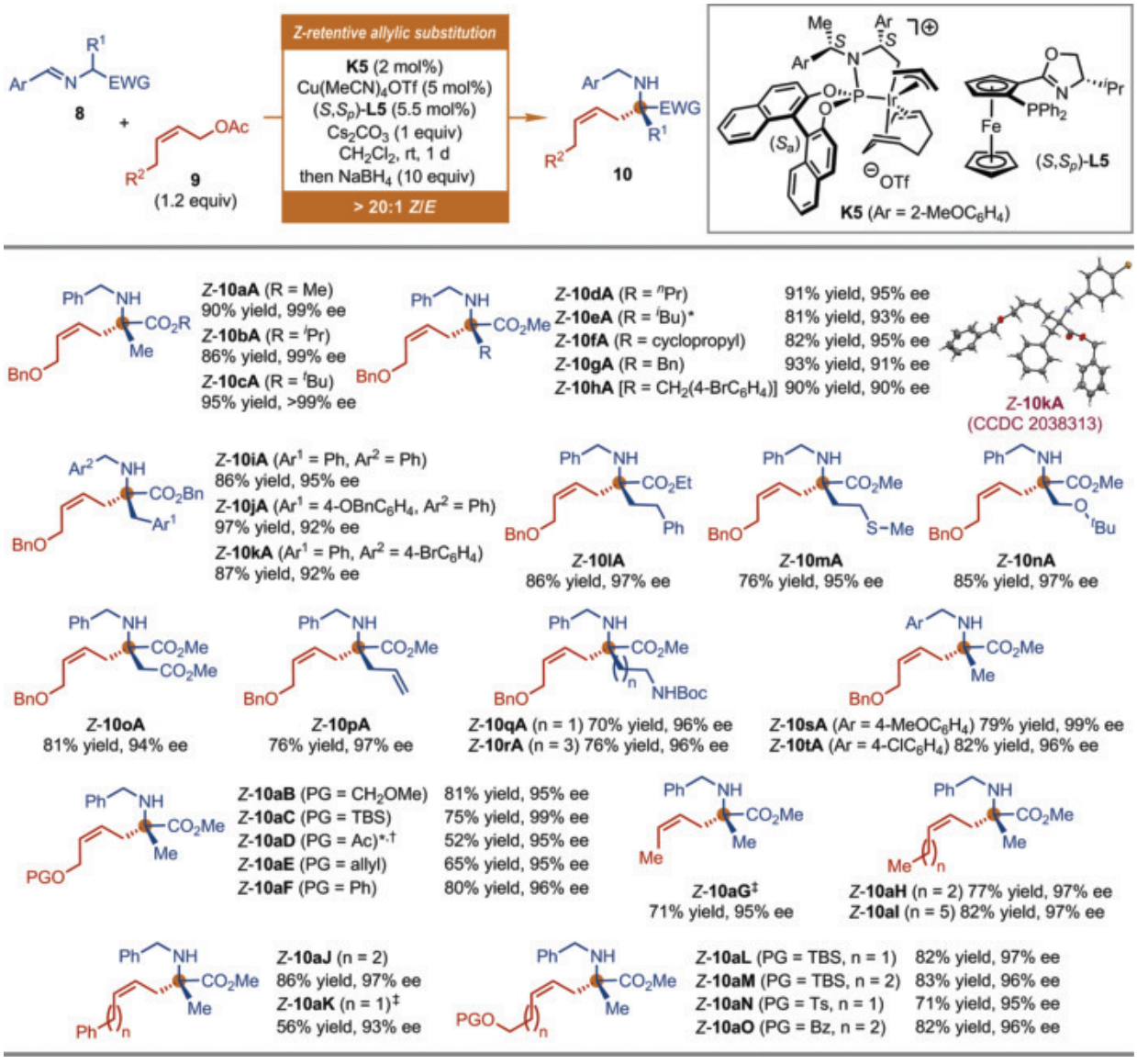


Fig. 5. Substrate scope of Ir-catalyzed *Z*-retentive asymmetric allylic substitution reactions of α-amino acid derivatives. All reactions were conducted under the following conditions. Step 1: **8** (0.2 mmol, 1 equiv), **9** (1.2 equiv), **K5** (2 mol %), $Cu(MeCN)_4OTf$ (5 mol %), (S,S_p)-**L5** (5.5 mol %), and Cs_2CO_3 (1 equiv) in CH_2Cl_2 (2 ml) at room temperature. Step 2: Reaction residue of step 1, $NaBH_4$ (10 equiv) in CH_2Cl_2/MeOH (1:1, 4 ml) at room temperature. The *Z*-type to *E*-type geometric ratio (*Z*/*E*) values were determined by ^1H NMR analysis of the crude mixtures of step 1. The isolated yields of linear *Z*-**10** are reported. The ee values were determined by HPLC analysis with a chiral stationary phase. The asterisk indicates that **9** (1.5 equiv) and **K5** (4 mol %) were used. The single-dagger symbol indicates a *Z*/*E* ratio of 5.3:1. The double-dagger symbol indicates that **9** (2 equiv) was used. Ac, acetyl; Boc, *tert*-butoxycarbonyl; Bz, benzoyl; EWG, electron-withdrawing group; PG, protecting group; Pr, propyl; TBS, *tert*-butyldimethylsilyl.

accommodate diverse prochiral nucleophiles in asymmetric allylic substitution reactions and thus can be expected to evolve into a broadly practical method for the synthesis of chiral *Z*-olefins.

REFERENCES AND NOTES

1. W.-Y. Siau, Y. Zhang, Y. Zhao, in *Stereoselective Alkene Synthesis*, J. Wang, Ed. (Springer, 2012), pp. 33–58.
2. J. B. Metternich, R. Gilmour, *Synlett* **27**, 2541–2552 (2016).
3. K. Singh, S. J. Staig, J. D. Weaver, *J. Am. Chem. Soc.* **136**, 5275–5278 (2014).
4. M. Bhadra *et al.*, *J. Am. Chem. Soc.* **141**, 6152–6156 (2019).
5. B. E. Maryanoff, A. B. Reitz, *Chem. Rev.* **89**, 863–927 (1989).
6. E. Negishi *et al.*, *Acc. Chem. Res.* **41**, 1474–1485 (2008).
7. K. Endo, R. H. Grubbs, *J. Am. Chem. Soc.* **133**, 8525–8527 (2011).
8. B. K. Keitz, K. Endo, P. R. Patel, M. B. Herbert, R. H. Grubbs, *J. Am. Chem. Soc.* **134**, 693–699 (2012).
9. S. J. Meek, R. V. O'Brien, J. Llaveria, R. R. Schrock, A. H. Hoveyda, *Nature* **471**, 461–466 (2011).
10. M. J. Koh *et al.*, *Nature* **517**, 181–186 (2015).
11. A. M. Kluwer, C. J. Elsevier, in *Handbook of Homogeneous Hydrogenation*, J. G. de Vries, C. J. Elsevier, Eds. (Wiley-VCH, 2007), pp. 375–411.
12. C. Oger, L. Balas, T. Durand, J.-M. Galano, *Chem. Rev.* **113**, 1313–1350 (2013).
13. S. Xu, H. Kamada, E. H. Kim, A. Oda, E.-i. Negishi, in *Metal-Catalyzed Cross-Coupling Reactions and More*, A. de Meijere, S. Bräse, M. Oestreich, Eds. (Wiley-VCH, 2014), pp. 133–278.
14. C. W. Cheung, F. E. Zhurkin, X. Hu, *J. Am. Chem. Soc.* **137**, 4932–4935 (2015).
15. M. K. Armstrong, M. B. Goodstein, G. Lalic, *J. Am. Chem. Soc.* **140**, 10233–10241 (2018).
16. F. W. van der Mei, C. Qin, R. J. Morrison, A. H. Hoveyda, *J. Am. Chem. Soc.* **139**, 9053–9065 (2017).
17. D. C. Fager, R. J. Morrison, A. H. Hoveyda, *Angew. Chem. Int. Ed.* **59**, 11448–11455 (2020).
18. R. J. Morrison, F. W. van der Mei, F. Romiti, A. H. Hoveyda, *J. Am. Chem. Soc.* **142**, 436–447 (2020).
19. C. Wasternack, B. Hause, *Ann. Bot.* **111**, 1021–1058 (2013).
20. C. Souccar, W. A. Varanda, J. W. Daly, E. X. Albuquerque, *Mol. Pharmacol.* **25**, 384–394 (1984).
21. A. Gonzaga dos Santos *et al.*, *Chem. Biodivers.* **7**, 205–215 (2010).

22. H. Tohgi, S. Konno, K. Tamura, B. Kimura, K. Kawano, *Stroke* **23**, 1400–1403 (1992).
23. W. Goessling *et al*., *Cell* **136**, 1136–1147 (2009).
24. U. Kazmaier, Ed., *Transition Metal Catalyzed Enantioselective Allylic Substitution in Organic Synthesis* (Springer, 2012).
25. B. M. Trost, M. L. Crawley, *Chem. Rev.* **103**, 2921–2944 (2003).
26. Z. Lu, S. Ma, *Angew. Chem. Int. Ed.* **47**, 258–297 (2008).
27. J. W. Faller, M. E. Thomsen, M. J. Mattina, *J. Am. Chem. Soc.* **93**, 2642–2653 (1971).
28. D. H. Gibson, D. K. Erwin, *J. Organomet. Chem.* **86**, C31–C33 (1975).
29. G. Consiglio, R. M. Waymouth, *Chem. Rev.* **89**, 257–276 (1989).
30. U. Kazmaier, F. L. Zumpe, *Angew. Chem. Int. Ed.* **39**, 802–804 (2000).
31. M. Yamamoto, S. Hayashi, K. Isa, M. Kawatsura, *Org. Lett.* **16**, 700–703 (2014).
32. H.-C. Lin *et al*., *J. Am. Chem. Soc.* **141**, 5824–5834 (2019).
33. T.-C. Wang, P.-S. Wang, L.-Z. Gong, *Sci. China Chem.* **63**, 454–459 (2020).
34. C. I. Jette, Z. J. Tong, R. G. Hadt, B. M. Stoltz, *Angew. Chem. Int. Ed.* **59**, 2033–2038 (2020).
35. R. Takeuchi, M. Kashio, *J. Am. Chem. Soc.* **120**, 8647–8655 (1998).
36. B. Bartels, G. Helmchen, *Chem. Commun.* **1999**, 741–742 (1999).
37. S. T. Madrahimov, J. F. Hartwig, *J. Am. Chem. Soc.* **134**, 8136–8147 (2012).
38. Q. Cheng *et al*., *Chem. Rev.* **119**, 1855–1969 (2019).
39. C.-X. Zhuo, W. Zhang, S.-L. You, *Angew. Chem. Int. Ed.* **51**, 12662–12686 (2012).
40. C.-X. Zhuo, C. Zheng, S.-L. You, *Acc. Chem. Res.* **47**, 2558–2573 (2014).
41. L. M. Repka, S. E. Reisman, *J. Org. Chem.* **78**, 12314–12320 (2013).
42. C. Defieber, M. A. Ariger, P. Moriel, E. M. Carreira, *Angew. Chem. Int. Ed.* **46**, 3139–3143 (2007).
43. S. L. Rössler, D. A. Petrone, E. M. Carreira, *Acc. Chem. Res.* **52**, 2657–2672 (2019).
44. J. F. Teichert, B. L. Feringa, *Angew. Chem. Int. Ed.* **49**, 2486–2528 (2010).
45. L. Wei, Q. Zhu, S.-M. Xu, X. Chang, C.-J. Wang, *J. Am. Chem. Soc.* **140**, 1508–1513 (2018).
46. X. Huo, J. Zhang, J. Fu, R. He, W. Zhang, *J. Am. Chem. Soc.* **140**, 2080–2084 (2018).

ACKNOWLEDGMENTS

Funding: Financial support for this work was provided by the National Key R&D Program of China (grant 2016YFA0202900), the National Natural Science Foundation of China (grants 21821002, 22031012 and 91856201), and the Chinese Academy of Sciences (grants XDB20000000 and QYZDY-SSW-SLH012). S.-L.Y. acknowledges support from the Tencent Foundation through the XPLORER PRIZE. **Author contributions:** The project was conceived of and supervised by S.-L.Y.; R.J. developed the Ir-catalyzed *Z*-retentive asymmetric allylic substitution reactions; L.D. contributed to expanding the scope for indole derivatives; R.J. and C.Z. performed the reaction mechanism studies; C.Z. wrote the manuscript and incorporated revisions suggested by all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Crystallographic data are available free of charge from the Cambridge Crystallographic Database Centre (CCDC) under CCDC 2007081 (for **K1'**, see also the x-ray crystallographic data section of the supplementary materials), 2007082 (for **K2**•H_2O), 2007085 (for *Z*-**3b'A**), 2007144 (for *Z*-**3hA**), and 2038313 (for *Z*-**10kA**). All other characterization data are in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/380/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S16
NMR and HPLC Data
References (*47*–*81*)

1 July 2020; resubmitted 3 November 2020
Accepted 1 December 2020
10.1126/science.abd6095

CONVERGENT EVOLUTION

Convergent evolution of pain-inducing defensive venom components in spitting cobras

T. D. Kazandjian[1]*, D. Petras[2,3]*, S. D. Robinson[4,5]*, J. van Thiel[6], H. W. Greene[7], K. Arbuckle[8], A. Barlow[9,10], D. A. Carter[5], R. M. Wouters[6], G. Whiteley[1], S. C. Wagstaff[1,11], A. S. Arias[12], L.-O. Albulescu[1], A. Plettenberg Laing[10], C. Hall[10], A. Heap[10], S. Penrhyn-Lowe[10], C. V. McCabe[13], S. Ainsworth[1], R. R. da Silva[2,14], P. C. Dorrestein[2], M. K. Richardson[6], J. M. Gutiérrez[12], J. J. Calvete[15], R. A. Harrison[1], I. Vetter[5,16], E. A. B. Undheim[4,5,17,18], W. Wüster[10], N. R. Casewell[1]†

Convergent evolution provides insights into the selective drivers underlying evolutionary change. Snake venoms, with a direct genetic basis and clearly defined functional phenotype, provide a model system for exploring the repeated evolution of adaptations. While snakes use venom primarily for predation, and venom composition often reflects diet specificity, three lineages of cobras have independently evolved the ability to spit venom at adversaries. Using gene, protein, and functional analyses, we show that the three spitting lineages possess venoms characterized by an up-regulation of phospholipase A_2 (PLA_2) toxins, which potentiate the action of preexisting venom cytotoxins to activate mammalian sensory neurons and cause enhanced pain. These repeated independent changes provide a fascinating example of convergent evolution across multiple phenotypic levels driven by selection for defense.

Convergent evolution, the independent emergence of similar traits across taxa, is a pervasive characteristic of biodiversity and provides natural replicates to enable understanding of key evolutionary processes (*1*). Given their discrete function and direct genotype-phenotype link, animal venoms are excellent systems for understanding the driving forces and underlying genetic mechanisms of molecular adaptation. Snake venoms consist of variable mixtures of proteinaceous components causing potent hemotoxic, neurotoxic, and/or cytotoxic pathologies in both prey and potential adversaries, including humans (*2*). Previous work suggests that venom variation is largely driven by dietary variation (*3*), but defensive drivers of snake venom evolution are rarely considered [although, see (*4*)].

The evolution of venom projection or "spitting" in cobras offers an ideal system for exploring the evolution of defensive toxins: This behavior plays no role in prey capture, targets specific sensory tissues, and is the only long-distance, injurious defensive adaptation among almost 4000 species of snakes. Unexpectedly, venom spitting evolved independently three times, all within a single clade of closely related elapid snakes (*5*, *6*): the African spitting cobras (*Naja*: subgenus *Afronaja*), Asian spitting cobras (*Naja*: subgenus *Naja*), and rinkhals (*Hemachatus*). All use fangs with modified orifices (*7*) to spray venom over distances of up to 2.5 m (*8*), targeting an aggressor's eyes (*9*) (fig. S1). These behavioral, morphological, and biochemical traits result in intense ocular pain and inflammation, which can lead to the permanent loss of eyesight (*10*). The three origins of spitting, solely within a clade more generally characterized by the visually defensive behavior of hooding (*5*, *6*), allow us to test whether similar selective pressures have resulted in convergent changes in venom composition coevolving with morphological and behavioral adaptations.

To investigate the evolution of venom spitting, we used a multidisciplinary approach consisting of transcriptomic, proteomic, functional, and phylogenetic comparisons of 17 widely distributed elapids: 14 *Naja* (true cobras); *Hemachatus haemachatus* (rinkhals); and two nonspitting immediate sister-group species, *Walterinnesia aegyptia* and *Aspidelaps scutatus*. First, we reconstructed the phylogeny

[1]Centre for Snakebite Research and Interventions, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK. [2]Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA. [3]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA. [4]Centre for Advanced Imaging, University of Queensland, St Lucia, QLD 4072, Australia. [5]Institute for Molecular Bioscience, University of Queensland, St Lucia, QLD 4072, Australia. [6]Institute of Biology, University of Leiden, Leiden 2333BE, Netherlands. [7]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA. [8]Department of Biosciences, College of Science, Swansea University, Swansea SA2 8PP, UK. [9]School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK. [10]Molecular Ecology and Fisheries Genetics Laboratory, School of Natural Sciences, Bangor University, Bangor LL57 2UW, UK. [11]Research Computing Unit, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK. [12]Instituto Clodomiro Picado, Facultad de Microbiología, Universidad de Costa Rica, San José 11501, Costa Rica. [13]School of Earth Sciences, University of Bristol, Bristol BS8 1RL, UK. [14]Núcleo de Pesquisa em Produtos Naturais e Sintéticos (NPPNS), Molecular Sciences Department, School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil. [15]Evolutionary and Translational Venomics Laboratory, Consejo Superior de Investigaciones Científicas, 46010 Valencia, Spain. [16]School of Pharmacy, University of Queensland, Woolloongabba, QLD 4102, Australia. [17]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway. [18]Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Blindern, 0316 Oslo, Norway.
†Corresponding author. Email: nicholas.casewell@lstmed.ac.uk
*These authors contributed equally to this work.

of these snakes using a multilocus coalescent species tree approach based on two mitochondrial and five nuclear genes. Fossil-calibrated molecular dating suggests that spitting originated in African spitting cobras between 6.7 million and 10.7 million years ago (Ma) and ~4 million years later in the Asian spitting cobras (2.5 to 4.2 Ma) (Fig. 1A). The origin of spitting in *Hemachatus* could not be dated beyond having occurred <17 Ma, after divergence from true cobras (*Naja*) (Fig. 1A).

Next, we used a top-down proteomics approach underpinned by venom gland transcriptomic data (*11*) to characterize the venom composition of each species. All cobra venoms are dominated by three-finger toxins (3FTXs), while in many species phospholipases A_2 (PLA_2) are the second most abundant toxin family (Fig. 1A and fig. S2). Principal coordinates analysis (PCoA, Bray-Curtis) of a proteomic data–derived venom composition matrix separated the spitting lineages into three clusters that are distinct from the homogeneous cluster of venoms from nonspitters (Fig. 1B). The sole exception, *Naja philippinensis*, has a purely neurotoxic venom despite being able to spit (*12*), and its venom composition placed it alongside nonspitting species (see Fig. 1B). Nonetheless, these findings demonstrate that each spitting cobra lineage exhibits distinct venom compositions that collectively differ from those of nonspitting cobras—a finding that is consistent with differences in venom-induced pathology observed after bites to humans (*13*).

In many elapid snakes, 3FTXs are major venom components (*14*) (table S1). They are encoded by a multilocus gene family, resulting in numerous functionally distinct isoforms, including neurotoxins and cytotoxins that disrupt cell membranes to cause cytotoxicity (*15*). Proteomic data revealed that cytotoxic 3FTXs (CTXs) are typically the most abundant toxins in *Naja* and *Hemachatus* venoms (mean: 57.7% of all toxins), contrasting with the sister-group species *W. aegyptia* and *A. scutatus* (*16*) and other elapids (Fig. 1A, figs. S2 and S3, and table S1). Although cytotoxins likely inflict defensive ocular pain, we found no significant difference in the abundances of CTXs between spitting and nonspitting species [phylogenetic generalized least squares (PGLS); $t = -0.83$, degrees of freedom (df) = 15, $P = 0.42$] (table S2), and ancestral state estimations suggest that the origin of CTX-rich venom preceded that of venom spitting (fig. S4) (*5*). Moreover, PCoA analysis of a CTX Euclidean distance matrix derived from venom proteomic data revealed that all highly abundant cobra CTXs cluster tightly together (Fig. 1C). Measures of irritation, stimulated by high doses (100 μg) of cobra venoms applied topically to nonsentient chick embryos (tables S3 and S4), also revealed no association between cytotoxicity and spitting (PGLS; $t = 1.08$, df = 15, $P = 0.30$) (figs. S5 and S6), consistent with prior reports of comparable cytotoxicity to mammalian cells across all cobra venoms (*5*).

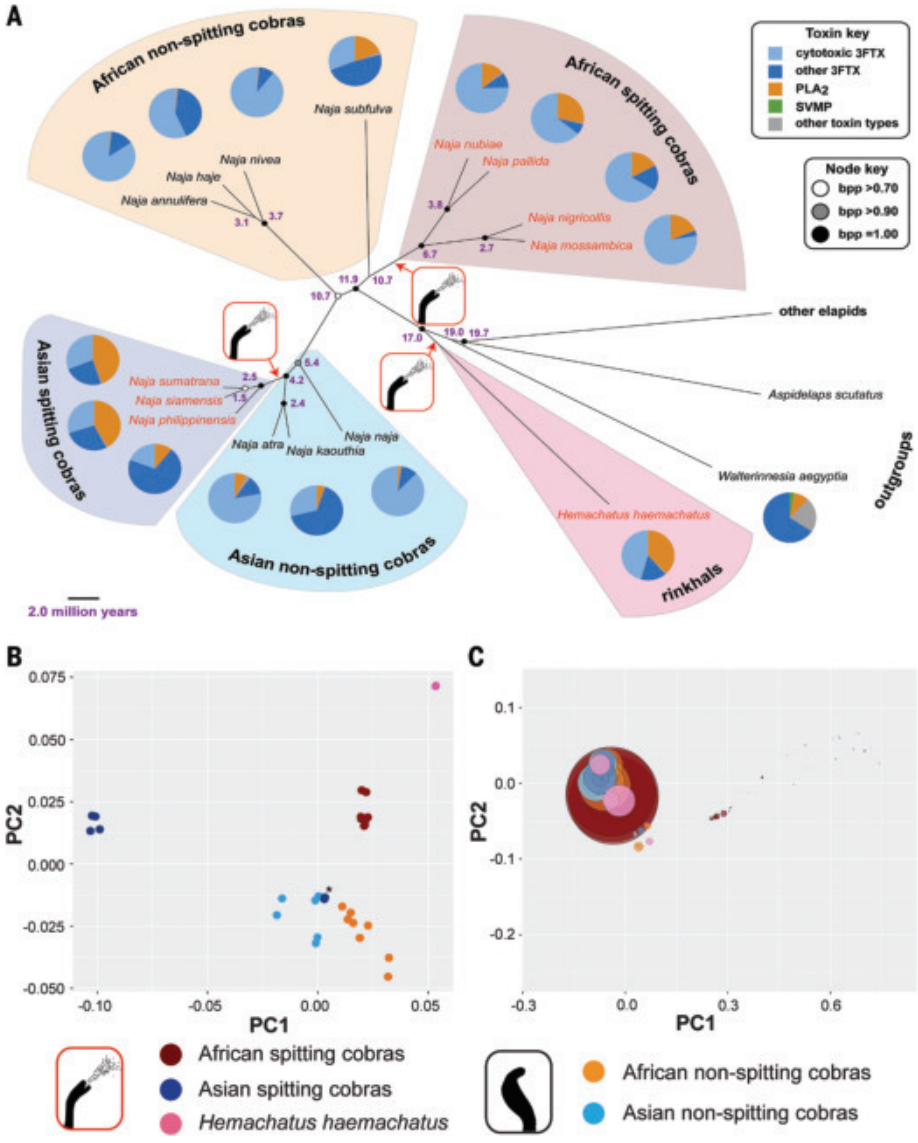


Fig. 1. Reconstruction of the evolutionary origin of venom spitting and comparative analysis of venom composition. (**A**) Multilocus-derived multispecies coalescent species tree, pruned to display the taxa whose venoms were analyzed in this study. Node support is indicated by colored circles, representing Bayesian posterior probabilities: black = 1.00, gray >0.90, white >0.70. Purple node labels indicate estimated divergence times (see fig. S14 for credibility intervals). Spitting species are highlighted by red tip labels, and the three independent origins of venom spitting are indicated by the red-boxed spitting images. Pie charts adjacent to tip labels represent proteomic toxin composition of each species as a percentage of total toxins. SVMP, snake venom metalloproteinases; bpp, Bayesian posterior probability. (**B**) Principal coordinates analysis (PCoA) of cobra (*Naja* spp.) and rinkhals (*H. haemachatus*) venom toxins reveal major distinctions between spitting and nonspitting lineages. Asterisk highlights the Asian spitting species *N. philippinensis*, which exhibits greater similarity to non-spitting species than to its nearest relatives. Each species is represented by two, typically overlapping, data points, which represent technical proteomic duplicates. (**C**) PCoA of cobra (*Naja* spp.) and rinkhals (*H. haemachatus*) cytotoxic three-finger toxins (CTXs) derived from top-down venom proteomics reveals that the most abundant CTXs detected in venom exhibit little sequence diversity among spitting and nonspitting lineages. Circle sizes reflect relative abundances of CTXs detected.

However, pain inflicted via slow-onset cytotoxicity may be less defensively relevant than rapid pain caused by direct algesic activity. To investigate venom-induced nociception, we assessed the activation of mammalian trigeminal neurons—sensory neurons derived from trigeminal ganglia that innervate the face and

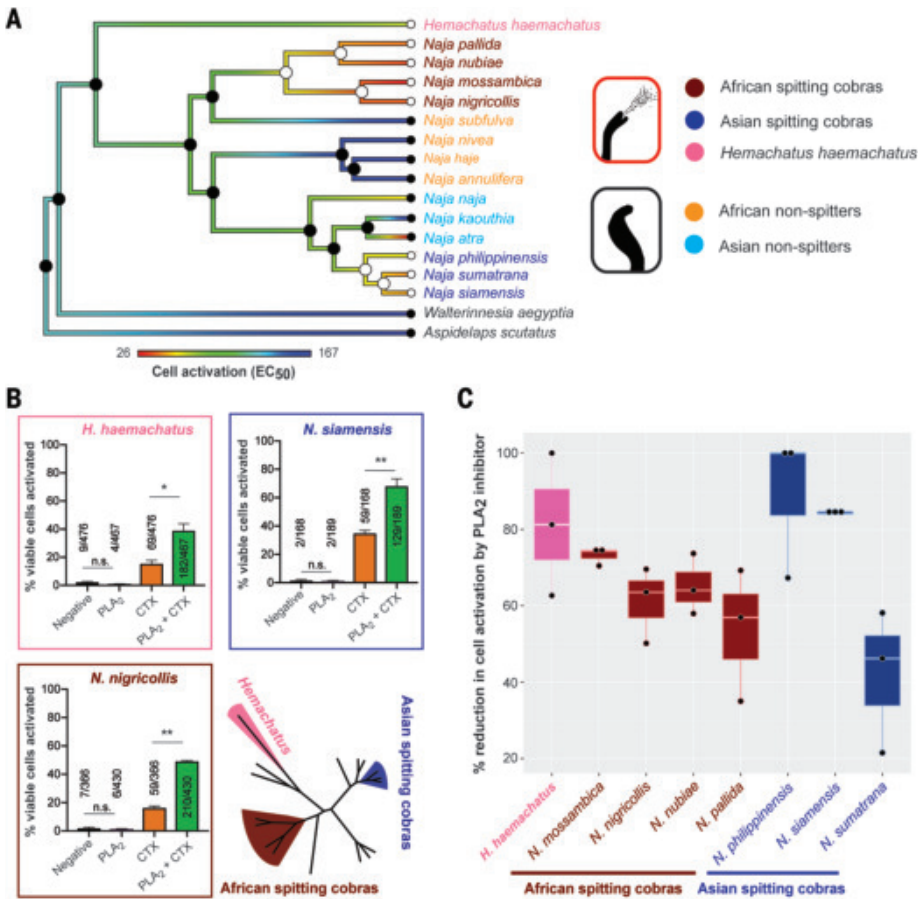


Fig. 2. Spitting cobra venoms cause significantly greater activation of sensory neurons than nonspitting cobras, mediated via potentiation by PLA_2 toxins. (**A**) Ancestral state estimation of half-maximal effective concentrations (EC_{50}) of venom-induced activation of neuronal cells shows a significant association between increased potency and venom spitting (PGLS, $t = -4.48$, $P = 0.0004$). EC_{50} values are expressed as the mean of triplicate measurements, and colored branches are scaled accordingly (red, low EC_{50} and thus high venom potency; blue, high EC_{50} and thus low venom potency). Filled or empty circles at nodes and tips represent estimated ancestral states of nonspitting or spitting, respectively, and colored tip labels correspond to the different lineages. (**B**) PLA_2 toxins in spitting cobra venoms potentiate the activating effect of CTXs on sensory neurons. A CTX fraction from each venom was added to dissociated mouse dorsal root ganglia neurons in the presence or absence of a corresponding PLA_2 fraction (added 1 min prior), neuronal activation (i.e., a rapid increase in $[Ca^{2+}]_i$) monitored, and data presented as mean ± SEM of the resulting percentage of viable cells from two to three independent experiments. The number of viable cells activated and the total number of cells used are also displayed above or within each bar. Statistical comparisons were performed using unpaired parametric *t*-tests; $^*P < 0.05$; $^{**}P < 0.01$; n.s., not significant. (**C**) The PLA_2 inhibitor varespladib reduces neuronal activation stimulated by spitting cobra venoms. Calcium influx in F11 cells was measured on a FLIPR instrument incubated in the presence of venom from spitting species (2.4 or 4.8 μg of venom in the case of *H. haemachatus* and *N. philippinensis*) and in the presence or absence of varespladib (13 μM). The data displayed represent the percentage of venom-only cell activation stimulated by treatment with venom and varespladib. Error bars represent SEM, and box midlines indicate median values.

eyes. All cobra venoms activated sensory neurons, with an observed mechanism of action consistent with nonspecific disruption of cell membranes, although activity was limited in African nonspitting cobras and the Asian nonspitter *Naja kaouthia* (fig. S7). Half-maximal effective concentrations (EC_{50}) of each venom in sensory neuron–derived F11 cells demonstrated significantly higher activity in spitting cobra venoms (PGLS; $t = -4.48$, df = 15, $P = 0.0007$) (Fig. 2A, fig. S8, and table S2). These findings support the hypothesis that venom spitting is associated with convergent elevations in venom-induced activation of mammalian sensory neurons, and that spitting cobra venoms are more effective in causing pain than their nonspitting counterparts.

To determine the toxins responsible for this effect, we repeated these experiments using fractionated venom from three representative spitting species (*N. nigricollis*, African; *N. siamensis*, Asian; and *H. haemachatus*) (fig. S9). For each species, only fractions corresponding to CTXs activated sensory neurons, while those corresponding to other toxins (e.g., neurotoxins, PLA_2s, etc.) were inactive (fig. S9). However, none of the CTX fractions completely recapitulated the effects of whole venom or re-pooled venom fractions, suggesting synergy between multiple venom components.

PLA_2s are nearly ubiquitous, typically enzymatic, multifunctional toxin components of snake venoms (*14*, *15*). As the hemolytic activity of CTXs was previously shown to be potentiated by PLA_2 toxins (*17*), we hypothesized that venom PLA_2s potentiate sensory neuron activation by CTXs. Consequently, we quantified the activation of sensory neurons stimulated by CTXs in the presence or absence of a corresponding PLA_2 fraction. The proportion of viable sensory neurons activated by each CTX fraction was significantly increased when combined with a PLA_2 fraction, and this result was consistent across representatives of the three spitting lineages (unpaired *t* test; *N. nigricollis*, $t = 18.77$, df = 2, $P = 0.003$; *N. siamensis*, $t = 5.75$, df = 4, $P = 0.005$; *H. haemachatus*, $t = 4.18$, df = 4, $P = 0.01$) (Fig. 2B, and fig. S10). Moreover, significant reductions in sensory neuron activation occurred in the presence of the PLA_2 inhibitor varespladib (unpaired *t* test; $t = 2.77$, df = 14, $P = 0.02$) (Fig. 2C and fig. S11), providing further compelling evidence that PLA_2s potentiate CTX effects on sensory neurons.

Consistent with the above findings, comparative analysis of the (i) proteomic abundance of PLA_2 toxins and (ii) enzymatic PLA_2 activity, determined via specific in vitro colorimetric assay, revealed that spitting cobra venoms have significantly higher PLA_2 abundance and activity than those of nonspitting species (PGLS; $t = 4.24$, df = 15, $P = 0.0007$, and $t = 2.24$, df = 15, $P = 0.04$, respectively) (Fig. 3, A and B; fig. S12; and table S2). Analysis of a PLA_2 Euclidean distance matrix derived from venom proteomic data revealed substantial differences between the different lineages of spitting and nonspitting cobras, particularly among the African species (Fig. 3C). Additionally, despite their divergence ~17 Ma, *H. haemachatus* PLA_2s clustered tightly with those from African spitting cobras, suggesting an element of molecular convergence. Interrogation of venom gland transcriptomic data revealed limited variation in PLA_2 gene number across the three spitting lineages, although

all exhibited increased PLA_2 abundance compared with nonspitting cobras (fig. S13A). Notably, phylogenetic analyses revealed a PLA_2 gene duplication event that coincided with the origin of venom spitting in the ancestor of African spitting cobras (fig. S13B). These data, alongside the sensory neuron assays, demonstrate that independent evolution of spitting is tightly linked with convergent increases in PLA_2 toxins, which cause increased algesic activity. Our findings therefore imply that evolution has funneled defensive venom phenotypes along repeatable and predictable pathways, although different molecular mechanisms likely underpin this convergence.

To exclude the possibility that functional distinctions simply reflect general differences in venom potency (e.g., venom lethality for prey capture), we tested our venoms in murine lethality assays (table S5) and found no significant differences between spitting and nonspitting species (PGLS; t = 0.86, df = 15, P = 0.40) (table S2). These results suggest that enhanced pain caused by spitting cobras is explicitly associated with defensive venom use rather than being an evolutionary by-product of selection for prey subjugation. Moreover, reanalysis of our data with the Asian species *N. kaouthia* and *N. atra* scored as spitting cobras [on the basis of recent reports (*18*, *19*)], did not alter our key findings relating to venom spitting and PLA_2-mediated enhanced activation of sensory neurons (table S6).

Our results detail the molecular and functional correlates of the evolution of venom spitting and demonstrate that defense can be a major driver of snake venom composition. Spitting likely only evolved within a single, relatively small clade of elapid snakes owing to the integrated exaptation of a distinctive combination of preexisting behaviors and cytotoxic venom activities. Early evolution of cytotoxic venom activity in cobras and near relatives (~26 Ma; fig. S4) has previously been linked to defense, as cytotoxicity co-originates with "hooding" (*5*), a long-distance visual aposematic display. This elevated posture directed toward predators, coupled with preemptive striking and occasional premature releases of venom, may have provided a behavioral precursor for the evolution of more-targeted venom spitting. Preexisting CTXs, largely absent from the other elapid venoms (table S1), likely provided the baseline ocular toxicity that favored inception and retention of spitting (*19*). Subsequent independent increases in PLA_2 toxins, which act in synergy with preexisting CTXs, resulted in increased venom-induced activation of nociceptors (Fig. 2). This potentiating effect of PLA_2s may be crucial for causing immediate pain of sufficient intensity to rapidly deter aggressors, allowing the snake to escape.

Rare but repeatedly evolved adaptations likely result from similar ecological circumstances. Most discussions of defensive behavior involve potential predators, and certain mammals and birds commonly eat snakes (*20*, *21*). However, predation on spitting cobras is evidently unremarkable in terms of evolutionary history and biogeography. Beyond predators, potential threats to snakes include inadvertent trampling and preemptive defensive killing. That spitting evolved to prevent snakes from being trampled by ungulates in African savannas (*22*) does not explain the existence of primarily forest-dwelling Asian spitting cobras (*23*). Moreover, large ungulates typically have lateral eyes, making them unlikely to be especially vulnerable to spitting.

Several considerations make ancient hominins a plausible and compelling candidate for favoring repeated evolution of spitting in the Afro-Asian cobras: (i) Growing evidence suggests that snakes have influenced primate neurobiology and behavior (*24*) and that interactions between these two lineages have been important throughout the 75-million-year history of primates (*25*, *26*). (ii) Compared with carnivorous mammals, anthropoid primates as a clade are visually acute, cognitively complex, and culturally sophisticated (*24*). (iii) Diverse anthropoids mob snakes, with some distinguishing between harmless and dangerous species, killing the latter from a distance with clubs or projectiles (*25*–*27*). (iv) The previous two characteristics are enhanced among bipedal, larger-brained hominins (*24*, *26*), which thus could have posed a considerable threat to snakes (*28*). The initial divergence of African spitting cobras as recently as 6.7 Ma (Fig. 1) occurred soon after the divergence of hominins from *Pan* (bonobos and chimpanzees) ~7 Ma (fig. S14), coinciding with early evolution of bipedalism, enlarged brains, tool use, and occupation of savannas by the former (*29*). Likewise, the origin of the Asian spitter clade ~2.5 Ma is approximately contemporaneous with the arrival in Asia of *Homo erectus* (*30*, *31*) (fig. S14). Although based on circumstantial

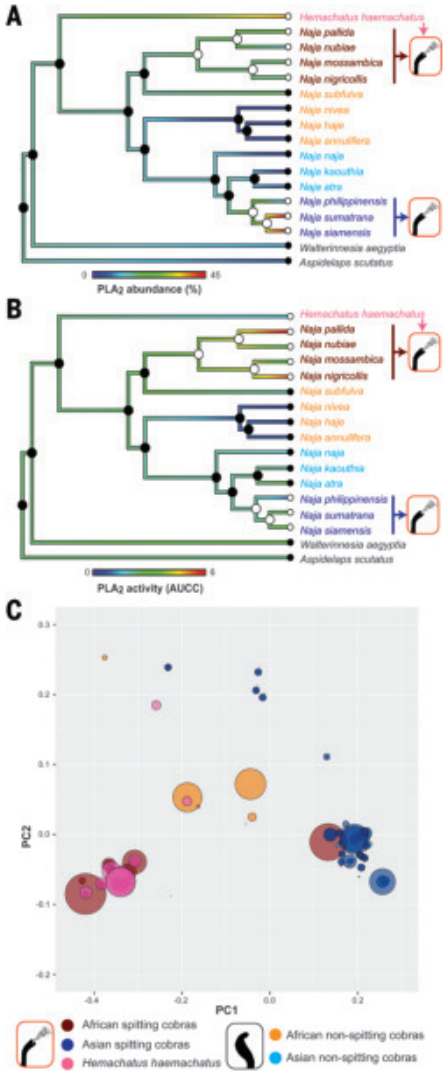


Fig. 3. The abundance, enzymatic activity, and diversity of PLA_2 toxins are associated with convergent evolution of venom spitting. (**A**) Ancestral state estimation of PLA_2 proteomic abundance, expressed as percentage of all toxins in venom proteomes, revealed a significant association with venom spitting (PGLS, t = 4.24, P = 0.0007). Colored branches are scaled according to PLA_2 abundance (blue, low abundance; red, high abundance), filled or empty circles at nodes/tips represent estimated ancestral states of nonspitting or spitting, respectively, and colored tip labels correspond to the different lineages. (**B**) Ancestral state estimation of enzymatic PLA_2 activity, expressed as area under the curve of concentration curves (AUCC) from kinetic in vitro colorimetric assay, revealed a significant association with venom spitting (PGLS, t = 2.24, P = 0.04). Colored branches are scaled according to PLA_2 activity (blue, low activity; red, high activity). Labels as in (A), and see fig. S12 for PLA_2 activity concentration curves. (**C**) PCoA of cobra (*Naja* spp.) and rinkhals (*H. haemachatus*) PLA_2 toxins derived from top-down venom proteomics reveals major variation between African spitting and nonspitting lineages but little variation between Asian cobras. Note the convergent placement of *Hemachatus* PLA_2 toxins with those of African spitting cobras. Circle sizes reflect relative abundances of PLA_2s detected in the venom proteomes.

evidence, additional fossils and more finely tuned dating of relevant cobra and primate divergences might allow further testing of this hypothesis.

Spitting cobras highlight how similar selection pressures and shared exaptations can drive convergence at molecular, morphological, behavioral, and functional levels. The outcome of these processes has resulted in the evolution of complex, integrated adaptations in ecologically important traits being funneled down repeatable pathways.

REFERENCES AND NOTES

1. J. Losos, *Improbable Destinies: How Predictable Is Evolution?* (Penguin Random House, 2017).
2. J. M. Gutiérrez *et al*., *Nat. Rev. Dis. Primers* **3**, 17063 (2017).
3. J. C. Daltry, W. Wüster, R. S. Thorpe, *Nature* **379**, 537–540 (1996).
4. H. Ward-Smith, K. Arbuckle, A. Naude, W. Wüster, *Toxins* **12**, 201 (2020).
5. N. Panagides *et al*., *Toxins* **9**, 103 (2017).
6. W. Wüster *et al*., *Mol. Phylogenet. Evol.* **45**, 437–453 (2007).
7. C. M. Bogert, *Bull. Am. Mus. Nat. Hist.* **81**, 285–360 (1943).
8. S. Rasmussen, B. Young, H. Krimm, *J. Zool.* **237**, 27–35 (1995).
9. G. Westhoff, M. Boetig, H. Bleckmann, B. A. Young, *J. Exp. Biol.* **213**, 1797–1802 (2010).
10. D. A. Warrell, L. D. Ormerod, *Am. J. Trop. Med. Hyg.* **25**, 525–529 (1976).
11. S. Ainsworth *et al*., *J. Proteomics* **172**, 173–189 (2018).
12. G. Watt, L. Padre, M. L. Tuazon, R. D. G. Theakston, L. Laughlin, *Am. J. Trop. Med. Hyg.* **39**, 306–311 (1988).
13. J. Meier, J. White, Eds., *Handbook of Clinical Toxicology of Animal Venoms and Poisons* (Informa Healthcare USA, Inc., 2008).
14. T. Tasoulis, G. K. Isbister, *Toxins* **9**, 290 (2017).
15. C. R. Ferraz *et al*., *Front. Ecol. Evol.* **7**, 218 (2019).
16. G. Whiteley *et al*., *J. Proteomics* **198**, 186–198 (2019).
17. E. Condrea, *Experientia* **30**, 121–129 (1974).
18. A. Paterna, *Herpetol. Bull.* **148**, 22–25 (2019).
19. V. Santra, W. Wüster, *Herpetol. Rev.* **48**, 455–456 (2017).
20. H. W. Greene, in *Biology of the Reptilia. Vol. 16: Ecology B: Defense and Life History*, C. Gans, R. B. Huey, Eds. (Alan R. Liss, Inc., 1988), pp. 1–152.
21. B. Van Valkenburgh, R. K. Wayne, *Curr. Biol.* **20**, R915–R919 (2010).
22. T. Barbour, *Copeia* **106**, 26–28 (1922).
23. E. R. Chu, S. A. Weinstein, J. White, D. A. Warrell, *Toxicon* **56**, 259–272 (2010).
24. L. A. Isbell, *The Fruit, the Tree, and the Serpent: Why We See So Well* (Harvard Univ. Press, 2009).
25. T. N. Headland, H. W. Greene, *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1470–E1474 (2011).
26. H. W. Greene, *Tracks and Shadows: Field Biology as Art* (Univ. of California Press, 2013).
27. H. W. Greene, *Am. Nat.* **190**, S69–S86 (2017).
28. S. Faurby, D. Silvestro, L. Werdelin, A. Antonelli, *Ecol. Lett.* **23**, 537–544 (2020).
29. L. Pozzi *et al*., *Mol. Phylogenet. Evol.* **75**, 165–183 (2014).
30. S. Prat, *C. R. Palevol* **17**, 6–16 (2018).
31. F. Han *et al*., *Quat. Int.* **434**, 75–83 (2017).

ACKNOWLEDGMENTS

The authors thank P. Rowley, E. Crittenden, and W. Grail for technical support and NVIDIA Corporation for donation of the TITAN-X GPU used for BEAST analyses. **Funding:** This work was funded from a studentship supported by Elizabeth Artin Kazandjian to T.D.K.; grant PE 2600/1 from the German Research Foundation (DFG) to D.P.; grant OPUS 1354156 from the U.S. National Science Foundation to H.W.G.; grants FAPESP 2017/18922-2 and 2019/05026-4 from the São Paulo Research Foundation to R.R.d.S.; grants RPG-2012-627 and RFG-10193 from the Leverhulme Trust to R.A.H. and W.W.; grant MR/L01839X/1 from the U.K. Medical Research Council to J.M.G., R.A.H., J.J.C., and N.R.C.; Career Development Fellowships APP1162503 and DP190103787 from the Australian National Health and Medical Research Council to I.V.; fellowship DE160101142 and grant DP160104025 from the Australian Research Council, and fellowship FRIPRO-YRT #287462 from the Research Council of Norway to E.A.B.U.; and a Sir Henry Dale Fellowship (200517/Z/16/Z) jointly funded by the Wellcome Trust and Royal Society to N.R.C. **Author contributions:** R.A.H., W.W., and N.R.C. conceived of the research. T.D.K., D.P., K.A., M.K.R., R.A.H., J.J.C., I.V., E.A.B.U., W.W., and N.R.C. designed the research. T.D.K., D.P., S.D.R., J.v.T., A.B., D.A.C., R.M.W., G.W., S.C.W., A.S.A., L.-O.A., A.P.L., C.H., A.H., S.P.-L., C.V.M., S.A., R.R.d.S., P.C.D., J.M.G., J.J.C., I.V., E.A.B.U., W.W., and N.R.C. performed the research. T.D.K. and N.R.C. wrote the manuscript with major input from D.P., S.D.R., H.W.G., I.V., E.A.B.U., and W.W. All authors discussed and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The molecular data associated with species tree generation have been deposited to the nucleotide database of NCBI and accession numbers are displayed in table S7. The transcriptome data have been deposited in the SRA and TSA databases of NCBI and are associated with the BioProject accession number PRJA506018. Mass spectrometry data and database search results for top-down and bottom-up proteomic experiments are publicly available in the MassIVE repository under accession number MSV000081885 and in ProteomeXchange with accession number PXD008597.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/386/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S10
References (*32*–*126*)
MDAR Reproducibility Checklist
Data S1 to S5

1 April 2020; accepted 7 December 2020
10.1126/science.abb9303

SOLAR CELLS

Nanoscale localized contacts for high fill factors in polymer-passivated perovskite solar cells

Jun Peng[1]*, Daniel Walter[1]*, Yuhao Ren[2], Mike Tebyetekerwa[1], Yiliang Wu[1], The Duong[1], Qiaoling Lin[2], Juntao Li[2]†, Teng Lu[3], Md Arafat Mahmud[1], Olivier Lee Cheong Lem[4], Shenyou Zhao[1], Wenzhu Liu[5], Yun Liu[3], Heping Shen[1], Li Li[4], Felipe Kremer[6], Hieu T. Nguyen[1], Duk-Yong Choi[7], Klaus J. Weber[1], Kylie R. Catchpole[1]†, Thomas P. White[1]†

Polymer passivation layers can improve the open-circuit voltage of perovskite solar cells when inserted at the perovskite–charge transport layer interfaces. Unfortunately, many such layers are poor conductors, leading to a trade-off between passivation quality (voltage) and series resistance (fill factor, FF). Here, we introduce a nanopatterned electron transport layer that overcomes this trade-off by modifying the spatial distribution of the passivation layer to form nanoscale localized charge transport pathways through an otherwise passivated interface, thereby providing both effective passivation and excellent charge extraction. By combining the nanopatterned electron transport layer with a dopant-free hole transport layer, we achieved a certified power conversion efficiency of 21.6% for a 1-square-centimeter cell with FF of 0.839, and demonstrate an encapsulated cell that retains ~91.7% of its initial efficiency after 1000 hours of damp heat exposure.

The rise in perovskite solar cell (PSC) performance has been marked by substantial increases in open-circuit voltage (V_{oc}), short-circuit current (J_{sc}), and fill factor (FF), as illustrated in Fig. 1 for reported values since 2014 (*1*–*9*). The steady increase in V_{oc} from ~1 to ~1.2 V (for simplicity, we ignore bandgap differences between the record cells) reflects efforts to reduce nonradiative recombination, especially at the perovskite–charge transport layer interfaces (*10*–*25*). For example, insertion of ultrathin passivation layers at one or both interfaces can lead to voltage gains of >100 mV in some cases (*11*, *22*). Reported passivation materials include insulating (*10*–*14*) and semiconducting (*15*–*17*) organic materials, inorganic dielectrics (*18*, *19*), and low-dimensional perovskites (*20*–*25*). Insulating organic materials and low-dimensional perovskites have provided some of the largest efficiency gains.

Two other important trends can be seen in Fig. 1. First, the J_{sc} of record cells has not increased substantially since 2015; a consequence of the near-unity internal quantum efficiency and low optical losses of state-of-the-art cells. Second, record FF values plateaued between 2015 and 2018, before a steep increase in 2019, despite the almost linear improvement in V_{oc} during the same period. Given that FF quantifies voltage and current losses at the maximum power point relative to the maximum

[1]Research School of Electrical, Energy and Materials Engineering, The Australian National University, Canberra Australian Capital Territory 2600, Australia. [2]State Key Laboratory of Optoelectronics Materials and Technologies, School of Physics, Sun Yat-sen University, Guangzhou 510275, China. [3]Reseach School of Chemistry, The Australian National University, Canberra Australian Capital Territory 2600, Australia. [4]Australian National Fabrication Facility, Research School of Physics, The Australian National University, Canberra Australian Capital Territory 2600, Australia. [5]Research Center for New Energy Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Jiading, Shanghai 201800, China. [6]Centre for Advanced Microscopy, The Australian National University, Canberra Australian Capital Territory 2600, Australia. [7]Australian National Fabrication Facility, The Australian National University, Canberra Australian Capital Territory 2600, Australia.
*These authors contributed equally to this work.
†Corresponding author. Email: lijt3@mail.sysu.edu.cn (J.L.); kylie.catchpole@anu.edu.au (K.R.C.); thomas.white@anu.edu.au (T.P.W.)

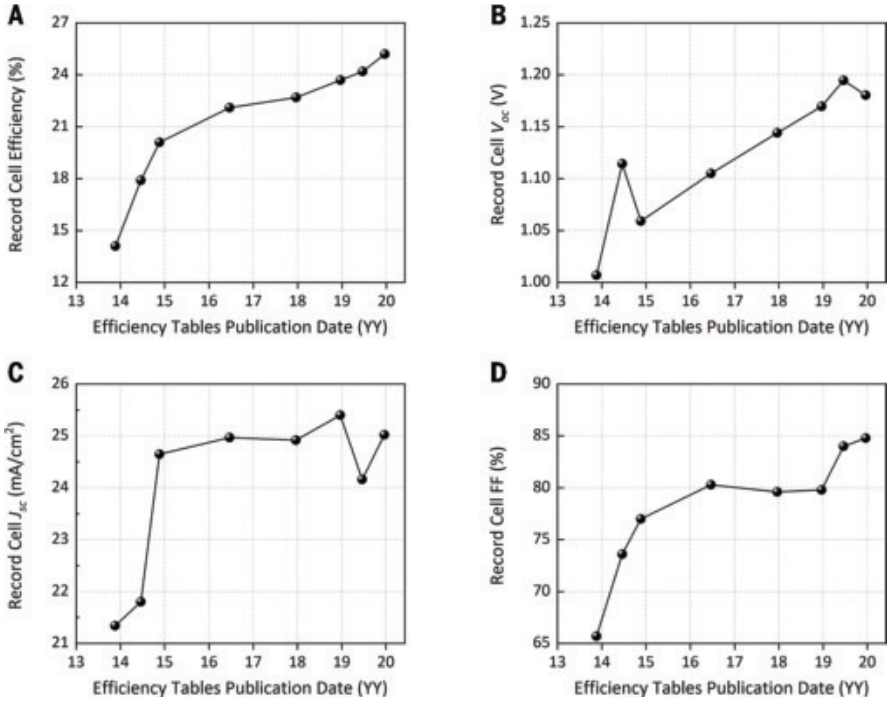


Fig. 1. Certified performance of record PSCs. (**A**) Record cell efficiency. (**B**) Record cell V_{oc}. (**C**) Record cell J_{sc}. (**D**) Record cell FF. Note that the effective area of all record cells plotted here was <0.2 cm^2.

values of both V_{oc} and J_{sc}, the theoretical FF limit is a function of V_{oc} and depends both on the active layer bandgap and the rate of nonradiative recombination that degrades V_{oc}. In practice, however, many cells perform well below this limit because of parasitic resistive losses (shunt and series). Thus, achieving a very high FF requires low recombination losses (high V_{oc}); smooth, uniform, and dense active layers (high shunt resistance); and very efficient charge extraction and transport (low series resistance). The lack of correlation between V_{oc} and FF in Fig. 1 suggests that the FF of record cells is not limited by recombination but rather by resistive losses. Indeed, the theoretical maximum FF for a solar cell with V_{oc} = 1.1 V is ~0.89, well above the highest values achieved so far.

Maximizing both V_{oc} and FF is a challenge faced by all high-efficiency solar cells because of the difficulties in reducing interface recombination without hindering charge extraction. This challenge is most apparent in the case of insulating and dielectric passivation layers that provide outstanding passivation of defects [e.g., SiO_2 passivation layers in Si cells (*26*) or poly(methyl methacrylate) (PMMA) passivation of perovskite cells (*10*)] but must be extremely thin to avoid detrimental FF losses from series resistance. As an extreme example, silicon solar cells with tunnel oxide passivated contacts can experience a 70% drop in FF for a 5-Å variation in oxide thickness (*27*). There are many examples for PSCs showing trade-offs between passivation quality (V_{oc}) and FF, not only for ultrathin layers of insulating polymers such as polystyrene and PMMA (*10*–*12*), but also for poorly conducting, low-dimensional perovskite passivation layers (*20*–*25*). Overcoming this trade-off with new materials or device architectures could improve PSC performance.

A further challenge facing perovskite technology is maintaining high efficiency, specifically FF, on larger cell areas (*5*, *6*, *28*, *29*). The record efficiencies quoted above are for very small cells (~0.1 cm^2) that are at least three orders of magnitude smaller than commercial Si cells. Increasing cell area increases the possibility of shunts created by film imperfections, whereas series resistance losses scale quadratically with current. Both contributions can rapidly degrade FF as cell areas increase, with record 1-cm^2 cells typically exhibiting FF values between 5 and 10% lower than their smaller counterparts (*5*).

In this work, we present a PSC architecture in which we replaced the commonly used mesoporous TiO_2 (meso-TiO_2) electron transport layer (ETL) with a sparse array of nanoscale TiO_2 cylinders (nanorods). The resulting nanopatterned ETL-perovskite interface could be effectively passivated with an ultrathin polymer passivation layer to achieve high V_{oc} that maintained outstanding charge collection and interfacial transport properties; this resulted in low series resistance and a high FF. This approach was demonstrated on a large-area (1-cm^2) cell with a certified power conversion efficiency (PCE) of 21.6% and FF = 0.839. A champion small-area (~0.165-cm^2) cell achieved a PCE of ~23.17% with V_{oc} = 1.240 V and FF = 0.845. We investigated and explained the performance of the nanostructured interface using three-dimensional (3D) numerical simulations that replicated accurately the FF enhancement and the observed experimental trends with nanopattern geometry. The simulations strongly suggested that the TiO_2 nanorods are incompletely coated with the PMMA:PCBM (phenyl-C_{61}-butyric acid methyl ester) passivation material, leaving exposed regions of local low-resistance contact directly analogous to local contact structures in high-efficiency silicon solar cells (*30*, *31*). Detailed comparisons between simulation and experiment placed plausible bounds on the nanorod area fraction that was left exposed and the density of recombination-active defects at this exposed ETL-perovskite interface. In addition to the nanopatterned ETL, we also introduce here an ionic dopant–free, blended hole transport layer (HTL) that offers similar cell performance to doped alternatives such as 2,2′,7,7′-tetrakis-(*N*, *N*-di-4-methoxyphenylamino)-9,9′-spirobifluorene (Spiro-OMeTAD) and poly(triaryl amine) (PTAA) but with much improved thermal stability. Encapsulated cells incorporating the new ETL and HTL combination retained >90% of their initial efficiency after 1000 hours of exposure to damp heat conditions of 85°C and 85% relative humidity.

Spiro-OMeTAD and PTAA are the most widely used HTL materials for high-efficiency n-i-p architecture PSCs. These layers are typically doped with bis(trifluoromethane)sulfonamide lithium salt (Li-TFSI) to improve the conductivity, but the hygroscopic Li-TFSI can accelerate moisture-induced perovskite degradation, and diffusion of Li ions throughout the cell further reduces performance and stability (*6*, *32*, *33*). These detrimental effects are exacerbated by high temperatures, contributing to the poor thermal stability of many perovskite cells. Alternative HTL layers that offer similar efficiency performance without affecting stability include poly(3-hexylthiophene) (P3HT), a semiconducting polymer used extensively in organic optoelectronics but only recently reported as an effective HTL for PSCs (*6*, *34*, *35*). Controlling the morphology and crystallinity of the P3HT film is essential for achieving a high charge carrier mobility without dopants. The hole mobility can range from ~10^{-5} cm^2 V^{-1} s^{-1} in an amorphous film up to ~0.1 cm^2 V^{-1} s^{-1} for a crystalline film with optimized π-π stacking of the polymer side chains (*6*, *35*). We show that blending P3HT with another thermally stable HTL material, copper phthalocyanine (CuPc), followed by a postdeposition solvent-annealing

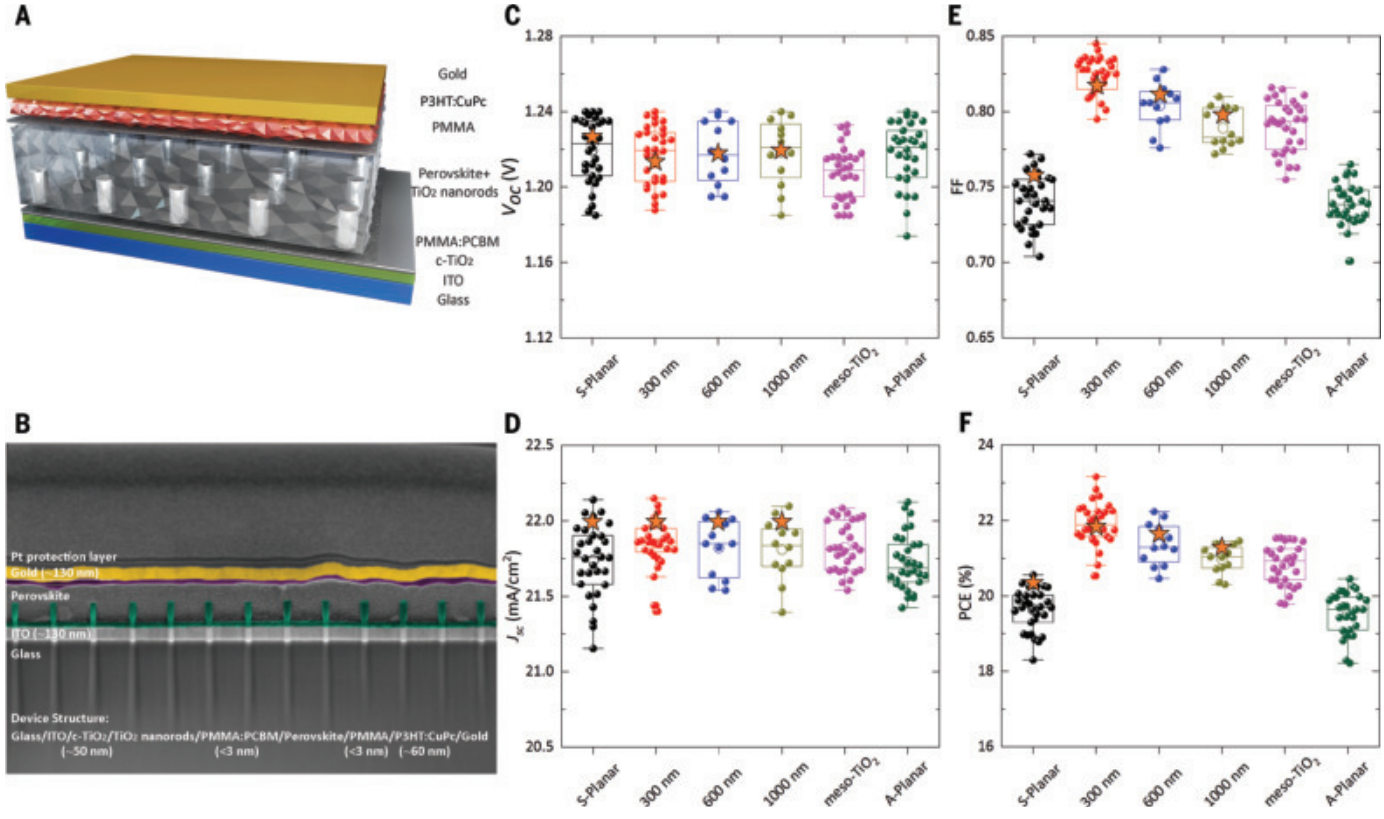


Fig. 2. Nanostructured TiO_2 ETL and statistical distribution of the photovoltaic parameters for nanopattern and control perovskite cells. (**A**) Schematic of nanopattern perovskite cell studied in this work. (**B**) Cross-sectional SEM image of the nanopattern cell structure [note that the perovskite layer was fabricated using ~1 mol/ml diluted precursor solution instead of the standard ~1.3 mol/ml (for cells); the Pt protection layer was only used to prepare the focused ion beam SEM cross-sectional image]. (**C**) Distribution of V_{oc}. (**D**) Distribution of J_{sc}. (**E**) Distribution of FF. (**F**) Distribution of PCE. Note that the control cells are S-Planar, meso-TiO_2, and A-Planar; the nanopattern cells are 300, 600, and 1000 nm. S-Planar represents planar cells based on solution-processed compact TiO_2 ETLs. meso-TiO_2 represents mesoporous cells based on solution-processed compact TiO_2/mesoporous TiO_2 ETLs. A-Planar represents planar cells based on ALD-deposited compact TiO_2 ETLs. "300 nm," "600 nm," and "1000 nm" represent the nanopattern cells based on nanostructured TiO_2 nanorods with a spacing of 300, 600, and 1000 nm, respectively. The orange star icons indicate the values calculated from the 3D numerical simulations.

step produced highly crystalline films with excellent hole transport characteristics.

Details of the HTL optimization process and film characterization are provided in the supplementary materials (see figs. S1 to S6). To summarize the key results, x-ray diffraction (XRD) analysis revealed that applying a solvent-annealing treatment with chlorobenzene to pure P3HT films improved crystallinity and conductivity compared with untreated films, and both properties were further improved by blending P3HT with CuPc (fig. S1). When incorporated into a baseline meso-TiO_2 PSC process as a substitute for Spiro-OMeTAD (fig. S4 shows the device structure), the best results were obtained for a 10:1 (v/v) ratio blend of P3HT (12 mg/ml):CuPc (5 mg/ml) (fig. S5). The average FF for cells with the optimized P3HT:CuPc blend was 0.790 ± 0.016, compared with 0.728 ± 0.016 for as-deposited P3HT films, and the average PCE for the optimized cells was 20.83 ± 0.55% (fig. S6). This performance compared favorably to conventional meso-TiO_2–based cells using Spiro-OMeTAD or PTAA as the HTL, as reported previously (*10*, *12*, *24*, *25*).

For electron extraction, the most widely used ETL configuration for n-i-p PSCs consists of a thin (~50 nm), compact TiO_2 layer (c-TiO_2) covered by a meso-TiO_2 layer of typical thickness >100 nm. Dense random arrays of TiO_2 nanowires grown by solution methods have also been explored as an alternative to meso-TiO_2 but with negligible performance benefit (*36*–*38*). The main role of the meso-TiO_2 (or random nanowire array) layer is to provide a scaffold for the perovskite film to form a dense, pinhole-free film. These layers may also contribute to charge carrier collection, but the large contact area between the meso-TiO_2 (or random nanowire array) and the perovskite can also result in substantial interface recombination. Therefore, some form of passivation is required to achieve high voltages. A thin passivating interlayer of PMMA:PCBM can effectively increase the V_{oc} of meso-TiO_2–based cells (*10*, *12*, *39*, *40*), but this layer must be ultrathin to limit resistive losses.

The average performance of meso-TiO_2–based cells with the structure glass/ITO/c-TiO_2/meso-TiO_2/PMMA:PCBM/$Cs_{0.05}FA_{0.88}MA_{0.07}PbI_{2.56}Br_{0.44}$/PMMA/P3HT:CuPc/Gold, where the ratio of PMMA:PCBM mixed precursor solution is 1:3 (w/w), is shown in Fig. 2 and replotted in fig. S6 for the optimized P3HT:CuPc blend discussed above (referred to as meso-TiO_2 in the remainder of this work). Here, ITO refers to indium-doped tin oxide, and MA and FA are methylammonium and formamidinium, respectively. The band gap of the perovskite we used throughout was ~1.6 eV (fig. S7). We also fabricated planar n-i-p cells with just the c-TiO_2 ETL layer, omitting the meso-TiO_2 layer altogether. Cell data are shown in Fig. 2 for planar cells fabricated with a solution-processed c-TiO_2 layer ("S-Planar") and an atomic-layer deposition (ALD) c-TiO_2 layer ("A-Planar"), with all other layers identical to the meso-TiO_2 cell. Note that the average V_{oc} of the planar cells is slightly higher than that of the meso-TiO_2 cells (Fig. 2C) and is likely the result of the smaller interface area over which recombination can occur. The reduced FF in the planar c-TiO_2 case (Fig. 2E), however, is likely the result of a more uniform passivation layer compared with the meso-TiO_2 ETL. The higher FF of the

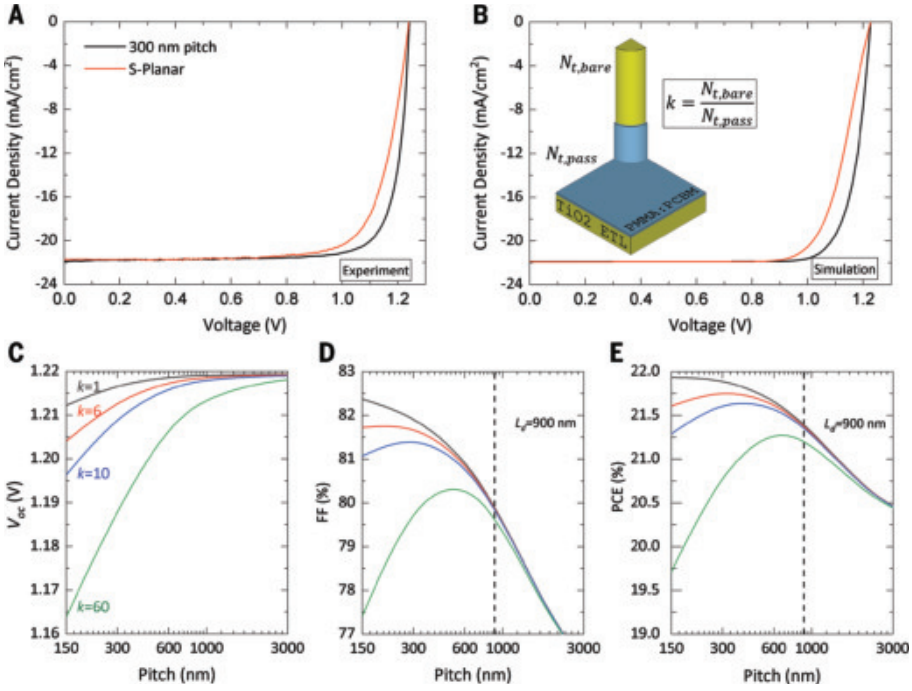


Fig. 3. Simulation results. (**A**) Experimental *J-V* curves of the S-Planar (reference cell) and the 300-nm pitch, representing midperformance cells from each batch. (**B**) Simulated *J-V* curves for the planar and nanostructured cell in which we presume that only 30% of the nanorod is coated with the insulating passivation layer (inset). The periodic separation of the nanorods, or pitch, governs V_{oc} (**C**), FF (**D**), and thus overall efficiency (**E**). The unpassivated and passivated ETL surfaces are characterized by defect concentrations, $N_{t,bare}$ and $N_{t,pass}$, respectively. (C) to (E) plot the dependence of cell performance on the ratio of unpassivated to passivated defect density, k, for CF = 30%, illustrating that the optimum pitch depends on the strength of unpassivated interface recombination. The electron diffusion length, L_d, is marked by vertical dashed lines.

meso-TiO_2 may result from partial or nonuniform coverage of the passivating layer on the textured ETL surface that has thinner regions with more efficient charge conduction through an otherwise poorly conducting film.

Comparison between the planar c-TiO_2 and the meso-TiO_2 ETL interfaces suggested that there exists a tradeoff between V_{oc} and FF that is mediated by the continuity of coverage of the passivation layer. On the basis of this finding, we developed a nanopatterned ETL consisting of an array of TiO_2 nanorods on a planar c-TiO_2 layer. This optimized texture enables the high V_{oc} of the planar architecture while providing efficient charge extraction and thus high FF. The ETL fabrication process is illustrated in fig. S8. Starting with a glass/ITO/c-TiO_2 substrate, we first deposited a PMMA resist layer by spin coating, followed by a 15-min annealing step. The nanorod array was defined in the resist using electron-beam lithography (EBL), and the pattern was developed in a solution of methyl isobutyl ketone (MIBK) and isopropyl alcohol (IPA). The entire sample was then coated with an ALD-deposited conformal TiO_2 layer that filled the holes in the PMMA resist and covered the resist between the holes. Finally, the TiO_2-capping layer was etched with inductively coupled plasma-reactive ion etching (ICP-RIE) until the resist surface was exposed, and the remaining resist was removed by oxygen plasma. After etching, the TiO_2 nanorod substrates were annealed at 500°C for 30 min under dry air before further processing, converting the as-deposited amorphous TiO_2 into a nanocrystalline anatase material.

Although the diameter, pitch (spacing), and height of the nanorods could all influence the cell performance, we set the diameter to 50 nm and the height to 200 nm and focused on investigating the impact of pitch. This approach was motivated by device simulations (described below) that predicted the largest changes in device performance through modification of local contact spacing. Figures S9 and S10 show the nanopatterned TiO_2 ETL with pitches of 300, 600, and 1000 nm. The nanopatterned perovskite cell structure was ITO/c-TiO_2/TiO_2 nanorods/PMMA:PCBM/$Cs_{0.05}FA_{0.88}MA_{0.07}PbI_{2.56}Br_{0.44}$/PMMA/P3HT:CuPc/Gold (Fig. 2, A and B).

To evaluate the performance enhancement provided by the nanopatterned ETL, in Fig. 2, we compare the nanopatterned cells with different pitches with three control cell architectures, meso-TiO_2, A-Planar, and S-Planar. The two different planar cell types were included to compare the solution and ALD-deposited materials because the nanorods were ALD deposited and both compact layers for the meso-TiO_2 and nanopatterned cells were solution deposited; the deposition method had a negligible impact on the cell performance. All of the cells include double-side passivation with ultrathin films of PMMA and PMMA:PCBM blends at the perovskite-HTL and perovskite-ETL interfaces, respectively (*10*, *11*).

All six cell architectures demonstrated excellent surface passivation quality (Fig. 2C), with an average V_{oc} > 1.2 V and maximum V_{oc} ~1.24 V. The nanopatterned cells achieved V_{oc} values equivalent to those of the planar cells for 300-, 600-, and 1000-nm nanorod pitches, with average V_{oc} values of 1.217 ± 0.015, 1.218 ± 0.017, and 1.218 ± 0.018 V, respectively, showing that the nanorods did not affect the passivation quality compared with planar cells. Also, J_{sc} was similar for all six cell architectures (Fig. 2D), with average values ranging from 21.73 ± 0.19 mA/cm^2 (A-Planar) to 21.85 ± 0.17 mA/cm^2 (300 nm).

A comparison of representative current-voltage (*J-V*) curves for planar and nanopatterned cells shows that the FF increase with nanorods was dominated by reduced series resistance (Fig. 3A). The planar cells had the lowest average FFs of 0.741 ± 0.017 (S-Planar) and 0.739 ± 0.013 (A-Planar), whereas the meso-TiO_2 cells had FF = 0.790 ± 0.017 (Fig. 2E). By contrast, the optimum nanopatterned cells with 300-nm pitch had an average FF = 0.824 ± 0.012, which decreased to 0.803 ± 0.015 for 600-nm pitch and 0.789 ± 0.013 for 1000-nm pitch. The 300-nm pitch nanopatterned cells had an average PCE of 21.91 ± 0.57%, compared with 19.64 ± 0.55% (19.54 ± 0.56%) for the S-Planar (A-Planar) cells and 20.83 ± 0.55% for the meso-TiO_2 cells (Fig. 2F). The best-performing nanopatterned cell had a PCE of 23.17% with V_{oc} = 1.240 V and FF = 0.845 (fig. S11).

There are several possible explanations for the combination of outstanding passivation (V_{oc}) and charge extraction/transport (FF) provided by the nanopatterned ETL, including: (i) improvements in the optoelectronic quality of the perovskite film when it is deposited on the nanopatterned substrate; (ii) modification of the passivation layer coverage and/or thickness at the ETL-perovskite interface resulting in an improved balance of passivation and charge transport across the interface; and (iii) improved charge collection and interface transport caused by the geometry of the nanorods. For possibility (i), the nanorods could scaffold the growth of the perovskite film, but we found no evidence of morphology changes from scanning electron microscope (SEM) images of perovskite films deposited on planar,

nanopatterned and meso-TiO_2 ETLs (fig. S12), nor did we find appreciable differences in crystal structure, orientation, or crystallinity from XRD measurements of films on the different substrates (fig. S13). Photoluminescence (PL; steady-state and time-resolved) measurements of films deposited on the different ETL structures were consistent with the experimental cell V_{oc} values: The planar and nanopatterned cells showed almost identical steady-state intensity and PL decay curves, whereas the meso-TiO_2 sample had a slightly lower PL intensity and faster decay, indicating increased recombination losses (fig. S14). Thus, the nanopatterned ETL had a minimal impact on the physical and optoelectronic properties of the perovskite film.

To test explanations (ii) and (iii) above, we developed a full 3D numerical drift-diffusion model of the n-i-p PSC including the periodic nanostructured TiO_2 nanorods (fig. S15). The resistive PMMA:PCBM layer at the perovskite-TiO_2 interface was introduced in the model as a 3-nm-thick insulating interlayer. The valence and conduction band energy levels of this interlayer are aligned to those of the perovskite and ETL, respectively, to avoid introducing energy barriers to charge transport. Bulk conductivity of the layer was reduced to match the FF of the S-Planar control samples. Nonradiative recombination at this interface and in the perovskite bulk was modeled with generalized Shockley-Read Hall statistics (*41*), parameterized by self-consistent time-resolved (TR) fitting to TRPL data measured for a bare perovskite film on glass, and a perovskite film deposited on a planar glass/ITO/TiO_2/PMMA:PCBM substrate (fig. S16).

From these fits, we extracted a bulk perovskite defect density of 2×10^{14} cm^{-3} for an effective electron lifetime of ~600 ns. This value corresponds well with recent results from Ni *et al.* (*42*), who extracted bulk defect concentrations of ~10^{14} to 10^{15} cm^{-3} in a range of multicrystalline perovskite films from independent capacitance-based techniques. We fit a charge-carrier mobility of 0.5 $cm^2/V{\cdot}s$ for a corresponding diffusion length of ~900 nm at 1 sun-equivalent injection levels (fig. S17). The fitted surface areal defect density at the PMMA:PCBM–passivated interface was 5×10^8 cm^{-2}, for an equivalent surface-recombination velocity of only 3 cm/s, indicative of the high-quality passivation achieved with this method. These parameters were integrated into the full 3D cell simulations, predicting V_{oc} and FF consistent with the experimental results (fig. S18). Details of the TRPL and cell models and associated simulation parameters are provided in the supplementary materials.

With this model, we initially tested possibility (iii), which presumed that the TiO_2 nanorods increased the effective surface area of the ETL and reduced resistance. Our simulations showed that if the TiO_2 nanorod was conformally coated by the passivation interlayer, then the increased surface area could not account for the enhancement in FF (fig. S19) and actually led to a marginal decrease in V_{oc} (–8 mV) and FF (–0.08%) for the 300 nm compared with the S-Planar cell because of the increase in surface area at which nonradiative recombination can occur. FF did not increase because collection efficiency in the perovskite bulk was near unity as the predicted charge carrier diffusion lengths (900 nm) are more than double the absorber thickness of 400 nm.

For possibility (ii), we presumed that the nanostructure influenced conformity of the PMMA:PCBM layer across the TiO_2-perovskite interface, which was deposited from solution by spin-coating. Some faces of the nanostructured rods likely had thinner or noncontinuous layers of PMMA:PCBM or both, for example, at the radially outward or top faces of the rod. We show this schematically in the inset of Fig. 3B. The uniformity of PMMA:PCBM at the TiO_2 nanorods could not be observed given that the PMMA:PCBM film was <3-nm thick and offered insufficient contrast to observe with electron microscopy. Instead, we simulated imperfect PMMA:PCBM coverage by assuming a coverage fraction (CF) of 30% of the TiO_2 nanorod sidewalls and presumed that the bare TiO_2:perovskite interface had a defect density one order of magnitude higher than the passivated surface at 5×10^9 cm^{-2}. A discussion of these assumptions is provided in the supplementary materials. Under these conditions, we found close agreement between simulation and experimental results for both the magnitude of the FF change between the S-Planar and 300-nm cases (Fig. 3B) and the change in FF with the nanorod pitch, as indicated by the star markers in Fig. 2E. The corresponding V_{oc} is only marginally affected, and overall device efficiency was consistent with experimental results (Fig. 2, C and F).

Our device simulations suggest that the formation of repeatable nanoscale regions of an unpassivated nanorod TiO_2-perovskite interface facilitate localized low-resistance charge transport with limited impact on effective interface recombination and V_{oc}. This localized contact structure approach is analogous to local contact structures used in high-efficiency laboratory and industrial silicon solar cells, in which local openings in electrically passivating and insulating dielectric layers are laser micromachined or chemically etched (*30*, *31*, *43*). In both cases, a global optimum trades off interface recombination (V_{oc}) and charge extraction (FF) (Fig. 3, C and D). When nonradiative recombination was increased at the unpassivated

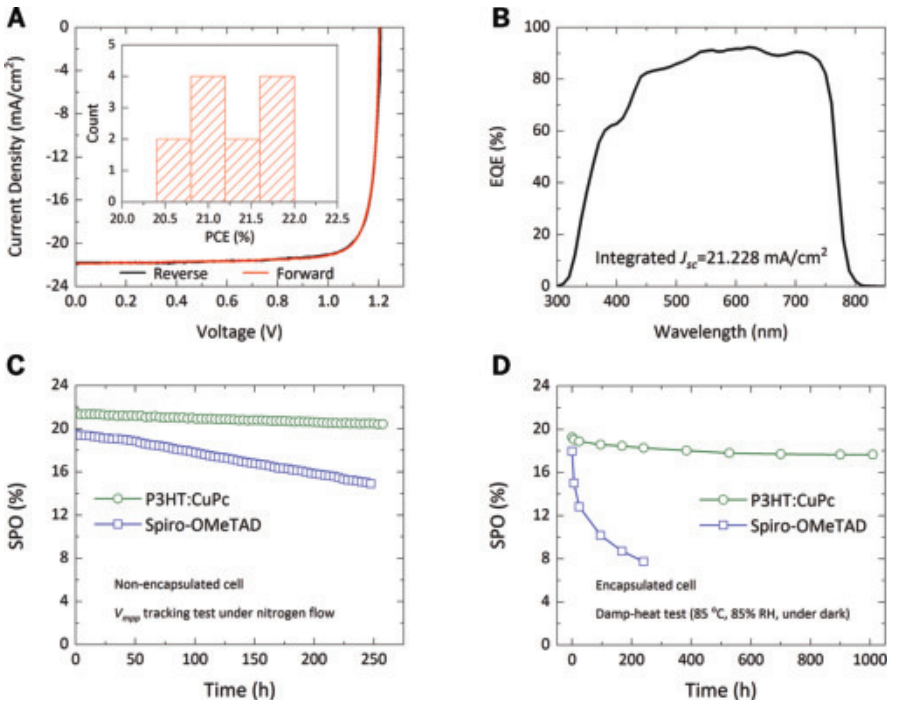


Fig. 4. Device performance and long-term stability tests for 1-cm^2 perovskite cells. (**A**) *J-V* performance. (**B**) EQE measurement. (**C**) SPO of nonencapsulated cells based on P3HT:CuPc and Spiro-OMeTAD HTL measured by V_{mpp} tracking under continuous 1 sun illumination intensity. (**D**) SPO of encapsulated cells based on P3HT:CuPc and Spiro-OMeTAD HTL. Note that the cell structure for the damp heat tests included a MoO_x-ITO barrier between the HTL and the gold electrode: ITO/c-TiO_2/TiO_2 nanorods/PMMA:PCBM/$Cs_{0.05}FA_{0.88}MA_{0.07}PbI_{2.56}Br_{0.44}$/PMMA/P3HT:CuPc (or Spiro-OMeTAD)/MoO_x (~10 nm)/IZO (~40 nm)/gold.

nanorod surfaces, expanding pitch reduced the effective interface recombination and increased V_{oc}. However, FF was maximized at an intermediate pitch. For small pitches, the increased recombination that reduced V_{oc} also reduced FF, whereas larger pitches reduced FF through increased transport resistance in the perovskite bulk. The convolution of FF and V_{oc} across pitches showed that a narrow pitch range optimized performance, which in our parameterization fell near the 300-nm pitch we used (Fig. 3E). Our experimental exploration of the cell structure focused on rod separation because in our device simulations (fig. S20) modification of rod height and diameter led to much smaller increases or, in some cases, decreases of V_{oc} and FF.

The results presented above were all obtained on cells of area 0.165 cm^2, but we also fabricated optimized 300-nm-pitch nanopatterned cells with an effective area of 1.2 cm^2 and *J-V* data measured with an ~1.02-cm^2 aperture mask. Our champion perovskite cell exhibited a PCE of 21.96% (21.86%) with V_{oc} ~ 1.205 V (1.204 V), J_{sc} ~ 21.80 mA/cm^2 (21.86 mA/cm^2), and FF ~ 0.836 (0.836) from reverse (forward) scans measured in-house (Fig. 4A). External quantum efficiency (EQE) spectra confirmed the measured photocurrent with a relative discrepancy of ~3% between the integrated J_{sc} from the EQE and the J_{sc} measured from the *J-V* curve (Fig. 4B). Another cell submitted to an independent testing center achieved a certified efficiency of ~21.698 ± 0.555% with FF of 83.87 ± 1.16% (fig. S21).

The histogram distribution figure (inset, Fig. 4A) presents the reproducibility of the large-area perovskite cells (details of the *J-V* parameters are provided in table S1). The ~30-mV decrease in V_{oc} when we used the aperture to measure the cell compared with the *J-V* performance measured without an aperture (fig. S22) was ascribed to carrier recombination within the cell's dark region shaded by the mask (*44*). Taking this deficit into account, we conclude that there was negligible loss in V_{oc} or FF between cells of area 0.165 and 1 cm^2 despite a >6× increase in photocurrent.

To test the light and thermal stability of the cells, we conducted two extended degradation tests: a steady-state power output (SPO) measurement [unencapsulated cells held at maximum power point voltage (V_{mpp}) under continuous illumination in a N_2 atmosphere] and a damp heat test (encapsulated cells exposed to 85°C and 85% relative humidity in the dark). For the SPO test, Fig. 4C shows that the cell with the P3HT:CuPC HTL retained ~95.3% of its initial efficiency ($PCE_{initial}$ ~ 21.45%) after 260 hours, whereas the control cell with a Spiro-OMeTAD HTL dropped to ~76.1% of its initial efficiency ($PCE_{initial}$ ~ 19.50%) after 250 hours. In the damp heat test results (Fig. 4D), the P3HT:CuPc–based cell retained ~91.7% of its initial efficiency after 1009 hours of damp heat exposure, whereas the cell with the Spiro-OMeTAD HTL exhibited substantial degradation after only 240 hours, which we attributed to the severe deterioration at the interface between perovskite and Spiro-OMeTAD caused by lithium ion diffusion (*6*, *32*, *33*). The cells used for the damp heat tests included an additional barrier layer consisting of MoO_x (~10 nm) and IZO (~40 nm) placed between the HTL and the gold electrode to prevent gold diffusion into the cell at elevated temperatures (*45*).

Although we only present here stability data for one control and one nanopatterned cell, the same tests were repeated on multiple cell types incorporating the P3HT:CuPC HTL, all of which exhibited a similar improvement in stability relative to Spiro-OMeTAD–based cells. The hydrophobic and thermally stable nature of the ionic dopant–free P3HT:CuPc HTLs could increase the stability of perovskite cells compared with ionic-doped organic HTL materials without any penalty in performance. Time-dependent power output and *J-V* curves showing the cell performance evolution during the two degradation tests are provided in figs. S23 to S26.

Although the nanopatterned ETL reported here was fabricated with EBL, the structure is compatible with commercial scale fabrication using nanoimprint techniques (*46*). Some form of self-assembled ETL nanostructure might also provide a sufficiently irregular surface to form local contacts through a passivation layer. More generally, our work demonstrates the important interaction among ETL morphology, interface passivation, and charge transport at the ETL-perovskite interface.

REFERENCES AND NOTES

1. National Renewable Energy Laboratory, U.S. Department of Energy, "NREL efficiency chart" (2020); www.nrel.gov/pv/assets/pdfs/best-research-cell-efficiencies.20200311.pdf.
2. J. Burschka *et al*., *Nature* **499**, 316–319 (2013).
3. N. J. Jeon *et al*., *Nature* **517**, 476–480 (2015).
4. W. S. Yang *et al*., *Science* **348**, 1234–1237 (2015).
5. W. S. Yang *et al*., *Science* **356**, 1376–1379 (2017).
6. E. H. Jung *et al*., *Nature* **567**, 511–515 (2019).
7. M. A. Green *et al*., *Prog. Photovolt. Res. Appl.* **27**, 3–12 (2019).
8. M. A. Green *et al*., *Prog. Photovolt. Res. Appl.* **27**, 565–575 (2019).
9. M. A. Green *et al*., *Prog. Photovolt. Res. Appl.* **28**, 3–15 (2019).
10. J. Peng *et al*., *Energy Environ. Sci.* **10**, 1792–1800 (2017).
11. J. Peng *et al*., *Adv. Energy Mater.* **8**, 1801208 (2018).
12. S.-H. Turren-Cruz, A. Hagfeldt, M. Saliba, *Science* **362**, 449–453 (2018).
13. Q. Wang, Q. Dong, T. Li, A. Gruverman, J. Huang, *Adv. Mater.* **28**, 6734–6739 (2016).
14. L. Zuo *et al*., *Sci. Adv.* **3**, e1700106 (2017).
15. C. Tao *et al*., *Energy Environ. Sci.* **8**, 2365–2370 (2015).
16. F. Zhang *et al*., *Adv. Mater.* **29**, 1606806 (2017).
17. J. Xu *et al*., *Nat. Commun.* **6**, 7081 (2015).
18. D. Koushik *et al*., *Energy Environ. Sci.* **10**, 91–100 (2017).
19. J.-P. Correa-Baena *et al*., *Energy Environ. Sci.* **10**, 1207–1212 (2017).
20. S. Gharibzadeh *et al*., *Adv. Energy Mater.* **9**, 1803699 (2019).
21. M. A. Mahmud *et al*., *Adv. Funct. Mater.* **30**, 1907962 (2020).
22. D. Luo *et al*., *Science* **360**, 1442–1446 (2018).
23. X. Zheng *et al*., *Nat. Energy* **5**, 131–140 (2020).
24. Q. Jiang *et al*., *Nat. Photonics* **13**, 460–466 (2019).
25. J. J. Yoo *et al*., *Energy Environ. Sci.* **12**, 2192–2199 (2019).
26. M. J. Kerr, A. Cuevas, *Semicond. Sci. Technol.* **17**, 35–38 (2001).
27. J. Shewchun, R. Singh, M. Green, *J. Appl. Phys.* **48**, 765–770 (1977).
28. Y. Wu *et al*., *Adv. Mater.* **29**, 1701073 (2017).
29. T. Matsui *et al*., *Adv. Mater.* **31**, e1806823 (2019).
30. E. Franklin *et al*., *Prog. Photovolt. Res. Appl.* **24**, 411–427 (2014).
31. M. A. Green, *Prog. Photovolt. Res. Appl.* **17**, 183–189 (2009).
32. H. Min *et al*., *Science* **366**, 749–753 (2019).
33. Y. Hou *et al*., *Science* **358**, 1192–1197 (2017).
34. Q.-Q. Chu *et al*., *J. Mater. Sci. Technol.* **35**, 987–993 (2019).
35. Q. Hu *et al*., *Solar RRL* **3**, 1800264 (2019).
36. L. Yang *et al*., *J. Colloid Interface Sci.* **534**, 459–468 (2019).
37. A. Fakharuddin *et al*., *ACS Nano* **9**, 8420–8429 (2015).
38. X. Li *et al*., *ACS Appl. Mater. Interfaces* **8**, 21358–21365 (2016).
39. A. J. Bett *et al*., *Prog. Photovolt. Res. Appl.* **28**, 99–110 (2020).
40. C. Wu *et al*., *Solar RRL* **2**, 1800052 (2018).
41. W. Shockley, W. T. Read Jr., *Phys. Rev.* **87**, 835–842 (1952).
42. Z. Ni *et al*., *Science* **367**, 1352–1358 (2020).
43. S. Zhang *et al*., *IEEE J. Photovolt* **6**, 145–152 (2016).
44. D. Kiermasch, L. Gil-Escrig, H. J. Bolink, K. Tvingstedt, *Joule* **3**, 16–26 (2019).
45. T. Duong *et al*., *Sol. Energy Mater. Sol. Cells* **188**, 27–36 (2018).
46. L. J. Guo, *Adv. Mater.* **19**, 495–513 (2007).

ACKNOWLEDGMENTS: Funding: This work was supported by the Australian Government through the Australian Renewable Energy Agency (ARENA) and the Australian Research Council. Responsibility for the views, information, or advice expressed herein is not accepted by the Australian Government. J.P. acknowledges the financial support of a postdoctoral fellowship from the Australian Centre for Advanced Photovoltaics (ACAP). T.P.W. is the recipient of an Australian Research Council Future Fellowship (project no. FT180100302) funded by the Australian Government. T.P.W. also acknowledges the support of the Open Fund of the State Key Laboratory of Optoelectronic Materials and Technologies (Sun Yat-sen University). Y.R., Q.L., and J.L. acknowledge funding from the National Natural Science Foundation of China (grant nos. 11974436 and 11674402) and the Guangdong Basic and Applied Basic Research Foundation (grant no. 2020B1515020019). **Author contributions:** J.P. conceived the idea, designed the overall experiments, and led the project. J.P., Y.W., T.D., M.A.M., and H.S. prepared and characterized the perovskite cell devices. D.W., T.P.W., and K.J.W. conducted the 3D numerical simulation. Y.R. and Q.L. performed the PMMA hole pattern fabrication. J.L. supervised the PMMA hole pattern fabrication and optimization. J.P. and D.-Y.C. performed ALD TiO_2 deposition and ICP-RIE etching. Y.W. and M.A.M. performed the GIXRD and XRD measurements and analysis. M.T. and H.T.N. performed the PL and TRPL measurements and analysis. T.D., O.L.C.L., and S.Z. conducted the SEM measurements and analysis. T.L. and Y.L. performed the AFM measurements and analysis. W.L. performed the EQE measurements. L.L. and F.K. performed the FIB-SEM measurements and analysis. J.P., D.W., and T.P.W. wrote and revised the manuscript. T.P.W. and K.R.C. supervised the project. All authors contributed to the discussion of the results and revision of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/390/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S27
Table S1
References (*47*–*60*)
3D Simulation Model File

24 March 2020; resubmitted 18 October 2020
Accepted 10 December 2020
10.1126/science.abb8687

SEXUAL DIMORPHISM

Transvection regulates the sex-biased expression of a fly X-linked gene

Charalampos Chrysovalantis Galouzis and Benjamin Prud'homme*

Sexual dimorphism in animals results from sex-biased gene expression patterns. These patterns are controlled by genetic sex determination hierarchies that establish the sex of an individual. Here we show that the male-biased wing expression pattern of the *Drosophila biarmipes* gene *yellow*, located on the X chromosome, is independent of the fly sex determination hierarchy. Instead, we find that a regulatory interaction between *yellow* alleles on homologous chromosomes (a process known as transvection) silences the activity of a *yellow* enhancer functioning in the wing. Therefore, this enhancer can be active in males (XY) but not in females (XX). This transvection-dependent enhancer silencing requires the *yellow* intron and the chromatin architecture protein Mod(mdg4). Our results suggest that transvection can contribute more generally to the sex-biased expression of X-linked genes.

Sexual dimorphism in morphology, physiology, and behavior is pervasive in animals. Sex-biased gene expression patterns, deployed during embryonic or adult development, direct the formation of these phenotypic differences between sexes (*1*). It is well established that the transcriptional regulators of the somatic sex determination hierarchies directly control sexually dimorphic gene regulation (*1*–*3*). However, the different tiers of these hierarchies seem to contribute to this control through a variety of regulatory mechanisms (*1*–*7*). To better understand the molecular mechanisms that govern sexually dimorphic gene regulation, we examined the dimorphic regulation of the *yellow* (*y*) gene in *Drosophila biarmipes*. This species has evolved a male-specific wing pigmentation spot (Fig. 1A, left) (*8*, *9*). During late pupal wing development in *D. biarmipes* males, Yellow spatial distribution prefigures the adult pigmentation spot (Fig. 1A, right). By contrast, almost no pigmentation pattern appears in adult females (Fig. 1B, left), and only a few wing cells produce a high level of Yellow, forming a typical dotted, stochastic pattern (Fig. 1B, right) (*8*, *10*).

We examined the contribution of the top tier of the somatic sex determination hierarchy to the regulation of *y* in *D. biarmipes*, focusing on *Sex-lethal* (*Sxl*) and *transformer* (*tra*), genes that initiate and maintain, respectively, female identity in a cell-autonomous manner in *Drosophila* (*11*). For this, we feminized the male wing by overexpressing *Sxl* and, conversely, masculinized the female wing by knocking down *tra* expression, using a wing-specific driver (*12*). In both cases, this altered the wing pigmentation pattern (Fig. 1, C and D, left, and fig. S1), revealing the conversion of the cells' sexual identity. Surprisingly, however, the spatial pattern of Yellow maintained its original sex-specific expression in both situations (Fig. 1, C and D, right). These data confirm that *y* expression, although necessary to form a wing spot, is not sufficient for this purpose, and indicate that shifting the sexual identity of the wing affects pigmentation, presumably through the modification of genes other than *y* that are also involved in wing spot formation (*8*, *12*). These results suggest that the sex determination hierarchy does not control the sexually dimorphic regulation of *y* in *D. biarmipes* wings.

To understand how *y* is sexually regulated, we sought to identify the cis-regulatory sequences that direct its dimorphic expression in the wing. *y* expression in the *D. biarmipes* wing is controlled by (at minimum) a pair of neighboring enhancers, the *wing* and the *spot* enhancers (Fig. 2A), which drive *y* expression throughout the wing at a low level and in the presumptive wing spot area at a higher level, respectively (*9*). The *spot* enhancer drives a sexually monomorphic activity in reporter assays in *D. biarmipes* (*12*) (fig. S2, A and B). This result suggests that the *spot* sequence does not receive female-specific regulatory

Aix-Marseille Université, CNRS, Institut de Biologie du Développement de Marseille (IBDM), Campus de Luminy Case 907, 13288 Marseille Cedex 9, France.
***Corresponding author. Email: benjamin.prudhomme@univ-amu.fr**

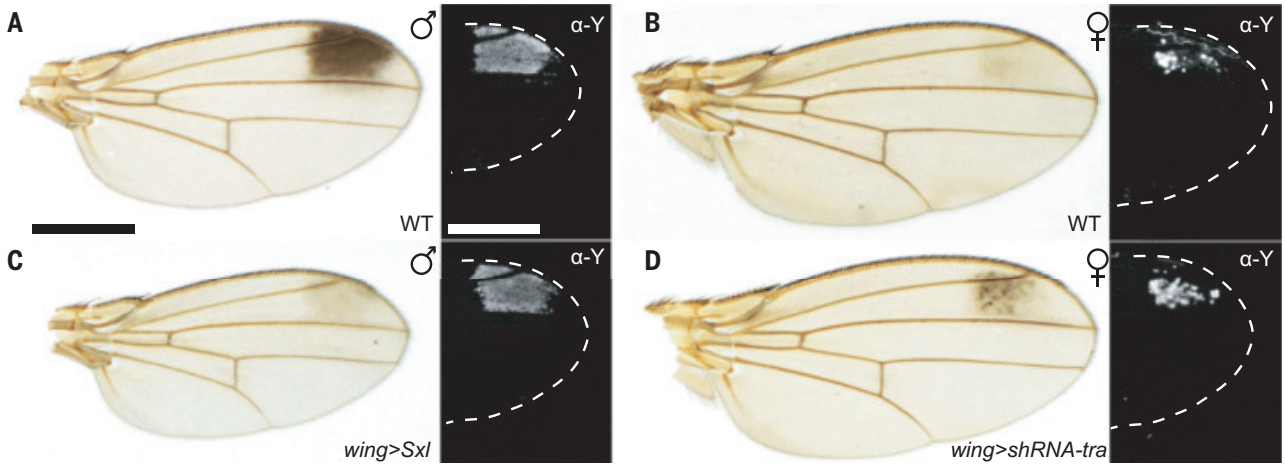


Fig. 1. Yellow sex-biased pattern in the *D. biarmipes* wing is independent of the sex determination hierarchy. (**A**) A male-specific wing pigmentation spot forms in adult *D. biarmipes* (left). The wing spot is prefigured during late pupal development by the spotted pattern of Yellow (revealed with an anti-Yellow antibody staining) (right). (**B**) In females, the wing pigmentation level is almost uniform (although a very faint spot appears in some individuals) (left), and Yellow is expressed at a high level in only a few cells (right). (**C**) Overexpression of *Sxl* in male wings reduces the black spot intensity to levels found in females (left), without altering Yellow spatial distribution (right). (**D**) Knocking down *tra* in females leads to increased pigmentation in the wing spot region (left), but does not affect the Yellow pattern (right). Antibody staining was performed on young adults, just after wing expansion (Figs. 1 to 3). White dashed lines indicate the wing contour. Scale bars, 500 μm (in Figs. 1 to 3, unless stated otherwise). WT, wild type; shRNA, short hairpin RNA.

inputs required for its silencing; another regulatory segment must be involved to control the sexually dimorphic expression of *y*.

To define a portion of the *y* locus that recapitulates its sexually dimorphic expression in the wing, we cloned a 25-kb-long fragment, y^{25kb}, that encompasses all coding and noncoding sequences of the gene. We also added a *mCherry* reporter in frame with the second exon of *y* to facilitate the detection of the protein product (Fig. 2A). Because *y* is X-linked, we inserted the *y::mCherry*^{25kb} reporter construct on the X chromosome in a *Drosophila melanogaster y* mutant. The *y::mCherry*^{25kb} construct fully rescues the *y* mutant phenotype in *D. melanogaster*, although it does not induce the formation of a pigmentation spot, as expected, because *y* is not sufficient to make dark pigment (*9*), demonstrating its full functionality (fig. S2, C to E). We then examined the distribution of Yellow::mCherry in the wings of freshly emerged adults. In males hemizygous for the transgene, Yellow::mCherry is produced moderately throughout the wing and at a higher level in the spot area, as expected (Fig. 2B). This pattern, hereafter described as a spotted pattern, is similar to the activity of the sum of the *wing* and *spot* enhancers in *D. melanogaster* (*9*). In females heterozygous for the transgene, *y::mCherry*^{25kb} drives spotted expression (Fig. 2C), as in hemizygous males. Surprisingly, in *D. melanogaster* females homozygous for the transgene, Yellow::mCherry is distributed uniformly throughout the wing, but no spotted expression is detected (Fig. 2D).

The different expression patterns driven by the transgene in hemi- and heterozygous as opposed to homozygous individuals could result from its insertion on the X chromosome

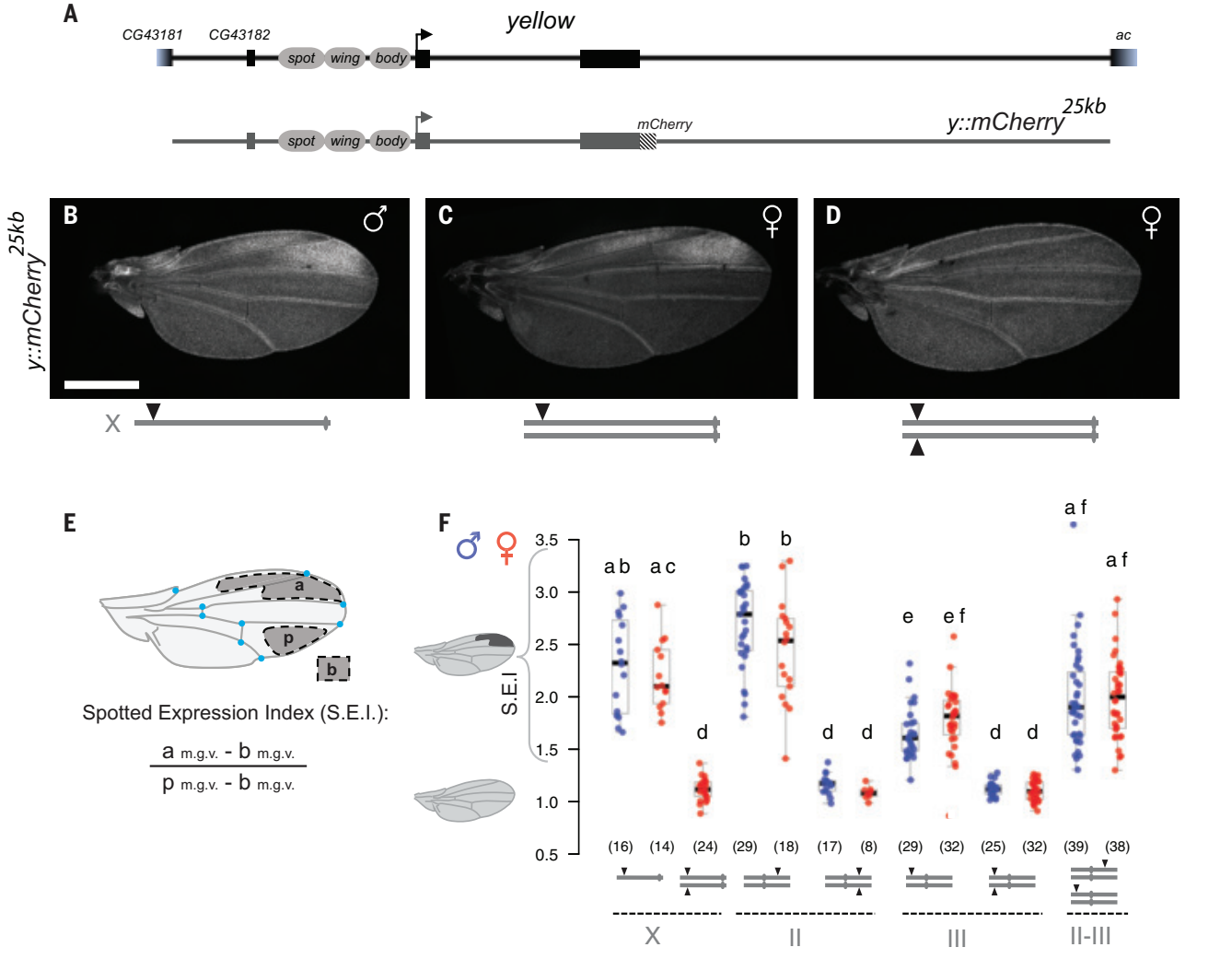


Fig. 2. Sexually dimorphic regulation of *y* requires functional homolog interaction. (**A**) Representation of the *D. biarmipes y* locus, with relative positions indicated for the *spot*, *wing*, and *body* enhancers and the 25-kb-long fragment cloned from the *y* locus tagged with *mCherry* (*y::mCherry*^{25kb}). (**B** and **C**) *y::mCherry*^{25kb} inserted on the X chromosome of *D. melanogaster* drives spotted expression in hemizygous males (B) and heterozygous females (C). (**D**) In homozygous females, the spotted expression of Yellow::Cherry is silenced. In Figs. 2 and 3, chromosomes are schematized below each genotype, with transgene insertion(s) indicated by black triangle(s). (**E**) We measured the mean gray value (m.g.v.) in three regions (a, anterior; p, posterior; b, background) to calculate the spotted expression index (S.E.I.) (*22*). (**F**) Quantification of the S.E.I. of *y::mCherry*^{25kb} insertions on different chromosomes in hemi-, hetero-, and homozygous and trans-heterozygous configurations. For all plots, each dot represents an individual wing (or abdomen); numbers in parentheses indicate sample sizes; and X, II, and III denote chromosomes. Males are in blue; females are in red. Also, for comparisons of multiple groups, we used letters (a to f) to describe the statistical relationship between them and used the following rule for all graphs: Groups that share at least one letter (e.g., ab versus bc) are statistically indistinguishable, and groups that have different letters (e.g., a versus b) are statistically different ($P < 0.05$; generalized linear model using a gamma distribution, followed by Tukey's contrasts for multiple comparisons of means).

and the existence of regulatory mechanisms specific to this sex chromosome (*13*). To explore this possibility, we inserted the *y::mCherry*^{25kb} construct on different autosomes (chromosomes II or III) and examined its activity in hetero- and homozygous flies, in both sexes. To quantitatively compare the degree of spotted expression between genotypes, we devised a spotted expression index (S.E.I.), which measures the mean intensity of reporter expression in the spot area relative to the mean intensity in a posterior region of the wing (Fig. 2E). Regardless of the sex of the individuals, chromosomal insertions, or vector backbone, *y::mCherry*^{25kb} drives spotted expression in flies heterozygous for the transgene, and uniform wing expression in homozygous individuals (Fig. 2F and fig. S2F). Together, these results enabled us to rule out the possibility that silencing of *y* spotted expression is controlled by an X chromosome–specific mechanism or even by female-specific regulators. Instead, they suggest that a regulatory mechanism (common to both sexes) that relies on locus copy number difference controls the spot pattern silencing.

To test whether the absolute copy number influences *y* regulation, we measured the spotted expression pattern of Yellow::mCherry in trans-heterozygous flies carrying two copies of the transgene inserted on chromosomes II and III and compared it with the pattern observed in heterozygous or homozygous individuals for each insertion. When the transgenes are present in two copies inserted on different autosomes, they drive a spotted expression pattern, in both sexes, similar to that in either of the heterozygous flies carrying each transgene, but in contrast to the silencing observed in homozygous individuals (Fig. 2F). We thus concluded that the silencing of the *y* spot pattern is not strictly due to a dose effect. Instead, the results suggest that the transgenes interact functionally, but only when they are on the same chromosome, to silence the *y* spot pattern.

To test this idea further, we inserted the *y::mCherry*^{25kb} transgene at two distinct sites on the X chromosome, separated by ~5 Mb, and measured their combined regulatory activity in the female wing in different configurations. As expected, females heterozygous for each transgene exhibited a spotted pattern, whereas homozygous females did not (fig. S2G). Notably, when the two copies are at different positions on the same chromosome in trans, or in cis on the same chromatid, the spotted expression is silenced, but not as strongly as in the homozygous flies (fig. S2G). Furthermore, the genomic sequences flanking the insertion sites are, on average, closer to one another in the nuclear space when the two transgenes are on the same chromatid than when only one transgene is present (fig. S2H). These results

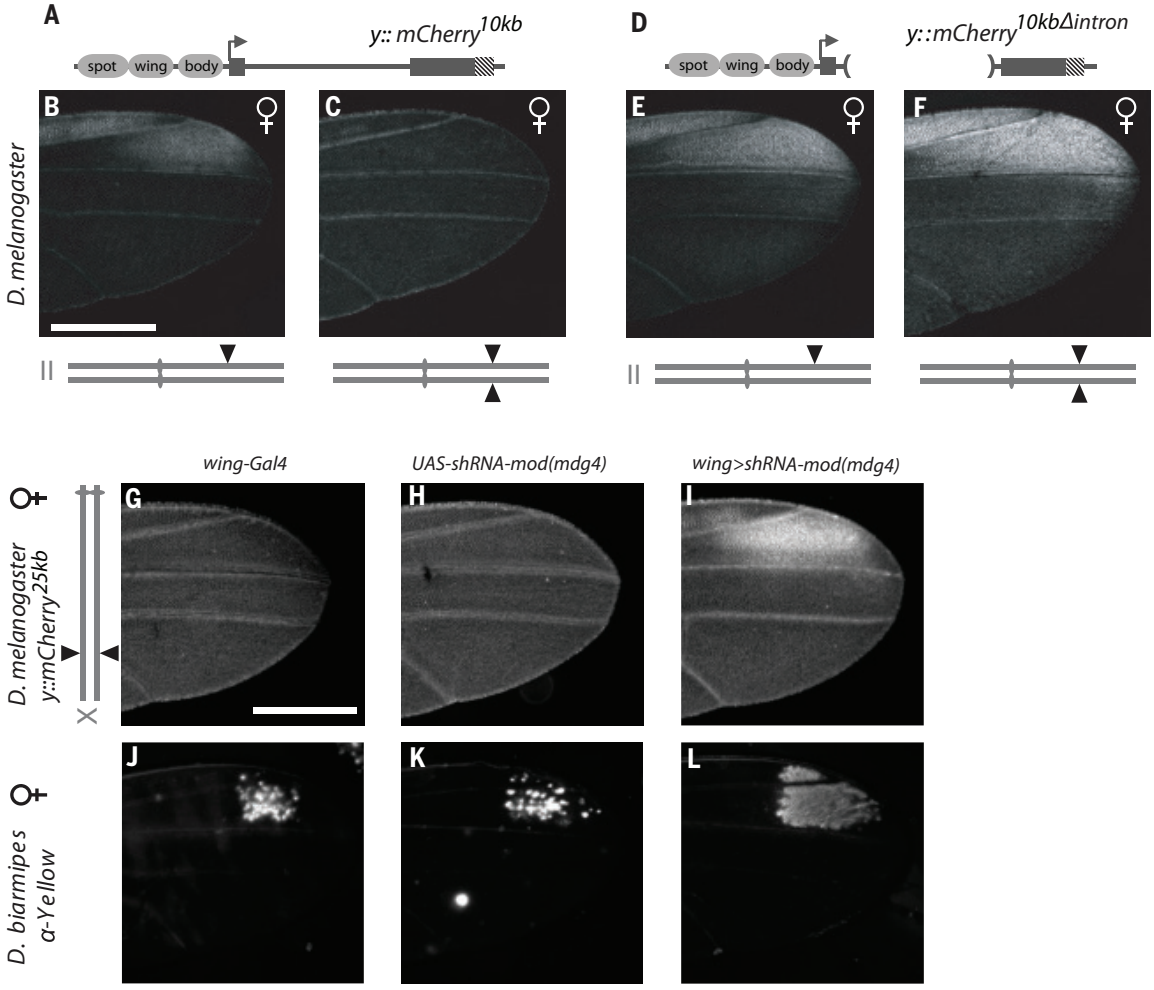


Fig. 3. The *y* intron and *mod(mdg4)* are required for the transvection-dependent silencing of the *y spot* enhancer. (**A** to **C**) *D. melanogaster* females carrying the *y::mCherry*^{10kb} reporter (A) on chromosome II display spotted expression when heterozygous for the transgene (B) and uniform wing expression when homozygous (C). (**D** to **F**) *D. melanogaster* females carrying *y::mCherry*^{10kbΔintron} (D) display spotted expression when heterozygous (E) or homozygous for the transgene (F). (**G** to **I**) Knocking down *mod(mdg4)* in *D. melanogaster* females abolishes silencing of *y::mCherry*^{25kb}. (**J** to **L**) Similarly, when knocking down *mod(mdg4)* in *D. biarmipes* females, the native Yellow wing pattern [(J) and (K)] becomes male-like (L). UAS, upstream activating sequence.

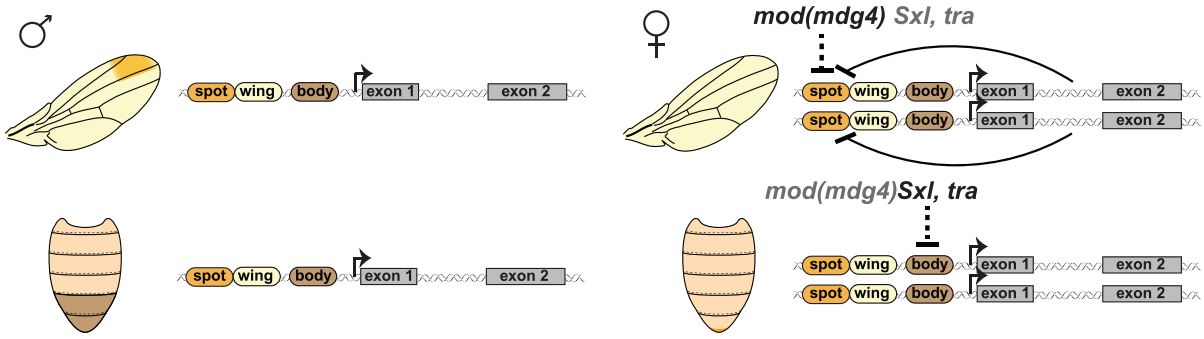


Fig. 4. Regulatory model of *y* sexually dimorphic expression in the wing and posterior abdomen of *D. biarmipes*. (**Left**) In males, where *y* is present in one copy because it is X-linked, the *spot* and the *body* enhancers are active. (**Right**) By contrast, in females, these enhancers are silenced by two distinct mechanisms: The *body* enhancer receives repressive inputs from targets of the sex determination hierarchy genes *Sxl* and *tra*, whereas the *spot* enhancer is silenced through transvection and uses the non–sex-specific Mod(mdg4) architectural protein.

suggest that functional interaction between the transgenes potentially requires their physical proximity. In the native context, interactions between *y* alleles are presumably facilitated by the tight pairing of homologous chromosomes, which are aligned end-to-end in all somatic cells in Diptera (*14*–*16*). From all of these results, we conclude that a regulatory interaction between *yellow* alleles on homologous chromosomes, a mechanism known as transvection (*17*–*19*), controls the sexually dimorphic regulation of *y* in the wing. Because *y* is X-linked, the transvection-dependent silencing of the spot pattern can occur only in females, which carry two X chromosomes.

We then wondered if the same mechanism also affects other sex-biased expression patterns of *y*. We examined the male-specific expression of *y* in the adult posterior abdominal segment (*20*), which is controlled by the *y body* enhancer, located near the *spot* and *wing* enhancers (Fig. 2A). The sexually dimorphic activity of the *body* enhancer receives indirect inputs from the sex determination hierarchy (*21*). We compared the reporter activity of the $y::mCherry^{25kb}$ transgene inserted on the X chromosome in hemizygous males, as well as hetero- and homozygous females (fig. S2, I to K). As expected, we observed a Yellow::mCherry signal in male posterior segments. By contrast, Yellow::mCherry intensity was low in females, in both hetero- and homozygous flies. This result suggests that the transvection-dependent silencing of *y* may affect only the *spot* enhancer.

To localize the sequences involved in the sexually dimorphic regulation of *y* in the wing, we first reduced the y^{25kb} region to a y^{10kb} fragment (fig. S3A) (*22*) and added *mCherry* in frame with the second exon of *y* (Fig. 3A). The $y::mCherry^{10kb}$ construct behaved like $y::mCherry^{25kb}$ in both hetero- and homozygous flies (Fig. 3, B and C, and fig. S4A). We observed similar results with a construct containing only the *y* 5′ and intron sequences of the y^{10kb} fragment (fig. S4, B to F). This result indicates that the *y* transcript or exonic sequences are dispensable for the transvection-dependent silencing of the spot pattern.

We then deleted the intron from $y::mCherry^{10kb}$ (Fig. 3D). Although $y::mCherry^{10kb\Delta intron}$ and $y::mCherry^{10kb}$ drive a similar expression pattern when present as a single copy (Fig. 3, B and E, and fig. S4A), we found that the spot pattern silencing is lost in flies homozygous for the $y::mCherry^{10kb\Delta intron}$ transgene (Fig. 3F and fig. S4A). This reveals that the *y* intron is necessary for the transvection-dependent silencing of the spot pattern.

Having identified that transvection shapes the sexually dimorphic regulation of *y*, we ran a genetic screen in *D. melanogaster* using available alleles or RNA interference (RNAi) lines for genes that have been previously associated with this phenomenon. We examined, in particular, members of the Polycomb group that have been recurrently involved in this mode of gene silencing (*23*) (table S2). The only candidate gene that, upon knockdown with two independent RNAi lines, affected the $y::mCherry^{25kb}$ sex-biased expression was *mod(mdg4)* (Fig. 3, G to I, and fig. S4G). We confirmed this result by knocking down *mod(mdg4)* in *D. biarmipes* females, which resulted in a male-like Yellow pattern in their wings (Fig. 3, J to L). These results suggest that *mod(mdg4)* is required for the transvection-dependent silencing of the *y spot* activity. Because Mod(mdg4) is involved in chromatin architecture and enhancer blocking (*24*–*26*), we speculate that it mediates the interactions between homologous *y* alleles, possibly by bridging genomic regions to bring the *spot* enhancers and the introns together.

Collectively, these results identify a regulatory mechanism for sex-biased expression of the X-linked gene *y* in the wing that is independent of the somatic sex determination hierarchy, which, by contrast, controls sex-biased *y* expression in the body (fig. S2, I to L) (*21*) and relies instead on transvection (Fig. 4). Transvection has been documented in different groups of organisms, but in most cases interactions are revealed only in mutant contexts or with transgenic constructs (*17*, *27*–*29*). Even in *D. melanogaster*, where transvection has been studied most extensively, the physiological relevance of this mechanism to gene regulation in a wild-type context has long remained questionable (*30*). Our results suggest that one notable contribution of transvection may be to shape the sexually dimorphic regulation of X-linked genes.

REFERENCES AND NOTES

1. T. M. Williams, S. B. Carroll, *Nat. Rev. Genet.* **10**, 797–804 (2009).
2. A. Prakash, A. Monteiro, *Curr. Opin. Insect Sci.* **17**, 40–48 (2016).
3. N. Camara, C. Whitworth, M. Van Doren, *Curr. Top. Dev. Biol.* **83**, 65–107 (2008).
4. J. L. Rinn, M. Snyder, *Trends Genet.* **21**, 298–305 (2005).
5. A. Sawala, A. P. Gould, *PLOS Biol.* **15**, e2002252 (2017).
6. X. Sun *et al.*, *G3* **5**, 2865–2874 (2015).
7. B. Hudry, S. Khadayate, I. Miguel-Aliaga, *Nature* **530**, 344–348 (2016).
8. P. J. Wittkopp, J. R. True, S. B. Carroll, *Development* **129**, 1849–1858 (2002).
9. N. Gompel, B. Prud'homme, P. J. Wittkopp, V. A. Kassner, S. B. Carroll, *Nature* **433**, 481–487 (2005).
10. A. Kopp, J. R. True, *Evol. Dev.* **4**, 278–291 (2002).
11. H. K. Salz, J. W. Erickson, *Fly* **4**, 60–70 (2010).
12. L. Arnoult *et al.*, *Science* **339**, 1423–1426 (2013).
13. J. C. Lucchesi, M. I. Kuroda, *Cold Spring Harb. Perspect. Biol.* **7**, a019398 (2015).
14. N. M. Stevens, *J. Exp. Zool.* **5**, 359–374 (1908).
15. J. AlHaj Abed *et al.*, *Nat. Commun.* **10**, 4485 (2019).
16. K. Viets *et al.*, *Dev. Cell* **51**, 341–356.e7 (2019).
17. E. B. Lewis, *Am. Nat.* **88**, 225–239 (1954).
18. J. A. Kennison, J. W. Southworth, *Adv. Genet.* **46**, 399–420 (2002).
19. I. W. Duncan, *Annu. Rev. Genet.* **36**, 521–556 (2002).
20. S. Jeong, A. Rokas, S. B. Carroll, *Cell* **125**, 1387–1399 (2006).
21. M. J. Roeske, E. M. Camino, S. Grover, M. Rebeiz, T. M. Williams, *eLife* **7**, e32273 (2018).
22. Materials and methods are available as supplementary materials.
23. Y. B. Schwartz, V. Pirrotta, *Nat. Rev. Genet.* **8**, 9–22 (2007).
24. O. Kyrchanova, P. Georgiev, *FEBS Lett.* **588**, 8–14 (2014).
25. L. Melnikova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14806–14811 (2004).
26. M. Savitsky, M. Kim, O. Kravchuk, Y. B. Schwartz, *Genetics* **202**, 601–617 (2016).
27. M. S. Hogan, D.-E. Parfitt, C. J. Zepeda-Mendoza, M. M. Shen, D. L. Spector, *Cell Stem Cell* **16**, 275–288 (2015).

28. D. Zhang, L. Bai, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4428–4433 (2016).
29. D. J. Mellert, J. W. Truman, *Genetics* **191**, 1129–1141 (2012).
30. M. S. Apte, V. H. Meller, *Genet. Res. Int.* **2012**, 430587 (2012).
31. C. C. Galouzis, B. Prud'homme, Transvection regulates the sex-biased expression of a fly X-linked gene, figshare (2021); https://doi.org/10.6084/m9.figshare.13239110.

ACKNOWLEDGMENTS

We are grateful to the Bloomington Drosophila Stock Center for fly stocks; Flybase for information support; N. Gompel for the anti-Yellow antibody; A. Sultanova for assistance in the collection of abdomens; and M. Paris, N. Gompel, J. Ewbank, and all laboratory members for helpful comments on the manuscript. **Funding:** This work was funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC grant agreement 615789 (B.P.). C.C.G. acknowledges funding from the Fondation pour la Recherche Médicale (FRM FDT201904008114). **Author contributions:** B.P. conceived the project. B.P. and C.C.G. designed the experiments. C.C.G. performed all experiments and statistical analyses. C.C.G. and B.P. analyzed the results and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The authors declare that the data supporting the findings of this study are available in the manuscript or supplementary materials, and raw image data are deposited in the figshare digital repository (*31*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/396/suppl/DC1
Material and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 and S2
References (*32*–*35*)
MDAR Reproducibility Checklist

15 April 2020; accepted 17 December 2020
10.1126/science.abc2745

HOST DEFENSE

Host succinate is an activation signal for *Salmonella* virulence during intracellular infection

Gili Rosenberg[1], Dror Yehezkel[1], Dotan Hoffman[1], Camilla Ciolli Mattioli[1], Moran Fremder[2], Hadar Ben-Arosh[1], Leia Vainman[1], Noa Nissani[1], Shelly Hen-Avivi[1], Shirley Brenner[1], Maxim Itkin[3], Sergey Malitsky[3], Ehud Ohana[2], Noa Bossel Ben-Moshe[1], Roi Avraham[1]*

Key to the success of intracellular pathogens is the ability to sense and respond to a changing host cell environment. Macrophages exposed to microbial products undergo metabolic changes that drive inflammatory responses. However, the role of macrophage metabolic reprogramming in bacterial adaptation to the intracellular environment has not been explored. Here, using metabolic profiling and dual RNA sequencing, we show that succinate accumulation in macrophages is sensed by intracellular *Salmonella* Typhimurium (*S.* Tm) to promote antimicrobial resistance and type III secretion. *S.* Tm lacking the succinate uptake transporter DcuB displays impaired survival in macrophages and in mice. Thus, *S.* Tm co-opts the metabolic reprogramming of infected macrophages as a signal that induces its own virulence and survival, providing an additional perspective on metabolic host–pathogen cross-talk.

Studies of the interactions between *Salmonella enterica* serovar Typhimurium (*S.* Tm) and macrophages have elucidated both a complex virulence program executed by the pathogen and the opposing ability of macrophages to recognize and kill invading bacteria (*1*). *S.* Tm has evolved to sense the transition from extracellular to intravacuolar environments, which is characterized by physiological host processes that aim to eliminate the invading pathogen (*2*). These intracellular cues provide signals that activate bacterial two-component systems, such as PhoPQ and PmrAB, which promote resistance to host antimicrobial peptides (*3*), and the type III secretion system, located on the *Salmonella* pathogenicity island 2 (SPI-2), which enables intracellular replication (*4*). Macrophage detection of invading bacteria via pattern recognition receptors results in disruption of the citric acid (TCA) cycle, which shifts host metabolism from oxidative phosphorylation (OXPHOS) toward aerobic glycolysis with consequent accumulation of several metabolic intermediates (*5*). Although it is increasingly understood that these metabolic changes are key to the regulation of host proinflammatory responses during infection (*6*), the impact of host metabolic reprogramming on intracellular pathogens is largely unknown.

To determine whether host-derived metabolites modulate the intracellular lifestyle of *S.* Tm, we sought to test their effects on *S.* Tm virulence and survival in macrophages. We used 2-deoxyglucose (2DG) to inhibit the shift from OXPHOS to glycolysis (*5*) in activated bone marrow–derived macrophages (BMDMs) infected with *S.* Tm (fig. S1A). We optimized an effective 2DG dose at a low concentration of 1 mM to minimize off-target effects (Fig. 1A and fig. S1, B and C). Additionally, 2DG may lead to a deficient glucose homeostasis of the bacteria that induces a stress response (*7*). 2DG treatment of *S.* Tm in culture medium caused their impaired growth and induced stress response genes, which was abolished in an *S.* Tm mutant lacking the transporter of 2DG, ManZ (*8*) (fig. S1, D to F, and table S1). Thus, 2DG and the Δ*manZ* mutant can serve as a tractable model to study the impact of macrophage metabolic reprogramming on intracellular *S.* Tm.

To assess the interplay of host metabolism and pathogen dynamics in our model, we applied a dual RNA sequencing (RNA-seq) method (*9*). We curated a list of 977 macrophage genes that were induced in infected cells (table S2). Host responses to infection with the wild-type (WT) or Δ*manZ* were similar (fig. S2A), and 2DG did not induce transcriptional changes in uninfected (naïve) cells (fig. S2B). We classified host genes that were up-regulated during infection into three classes, on the basis of their expression changes in the presence of 2DG (Fig. 1B, fig. S2C, and table S2). Notably, class 2 comprised genes down-regulated by 2DG, including genes differentially expressed by metabolic reprogramming, verified glycolysis inhibition by 2DG (Fig. 1C and fig. S2D).

To test the impact of inhibiting the host OXPHOS-to-glycolysis shift on intracellular *S.* Tm, we curated 960 bacterial genes that were up-regulated during infection and differentially expressed with 2DG treatment (fig. S3A and table S3). Expression of stress response genes in WT 2DG-treated cells (fig. S3B), and of *ripA*, part of the itaconic acid degradation operon (*10*) in the Δ*manZ* strain (fig. S3C), suggested that the macrophage metabolic state is sensed by intracellular *S.* Tm. Notably, global regulon analysis (table S4) indicated decreased expression in Δ*manZ*-infected, 2DG-treated cells of regulons that mediate intracellular survival in macrophages, most prominently the SPI-2 regulon (hypergeometric test, $P < 0.007$, $Q < 0.12$) (Fig. 1, D and E, and fig. S3, D and E). Thus, there is a clear correlation between the metabolic state of infected macrophages and *S.* Tm virulence gene expression.

We next analyzed intracellular accumulation of host metabolites during infection with *S.* Tm. Several metabolites, including primary metabolites of the TCA cycle, were significantly altered upon *S.* Tm infection and also reduced by 2DG treatment (Fig. 2A, fig. S4A, and table S5). To test whether *S.* Tm responds to the accumulation of these host metabolites, we performed a virulence assay utilizing the antimicrobial peptide polymyxin B (PMB). Survival of a Δ*phoP* strain which is sensitive to PMB challenge (*11*) was significantly

[1]Department of Biological Regulation, Weizmann Institute of Science, 7610001 Rehovot, Israel. [2]Department of Clinical Biochemistry and Pharmacology, Faculty of Health Sciences, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel. [3]Life Sciences Core Facilities, Weizmann Institute of Science, 7610001 Rehovot, Israel.
***Corresponding author. Email: roi.avraham@weizmann.ac.il**

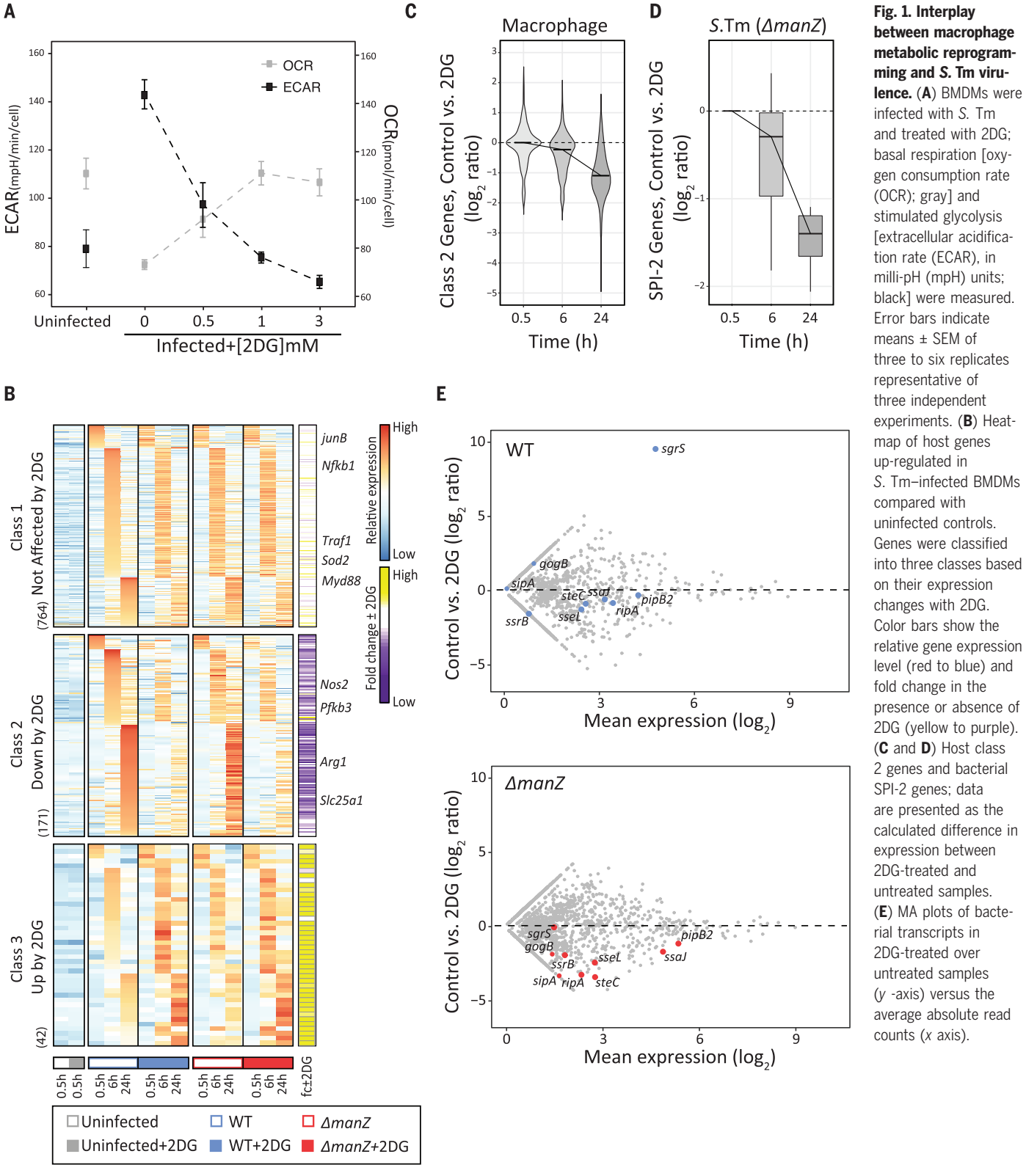


Fig. 1. Interplay between macrophage metabolic reprogramming and *S.* Tm virulence. (**A**) BMDMs were infected with *S.* Tm and treated with 2DG; basal respiration [oxygen consumption rate (OCR); gray] and stimulated glycolysis [extracellular acidification rate (ECAR), in milli-pH (mpH) units; black] were measured. Error bars indicate means ± SEM of three to six replicates representative of three independent experiments. (**B**) Heatmap of host genes up-regulated in *S.* Tm–infected BMDMs compared with uninfected controls. Genes were classified into three classes based on their expression changes with 2DG. Color bars show the relative gene expression level (red to blue) and fold change in the presence or absence of 2DG (yellow to purple). (**C** and **D**) Host class 2 genes and bacterial SPI-2 genes; data are presented as the calculated difference in expression between 2DG-treated and untreated samples. (**E**) MA plots of bacterial transcripts in 2DG-treated over untreated samples (*y* -axis) versus the average absolute read counts (*x* axis).

restored only in the presence of succinate and itaconate (Fig. 2B and fig. S4B). *S.* Tm grew normally in medium supplemented with succinate, but not in medium with itaconate, at concentrations found within activated macrophages (*12*) (fig. S4, C and D). Thus, we hypothesized that succinate serves as a signal to induce *S.* Tm virulence.

With ^{14}C-labeled succinate, we found that *S.* Tm took up and accumulated exogenous succinate (Fig. 2C). We selected a subset of SPI-2 candidate genes whose expression correlated with metabolic remodeling of infected macrophages (Fig. 1D). We measured induced mRNA expression and fluorescence from

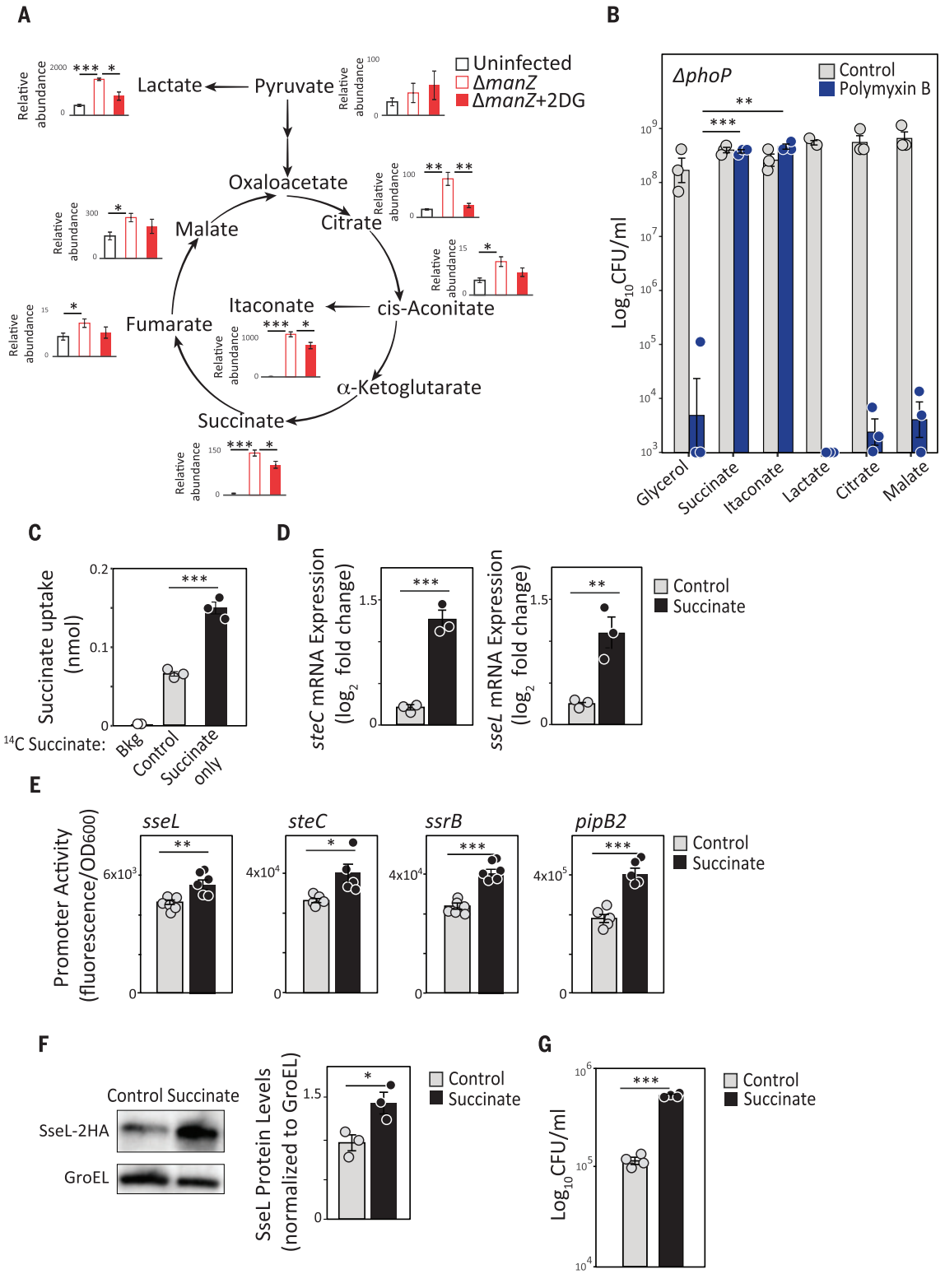


Fig. 2. *S.* Tm virulence is induced in response to succinate. (**A**) BMDMs were infected with a Δ*manZ* strain and treated with 2DG, and intracellular levels of central metabolites were measured by mass spectrometry. (**B**) Δ*phoP* mutant grown without (control; gray) or with PMB (blue) supplemented with 10 mM of the indicated metabolites. Survival was measured by colony-forming units (CFU). (**C**) *S.* Tm was cultured in the indicated medium supplemented with ^{14}C-labeled succinate, and succinate uptake was measured. Background indicates heat-killed bacteria. (**D** to **G**) *S.* Tm was grown in control medium or medium supplemented with 10 mM succinate. The expression of *steC*, *sseL*, and *gapA* (normalizing control) was measured by quantitative real-time polymerase chain reaction(D). Fluorescent signal was measured for *S.* Tm strains expressing a transcriptional reporter fused to selected SPI-2 promoters [normalized to the optical density at 600 nm (OD_{600})] (E). Immunoblot of SseL protein (SseL-2HA) and GroEL (F). BMDMs were infected with *S.* Tm, and intracellular survival was measured by CFU (G). Error bars indicate means ± SEM, representative of two [(C)] or three [(B), (D to G)] independent experiments. *P < 0.05; **P < 0.01; ***P < 0.001 (unpaired t test).

transcriptional reporters of these genes in succinate-containing medium (Fig. 2, D and E, and fig. S4E). Succinate also induced protein levels of one of these SPI-2 effectors, SseL (Fig. 2F). Finally, *S.* Tm grown in succinate-containing medium showed increased survival in BMDMs compared to control (Fig. 2G). Thus, succinate is sufficient to mount a bacterial pathogenic program.

Using RNA-seq, we comprehensively analyzed the direct effects of succinate on *S.* Tm (table S6). Succinate induced the expression of several regulons known to mediate *S.* Tm virulence within macrophages (Fig. 3A, fig. S5A, and table S7). In line with the observed effect of host metabolic shift on induction of SPI-2 genes during intracellular infection, growth in succinate-containing medium significantly

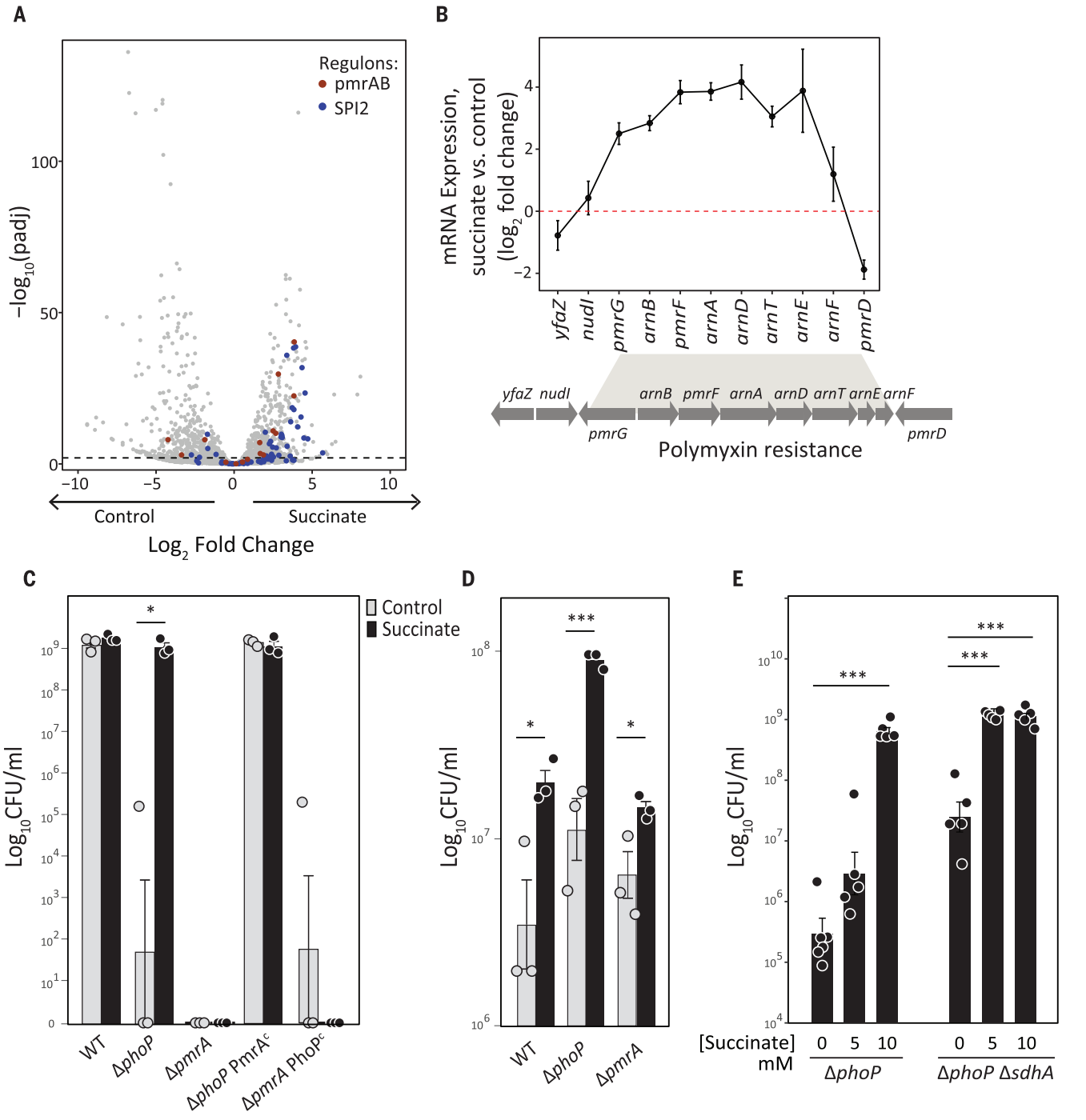


Fig. 3. Succinate induces SPI-2 and antimicrobial resistance. (**A**) RNA-seq analysis of *S.* Tm grown in control medium or succinate-containing medium. SPI-2 and PmrAB regulons were significantly induced in succinate. (**B**) mRNA expression of genes induced by succinate [taken from (A)] across the PMB resistance locus. (**C** to **E**) *S.* Tm strains were grown in the presence of PMB (C and E) or defensin (D), and survival was measured by CFU. Error bars indicate means ± SEM, representative of three independent experiments. *P < 0.05; **P < 0.01; ***P < 0.001 (unpaired t test).

induced expression of the SPI-2 regulon (hypergeometric test, Q < 0.001). Furthermore, succinate induced the expression of the PmrAB regulon (hypergeometric test, Q < 0.001), which is important for antimicrobial peptide resistance (*3*) (Fig. 3B). We tested whether succinate-induced PMB resistance (Fig. 2B) is mediated by PmrA-regulated genes (Fig. 3C and fig. S5B). Similar to the Δ*phoP* strain, the Δ*pmrA* strain was susceptible to PMB in control media. However, succinate exposure did not induce PMB resistance in Δ*pmrA*. Resistance was preserved in a Δ*phoP* $PmrA^c$ strain, but not in a Δ*pmrA* $PhoP^C$ strain. Thus, succinate-mediated PMB resistance appears to occur by the induction of PmrA-regulated virulence genes; this deviates from the canonical activation by PhoPQ (*13*). Succinate also induced resistance to the intracellular antimicrobial peptide defensin (*14*) (Fig. 3D and fig. S5C), suggesting a role for succinate in protection against antimicrobial peptides during intracellular infection within macrophages.

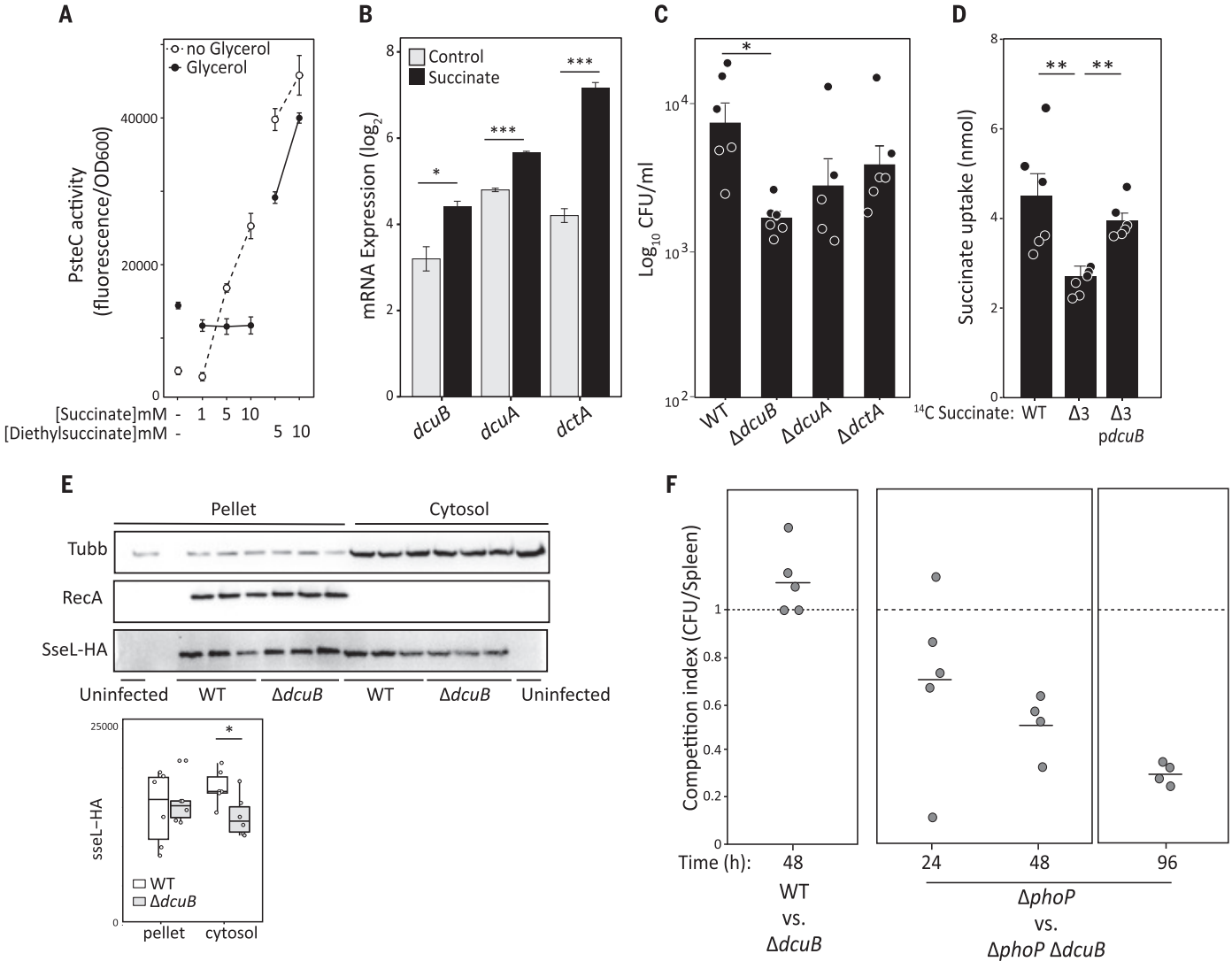


Fig. 4. Uptake of host succinate is required for intracellular survival of *S.* Tm. (**A**) *S.* Tm harboring a PsteC-*cfp* reporter was grown in media without or with glycerol and with succinate or diethylsuccinate as indicated, and fluorescence was measured (normalized to the OD_{600}). (**B**) Expression of the C_4-dicarboxylate transporter genes, taken from RNA-seq data of Fig. 3A. (**C**) BMDMs were infected with the indicated *S.* Tm strains, and intracellular survival was measured by CFU. (**D**) *S.* Tm strains were cultured in succinate-containing medium, and succinate uptake was measured with ^{14}C-labeled succinate. (**E**) Macrophages were infected with WT or Δ*dcuB* strains. SseL protein was quantified by protein immunoblotting of bacterial pellet and host cell cytosol fractions. Boxplots show SseL protein abundance. (**F**) Groups of mice (boxes) were infected i.p. with either WT and Δ*dcuB* or Δ*phoP* and Δ*phoP* Δ*dcuB* mutants, and competitive index was measured by CFU. Error bars indicate means ± SEM, representative of two [(B), (D), and (F)] or three [(A), (C), and (E)] independent experiments. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (unpaired *t* test).

During gut colonization, *S.* Tm can utilize succinate generated by the gut microbiota to complete the TCA cycle (*15*). To test this possibility, we generated an *S.* Tm double mutant lacking *phoP* and the succinate dehydrogenase complex subunit A (Δ*sdhA* Δ*phoP*). Similar to Δ*phoP*, Δ*sdhA* Δ*phoP* bacteria grown with succinate showed restored resistance to PMB (Fig. 3E and fig. S5, D and E). This result is consistent with succinate acting as a signal for *S.* Tm virulence independent of its role in central metabolism.

Shifts in host succinate levels might be sensed in the bacterial outer membrane or accumulate within the bacterial cytoplasm. Preferred sugars, such as glycerol, induce carbon catabolite repression, which prevents succinate uptake (*16*). Bacteria grown with succinate alone exhibited a dose-dependent induction of the SPI-2 gene *steC*, which was lost in medium containing glycerol (Fig. 4A and fig. S6, A and B). The addition of permeable succinate (diethylsuccinate) resulted in elevated *steC* expression even in the presence of glycerol or of media from activated macrophages (fig. S6C). Thus, succinate is actively transported into bacteria to induce *S.* Tm pathogenicity.

Succinate transport into bacteria is mediated by C_4-dicarboxylate transporters (*17*). Growth in succinate-containing medium increased expression of three C_4-dicarboxylate transporters (Fig. 4B). We examined intracellular survival of the three transporter mutants and measured a significantly lower survival of the Δ*dcuB* mutant (Fig. 4C and fig. S6D). Expression of *dcuB* was induced by growth in succinate-containing medium and within BMDMs (fig. S6, E and F). Because genetic redundancy may mask individual transporter phenotypes (*18*), we assessed succinate uptake in a triple *dcu* mutant (Δ3) background. Succinate uptake was inhibited in the Δ3 mutant and was significantly restored by complementation of DcuB (Fig. 4D). Similar effects were measured for the expression of the SPI-2 effector *steC* (fig. S6G). In macrophages

infected with WT or Δ*dcuB* strains, secretion of the SseL SPI-2 effector to the host cytosol was significantly reduced in the Δ*dcuB* mutant (Fig. 4E). Thus, active transport of succinate through DcuB is necessary for *S*. Tm virulence within macrophages.

To test the role of succinate uptake during in vivo infection, mice were intraperitoneally (i.p.) infected either with WT and Δ*dcuB* mutant or with Δ*phoP* and Δ*phoP* Δ*dcuB* mutants (Fig. 4F). There was no significant difference in the competitive index between the WT and Δ*dcuB*. By contrast, the Δ*phoP* Δ*dcuB* had significantly lower survival compared to the Δ*phoP* mutant at 48 and 96 hours after infection (paired *t* test $P < 0.012$, $P < 0.0003$, respectively). Thus, the relationship between host accumulation of succinate and bacterial virulence, mediated by DcuB, is also relevant for in vivo infection.

This work indicates that *S*. Tm has co-opted the metabolic reprogramming of infected macrophages as a signal that induces its own virulence (fig. S7). Other studies reinforce the importance of host and microbiota metabolic products that promote colonization by *S*. Tm and other pathogens (*19*–*21*). Our study indicates that during intracellular infection, uptake of host succinate by the DcuB transporter converges on a distinctive pathway for virulence induction. One plausible mechanism could be by direct interaction of metabolites with bacterial proteins, as with host proteins (*5*). We suggest a model whereby bacterial virulence is regulated by sensing of TCA intermediates of the macrophage metabolic shift, and this may be a general theme required for the pathogenic success of diverse intracellular bacteria.

REFERENCES AND NOTES

1. E. Mills, R. Avraham, *Curr. Opin. Microbiol.* **36**, 69–75 (2017).
2. D. Drecktrah, L. A. Knodler, R. Ireland, O. Steele-Mortimer, *Traffic* **7**, 39–51 (2006).
3. J. S. Gunn, *Trends Microbiol.* **16**, 284–290 (2008).
4. E. Jennings, T. L. M. Thurston, D. W. Holden, *Cell Host Microbe* **22**, 217–231 (2017).
5. D. G. Ryan, L. A. J. O'Neill, *Annu. Rev. Immunol.* **38**, 289–313 (2020).
6. L. E. Gleeson *et al*., *J. Immunol.* **196**, 2444–2449 (2016).
7. M. Bobrovskyy, C. K. Vanderpool, *Front. Cell. Infect. Microbiol.* **4**, 61 (2014).
8. B. Erni, B. Zanolari, H. P. Kocher, *J. Biol. Chem.* **262**, 5238–5247 (1987).
9. R. Avraham *et al*., *Nat. Protoc.* **11**, 1477–1491 (2016).
10. S. Srikumar *et al*., *PLOS Pathog.* **11**, e1005262 (2015).
11. Y. Shi, M. J. Cromie, F.-F. Hsu, J. Turk, E. A. Groisman, *Mol. Microbiol.* **53**, 229–241 (2004).
12. L. Tretter, A. Patocs, C. Chinopoulos, *Biochim. Biophys. Acta Bioenerg.* **1857**, 1086–1101 (2016).
13. L. F. F. Kox, M. M. S. M. Wösten, E. A. Groisman, *EMBO J.* **19**, 1861–1872 (2000).
14. P. I. Fields, E. A. Groisman, F. Heffron, *Science* **243**, 1059–1062 (1989).
15. L. Spiga *et al*., *Cell Host Microbe* **22**, 291–301.e6 (2017).
16. J. L. Botsford, R. D. DeMoss, *J. Bacteriol.* **105**, 303–312 (1971).
17. I. G. Janausch, E. Zientz, Q. H. Tran, A. Kröger, G. Unden, *Biochim. Biophys. Acta Bioenerg.* **1553**, 39–56 (2002).
18. A. Kleefeld, B. Ackermann, J. Bauer, J. Krämer, G. Unden, *J. Biol. Chem.* **284**, 265–275 (2009).
19. P. Thiennimitr *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17480–17485 (2011).
20. A. Jacobson *et al*., *Cell Host Microbe* **24**, 296–307.e7 (2018).
21. M. M. Curtis *et al*., *Cell Host Microbe* **16**, 759–769 (2014).

ACKNOWLEDGMENTS

We thank A. Barczak, E. Gottlieb, and A. Vardi for their valuable advice and their reading and commentary on this manuscript. **Funding:** This work was supported by the European Research Committee (ERC grant no. 756653), the Israel Science Foundation (grant no. 1890/17), the Minerva Foundation with funding from the Federal Ministry for Education and Research, the Estate of David Turner, the Merle S. Cahn Foundation, the Estate of Sylvia Holder, the Estate of Zvia Zeroni, and the Estate of Leah Arbel. R.A. is the Incumbent of the Philip Harris and Gerald Ronson Career Development Chair. Partial funding was provided to E.O. by the Israeli Science Foundation (ISF 271/16 and 2164/16).

Author contributions: G.R. and R.A. designed the study. G.R., D.H., C.C.M., H.B.-A., L.V., S.H.-A., and S.B. performed the experiments. M.F. and E.O. performed ^{14}C-succinate uptake experiments. M.I. and S.M. performed metabolomics experiments. D.Y. and N.N. analyzed the data. G.R. and R.A. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All RNA-seq data have been deposited in NCBI's Gene Expression Omnibus (GEO) under the super-series accession number GSE139208. All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/400/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S10
References (*22*–*40*)
MDAR Reproducibility Checklist

7 January 2020; resubmitted 6 August 2020
Accepted 20 November 2020
10.1126/science.aba8026

IMMUNOMETABOLISM

Glycolysis fuels phosphoinositide 3-kinase signaling to bolster T cell immunity

Ke Xu[1,2], Na Yin[1]*, Min Peng[1]*, Efstathios G. Stamatiades[1,3], Amy Shyu[1,4], Peng Li[1], Xian Zhang[1], Mytrang H. Do[1,2], Zhaoquan Wang[1,2], Kristelle J. Capistrano[1], Chun Chou[1], Andrew G. Levine[1,2], Alexander Y. Rudensky[1,2,5], Ming O. Li[1,2]†

Infection triggers expansion and effector differentiation of T cells specific for microbial antigens in association with metabolic reprograming. We found that the glycolytic enzyme lactate dehydrogenase A (LDHA) is induced in CD8$^+$ T effector cells through phosphoinositide 3-kinase (PI3K) signaling. In turn, ablation of LDHA inhibits PI3K-dependent phosphorylation of Akt and its transcription factor target Foxo1, causing defective antimicrobial immunity. LDHA deficiency cripples cellular redox control and diminishes adenosine triphosphate (ATP) production in effector T cells, resulting in attenuated PI3K signaling. Thus, nutrient metabolism and growth factor signaling are highly integrated processes, with glycolytic ATP serving as a rheostat to gauge PI3K-Akt-Foxo1 signaling in the control of T cell immunity. Such a bioenergetic mechanism for the regulation of signaling may explain the Warburg effect.

In response to antigen stimulation, naïve T (Tn) cells differentiate into effector T (Teff) cells with rewired metabolic pathways (*1*–*4*). Lactate dehydrogenase A (LDHA), a glycolysis enzyme that converts pyruvate to lactate, promotes Teff cytokine expression (*5*). However, glycolytic control of Teff responses to infection has yet to be defined, with a recent study questioning its importance (*6*). To this end, we infected mice with *Listeria monocytogenes* expressing chicken ovalbumin protein as an antigen (LM-OVA) (*7*). The expression of LDHA in H-2K^b-OVA$^+$ Teff cells exceeded that in CD8$^+$ Tn cells by a factor of ~5 (Fig. 1A and fig. S1). Activation of Tn cells with increasing doses of anti-CD3 progressively induced LDHA expression (Fig. 1B and fig. S2, A and B), in line with induction of the LDHA transcriptional regulator c-Myc (Fig. 1B and fig. S2B) (*8*). Notably, although proximal T cell receptor (TCR) signaling, marked by phospho-Zap70 and phospho-LAT, was saturated at low concentrations of anti-CD3 (fig. S2C), phosphoinositide 3-kinase (PI3K)–dependent Akt phosphorylation showed a response curve similar to that of c-Myc and LDHA (Fig. 1B and fig. S2B). Indeed, blockade of PI3K signaling repressed c-Myc and LDHA expression (Fig. 1C and fig. S2D). By contrast, phospho-Zap70 and phospho-LAT were minimally affected (fig. S2E). These observations, together with previous findings that PI3K-Akt signaling promotes expression of the glucose transporter Glut1 (*9*), support a

[1]Immunology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [2]Immunology and Microbial Pathogenesis Graduate Program, Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, NY 10065, USA. [3]Institute of Microbiology, Infectious Diseases and Immunology, Charité University Medical Centre, 12203 Berlin, Germany. [4]Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [5]Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
*Present address: Institute for Immunology and School of Medicine, Tsinghua University, Beijing 100084, China.
†Corresponding author. Email: lim@mskcc.org

multifaceted role for PI3K in the control of glucose metabolism in Teff cells.

To investigate LDHA function in Teff cells, we used *Tbx21*^{Cre} mice that targeted H-2K^{b}-OVA^{+} Teff cells (fig. S3, A to C). *Ldha*$^{fl/fl}$ [designated as wild-type (WT)] and *Tbx21*^{Cre}*Ldha*$^{fl/fl}$ [designated as knockout (KO)] mice were infected with LM-OVA followed by rechallenge, as protective immunity is dependent on T cells during a recall response (*7*). KO mice had barely detectable antigen-specific Teff cells, in contrast to robust expansion in WT mice (Fig. 1D). Defective T cell responses were observed throughout the course of bacterial infection (Fig. 1E), and pathogen was

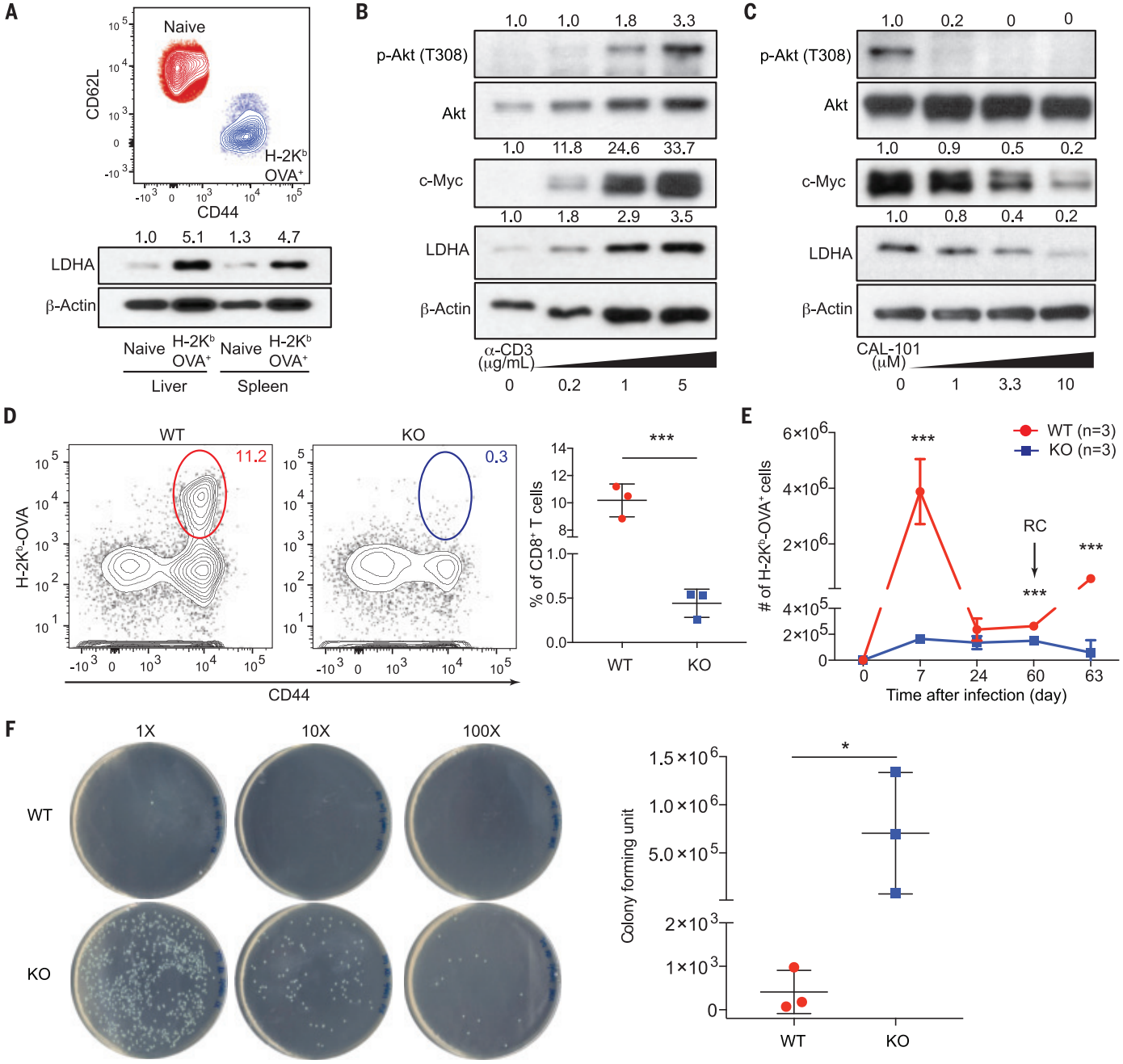


Fig. 1. PI3K-dependent expression of LDHA in CD8^{+} effector T cells is essential for antibacterial immunity. (**A**) Flow cytometric analysis of CD44 and CD62L expression in splenic naïve and H-2K^{b}-OVA^{+} CD8^{+} T cells 7 days after LM-OVA infection. The immunoblot shows normalized expression of LDHA to β-actin. (**B**) Naïve CD8^{+} T cells were stimulated with anti-CD3 in the presence of anti-CD28 and IL-2. Immunoblotting and normalized expression of p-Akt (T308) to Akt as well as c-Myc and LDHA to β-actin. (**C**) Naïve CD8^{+} T cells were stimulated with anti-CD3, anti-CD28, and IL-2 in the presence of the PI3K inhibitor CAL-101. The immunoblot shows normalized expression of p-Akt (T308) to Akt as well as that of c-Myc and LDHA to β-actin. (**D** to **F**) *Ldha*$^{fl/fl}$ (wild-type, WT) and *Tbx21*^{Cre}*Ldha*$^{fl/fl}$ (knockout, KO) mice were infected, followed (or not) by rechallenge (RC) with LM-OVA. (D) Representative flow cytometry plots (left) and frequencies (right) of splenic H-2K^{b}-OVA^{+} CD8^{+} T cells 7 days after LM-OVA infection. (E) H-2K^{b}-OVA^{+} CD8^{+} T cells were enumerated 7, 24, and 60 days after infection and on day 3 after secondary infection. (F) Splenic bacterial burden from day 3–RC mice. Data are representative of three [(A) to (C)] or two [(D) to (F), n = 3 per genotype, mean ± SD] independent experiments. *P < 0.05, ***P < 0.001 [unpaired t tests for the measurements between the two groups in (D) and (F); multiple t tests between the two groups in (E)].

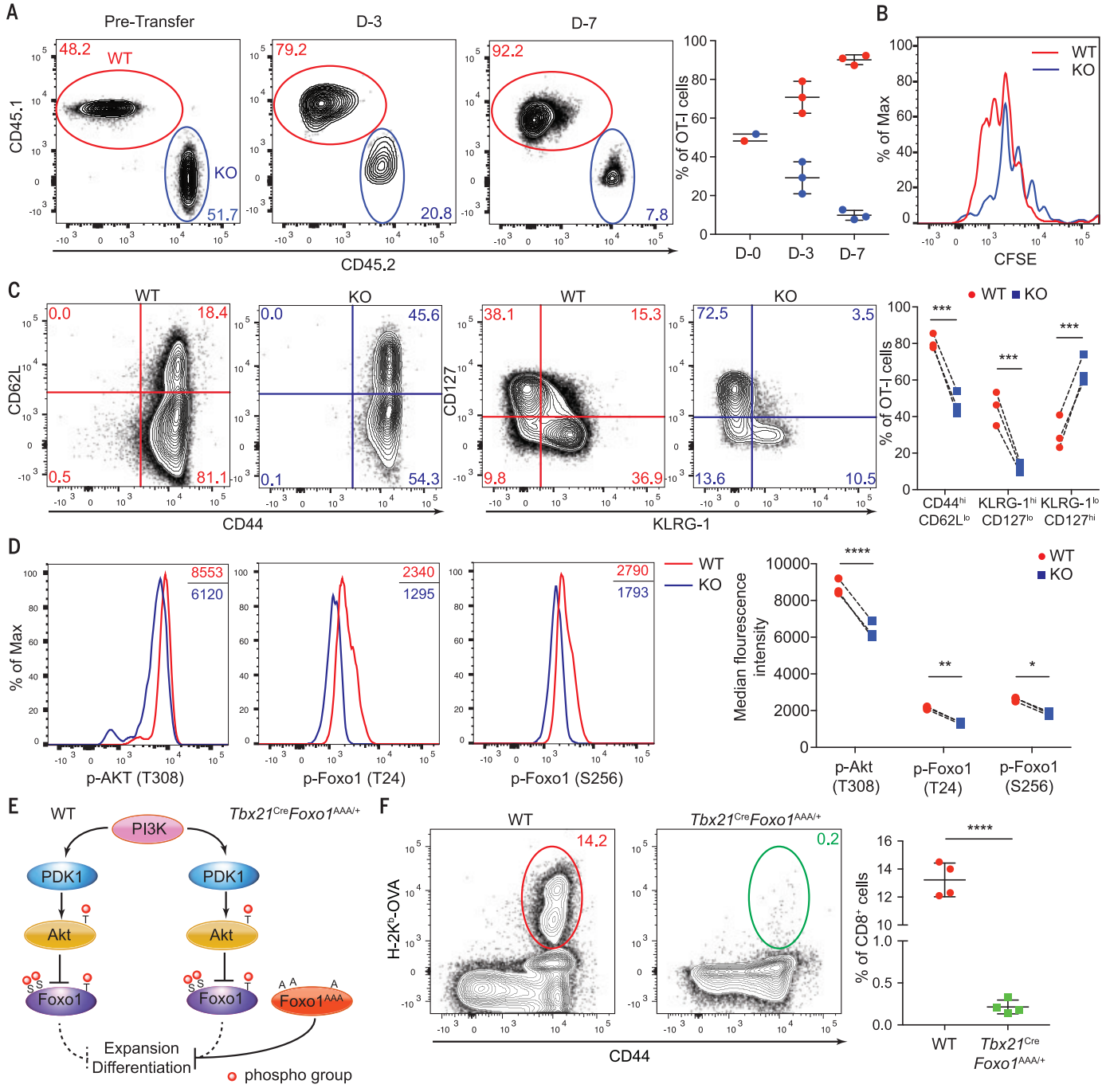


Fig. 2. LDHA deficiency impairs Akt and Foxo1 phosphorylation, thereby causing defective CD8$^+$ T cell expansion and differentiation. Congenically marked naïve *Ldha*$^{fl/fl}$ (WT) and *Tbx21*^{Cre}*Ldha*$^{fl/fl}$ (KO) OT-I T cells were labeled with carboxyfluorescein succinimidyl ester (CFSE), mixed at a 1:1 ratio, and transferred into WT recipient mice followed by infection with LM-OVA. (**A**) Representative plots and ratios (mean ± SD) of splenic WT and KO OT-I T cells at day 3 and day 7 after infection. (**B**) Proliferation of splenic WT and KO OT-I T cells at day 3 after infection, as assessed by CFSE dilution. (**C**) Representative flow cytometry plots of CD44, CD62L, CD127, and KLRG-1 expression and the percentages of CD44^{hi}CD62L^{lo}, KLRG-1^{hi}CD127^{lo}, and KLRG-1^{lo}CD127^{hi} populations among splenic WT and KO OT-I T cells from recipient mice at day 7 after infection. (**D**) Representative flow cytometry plots and median fluorescence intensities of p-Akt (T308), p-Foxo1 (T24), and p-Foxo1 (S256) in splenic WT and KO OT- I T cells. (**E**) Schematic of the PI3K-Akt-Foxo1 signaling pathway. *Foxo1*^{AAA} was engineered to resist AKT-mediated repression by replacing three phosphorylation sites [Thr^{24} (T24), Ser^{256} (S256), and Ser^{319} (S319)] with alanine (A). (**F**) Representative flow cytometry plots and frequencies of splenic H-2K^b-OVA$^+$ CD8$^+$ T cells in WT and *Tbx21*^{Cre}*Foxo1*$^{AAA/+}$ mice 7 days after LM-OVA infection (n = 4 per genotype, mean ± SD). Data in all panels are representative of at least three independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001 [paired t tests for the measurements between the two groups in (C) and (D); unpaired t test for the measurements between the two groups in (F)].

efficiently cleared in WT but not KO mice (Fig. 1F and fig. S3, D to G). Similarly diminished expansion of Teff cells was observed in KO mice infected with lymphocytic choriomeningitis virus (fig. S4, A and B). Furthermore, although WT Teff cells displayed a fully activated $CD44^{hi}CD62L^{lo}$ phenotype, KO Teff cells expressed high levels of CD62L (fig. S4, C to F). LDHA deficiency also resulted in higher proportions of $KLRG\text{-}1^{lo}CD127^{hi}$ memory precursor effector cells (MPECs) at the expense of $KLRG\text{-}1^{hi}CD127^{lo}$ short-lived effector cells (SLECs) (fig. S4, C to F) (*10*). Similar dysregulation of CD62L and KLRG1 was observed in Teff cells in KO mice infected with LM-OVA

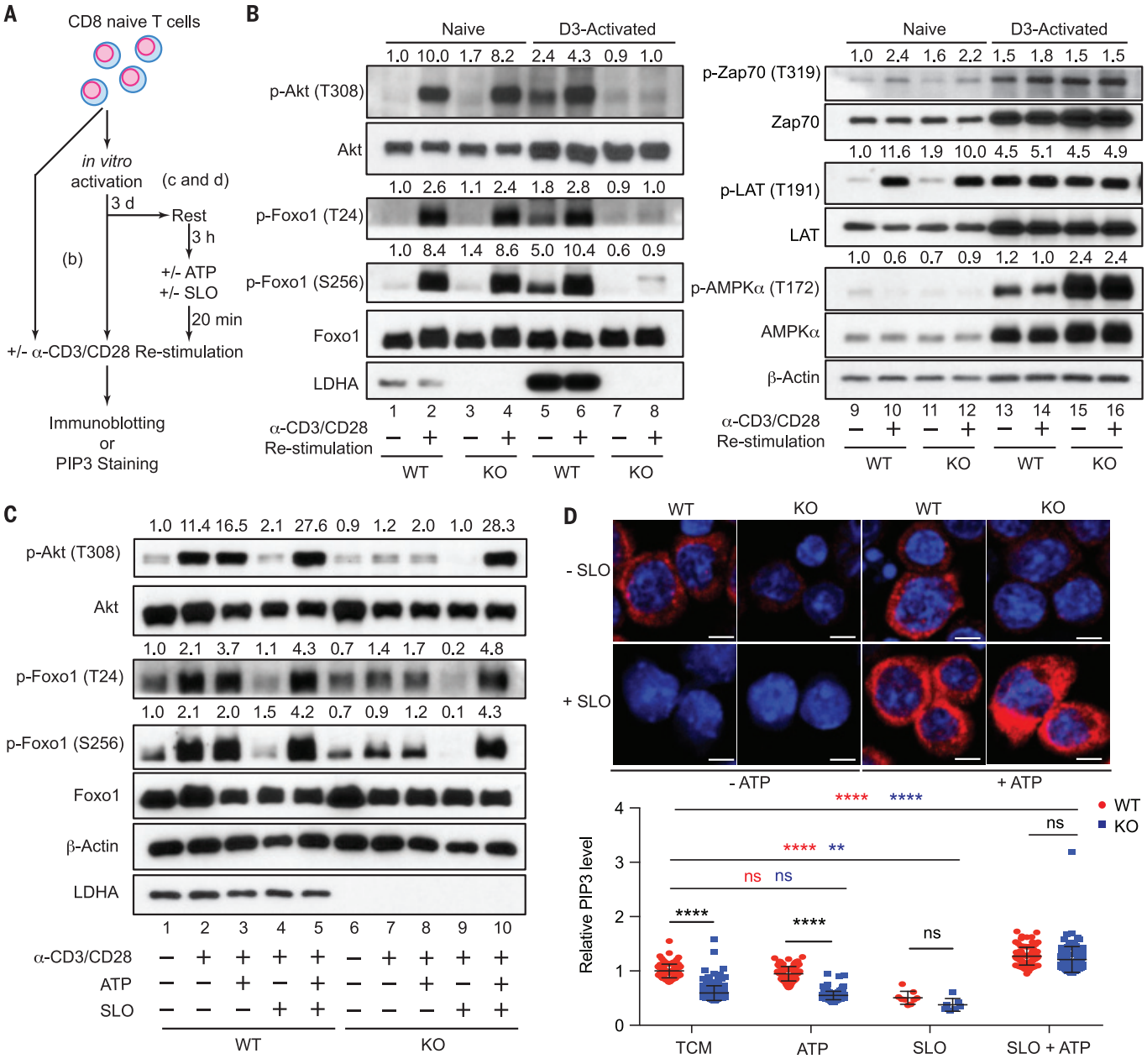


Fig. 3. Reduced glycolytic ATP production accounts for defective PI3K-Akt-Foxo1 signaling in LDHA-deficient T cells after antigen stimulation. (**A**) Schematic depicting how naïve $CD8^+$ T cells were treated for assessment of kinase signaling and PIP_3 generation. (**B**) Naïve *Ldha*$^{fl/fl}$ (WT) and *Cd4*^{Cre}*Ldha*$^{fl/fl}$ (KO) $CD8^+$ T cells were stimulated with anti-CD3, anti-CD28, and IL-2 for 3 days. Day 3–activated T cells and freshly isolated naïve WT and KO $CD8^+$ T cells were left untreated or labeled with biotinylated anti-CD3 and anti-CD28 and activated by streptavidin cross-linking. Immunoblots show normalized expression of p-Akt (T308) to Akt, p-Foxo1 (T24), or p-Foxo1 (S256) to Foxo1, p-Zap70 (T319) to Zap70, p-LAT (T191) to LAT, and p-AMPKα (T172) to AMPKα. (**C**) Day 3–activated WT and KO $CD8^+$ T cells were collected, rested in RPMI-1640 medium, and incubated in the presence or absence of ATP and/or streptolysin O (SLO). T cells were later restimulated with biotinylated anti-CD3 and anti-CD28 through streptavidin cross-linking. Immunoblots show normalized expression of p-Akt (T308) to Akt and p-Foxo1 (T24) or p-Foxo1 (S256) to Foxo1. (**D**) Representative immunofluorescent images of PIP_3 and its quantification in WT and KO $CD8^+$ T cells after receiving the indicated treatments as described in (C). Each dot represents one cell; scale bar, 5 μm. Data are representative of three [(A) to (C)] or two (D) independent experiments. $^{**}P < 0.01$, $^{****}P < 0.0001$ [two-way ANOVA test for the measurements between the two groups in (D)]; ns, not significant.

(fig. S5). Notably, attenuation of glucose metabolism with 2-deoxy-D-glucose affected Teff cell expansion, but not CD62L or KLRG1 expression (fig. S5), revealing a qualitatively distinct function for LDHA in the control of T cell activation and differentiation.

To probe $CD8^+$ T cell–intrinsic defects caused by LDHA deficiency, we crossed *Tbx21*^{Cre}*Ldha*$^{fl/fl}$ mice to the OT-I TCR-transgenic background. In a co-transfer setting, time-dependent outcompetition of KO OT-I T cells by WT OT-I T cells was observed in association with diminished proliferation (Fig. 2, A and B, and fig. S6A). Mitochondrial reactive oxygen species (ROS) have been implicated in the control of T cell proliferation (*11*). Although WT OT-I T cells had higher levels of total ROS, mitochondrial ROS was higher in KO OT-I T cells (fig. S6, B and C). This suggested that their diminished expansion was unlikely to have been caused by defective mitochondrial activity. In addition, KO OT-I T cells expressed lower levels of the effector molecules interferon-γ, tumor necrosis factor–α, and granzyme B (fig. S7). Furthermore, KO OT-I T cells failed to down-regulate CD62L and did not differentiate to SLECs (Fig. 2C and fig. S8A). Thus, LDHA deficiency causes a multitude of T cell–intrinsic defects in activation, proliferation, and differentiation.

Teff cell heterogeneity is governed by growth factor signaling that modulates the transcriptional programs of T cells (*12*). Notably, the transcription factor Foxo1 promotes CD62L expression and MPEC generation (*7*, *13*, *14*), with its activity negatively regulated by Akt-induced phosphorylation (*15*, *16*). Indeed, diminished phosphorylation of Akt and Foxo1 was observed in KO OT-I T cells (Fig. 2D and fig. S8B). In addition, defective phosphorylation of the mTORC1 target S6 was observed in KO OT-I T cells (fig. S8C). However, this would be unlikely to account for the severe defects of LDHA-deficient T cells, as ablation of mTORC1 causes a reduction of antigen-specific T cells by a factor of ~2 (*17*). To determine whether compromised Akt-induced Foxo1 phosphorylation might account for Teff

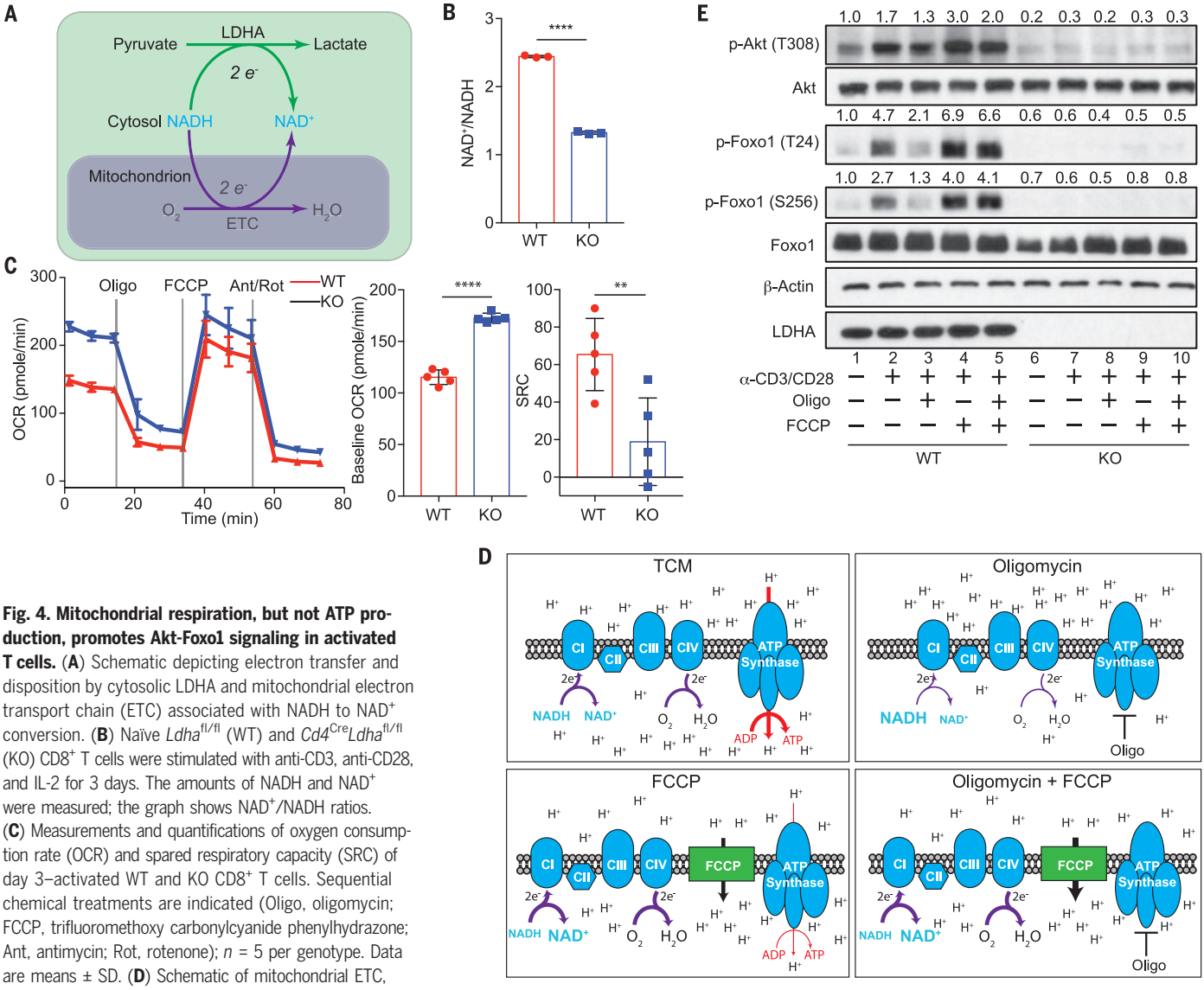


Fig. 4. Mitochondrial respiration, but not ATP production, promotes Akt-Foxo1 signaling in activated T cells. (**A**) Schematic depicting electron transfer and disposition by cytosolic LDHA and mitochondrial electron transport chain (ETC) associated with NADH to NAD^+ conversion. (**B**) Naïve *Ldha*$^{fl/fl}$ (WT) and *Cd4*^{Cre}*Ldha*$^{fl/fl}$ (KO) $CD8^+$ T cells were stimulated with anti-CD3, anti-CD28, and IL-2 for 3 days. The amounts of NADH and NAD^+ were measured; the graph shows NAD^+/NADH ratios. (**C**) Measurements and quantifications of oxygen consumption rate (OCR) and spared respiratory capacity (SRC) of day 3–activated WT and KO $CD8^+$ T cells. Sequential chemical treatments are indicated (Oligo, oligomycin; FCCP, trifluoromethoxy carbonylcyanide phenylhydrazone; Ant, antimycin; Rot, rotenone); n = 5 per genotype. Data are means ± SD. (**D**) Schematic of mitochondrial ETC, proton distribution, and ATP synthetase activity under the indicated treatment conditions. (**E**) Day 3–activated WT and KO $CD8^+$ T cells were collected and incubated in RPMI-1640 medium in the absence or presence of oligomycin and/or FCCP. T cells were later restimulated with biotinylated anti-CD3 and anti-CD28 through streptavidin cross-linking. Immunoblots show normalized expression of p-Akt (T308) to Akt and p-Foxo1 (T24) or p-Foxo1 (S256) to Foxo1. Data are representative of three independent experiments [(B), (C), and (E)]. $^{**}P < 0.01$, $^{****}P < 0.0001$ [unpaired t tests for the measurements between the two groups in (B) and (C)].

cell defects observed in KO mice, we used a mouse strain expressing a phosphorylation-resistant form of Foxo1 (Fig. 2E and fig. S9A) (*18*). Relative to WT mice, *Tbx21*^{Cre}*Foxo1*$^{AAA/+}$ mice had barely detectable H-2K^{b}-OVA^{+} Teff cells, which also failed to down-regulate CD62L expression or up-regulate KLRG-1 expression (Fig. 2F and fig. S9, B to D). Thus, LDHA promotes Akt-Foxo1 signaling in CD8^{+} Teff cells, with ectopic activation of Foxo1 recapitulating the defects associated with LDHA deficiency.

To determine whether LDHA broadly promoted Akt-Foxo1 signaling in CD8^{+} T cells, we crossed *Ldha*$^{fl/fl}$ mice with *Cd4*^{Cre} mice to delete *Ldha* in Tn cells (Fig. 3A). Notably, LDHA deficiency had negligible effects on Akt-Foxo1 signaling in Tn cells. However, activated KO T cells had diminished phosphorylation of Akt and Foxo1 but not of Zap70 or LAT (Fig. 3B and fig. S10A). This was then recapitulated with a LDHA inhibitor in WT T cells (fig. S10B). LDHA-catalyzed pyruvate-to-lactate conversion defines an efficient pathway of carbon disposal, with its deficiency predicted to lead to diminished glycolysis-associated ATP production. Indeed, phosphorylation of the cellular energy sensor AMPKα (*19*) was higher in KO Teff cells (Fig. 3B and fig. S10A), in association with reduced ATP levels and glucose usage (fig. S11, A and B).

To investigate whether reduced ATP production in antigen-experienced KO T cells accounted for diminished Akt-Foxo1 signaling, we restored ATP with streptolysin O (SLO) (fig. S11, C and D) (*20*). As expected, anti-CD3/CD28-induced Akt and Foxo1 phosphorylation was inhibited in the absence of LDHA, and this was unaltered by the addition of ATP (Fig. 3C and fig. S11E). SLO alone inhibited Akt and Foxo1 phosphorylation (Fig. 3C and fig. S11E). By contrast, ATP supplementation together with SLO enhanced Akt-Foxo1 signaling and corrected the defects associated with LDHA deficiency (Fig. 3C and fig. S11E). Because the proximal TCR signaling was unaffected (fig. S12), we examined whether the differential bioenergetic states had an impact on the generation of PI3K-catalyzed phosphatidylinositol 3,4,5-trisphosphate (PIP_3), a rate-limiting step for PDK1-mediated Akt phosphorylation (*21*). Indeed, KO T cells produced lower amounts of PIP_3, and this was unaffected by ATP supplementation (Fig. 3D). Inclusion of SLO in the absence or presence of ATP inhibited or enhanced PIP_3 generation, respectively, nullifying the differences between WT and KO T cells (Fig. 3D). Similar to murine T cells, blockade of LDHA-mediated aerobic glycolysis almost completely inhibited Akt-Foxo1 signaling in antigen-experienced human T cells, which could be corrected by SLO-assisted intracellular delivery of ATP (fig. S13). Thus, the PI3K pathway has an evolutionarily conserved ATP-sensing function.

LDHA-catalyzed conversion of pyruvate to lactate regenerates cytosolic oxidized nicotinamide adenine dinucleotide (NAD^{+}) to support glycolytic ATP generation (Fig. 4A). Indeed, the ratio of NAD^{+} to NADH was lower in activated KO T cells (Fig. 4B), in association with a lower extracellular acidification rate (ECAR) but a higher oxygen consumption rate (OCR) (Fig. 4C and fig. S14, A and B). The mitochondrial electron transport chain (ETC)–mediated oxygen reduction promotes ATP synthesis through oxidative phosphorylation as well as glycolysis when coupled with redox shuttle systems (Fig. 4A and fig. S14C). To delineate the contribution of these ETC-dependent sources of ATP to Akt-Foxo1 signaling, we used an ATP synthase inhibitor, oligomycin, and an uncoupler, FCCP (Fig. 4, C and D, and fig. S14C). Oligomycin attenuated Akt and Foxo1 phosphorylation in WT T cells, whereas Akt-Foxo1 signaling in KO T cells was barely detectable (Fig. 4E and fig. S14D). Notably, treatment with FCCP or oligomycin plus FCCP enhanced Akt-Foxo1 signaling in WT T cells in an ETC-dependent manner, but not in KO T cells (Fig. 4E and figs. S14D and S15), in agreement with their higher spare respiratory capacity (Fig. 4C). In contrast to activated T cells, WT and KO Tn cells exhibited comparable ECAR and OCR (fig. S16, A to C) and were sensitive to oligomycin- and FCCP-mediated inhibition of Akt and Foxo1 phosphorylation (fig. S16D). Thus, Akt-Foxo1 signaling is supported by mitochondrial and glycolytic ATP in Tn and Teff cells, respectively. Furthermore, glycolytic ATP production is enabled primarily by LDHA, with ETC playing a subsidiary role (fig. S17).

Antigen stimulation of CD8^{+} T cells induces glucose uptake to support effector functions (*8*, *22*–*25*). A selective role for glycolytic ATP to fuel PI3K signaling in Teff cells can be consequent to remodeling of the mitochondrial network (*26*), which may spare it as an energy house for glucose anabolism including acetyl–coenzyme A–mediated gene regulation (*5*). Diminished glycolytic ATP production in LDHA-deficient Teff cells impairs PI3K but not proximal TCR signaling. Notably, Zap70 has an ATP Michaelis constant K_m of ~3 μM (*27*), whereas PI3Kδ, the major isoform expressed in T cells, has an ATP K_m of ~118 μM (*28*), which may cause PI3Kδ to function as an ATP sensor in gauging the bioenergetic state of Teff cells. As enhanced PI3K signaling is a major oncogenic event (*29*), future studies will be needed to determine whether the PI3K-LDHA positive feedback circuit provides a mechanistic explanation for the Warburg effect that is observed in cancer cells (*30*). Such a bioenergetic mechanism for the regulation of signaling may also guide the development of effective vaccines for infectious diseases and therapeutics for cancer.

REFERENCES AND NOTES

1. R. Wang, D. R. Green, *Nat. Immunol.* **13**, 907–915 (2012).
2. N. J. MacIver, R. D. Michalek, J. C. Rathmell, *Annu. Rev. Immunol.* **31**, 259–283 (2013).
3. N. M. Chapman, M. R. Boothby, H. Chi, *Nat. Rev. Immunol.* **20**, 55–70 (2020).
4. E. L. Pearce, M. C. Poffenberger, C. H. Chang, R. G. Jones, *Science* **342**, 1242454 (2013).
5. M. Peng *et al.*, *Science* **354**, 481–484 (2016).
6. E. H. Ma *et al.*, *Immunity* **51**, 856–870.e5 (2019).
7. M. V. Kim, W. Ouyang, W. Liao, M. Q. Zhang, M. O. Li, *Immunity* **39**, 286–297 (2013).
8. R. Wang *et al.*, *Immunity* **35**, 871–882 (2011).
9. J. A. Wofford, H. L. Wieman, S. R. Jacobs, Y. Zhao, J. C. Rathmell, *Blood* **111**, 2101–2111 (2008).
10. N. S. Joshi *et al.*, *Immunity* **27**, 281–295 (2007).
11. L. A. Sena *et al.*, *Immunity* **38**, 225–236 (2013).
12. S. M. Kaech, W. Cui, *Nat. Rev. Immunol.* **12**, 749–761 (2012).
13. R. R. Rao, Q. Li, M. R. Gubbels Bupp, P. A. Shrikant, *Immunity* **36**, 374–387 (2012).
14. R. Hess Michelini, A. L. Doedens, A. W. Goldrath, S. M. Hedrick, *J. Exp. Med.* **210**, 1189–1200 (2013).
15. W. Ouyang, M. O. Li, *Trends Immunol.* **32**, 26–33 (2011).
16. S. M. Hedrick, R. Hess Michelini, A. L. Doedens, A. W. Goldrath, E. L. Stone, *Nat. Rev. Immunol.* **12**, 649–661 (2012).
17. K. N. Pollizzi *et al.*, *J. Clin. Invest.* **125**, 2090–2108 (2015).
18. W. Ouyang *et al.*, *Nature* **491**, 554–559 (2012).
19. J. Rolf *et al.*, *Eur. J. Immunol.* **43**, 889–896 (2013).
20. I. Walev *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3185–3190 (2001).
21. P. S. Costello, M. Gallagher, D. A. Cantrell, *Nat. Immunol.* **3**, 1082–1089 (2002).
22. C. M. Cham, T. F. Gajewski, *J. Immunol.* **174**, 4670–4677 (2005).
23. P. M. Gubser *et al.*, *Nat. Immunol.* **14**, 1064–1072 (2013).
24. P. C. Ho *et al.*, *Cell* **162**, 1217–1228 (2015).
25. A. V. Menk *et al.*, *Cell Rep.* **22**, 1509–1521 (2018).
26. M. D. Buck *et al.*, *Cell* **166**, 63–76 (2016).
27. N. Isakov, R. L. Wange, J. D. Watts, R. Aebersold, L. E. Samelson, *J. Biol. Chem.* **271**, 15753–15761 (1996).
28. J. R. Somoza *et al.*, *J. Biol. Chem.* **290**, 8439–8446 (2015).
29. M. D. Goncalves, B. D. Hopkins, L. C. Cantley, *N. Engl. J. Med.* **379**, 2052–2062 (2018).
30. O. Warburg, *J. Cancer Res.* **9**, 148–163 (1925).

ACKNOWLEDGMENTS

We thank members of the Li laboratory and M. Huse for helpful discussions. **Funding:** Supported by NIH grant R01 AI 102888 and a Howard Hughes Medical Institute Faculty Scholar Award (M.O.L.) and by Memorial Sloan Kettering Cancer Center Support Grant/Core Grant P30 CA08748. **Author contributions:** M.O.L. conceived the project; K.X. designed and performed most experiments with input from M.O.L., N.Y., and M.P.; K.X. and E.G.S. conducted the immunofluorescence imaging experiments; K.X. and A.S. performed cell sorting experiments; P.L., X.Z., M.H.D., Z.W., and C.C. assisted with the design of experiments; K.J.C. helped managing the mouse colony; A.G.L. and A.Y.R. provided critical mouse strains; and K.X. wrote and M.O.L. edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. The mouse strains generated for this study, *Tbx21*^{Cre}*Ldha*$^{fl/fl}$, *Tbx21*^{Cre}*Foxo1*$^{AAA/+}$, and *Cd4*^{Cre}*Ldha*$^{fl/fl}$, are available from the corresponding author under a material transfer agreement with the Memorial Sloan Kettering Cancer Center.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/405/suppl/DC1
Materials and Methods
Figs. S1 to S17
Reference (*31*)

11 February 2020; resubmitted 29 August 2020
Accepted 11 December 2020
10.1126/science.abb2683

MOSQUITO BIOLOGY

Clock genes and environmental cues coordinate *Anopheles* pheromone synthesis, swarming, and mating

Guandong Wang[1,2]*, Joel Vega-Rodríguez[3]*†, Abdoulaye Diabate[4]*, Jingnan Liu[5]*, Chunlai Cui[1,2], Charles Nignan[4], Ling Dong[1,2], Fang Li[1,2], Cheick Oumar Ouedrago[4], Abdoul Malik Bandaogo[4], Péguédwindé Simon Sawadogo[4], Hamidou Maiga[4], Thiago Luiz Alves e Silva[6], Tales Vicari Pascini[6], Sibao Wang[1,2]‡§, Marcelo Jacobs-Lorena[3]‡§

Anopheles mating is initiated by the swarming of males at dusk followed by females flying into the swarm. Here, we show that mosquito swarming and mating are coordinately guided by clock genes, light, and temperature. Transcriptome analysis shows up-regulation of the clock genes *period* (*per*) and *timeless* (*tim*) in the head of field-caught swarming *Anopheles coluzzii* males. Knockdown of *per* and *tim* expression affects *Anopheles gambiae s.s.* and *Anopheles stephensi* male mating in the laboratory, and it reduces male *An. coluzzii* swarming and mating under semifield conditions. Light and temperature affect mosquito mating, possibly by modulating *per* and/or *tim* expression. Moreover, the desaturase gene *desat1* is up-regulated and rhythmically expressed in the heads of swarming males and regulates the production of cuticular hydrocarbons, including heptacosane, which stimulates mating activity.

The malaria parasite, transmitted by *Anopheles* mosquitoes, infects more than 200 million people and causes nearly half a million deaths annually (*1*). In the absence of an effective vaccine, mosquito control using insecticide-treated nets and indoor residual spraying remains the most effective means of combating the disease (*2*). However, mosquito control is being affected by resistance to commonly used insecticides (*3*, *4*). Thus, there is an urgent need to develop alternative tools to control mosquito populations (*5*). Male mating behavior is a key aspect of the mosquito life cycle, and targeting this behavior has shown promise for vector control (*6*).

Males of numerous dipteran species, including many anopheline and culicine mosquitoes, form flight swarms that contain tens to thousands of males as a prerequisite for mating (*7*–*13*). Females fly into the swarm, where they select a male for copulation (*14*). After insemination, the female is generally not receptive to copulation for the remainder of the female's life (*15*).

Day-night cycle, light intensity, and temperature regulate the circadian rhythmicity of behavior and physiology of most organisms (*16*–*18*). For example, *Anopheles* mosquitoes swarm and mate daily at dusk (*8*). However, the molecular mechanisms that modulate mosquito swarming and mating behavior have yet to be clarified. Understanding mating biology is necessary to guide the development and implementation of vector control programs by means of the release of either conventional sterile or genetically engineered males (*19*). In this work, we undertook experiments in the laboratory and in near-field conditions to identify components of the circadian clock apparatus that regulate male swarming and mating behavior in *Anopheles* mosquitoes.

To identify genes involved in swarming and mating behaviors, we collected *Anopheles coluzzii* males at dusk from swarms and from houses (resting, nonswarming) in Vallée du Kou, Burkina Faso, and isolated RNA from their heads. *Anopheles gambiae s.s.* genome microarray analysis showed that 87 genes significantly differed (≥1.75-fold change, Student's *t* test, $P < 0.01$) between the two groups, including 45 up-regulated and 42 down-regulated genes in the heads of swarming versus nonswarming males (Fig. 1A and fig. S1). Notably, up-regulated genes in swarming males were overrepresented in the functional categories of circadian clock and metabolism, including two canonical clock genes—*period* (*per*) (AGAP001856) and *timeless* (*tim*) (AGAP006376)—and nine genes involved in carbohydrate, amino acid, or lipid metabolism (Fig. 1B). By contrast, all four differentially expressed genes (DEGs) encoding immunity were down-regulated in swarming male heads compared with nonswarming male heads (Fig. 1B). We used quantitative reverse transcription polymerase chain reaction (qRT-PCR) to validate DEGs associated with circadian clock and metabolism. Most genes tested had the expected directional changes (table S1). Collectively, global transcription analysis indicates that circadian clock and metabolism genes are associated with male swarming behavior.

To address the hypothesis that clock genes influence male swarming and mating activity, we silenced *per* and *tim* in *Anopheles stephensi* virgin males by systemic injection of double-stranded RNA (ds*per* or ds*tim*) (fig. S2). Injected males were allowed to mate with virgin females, and the presence of sperm in the female spermatheca was used as a mating indicator (fig. S3). Although swarming is not a common behavior in laboratory small cages, male *Anopheles* mosquitoes still exhibit swarm-like mating flight activity (*14*, *20*). We confirmed that *Anopheles* males display evening peak flight activity in laboratory cages and that knockdown of *per* or *tim* gene expression in *An. stephensi* males significantly reduces flying activity (fig. S4). Silencing of *per* or *tim* significantly affected male mating flight activity around Zeitgeber time (ZT) ZT13 to ZT15, when peak mating activity of double-stranded green fluorescent protein gene (ds*GFP*)–treated mosquitoes (control) occurs in small cages (fig. S4 and movies S1 and S2). Reduction in flight activity correlated with significant reduction of mating (female insemination rate) (Fig. 1, C and D). These results suggest that the two core clock genes *per* and *tim* regulate mating activity in *Anopheles* mosquitoes.

To determine the role of *per* and *tim* in mosquito swarming and mating activity in a semifield setting, we performed experiments in a large outdoor mosquito sphere in the Vallée du Kou in Burkina Faso. Locally captured and reared *An. coluzzii* virgin male mosquitoes were injected with double-stranded RNA encoding *per*, *tim*, or *GFP* as control. Three days later, the injected males and virgin females were released before sunset in compartments with a swarming marker (1-m^2 black cloth) (*21*), and swarming behavior was monitored for 1 hour. We observed a significant reduction in swarm maximum height and swarm size (number of swarming males) in male mosquitoes injected with ds*per* or ds*tim* (Fig. 1, E and F). Moreover, mating was significantly reduced in males injected with ds*per* (40% insemination) or ds*tim* (47% insemination) compared with male mosquitoes injected with ds*GFP* (70% insemination) (Fig. 1G). These reductions are consistent with those observed in a laboratory setting (Fig. 1C). Taken together, these results indicate that the circadian genes *per* and *tim* regulate male *Anopheles* swarming and mating activity.

[1]CAS Key Laboratory of Insect Developmental and Evolutionary Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences (CAS), Shanghai, China. [2]CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, China. [3]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health and Malaria Research Institute, Baltimore, MD, USA. [4]Institut de Recherche en Science de la Santé (IRSS), Bobo-Dioulasso, Burkina Faso. [5]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [6]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, USA.

*These authors contributed equally to this work. †Present address: Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, USA. ‡These authors contributed equally to this work.

§Corresponding author. Email: sbwang@cemps.ac.cn (S.W.); ljacob13@jhu.edu (M.J.-L.)

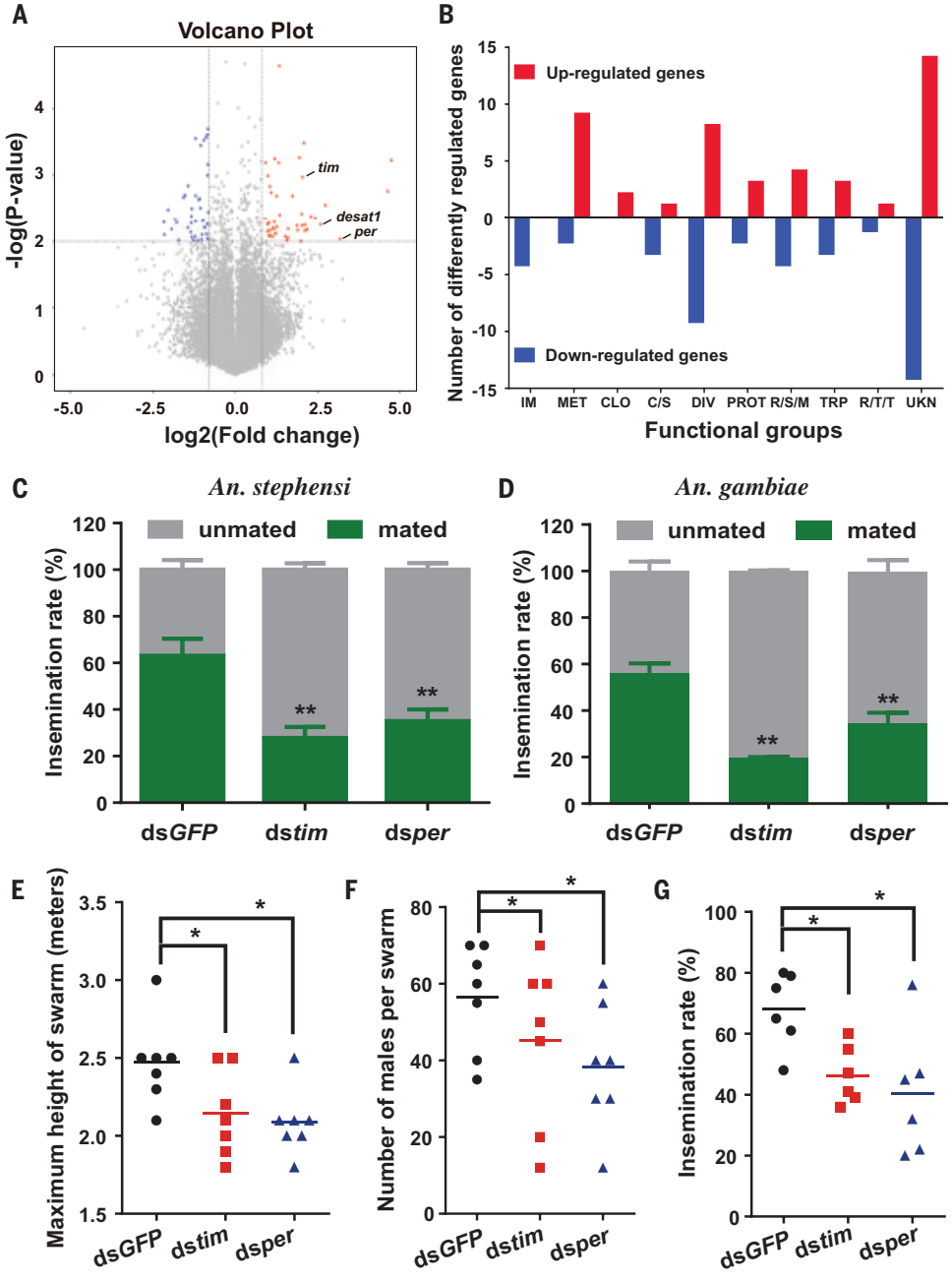


Fig. 1. Mosquito swarming and mating are regulated by the clock genes *per* and *tim*. (**A** and **B**) Transcriptome analysis of *An. coluzzii* head gene expression between swarming and nonswarming male mosquitoes. (A) Volcano plot of DEGs. Significantly up-regulated and down-regulated genes are marked with red and blue spots, respectively. (B) DEGs by gene ontology category in swarming male heads compared with non-swarming males. Differential gene expression was considered significant when the expression fold change (swarming versus nonswarming) was ≥0.75 on a log2 scale ($P < 0.01$). IM, immunity; MET, metabolism; CLO, clock; C/S, cytoskeletal and structural; DIV, diverse; PROT, proteolysis; R/S/M, oxidoreductive, stress-related, and mitochondrial; TRP, transport; R/T/T, replication, transcription, and translation; UKN, unknown function. (**C** and **D**) Silencing of *tim* and *per* in virgin *An. stephensi* (C) or *An. gambiae s.s.* (D) males causes a reduction in the rate of female insemination, as determined by light microscopy examination for the presence of sperm in the dissected female spermatheca (fig. S3). A total of 60 virgin males were injected with double-stranded RNA for each treatment. Error bars indicate SEMs. Similar results were obtained from two biological repeats. (**E** to **G**) Knockdown of *tim* or *per* in virgin *An. coluzzii* males affects the swarm maximum height (E), swarm size (number of swarming males per swarm) (F), and mating frequency (G) under semifield conditions (Burkina Faso). Each symbol denotes the result of an independent experiment ($n = 7$). Horizontal lines represent the means. Statistics were performed with Student's *t* test or one-way analysis of variance (ANOVA). *$P < 0.05$; **$P < 0.01$.

The circadian clock is intimately regulated by temperature and light fluctuations (*22*). To investigate the effect of temperature on mosquito mating, 3-day-old virgin male and female *An. stephensi* mosquitoes were pooled and kept overnight at 27°, 19°, or 34°C. We found that low temperature (19°C) and high temperature (34°C) significantly inhibited mosquito mating activity compared with the optimum temperature of 27°C (Fig. 2A).

We next examined whether light influences mosquito mating activity by comparing mating frequencies of mosquitoes that were maintained either in light-dark (LD) conditions (lights off at night) or light-light (LL) conditions (lights on at night). The mating frequency of mosquitoes in the LL condition was significantly lower than that of mosquitoes in the LD condition (Fig. 2B). These results suggest that adverse temperatures and continuous light strongly affect mosquito mating. Moreover, the rhythm of *per* and *tim* transcripts in *An. stephensi* male heads is strongly affected by changes in ambient temperature. Peak levels of *per* mRNA were higher at 19°C and lower at 34°C compared with those at 27°C (Fig. 2C). Contrastingly, the highest levels of *tim* expression were observed at 34°C with successive reduction at lower temperatures (Fig. 2D). Furthermore, the transcription abundance of *per* and *tim* in male heads was significantly lower under the continuous-light condition (Fig. 2, E and F), which indicates that prolonged exposure to LL conditions reduces the expression of *per* and *tim* in the evening.

Of 86 DEGs, one unknown gene (AGAP002588) and a cuticular protein gene *CPR30* (AGAP006009) were among the most up-regulated genes (~25-fold) in the heads of swarming males compared with nonswarming males (table S1). Additionally, an ortholog to the *Drosophila desaturase 1* (*desat1*) gene required for pheromone production (*desat1*, AGAP001713) and a P450 gene *CYP325G1* (AGAP002196) were also highly up-regulated in swarming male heads (table S1).

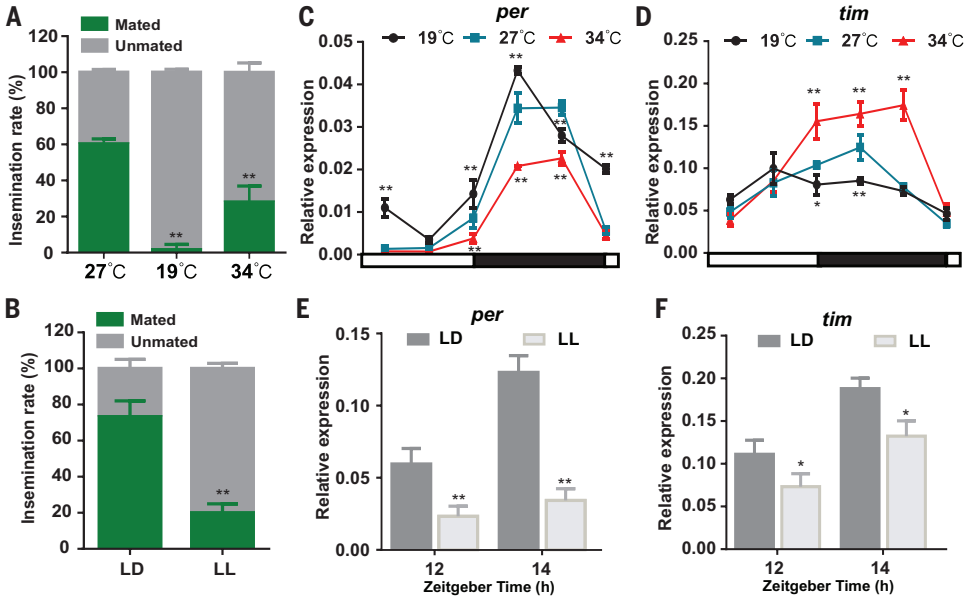


Fig. 2. Temperature and light affect mating activity. Newly emerged male and female mosquitoes were kept separately for 3 days at 27°C under 12-hour–12-hour LD cycling before performing the experiment. (**A**) Temperature affects mosquito mating activity. Males and females were pooled; kept overnight at 19°, 27°, or 34°C; and female insemination rates were determined. (**B**) Light inhibits mosquito mating activity. Males and females were pooled, kept in LD or LL conditions, and female insemination rates were determined. (**C** and **D**) Effect of temperature on *per* (C) and *tim* (D) transcript abundance in male *An. stephensi* heads under LD cycling. Mosquitoes were kept under LD cycles at 19°, 27°, or 34°C. *per* and *tim* transcript rhythm in male heads was measured by qRT-PCR. The housekeeping gene *RPS7* (*AsS7*) was used as an internal reference in qRT-PCR assays. (**E** and **F**) Effect of prolonged light on *per* (E) and *tim* (F) transcript abundance in male *An. stephensi* heads. Male mosquitoes were maintained for 3 days at 27°C under LD cycling and then either kept in the LD condition or switched to the LL condition. Mosquitoes were collected at ZT12 (lights off) and ZT14 (2 hours after lights off) for detection of *per* and *tim* expression levels. Error bars indicate SEMs. *$P < 0.05$; **$P < 0.01$ (Student's *t* test). Similar results were obtained from two biological repeats.

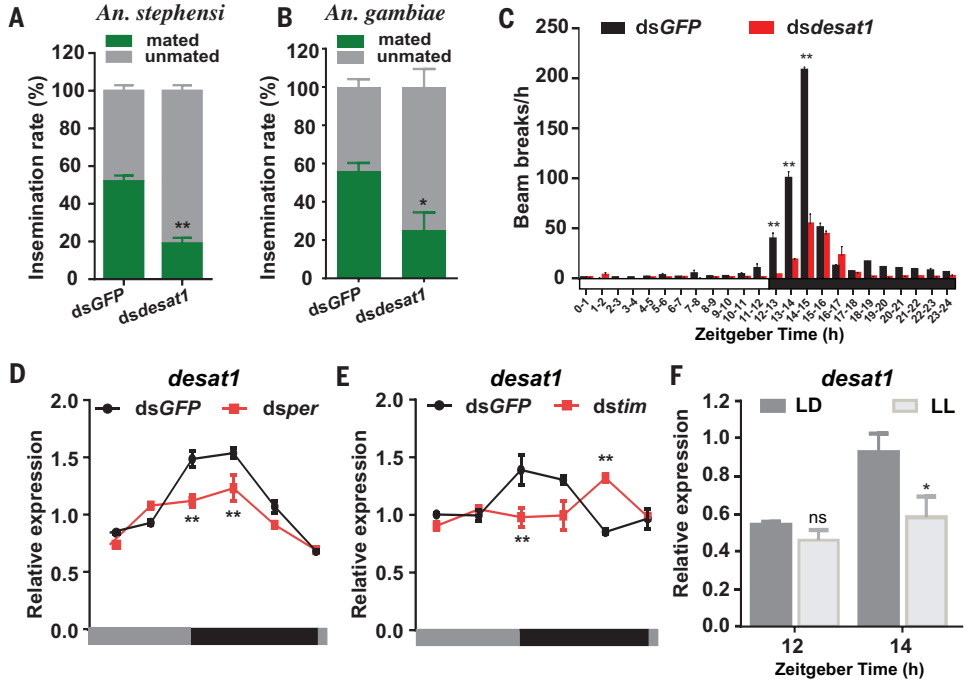


Fig. 3. *Desat1* regulates male flight activity and mating. (**A** and **B**) Silencing *desat1* expression in *An. stephensi* (A) and *An. gambiae s.s.* (B) virgin males causes reduction of female insemination. (**C**) Flight activity of *An. stephensi* virgin male mosquitoes injected with ds*GFP* (black) or ds*desat1*(red) and maintained under 12-hour–12-hour LD cycles. Values show the total activity within each hourly time bin (mean ± SEM) of 16 mosquitoes in one treatment. The white and black bar below the graph denotes when lights were on and off, respectively. Lights on occurred at ZT0 and lights off occurred at ZT12. (**D** and **E**) Silencing *per* (D) or *tim* (E) influences the expression pattern of *destal1* in *An. stephensi* male heads maintained under DD (constant dark) condition. The housekeeping gene *RPS7* (*AsS7*) was used as the internal control for qRT-PCR. Data were normalized to median fold change. Circadian day and night are indicated by the horizontal gray and black (indicating subjective day and subjective night, respectively) bars below the charts. (**F**) Effect of prolonged light on *desat1* transcript abundance in male *An. stephensi* heads. After being maintained at 27°C under LD cycling for 3 days, male mosquitoes were either kept in the LD condition or switched to the LL condition. Mosquitoes were collected at ZT12 (lights off) and ZT14 (2 hours after lights off) for detection of *desat1* expression levels. Error bars indicate SEMs. ns, not significant; *$P < 0.05$; **$P < 0.01$ (Student's *t* test). Similar results were obtained from three experimental repeats.

The expression of *desat1*, the AGAP002588 gene homolog, and *CPR30*—but not *CYP325G1*—exhibited robust, high-amplitude rhythm in synchrony with the two canonical clock genes *per* and *tim* (fig. S5). The expression patterns of these DEGs in different mosquito tissues is presented in fig. S6.

We next determined whether these genes play any role in mosquito mating. Knockdown of *desat1* expression in *An. stephensi* and *An. gambiae s.s.* virgin males significantly reduced female insemination rates (Fig. 3, A and B) and markedly reduced male mating flight activity (Fig. 3C and movie S3), whereas silencing of the AGAP002588 gene homolog, *CPR30*, and *CYP325G1* did not affect female insemination rate (fig. S7). Moreover, we found that knockdown of *per* or *tim* significantly altered the amplitude or rhythm of *desat1* expression in *An. stephensi* male heads under a dark-dark (DD) (constant dark) condition

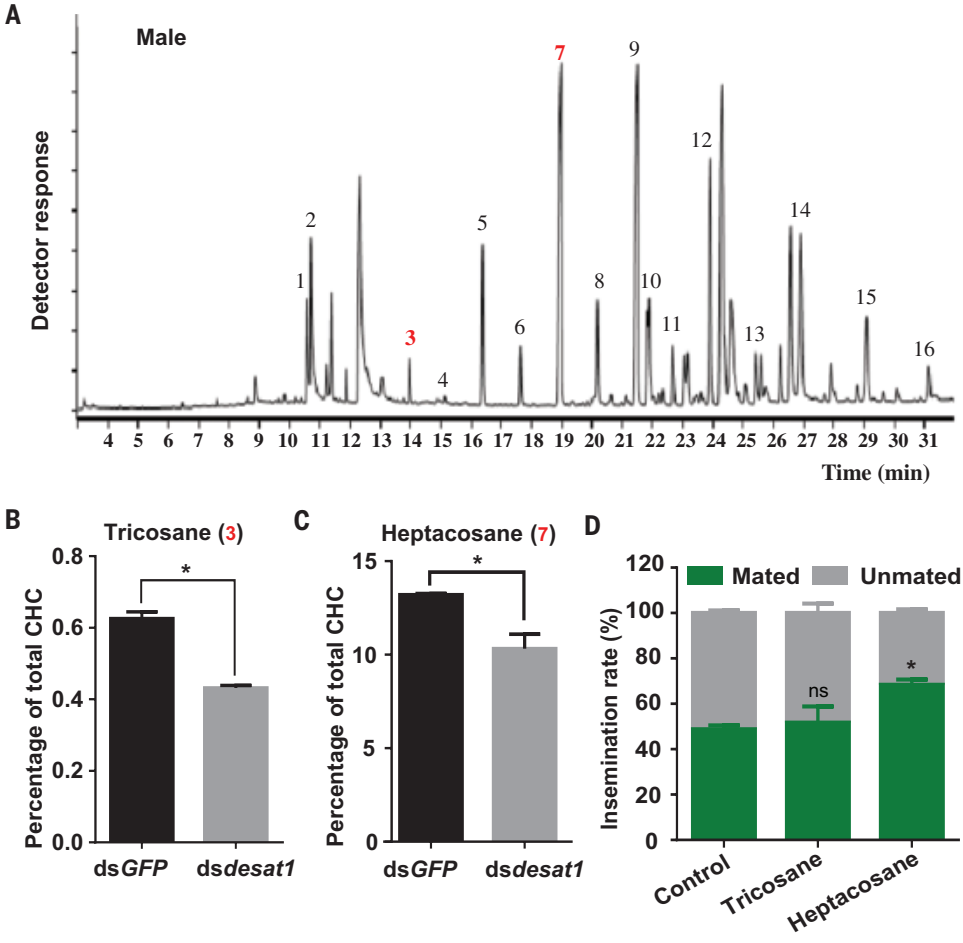


Fig. 4. Silencing of *desat1* changes the CHC profile of male mosquitoes. (**A**) Representative CHC profile of virgin male *An. stephensi* mosquitoes. Numbers above the peaks correspond to peak numbers given in table S2. (**B** and **C**) Comparison of CHCs between mosquitoes injected with ds*GFP* and ds*desat1*. Silencing of *desat1* resulted in a significant reduction of the relative percentage of tricosane (peak 3) (B) and heptacosane (peak 7) (C). (**D**) Effect of the CHCs tricosane and heptacosane on mosquito mating activity. Perfuming *Anopheles* virgin males with tricosane did not alter mating activity, whereas perfuming males with heptacosane significantly increased female insemination. Error bars indicate SEMs. *$P < 0.05$ (Student's *t* test).

(Fig. 3, C and D), which indicates that *desat1* is a clock-controlled gene.

We next investigated whether temperature and light affect the expression of *desat1*. We found that *desat1* transcript rhythm in the *An. stephensi* male heads is not affected by changes in ambient temperature (fig. S8), but under prolonged light exposure, *desat1* transcript abundance decreases significantly (Fig. 3F). These results indicate that light exposure in the evening inhibits *desat1* expression, which may be related to the inhibition of mosquito mating by light (Fig. 2B).

Many dipterans use cuticular hydrocarbons (CHCs) as sex pheromones for species and sex recognition and courtship (*23*–*25*), but little is known about the possible role of CHCs in mosquito mating and chemical communication. In *Drosophila*, *desat1* regulates the biosynthesis of CHCs (*26*), several of which function as sex pheromones (*27*). To examine whether *desat1* regulates the biosynthesis of *Anopheles* CHCs, we determined the hydrocarbon profile by gas chromatography of cuticular extracts of virgin male *An. stephensi* whole body. Sixteen major CHCs were identified (Fig. 4A and table S2). Silencing of *desat1* resulted in significant reduction of tricosane and heptacosane representation (Fig. 4, B and C), which indicates that *desat1*, a circadian regulated gene, is involved in the production of *Anopheles* CHCs. We next examined the role of tricosane and heptacosane in *Anopheles* mating activity and found that perfuming *Anopheles* males with tricosane did not alter mating activity, whereas perfuming males with heptacosane lead to a significant increase in the rate of female insemination (Fig. 4D and fig. S9). These results suggest that heptacosane enhances the interaction between males and courting females.

Understanding mosquito mating biology is central to any successful genetic-control strategies. Our transcriptome analysis showed enrichment of up-regulated metabolism and circadian clock genes and down-regulated immune genes in swarming mosquitoes, as swarming is an energy-consuming process accompanied by increased metabolic activity that occurs at dusk.

The circadian clock is cell-autonomous and consists of a series of interlocking transcriptional feedback loops requiring *per* and *tim*. Our work shows that these core clock genes are markedly up-regulated in swarming male mosquitoes compared with nonswarming males, and knocking down their expression affects male swarming (lowering swarm height and size) and mating in the laboratory and in semifield conditions.

Light has been shown to influence mosquito blood feeding, gene expression, fecundity, and reproductive barrier (*14*, *28*–*33*). Our study shows that adverse temperature (lower or higher) and increased light exposure during dark periods inhibit mating activity and influence the expression of *per* and *tim*, which indicates that the two factors interact to regulate the timing of male swarming. In *Drosophila*, temperature regulates the levels and activity of PER and TIM by modulating the genes' splicing patterns (*34*, *35*). Similarly, our study shows that the *per* and *tim* transcript levels in mosquitoes are affected by ambient temperatures. Therefore, light and temperature together entrain the circadian clock, influencing the programming of mosquito swarming and mating (fig. S10).

In many dipteran species, sex pheromones—together with visual, tactile, and acoustic cues—play important roles in courtship behavior preceding mating (*36*–*38*). CHCs act as sex pheromones in mate recognition and chemical communication for many insect genera (*23*, *27*). In *Drosophila*, these CHC pheromones are synthesized by *desat1* in secretory cells

called oenocytes in the abdominal epidermis, and they are then deposited on the cuticle (*23*, *26*). No *Anopheles* pheromone has been identified (*39*). Notably, we found in mating bioassays that perfuming virgin males with the C27 linear heptacosane promotes mosquito mating activity. Heptacosane is known to be a contact sex pheromone that facilitates mating activity of the tea weevil *Myllocerinus aurolineatus* (*40*). We found that *desat1* is a clock-controlled gene in mosquitoes and that light inhibits its transcription in the evening, reducing the production of CHCs and mating activity. A previous study has shown that the abundance of CHCs, including heptacosane (nC27), in male *Drosophila* varies in response to light and time of day (*41*), which indicates that male CHC profiles are a dynamic trait.

This work provides mechanistic insights on the molecular, chemical, and environmental factors regulating *Anopheles* swarming and mating behaviors. It contributes to a better understanding of mating mechanisms, which might lead to vector control strategies that target insect reproductive behavior.

REFERENCES AND NOTES

1. World Health Organization (WHO), World malaria report 2018 (2018); www.who.int/malaria/publications/world-malaria-report-2018/report/en/.
2. E. Chanda, B. Ameneshewa, M. Bagayoko, J. M. Govere, M. B. Macdonald, *Trends Parasitol.* **33**, 30–41 (2017).
3. H. Ranson, N. Lissenden, *Trends Parasitol.* **32**, 187–196 (2016).
4. G. Harrison, *Mosquitos, Malaria, and Man: A History of the Hostilities Since 1880* (Dutton, 1978).
5. C. Cui *et al*., *Nat. Commun.* **10**, 4298 (2019).
6. A. Diabate, F. Tripet, *Parasit. Vectors* **8**, 347 (2015).
7. A. Diabaté *et al*., *BMC Evol. Biol.* **11**, 184 (2011).
8. P. S. Sawadogo *et al*., *Acta Trop.* **132**, S42–S52 (2014).
9. S. Butail *et al*., *J. R. Soc. Interface* **9**, 2624–2638 (2012).
10. J. A. Downes, *Annu. Rev. Entomol.* **14**, 271–298 (1969).
11. E. W. Kaindoa *et al*., *Wellcome Open Res.* **2**, 88 (2017).
12. N. C. Manoukis *et al*., *J. Med. Entomol.* **46**, 227–235 (2009).
13. B. S. Assogba *et al*., *Acta Trop.* **132**, S53–S63 (2014).
14. J. D. Charlwood, M. D. R. Jones, *Physiol. Entomol.* **5**, 315–320 (1980).
15. P. Gabrieli *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16353–16358 (2014).
16. J. M. Hurley, J. J. Loros, J. C. Dunlap, *Trends Biochem. Sci.* **41**, 834–846 (2016).
17. Y. Zhang, P. Emery, in *Insect Molecular Biology and Biochemistry*, L. Gilbert, Ed. (Academic Press, 2012), pp. 513–551.
18. T. Sakai, N. Ishida, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9221–9225 (2001).
19. B. G. J. Knols, H. C. Bossin, W. R. Mukabana, A. S. Robinson, *Am. J. Trop. Med. Hyg.* **77**, 232–242 (2007).
20. L. Facchinelli *et al*., *Malar. J.* **14**, 271 (2015).
21. A. Niang *et al*., *Parasit. Vectors* **12**, 446 (2019).
22. P. Lamba, D. Bilodeau-Wentworth, P. Emery, Y. Zhang, *Cell Rep.* **7**, 601–608 (2014).
23. C. Wicker-Thomas, I. Guenachi, Y. F. Keita, *BMC Biochem.* **10**, 21 (2009).
24. A. R. Polerstock, S. D. Eigenbrode, M. J. Klowden, *J. Med. Entomol.* **39**, 545–552 (2002).
25. G. J. Blomquist, J. A. Tillman-Wall, L. Guo, D. R. Quilici, P. Gu, C. Schal, in *Insect Lipids: Chemistry, Biochemistry and Biology*, D. W. Stanley-Samuelson, D. R. Nelson, Eds. (Univ. Nebraska Press, 1993), pp. 317–351.
26. F. Bousquet *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 249–254 (2012).
27. R. R. H. Anholt, P. O'Grady, M. F. Wolfner, S. T. Harbison, *Genetics* **214**, 49–73 (2020).
28. S. Das, G. Dimopoulos, *BMC Physiol.* **8**, 23 (2008).
29. A. D. Sheppard *et al*., *Parasit. Vectors* **10**, 255 (2017).
30. A. C. Honnen, J. L. Kypke, F. Holker, M. T. Monaghan, *Sustainability* **11**, 6220 (2019).
31. S. S. Rund, J. E. Gentile, G. E. Duffield, *BMC Genomics* **14**, 218 (2013).
32. S. P. Sawadogo *et al*., *Parasit. Vectors* **6**, 275 (2013).
33. S. S. Rund, S. J. Lee, B. R. Bush, G. E. Duffield, *J. Insect Physiol.* **58**, 1609–1619 (2012).
34. A. Martin Anduaga *et al*., *eLife* **8**, e44642 (2019).
35. J. Majercak, D. Sidote, P. E. Hardin, I. Edery, *Neuron* **24**, 219–230 (1999).
36. C. Wicker-Thomas, *J. Insect Physiol.* **53**, 1089–1100 (2007).
37. L. J. Cator, B. J. Arthur, L. C. Harrington, R. R. Hoy, *Science* **323**, 1077–1079 (2009).
38. M. P. Su, M. Andrés, N. Boyd-Gibbins, J. Somers, J. T. Albert, *Nat. Commun.* **9**, 3911 (2018).
39. M. Gendrin, *Cell Host Microbe* **22**, 577–579 (2017).
40. X. Sun *et al*., *J. Chem. Ecol.* **43**, 557–562 (2017).
41. C. Kent, R. Azanchi, B. Smith, A. Chu, J. Levine, *PLOS ONE* **2**, e962 (2007).

ACKNOWLEDGMENTS

We thank S. Zhan and Y. Lai for heatmap analysis, W. Hu for gas chromatography–mass spectrometry analysis, and H. Liu and Y. Zhang for discussions. We thank N. C. Manoukis for comments on the manuscript. **Funding:** This work was supported by grants from the National Natural Science Foundation of China (grants 32021001, 31830086, and 31772534), the National Key R&D Program of China (grants 2020YFC1200100, 2018YFA0900502, 2017YFD0200400, and 2017YFD0201202), the Strategic Priority Research Program of Chinese Academy of Sciences (grant XDB11010500), NIH grant R01AI031478, the NIH Distinguished Scholars Program and the Division of Intramural Research AI001250-01, NIAID, and the Bloomberg Philanthropies. **Author contributions:** S.W. and M.J.-L. conceived the project; S.W. designed the study and performed microarray analysis; G.W. performed most of the experiments; A.D. and H.M. collected swarming mosquitoes; J.V.-R., A.D., C.N., C.O.O., A.M.B., and P.S.S. conducted semifield experiments; J.L. and G.W. designed and conducted locomotor activity assays; G.W., C.C., and L.D. conducted RNA interference (RNAi) and mating assays in *An. stephensi*; G.W. performed *desat1* and CHCs assays; F.L. reared mosquitoes; J.V.-R., T.L.A.S., and T.V.P. conducted RNAi and mating assays in *An. gambiae s.s.*; S.W. wrote the original draft of the manuscript; S.W., J.V.-R., and M.J.-L. reviewed and edited the manuscript; and S.W. is the lead contact. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and code to understand and assess the conclusions of this research are available in the main text, supplementary materials, and via the Gene Expression Omnibus (GEO) database under accession no. GSE150971.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/411/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S3
References (*42*–*46*)
MDAR Reproducibility Checklist
Movies S1 to S3

20 June 2020; accepted 18 November 2020
10.1126/science.abd4359

METABOLIC EVOLUTION

Ancient balancing selection maintains incompatible versions of the galactose pathway in yeast

James Boocock[1,2,3,4], Meru J. Sadhu[1,2,3,4]*, Arun Durvasula[1], Joshua S. Bloom[1,2,3,4]†, Leonid Kruglyak[1,2,3,4]†

Metabolic pathways differ across species but are expected to be similar within a species. We discovered two functional, incompatible versions of the galactose pathway in *Saccharomyces cerevisiae*. We identified a three-locus genetic interaction for growth in galactose, and used precisely engineered alleles to show that it arises from variation in the galactose utilization genes *GAL2*, *GAL1/10/7*, and phosphoglucomutase (*PGM1*), and that the reference allele of *PGM1* is incompatible with the alternative alleles of the other genes. Multiloci balancing selection has maintained the two incompatible versions of the pathway for millions of years. Strains with alternative alleles are found primarily in galactose-rich dairy environments, and they grow faster in galactose but slower in glucose, revealing a trade-off on which balancing selection may have acted.

Variation in nutrient availability between environments has led to the evolution of diverse metabolic pathways. In humans, mutations in these pathways give rise to diseases known as inborn errors of metabolism (*1*). The budding yeast *Saccharomyces cerevisiae* is commonly used for studying eukaryotic metabolism (*2*). A classic well-studied pathway for galactose metabolism includes a galactose transporter, encoded by the gene *GAL2*, and the enzymes encoded by *GAL1*, *GAL10*, and *GAL7*, which convert galactose to glucose-1-phosphate (*3*). Phosphoglucomutase, encoded by *PGM1* and *PGM2*, then converts glucose-1-phosphate to glucose-6-phosphate—the substrate for glycolysis.

Within the same genus, some strains of *Saccharomyces kudriavzevii* can metabolize galactose, whereas others have lost this ability through pseudogenization of multiple genes in the pathway, and it was proposed that the two versions of the pathway have been maintained through multiloci balancing selection (*4*). Balancing selection maintains genetic diversity against the forces of genetic drift and

[1]Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA, USA. [2]Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, CA, USA. [3]Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, CA, USA. [4]Institute for Quantitative and Computational Biology, University of California, Los Angeles, Los Angeles, CA, USA.
*Present address: Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA.
†Corresponding author. Email: jbloom@mednet.ucla.edu (J.S.B.); lkruglyak@mednet.ucla.edu (L.K.)

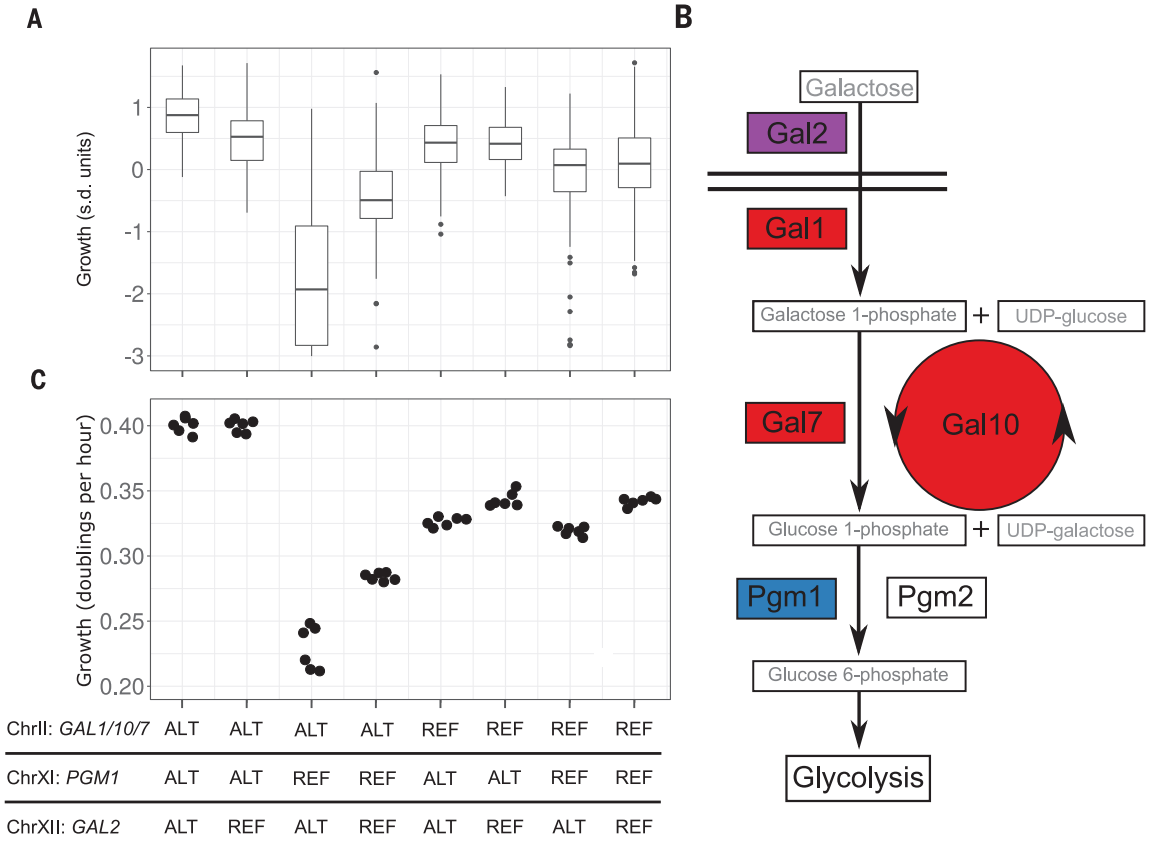


Fig. 1. Three-locus genetic interaction for growth in galactose. (**A**) Boxplots show growth in galactose of yeast segregants (n = 867) derived from a cross between CBS2888 and YJM981. Each boxplot corresponds to segregants with one of eight distinct combinations of alleles at the three loci (ChrII, *GAL1/10/7*; ChrXI, *PGM1*; and ChrXII, *GAL2*). (**B**) Galactose metabolic pathway. Components of the pathway corresponding to the three loci are shown in different colors. UDP, uridine 5′-diphosphate. (**C**) Growth of allele replacement strains in galactose. BY alleles [reference (REF)] and CBS2888 alleles [alternative (ALT)].

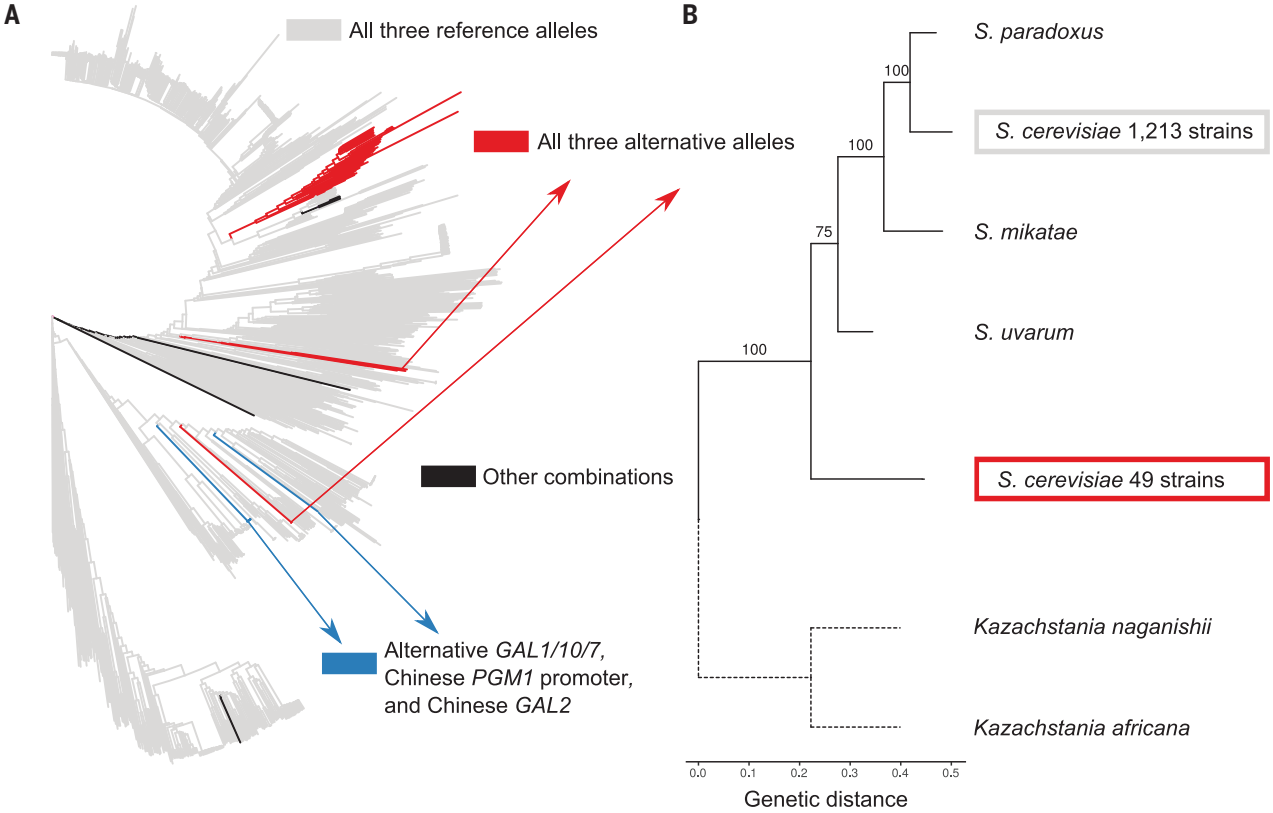


Fig. 2. The alternative galactose alleles are broadly distributed and fall outside the *Saccharomyces* genus. (**A**) Genome-wide neighbor-joining tree of 1276 sequenced yeast isolates. (**B**) Bootstrapped maximum likelihood phylogenetic tree of the *GAL1/10/7* alleles from CBS2888 (alternative), BY (reference), other species in the *Saccharomyces* genus, and two outgroup species. The outgroup branches (dotted lines) were rescaled to the average branch length.

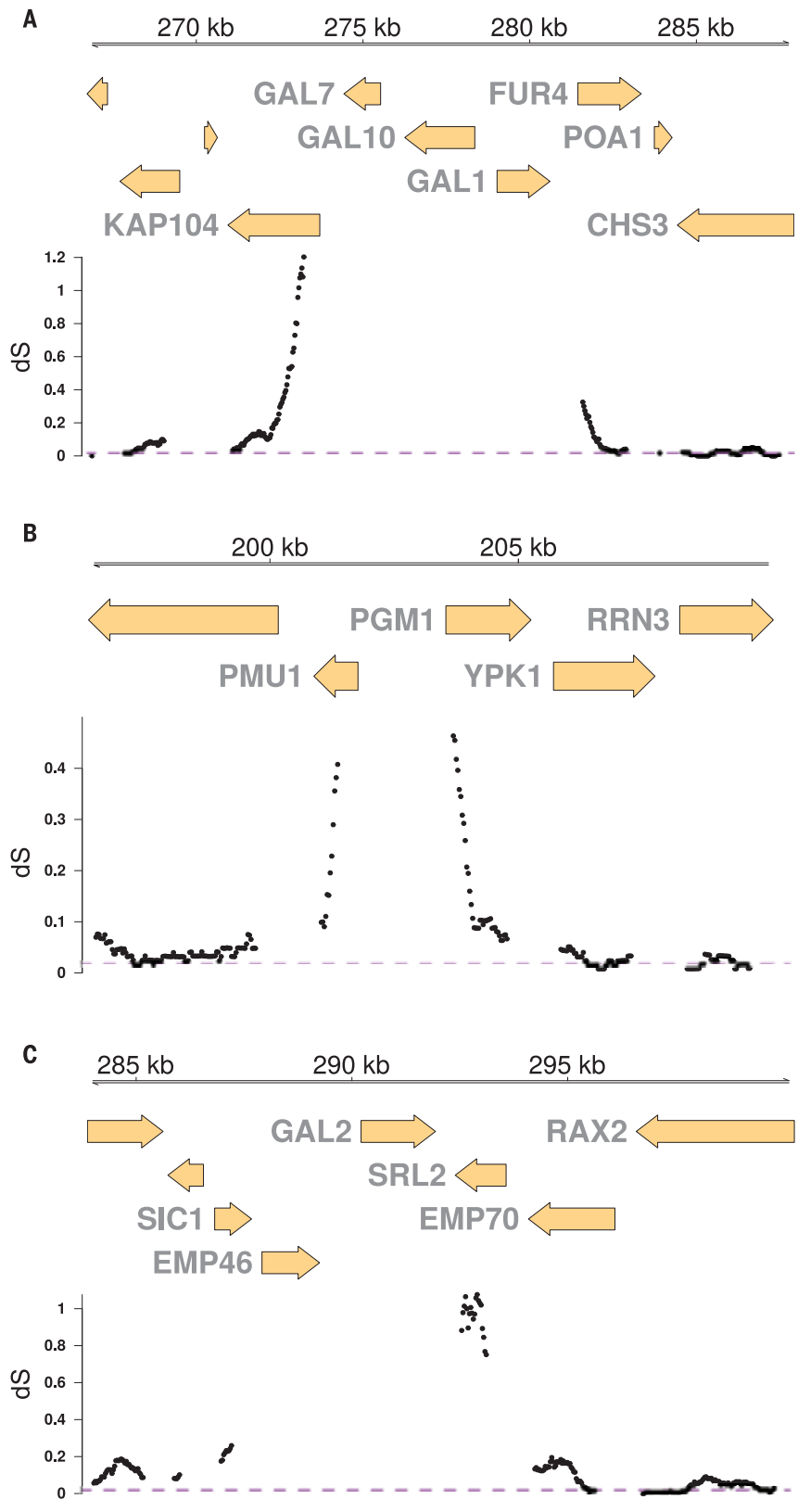


Fig. 3. A signature of ancient balancing selection. Estimated rate of synonymous substitutions per site (dS) in 200-codon windows stepped every 10 base pairs between CBS2888 (alternative) and BY (reference) genes surrounding the galactose loci. (**A** to **C**) Genes adjacent to *GAL1/10/7* (A), genes adjacent to the *PGM1* promoter (B), and genes adjacent to *GAL2* (C). The purple dashed line shows the genome-wide average dS of 0.014. The moderately elevated dS in some genes (*SIC1* and *KAP104*) provides evidence of ongoing balancing selection (*4*).

has typically been observed to act on single loci (*5*). Multiloci balancing selection is expected to be extremely rare because it has to overcome the independent segregation of alleles at the different loci.

We studied growth in galactose in a large set of crosses in *S. cerevisiae* (*6*) and observed a genetic interaction among three loci in crosses involving the soil strain CBS2888 (three-way effect size 0.19 SD units, chi-square test, $P < 10^{-15}$) (Fig. 1A, figs. S1 and S2, and tables S1 to S3). The nonadditive nature of the effects of the three loci is best illustrated by the phenotype of segregants that inherit the CBS2888 allele at the loci on chromosome II (ChrII) and ChrXII and the non-CBS2888 allele at the locus on ChrXI; these segregants grow much slower in galactose than those with any other combination of alleles (Fig. 1A). The three loci contain genes that encode components of galactose metabolism: *GAL1*, *GAL10*, and *GAL7* on ChrII; *PGM1* on ChrXI; and *GAL2* on ChrXII (Fig. 1B and fig. S1) (*3*). The CBS2888 alleles of these genes were highly diverged from the reference (fig. S3 and table S4). We hereafter refer to the divergent galactose alleles found in CBS2888 as the alternative alleles and alleles observed in the other strains as the reference alleles.

We used CRISPR-Cas9 to engineer strains with all eight possible combinations of the three alternative and three reference galactose alleles in a common genetic background (fig. S4 and tables S5 to S7) (*7*, *8*). We then measured the growth rates of the eight engineered strains in galactose and recapitulated the mapping results (Fig. 1C, fig. S5, and tables S1 and S2), demonstrating that variants in the coding and intergenic regions of *GAL1/10/7* and *GAL2* and in the promoter region of *PGM1* are responsible for the observed genetic interaction. In particular, the strain with the reference *PGM1* promoter allele and the alternative *GAL1/10/7* and *GAL2* alleles exhibited a severe growth defect in galactose, confirming that the components of the reference and alternative pathways are incompatible.

To better understand cis-acting regulatory differences between the alternative and reference galactose alleles (*9*), we grew a diploid hybrid strain (CBS2888xBY) in glucose, transferred it to galactose medium, and sequenced RNA from samples collected throughout a growth time course (*7*). In glucose, the expression of the CBS2888 allele of *PGM1* was slightly lower than that of the reference allele (fig. S6). By contrast, 1 hour after the switch to galactose, the expression of the CBS2888 allele of *PGM1* was 15.5 times greater than that of the reference allele (binomial test, $P < 10^{-100}$), and this difference persisted for the rest of the time course (fig. S6 and table S8).

The alternative *PGM1* promoter allele contains a *GAL4* upstream activating sequence (UAS) (*10*), whereas the reference allele does

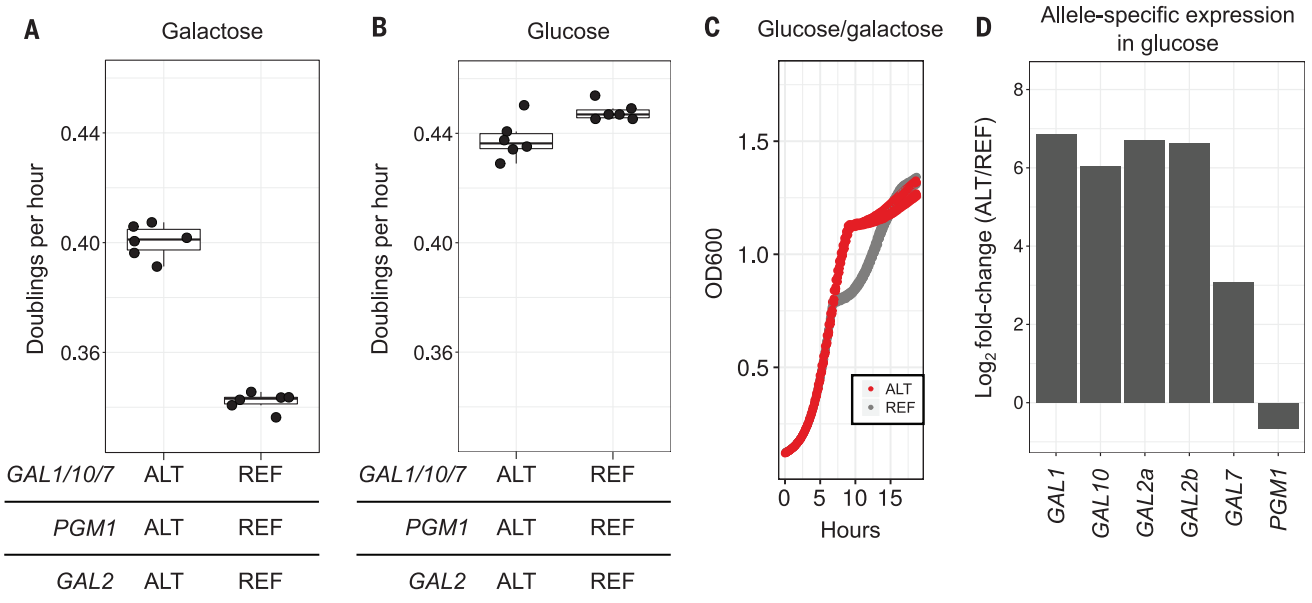


Fig. 4. Trade-offs between the alternative and reference alleles in the galactose pathway. (**A**) Growth rates of allele replacement strains (n = 6) with all three reference (right) or all three alternative (left) alleles in galactose as a sole carbon source. (**B**) As in (A), but for cells grown in glucose. (**C**) Allele-specific expression of the galactose genes from a diploid hybrid (CBS2888xBY) strain grown in glucose. OD600, optical density at 600 nm. (**D**) Growth curves of allele replacement strains with all three alternative (red) or reference (gray) galactose alleles in mixed glucose and galactose medium.

not (fig. S4). We engineered a point mutation disrupting the UAS in a strain with all three alternative galactose alleles (*7*). This single mutation recapitulated the growth defect in galactose observed in a strain with a combination of the reference allele of the *PGM1* promoter and alternative alleles of the other *GAL* genes (fig. S7). We conclude that the induction of *PGM1* in galactose, mediated through a *GAL4* UAS, is critical for the proper functioning of the alternative galactose pathway.

We searched for the alternative and reference galactose alleles in worldwide collections of sequenced *S. cerevisiae* isolates comprising 1276 strains (*11*, *12*) and found three common combinations: only reference alleles (1213 strains), only alternative alleles (49 strains), and 8 strains from China with the alternative *GAL1/10/7* allele and alleles of *GAL2* and the *PGM1* promoter that differ from both the reference and the alternative alleles (Fig. 2A, fig. S8, and table S9) (*7*). No strains carried the reference *PGM1* promoter allele and the alternative *GAL1/10/7* and *GAL2* alleles, suggesting that this combination causes a fitness disadvantage in natural environments and has been purged by selection. This hypothesis is further supported by a high linkage disequilibrium index (ε = 0.59) for the three loci (fig. S9) (*13*).

The alternative galactose alleles are fixed in two lineages of strains found in dairy products, including Camembert cheese from France, kefir grains from Japan, and fermented yak and goat milk from China (table S9). These environments are rich in lactose, a disaccharide of glucose and galactose. *S. cerevisiae* relies on the activity of other fungi and bacteria to break down lactose into glucose and galactose, which it then metabolizes (*14*). This observation suggests that the alternatives alleles are maintained by natural selection in dairy environments.

We dated the split between the alternative and reference galactose alleles to ~3.2 billion generations ago (95% confidence interval = 2.5 to 4.5 billion generations), which predates the most recent common ancestor of the *Saccharomyces* genus (figs. S10 and S11 and tables S4 and S10) (*7*, *15*). Phylogenetic clustering placed the alternative galactose alleles outside the *Saccharomyces* genus and supports an ancient origin of the alternative alleles (Fig. 2B and figs. S12 and S13) (*7*, *16*).

One force that can maintain highly diverged alleles within a species is balancing selection. This process is expected to generate a signature of elevated sequence divergence at linked neutral sites that decays with genetic distance from the selected variant (*5*). We examined the rate of synonymous substitutions per site (dS) across the CBS2888 genome relative to the reference and observed a strong signature of ancient balancing selection at all three galactose loci (Fig. 3 and figs. S14 to S18) (*7*).

The strains with the alternative or Chinese alleles contain *GAL2* genes duplicated in tandem, and *GAL2* is also duplicated in two other yeast species: *Saccharomyces uvarum* and *Saccharomyces eubayanus*. We aligned all the *GAL2* paralogs and observed that the N-terminal cytosolic regions (amino acids 1 to 67) were highly dissimilar within species and phylogenetically clustered across species (fig. S19). These results suggest that the N-terminal regions of the *GAL2* paralogs are functionally distinct and maintained by selection, and they also provide evidence that the alternative alleles have an ancient origin in *Saccharomyces* (fig. S20).

It has been proposed that the alternative galactose alleles arose through introgression around the time humans domesticated milk-producing animals, but no species that could have donated the alleles has been identified (*7*, *17*). A relatively recent introgression would generate a sharp boundary between dS at the *GAL* genes and the rest of the genome. Instead, our data suggest that the variation at these loci has accumulated within *S. cerevisiae* over time.

We performed forward genetic simulations to distinguish between scenarios that could have given rise to the observed signatures of balancing selection (figs. S21 to S23) (*7*). Models of recent introgression (<50 million generations ago), with or without balancing selection, were not well-supported when compared with a model of ancient balancing selection (figs. S24 to S27 and table S11) (*7*).

Balancing selection can act on fitness tradeoffs, in which alleles with higher fitness in one environment have lower fitness in another (*5*). Although all strains grow faster in glucose than in galactose [t statistic (T) = 7.80, t test, $P < 10^{-5}$],

the strains with the alternative alleles grow faster in galactose than the strains with the reference alleles (Figs. 1C and 4A). *S. cerevisiae* encounters and metabolizes a wide variety of sugars (*18*) but prefers glucose (*19*). In glucose, the strains with the reference alleles grow 2% faster than strains with the alternative alleles ($T = -3.12$, t test, $P = 0.017$) (Fig. 4B and fig. S28). This faster growth provides an explanation for the maintenance of the reference alleles in the strains that do not frequently encounter galactose.

In the strains with reference alleles, the *GAL* genes are robustly repressed by glucose and induced by galactose (*3*). This leads to a pause in growth known as the diauxic shift, when yeast switch from metabolizing glucose to metabolizing galactose. Strains with the three alternative galactose alleles do not undergo a diauxic shift (Fig. 4C and fig. S29). RNA sequencing showed that in glucose, the reference alleles are repressed, whereas the alternative *GAL* alleles are constitutively expressed (fold change = 40.6, binomial test, $P < 10^{-16}$) (Fig. 4D, fig. S30, and table S8) (*7*). The constitutive expression of the *GAL* genes eliminates the diauxic shift (*20*), providing a fitness benefit when galactose is encountered. However, gene expression can be costly (*21*), and this could explain why the alternative galactose pathway leads to a growth disadvantage in glucose.

The incompatible allele combinations we identified may provide a model for classical galactosemia, an inborn error of metabolism caused by recessive mutations in *GALT*, the human homolog of *GAL7* (*22*), that can lead to life-threatening symptoms if galactose is not eliminated from diet. The precise molecular mechanisms of galactosemia are not well understood (*23*), but yeast models of galactose toxicity suggest that the incompatibility observed in this work arises from the same metabolic defect that underlies galactosemia. Finally, our results go beyond previous findings (*4*) in showing that balancing selection can preserve two alternate, functionally distinct states of a multiloci genetic network, providing a general mechanism for the maintenance of complex, interacting genetic variation at coadapted alleles.

REFERENCES AND NOTES

1. J. M. Saudubray, À. Garcia-Cazorla, *Pediatr. Clin. North Am.* **65**, 179–208 (2018).
2. J. Nielsen, *Annu. Rev. Biochem.* **86**, 245–275 (2017).
3. C. A. Sellick, R. N. Campbell, R. J. Reece, *Int. Rev. Cell Mol. Biol.* **269**, 111–150 (2008).
4. C. T. Hittinger *et al.*, *Nature* **464**, 54–58 (2010).
5. D. Charlesworth, *PLOS Genet.* **2**, e64 (2006).
6. J. S. Bloom *et al.*, *eLife* **8**, e49212 (2019).
7. Materials and methods and supplementary text are available as supplementary materials.
8. M. J. Sadhu *et al.*, *Nat. Genet.* **50**, 510–514 (2018).
9. F. W. Albert, L. Kruglyak, *Nat. Rev. Genet.* **16**, 197–212 (2015).
10. A. Traven, B. Jelicic, M. Sopta, *EMBO Rep.* **7**, 496–499 (2006).
11. S.-F. Duan *et al.*, *Nat. Commun.* **9**, 2690 (2018).
12. J. Peter *et al.*, *Nature* **556**, 339–344 (2018).
13. Y. Okada, *Hum. Genome Var.* **5**, 29 (2018).
14. J. L. Legras *et al.*, *Mol. Biol. Evol.* **35**, 1712–1727 (2018).
15. M. Kellis, N. Patterson, M. Endrizzi, B. Birren, E. S. Lander, *Nature* **423**, 241–254 (2003).
16. D. R. Scannell *et al.*, *Genes Genomes Genet.* **1**, 11–25 (2011).
17. S. F. Duan *et al.*, *Curr. Biol.* **29**, 1126–1136.e5 (2019).
18. B. Turcotte, X. B. Liang, F. Robert, N. Soontorngun, *FEMS Yeast Res.* **10**, 2–13 (2010).
19. J. H. Kim, A. Roy, D. Jouandot II, K. H. Cho, *Biochim. Biophys. Acta* **1830**, 5204–5210 (2013).
20. J. I. Roop, K. C. Chang, R. B. Brem, *Nature* **530**, 336–339 (2016).
21. G. I. Lang, A. W. Murray, D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5755–5760 (2009).
22. A. I. Coelho, M. E. Rubio-Gozalbo, J. B. Vicente, I. Rivera, *J. Inherit. Metab. Dis.* **40**, 325–342 (2017).
23. K. Lai, L. J. Elsas, K. J. Wierenga, . *IUBMB Life* **61**, 1063–1074 (2009).
24. J. Boocock, M. J. Sadhu, A. Durvasula, J. S. Bloom, L. Kruglyak, Ancient balancing selection maintains incompatible versions of the galactose pathway in yeast (Figure creation), Dryad (2020); https://doi.org/10.5068/D14370.
25. J. Boocock, theboocock/ancient_bal_scripts: submission, Zenodo (2020); https://doi.org/10.5281/zenodo.4132713.
26. J. Boocock, theboocock/popgen_utilities: submission, Zenodo (2020); https://doi.org/10.5281/zenodo.4131787.
27. J. Boocock, theboocock/gal_bal: submission, Zenodo (2020); https://doi.org/10.5281/zenodo.4107954.

ACKNOWLEDGMENTS

We thank O. Schubert, E. Ben-David, L. Guo, and S. Zdraljevic for helpful manuscript feedback and edits. **Funding:** This work was supported by funding from the Howard Hughes Medical Institute and an NIH grant (2R01GM102308-06 to L.K.). A.D. is supported by the NSF Graduate Research Fellowship (DGE-1650604). **Author contributions:** J.B. performed experiments with assistance from M.J.S. and J.S.B.; J.B. analyzed data with assistance from J.S.B. and A.D.; L.K. and J.S.B. supervised the project. J.B., A.D., J.S.B., and L.K. wrote the manuscript. All authors discussed and agreed on the final version of the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Sequencing data are available under NCBI BioProject PRJNA575066. The processed datasets are available on Dryad (*24*). The following code is available on Zenodo: custom scripts used to analyze processed data and create figures (*25*), a custom Python package used with our population genetic analysis (*26*), and the code used to perform and analyze the forward genetic simulations (*27*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6527/415/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S30
References (*28*–*58*)
Tables S1 to S11
MDAR Reproducibility Checklist

1 November 2019; resubmitted 8 May 2020
Accepted 17 December 2020
10.1126/science.aba0542

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

spj.sciencemag.org/bmef



BME Frontiers

OPEN ACCESS

Biomedical Engineering (BME) Frontiers is a **Science Partner Journal** distributed by the **American Association for the Advancement of Science** (AAAS) in collaboration with the **Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences** (SIBET CAS). *BME Frontiers* aims to serve as an effective platform for the multidisciplinary community of biomedical engineering. The journal will publish breakthrough research in the fields of pathogenic mechanisms as well as disease prevention, diagnosis, treatment, and assessment.

The Science Partner Journals (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, open access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS expands its efforts to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility and publishing expertise that AAAS is uniquely positioned to **offer as the world's largest general science membership society**.

Submit your research to *Biomedical Engineering Frontiers* today!
Learn more at: **spj.sciencemag.org/bmef**

ARTICLE PROCESSING CHARGES WAIVED UNTIL 2021



UNDERSTANDING DYNAMIC ECOSYSTEMS

AAAS
ANNUAL MEETING

Climate Change | Sustainability | Artificial Intelligence | International S&T | Resource Management
Public Engagement | Technology | Diversity | Public Health | Infectious Disease | Innovation
Ecology | Workforce Developtment | Big Data | Anthropogenic Processes | Microbiome

The 2021 Annual Meeting will convene entirely online, February 8-11, with pre-recorded videos available starting on January 18. Join us!

Meeting registration and program now available online.

For more information, please visit:

aaas.org/meetings | #AAASmtg

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE



Dear Science Enthusiast:

It is difficult to believe how much our world has changed since our 2020 AAAS Annual Meeting in February. The theme of the 2021 Annual Meeting, Understanding Dynamic Ecosystems, took on a new and unexpected meaning, especially in light of how vital advances in science and technology come together to meet and to overcome the challenges of today's rapidly changing world.

What 2020 has demonstrated is that coming together virtually to highlight and discuss advances and collaborations across the STEM ecosystem is needed now more than ever—so much so that we are radically reducing the registration fee for our scientific program, thanks in part to the generous support of our sponsors. We continue to offer workshops, e-posters, and other complimentary programming.

Just as the rest of the world has adapted to a new normal, so too has your Annual Meeting team. Our new virtual format has transformed many aspects of the meeting, and we extend our gratitude to our attendees, speakers, sponsors, and volunteers for their patience and flexibility. This February, we hope you will consider sharing your voice with our diverse community of scientists, educators, policymakers, journalists, and many others.

Sincerely,
Claire M. Fraser, Ph.D.
President, AAAS
Chair, 2021 Annual Meeting
Dean's Endowed Professor and Director, Institute for Genome Sciences
University of Maryland School of Medicine

Starting January 18

Look out for the pre-release of meeting content:

- Hundreds of Scientific Session videos highlighting the work of individual panelists will be available to stream.
- E-poster galleries will showcase research, including highlights from early career researchers.

Live February 8–11

- Live updates throughout each day. (See back page)
- Invited Topical Lecturers will be available for Q&A.
- Scientific Sessions + Panels will convene to discuss how each presenter's work interacts with others' and for Q&A.
- Public Workshops will focus on:
  - Advancing your career;
  - Advocating and building alliances; and
  - Diversifying and enhancing equity in the scientific enterprise.
- And, of course, leading scientists will address attendees and the public during Plenary Lectures.

CONNECT WITH US!
#AAASmtg

aaas.org/meetings

@AAASmeetings

/AAAS.Science

@aaasorg

/aaasorg

Thanks to the generous support of our sponsors, access to the scientific program for AAAS members is available at a deeply discounted rate of $50.

Not a AAAS Member?
Now's the time to join. We've made joining easier than ever with new membership and registration packages starting at $75.* Simply choose the membership option for you and complete your meeting registration.

*Offer valid from November 6, 2020 to February 11, 2021, for new individual members only. There is a limit of one AAAS "Future of Science" T-shirt per membership order. Members with an international shipping address and student members who select a digital subscription are not eligible to receive the T-shirt. Please allow up to ten weeks for delivery of the T-shirt. The AAAS "Future of Science" T-shirt is provided as is, without any guarantees or warranty, and cannot be exchanged or returned.

In association with the product, AAAS makes no warranties of any kind, either express or implied, including but not limited to warranties of merchantability or fitness for a particular purpose. AAAS members are entitled to receive issues published only during their 12-month membership term.

Plenary Lectures



Claire M. Fraser
President, AAAS
University of Maryland School of Medicine

Presented by: Arizona State University
Host University



Ruha Benjamin
Princeton University



Anthony S. Fauci
U.S. National Institute of Allergy and Infectious Diseases



Mary L. Gray
Microsoft Research



Sethuraman Panchathan
U.S. National Science Foundation

Topical Lectures



Anne H. Charity Hudley
University of California,
Santa Barbara



Joseph M. DeSimone
Carbon, Inc. and
Stanford University



Ayanna Howard
Georgia Institute of Technology



Yalidy Matos
Rutgers University



Nalini M. Nadkarni
University of Utah

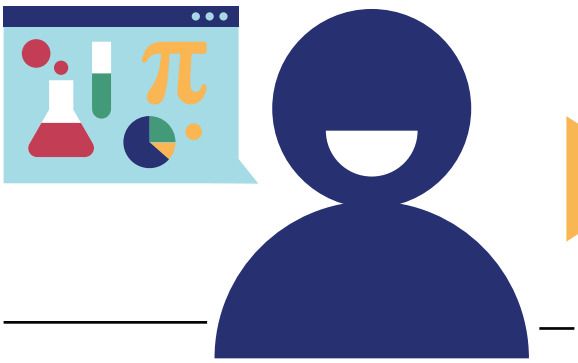


Yohannes Haile-Selassie
Cleveland Museum of Natural History



Patrick O. Brown
Impossible Foods

As of January 4, 2021



AAAS Live!

February 8-11

TUNE IN to catch quick updates, social media news, and important discussions throughout the day.

To see the schedule and enter the conference site, visit:
aaas.org/meetings

EVENT SCHEDULE

VIDEO LIBRARY

THE EXCHANGE

AAAS E-POSTERS

SOCIALLY DISTANCED MINGLE

SCI-MIC PODCASTS

PERSONALIZED BACKGROUND & BINGO

JOIN OUR GROUP PICTURE

SWAG STORE

CONNECT WITH US!
#AAASmtg

@AAASmeetings

/AAAS.Science

@aaasorg

/aaasorg



Diagnostic tests for COVID-19 help stave off the pandemic and buy time until more vaccines reach the population.

Pandemic pivot: How scientists answered the call for diagnostic tests

In medicine as well as in sports, the best defense is a good offense. When it comes to the current COVID-19 pandemic, the offense entails large-scale diagnostic testing that can rapidly, accurately, and reliably detect SARS-CoV-2. Although vaccines are currently being distributed, diagnostic tests–many of which are produced by companies that have already used similar technology to combat other diseases–remain key to containing the pandemic and preventing more deaths. **By Alan Dove**

At the turn of the 21st century, coronaviruses were among the most obscure subjects in virology–it's hard to draw much public interest in pathogens that cause little more than minor sniffles. Struggling for funding and toiling in obscurity, a handful of scientists nonetheless produced a substantial body of basic research on these RNA viruses. Then, in 2002, clusters of patients suddenly developed a novel form of severe acute respiratory syndrome (SARS), a life-threatening condition subsequently linked to a new coronavirus that had leapt from animals into humans. Several thousand cases and hundreds of deaths later, public health officials had contained the spread of SARS, but the field of coronavirology had been transformed forever.

After the SARS crisis, wildlife surveys and stepped-up surveillance revealed numerous coronaviruses in nature that could, in theory, spill into humans under the right circumstances. By 2019, those findings and the appearance of Middle East Respiratory Syndrome (MERS), another serious respiratory disease caused by a zoonotic coronavirus infection, made it clear that coronaviruses posed a serious threat to global health. Then SARS-CoV-2 emerged in 2020, and the global COVID-19 pandemic began. As the scale of the crisis became clear, scientists and pharmaceutical companies around the globe rallied to address it, often repurposing technologies developed to combat other diseases. 2020 was the year everyone became a coronavirologist.

IMAGE: © OSORIOARTIST\SHUTTERSTOCK.COM

Turning on the LAMP

For vaccine developers, the path ahead was as obvious as it was arduous. Even before the World Health Organization's declaration of a global pandemic on March 11, hundreds of academic and corporate scientists were already turning their existing tools and techniques toward developing SARS-CoV-2 vaccines. As that effort ground ahead, though, public health officials had a desperate need for diagnostic tests. Because many SARS-CoV-2 infections are asymptomatic or mild, tracking the virus's spread required fast, reliable testing, and existing resources weren't up to the task. Fortunately, reinforcements were on the way.

"Before the crisis, we didn't work at all on coronaviruses," says Franck Molina, director of the **Sys2Diag** program at CNRS in Montpellier, France. Molina's team was nevertheless in an ideal position to develop a new generation of SARS-CoV-2 tests. The lab's academic–corporate collaboration gives the team a multidisciplinary edge, encompassing everything from laboratory researchers to artificial intelligence and bioinformatics engineers. Having previously developed liquid biopsy tests for customized cancer therapies, Sys2Diag pivoted to measuring viral RNA instead of tumor biomarkers.

When the team's researchers began developing their new SARS-CoV-2 diagnostic system, established medical labs were already using sensitive, specific PCR tests. The key drawback to those tests is that they require sophisticated labs, trained personnel, and expensive thermal cyclers, all of which were in short supply. Molina and his colleagues focused on creating a much simpler, cheaper test using loop-mediated isothermal amplification (LAMP). "The machine [for LAMP] is just a heating machine, and we knew the properties of this technology," says Molina.

Besides using simpler equipment for DNA amplification, the investigators also streamlined everything from sample collection to results cont.>

reporting as much as possible. The final iteration of the system, which received marketing approval in the United States and European Union in June 2020, yields robust results only 40 minutes after the patient spits into a tube. At the end of the process, a smartphone application reads the colorimetric result. "Then immediately you have the formatted readout of the test, positive or negative, and it automatically sends the information to a database," says Molina. Minimally trained technicians can conduct the tests almost anywhere.

SkillCell, a company spun off from Sys2Diag, now sells this testing system as EasyCOV. Molina says it's in use worldwide in airports, hospitals, clinics, and various industries that need to test their employees. In those settings, it competes with other rapid viral RNA testing products, such as **Abbott's** ID NOW COVID-19 system, which also uses isothermal amplification.

Immunity counts

While many biologists and life science companies changed focus to help fight the new virus, others simply continued their regular work as the pandemic came to them. That's what happened at PerkinElmer, one of the world's largest suppliers of laboratory and diagnostic reagents and equipment. "We have several thousand products in our catalog, and when the pandemic happened, this was immediately on our radar," says Iswariya Venkataraman, scientific affairs manager at **EUROIMMUN US**, a PerkinElmer subsidiary.

When the scope of the SARS-CoV-2 outbreak in China became clear in December 2019, EUROIMMUN began developing PCR tests to detect the virus, as well as an enzyme-linked immunosorbent assay (ELISA) for antibodies against it. "The priority for us was in [being] able to first diagnose infection, and then also understand the humoral immune response following this infection," says Venkataraman. While the immediate need to identify and quarantine infected individuals claimed much of the attention early in the pandemic, public health officials quickly realized the importance of antibody screening as well, to identify those who'd had the infection already.

By early May 2020, EUROIMMUN had gained approval for its rapid qualitative test to detect antibodies against the SARS-CoV-2 spike glycoprotein. The test uses a purified preparation of the S1 domain of the spike, generally considered the most specific antigenic target on the virus, to bind immunoglobulin G (IgG) antibodies in a standard ELISA procedure familiar to any medical testing lab. IgG arises late in an infection and is considered a good long-term marker of antibody-based humoral immunity. "For IgG antibodies, the median onset is about day 14 or 15 [after infection]," says Venkataraman.

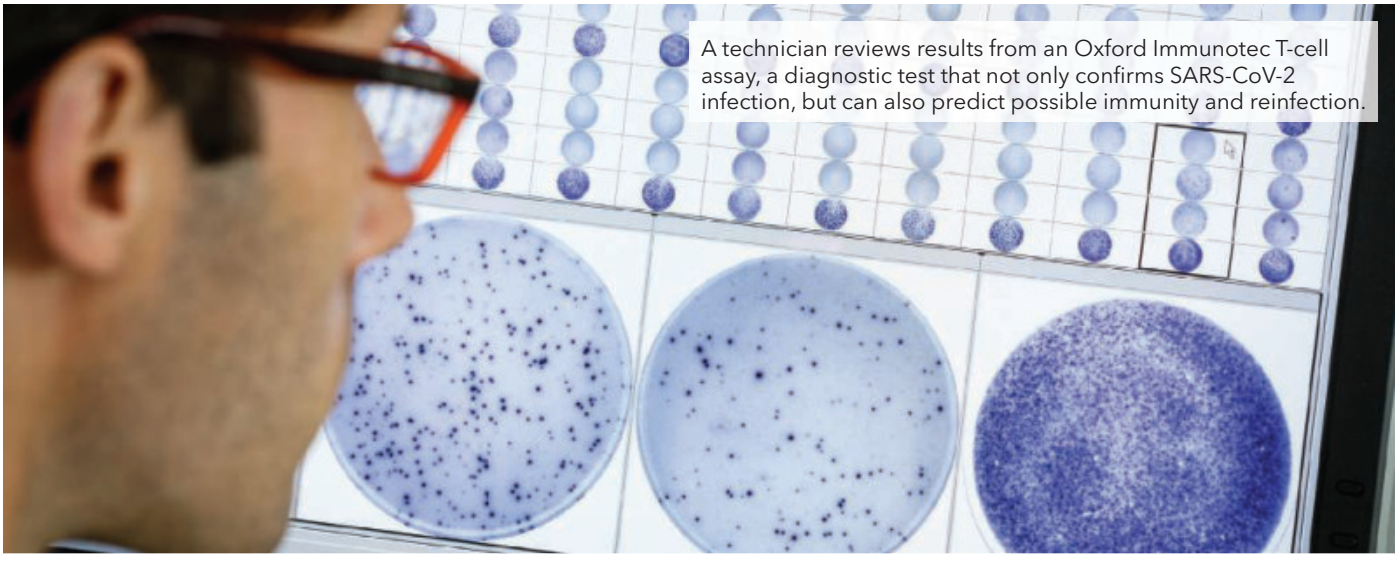
As the pandemic continued, qualitative testing revealed wide variation in individual immune responses against the virus, with some people testing negative for antibodies within weeks of infection, while others remained positive for months. At the same time, vaccine developers were gearing up for clinical trials and needed to know not only that their products elicited antibody responses, but how strong those responses were. EUROIMMUN answered this need by calibrating their now-established qualitative test. "It's exactly the same antigen target, just with the addition of calibrators to make it quantitative," says Venkataraman. The company released the new test, called QuantiVac, in December 2020.

QuantiVac is now marketed as a quantitative test in Europe and a semiquantitative test in the United States. The distinction is due to the U.S. Food and Drug Administration (FDA) requirement that quantitative tests be based on established international reference standards, which don't yet exist for SARS-CoV-2 antibodies. "When there is an international standard available, we will calibrate our assays to that," says Venkataraman.

The QuantiVac kit includes a set of six reference samples representing a range of antibody concentrations. "With that, you're able to plot a standard curve, and once you have a standard curve in place, you're able to determine the unknown concentration in your sample," says Venkataraman, adding that vaccine developers and immunologists should find the quantitative data useful in detailing patients' immune responses.

Antibody can see the difference

Other diagnostics-focused companies have followed a similar path, with similar timelines. "It was around February [2020] when it became clear that the pandemic was evolving, and then we had discussions cont.>



A technician reviews results from an Oxford Immunotec T-cell assay, a diagnostic test that not only confirms SARS-CoV-2 infection, but can also predict possible immunity and reinfection.

IMAGE: COURTESY OF OXFORD IMMUNOTEC

Featured participants

Abbott
www.diagnostics.abbott/us/en/home

EUROIMMUN US
www.euroimmun.com

Oxford Immunotec
www.oxfordimmunotec.com/north-america

Roche Diagnostics
diagnostics.roche.com

SkillCell
www.skillcell-alcen.com

Sys2Diag
www.alcediag-alcen.com/alcediag/sys2diag

on what kind of role an antibody assay may play in the management of the pandemic," says Beatus Ofenloch-Haehnle, head of immunoassay development and system integration at **Roche Diagnostics** in Penzberg, Germany. By the second week of March, he and his colleagues were working on adapting established Roche assays to detect anti-SARS-CoV-2 antibodies. The resulting qualitative assay for pandemic virus exposure launched in Europe and the United States 6 weeks later.

Unlike the EUROIMMUN assay, Roche's qualitative test identifies antibody responses against the SARS-CoV-2 nucleocapsid protein instead of the spike glycoprotein. Rather than a standard ELISA, the test uses Roche's proprietary Elecsys electrochemiluminescence immunoassay. In this technique, two different antigen preparations, each bound to a different conjugate, bind to the antibodies in the sample and attach them to magnetic beads and luminescent markers. A magnet then precipitates the antibody-bound beads, and an applied voltage causes the precipitated antibody sandwiches to glow.

The entire test runs on Roche's well-established cobas automation platform. "We have twenty years of experience in developing all kinds of immunoassays [on it]," says Ofenloch-Haehnle. Because Elecsys hardware is already used for other assays in more than 40,000 diagnostic and research labs around the world, the new SARS-CoV-2 exposure test was easy for users to adopt.

High levels of automation and decades of user-interface refinements mean that the cobas-driven Elecsys assay is also easy for new entrants to the field to pick up. Indeed, Ofenloch-Haehnle says many labs have bought cobas systems for the first time just to run the new SARS-CoV-2 test. Using a 12-microliter blood sample, the entire process takes 18 minutes. The largest cobas systems can run 300 tests an hour, requiring the technician to do little more than load and unload samples.

Because it detects the viral nucleocapsid, the Roche qualitative test is useful for identifying individuals who've been infected with the virus. All the vaccines currently in late-stage clinical development, however, are designed to stimulate immune responses only against the external spike glycoprotein. The Elecsys qualitative SARS-CoV-2 test doesn't detect those antibodies. "So, we decided to develop a second assay, and this time a quantitative assay, because you need quantitation to assess antibody levels and come up with potential thresholds for revaccination or immunity," says Ofenloch-Haehnle.

The second assay, released in the United States and Europe in September 2020, uses the same Elecsys platform to provide quantitative measurements of anti-spike IgG antibodies. The two assays could also be useful together in some settings. "While nucleocapsid is solely from natural infection, having the [anti-spike assay], one could see in vaccinated individuals whether some natural infection was going on and also differentiate between the response to the vaccine and the natural infection," explains Ofenloch-Haehnle.

Time for T

While antibody tests have been crucial for monitoring the scope of the pandemic and tracking vaccine responses, they measure only one side of the immune response to SARS-CoV-2. Besides acquiring humoral immunity with antibodies, people who contract the virus also develop cellular immunity mediated by killer T cells. "T cells are very relevant in how your body responds to SARS-CoV-2," says Peter Wrighton-Smith, chief executive officer of **Oxford Immunotec** in Abingdon, United Kingdom.

T-cell responses are more difficult to measure, though, and both research and clinical laboratories have struggled to determine exactly how important cellular immunity is in COVID-19. "One of the problems we have in the academic literature is that the findings are ... all using different methods," says Wrighton-Smith. To address that, his company began producing standardized kits for research labs to test SARS-CoV-2 T-cell responses. Released in May 2020, those kits then became a centerpiece of a study the company undertook with Public Health England, analyzing T-cell responses in health care workers and first responders.

"We tested people at the beginning, and then we followed them up for months afterwards to see who was getting SARS-CoV-2," says Wrighton-Smith. The results revealed that T cell–mediated immunity was a common feature in many infections, including some that didn't produce robust antibody responses. Critically, a robust T-cell response correlated strongly with immunity from reinfection; people who failed to develop such responses were significantly more likely to get reinfected. "T cells are probably more important than antibodies in determining who has immunity," says Wrighton-Smith.

To measure cellular immunity in clinical settings, Oxford Immunotec turned to T-Spot, the platform they've been developing for the past 18 years. With this system, technicians isolate T cells from a blood sample, count and distribute them into multiwell plates, and then perform an assay to measure the number of cells that react to the test antigen. The company's first product, T-SPOT.*TB*, is already in use worldwide for tracking cellular responses to tuberculosis. "By discerning whether you have these cells, we're able to confirm whether you currently have the infection or have recently seen it, and by counting how strong the response is, we may also be able to tell ... the extent of your immunity," says Wrighton-Smith.

The T-SPOT *Discovery* SARS-CoV-2 test can be used as either a qualitative or quantitative test, and the company expects to get it licensed by early 2021. "We're working feverishly to get that to happen, [because] we have a technology that we think will be really important in the next phase of the pandemic," says Wrighton-Smith. Like others working on pandemic-related technologies, he emphasizes that while the arrival of effective vaccines promises to bring the virus under control, the need for new innovations–both for this pandemic and future ones–will persist.

Alan Dove is a science writer and editor based in Massachusetts.



SARS-CoV-2 Testing Microplate
Porvair Sciences has introduced a new 2.2-mL 96 deep-well plate that is fully compatible with the Thermo Scientific KingFisher range of nucleic acid purification systems. These systems are reported to drastically reduce processing time and increase productivity. Designed for compatibility and efficiency, each v-shaped bottom well in Porvair's new plate supports the specialized magnetic tips of all KingFisher instruments, maximizing liquid-sample collection, mixing, and uptake during the purification process, and ensuring reproducible purification of nucleic acids from SARS-CoV-2 patient samples.
Porvair Sciences
For info: 800-552-3696
www.microplates.com/kingfisher-compatible-96-well-microplate

SARS-CoV-2 Monoclonal Antibodies
The Native Antigen Company introduces 10 new monoclonal antibodies recognizing SARS-CoV-2, the causative agent of COVID-19. These antibodies recognize different epitopes of the SARS-CoV-2 spike glycoprotein, nine of which are SARS-CoV-2 specific and do not cross-react with other human coronaviruses; they extend our existing range and offer researchers a source of high-quality reagents for R&D of in vitro diagnostics, therapeutics, and vaccines for COVID-19. The SARS-CoV-2 antibodies were raised using our own S1 and S2 recombinant spike proteins as immunogens. Produced using our proprietary mammalian expression system, these antigens exhibit full glycosylation and proper folding to ensure that antibodies raised against them are highly specific. The antibodies have been screened against our other coronavirus proteins, including SARS, MERS, NL63, OC43, 229E, and HKU1, to demonstrate specificity. Three of the antibodies have been shown to be specific for the receptor-binding domain (RBD) of the spike protein of SARS-CoV-2 bound to the cellular angiotensin-converting enzyme 2 (ACE2) receptor.
The Native Antigen Company
For info: +44-(0)-1865-595230
thenativeantigencompany.com

SARS-CoV-2 Antibodies and Antigens
AMS Biotechnology has further expanded its product range to help in the fight against COVID-19, with the introduction of a new range of SARS-CoV-2 antibodies and antigens to aid development of new immunoassay diagnostic kits. A suite of SARS-CoV-2 mutant spike and nucleocapsid antigens is now available. Perfect for use in a wide range of different antibody detection methods, these antigens are customizable and available in bulk. In addition, we have also introduced a new range of matched antibody pairs specifically designed and screened for antigen detection kit development. Verified against inactivated COVID-19 virus samples, these antibody pairs are ideal for development of diagnostic assays, including chemiluminescent immunoassays, ELISAs, and gold immunochromatographic assays [also known as lateral flow assays (LFAs)]. The sensitivity of these matched antibody pairs when used in a sandwich ELISA kit has been shown to be as low as 12 pg/mL. The virus titer, which was determined using an LFA, is as low as 102 cell culture infectious dose 50% (CCID50)/mL.
AMS Biotechnology
For info: 617-945-5033
www.amsbio.com/antibodies-antigens-sars-cov-2-diagnostic-kits

PCR Test to Differentiate Between COVID-19 and Flu
EUROIMMUN, a PerkinElmer company, announces the launch of the European Conformity (CE)-marked EURORealTime SARS-CoV-2/Influenza A/B PCR test for direct detection of SARS-CoV-2, influenza virus type A, and influenza virus type B. It expands the testing portfolio for acute COVID-19 diagnostics by supporting differential diagnostics between SARS-CoV-2 infections and the common flu, and is available in countries accepting the CE mark. PCR technology is considered the gold standard for direct pathogen detection. The EURORealTime test allows fast detection and differentiation of genetic material using throat-swab samples of patients with acute symptoms. Validation efforts revealed high agreement of results obtained with the EURORealTime test and those obtained with reference PCR tests for SARS-CoV-2 and influenza A/B. No cross-reactions with other common respiratory pathogens were detected. The assay is compatible with common real-time PCR thermal cyclers, and the EURORealTime Analysis Software allows for reliable, standardized evaluation of test results.
PerkinElmer
For info: 800-762-4000
www.perkinelmer.com

Rapid Antibody Test Kit for SARS-COV-2
MP Biomedicals Asia Pacific announces the successful development of the ASSURE SARS-CoV-2 IgG/IgM rapid test, a rapid antibody test kit for COVID-19, in collaboration with the Agency for Science, Technology and Research (A*STAR). It produces accurate results in as little as 15 min and employs a lateral flow format similar to those used in home pregnancy tests. The test kit detects immunoglobulin G (IgG) and immunoglobulin M (IgM) antibodies, which are produced by the body to combat SARS-CoV-2, from samples of human blood, plasma, or serum. ASSURE should not be used in the clinical diagnosis of COVID-19 infections or in the evaluation of persons with acute respiratory symptoms, especially within the first 14 days of illness. However, ASSURE can help to determine whether an individual has been previously exposed to the virus and has generated antibodies as a result. This can help identify asymptomatic individuals or those with only mild symptoms who were not subjected to reverse transcription-PCR testing.
MP Biomedicals
For info: 800-854-0530
www.mpbio.com/about-us

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Confused about
your next career move?

Download Free Career
Advice Booklets!

ScienceCareers.org/booklets



DIRECTOR OF MANUFACTURING AUTOMATION

Takeda seeks a **Director of Manufacturing Automation** to lead our Cell Therapy program, including both research and development efforts as well as cell manufacturing. The successful candidate would continue to build and expand upon our research program, developing automated manufacturing processes to produce cell therapy product for both liquid and solid tumors. Experience in engineering immunological cell therapies such as CAR-Ts is required, as is familiarity with cutting-edge new technologies, such machine learning-based processing technologies and closed-loop adaptive manufacturing. The ideal candidate would have experience in leading and managing research, development, and manufacturing of anti-cancer cell therapies, as well as experience with developing and testing cell therapy agents in in vivo and in vitro studies, and collaborating with research partners in both academic and clinical settings. Competitive salary of $225,000-$273,000 and benefits.

Takeda is a company devoted to developing a new class of cell therapy product to dramatically expand cell therapy cures for cancer. Our pipeline includes both autologous and allogeneic therapies, spanning the fields for CAR-T, CAR-NK, stem cells and iPSC. At Takeda, we believe that working in multidisciplinary teams is essential to innovation. We have diverse backgrounds in cell therapy, immunology, protein engineering, nanoparticle engineering, computational biology, drug development, and medicine. We are always looking for talented individuals with an entrepreneurial spirit.

Takeda is an affirmative action/equal opportunity employer and encourages applicants of all backgrounds to apply. Interested applicants should apply online by sending a letter of interest, CV, and three letters of reference to **www.takedajobs.com** and search for **Job ID #R0027222**

SOMETIMES
THE GRASS REALLY
IS GREENER
SOMEPLACE NEW.

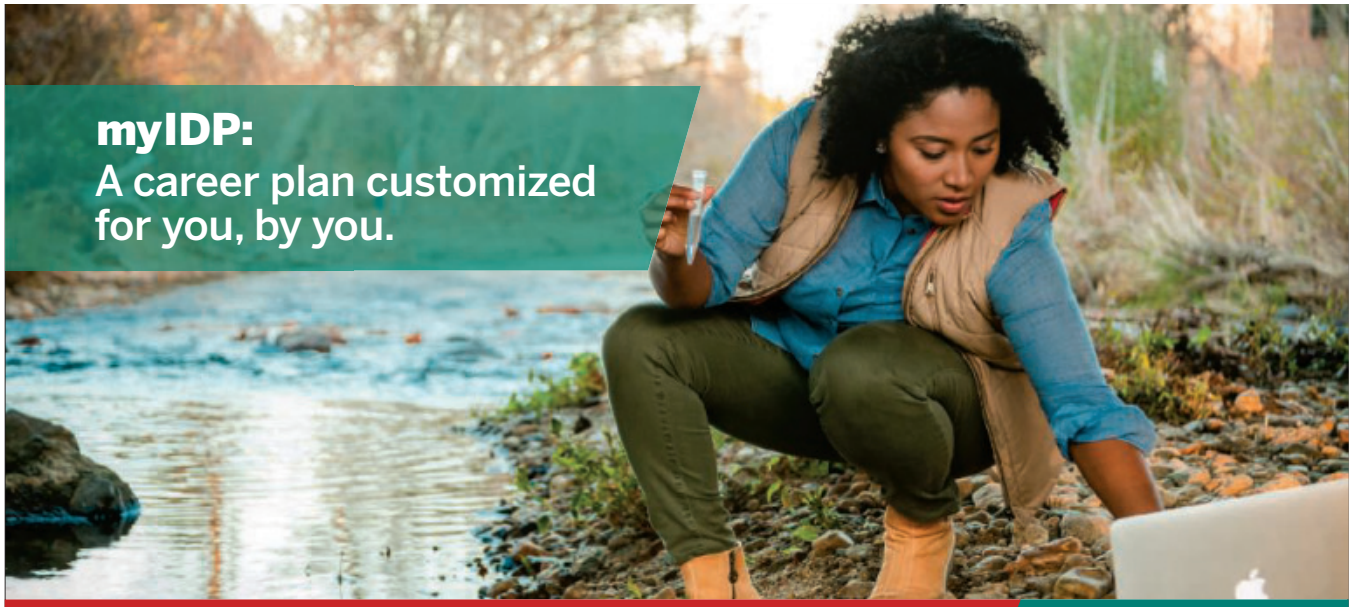


Find your next job at ScienceCareers.org

There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: sciencecareers.org

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Science Careers online @sciencecareers.org



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.
- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- Ability to select which portion of your IDP you wish to share with advisors, mentors, or others.
- A certificate of completion for users that finish myIDP.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers AAAS In partnership with:

FASEB Federation of American Societies for Experimental Biology
University of Massachusetts Medical School
UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND

Who's the top employer for 2020?

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles at sciencecareers.org/topemployers

Science 2020
TOP EMPLOYER



By **Paul Bierman**

A department terminated

I logged onto a video call last month, not knowing what to expect. Faculty members in my department had just 1 day's notice of the meeting with our dean, which made me wonder, "Is it another budget cut—or worse?" After a quick apology, the dean pivoted to the news: Given the crunch on university finances amid the COVID-19 pandemic, he'd made a difficult decision to terminate the geology department. If we wanted to keep our jobs, we'd have to find another department willing to take us in. His words hit me like a ton of bricks. For 20 years—ever since I received tenure—I assumed I'd always have a job if I kept up my teaching and research. But the pandemic had brought long-standing problems to a head.

Three years ago, the same dean had asked us for help. His budget was in the red because enrollments were falling and a new budget system had moved money out of his college. The geology department had a few large, popular courses, but many of our upper-level classes were highly specialized, attracting only a handful of students. Fewer than 10 students a year majored in geology. We needed larger class sizes and more students to enroll as majors, the dean told us.

We convened faculty meetings to discuss how to move forward. Our curriculum hadn't changed much in decades. Our students learned how to identify rocks and stare down microscopes, but they weren't much exposed to many of the more pressing problems in geosciences, such as climate change and groundwater pollution. Some of us, including me, wanted to overhaul the curriculum. But others argued against abandoning our focus on traditional skills and concepts.

We ended up making changes only around the edges. We added new elective courses on climate, medical geology, and extraterrestrial life, which attracted hundreds of students—many from other departments. But we didn't change the course requirements for geology majors, and many professors continued to teach the same material in the same way. The number of majors didn't budge.

Then came COVID-19. We stayed open and taught students on campus, but enrollments continued to fall. We expected administrators to make cuts, but we did not think entire departments would be axed—especially not ours. Our seven professors collectively bring in hundreds of thousands of federal research dollars every year and publish dozens of papers. But the overhead from our grants did not make up for the lack of tuition dollars. The quality of our research wasn't enough to save us.



"The quality of our research wasn't enough to save us."

After the meeting with our dean, my mind raced. What would become of our labs, our grad students, and all the projects in full swing? I had a secondary faculty appointment in a different college on campus, so I might be able to work there and avoid leaving the town where I'd lived for 27 years. But I didn't want to leave my geology colleagues. I wanted a solution that kept us together as a team.

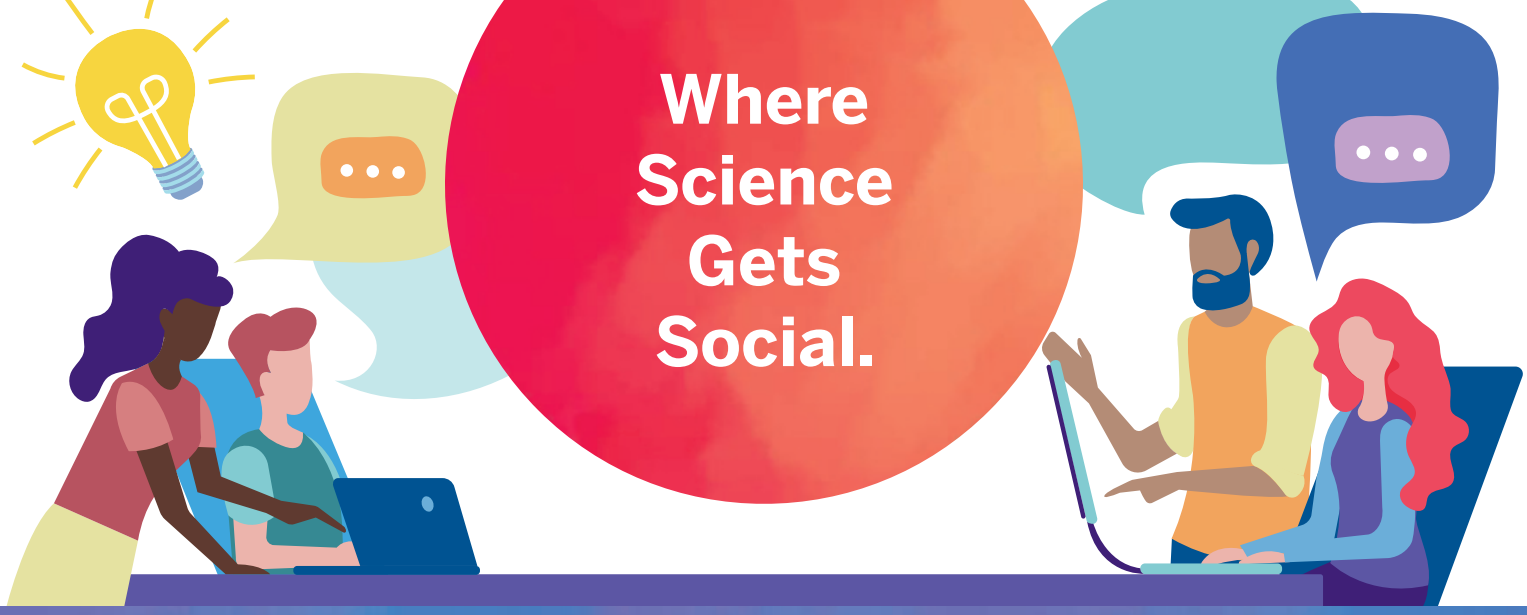
Two days after the dean's call, I spoke with a university administrator who prodded me to not give up and asked, "What's your vision?" I spent the weekend speaking with my colleagues and brainstorming a plan. Together, we imagined a wholly new earth science curriculum, one that would prepare students for the challenges of today. We'd rethink all of our courses from scratch, focusing on problems that students—and their potential employers—care about. For instance, instead of simply teaching how rocks crack and weather away, we'd explore how those cracks affect the movement of polluted groundwater.

Our new vision looks to the future and leaves the past behind, something we could never bring ourselves to do before the prospect of termination forced us to spring into action. We're not sure how the university will react to our plan. Hopefully, our efforts aren't coming too late.

I hope academics elsewhere can learn from our mistakes and take a hard look at their own offerings to make sure they're serving the current generation of students. It's hard to overhaul a department and develop entirely new courses. But it's much harder if you wait until a crisis to act. ■

Paul Bierman is a Gund Fellow and professor at the University of Vermont.

ILLUSTRATION: ROBERT NEUBECKER



AAAS.ORG/COMMUNITY







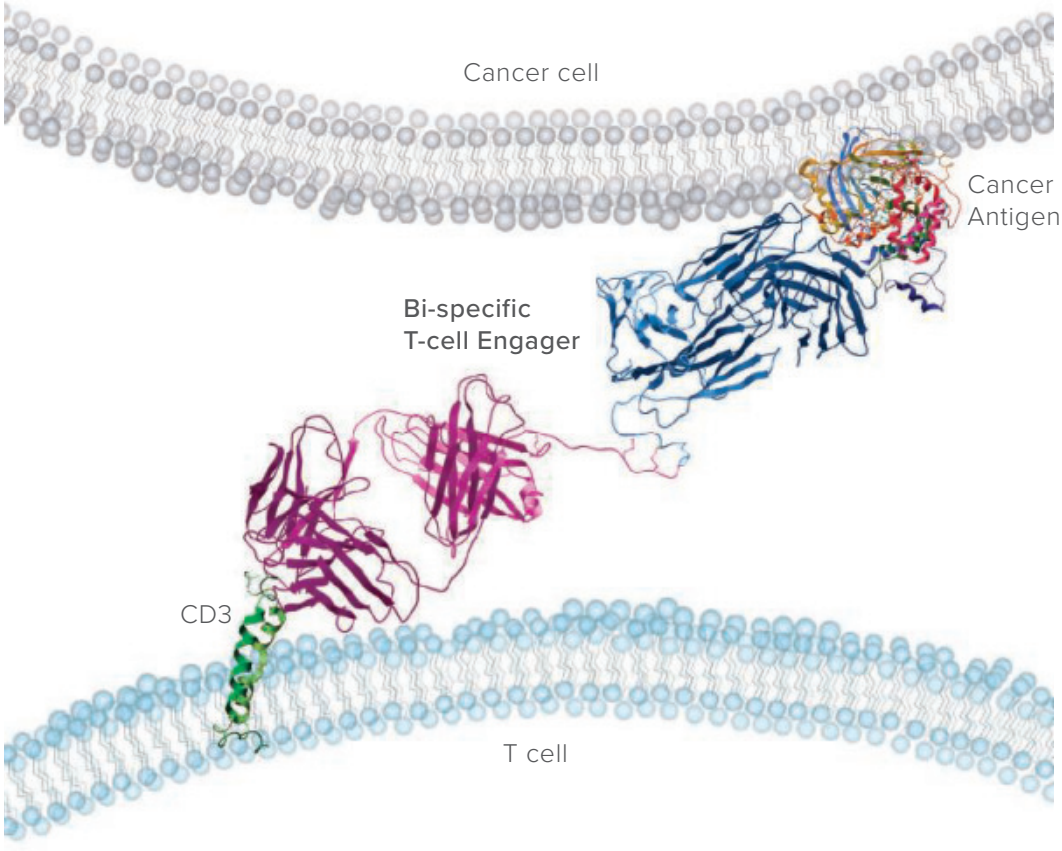
AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Bispecific Antibody Development

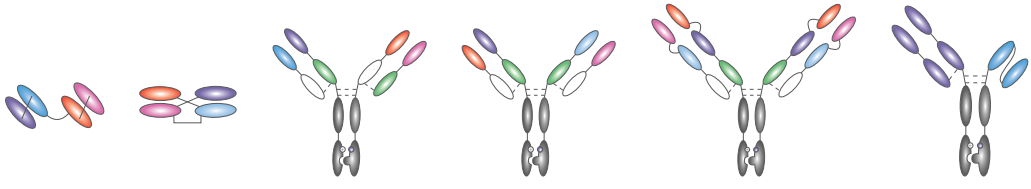
Custom Bispecific Antibodies for Research Use



Learn how ProMab's services can help you establish a proof-of-concept for your bispecific antibody and accelerate your antibody therapeutics research.

Personalized Antibody Solutions

- Bispecific Antibody Design
- Bispecific Antibody Development
- Bispecific Antibody Validation
- Monoclonal Antibody Development
- Antibody Humanization
- Antibody Producing Stable Cell Line



Stylized overview of the numerous designs used to generate bispecific antibodies.

All products are for research only

Discover more | www.promab.com

ProMab Biotechnologies, Inc.
2600 Hilltop Dr, Building B, Richmond, CA 94806
1.866.339.0871 | info@promab.com