Ecologists battle invasive coconut palms p. 1046

Digital health transformation in India p. 1050

Uncovering the origin of Earth's water pp. 1058 & 1110

Science

$15
28 AUGUST 2020
sciencemag.org

AAAS



CHEMISTRY BY

QUBIT

Quantum chip simulates molecular isomerization

pp. 1054 & 1084

CONTENTS

28 AUGUST 2020 • VOLUME 369 • ISSUE 6507



1046

On Palmyra Atoll in the central Pacific, a restoration project seeks to remove millions of invasive coconut palms.

NEWS

IN BRIEF

1036 News at a glance

IN DEPTH

1038 In plasma OK, critics see politics, not science
FDA's emergency move raises fears of rushed approval for other experimental medicines
By K. Kupferschmidt and J. Cohen

1039 Critics slam NIH for demands on bat grant targeted by Trump
Nobelists, former agency leaders deplore politicization *By M. Wadman*

1040 How Italy's 'father of the swabs' fought the virus
Microbiologist Andrea Crisanti defied the authorities to launch his testing campaign
By D. Starr

1041 Spit shines for easier coronavirus testing
Tests using saliva are cheaper and faster than those with nasal swabs—and can be just as accurate *By R. F. Service*

1042 Hidden web of fungi could shape the future of forests
Plants' fungal partners can help their hosts resist drought and heat—or make them more vulnerable *By E. Pennisi and W. Cornwall*

1043 New feedbacks speed up the demise of Arctic sea ice
Wave action and a growing underwater blob of heat add to impact of warming air
By P. Voosen
PODCAST

1045 How 'elite controllers' tame HIV without drugs
Rare people who naturally subdue the virus may weed out cells with active HIV genomes
By J. Cohen

FEATURES

1046 Ridding paradise of palms
On Pacific islands, coconut palms are an aggressive invader. Biologists are trying to eradicate millions of palms from one atoll
By K. R. Weiss

INSIGHTS

PERSPECTIVES

1050 Bridging digital health divides
COVID-19 is accelerating digital transformation in the Indian health sector
By A. Agrawal

1052 Restoring neuron connections
A synthetic peptide promotes functional recovery in neurodegeneration and spinal cord injury models *By P. C. Salinas*
RESEARCH ARTICLE p. 1074

1054 A quantum-computing advantage for chemistry
Error-mitigated Hartree-Fock computation is performed with up to 12 superconducting qubits *By X. Yuan*
RESEARCH ARTICLE p. 1084

1055 Zika virus increases risk of dengue disease
Interactions between Zika virus immunity and dengue disease have implications for vaccination *By H. Clapham*
REPORT p. 1123

1056 Strategies for navigating a dynamic world
A rich neural representation of its environment enables an adaptable organism to respond to changes
By S. Steixner-Kumar and J. Gläscher
RESEARCH ARTICLE p. 1076

1058 The origins of water
New measurements of Earth's building blocks point to a simpler water source *By A. H. Peslier*
REPORT p. 1110

1059 Zena Werb (1945–2020)
Pioneering cancer biologist and beloved mentor *By N. Boudreau and M. Bissell*

POLICY FORUM

1060 Computational social science: Obstacles and opportunities
Data sharing, research ethics, and incentives must improve *By D. M. J. Lazer* et al.

PHOTO: FARLOW/OLSON/NATIONAL GEOGRAPHIC

Published by AAAS

BOOKS *ET AL.*

1063 Secrets of the Vikings
The violent seafarers are brought to vivid life in a rich synthesis of old and new scholarship
By T. H. McGovern

1064 A history of insatiable intellectuals
Proficient in a range of topics, polymaths are overlooked as often as they are celebrated
By A. Robinson

LETTERS

1065 Development projects jeopardize India's forests
By S. Khadakkar

1065 North African forests falling to charcoal
By R. Khelifa

1066 Protect Iran's Zagros forests from wildfires
By M. Kheshti

1066 Errata

RESEARCH

IN BRIEF

1069 From *Science* and other journals

REVIEW

1072 Paleoecology
Using paleo-archives to safeguard biodiversity under climate change *D. A. Fordham* et al.
REVIEW SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABC5654
PODCAST

RESEARCH ARTICLES

1073 Microbial engineering
A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose *R. L. Shahab* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB1214



PHOTO: REUTERS/OSWALDO RIVAS

1074 Neuroscience
A synthetic synaptic organizer protein restores glutamatergic neuronal circuits
K. Suzuki et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB4853
PERSPECTIVE p. 1052

1075 Cell biology
Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown *L. Tweedy* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY9792

1076 Neuroscience
Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex
P. Domenech et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB0184
PERSPECTIVE p. 1056

1077 Microbiology
Diverse enzymatic activities mediate antiviral immunity in prokaryotes
L. Gao et al.

1084 Quantum computing
Hartree-Fock on a superconducting qubit quantum computer
Google AI Quantum and Collaborators
PERSPECTIVE p. 1054

1089 Structural biology
Architecture of a catalytically active homotrimeric plant cellulose synthase complex *P. Purushotham* et al.

REPORTS

1094 Bacterial metabolism
A nitrogenase-like enzyme system catalyzes methionine, ethylene, and methane biogenesis
J. A. North et al.

1099 Electrochemistry
Accelerating water dissociation in bipolar membranes and for electrocatalysis
S. Z. Oener et al.

1104 Carbon nanotubes
Super-durable ultralong carbon nanotubes
Y. Bai et al.

1106 Coronavirus
Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions
S. T. Ali et al.

1110 Cosmochemistry
Earth's water may have been inherited from material similar to enstatite chondrite meteorites *L. Piani* et al.
PERSPECTIVE p. 1058

1113 Organic chemistry
Static to inducibly dynamic stereocontrol: The convergent use of racemic β-substituted ketones *J. S. DeHovitz* et al.

1119 Coronavirus
Structural basis of a shared antibody response to SARS-CoV-2
M. Yuan et al.

1123 Flaviviruses
Zika virus infection enhances future risk of severe dengue disease
L. C. Katzelnick et al.
PERSPECTIVE p. 1055

1128 Immunogenomics
Mosquito cellular immunity at single-cell resolution *G. Raddi* et al.

1132 Protein design
Expanding the space of protein geometries by computational design of de novo fold families *X. Pan* et al.

DEPARTMENTS

1035 Editorial
The danger of DIY vaccines
By Arthur L. Caplan and Alison Bateman-House

1138 Working Life
One, two, three, thrive *By Sieun Ruth Lee*

ON THE COVER



The Sycamore quantum processor, composed of an array of superconducting transmon qubits, simulates the energy of the diazene molecule in various conformations. By combining error mitigation and improved qubit calibration techniques, the simulation accurately models the isomerization pathway that reorients the hydrogen atoms (white) about the central bonds between nitrogen atoms (blue). See pages 1054 and 1084. *Illustration: C. Bickel/Science*

Science Staff 1034
AAAS News & Notes 1067
Science Careers 1137

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.

Copyright © 2020 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS.

Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Published by AAAS

Science AAAS

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

Executive Editor Monica M. Bradford

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

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

News Editor Tim Appenzeller

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

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Xing Liu **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **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 **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

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

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

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

---

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

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

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

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

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

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

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

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

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

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

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

Adriano Aguzzi, *U. Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Judith Allen, *U. of Manchester*
Sebastian Amigorena, *Institut Curie*
James Analytis, *UC Berkeley*
Trevor Archer, *NIEHS, NIH*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Clare Baker, *U. of Cambridge*
Nenad Ban, *ETH Zürich*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *U. of Texas at Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Gabriele Bergers, *VIB*
Bradley Bernstein, *Mass. General Hospital*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med. School*
Peer Bork, *EMBL*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *U. of California, Irvine*
György Buzsáki, *New York U. School of Med.*
Blanche Capel, *Duke U.*
Annmarie Carlton, *U. of California, Irvine*
Nick Chater, *U. of Warwick*
M. Keith Chen, *UCLA*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Carolyn Coyne, *U. of Pittsburgh*
Roberta Croce, *VU Amsterdam*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *U. of North Carolina*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Christian Davenport, *U. of Michigan*
Frans de Waal, *Emory U.*
Claude Desplan, *New York U.*
Sandra Díaz, *U. Nacional de Córdoba*
Ulrike Diebold, *TU Wien*
Hong Ding, *Inst. of Physics, CAS*
Jennifer Dionne, *Stanford U.*
Dennis Discher, *U. of Penn.*
Gerald Dorn, *Washington U. in St. Louis*
Jennifer A. Doudna, *UC Berkeley*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Christopher Dye, *U. of Oxford*
Todd Ehlers, *U. of Tübingen*
Jennifer Elisseeff, *Johns Hopkins U.*
Tim Elston, *U. of North Carolina*
Andrea Encalada, *U. San Francisco de Quito*
Nader Engheta, *U. of Penn.*
Karen Ersche, *U. of Cambridge*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Michael Feuer, *The George Washington U.*
Toren Finkel, *U. of Pittsburgh Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzl, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Eileen Furlong, *EMBL*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *UCLA*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *U. of South Florida*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Kip Guy, *U. of Kentucky College of Pharmacy*
Taekjip Ha, *Johns Hopkins U.*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. College London*
Jian He, *Clemson U.*
Carl-Philipp Heisenberg, *IST Austria*
Ykä Helariutta, *U. of Cambridge*
Janet G. Hering, *Eawag*
Hans Hilgenkamp, *U. of Twente*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Lora Hooper, *UT Southwestern Med. Ctr.*
Fred Hughson, *Princeton U.*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS and U. of Arizona*
Kai Johnsson, *EPFL*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *UC Berkeley*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med. School*
Nancy Knowlton, *Smithsonian Institution*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Thomas Langer, *Max Planck Inst. Cologne*
Mitchell A. Lazar, *U. of Penn.*
Wendell Lim, *UC San Francisco*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *Washington U. in St. Louis*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Asifa Majid, *U. of York*
Oscar Marín, *King's College London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth College*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Graham Medley, *LSHTM*
Jane Memmott, *U. of Bristol*
C. Jessica Metcalf, *Princeton U.*
Baoxia Mi, *UC Berkeley*
Edward Miguel, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Yasushi Miyashita, *U. of Tokyo*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci., Res. & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Harry Orr, *U. of Minnesota*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Margaret Palmer, *U. of Maryland*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Daniel Pauly, *U. of British Columbia*
Samuel Pfaff, *Salk Inst. for Biological Studies*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Institut Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UC San Francisco*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Félix A. Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial College London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *U. of Texas at Austin*
Shimon Sakaguchi, *Osaka U.*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Jürgen Sandkühler, *Med. U. of Vienna*
Erica Ollmann Saphire, *La Jolla Inst.*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Rebecca Sear, *LSHTM*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Washington*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UC San Francisco*
Robert Siliciano, *Johns Hopkins U. School of Med.*
Lucia Sivilotti, *U. College London*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina* **(S)**
Mark Smyth, *QIMR Berghofer*
Pam Soltis, *U. of Florida*
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Dartmouth College*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Wim van der Putten, *Netherlands Inst. of Ecology*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Kathleen Vohs, *U. of Minnesota*
David Wallach, *Weizmann Inst. of Sci.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Jonathan Weissman, *UC San Francisco*
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, *U. of Penn. School of Med.*
Jonathan Zehr, *UC Santa Cruz*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

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

Published by AAAS

EDITORIAL

The danger of DIY vaccines

The world needs to trust science if vaccines are to prove useful, particularly those being developed to combat coronavirus disease 2019 (COVID-19). That is what makes the recent appearance of highly visible "do-it-yourself" (DIY) vaccine research so morally troubling. It's an obstacle to securing this trust.

As reported in last month's MIT Technology Review, at least 20 people are following the lead of geneticist and entrepreneur Preston Estep to take and promote a homebrew potential vaccine for COVID-19. They have formed a group, Rapid Deployment Vaccine Collaborative (RaDVaC), with the avowed mission of rapidly developing and sharing a vaccine recipe simple enough to be produced and administered by the public. On RaDVaC's website, the group describes itself as citizen scientists, although "many of us are trained scientists and engineers."

RaDVaC describes the formulation of its vaccine as containing peptides that prompt the user's body to generate antibodies to severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2, the virus that causes COVID-19). Furthermore, chitosan, a substance found in the shells of crustaceans, is added to coat the peptides and ease their delivery through mucosal tissue in the nose. The RaDVaC developers deliver their vaccine in a nasal spray in an attempt to trigger a strong, localized immune response in an area of the body where COVID-19 infection often takes hold. An untold number of them have taken one or more doses of this concoction, including the renowned Harvard geneticist George Church.

So, what is wrong with a tiny group of high-profile scientists and their admirers developing a vaccine, administering it to themselves, and distributing the formulation to others who choose to try it? A great deal is wrong.

The DIY RaDVaC initiative is far more likely to contribute to growing public mistrust of all vaccines than it is to provide a path forward to combating the pandemic. Those who are increasingly mistrustful of all the talk of "warp speed" in promising a COVID-19 vaccine are hardly going to be encouraged to change their minds by rogue scientists experimenting with no oversight at the fringes of what is ethically acceptable.

The DIY effort has no animal or safety trials; no confirmation of safety by closely monitoring healthy volunteers; no dosage studies; no effort to review the proposed science or recruitment of volunteers by an outside, independent ethics review committee; no plan to record all users, to encourage diversity among users, or for systematic follow-up; and no plan to provide help or compensation to anyone harmed by their participation. Moreover, there have been no papers published or data released in peer-reviewed outlets about the vaccine. The research is rife with conflicts of interest in that those making the vaccine are recruiting friends to try it while promoting their actions in the media. They are not selling their vaccine but stand to benefit from attention in the media and any resulting philanthropic support.

> "The DIY... initiative is...likely to contribute to growing public mistrust of all vaccines..."

There are by most estimates around 200 COVID-19 vaccines in development around the world. Some three dozen are in human trials. A handful have progressed to full-fledged clinical trials. News accounts report that large-scale vaccination efforts without full human safety testing or ethical review are underway in Russia. Given the horrors inflicted on people around the world by the pandemic, it would be reasonable to expect that concerted efforts to find a vaccine would have huge popular support, but that is not the case. Large percentages of people polled in many nations say they will not use or are worried about the safety of a COVID-19 vaccine. Nearly half of all those polled in the United States and United Kingdom in recent months said they would refuse vaccination.

COVID-19 vaccine skepticism has many sources, including doubts about the trustworthiness of safety claims made by government leaders whom many distrust; the perceived inability of government regulatory agencies to maintain independence from political pressure to quickly approve products; and impressions of financial conflicts of interest on the part of those making vaccines.

Trust is the key ingredient in any effort to facilitate a vaccine solution to the current pandemic. Peer-reviewed science transparently assessed in carefully controlled trials by independent experts is the only way to cement that trust. DIY vaccinology is dangerous at a time when non-evidence–based claims of COVID-19 "cures" have done little but sow mistrust of science and public health.

–Arthur L. Caplan and Alison Bateman-House

Arthur L. Caplan
is the director of the Division of Medical Ethics at New York University Grossman School of Medicine, New York, NY, USA. arthur.caplan@nyumc.org

Alison Bateman-House
is an assistant professor in the Division of Medical Ethics at New York University Grossman School of Medicine, New York, NY, USA. alison.bateman-house@nyulangone.org

10.1126/science.abe4440

Published by AAAS

NEWS

IN BRIEF

Edited by **Jeffrey Brainard**



COVID-19

Pop concertgoers help test virus risks

About 1400 volunteers showed their support of music and science by attending an unusual, experimental concert in Leipzig, Germany, last week. Wearing masks, they gathered in a 12,000-seat arena to listen to pop singer Tim Bendzko and his band—and to help researchers test ways to keep audiences in large indoor gatherings safe from the spread of the COVID-19 virus. The volunteers had tested negative for the virus and passed a temperature check before entering the venue. Scientists from Martin Luther University of Halle-Wittenberg gave each participant a tracking device that measured how close they came to others. Attendees used a fluorescent-tagged disinfectant on their hands, allowing the researchers to see which surfaces were touched the most. The scientists also used smoke machines and carbon dioxide sensors to trace the movement of air throughout the arena—including when Bendzko encouraged the audience to sing along. During the extended concert, which lasted 10 hours, they tested three patterns of entering and exiting the hall and varying distances between seats. They hope to release results of their study by early October.

First COVID-19 reinfection ID'd

INFECTIOUS DISEASES | Scientists have found the first solid evidence that people can be reinfected with the virus that causes COVID-19, raising questions about immunity to the disease. A study published on 25 August in *Clinical Infectious Diseases* describes a 33-year-old man in Hong Kong who was treated for a mild case in March; he harbored the virus again when he was tested at the airport after returning from Europe on 15 August, when he showed no symptoms. The university team sequenced the virus, SARS-CoV-2, from both infections and found numerous differences between the two, bolstering the case for a second infection. The report has researchers wondering how common such cases are, whether their symptoms tend to be mild, and whether they can infect others.

Neutrino hunter gets upgrade

ASTROPHYSICS | Scientists announced last week a completed upgrade to the Super-Kamiokande (Super-K) detector in central Japan that will help it spot neutrinos emanating from supernovae. For the modification, researchers added the rare-earth metal gadolinium to the purified water in Super-K's large observation tank. Each year, they expect, a few supernova antineutrinos will hit water molecules, producing neutrons. Gadolinium nuclei are effective at absorbing the neutrons and producing gamma rays that Super-K can detect. Such observations will provide new clues in understanding the physics and history of supernovae.

Playing around with ecosystems

ECOLOGY | A new video game inspired by the reintroduction of wolves to Yellowstone National Park allows players to build realistic virtual ecosystems while contributing to ecological research. EcoBuilder differs from similar games, such as Niche, because it runs on mathematical equations that researchers use to simulate species interactions and changes to ecosystems, says creator Jonathan Zheng, a graduate

PHOTO: SEAN GALLUP/GETTY IMAGES

Published by AAAS

student at Imperial College London. For example, the game can predict how changes in the abundances of predators like the Yellowstone wolves affect the survival of herbivores they prey on and plants eaten by the herbivores. Researchers will study the players' game results with their consent and use the data to help refine the equations used in the simulations.

Gold mine faces new hurdle

FISHERIES | Following criticism by prominent Republicans, U.S. regulators on 24 August said they could not issue permits for a massive gold and copper mine in the headwaters of the world's largest sockeye salmon run because of the "substantial environmental impacts." The move by the U.S. Army Corps of Engineers appeared to reverse findings from its report in July, which found the Pebble Mine near Bristol Bay, Alaska, would not measurably affect the salmon fishery. Since then, several Republicans close to President Donald Trump have spoken out against the project, including his son, Donald Trump Jr.; Fox News commentator Tucker Carlson; and Bass Pro Shops founder Johnny Morris. Pebble Partnership CEO Tom Collier said the company was working on a mitigation plan that would answer the Corps's concerns. But environmentalists said the agency's new position could kill the mine.

Fetal tissue proposals nixed

BIOMEDICINE | A Trump administration board last week recommended against funding 13 of 14 research proposals that would have used human fetal tissue donated after elective abortions. The applications, submitted to the U.S. National Institutes of Health (NIH), had been recommended for funding after scientific and ethical reviews. But in 2019, President Donald Trump imposed an extra layer of ethical review on such proposals, and on 31 July the NIH Human Fetal Tissue Research Ethics Advisory Board met to consider the first proposals submitted under the new regime. The panel released its recommendations last week to U.S. Health and Human Services Secretary Alex Azar; the recommended rejections take effect automatically. The panel's objections included insufficient justifications for the use of the tissue and purported inadequacies in informed consent forms. At least 10 of the board's 15 members have publicly opposed abortion, fetal tissue research, human embryonic stem cell research, contraception, or a combination of these.

Fixes for satellite tracks proposed

ASTRONOMY | Planned constellations of thousands of communications satellites in low orbits will mar many images from ground-based optical and infrared telescopes and may alter the night sky for naked-eye watchers, a report concludes. SpaceX already has 655 satellites in orbit and others, including OneWeb and Amazon, will soon follow, possibly adding tens of thousands. A U.S. National Science Foundation workshop several weeks ago brought together 250 astronomers and engineers; their final report, released on 25 August, says twilight observations, such as searches for Earth-threatening asteroids and optical signals from gravitational wave events, will be worst affected. The report suggests operators could reduce problems by launching fewer satellites at lower altitudes and lessening their reflectivity. And astronomers could help design digital tools to mask or remove trails of satellites in images.

Release of GM mosquitoes OK'd

BIOTECHNOLOGY | Local officials in Florida last week cleared the way for the first release of genetically modified mosquitoes. The strain of *Aedes aegypti* to be let go was developed by the U.K.-based company Oxitec to combat the spread of viruses including dengue and Zika. The green light from the Florida Keys Mosquito Control District Board of Commissioners follows reviews by state regulators and a determination by the U.S. Environmental Protection Agency that the release poses "no unreasonable adverse effects" to humans or the environment. Starting in 2021, the company plans to put out 750 million engineered male mosquitoes to breed with wild females, spreading a gene that kills female offspring before they reach maturity and can bite people and spread disease. (Male offspring survive to pass on the lethal gene to the next generation.) The planned U.S. release has long faced local opposition.



The interior of a tree burns in California's Big Basin Redwoods State Park.

PHOTO: MAX WHITTAKER/*THE NEW YORK TIMES*

CONSERVATION

Climate change blamed for California wildfires

Two wildfires that are among California's largest ever have wreaked havoc since last week on people and nature. A record 11,000 lightning strikes in 72 hours helped ignite dry vegetation. At least seven people had died as of 25 August, and air pollution reached health-threatening levels in places. The University of California, Santa Cruz, and its surrounding community were evacuated as one fire threatened. In Big Basin Redwoods State Park, flames damaged the massive trees; scientists expected that the thick bark of the fire-adapted species would protect older ones, but worried about younger redwoods growing elsewhere in denser plots, where fires burned hotter. Other researchers anxiously monitored hatchlings in nests of endangered California condors. Climate scientists suspect drier conditions created by global warming are driving the blazes; a 2019 study found an eightfold increase in the annual area burned by summer fires in California from 1972 to 2018.

Published by AAAS



IN DEPTH

After echoing Trump's claims about convalescent plasma, the Food and Drug Administration's Stephen Hahn (left) backtracked.

COVID-19

In plasma OK, critics see politics, not science

FDA's emergency move raises fears of rushed approval for other experimental medicines

By **Kai Kupferschmidt** *and* **Jon Cohen**

It was, U.S. President Donald Trump said, "a very historic breakthrough" in the fight against COVID-19 that "would save countless lives." To critics, the emergency use authorization (EUA) that the U.S. Food and Drug Administration (FDA) issued last Sunday for the use of convalescent plasma to treat people with severe COVID-19 looked more like a breakdown, with the evidence overstated and potentially worrisome consequences for the evaluation of other experimental medicines.

Flanked at a Sunday evening press conference by FDA Commissioner Stephen Hahn, Trump claimed the antibody-rich plasma, which comes from people who have recovered from the disease, was "safe and very effective," and had been "proven to reduce mortality by 35%"—what he called "a tremendous number." But that number has received a tremendous amount of scrutiny. It surprised even the researchers who conducted the study on which Trump apparently based the declaration. "I don't know where the 35% number comes from," says Arturo Casadevall of Johns Hopkins University's Bloomberg School of Public Health, an author of the study, which has been posted as a preprint but has yet to be peer reviewed. Researchers' best guess was that it reflected cherry picking of data from a small subset of the study—and that the EUA, issued on the eve of the Republican National Convention, represented as much political theater as science.

Randomized, controlled trials (RCTs), the gold standard for assessing experimental treatments, haven't yet shown that COVID-19 patients benefit from convalescent plasma. One such study, which includes thousands of patients, is underway as part of the Recovery trial in the United Kingdom. The EUA rests heavily on data from the COVID-19 Plasma Consortium, funded by the Department of Health and Human Services to provide access to the treatment and assess its safety. The group has treated more than 90,000 patients at nearly 3000 sites. That study has no untreated control group, however, so it cannot determine whether convalescent plasma has any risks that outweigh benefits.

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

The treatment was allowed under FDA's Expanded Access program, which makes experimental treatments available to patients. Casadevall says this program requires burdensome, time-consuming paperwork that the EUA will do away with. That will be helpful for understaffed hospitals that treat "underserved populations," he says. "I think the FDA made the right call. The political noise is unfortunate."

Others say the political noise is drowning out the science.

Hahn claimed at the press conference that Trump's 35% figure translates to 35 lives saved per 100 sick people. Casadevall and others suspect that number is based on a small part of the study. One possibility is an analysis involving about 3000 patients that assesses the impact of different COVID-19 antibody levels in the plasma. In this group, 8.9% of those who received plasma containing high antibody levels died within 7 days, versus 13.7% of those who received plasma with low levels. Although that is a 35% relative difference, the absolute difference, 4.8%, amounts to 4.8 lives saved per 100 sick people—not 35. And after 30 days, the relative difference had dropped to 25%. At that point, the mortality rates were 22.3% and 29.6%, respectively, among recipients of plasma with high and low levels of antibodies—hardly a historic breakthrough.

The White House claim is "statistical acrobatics," says Luciana Borio, a former chief scientist at FDA who also worked on the Trump White House's National Security Council. On Monday evening, Hahn acknowledged he had made a mistake, tweeting that the criticism was "entirely justified. What I should have said better is that the

PHOTO: PETE MAROVICH/GETTY IMAGES

Published by AAAS

data show a relative risk reduction not an absolute risk reduction."

Nicole Bouvier, a virologist and infectious disease clinician at the Icahn School of Medicine at Mount Sinai—a participant in the Mayo Clinic–led consortium—says the data gathered so far do suggest convalescent plasma helps some COVID-19 patients. "There is probably a benefit to it, but it's probably not a major breakthrough," Bouvier says. She wishes the U.S. government had organized large-scale RCTs for the intervention. "We're trying to build a wall of evidence and we're putting in pebbles," she says.

Borio and others worry the EUA may make it harder to gather that kind of evidence for other treatments. Convalescent plasma contains a mixture of different antibodies, only some of which hinder the virus, and Borio and many others have high hopes for a more targeted and standardized intervention: "monoclonal" preparations that contain high levels of the most potent antibodies only. But now that the EUA has put convalescent plasma within reach of more patients, it may become harder to enroll people in RCTs for monoclonal antibodies, Borio says.

Myron Cohen of the University of North Carolina, Chapel Hill, who oversees large-scale RCTs of monoclonals now taking place under the aegis of the U.S. National Institutes of Health, agrees. "If potential study participants for a COVID-19 treatment … believe there is a beneficial and safe agent they can receive without randomization to a placebo, they may logically and often choose that path," Cohen says.

Given that tens of thousands of Americans have already received convalescent plasma through FDA's Expanded Access program, it's unclear how many more people the EUA will actually benefit. Bouvier says her hospital may not even be able to use the EUA, because it requires that convalescent plasma be first characterized with a specific antibody test that it does not have.

Eric Topol, a cardiologist who directs the Scripps Research Translational Institute, says the EUA "again represents the FDA caving directly to Trump pressure," as he believes it did when it issued an EUA (later rescinded) for hydroxychloroquine treatment for COVID-19. "It sadly and unacceptably exemplifies loss of independent FDA assessment of evidence and data overridden by political pressure," Topol says.

Many scientists worry the same could happen in future decisions about EUAs for COVID-19 vaccines—with far greater potential consequences because vaccines presumably will be given to hundreds of millions of healthy people. ■

COVID-19

Critics slam NIH for demands on bat grant targeted by Trump

Nobelists, former agency leaders deplore politicization

By **Meredith Wadman**

Prominent scientists are condemning the U.S. National Institutes of Health (NIH) for imposing unprecedented conditions before it will revive a grant to a small nonprofit research institute. The agency killed the grant in April, days after U.S. President Donald Trump promised to "end [it] very quickly." The grant, which funded studies of how bat coronaviruses spread to humans, included work done by researchers in Wuhan, China, where the pandemic coronavirus was first identified. Trump has alleged, without evidence, that the virus escaped from their laboratory (*Science*, 8 May, p. 561).

In an 8 July letter, Michael Lauer, NIH's deputy director for extramural research, detailed the unusual steps the EcoHealth Alliance must take to restore funding for its grant. (The letter was first reported on 19 August by *The Wall Street Journal*.) They include explaining an alleged decrease in cellphone traffic and possible roadblocks near the Wuhan Institute of Virology in October 2019, compelling WIV to submit to an inspection of its facilities and records by the U.S. National Academy of Sciences or the National Institutes of Allergy and Infectious Diseases focusing on whether WIV staff had the pandemic coronavirus in their possession prior to December 2019, providing a vial of the virus that WIV scientists used to sequence the pandemic coronavirus in January, explaining the "apparent disappearance" of a technician from WIV, and providing WIV's responses to 2018 U.S. State Department cables that discussed "safety concerns" at the lab. "We have concerns that WIV has not satisfied safety requirements," Lauer wrote.

> **"This is … an unconscionable action."**
> **Gerald Keusch**, Boston University

Many of the conditions are "impossible" to fulfill, the EcoHealth Alliance said, and "will effectively block us from continuing this critical work." And outside researchers are angered by what they see as NIH's complicity in political meddling. "This is … an unconscionable action" and "a threat to the entire scientific community," says Gerald Keusch, an emerging infectious disease expert at Boston University who was previously the head of NIH's Fogarty International Center.

"If I was a scientist wanting to pursue the agenda that NIH laid out, I would join the CIA," says Richard Roberts, chief scientific officer at New England Biolabs, a research supply company. "It is totally outrageous when political considerations steamroller over science." Roberts organized a letter that he and 76 other Nobel laureates sent to NIH Director Francis Collins in May, protesting the grant's cancellation. "Collins should have resigned over the original cancellation," Roberts says, "but he appears totally spineless."

NIH declined requests to interview Collins and Lauer, saying the agency "does not discuss internal deliberations on specific grants." But Richard Ebright, a molecular biologist at Rutgers University, New Brunswick, says NIH is within its rights to impose most or all of the conditions. He has criticized work proposed by the grant, asserting it could produce dangerous new viruses. "It is appropriate for the NIH to enforce biosafety standards, biosafety-reporting, and data-availability conditions," Ebright says.

NIH first funded the EcoHealth Alliance grant in 2014 and renewed it in 2019 after it scored in the top 3% of grants reviewed. During the first 5 years, the EcoHealth Alliance sent $599,000 from the $3.1 million grant to a lab led by Shi Zhengli, a WIV virologist, to identify bat coronaviruses at high risk of jumping to humans. The grant "addresses one of the most important problems in the world right now," former NIH Director Harold Varmus, who signed the Nobel laureates' letter, told *The Wall Street Journal*. He called NIH's conditions "outrageous."

In the grant's next term, through 2024, scientists were planning human, wildlife, and lab-based studies to pinpoint places in southern China where bat coronaviruses are at high risk to jump to humans. The EcoHealth Alliance says none of the money awarded in 2019 and 2020 was sent to WIV. The group says that, in response to NIH's actions, it has received $751,000 in unsolicited donations. ■

Published by AAAS

VOICES OF THE PANDEMIC

How Italy's 'father of the swabs' fought the virus

Microbiologist Andrea Crisanti defied the authorities to launch his testing campaign

By **Douglas Starr**

Andrea Crisanti was on a 30-hour flight from Italy to Australia for a conference on 22 February when some disturbing news appeared on his phone. Italy had just had its first COVID-19 death, and more cases were accumulating fast. He asked conference organizers to move his talk to the first day, and made the grueling trip back home after that. "It's something I do not recommend," he says.

Crisanti, head of the microbiology department at the University of Padua, already knew trouble was coming and had geared up his lab to do large-scale testing for the new coronavirus. As it began to devastate his nation, Crisanti put his university and region at the forefront of the fight with an all-out campaign of testing and quarantine, even when that meant defying conventional wisdom.

A soft-spoken 65-year-old with graying hair and soft brown eyes, Crisanti has a matter-of-fact way of stating his opinions—even when he opines that something is "bullshit." "He's an innovative person who knows his own worth and has confidence in his judgments," says Jules Hoffmann, a Nobel Prize winner and professor of integrative biology at the University of Strasbourg. His decisiveness helped rein in his region's outbreak and show the rest of Italy how to tame the virus, which hit the country early and hard.

Crisanti, who trained in immunology and biotechnology in Rome before spending 25 years at Imperial College London, was used to fighting another scourge: malaria. Last fall, the University of Padua recruited him to continue his research on genetic strategies to block mosquito reproduction. But when news about the coronavirus began to emerge from China, Crisanti immediately shifted his focus.

In late January, when Chinese scientists published the genetic sequence of the new coronavirus, Crisanti began to test university students returning from China, symptomatic or not. He had conducted a few hundred tests when the regional health department told him to stop. Guidelines from the World Health Organization and Italy's National Institute of Health said to test only patients with symptoms, he was told. Crisanti says the restriction made no sense: "I know very few infectious diseases where asymptomatic people do not play a major role."

That's where things stood when he got word of the first Italian COVID-19 fatality. The patient was from Vo, a prosperous village in the region of Veneto, about 50 kilometers west of Venice. The region's governor ordered a 2-week quarantine of the town and testing of almost all 3300 residents. Anyone who tested positive was put on lockdown.



"Lock down the village, test everybody, and isolate the positives. It really works."

Andrea Crisanti, University of Padua

At the time, anecdotal reports were emerging from China about asymptomatic transmission, but no one had produced definitive evidence. Crisanti saw Vo as an ideal place to conduct an epidemiological experiment: a small population, universally tested, whose progress could be monitored closely. He got approval to retest everyone in the village 9 days after the first round of testing.

The numbers confirmed his thinking about asymptomatic transmission. In the first round of testing, 73 residents were positive for the virus. More than 40% of them had no symptoms yet had levels of the virus similar to those who were visibly ill. The Vo study also confirmed that isolating people helps stem transmission. Everyone who had tested positive was confined to their home, regardless of whether they had symptoms. By the second round of testing, a week and a half later, the number of positives had dropped to 29; they, too, were isolated. A third round of tests 2 months after the second found no positive cases.

"If you want to eliminate a cluster you have to lock down the village [or neighborhood], test everybody, and isolate the positives," Crisanti says. "It really works."

Crisanti persuaded the regional government of Veneto to test anyone with even the mildest of symptoms, and to trace and test their contacts as well. The effort targeted medical personnel and essential workers, such as supermarket cashiers. It helped that Veneto has a long tradition of taking strong public health measures, dating back to the invention of the quarantine during the 14th century plagues. (The word "quarantine" is derived from the word for "40 days" in an old Venetian dialect—the period for which incoming ships had to anchor in the harbor to avoid bringing in plague.) The region's infrastructure was ready for a pandemic, with a health care policy that emphasizes decentralized primary care. In this case, that meant sending well-equipped nurses to test people at home or admitting them to small local hospitals with dedicated COVID-19 units.

In contrast, neighboring Lombardy, the prosperous region in which Milan is located, has emphasized large, urban hospitals offering first-rate surgical and

ILLUSTRATION: KATTY HUERTAS

Published by AAAS

specialty care. That system backfired in the pandemic, funneling sick people into the hospitals, which in turn became sources of infection. Lombardy became the worst affected region of Italy, with 2.5 times the number of cases and four times the number of deaths per capita as Veneto.

From the beginning, Crisanti was prescient. In late January he ordered enough reagent to process half a million swabs; then had his lab analyze the reagents and begin to produce its own. Thus, when other regions were running short, Veneto had a surplus of reagents. Later he ordered a piece of equipment that could process tests at high speed, tracking down a demo machine in London when he couldn't procure one through the usual means because of heightened demand from the pandemic. "We got the only one in Italy," he says. The machine quadrupled his laboratory's throughput to more than 6000 swabs per day. Along the way Veneto became an example of the value of extensive testing, tracing, and isolation—and ensuring the means to do it.

Newspapers hailed Crisanti as the "father of the swabs," and the "rebel scientist," for his defiance of official policy in the early days of the pandemic. He received the Lion of Veneto award for his service to the region, the seal of the city of Padua, and was honored by a special concert in Vo. Yet it hasn't all been smooth. As the outbreak began to abate, the region's governor, Luca Zaia, downplayed Crisanti's contribution in comments to the press and claimed that he and his government deserved credit for taming the virus. Eager to reopen Veneto for tourism, Zaia became irritated by Crisanti's insistence to go slow and turned to other scientists for advice. The freeze-out became so severe that in July, Crisanti said he would resign from the region's advisory board, only to be talked out of it by colleagues and admirers.

Now, there's a truce between the scientist and the politician. "It may have been a joint effort," says Antonio Cassone, professor emeritus of medical microbiology at the University of Perugia. "But Andrea proved essential."

Moving forward, Crisanti is analyzing the genetic and blood samples his team collected during the Veneto outbreak to learn more about individual susceptibility and antibody response. He remains undaunted by his encounter with politics. "The most important thing is to convey simple, clear, and honest messages," he says. "And if you don't know something just say it openly. People need to know the truth." ■

Douglas Starr is a journalist in the Boston area.



A woman spits into a tube so that her saliva can be tested for the presence of the novel coronavirus.

PHOTO: UNIVERSITY OF ILLINOIS, URBANA-CHAMPAIGN

COVID-19

Spit shines for easier coronavirus testing

Tests using saliva are cheaper and faster than those with nasal swabs—and can be just as accurate

By **Robert F. Service**

First, a technician pushes a pencil-length swab to the very back of your nasal passages. Then you pay $100 or more, and wait days for an answer. But faster, cheaper, more pleasant ways to test for the novel coronavirus are coming online. This month, the U.S. Food and Drug Administration granted emergency use authorization for two tests that sample saliva instead of nasal fluid, and more innovations are likely after FDA relaxed rules to allow new tests to be adopted more quickly. One candidate was announced last week: an experimental test, potentially faster and cheaper, that analyzes saliva in a new way.

"There is real promise here," says Anne Wyllie, a microbiologist at Yale University who helped develop one of the new tests authorized this month. Takanori Teshima, chief of laboratory medicine at Hokkaido University, who also reported successful results testing saliva, agrees. Saliva tests "will have a big impact worldwide."

When SARS-CoV-2, the respiratory virus that causes COVID-19, emerged in December 2019, researchers scrambled to develop tests to detect the virus. Initially, they turned to a long-trusted technique for diagnosing respiratory infections: looking for viral genetic material in mucosal fluid, thought to be the best hunting ground for a respiratory virus, collected from deep in a patient's nasal passages. That's where the 15-centimeter swab comes in. The swab goes into a plastic tube with a chemical mixture that stabilizes the virus during transport to a diagnostics lab. There, technicians extract its genetic material and load it into a machine to carry out the polymerase chain reaction (PCR), which amplifies snippets of genetic material unique to the virus.

The procedure accurately identifies infections about 95% of the time. But the test is uncomfortable and, because collecting the swab requires close contact with patients, it puts medical personnel at risk of contracting the virus. "Nobody wants to do that job," Teshima says.

Testing saliva for SARS-CoV-2 was no sure thing. Studies with other respiratory diseases showed saliva tests identified only about 90% of people for whom swab tests indicated an infection. But the appeal of an easier and safer test for the new coronavirus led researchers to try. People being tested simply drool into a bar-coded plastic tube, seal it, and drop it in a pouch that's shipped to a lab for PCR analysis. Because the procedure directly tests the fluid responsible for transmitting the virus between people, it may give a better indication of who is most contagious, says Paul Hergenrother, a chemist at the University of

Published by AAAS

Illinois, Urbana-Champaign (UIUC), who led his university's saliva test development.

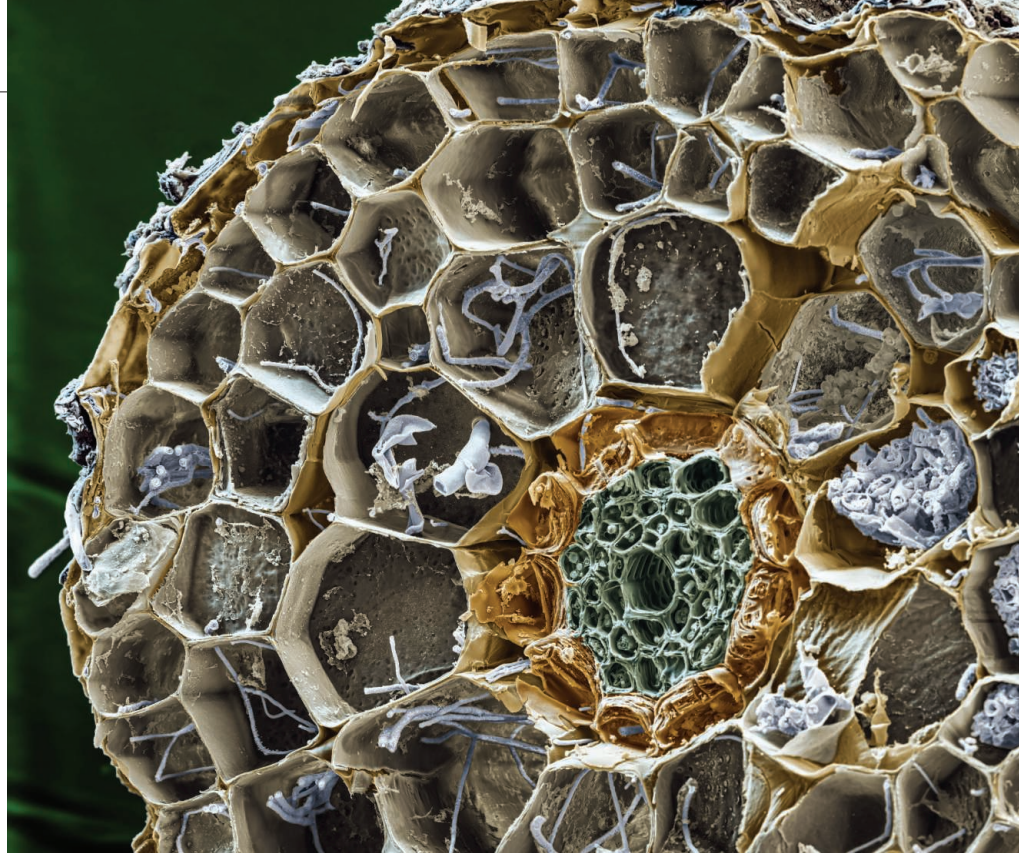
As early as 12 February, researchers in Hong Kong and China reported in *Clinical Infectious Diseases* that they could identify SARS-CoV-2 from saliva in 11 of 12 patients whose swabs showed virus. Since then, groups in the United States, Singapore, and Japan have confirmed and further simplified the procedures, cutting out costly steps such as adding specialized reagents to stabilize the virus and extract the genetic material.

In May, Wyllie and Yale colleagues teamed up with the National Basketball Association, which provided $500,000 to develop Yale's saliva diagnostic; the test is now used for frequently testing players. On 4 August, the Yale team posted a preprint on medRxiv that said its saliva test agreed with swab results 94% of the time, at a cost of as little as $1.29 per sample, roughly 1/100 as much as commercial swab-based tests. On 15 August, FDA granted emergency approval for the test, called SalivaDirect, so that other FDA-approved labs can use the protocol. Last week, the agency extended approval to the UIUC test given its similarity to the Yale test. UIUC is now using its saliva test with all 60,000 students, faculty, and staff twice a week, in order to isolate infected individuals as quickly as possible. "Testing saliva makes sense scientifically, and it makes sense logistically," Hergenrother says.

A new saliva test for RNA viruses, such as Zika and SARS-CoV-2, was reported last week in *Science Advances* by researchers at the University at Albany. It could be even faster and cheaper because it does not need expensive lab equipment such as PCR machines. Rather than amplifying RNA to identify the virus, the approach uses snippets of DNA that bind to short, unique sections of RNA and change them from linear strands to loops. That alters how the RNA behaves in a common lab procedure known as gel electrophoresis, making it easy to detect. "This is innovative," Wyllie says.

A relaxation of FDA rules announced last week could lead to still more variants. The new rules allow approved clinical labs to use tests they have developed without any additional approval step. In a tweet, Michael Mina, an epidemiologist at Harvard University's T.H. Chan School of Public Health, called FDA's decision "Huge news!!" because it would encourage labs to develop novel tests. It may also help speed development of rapid tests that look for viral proteins rather than genetic material—an efficient way to screen large numbers of asymptomatic people.

"We don't need one test to be the end all and be all," Wyllie says. "We just want options." ■



In an electron micrograph, mycorrhizal fungi (bright gray) penetrate root cells, where the fungi supply water and nutrients in exchange for carbon.

IMAGE: EYE OF SCIENCE/SCIENCE SOURCE

ECOLOGY

Hidden web of fungi could shape the future of forests

Plants' fungal partners can help their hosts resist drought and heat—or make them more vulnerable

By **Elizabeth Pennisi** *and* **Warren Cornwall**

The future of the world's flora may depend as much, if not more, on what's below the ground as what's above. Beneath 90% of all plants lies an invisible support system—subterranean fungal partners that form a network of filaments connecting plants and bringing nutrients and water to their roots. In return, the plants provide a steady supply of carbon to the fungi. Now, researchers are learning that these hidden partners can shape how ecosystems respond to climate change.

The right fungal partners can help plants survive warmer and drier conditions, according to a study reported earlier this month at the online annual meeting of the Ecological Society of America. But other studies at the meeting showed climate change can also disrupt these so-called mycorrhizal fungi, possibly speeding the demise of their host plants. "The picture is becoming clearer that we really cannot ignore the responses of mycorrhizal fungi to climate change," says Matthias Rillig, an ecologist at the Free University of Berlin.

These fungal associates come in two forms. Arbuscular mycorrhizae (AM), common in tropical and some temperate forests as well as fields and meadows, invade root cells and extend thin hairs called hyphae into the soil. Ectomycorrhizal (EM) fungi, in contrast, associate with conifers as well as oak, hickory, alder, and beech. They settle on the outside of roots, and their networks of hyphae give rise to the mushrooms that pop up on moist forest floors.

Both types absorb phosphorus and other nutrients, capture nitrogen from decaying organic matter, and help store carbon in the soil. "Mycorrhizal associations are arguably the most important symbioses in terrestrial ecosystems because of their importance for plant productivity," says

Published by AAAS

Christopher Fernandez, a soil ecologist at the University of Minnesota, Twin Cities.

Climate change might alter these associations, he says. Fernandez is part of the B4WARMED (Boreal Forest Warming at an Ecotone in Danger) project, a large-scale effort to monitor the impact of warming and drying on the boreal forests that stretch across northern latitudes. The study is also artificially warming and drying plots of forest, and Fernandez studied the impact of the simulated climate change on the hidden fungi by sequencing soil and root samples from plots.

As conditions became warmer and drier, the diversity of the EM fungi fell, and "weedy" EM fungi took over, Fernandez reported at the meeting. These "weeds" don't devote a lot of energy to building extensive underground networks, causing their connectivity to break down. If the same disruption happens as climate change unfolds, fewer seedlings might establish their critical partnerships with fungi, which could deprive the trees of nutrients.

Monitoring conducted by B4WARMED has already shown that the warmer, drier climate of recent years is taking a toll on the boreal forest. What role any change in the EM fungi is playing is not yet clear, but "the alteration of mycorrhizal fungal communities in response to climate change is deeply concerning," Fernandez says.

The B4WARMED results "show that there may be some major shifts in both above- and belowground communities in the future," says Sarah Sapsford, an ecologist at the University of Canterbury. "What we see now we may not see again."

A different ecosystem, the pinyon pine forest of the U.S. Southwest, shows that existing variations in mycorrhizae can also influence trees' resilience. Decades ago, a moth infestation stunted some of the pinyons (*Pinus edulis*) growing near Arizona's Sunset Crater Volcano National Monument. The hard-hit trees have different EM fungi from their taller neighbors, suggesting the two types of trees may differ genetically.

When a megadrought struck the region in 2002 and 2003, twice as many of the taller trees died. To see whether the mycorrhizae made a difference, Catherine Gehring, an ecologist at Northern Arizona University, Flagstaff, and her colleagues grew seedlings from the two types of plants with and without their fungi in a greenhouse, under different watering regimens.

"We found that ectomycorrhizal fungi played a critical role in drought tolerance," Gehring reported at the meeting. Sanna Sevanto, a biophysicist at Los Alamos National Laboratory, watched the fungi in action by dousing seedlings' roots in heavy water, which served as a tracer. Water moved into the drought-tolerant roots infected with their fungus much faster than when they were sterile, Gehring reported.

She and her colleagues are working with local researchers to replant pines in the Navajo Nation, not far from Sunset Crater, and they are applying what they learned. Because genetic differences among the trees seem to determine which of the two groups of EM fungi colonizes them, the team will be careful to plant seedlings with the right genotype to attract the drought-resistant fungi. That "could mean the difference between life and death during drought," Gehring says.

A third study suggests mycorrhizae could shape how not just trees but entire ecosystems respond to environmental change. Because AM and EM fungi are associated with different tree species, Colin Averill, an ecologist at ETH Zurich, wondered whether the fungi themselves help determine which kind of forest grows in a particular region.

Earlier researchers had suspected they do, but Averill and his colleagues searched for evidence in a vast set of U.S. Forest Service data tracking the species, growth, and death of every tree bigger than a sapling in 6965 forest plots across the eastern United States. They found that many plots are dominated either by AM-related trees or by EM-related trees; a mix of the two is rarer.

A statistical analysis found fungi were key to this stark split, and the researchers found a likely reason: The dominant mycorrhizae can help lock a forest into a stable state. Tree measurements taken at 5-year intervals at each plot showed an AM tree was at least 10 times more likely to take root in an AM forest than in an EM forest and twice as likely to survive. Meanwhile, EM trees were more likely to thrive in EM forests.

The fungi may enforce this monopoly by altering the soil in ways that favor specific species—for example by controlling nitrogen levels. A more established fungal network could also help young trees endure heavy shade that interferes with photosynthesis, or older trees withstand drought or diseases. "If you're an EM tree you can plug into an EM network, and that can help you survive," Averill says.

The grip of those fungi could slow forests' response to external pressures such as climate change. AM-affiliated trees tend to fare better in hotter temperatures, for example, but they might be slower than expected to colonize an EM-dominated forest as the climate warms, Averill says. "These types of dynamics might become really important as we try to forecast how the global forest system will change in the future." ■

CLIMATE CHANGE

New feedbacks speed up the demise of Arctic sea ice

Wave action and a growing underwater blob of heat add to impact of warming air

By **Paul Voosen**

In March, soon after arriving aboard the *Polarstern*, a German icebreaker frozen into Arctic sea ice, Jennifer Hutchings watched as ice broke up around the ship, weeks earlier than expected. Even as scientists on the research cruise scrambled to keep field instruments from plunging into the ocean, Hutchings, who studies ice deformation at Oregon State University, Corvallis, couldn't suppress a thrill at seeing the crack up, as if she had spotted a rare bird. "I got to observe firsthand what I studied," she says.

Arctic sea ice is itself an endangered species. Next month its extent will reach its annual minimum, which is poised to be among the lowest on record. The trend is clear: Summer ice covers half the area it did in the 1980s, and because it is thinner, its volume is down 75%. With the Arctic warming three times faster than the global average, most scientists grimly acknowledge the inevitability of ice-free summers, perhaps as soon as 2035. "It's definitely a when, not an if," says Alek Petty, a polar scientist at NASA's Goddard Space Flight Center.

Now, he and others are learning that a warming atmosphere is far from the only factor speeding up the ice loss. Strengthening currents and waves are pulverizing the ice. And a study published last week suggests deep heat in the Arctic Ocean has risen and is now melting the ice from below.

Ice has kept its grip on the Arctic with the help of an unusual temperature inversion in the underlying waters. Unlike the Atlantic or Pacific oceans, the Arctic gets warmer as it gets deeper. Bitter winters and chilly, buoyant freshwater from Eurasian rivers cool its surface layers, which helps preserve the underside of the ice. But at greater depths sits a warm blob of salty Atlantic water, thought to be safely separated from the sea ice.

As the reflective ice melts, however, it is

Published by AAAS



The *Polarstern*, released too early from a floe, returned to the North Pole in August amid thin ice.

replaced by darker water, which absorbs more of the Sun's energy and warms. Those warming surface waters are likely migrating down into the blob, which robotic temperature probes, moorings, and oceanographic surveys show is steadily warming and growing. With enough heat to melt the Arctic's ice three to four times over, the blob could devour the ice from below if the barrier of the cold surface layers ever dissipates.

Measurements from the eastern Arctic Ocean, published last week in the *Journal of Climate*, show the blob, usually found 150 meters below or deeper, has recently moved up to within 80 meters of the surface. Increased turbulence means some of that heat is now melting ice, says Igor Polyakov, an oceanographer at the University of Alaska, Fairbanks. "This heat has become, regionally, the key forcing for sea ice decay."

The process, called "Atlantification," is already well underway in the Barents Sea, north of Norway, where fingers of warm Atlantic water have spread north and risen, melting sea ice even in winter months. The invasion shows no sign of stopping, says Helene Asbjørnsen, an oceanographer at the University of Bergen who has helped chart this migration. "Ultimately we expect it to extend into the Arctic more."

The $134 million Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAiC), based on the *Polarstern*, is exploring another ice-destroying feedback. The ship froze itself into a floe in October 2019, to give the team a chance to observe the floe for one full year as the summer melt season shifted back into freezing. But the project ran into challenges. First came the COVID-19 pandemic, which made planned personnel rotations difficult. Then the ice drifted too far south too quickly. In late July, the day after the team pulled up its remaining instruments, the floe broke up and melted. "To me that is a big loss, and I'm pretty bummed about it," says Matthew Shupe, a climate scientist at the University of Colorado, Boulder, who helped lead U.S. contributions to the cruise. But, he added, there was a bonus: "We never planned to be around for that 'death of an ice floe' process."

The *Polarstern*'s floe is not an isolated case. Remote sensing satellites show that over the past 20 years, ice has been drifting faster, potentially sweeping it into warmer waters, says Sinéad Farrell, a sea ice scientist at the University of Maryland, College Park. One reason for the change in pace could be faster currents in the Arctic Ocean, as ice melt exposes more water to the push of the wind, says Arild Sundfjord, a physical oceanographer at the Norwegian Polar Institute. "We think we see signs of that."

Another factor could be an increase in the roughness of the sea ice, which allows wind to catch and propel it. MOSAiC scientists deployed GPS stations across the floe's melange of first-year and thicker multiyear ice to monitor its speed and deformation. They suspect that as the ice becomes thinner and weaker, it is more prone to the crunch and crumble that builds up wind-catching ridges, Hutchings says, but they're still resolving whether that is true. The turmoil took a heavy toll on the expedition, crushing some instruments like aluminum cans and destroying snow sampling sites. It was frustrating, Shupe says. "We don't really control anything here," he says. "The Arctic is telling us its story and we just need to be clever enough to document it."

ICESat-2, a laser altimeter launched by NASA in 2018, will help extrapolate findings from MOSAiC to the rest of the Arctic. Unlike previous satellites, ICESat-2 can distinguish between ice floe cracks and melt ponds on top, and it is already showing stark differences between multiyear and first-year ice, Farrell says. In a surprise, the ICESat-2 team is finding that the multiyear ice overall is twice as rough as first-year ice. "It's kind of like aging skin," she says. "They get more wrinkly over time." The satellite also seems to be capable of capturing waves amid the ice, and linking them to nearby storms, Petty says. It's another worrying mechanism that could speed up ice loss, he says. "As waves break the ice apart, it gets more exposed to heat—and melts further."

The retreat of the ice bodes ill for global climate, but it is making the Arctic easier to study. This month saw the start of the Synoptic Arctic Survey, which will knit together more than a dozen national Arctic cruises by ice breakers and other research ships. The survey will cover the Arctic's entirety, providing a near-simultaneous picture of currents, life, and water conditions and chemistry, rather than a collection of regional snapshots over time. The pandemic delayed all but two of the cruises, which were planned for this summer: those of Japan's *Mirai* and South Korea's *Aron*. But once completed, the survey could answer basic questions, such as whether the Arctic is a net source or sink of carbon dioxide.

And it could not have been done in the ice-bound Arctic of old. "Now," Sundfjord says, "we can go wherever, and whenever, we want." ■

Going, going ...

Summer Arctic sea ice covers half the area it did in the 1980s, and it could disappear by 2035. The ice faces threats not only from warming air, but also from waves, currents, and melting from below.



CREDITS: (PHOTO) STEFFEN GRAUPNER; (MAP) NATIONAL SNOW AND ICE DATA CENTER, UNIVERSITY OF COLORADO, BOULDER. ADAPTED BY N. DESAI/*SCIENCE*

Published by AAAS

BIOMEDICINE

How 'elite controllers' tame HIV without drugs

Rare people who naturally subdue the virus may weed out cells with active HIV genomes

By **Jon Cohen**

A tiny fraction of the 38 million HIV-infected people in the world have what seems like a superpower. Without the help of antiretroviral (ARV) drugs, they keep the AIDS virus at undetectable levels in their blood, sometimes for many years, even though they still have HIV genes woven into their chromosomes. Now, the most in-depth genomic analysis of these rare individuals, who account for less than 0.5% of all HIV infections, reveals a clue to their success, which scientists hope will ultimately lead to new strategies to corral the virus in others.

All but one of the 64 "elite controllers" analyzed in the new research have abundant and intact HIV genomes, so-called proviruses, integrated in their cells. But compared with people who have to take ARVs, elite controllers have far more proviruses in chromosome regions where little gene activity occurs, the study finds. Somehow these people have eliminated infected cells with proviruses parked in areas where they more readily co-opt the cellular machinery needed to copy themselves. As a result, elite controllers' immune systems can handle for prolonged periods—decades in some cases—the low levels of HIV they presumably do produce.

The "comprehensive and elegant study" provides "nature's proof of principle of how a functional cure [for HIV] is possible," says Beatrice Hahn, a virologist at the University of Pennsylvania. The challenge now is to identify how to translate this elite controller trick to the far larger HIV-infected population, but the new study "gives us a blueprint," says Mathias Lichterfeld, an infectious disease physician at Brigham and Women's Hospital and a co-author of the work, published in *Nature* this week.

When people on ARVs stop the drugs, HIV typically floods their blood within weeks, as "reservoirs" of proviruses in various cells and tissues spring into action. As a result, HIV cure researchers long have attempted to reduce the size of these reservoirs. But the *Nature* study, led by immunologist Xu Yu of the Ragon Institute of MGH, MIT and Harvard, suggests the size of a reservoir may be less important than limiting proviruses to quiescent chromosome homes.

HIV typically splices its provirus into a cell's genes, which make up only 1% of the human genome. That helps the virus co-opt the gene transcription machinery needed to churn out new HIVs. But when Yu and colleagues pinpointed the location of nearly 4000 proviruses in cells from the elite controllers and 42 people on ARVs, they found a striking difference. In elite controllers, 45% of the proviruses reside in "gene deserts," regions of chromosomes where little transcription occurs. In people on ARVs, the number was 17%.

Studies probing how elite controllers manage HIV often are more intriguing than convincing because they rarely include more than a handful from that group. The fact that dozens were examined in the new work impresses Hahn. "Whoever decided to make a concerted effort to compile this cohort is a genius," she says.



At a 2019 HIV cure meeting, "elite controller" Loreen Willenberg (right) speaks with Rowena Johnston, research director of amfAR.

PHOTO: AMFAR/THE FOUNDATION FOR AIDS RESEARCH

One member of the cohort presents an extreme case: what amounts to a natural cure. The woman, Loreen Willenberg, who was diagnosed as HIV positive in 1992 yet remains healthy without ARVs, had the lowest levels of proviruses of anyone in the study. What's more, the only proviruses found among the billions of her cells studied were incapable of producing new HIVs. "We did not do integration site analysis on [her] samples, because there's no intact provirus, and even the defective proviruses were so rare in her genome," Yu says.

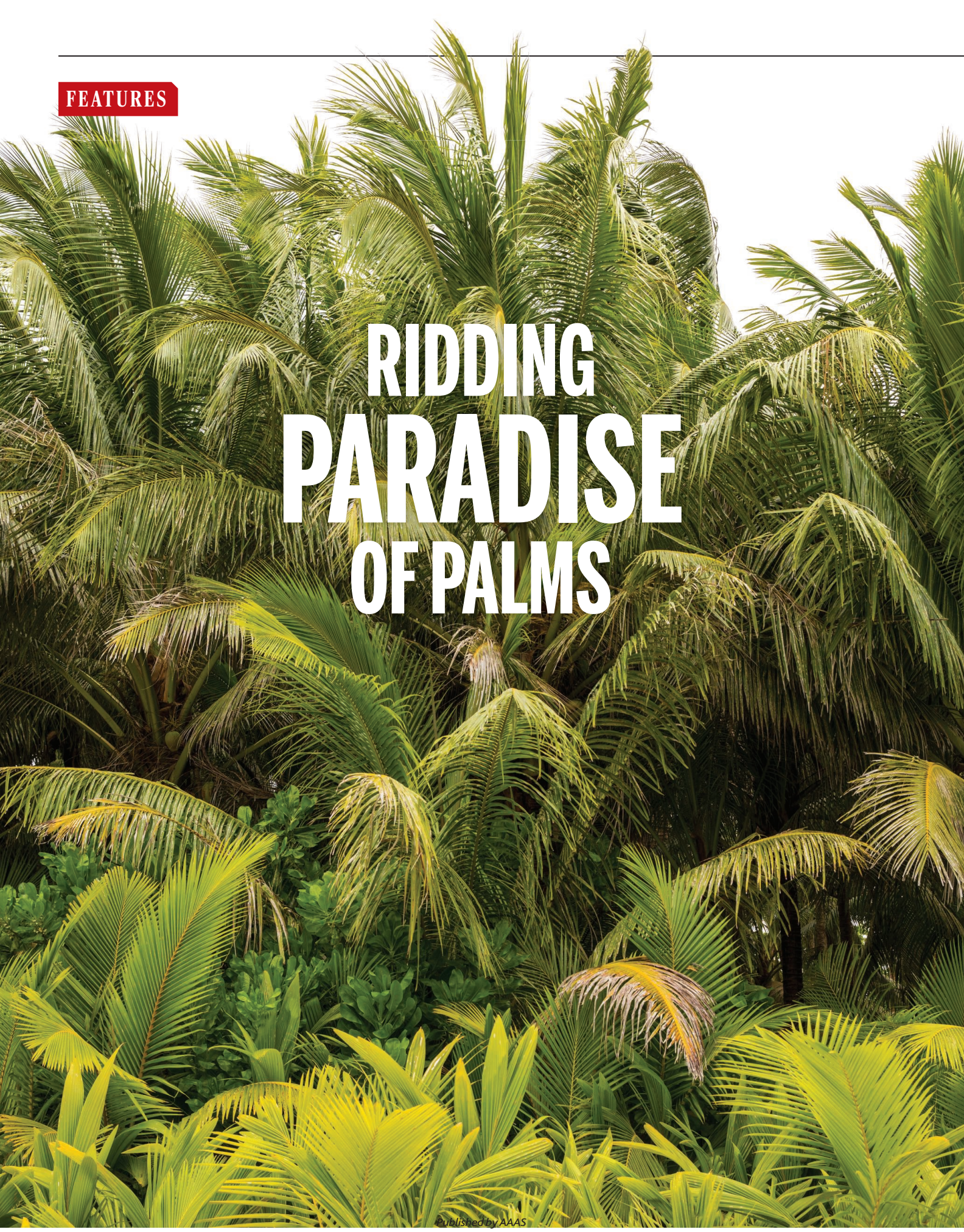
Willenberg has a seemingly unharmed immune system despite her untreated HIV infection. She has had her blood levels of HIV monitored for 24 years and has only had one detectable burst of viral replication, when she had a bad bout of the flu. Intensively studied by many HIV scientists, Willenberg nearly 15 years ago established a foundation that helps elite controllers, some of whom joined this clinical study. "If there were some way that this unique state of natural control of HIV infection could be somehow understood and then translated to individuals who did not spontaneously control HIV infection, then that's what I had to do," she says.

Previous studies of Willenberg and other elite controllers suggest they frequently have a genetic predisposition to mount stronger immune responses against the virus. Yu and her co-authors suspect this response selectively targeted cells with the most fit proviruses, tilting controllers' reservoirs toward their hamstrung relatives. A better understanding of this evolutionary immune process might be key to interventions that, in effect, create more elite controllers.

HIV cure studies to date have largely attempted to prod latent proviruses into making new viruses, theoretically setting up those cells for elimination and steadily reducing the size of the reservoir. But this "shock and kill" strategy has had little success, and Lichterfeld would now like to assess cure interventions by quantifying how much they shift the proviral reservoir from gene regions to gene deserts.

Steven Deeks of the University of California, San Francisco, another author on the paper, says the new findings could help the elite controllers. Because of worry that they may suffer chronic inflammation from low levels of HIV replication, many doctors advise this group to still take ARVs. If they have a relatively high level of proviruses in gene deserts, it may be safe for them to stop treatment. "A lot of them now are getting medication they don't need," Deeks says. ■

Published by AAAS



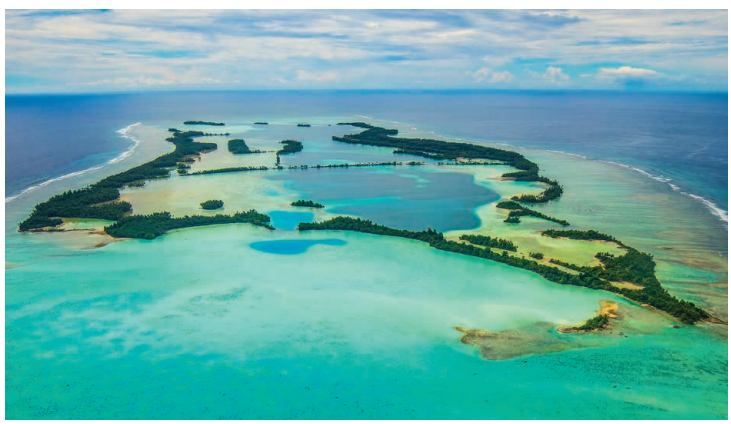
RIDDING PARADISE OF PALMS

On Pacific islands, coconut palms are an aggressive invader. Biologists are trying to eradicate millions of palms from one atoll

By **Kenneth R. Weiss**, *on Palmyra Atoll in the central Pacific Ocean; Photography by* **Dana Edmunds**

Duncan Coles lops the head off a juvenile coconut tree with a practiced swing of his machete. He and nine other volunteers chop away at the thicket, stepping over piles of fallen coconuts and fist-size hermit crabs. Soon they return to their decapitated kills, dousing each stump with blue-dyed herbicide. Two other volunteers use power drills to set upon the mature coconut trees towering overhead, boring holes and filling them with shots of herbicide.

The slashing and poisoning is part of an unprecedented endeavor to rid this remote atoll of all but a few coconut palms (*Cocos nucifera*). The gangly tree is an icon of idyllic tropical islands, but also an aggressive invasive species that crowds out native plants and animals. By removing 99% of Palmyra's millions of palms, biologists hope to create more room on the atoll's three dozen islets for indigenous forests and seabirds, including the world's second largest colony of red-footed boobies. If all goes as intended, the restoration effort could help make this coral-ringed atoll, which has an elevation of just 2 meters, more resilient to sea-level rise and other ravages of climate change.



Palmyra Atoll (above and opposite page) sits in isolation 1600 kilometers south of Honolulu, just north of the equator.

The researchers are now nearly halfway to their goal, and are intently tracking whether their tactics are producing the predicted ecological effects. If the restoration methods work, they could ultimately be replicated on other islands with abandoned coconut plantations run amok. On Palmyra, however, the team aims to transform virtually the entire landscape and ecosystem, something that's only possible because the atoll has no human inhabitants relying on coconut palms to produce food, water, fuel, or building materials, says ecologist Alex Wegmann of the Nature Conservancy (TNC), which co-owns the atoll with the U.S. Fish and Wildlife Service (USFWS). "If there is anywhere safe to test this," Wegmann says, "Palmyra is that place."

Palmyra's transformation into a laboratory for ecological restoration marks yet another twist in its rich history, which has included serving as a U.S. Navy base during World War II and a coconut plantation. More recently it was eyed as the potential home of a tropical nudist colony, a nuclear waste dump, and a launch pad to shoot rockets into space. Located 1600 kilometers south of Honolulu, the atoll sits just shy of 6° north of the equator, on the edge of the intertropical convergence zone, a band of ocean known for slack winds and abundant rainfall. Palmyra receives 4.3 meters of rain per year.

The atoll got its name in 1802 after its namesake, a U.S. sailing ship, splintered on its reef, a mere 4 years after the atoll's first reported sighting. Sixty years later, Hawaii's King Kamehameha IV declared Palmyra part of his kingdom, and it became a U.S. territory in 1898 when the United States annexed Hawaii. Since then, it has been mostly in private hands.

In 2000, TNC purchased the atoll for $30 million and sold most of the outer islets and lands that emerge at low tide to the U.S. government for $9 million. On Cooper Island, which has an airstrip, TNC built a research station, which has helped the atoll become a prime destination for scientists seeking to study marine ecosystems largely free of human impacts.

Others have looked at the ecological connections between land and sea. Researchers led by marine ecologist Douglas McCauley and terrestrial ecologist Hillary Young, both of the University of California, Santa Barbara, have traced how seabirds that roost and nest in Palmyra's native forests act as nutrient transporters. The birds catch fish and squid in the surrounding ocean, then rain nutrient-rich guano onto the soil, nurturing vegetation and the animals that feed on the plants. The nutrients also wash into the atoll's lagoon, producing blooms of plankton that feed manta rays and other marine life.

But that virtuous circle is weaker in forests dominated by the coconut palms. Originally from Thailand and the Philippines, according to genetic evidence, the palms were brought to the central Pacific Ocean by Polynesian seafarers an estimated 1500 years ago, as a fast-growing source of food and other resources. There are no signs of ancient human settlements on Palmyra, so the trees may have arrived on the atoll when currents washed their nuts—which can survive for months at sea thanks to their thick buoyant husks and fat reserves—from some distant inhabited island.

By the early 1800s, a passing sailor wrote that some of Palmyra's "islets seemed exclusively appropriated to the cocoa-nut tree." In 1885, entrepreneurs began to plant commercial groves, which continued intermittently through the 1950s. In a 1913 survey of the atoll, botanist Joseph Rock counted some 25,000 palms, bearing nuts which were "the finest and the biggest the writer has ever seen." He sent some to an expert in Italy, who was impressed by "their large size … [and] distinctly trigonal shape … the three edges almost take the shape of wings."

But although the palms have flourished, the forests they dominate have far fewer birds and fewer manta rays foraging along their shores. Seabirds probably avoid nesting in coconut palms, Young says, because of their branchless trunks, relatively small canopies, and swaying fronds. They prefer a more stable architecture—the forks, nooks, and crannies of native trees.

As a result, coconut-dominated forests tend to have nutrient-poor soils, less biological diversity, and a different suite of insects (many of them invasive species) compared with native forests. To better understand such differences, ecologist John McLaughlin, a postdoctoral researcher in Young's lab, has fogged palms and native trees with pesti-

Published by AAAS



Red-footed boobies roost in a native *Pisonia* tree. Palmyra's seabirds tend to shun introduced palms in favor of native species.

cides, using giant funnels to catch the insects that fall out. It's part of an ambitious food web project to reveal the abundance, diets, and interactions of 300 species. On a bad day in the palm forest, McLaughlin says, the funnels will thwack with the sound of palm-size cane spiders and invasive giant cockroaches hitting the Teflon-coated plastic cones. That's when, he says, "You know it's time to get out of there."

Research has shown that fast-growing coconut palms can suck up moisture so fast that they increase water stress in nearby native trees. The palms are more drought and salt tolerant than native trees—advantages that help them establish dominance and spread as a monoculture. Then, in a coup de grâce, they crush competitors with falling nuts and fronds.

Ripping out the palms has long been on the list of restoration projects on Palmyra. First, however, managers decided to attack another invader, black rats (*Rattus rattus*), which likely arrived on ships during World War II. With no predators, rats multiplied into the tens of thousands. They ate the seeds and gnawed the saplings of native trees and attacked seabird colonies, including those of sooty terns, which nest on the ground. Rats are the key suspects behind the absence on Palmyra of eight other species of ground or burrow-nesting birds, including shearwaters and petrels, all found on central Pacific islands that have remained rat-free.

The first attempt to eradicate the rats in 2002 failed, partly because Palmyra's abundant land crabs outcompeted the rodents for the poisonous bait. The crabs' physiology allowed them to eat the poison—the anticoagulant brodifacoum—without ill effect.

The second effort was successful only after Wegmann and colleagues radio-collared rats and discovered that the rodents liked to hang out in the crowns of coconut palms. The crowns became a convenient platform for stashing cotton gauze sacks of poison bait, delivered by workers firing slingshots or dangling from helicopters. Crabs do not reach the palm tops.

Once rats were exterminated in 2011, researchers watched with delight as native tree saplings began to spring from the forest floor. There were also happy surprises. Scientists discovered two additional species of land crabs that had likely gone undetected because voracious rats suppressed their numbers. And researchers realized they were no longer being bitten by Asian tiger mosquitoes, a pest that attacks during the day and can carry dengue and yellow fever. It appears the mosquitoes depended on rats rather than humans or birds for blood meals, says parasitologist Kevin Lafferty of the U.S. Geological Survey.

The eradication of the rodents, however, only made the palm problem more apparent. Without rats devouring palm flowers, nuts, and sprouts, mounds of coconuts built up on the forest floor, and lagoons became lumpy with floating nuts, each a tree in waiting. Falling coconuts and fronds posed such a threat that managers, volunteers, and others were required to wear hardhats while in the forests, or to enter with a buddy, so they wouldn't get knocked out and potentially picked apart by crabs scavenging anything that falls to the ground. The largest scavenger, the coconut or robber crab, can weigh up to 4 kilograms, hoisted by legs spanning 1 meter.

With such obstacles in mind, managers began to plan a $2.2 million campaign to obliterate most of the palms, its funding largely split between USFWS and TNC, with a $250,000 grant from the nonprofit Wildlife Conservation Society. The nonprofit Island Conservation promises $100,000.



Coconut crabs, which can reach 4 kilograms, scavenge what they like.

PHOTO: DANA EDMUNDS

Managers knew it would be a monumental task: Mapping revealed that 42% of Palmyra's 248-hectare landmass was covered with palm forests, holding some 29,000 adult trees and 2.1 million juveniles. "The numbers seemed daunting," recalls Wegmann, the program's science director. "But at least, unlike the rats, we knew that the palms couldn't hide or run away."

Felling Palmyra's dense palm groves with chainsaws was a no-go: The top-

Published by AAAS



Nick Holmes (left) and Alex Wegmann (right), scientists with the Nature Conservancy, look for seabirds after volunteers have cut down and poisoned palms.

PHOTO: DANA EDMUNDS

pling trunks would crush native vegetation and wildlife, and release a potentially harmful pulse of carbon and other nutrients into the lagoon and coral reef as the vegetation decomposed. Instead, the researchers settled on cutting down the saplings and poisoning the adult palms with the herbicide glyphosate. The dead trees, still standing, would decompose more slowly, reducing nutrient pulses—at the cost of making some islets look like sand ashtrays filled with giant half-burned match sticks.

Project managers had hoped the palm cull, which began in 2019, would be done by the end of this year. But the coronavirus pandemic has put the program on pause; the most recent team of volunteers was airlifted off the atoll at the end of May. The crews had made progress by then, poisoning 54% of the adult palms and felling 40% of the sprouts. They hope to be able to come back and finish the job next year.

"We are not trying to eradicate coconuts," notes Stefan Kropidlowski, manager of the USFWS refuge on Palmyra. "We are trying to remove the coconut monoculture and allow the native forests to get a foothold." Managers expect to retain 1% of the palms, mainly on islet edges for erosion control. They'll also preserve some trees that produce the unusual trigonal nuts, and leave plots of palms that will serve as controls in planned studies, including examinations of changes in soil carbon storage and food web diversity.

Once the palms are mostly gone, crews will plant native trees, especially the giant *Pisonia* and the tree heliotrope. The hope is that the native plants will shade out invasive palms and become the dominant species, Wegmann says. Yet some scientists who do fieldwork on Palmyra fear managers may end up on a conservation treadmill, endlessly cutting back invasive palms. Responding to such criticism, project managers say they have found it's easy to kill palm sprouts with a machete before they start to produce nuts. "It takes coconuts 10 years to reach reproduction age," Wegmann says. "If there is maintenance, it will be minimal."

Whether the palm cull will boost seabird populations remains to be seen, says Young, whose students and colleagues are helping monitor the effects. She suspects seabird numbers are limited more by the availability of food than nesting sites. "It may benefit the birds, but there are a lot of unknowns."

Wegmann is optimistic. "We are realigning the atoll to be as seabird friendly as possible," he says. Palmyra currently hosts 11 seabird species, including a colony of some 25,000 red-footed boobies (only the Galápagos have a bigger colony). There are also nesting colonies of roughly 100,000 sooty terns, smaller groups of great frigatebirds, black and brown noddies, white terns, and tropicbirds. Biologists believe perhaps eight other seabird species once used the atoll, including wedge-tailed shearwaters, tropical shearwaters, Phoenix petrels, and white-throated storm petrels. In hopes of attracting them back, researchers have set up a sound system that broadcasts their calls.

They are also considering relocating some very rare species to Palmyra. Its remoteness and protected status could make it a safe haven for endangered birds such as the Kiritimati (Christmas Island) warbler and the Guam kingfisher, which is now kept alive through a captive breeding program.

Larger bird populations spread across the atoll, and their output of guano, could help Palmyra's surrounding coral reefs become more resilient to climate change. Case studies from the Indian Ocean, focusing on rat-free islands with larger seabird populations, have found that the extra guano produces more plankton and algae, as well as faster growing fish that keep coral reefs clean and more productive. A thriving reef can recover more quickly after coral bleaching events caused by marine heat waves, researchers note, and coral growth may be better at keeping pace with rising seas.

Here on Palmyra Atoll, the coral reefs are already in good shape. Earlier restoration work has removed rusting shipwrecks that were leaching iron onto the reefs, fueling the growth of an invasive anemonelike animal called a corallimorph, which smothers corals. Someday, managers hope to remove the last of the causeways built by the U.S. Navy. They impede the flushing of ocean water, leaving parts of lagoons with hot, stagnant water that can kill corals.

Palmyra is in a better position than many of the world's 339 low-lying atolls to cope with sea-level rise. Even under a worst case scenario of a 1-meter rise in the 21st century, models suggest nearly half of its land would remain above water. Still, managers of other atolls overrun by coconut palms and rats—including one in French Polynesia and another in American Samoa—are keeping a close eye on Palmyra's restoration experiment. Its success, they say, could provide clues to how to navigate an uncertain future. ■

Kenneth R. Weiss is a writer in Carpinteria, California.

Published by AAAS

INSIGHTS

PERSPECTIVES



PHOTO: HINDUSTAN TIMES/GETTY IMAGES

VIEWPOINT: COVID-19

Bridging digital health divides

COVID-19 is accelerating digital transformation in the Indian health sector

By **Anurag Agrawal**

On 23 March 2020, India announced a national lockdown to contain the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic (*1*). Routine health care virtually came to a standstill, with only emergency care being provided. In view of the high infectivity of SARS-CoV-2 and scarcity of quality personal protective equipment, patients and health care workers (HCWs) both looked for alternatives to face-to-face care (*2*). Telemedicine and digital health, oversold and underdelivered for the past two decades, found a new impetus driven by coalescence of often antagonistic viewpoints on issues such as rights of medical practitioners versus digital health companies, regulatory standards to ensure quality and data security, and financial models of public versus private goods. However, whether the current urgency will succeed in bringing about digital health transformation, with quality care and seamless connections across spheres of life, will depend on many factors.

The old adage of crisis triggering reform came true in India, with pending matters being rapidly cleared. On 25 March, telemedicine practice guidelines were released, removing uncertainty about the legitimacy of telemedicine since 2018 due to an adverse court decision (*3*). In May, financial approval was granted to the 2019 national digital health blueprint that advocates a separate national digital health authority to develop and administer high-value digital public health goods in the form of a national health stack containing master health data of the nation and necessary tools for authorized access (*4*). These far-reaching decisions are the culmination of a long process, and the rapid timeline of their clearance is testimony to the catalyst effect of the SARS-CoV-2 pandemic on digital health transformation. However, during times of rapid acceleration, it is important to steer precisely and with anticipation to avoid crashes. The organic growth plan contained in the "blueprint" should not become the "wild west," where concern for data integrity, privacy, and ethics gets lost.

India has a well-recognized need and capacity for digital health services such as telemedicine, owing to large underserved rural areas, as well as domestic strength in information technology. There has been continuing investment in general digital infrastructure and governance, and India is one of the fastest-growing digital consumer markets, with more than half a billion internet subscribers using over 8 GB data monthly, at some of the lowest data costs in the world (*5*). Data speed is sufficient for video communication well beyond the megacities and cities, into the rural hinterlands.

Yet, telemedicine is one of the least used digital services despite almost two decades of planning and pilot studies. Satellite-based telemedicine was launched in 2000 by Apollo Hospitals (a large corporate) and the Indian Space Research Organization through a public-private partnership (*6*). A national telemedicine task force was established as early as 2005. Unfortunately, scalable and sustain-

Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research India, New Delhi, India. Email: a.agrawal@igib.in

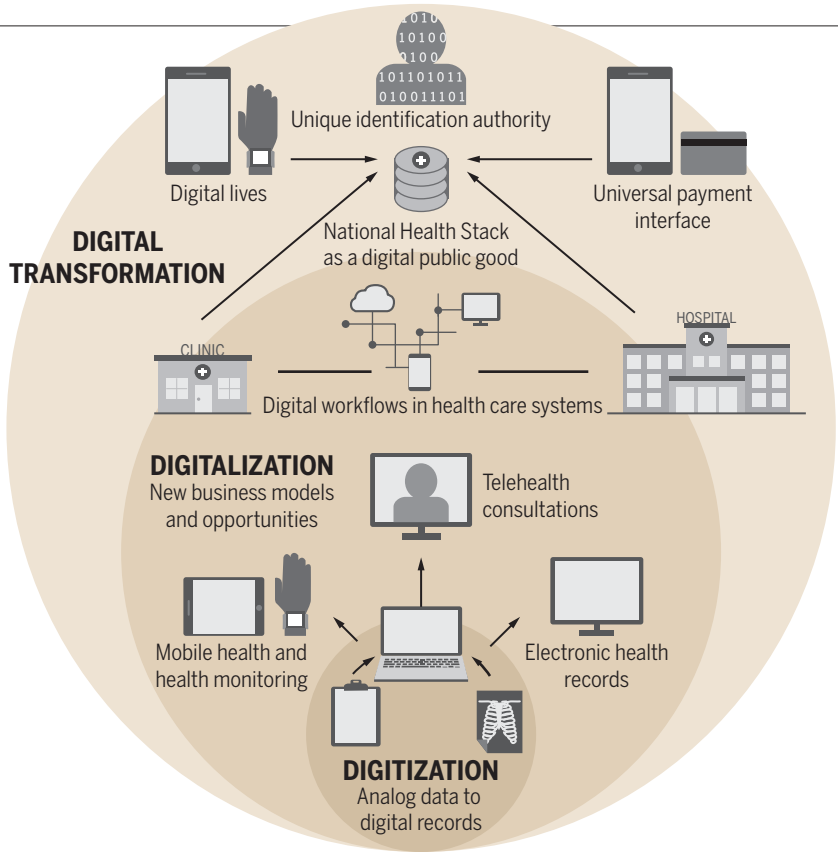
Published by AAAS

The COVID-19 pandemic has accelerated digital health implementation in India, such as this telemedicine hub in New Delhi, expanding the reach of health care.

able models never emerged, partly because of technological constraints such as limited internet speed and partly owing to lack of economic incentives. The few notable successes, such as teleradiology, were due to a naturally digital workflow that required little additional infrastructure, coupled with sustainable economic models. Imaging services became possible in regions that had been unable to recruit radiologists, opening new markets and attracting private investment. Teleradiology was also used to optimally distribute cases to experts across the globe, allowing use of time-zone differences and price differentials to create value-added consultation services from India (*7*). In other medical fields, where workflow was not natively digital, it was difficult to even get to digitization of analog data, let alone digitalization, where core health care processes are unified digitally. Digitalization can improve quality, efficiency, and accessibility but requires reimagination of existing systems (see the figure).

An important lesson from India's experience is that citizen-centric digital health policies, such as increasing accessibility for underserved areas with low sociodemographic development, require simultaneous investments in physical health infrastructure. A major challenge is that more than 80% of outpatient care is currently delivered by the private sector, with wide variations in quality and cost (*8*). However, the government-provided care sector dominates in terms of inpatient beds and perinatal care. Reimagining digital health in India must thus be a multipronged strategy.

Individual practitioners, working on a fee-for-service basis, are the backbone of Indian health care, especially in smaller cities and rural areas. These are typically low-cost, high-volume practices with an average consultation time of less than 2 min, much of which goes into prescriptions and refills (*9*). There is a sense of patient ownership and a justifiable fear that transition to digital health may lead to weakening of relationships with patients, divert referrals, create additional burden on time and infrastructure, and reduce financial returns. In the time of coronavirus disease 2019 (COVID-19), fear has driven consensus regarding the viability and desirability of digital interactions, given that HCWs figured prominently among the global infected and the dead, accounting for a large proportion of young and healthy COVID-19 patients with severe illness (*2*). This has accelerated initial digital adoption by individual practitioners, but there is a danger that these will be stopgap solutions of insufficient quality and without measures for data privacy and protection.



Stages in digital health evolution

The first stage of digitization has been crossed in many places in India. The next stage of digitalization has started in megacities but is yet to percolate nationally. The final stage of digital transformation is envisioned in a national digital health blueprint advocating a fully connected open Health Stack securely aggregating patient, provider, and payer data. Necessary elements such as the world's largest biometrically enabled and cloud-based national unique identification authority and a linked universal payment interface raise hope for successful convergence.

Conversely, digitalization has already happened in corporate health care organizations that employ many physicians, cater to a higher socioeconomic stratum with digital access, and receive payment from insurance companies for documented services. These organizations exercise high control over customized digital platforms that are professionally managed. In contrast to the corporate sector, government health care institutions are largely still analog, but there is a strong digital commitment in the "blueprint." For example, India aims to create 150,000 wellness centers and provide health insurance to 100 million families under the "Ayushman Bharat" scheme. The blueprint envisions a central digital repository and secure access for every citizen to their medical records. This seems possible because necessary foundations have been laid. The world's largest biometric unique identity platform (Aadhaar), covering 1.3 billion Indians, is fully cross-functional with IndiaStack, a set of cross-domain generic building blocks that allows government, businesses, start-ups, and developers to formulate presence-less, paperless, and cashless service delivery solutions. Building a Health Stack of registries, data, and freely available tools on top of the IndiaStack is on the horizon. To ensure affordability and ease of participation by individual practitioners, a digital public goods philosophy is essential. The current greenfield state, with strong foundations, is an excellent opportunity to leapfrog the complex, restrictive, digital health systems that have evolved elsewhere.

In the beginning of the lockdown, India saw an acceleration of messaging-based consultations, largely on the WhatsApp platform. This is easy for patients and individual HCWs, but unsuited to quality, scale, or data security. eSanjeevani is a government-built public telemedicine platform that is expected to nationally support Ayushman Bharat wellness centers in providing free universal health coverage (*10*). COVID-19 has accelerated its uptake and use by multiple states. The central point of accessing COVID-19–related information and care is a multilingual personal risk assessment and

GRAPHIC: MELISSA THOMAS BAUM/*SCIENCE*

Published by AAAS

contact tracing app called Aarogya Setu (AS), which became the world's fastest-growing mobile app after its launch on 2 April (*11*). AS links out to a web portal, AS Mitr (ASM), that permits users to obtain telemedicine consultations, necessary testing, and home delivery of medicines (*12*). All consultations are free, whether through eSanjeevani or one of many private providers. ASM has become a fast-tracked entry point into digital health for well-known tech companies, as well as new coalitions that have sprung up almost overnight between venture capitalists, hospitals, and online marketplaces.

Mobile phone penetration in India is extremely high, but not everyone has a smartphone or internet access suitable for ASM. Under project StepOne, Indian start-ups specialized in cloud telephony, language processing, call handling, and software development came together to cater to people without smartphones or internet, through an automated interactive voice response system that guides callers and enables consultations with volunteer doctors, if needed (*13*). Within a few months, this has become one of the largest such efforts, handling ~40,000 calls daily and spanning 10 states, with volunteers who can handle the 30 major Indian languages.

In India, both the challenges and opportunities are immense. With COVID-19 providing immediate traction and potential to develop very large user bases, many of these entrants are almost certain to extend digital offerings to other health care sectors. There has also been a market shift with digital products created for affordable health care access becoming a preferred option with wider markets. For instance, CogniABle, an Indian digital health company, has been providing early screening and affordable behavioral intervention for autism (*14*). The demand of such niche products has moved from local to global, as the world suffers through lockdowns and looks for digital alternatives. Similar experiences abound in every sector from diabetes to women's health, and innovative digital health products have a great opportunity to scale. Ensuring quality data flow across homes, diagnostic centers, clinics, and hospitals is a major challenge requiring low-cost, high-quality connected health devices and wearables, an area that has now been prioritized by the Council of Scientific and Industrial Research, India. Artificial intelligence (AI) tools are being developed to process the increasing digital data inflows, from chest x-rays to videos. Although the immediate impact of such tools on COVID-19 is unclear, rapid growth of the AI-health interface is likely.

It remains to be seen how conventional health systems, medical providers, and consumers react to these changes after COVID-19. Although health corporates and an emerging class of digital natives in Indian megacities may see immediate benefits, the case is not so clear elsewhere. Increased affordability and access to citizens everywhere is the real touchstone of success. An obvious area of concern, given the composition of the new coalitions, is the ownership and use of data. Use of data to establish dominance in a fledgling market or for exploitative practices is not unheard of in other sectors. Health being a human right, regulatory policies for use, storage, access, and transmission of health data must firmly avoid both paralysis and colonialism. Creation of authorized health-data exchanges, with transparent transactions, is one of the possible solutions (*15*). The purpose of digital health care remains better care of people's health; a vibrant data or digital economy is only a desirable side product.

The future of digital health in India looks bright, with COVID-19 providing urgency to a sector that has seen much deliberation with limited action. It is important to now focus on building trust and reducing friction, between the existing health system and digital entrants. Patient primacy, ethical data use, respect for conventional health care, and sustainable growth models must be integral parts of the Indian digital health movement. For the foreseeable future, babies will still be delivered by humans, and public health infrastructure, strong social contracts, and thoughtful leadership will still matter most. ■

REFERENCES AND NOTES

1. P. Pulla, *BMJ* **368**, m1251 (2020).
2. P. Lapolla *et al.*, *Infect. Control Hosp. Epidemiol*. 10.1017/ice.2020.241 (2020).
3. Ministry of Health and Family Welfare (India), Telemedicine Practice Guidelines (2020); www.mohfw.gov.in/pdf/Telemedicine.pdf.
4. Ministry of Health and Family Welfare (India), National Digital Health Blueprint (2019); www.main.mohfw.gov.in/sites/default/files/National_Digital_Health_Blueprint_Report_comments_invited.pdf.
5. McKinsey Global Institute, Digital India: Technology to transform a connected nation (2019); www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation.
6. V. G. Chellaiyan, A. Y. Nirupama, N. Taneja, *J. Family Med. Prim. Care* **8**, 1872 (2019).
7. A. Agrawal, D. B. Koundinya, J. S. Raju, A. Agrawal, A. Kalyanpur, *Emerg. Radiol.* **24**, 157 (2017).
8. T. Jayakrishnan, B. Thomas, B. Rao, B. George, *Int. J. Med. Public Health* **3**, 225 (2013).
9. G. Irving *et al.*, *BMJ Open* **7**, e017902 (2017).
10. eSanjeevani; https://esanjeevani.in.
11. A. Jhunjhunwala, *Indian Natl. Acad. Eng.* **5**, 157 (2020).
12. Aarogya Setu Mitr; www.aarogyasetumitr.in.
13. Project StepOne; www.projectstepone.org.
14. CogniAble; https://cogniable.tech.
15. S. Balsari *et al.*, *J. Med. Internet Res.* **20**, e10725 (2018).

ACKNOWLEDGMENTS

Thanks to Lipsa for help with artwork and The Lancet–Financial Times commission for Governing Health Futures 2030 for discussions. A.A. is supported by the Council of Scientific and Industrial Research (India), Wellcome Trust-DBT India Alliance, and Fondation Botnar.

10.1126/science.abc9295

MEDICINE

Restoring neuron connections

A synthetic peptide promotes functional recovery in neurodegeneration and spinal cord injury models

By **Patricia C. Salinas**

The structural and functional integrity of synapses is crucial for neuronal circuit function and for the ability of animals to respond and adapt to their environment. However, synapses become dysfunctional or are lost in developmental and neurodegenerative disorders, and after brain and spinal cord injury (SCI). Deficiency in synapse integrity contributes to cognitive impairment and motor disability. To restore damaged neuronal circuits after SCI, antibody therapy against molecules such as Nogo, which inhibits axon growth, are in clinical trials (*1*). For other conditions such as Alzheimer's disease (AD), blockade of the pathogenic molecule amyloid-β or its downstream effects are being tested as potential therapies (*2*). To develop effective approaches for restoring impaired synapses, careful consideration should be given to the diversity of synaptogenic factors. To overcome this, on page 1074 of this issue, Suzuki *et al.* (*3*) report a structurally guided synthetic peptide to restore synaptic integrity of damaged and impaired circuits in mice.

The formation of synapses requires the combinatorial function of several molecules, including extracellular secreted molecules [Wnts, fibroblast growth factors (FGFs), and semaphorins], transmembrane proteins [neuroligin (NLGN), neurexin (NRX), ephrin receptors, and synaptic cell adhesion molecule (SynCAM)], and also components of the extracellular matrix present at the synaptic cleft, such as extracellular scaffolding proteins (ESPs) (*4, 5*). ESPs, which include cerebellin (CBLN1) and neuronal pentraxin 1 (NP1) (*6*), form bridges across the synaptic cleft, allowing

Department of Cell and Developmental Biology, University College London, London, UK. Email: p.salinas@ucl.ac.uk

Published by AAAS

the coordinated assembly of the presynaptic terminal that contains the machinery for neurotransmitter release and the postsynaptic side with neurotransmitter receptors, ion channels, and signaling molecules. These synaptic organizers could be ideal molecules to restore defective synaptic connections. However, the diversity of these molecules in different brain regions poses a challenge for designing strategies to repair damaged neuronal circuits.

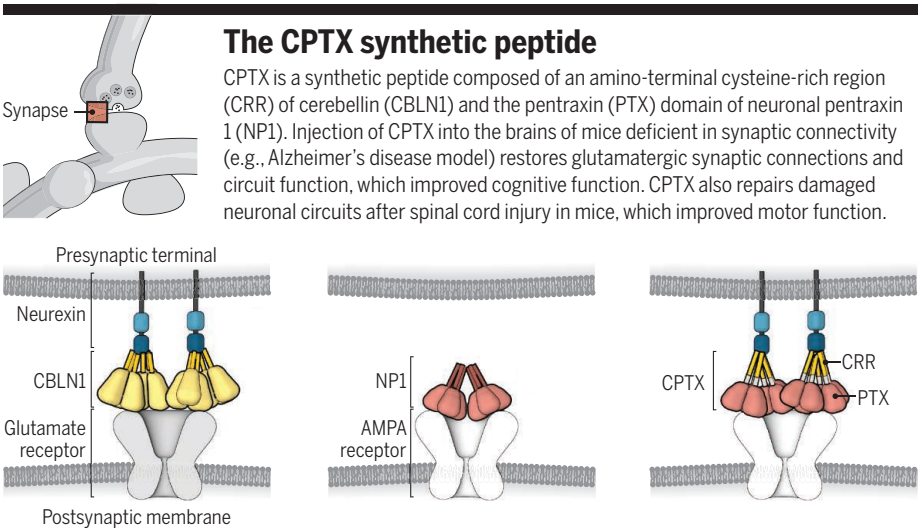
Suzuki *et al.* designed a molecule that could promote the formation of synapses in various brain regions by taking advantage of the structure of two widespread ESPs, CBLN1 and NP1 (see the figure). The synthetic peptide, CPTX, contains a key binding domain from each of these proteins: the amino-terminal cysteine-rich region (CRR) of CBLN1 and the pentraxin (PTX) domain of NP1 (NP1-PTX). The CRR binds specific ectodomains of NRX1β, a protein present at most presynaptic terminals, whereas NP1-PTX binds to the extracellular regions of postsynaptic glutamate receptors. CPTX therefore forms a molecular bridge between the pre- and postsynaptic sides, promoting the assembly of glutamatergic synapses.

Suzuki *et al.* evaluated the effectiveness of CPTX in promoting synapse formation in vivo. CPTX injection into the adult hippocampus of wild-type mice promotes the formation of functional glutamatergic synapses 3 days after injection. To address the potential beneficial effect of CPTX on synaptic recovery, the authors injected CPTX in mice lacking the glutamate receptor delta-2 subunit (*GluD2*) or *Cbln1* genes, which exhibit defects in synapse number, synaptic function, and ataxia (*7, 8*). Three days after injection of CPTX into the cerebellum of these adult mice, synaptic connectivity was evaluated by using cellular, electrophysiological, and behavioral approaches. In both models, CPTX increased the number of connections between Purkinje cell (PC) neurons and the parallel fibers (PF) axons of granule cell neurons, the postsynaptic partners of PCs. CPTX also restored the functional connectivity between PFs with PCs and improved motor coordination. Thus, CPTX promotes structural and functional recovery of neuronal circuits in cerebellar dysfunction in mice.

To further test the utility of CPTX for therapeutic use, the authors investigated its effect on a model of neurodegeneration and also on a model of SCI. They used 5xFAD mice, an established mouse model of early familial AD. These mice exhibit synapse degeneration and defects in both long-term potentiation [(LTP), which reflects persistent strengthening of synapses], and hippocampal-mediated memory (*9*). Injection of CPTX into the hippocampus of 15- to 18-month-old mice—the time when synaptic connectivity is severely impaired—completely restored the number of postsynaptic structures (dendritic spines), LTP, and hippocampal-mediated memory. The effectiveness of CPTX was also tested in a mouse model of SCI where injury was caused by hemisection at thoracic level, a procedure that damages the motor nerve tracks that innervate skeletal muscles (*10*). Injection of CPTX at the time of the hemisection or a week later resulted in substantial functional recovery. Suzuki *et al.* also induced injury by contusion or compression, which is considered the best model for SCI in humans. Injection of CPTX into the spinal cord induces functional recovery even when tested 8 weeks after the injury. These exciting findings demonstrate that CPTX induces functional recovery after a single injection and its effect persists for several days or weeks.

Analyses of the functional recovery induced by CPTX reveal two interesting findings. One is that CPTX induces functional recovery even when the protein is cleared from the tissue (after 3 days). These findings suggest that CPTX triggers a recovery cascade that persists beyond its presence in the tissue. In addition, the long-term effects of CPTX vary in different brain regions and/or the nature of the injury. In the cerebellum, CPTX injection in *Cbln1* mutant mice induces a transient functional recovery that lasts only a few days. By contrast, CPTX injection into the spinal cord induces functional recovery that lasts several weeks. Further studies are needed to elucidate the molecular mechanisms by which CPTX regulates synapse recovery in different regions of the central nervous system or mode of injury.

The study of Suzuki *et al.* demonstrates that CPTX is a powerful molecule that restores synaptic connectivity and behavioral function in different brain regions and pathological conditions. The biggest challenge now is translating these promising and exciting results into medical applications. Several aspects need to be considered. CPTX could have side effects—for example, by inducing an imbalance between excitatory and inhibitory synapses. Although this seems unlikely as the animals do not exhibit signs of increased excitability, further studies are needed to test potential adverse effects. The mode of delivery of the synthetic peptide is important, as direct injection into the brain is an invasive procedure. Alternatively, intravenous injection could be used, but then the peptide (or analogs) needs to cross the blood–brain barrier to access the brain. Given the size and structure of CPTX, it is unlikely to do so. In the future, CPTX or newly designed analogs that can access the brain could provide an effective treatment for neurodevelopmental disorders, neurodegenerative conditions, and nerve injury. ■



The CPTX synthetic peptide

CPTX is a synthetic peptide composed of an amino-terminal cysteine-rich region (CRR) of cerebellin (CBLN1) and the pentraxin (PTX) domain of neuronal pentraxin 1 (NP1). Injection of CPTX into the brains of mice deficient in synaptic connectivity (e.g., Alzheimer's disease model) restores glutamatergic synaptic connections and circuit function, which improved cognitive function. CPTX also repairs damaged neuronal circuits after spinal cord injury in mice, which improved motor function.

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. K. Kucher *et al.*, *Neurorehabil. Neural Repair* **32**, 578 (2018).
2. Alzforum (2020); www.alzforum.org/therapeutics.
3. K. Suzuki *et al.*, *Science* **369**, eabb4853 (2020).
4. T. C. Südhof, *Neuron* **100**, 276 (2018).
5. V. Budnik, P. C. Salinas, *Curr. Opin. Neurobiol.* **53**, 90 (2011).
6. M. Yuzaki, *Annu. Rev. Physiol.* **80**, 243 (2018).
7. A. Ito-Ishida *et al.*, *J. Neurosci.* **28**, 5920 (2008).
8. K. Matsuda *et al.*, *Science* **328**, 363 (2010).
9. H. Oakley *et al.*, *J. Neurosci.* **26**, 10129 (2006).
10. G. Courtine, M. V. Sofroniew, *Nat. Med.* **25**, 898 (2019).

ACKNOWLEDGMENTS

The author is supported by the MRC, BBSRC, Alzheimer's Research UK, and Alzheimer's Society.

10.1126/science.abd4762

Published by AAAS

CHEMICIAL PHYSICS

A quantum-computing advantage for chemistry

Error-mitigated Hartree-Fock computation is performed with up to 12 superconducting qubits

By **Xiao Yuan**

Quantum computers potentially have computational power greater than that of their classical counterparts. The recent demonstration of "quantum supremacy" on Google's 53-qubit Sycamore quantum processor (*1*) has reinforced this idea, but it remains unknown whether the next generation of quantum computers will be able to solve classically intractable problems of practical interest. On page 1084 of this issue, Rubin *et al.* (*2*) take steps toward answering this question with an experimental implementation of Hartree-Fock calculations of molecular electronic energies on a superconducting processor. Although the calculations performed are also efficient to run on classical computers, the experiment demonstrates many of the key building blocks for quantum chemistry simulation and paves the way toward achieving quantum advantage for problems of chemical interest.

Using controllable quantum systems to simulate quantum mechanical problems in chemistry and physics was the brainchild of Richard Feynman, who remarked in the 1980s that "If you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy" (*3*). Since then, there has been substantial theoretical and experimental progress toward this goal. In particular, the rapid recent development of superconducting qubits, such as Google's Sycamore quantum processor, has enabled quantum supremacy, which samples from the outputs of random quantum circuits more efficiently than appears possible with even the largest classical supercomputers (*1*).

Rubin *et al.* investigate the performance of this same processor to determine the electronic structure of molecular systems. Such an accomplishment would have academic as well as commercial value, as it could enable the design of improved catalysts or new medicines. Since the first quantum algorithm for quantum computational chemistry was proposed in 2005 (*4*), there have been numerous developments to reduce its computational cost (*5*, *6*). One of the most influential developments is that of the variational quantum eigensolver (VQE), which reduces the burden on the quantum processor by leveraging a classical coprocessor (*7*) (see the figure).

Prior proof-of-principle VQE experiments have realized electronic structure calculations with up to six qubits (*8*). It is still an open question as to whether the VQE can solve classically intractable instances of the electronic structure problem, which may require on the order of 100 qubits. As the problem size increases, so too does the quantum circuit depth (the number of layers of gates; five in the figure) required to realize the quantum algorithm. Even if the quality of the qubits is maintained while the processor is scaled up, larger processors with deeper circuits will lead to an increased error rate for the calculation. Assessing whether this build-up of errors is fatal for the VQE is one of the most pressing open questions in the field of quantum computing.

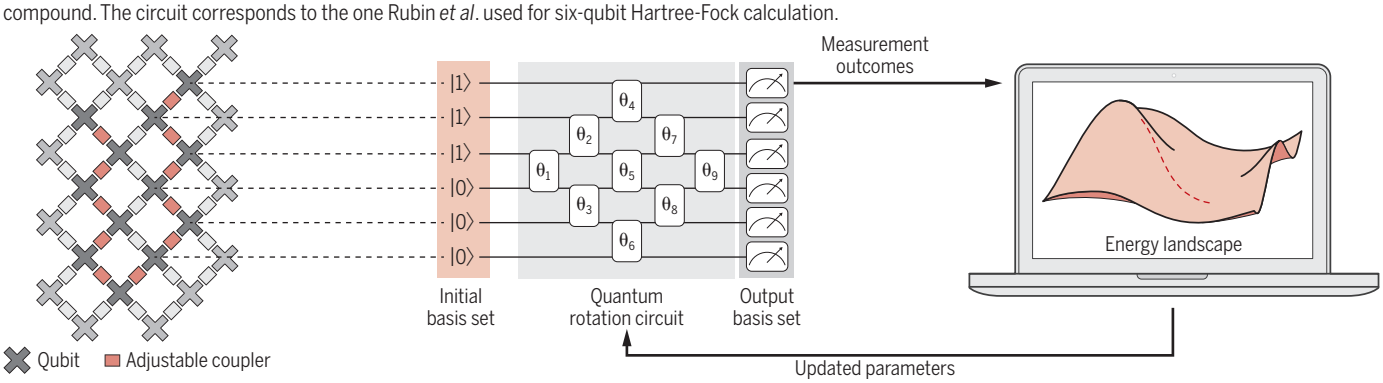
Rubin *et al.* take steps to address these open questions through an experimental VQE implementation using 6 to 12 qubits. The experiment implements the Hartree-Fock method for calculating the binding energy of hydrogen chains and the isomerization of diazene. The Hartree-Fock method provides approximate solutions to the electronic structure problem and is a classically tractable calculation. It is typically used as an initial step in quantum computational approaches to solving the electronic structure problem.

Nonetheless, this VQE experiment demonstrates many of the key components for large-scale VQE implementations, including electronic state preparation, Hamiltonian measurement for any one- and two-particle reduced-density matrix elements, two error mitigation techniques, and outer-loop classical optimization. Together, these features lead to the successful extension of prior investigations into quantum computational

Stanford Institute for Theoretical Physics, Stanford University, Stanford, CA, USA. Email: xiao.yuan.ph@gmail.com

A variational quantum eigensolver

A parameterized quantum circuit, with properly prepared initial states and with the aid of a classical co-processer, approximates the wave function of a chemical compound. The circuit corresponds to the one Rubin *et al.* used for six-qubit Hartree-Fock calculation.



Google Sycamore
The schematic shows the superconducting qubits and the adjustable couplers that connect them.

Quantum circuit
The input states of the wave function are initialized as $|0\rangle$ or $|1\rangle$. Each box with a parameter θ is a rotation gate that is compiled into two two-qubit gates and three single-qubit gates.

Classical calculation
The outcomes are fed into a classical computer that trains the parameters to learn the ground state of the system. The classical parameters that determine the wave function are fed back to the quantum computer, and the calculation is rerun.

GRAPHIC: JOSHUA BIRD/*SCIENCE*

Published by AAAS

chemistry. The techniques demonstrated in this work will likely form the foundation of future VQE experiments targeting classically intractable systems.

Perhaps the most important conclusions from the work of Rubin *et al.* are the necessity of tailoring algorithms to the quantum processor and the importance of error mitigation techniques. Because near-term quantum devices typically have restricted realizable gates, the ability to compile the circuit for a given architecture is crucial for simulation efficiency and accuracy. The methods showcased by Rubin *et al.* for realizing electronic states with nearest-neighbor gates shed light on how to implement more complicated calculations with restricted hardware topologies.

Even with the compilation methods discussed above, the presence of noise in these calculations is still a pressing issue. Rubin *et al.* show how to obtain accurate results despite this noise through the use of error mitigation strategies. The techniques used in this work are specialized for quantum simulations, exploiting particle conservation through local density matrix information and the *N*-representability conditions of fermionic systems. For the 12-qubit calculation with 72 two-qubit gates, the combined error mitigation techniques effectively improve the raw-state fidelity to >99%, which represents an increase of about two orders of magnitude.

Whether noisy intermediate-scale quantum computers will be able to surpass classical supercomputers in solving chemistry problems has become one of the most exciting questions in quantum computing. Preliminary calculations on small- and intermediate-sized systems have verified the feasibility of the most promising quantum algorithms. However, further work is needed to enable similar calculations to be performed for system sizes that are greater by one to two orders of magnitude. Experimentally, quantum devices need to be scaled up to hundreds or even thousands of qubits. The characterization of errors should be improved alongside a reduction in error rates. Theoretically, we seek more efficient algorithms and more effective error mitigation techniques, especially ones tailored for specific problems and quantum devices. ■

REFERENCES AND NOTES

1. F. Arute *et al.*, *Nature* **574**, 505 (2019).
2. N. C. Rubin *et al.*, *Science* **369**, 1084 (2020).
3. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982).
4. A. Aspuru-Guzik *et al.*, *Science* **309**, 1704 (2005).
5. S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, X. Yuan, *Rev. Mod. Phys.* **92**, 015003 (2020).
6. Y. Cao *et al.*, *Chem. Rev.* **119**, 10856 (2019).
7. A. Peruzzo *et al.*, *Nat. Commun.* **5**, 4213 (2014).
8. A. Kandala *et al.*, *Nature* **549**, 242 (2017).

10.1126/science.abd3880



A health worker fumigates homes in Managua, Nicaragua, to reduce the risk of mosquito-borne dengue and Zika viruses.

PHOTO: REUTERS/OSWALDO RIVAS

INFECTIOUS DISEASE

Zika virus increases risk of dengue disease

Interactions between Zika virus immunity and dengue disease have implications for vaccination

By **Hannah Clapham**

The dengue flavivirus (DENV) is a mosquito-borne virus that occurs in more than 100 countries and is known for inducing complex immunity in human infections. The first infection with one of the four DENV serotypes generally has a low risk of severe outcome. Subsequently, it is thought that there is a period of cross-protection in which infection with any serotype results in even lower risk of severe disease. After this period, an infection with a different serotype has a higher risk of severe disease (*1*). In 2015 to 2016, Zika virus (ZIKV, also mosquito-borne) spread rapidly across Central and South America and the Caribbean, where DENV is also endemic (*2*). Because ZIKV is also a flavivirus, whether prior infection with DENV alters Zika disease risk and vice versa became an important question as ZIKV spread. On page 1123 of this issue, Katzelnick *et al.* (*3*) studied the long-running dengue cohort in Nicaragua and found evidence that prior ZIKV infection increases dengue disease risk.

DENV infection can result in a range of outcomes, including asymptomatic, mild flu-like illness and severe dengue, which is characterized by complications of bleeding and organ damage. ZIKV infections are mostly mild, except for a small proportion that result in Guillain-Barre syndrome (autoimmunity to peripheral nerves) or birth defects in infants born to mothers infected during pregnancy.

An initial DENV infection generates high levels of serotype cross-reactive neutralizing antibodies. However, these antibodies wane over time and are thought to reach "subneutralizing" levels. The more severe disease in subsequent infections is thought to arise in part from antibody-dependent enhancement, whereby these antibodies are able to bind to, but not neutralize, the virus, which can promote virus cell entry. To study the relationship between prior DENV infections and subsequent dengue outcome, a long-running cohort study is ideal. The cohort study in Managua, Nicaragua, running since 2004, has led to numerous insights into dengue immunity and disease. With the arrival of ZIKV, this study effectively expanded to include ZIKV infections, leading to the findings of Katzelnick *et al.*

For ZIKV infection after a DENV infection, in vitro testing of sera showed that previous DENV infection led to both neutralizing and enhancing responses to ZIKV (*4*).

Saw Swee Hock School of Public Health, National University Singapore, Singapore. Email: hannah.clapham@nus.edu.sg

Published by AAAS

In nonhuman primates, DENV infection did not affect the outcome of subsequent ZIKV infection (*5*). However, population studies showed that previous DENV immunity correlated with lower risk of ZIKV infection (*6*). In the Nicaraguan cohort, there was evidence that previous DENV infection may protect against subsequent symptomatic ZIKV infection. However, there was no relationship with DENV immunity, and no enhancing effect was observed (*7*).

Therefore, studies on how DENV infection affects subsequent ZIKV infection outcome have shown varied results. The reasons for this could include differences in the number of prior DENV exposures and serotypes; time between DENV and ZIKV infections; and differences between animal model, in vitro, and human responses. In addition, the range of outcomes for Zika disease is not as wide as for dengue, so differences due to any enhancement could be harder to detect. There are also remaining questions about how prior DENV infection modulates the risk of birth defects to infants born to mothers infected with ZIKV during pregnancy.

What about how ZIKV infection modulates subsequent DENV infection outcome? Before the 2015–2016 outbreak, there had been little research on ZIKV and few reported cases, so the long-term impact of ZIKV infection could not be assessed. Initially after the outbreak, it was possible to show that individuals who were exposed to ZIKV generated both neutralizing and enhancing antibodies for DENV (*8*). In nonhuman primates, there was some evidence of enhancement upon infection with DENV a short time after ZIKV infection (*9*). Now, 5 years after the outbreak, it is possible to assess the longer-term effects of ZIKV on subsequent dengue disease risk in humans. A previous modeling study using reported dengue cases from Colombia and Brazil estimated that ZIKV infection led to cross-protection to DENV, resulting in low reported dengue case numbers in the 2 years after the ZIKV outbreak (*10*). This study predicted future higher numbers of dengue cases after this period, which will need to be assessed.

The study of Katzelnick *et al.* similarly shows very low DENV transmission in the 2 years after the ZIKV outbreak. The important aspect of this study is in the assessment of individual disease risk in the 2019 dengue outbreak, the largest on record. During this epidemic, Katzelnick *et al.* clearly show that on an individual level, there is an increased risk of symptomatic and severe dengue disease in individuals with prior ZIKV immunity. Together, these data show that ZIKV infection modulates the disease severity of subsequent DENV infection in both negative and positive ways, on an individual and population level. This suggests that previous ZIKV infection acts similarly to previous DENV infection. However, there are still unknowns, such as whether risk differs by DENV serotype and the longer-term impact of ZIKV infection on dengue.

Another reason this is an important result is for vaccination. Follow-up of participants in the Dengvaxia vaccine trial showed that dengue-unexposed individuals who were vaccinated showed an increased risk of severe disease in the longer term upon the next DENV infection, similar to individuals experiencing a second natural DENV infection (*11*). Dengvaxia is the only licensed dengue vaccine to date, but there are others in development. For ZIKV, there are also several vaccines currently in development, but none are licensed. It is now important to establish whether dengue vaccine efficacy varies according to previous ZIKV infection. For Dengvaxia, vaccination in the initial trials was before the 2015–2016 ZIKV outbreak. It is also possible that ZIKV vaccination could lead to more severe subsequent DENV infections. It would therefore seem prudent for both DENV and ZIKV vaccine studies to measure immunity to both viruses before and after vaccination and to follow up for both diseases within any trial. Because viral immunity will be changing over time, long-term follow-up will be required.

For other flaviviruses, this interaction with vaccination and risk has also been considered, such as for Japanese encephalitis virus and DENV (*12*). Interactions with the yellow fever flavivirus and vaccinations should also be considered in relevant countries. More generally, the cross-reactive properties of viral immunity, how to measure it, how it changes over time, and the impact on disease risk are important for a number of viral infections, and the study of Katzelnick *et al.* serves as a reminder that issues of cocirculation and cross-immunity should remain in view for understanding transmission dynamics and vaccination. ■

REFERENCES AND NOTES

1. W. Dejnirattisai *et al.*, *Science* **328**, 745 (2010).
2. WHO, Zika virus and complications: 2016 Public Health Emergency of International Concern; www.who.int/emergencies/zika-virus-tmp/en/.
3. L. C. Katzelnick *et al.*, *Science* **369**, 1123 (2020).
4. W. Dejnirattisai *et al.*, *Nat. Immunol.* **17**, 1102 (2016).
5. M. K. McCracken *et al.*, *PLOS Pathog.* **13**, e1006487 (2017).
6. I. Rodriguez-Barraquer *et al.*, *Science* **363**, 607 (2019).
7. A. Gordon *et al.*, *PLOS Med.* **16**, e1002726 (2019).
8. K. Stettler *et al.*, *Science* **353**, 823 (2016).
9. J. George *et al.*, *Sci. Rep.* **7**, 10498 (2017).
10. R. K. Borchering *et al.*, *Nature* **10**, 5730 (2019).
11. S. Sridhar *et al.*, *N. Engl. J. Med.* **379**, 327 (2018).
12. K. B. Anderson *et al.*, *PLOS Negl. Trop. Dis.* **5**, e1311 (2011).

10.1126/science.abd5922

NEUROSCIENCE

Strategies for navigating a dynamic world

A rich neural representation of its environment enables an adaptable organism to respond to changes

By **Saurabh Steixner-Kumar** *and* **Jan Gläscher**

One of the most difficult problems for an adaptable agent is gauging how to behave in a nonstationary environment. When conditions are stable, an organism generally pursues a strategy known to provide the best outcome. However, when environmental conditions change, an organism abandons the current action plan and searches for a new best option. The most challenging aspect of this search—calculating the exact time point at which to change strategies—requires the brain to integrate past and present observations and evaluate whether they remain consistent with current environmental conditions. On page 1076 of this issue, Domenech *et al.* (*1*) report on the modeling of rare direct electrical recordings from the prefrontal cortices (PFCs) of a small group of human epilepsy patients as they flexibly negotiated a nonstationary environment.

To understand the brain's mode of navigation, consider for example a sailor at sea (see the figure). The winds and the currents determine the waves that drive the sailor to continuously adjust the rudder so as to stay on course. By observing the wave patterns, he can anticipate the navigational effects of his actions and adapt accordingly. But when the currents or the weather changes, the sailor must adapt his course to reach the next port of call. At that time, the sailor observes essentially the same stimulus (the waves) but has to remap his action plan (rudder adjustments) to the new wind conditions and currents.

This difficult-decision problem—how to detect and then adapt to a nonstationary environment—is captured perfectly in the exploration-exploitation dilemma: When

Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. Email: glaescher@uke.de; s.steixner-kumar@uke.de

Published by AAAS

should I stop exploiting my current action plan and start exploring different ways to reach my goals? An optimal solution tracks the discounted sum of normalized future rewards. However, this approach applies strictly to stationary environments and thus does not capture the dynamic changes that organisms encounter in their daily lives (*2*). Yet the human brain and those of other species seem to smoothly solve the exploration-exploitation dilemma in nonstationary environments.

Decision neuroscience has investigated the flexible adaptation to changing environmental contingencies with diverse experimental paradigms and assorted computational models. The simplest paradigm is probabilistic reversal learning, in which the agent has to search for reward among two options with complementary reward probabilities. This adaptation problem can be solved by hidden Markov models (*3*), which are well-approximated by reinforcement learning (RL) models that also update nonchosen actions (*4*). Extension of this paradigm to include independently changing reward probabilities reveals two distinct neural responses: Expected-value signals, which reflect "exploitative" choices, spur activation of the ventromedial prefrontal cortex (vmPFC); and "explorative" choices (that is, the choosing of a currently lesser valued option) activate the frontopolar cortex (*5*).

Another task with both rapid and slow changes in the reward probabilities of various options was used to develop a hierarchical Bayesian model that estimates the volatility of the environment and adjusts the learning rate accordingly (*6*). This model has found its generalization in the hierarchical Gaussian filter (HGF) framework (*7*), which is widely used in modeling social and nonsocial human decision-making in nonstationary environments. Although these computational modeling frameworks differ, all are trying to solve similar problems: How to infer the latent structure of the world from discrete observations and how to detect transitions between different states of the world.

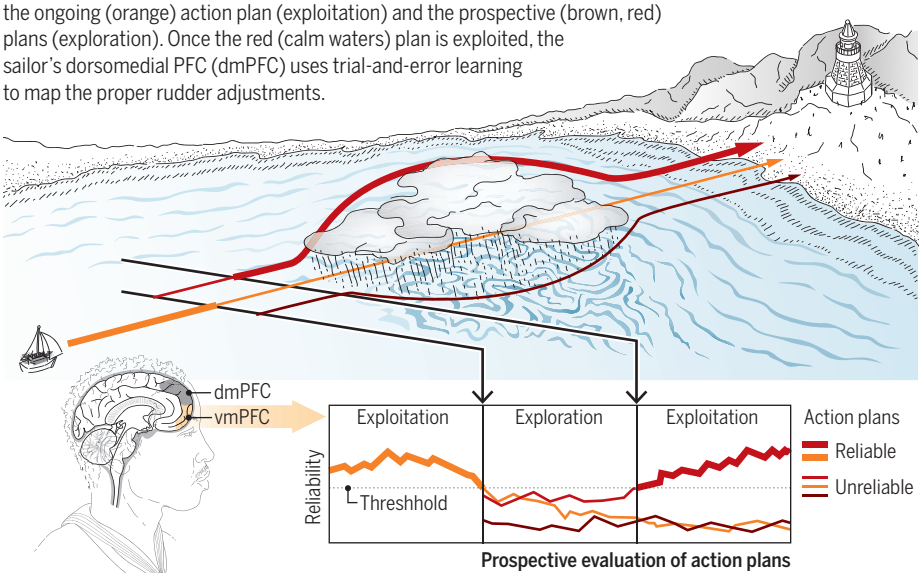
Domenech *et al*. address the same problems with yet another experimental paradigm, this one carried out with a small group of human epilepsy patients. Electrodes deeply implanted in the patients' PFCs delivered direct electrical recordings from the vmPFC and dorsomedial PFC (dmPFC) while the patients performed a multioption decision task. The participants had to associate three different stimuli with three distinct actions, thus constituting an action plan. The mapping changed every 33 to 57 trials, and participants had to relearn the association of the same stimuli with a different combination of actions, much like our sailor at sea who faces changes in weather and currents that alter wave patterns. The computational model (*8*) generates a reliability value for the ongoing action plan and other concurrently monitored plans. When the ongoing action plan is deemed reliable, the model is in "exploitation" mode and learns the stimulus-action mapping through RL mechanisms. When the ongoing action plan is deemed unreliable, the model switches to "exploration" mode. New provisional action plans are created and evaluated, until one emerges as a reliable predictor for successful stimulus-action mapping (see the figure).

Using a state-of-the-art model-based analysis that associates the model-derived variables with the brain activity in various frequency bands of the neural recordings, the authors found a delicate interplay between the vmPFC and dmPFC that supports a predictive coding interpretation for resolution of the exploration-exploitation dilemma. vmPFC monitors and represents the reliability of the ongoing action plan. vmPFC relays the ongoing action plan to the dmPFC as either a "stay" or "switch" trial. A stay trial triggers additional learning through RL mechanisms in the dmPFC. In contrast, the dmPFC responds to a switch trial by suppressing activity related to maintaining the ongoing action plan. These findings resonate with and extend earlier results obtained with functional neuroimaging (*5, 9*).

These computational approaches to the problem of behavioral flexibility in a nonstationary environment share one commonality: They are all building a model of the environment and the transition therein, either explicitly (as in the HGF framework) or implicitly (by evaluating the ongoing action plan, as in the Domenech *et al*. study). Although all of these models strive for generality, each was developed for a specific experimental context. It remains to be seen which of these provides the best account of flexible decision-making in humans and other species, preferably using a unified experimental paradigm. A model-free RL account (*10*) likely will not suffice, as several studies have demonstrated the superiority of more-complex models over this "vanilla" RL model. Rather, an agent requires a rich representation of the environment and its dynamic transitions (often referred to as model-based learning) (*10*) to solve the exploration-exploitation dilemma and flexibly respond to a changing world. ■

A sailor solves a dilemma at sea

As the ship nears bad weather, the sailor's ventromedial prefrontal cortex (vmPFC) evaluates the ongoing (orange) action plan (exploitation) and the prospective (brown, red) plans (exploration). Once the red (calm waters) plan is exploited, the sailor's dorsomedial PFC (dmPFC) uses trial-and-error learning to map the proper rudder adjustments.



GRAPHIC: A. KITTERMAN/*SCIENCE*

REFERENCES AND NOTES

1. P. Domenech, S. Rheims, E. Koechlin, *Science* **369**, eabb0184 (2020).
2. J. D. Cohen, S. M. McClure, A. J. Yu, *Philos. Trans. R. Soc. London Ser. B* **362**, 933 (2007).
3. A. N. Hampton, P. Bossaerts, J. P. O'Doherty, *J. Neurosci.* **26**, 8360 (2006).
4. J. Gläscher, A. N. Hampton, J. P. O'Doherty, *Cereb. Cortex* **19**, 483 (2009).
5. N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, R. J. Dolan, *Nature* **441**, 876 (2006).
6. T. E. J. Behrens, M. W. Woolrich, M. E. Walton, M. F. S. Rushworth, *Nat. Neurosci.* **10**, 1214 (2007).
7. C. Mathys, J. Daunizeau, K. J. Friston, K. E. Stephan, *Front. Hum. Neurosci.* **5**, 39 (2011).
8. A. Collins, E. Koechlin, *PLOS Biol.* **10**, e1001293 (2012).
9. M. Donoso, A. G. E. Collins, E. Koechlin, *Science* **344**, 1481 (2014).
10. N. D. Daw, P. Dayan, *Philos. Trans. R. Soc. London Ser. B* **369**, 20130478 (2014).

10.1126/science.abd7258

Published by AAAS

COSMOCHEMISTRY

The origins of water

New measurements of Earth's building blocks point to a simpler water source

By **Anne H. Peslier**

Our blue planet having water seems such a simple and obvious fact that the question of why Earth has water at all feels like a trivial one. However, the origin of this key ingredient for life has remained a long-standing topic of debate. According to models of Solar System formation, Earth, as an inner Solar System planet, should have little to no water. On page 1110 of this issue, Piani *et al.* (*1*) analyze enstatite chondrite meteorites, a material similar to Earth's main building blocks, and address the origins of Earth's water.

Early models of planetary formation predicted that the nebular gas near our young Sun was too hot to form ice. Water as vapor therefore could not have been easily incorporated into the materials that formed the inner rocky planets Mercury, Venus, Earth, and Mars (*2*). Only the outer Solar System planets, such as Jupiter, contain an abundance of water and other volatiles, because they formed beyond the "snow line," the imaginary boundary between water vapor and water ice in the solar nebula (*2*, *3*). These planets were born far enough from the Sun to be able to easily incorporate frozen water.

Earth not only has water in its oceans and atmosphere but also contains the equivalent of several oceans' worth of water locked inside the rocks of its deep interior (*4*). To solve this puzzle, special types of meteorites that formed at the very beginning of the Solar System can be analyzed. These space rocks, called chondrites, have not been incorporated and chemically modified into a large planet. Chondrites represent condensates of the solar nebular gas and are the best representatives of the type of material that aggregated to eventually form planets. Earth's isotopic composition indicates that the main building blocks of rocky planets are enstatite chondrite–like materials [e.g., (*5*)].

However, enstatite chondrites were believed to be too dry to account for Earth's water. Instead, a water-rich type of chondrite called carbonaceous chondrite was suggested to have brought most of the planet's water. Carbonaceous chondrites, however, are not ideal because they have a very different isotopic composition than Earth and formed outside of the snow line (*6*). To reconcile these apparent inconsistencies, carbonaceous chondrite–like material was proposed to have been added to Earth from beyond the snow line when the planet was already mostly formed [e.g., (*4*)]. For these chondrites to pelt the early Earth, either the formation of Jupiter or its inner Solar System wandering has been suggested to have disturbed the orbits of asteroids, ejecting them from the outer part of the asteroid belt (*6*).

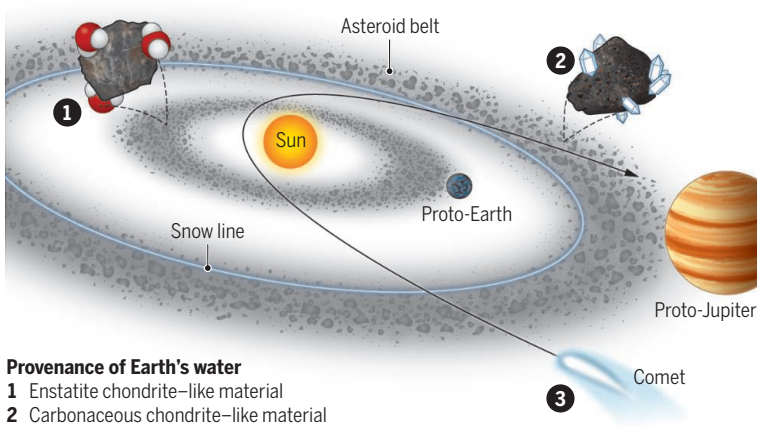
Piani *et al.* offer a simpler explanation. The authors convincingly argue that water could come from enstatite chondrites (see the figure). The authors measured the water concentration and deuterium to hydrogen (D/H) ratios of these meteorites. Enstatite chondrites apparently do have enough water to provide Earth with the mass equivalent of several surface oceans' worth of water. In addition, the D/H ratio and nitrogen isotope composition match those of Earth's interior.

It would now seem that Earth could have acquired its water during accretion after all. Alternative models suggest the possibility of retaining water in the inner planets' building blocks. Surface adsorption of water on nebular condensates (*7*), retention of nebular gas around planet embryos (*8*), or inward migration of the snow line (*2*) have all been proposed. That enstatite chondrites formed in the inner Solar System (*5, 6*) relaxes the challenging requirement of bringing water to Earth from farther away from the Sun.

Piani *et al.*'s observations do not provide any constraints on timing for the incorporation of water into Earth. Early incorporation during accretion could be irrelevant if most water degassed during early events in the planet's history. Formation of magma oceans and intense asteroid bombardment would both cause water to be lost to space (*9, 10*). Alternatively, enstatite chondrite–like material could have been incorporated into Earth after accretion (*11*). After Earth's accretion, it appears that water was added to its oceans and atmosphere (not its interior), with contributions from comets and carbonaceous chondrite–like material (*3*, *12*). Nevertheless, the authors' work brings a crucial and elegant element to this puzzle. Earth's water may simply have come from the nebular material from which the planet accreted. ■

Sources of Earth's water

New measurements of enstatite chondrites indicate that water could have been primarily acquired from Earth's building blocks. Additional water was delivered to Earth's early oceans and atmosphere by water-rich material from comets and the outer asteroid belt.



Provenance of Earth's water
1 Enstatite chondrite–like material
2 Carbonaceous chondrite–like material
3 Comets contributing to the atmosphere's composition

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

Jacobs, NASA Johnson Space Center, Houston, TX 77058, USA. Email: anne.h.peslier@nasa.gov

REFERENCES AND NOTES

1. L. Piani *et al.*, *Science* **369**, 1110 (2020).
2. L. Hartmann, F. Ciesla, O. Gressel, R. Alexander, *Space Sci. Rev.* **212**, 813 (2017).
3. O. Grasset, J. Castillo-Rogez, T. Guillot, L. N. Fletcher, F. Tosi, *Space Sci. Rev.* **212**, 835 (2017).
4. A. H. Peslier, M. Schönbächler, H. Busemann, S.-I. Karato, *Space Sci. Rev.* **212**, 743 (2017).
5. N. Dauphas, *Nature* **541**, 521 (2017).
6. P. H. Warren, *Earth Planet. Sci. Lett.* **311**, 93 (2011).
7. M. J. Drake, *Meteorit. Planet. Sci.* **40**, 519 (2005).
8. Z. D. Sharp, *Chem. Geol.* **448**, 137 (2017).
9. L. T. Elkins-Tanton, *Annu. Rev. Earth Planet. Sci.* **40**, 113 (2012).
10. H. Lammer *et al.*, *Astron. Astrophys. Rev.* **26**, 2 (2018).
11. M. Fischer-Gödde, T. Kleine, *Nature* **541**, 525 (2017).
12. B. Marty *et al.*, *Earth Planet. Sci. Lett.* **441**, 91 (2016).

10.1126/science.abc1338

Published by AAAS

RETROSPECTIVE

Zena Werb (1945–2020)

Pioneering cancer biologist and beloved mentor

By **Nancy Boudreau**[1] *and* **Mina Bissell**[2]

Zena Werb, renowned cancer biologist, passed away on 16 June. She was 75 years old. Zena was recognized internationally in the fields of proteolysis, development, and inflammation in breast cancer. Her studies of the development of the mammary gland and its neoplastic conversion during the initiation and progression of breast cancer revolutionized the fields of development and tumor biology and metastasis. Zena also served as a mentor and role model for myriad scientists.

Zena was born in the Bergen-Belsen concentration camp in Germany in March 1945. Her parents had been forced to separate during World War II, but after a fortunate reunion, they moved to Canada in 1948. Zena earned a bachelor's degree in biochemistry and physiology from the University of Toronto in 1966 and a Ph.D. in cell biology from the Rockefeller University in New York City in 1971. She conducted postdoctoral research at the Strangeways Research Laboratory in Cambridge, United Kingdom, where she studied matrix metalloproteinases (MMPs)—a type of enzyme that breaks down proteins—and then she taught briefly at Dartmouth College in New Hampshire. In 1976, she launched her own laboratory at the University of California, San Francisco (UCSF), where she would spend the rest of her career, most recently as professor and vice-chair in the department of anatomy and the associate director for basic science at the UCSF Helen Diller Family Comprehensive Cancer Center.

Through her pioneering work, Zena demonstrated the critical roles of MMPs in tissue functions during development and in disease. In the late 1970s and early 1980s, emerging evidence suggested that the extracellular proteins, which were targets of the lytic activity of MMPs, were providing not just structural support for cells but also cues that directly influence cellular signaling and behavior. These interactions between cells and extracellular proteins were suspected to be both dynamic and reciprocal. Zena elegantly showed that when cells bind to the extracellular protein fibronectin through a specific cell-surface receptor, intracellular signaling changes and the cell's production of MMPs increases. This critical piece of data demonstrated the bidirectional nature of interactions between cells and their extracellular environment (i.e., dynamic reciprocity).

Knowing that a complete understanding of MMPs would require the study of the complex biology of a variety of organs and conditions, Zena immersed herself in the biology of bone, mammary and salivary glands, embryonic development, wound healing, and then cancer. When the cancer field was coalescing around the notion that MMP inhibitors might prevent tumor metastasis, Zena was one step ahead. She knew that this strategy would wreak havoc on normal tissue function and potentially exacerbate the disease and also that metastasis was much more involved than simple protease-assisted breaching of basement membranes.



PHOTO: SONYA YRUEL FOR UCSF

Zena's research provided a molecular and mechanistic framework for understanding how the extracellular matrix contributed to tissue morphogenesis and remodeling and the tumor microenvironment. Her groundbreaking proposals required both courage and insight. In recognition of her pioneering work, she received the Excellence in Science Award from the Federation of American Societies for Experimental Biology, the E. B. Wilson Medal awarded by the American Society for Cell Biology (ASCB) "for far-reaching contributions to cell biology over a lifetime in science," and the ASCB Sandra K. Masur Senior Leadership Award. In 2010, Zena was elected to the National Academy of Sciences.

I (M.B.) became a close friend and collaborator of Zena's in 1987. I (N.B.) met Zena when I was a postdoctoral fellow in M.B.'s lab; Zena was a co-mentor for our work on cell death and the extracellular matrix. We both remember how excited Zena was about every experiment, regardless of success or failure. She had a passion for exploration and an ability to distill highly complex biological questions into manageable and testable hypotheses.

Sharing knowledge with students was a thrill for Zena. When she taught, her mesmerizing encyclopedic knowledge was on full display. Those who did not know Zena personally were sometimes intimidated by her direct style, which was evident in her keynote lectures and her questions to other presenters. However, her eagerness to ask the difficult questions and her expectation of thoughtful answers were rooted in her passion for good science.

Those of us who were fortunate to know Zena personally found her to be kind-hearted and generous. She was never too busy to provide advice to the many who sought her input. She was sincerely interested in her mentees and loved to engage in lively discussions with colleagues, always showing an interest that extended beyond their ideas to their personal well-being. Her support for students and colleagues was well recognized: The American Association for Cancer Research and its Women in Cancer Research member group gave her the Charlotte Friend Memorial Lectureship for her meritorious contributions to the field of cancer research and her work advancing women in science. UCSF also honored her with a Lifetime Achievement in Mentoring Award.

Zena's extraordinary career and research were complemented by her humility and humanity. She will be remembered for her original, creative, and fearless thinking, which led to her seminal work demonstrating how the extracellular environment influences the function of normal and malignant cells. Her rich legacy supporting the careers of others will live on in the scientists she mentored. ■

[1]Former Chair, Tumor Metastasis Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [2]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94710, USA. Email: mjbissell@lbl.gov

10.1126/science.abe0952

Published by AAAS

POLICY FORUM

SOCIAL SCIENCE

Computational social science: Obstacles and opportunities

Data sharing, research ethics, and incentives must improve

By **David M. J. Lazer[1,2], Alex Pentland[3], Duncan J. Watts[4], Sinan Aral[3], Susan Athey[5], Noshir Contractor[6], Deen Freelon[7], Sandra Gonzalez-Bailon[4], Gary King[2], Helen Margetts[8,9], Alondra Nelson[10,11], Matthew J. Salganik[12], Markus Strohmaier[13,14], Alessandro Vespignani[1], Claudia Wagner[14,15]**

The field of computational social science (CSS) has exploded in prominence over the past decade, with thousands of papers published using observational data, experimental designs, and large-scale simulations that were once unfeasible or unavailable to researchers. These studies have greatly improved our understanding of important phenomena, ranging from social inequality to the spread of infectious diseases. The institutions supporting CSS in the academy have also grown substantially, as evidenced by the proliferation of conferences, workshops, and summer schools across the globe, across disciplines, and across sources of data. But the field has also fallen short in important ways. Many institutional structures around the field—including research ethics, pedagogy, and data infrastructure—are still nascent. We suggest opportunities to address these issues, especially in improving the alignment between the organization of the 20th-century university and the intellectual requirements of the field.

We define CSS as the development and application of computational methods to complex, typically large-scale, human (sometimes simulated) behavioral data (*1*). Its intellectual antecedents include research on spatial data, social networks, and human coding of text and images. Whereas traditional quantitative social science has focused on rows of cases and columns of variables, typically with assumptions of independence among observations, CSS encompasses language, location and movement, networks, images, and video, with the application of statistical models that capture multifarious dependencies within data. A loosely connected intellectual community of social scientists, computer scientists, statistical physicists, and others has coalesced under this umbrella phrase.

MISALIGNMENT OF UNIVERSITIES

Generally, incentives and structures at most universities are poorly aligned for this kind of multidisciplinary endeavor. Training tends to be siloed. Integrating computational training directly into social science (e.g., teaching social scientists how to code) and social science into computational disciplines (e.g., teaching computer scientists research design) has been slow. Collaboration is often not encouraged, and too often is discouraged. Computational researchers and social scientists tend to be in different units in distinct corners of the university, and there are few mechanisms to bring them together. Decentralized budgeting models discourage collaboration across units, often producing inefficient duplication.

Research evaluation exercises such as the United Kingdom's Research Excellence Framework, which allocate research funding, typically focus within disciplines, meaning that multidisciplinary research may be less well recognized and rewarded. Similarly, university promotion procedures tend to underappreciate multidisciplinary scholars. Computational research infrastructures at universities too often cannot fully support analysis of large-scale, sensitive data sets, with the requirements of security, access to a large number of researchers, and requisite computational power. To the extent these issues have been partially resolved in the academy (e.g., with genomic data), lessons have not fully made their way into practice in CSS.

INADEQUATE DATA-SHARING PARADIGMS

Current paradigms for sharing the kinds of large-scale, sensitive data used in CSS offer a mixed bag. There have been successes built on partnerships with government, especially in economics, from the study of inequality (*2*) to the dynamics of labor markets (*3*). There are emerging, well-resourced models of administrative data research facilities serving as platforms for analyzing microlevel data while preserving privacy (*4*). These offer important lessons for potential collaboration with private companies, including the development of methodologies to keep sensitive data secure, yet accessible for analyses (e.g., innovations in differential privacy).

The value proposition for private companies is different and there has been predictably less progress. Data possessed by government agencies are held in trust for the public, whereas data held by companies are typically seen as a key proprietary asset. Public accountability inherent in sharing data is likely seen as a positive for the relevant stakeholders for government agencies, but generally, far less so for shareholders for private companies. Access to data from private companies is thus rarely available to academics, and when it is, it is typically granted through a patchwork system in which some data are available through public application programming interfaces (APIs), other data only by working with (and often physically in) the company in question, and still other data through personal connections and one-off arrangements, often governed by nondisclosure agreements and subject to potential conflicts of interest. An alternative has been to use proprietary data collected for market research (e.g., Comscore, Nielsen), with methods that are sometimes opaque and a pricing structure that is prohibitive to most researchers.

We believe that this approach is no longer acceptable as the mainstay of CSS, as pragmatic as it might seem in light of the apparent abundance of such data and limited resources available to a research community in its infancy. We have two broad concerns about data availability and access.

First, many companies have been steadily cutting back data that can be pulled from their platforms (*5*). This is sometimes for good reasons—regulatory mandates (e.g., the European Union General Data Protection Regulation), corporate scandal (Cambridge Analytica and Facebook)—however, a side effect is often to shut down avenues of potentially valuable research. The susceptibility of data availability to arbitrary and unpredictable changes by private actors, whose cooperation with scientists is strictly voluntary, renders this system intrinsically unreliable and potentially biased in the science it produces.

[1]Northeastern University, Boston, MA, USA. [2]Harvard University, Cambridge, MA, USA. [3]Massachusetts Institute of Technology, Cambridge, MA, USA. [4]University of Pennsylvania, Philadelphia, PA, USA. [5]Stanford University, Stanford, CA, USA. [6]Northwestern University, Evanston, IL, USA. [7]University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [8]University of Oxford, Oxford, UK. [9]The Alan Turing Institute, London, UK. [10]Institute for Advanced Study, Princeton, NJ, USA. [11]Social Science Research Council, New York, NY, USA. [12]Princeton University, Princeton, NJ, USA. [13]RWTH Aachen University, Aachen, Germany. [14]GESIS–Leibniz Institute for the Social Sciences, Cologne, Germany. [15]University of Koblenz-Landau, Landau, Germany. Email: d.lazer@northeastern.edu

Published by AAAS

Second, data generated by consumer products and platforms are imperfectly suited for research purposes (*6*). Users of online platforms and services may be unrepresentative of the general population, and their behavior may be biased in unknown ways. Because the platforms were never designed to answer research questions, the data of greatest relevance may not have been collected (e.g., researchers interested in information diffusion count retweets because that is what is recorded), or may be collected in a way that is confounded by other elements of the system (e.g., inferences about user preferences are confounded by the influence of the company's ranking and recommendation algorithms). The design, features, data recording, and data access strategy of platforms may change at any time because platform owners are not incentivized to maintain instrumentation consistency for the benefit of research.

For these reasons, research derived from such "found" data is inevitably subject to concerns about its internal and external validity, and platform-based data, in particular, may suffer from rapid depreciation as those platforms change (*7*). Moreover, the raw data are often unavailable to the research community owing to privacy and intellectual property concerns, or may become unavailable in the future, thereby impeding the reproducibility and replication of results.

INADEQUATE RULES

Finally, there has been a failure to develop "rules of the road" for scientific research. Despite prior calls to develop such guidance, and despite major lapses that undermined public trust, the field has failed to fully articulate clear principles and mechanisms for collecting and analyzing digital data about people while minimizing the potential for harm. Few universities provide technical, legal, regulatory, or ethical guidance to properly contain and manage sensitive data. Institutional Review Boards are still generally not attuned, and consistent in their response, to the distinct ethical challenges around digital trace data. The recent modification of the Common Rule in the United States, which concerns ethics of human subjects research, did not fully address these problems.

For example, in a networked world, how should we deal with the fact that sharing information about oneself intrinsically provides signals about those with whom one is connected? The challenges around consent highlight the importance of managing security of sensitive data and also of reimagining institutional review processes and ethical norms; yet few universities integrate infrastructure and oversight processes to minimize the risks of security lapses.

Cambridge Analytica, and other, similar, events, have engendered an impassioned debate around data sovereignty. Battle lines have been drawn between privacy advocates and companies, where the former seek to minimize the collection and analysis of all individual data, whereas the latter seek to justify their collection strategies on the grounds of providing value to consumers.

Often missing in public debates are voices for policies that would encourage or mandate the ethical use of private data that preserves public values like privacy, autonomy, security, human dignity, justice, and balance of power to achieve important public goals—whether to predict the spread of disease, shine a light on societal issues of equity and access, or the collapse of the economy. Also often missing are investments in infrastructures in the academy that could power knowledge production and maintain privacy.

Resources and rules, incentives and innovations

Strengthen collaboration

- Develop enforceable guidelines in collaborations with industry around research ethics, transparency, researcher autonomy, and replicability.
- Develop secure data centers supplemented by an administrative infrastructure for granting access, monitoring outputs, and enforcing privacy and ethics rules.

New data infrastructures

- Develop large-scale, secure, privacy-preserving, shared infrastructures driven by citizen contributions of time and/or data. Capture the metadata that describe the collection process.
- Develop infrastructure to capture the dynamic, algorithm-driven behavior of the major platforms over time.
- Promote legal frameworks that allow and mandate ethical data access and collection about individuals and rigorous auditing of platforms.

Ethical, legal, and social implications

- Professional associations should help develop ethical guidelines.
- Large investments are needed to develop regulatory frameworks and ethical guidance for researchers.

Reorganize the university

- Develop structures that connect researchers having shared interests in computational approaches.
- Fundamentally reconceive graduate and undergraduate curricula.
- Reward collaboration across silos.
- Appoint faculty with multi-unit affiliations
- Physically collocate faculty from different fields
- Allocate internal funding to support multidisciplinary collaboration.
- Empower and enforce ethical research practices—e.g., centrally coordinated, secure data infrastructures.

RECOMMENDATIONS

In response to these problems, we make five recommendations.

Strengthen collaboration

Despite the limitations noted above, data collected by private companies are too important, too expensive to collect by any other means, and too pervasive to remain inaccessible to the public and unavailable for publicly funded research (*8*). Rather than eschewing collaboration with industry, the research community should develop enforceable guidelines around research ethics, transparency, researcher autonomy, and replicability. We anticipate that many approaches will emerge in coming years that will be incentive compatible for involved stakeholders.

The most widespread and longest-standing model is open, aggregated data such as Census data. The aforementioned models developed to share government data, with an emphasis on security and privacy, offer promise in working with corporate data. The United Nations Sustainable Development Goals call for partnerships on public-private data sources to provide a wide variety of new, very rich neighborhood-by-neighborhood measures across the entire world (*9*), and national statistical offices in every corner of the world are quietly working on producing such products, but progress is slow owing to lack of funding. The development of secure administrative data centers supplemented by an administrative infrastructure for granting access, monitoring outputs, and enforcing compliance with privacy and ethics rules offers one model for moving forward. As noted above, this model has already been demonstrated in the domain of government administrative data; as well as in a few cases, by telecommunications and banking companies.

Similar models are rare—but becoming more common—for academic research. The Open Data Infrastructure for Social Science and Economic Innovations in the Netherlands is one example. Facebook has iterated through multiple models for collaboration with academics. In its early years, it focused on one-off collaborations, largely in-

Published by AAAS

formally negotiated. After the 2016 election, it launched Social Science One, providing access to aggregate data of news consumption, which, despite being well resourced, has faced challenges in providing data (*10*).

Coronavirus disease 2019 (COVID-19) has played a particular role in creating partnerships between researchers and companies to produce insights regarding the trajectory of the disease. (COVID-19 has, in many countries, including the United States, also illuminated the fractured and politically contingent nature of much public data regarding the disease.) Twitter has provided a streaming API regarding COVID-19 for approved researchers. Similarly, location data companies such as Cuebiq have provided access to anonymized mobility data. There remain open questions as to what extent these data-sharing efforts will continue after the disease settles into the history books and, if so, how to robustly align them with critical research norms in academia, such as transparency, reproducibility, replication, and consent.

The election examples with respect to Facebook highlight the potentially adversarial role between researchers and corporations. A central contemporary question for the field of CSS (as discussed below) is in what ways particular sociotechnical systems are playing positive and negative roles in society. This tension may partially (but not entirely) be resolved if companies feel that it is in their long-term interest to transparently study and anticipate these issues. Even in the most optimistic scenario, however, there will be a disjuncture between the public interest in the insights that research could produce, and corporate interests.

Academia, more generally, needs to provide carefully developed guidelines for professional practice. What control can companies have over the research process? It clearly is not acceptable for a company to have veto power over the content of a paper; but the reality of any data-sharing agreement is that there are negotiated domains of inquiry. What are the requirements for providing data for replication? What are the needs of researchers for access to a company's internal data management and curation processes?

New data infrastructures

Privacy-preserving, shared data infrastructures, designed to support scientific research on societally important challenges, could collect scientifically motivated digital traces from diverse populations in their natural environments, as well as enroll massive panels of individuals to participate in designed experiments in large-scale virtual labs. These infrastructures could be driven by citizen contributions of their data and/or their time to support the public good, or in exchange for explicit compensation. These infrastructures should use state-of-the-art security, with an escalation checklist of security measures depending on the sensitivity of the data. These efforts need to occur at both the university and cross-university levels. Finally, these infrastructures should capture and document the metadata that describe the data collection process and incorporate sound ethical principles for data collection and use. The Secure Data Center at the GESIS Leibniz Institute for the Social Sciences is an example of shared infrastructure for research on sensitive data. Further, it is important to capture the algorithm-driven behavior of the major platforms over time (*11*, *12*), both because algorithmic behavior is of increasing importance and because algorithmic changes create enormous artifacts in platform-based data collection. It is critical that legal frameworks allow and mandate ethical data access and collection about individuals and rigorous auditing of platforms.

Ethical, legal, and social implications

We need to develop ethical frameworks commensurate with scientific opportunities and emergent risks of the 21st century.

Social science can help us understand the structural inequalities of society, and CSS needs to open up the black box of the data-driven algorithms that make so many consequential decisions, but which can also embed biases (*13*). The Human Genome Project devoted more than $300 million as part of its Ethical, Legal, and Social Implications program "to ensure that society learns to use the information only in beneficial ways" (*14*). There are no off-the-shelf solutions on ethical research. Professional associations need to work on the development of new ethical guidelines—the guidelines developed by the Association of Internet Researchers offer one example of an effort to address a slice of the issue. Large investments, by public funders as well as private foundations, are needed to develop informed regulatory frameworks and ethical guidance for researchers, and to guide practice in the field in government and organizations.

Reorganize the university

Computation is adjacent to an increasing number of fields—from astronomy to the humanities. There needs to be innovation in the organization of typically siloed universities to reflect this, with the development of structures that connect diverse researchers, where collaborating across silos is professionally rewarded. Successful examples of institutional practice include the appointment of faculty with multi-unit affiliations (e.g., across computer science and social science disciplines) and of research centers that physically collocate faculty from different fields, as well as allocation of internal funding to support multidisciplinary collaboration. There needs to be a fundamental reconceiving of the development of undergraduate and graduate curricula for training a new generation of scientists. There must be pervasive efforts within the university to empower and enforce ethical research practices—e.g., centrally coordinated, secure data infrastructures.

Solve real-world problems

The preceding recommendations will require resources, from public and private sources, that are extraordinary by current standards of social science funding. To justify such an outsized investment, computational social scientists must make the case that the result will be more than the publication of journal articles of interest primarily to other researchers. They must articulate how the combination of academic, industrial, and governmental collaboration and dedicated scientific infrastructure will solve important problems for society—saving lives; improving national security; enhancing economic prosperity; nurturing inclusion, diversity, equity, and access; bolstering democracy; etc. Current applications of CSS in the global response to the pandemic are emblematic of the broader potential of the field. Beyond generating results that are meaningful outside of academia, the pursuit of this objective may also lead to more replicable, cumulative, and coherent science (*15*). ■

REFERENCES AND NOTES

1. D. Lazer *et al.*, *Science* **323**, 721 (2009).
2. R. Chetty, N. Hendren, P. Kline, E. Saez, *Q. J. Econ.* **129**, 1553 (2014).
3. J. J. Abowd, J. Haltiwanger, *J. Lane, Am. Econ. Rev.* **94**, 224 (2004).
4. A. Reamer, J. Lane, A Roadmap to a Nationwide Data Infrastructure for Evidence-Based Policymaking (2018); https://journals.sagepub.com/doi/abs/10.1177/0002716217740116.
5. D. Freelon, *Polit. Commun.* **35**, 665 (2018).
6. M. J. Salganik, Bit by Bit: Social Research in the Digital Age (Princeton Univ. Press, 2017).
7. K. Munger, *Soc. Media Soc* **5**, 205630511985929 (2019).
8. Social Science Research Council, To Secure Knowledge: Social Science Partnerships for the Common Good (2018); www.ssrc.org/to-secure-knowledge/.
9. IEAG, UN, "A World that Counts—Mobilising the Data Revolution for Sustainable Development." Independent Expert Advisory Group on a Data Revolution for Sustainable Development (2014).
10. G. King, N. Persily, "A New Model for Industry-Academic Partnerships" (Working Paper, 2018); http://j.mp/2q1lQpH.
11. A. Hannák *et al.*, in Proceedings of the 22nd International Conference on World Wide Web (ACM Press, New York, 2013), pp. 527–538.
12. I. Rahwan *et al.*, *Nature* **568**, 477 (2019).
13. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, *Science* **366**, 447 (2019).
14. J. E. McEwen *et al.*, *Annu. Rev. Genomics Hum. Genet.* **15**, 481 (2014).
15. D. J. Watts, *Nat. Hum. Behav.* **1**, 0015 (2017).

10.1126/science.aaz8170

Published by AAAS



The flags of Norway and Iceland adorn a replica Viking ship called *Gaia*.

PHOTO: ROBERT HARDING PICTURE LIBRARY/NAT GEO IMAGE COLLECTION

ARCHAEOLOGY

Secrets of the Vikings

The violent seafarers are brought to vivid life in a rich synthesis of old and new scholarship

By **Thomas H. McGovern**

Neil Price's *The Viking Way,* now in its second edition, was an instant hit in the Viking studies community when it published in 2002 and rapidly became a collector's item. Similarly, his edited volume with Stefan Brink, *The Viking World* (published in 2008), remains a valuable sourcebook. Price's latest book, *Children of Ash and Elm*, brings together multiple themes from previous scholarship and provides a narrative overview of new scholarship—which now spans disciplines from geoscience to saga studies—to provide previously untold stories and insightful perspectives on the Viking people.

Children of Ash and Elm expands the chronological boundaries of the Viking Age to include the late Roman sixth and seventh centuries and (in places) the early Middle Ages. Price does this by placing the rise of the Vikings in the context of Scandinavia's recovery from an unusually harsh winter in 535–536 CE, the impacts of the first plague pandemic (the Justinianic plague; 541–549 CE), and the rise in popularity of large dining halls as centers of cultural exchange during the Vendel era (550–790 CE), combining new geoscience data and paleoepidemiological observations with some well-chosen insights from *Beowulf.* Price also expands the usual geographic boundaries that defined the Viking era and does a fine job of connecting events in the Baltic, North Sea, Atlantic, and Mediterranean worlds to reveal how physical proximity and intercultural dynamics contributed to the Viking proclivities for alternately trading and raiding and their seemingly inconsistent practice of both building and burning urban centers.

Price is well known for making Vikings fun. Filed teeth, eye shadow, sex magic, bloody sacrifices, shamanic female sorcery, shape-changing berserkers, and warrior women are now firmly part of our view of the Vikings thanks to his efforts. This book will not disappoint if you are looking for more on any of these topics; however, Price's Vikings remain violent and lethal, and he is clear about their central role in the slave trade. He is likewise unflinching in confronting and debunking the invocations of the Viking legacy by Nazis and white supremacists.

An ongoing challenge for Viking Age researchers is the need to integrate rich Icelandic literary records composed centuries after the Vikings with contemporary written sources, the ever-growing archives of archaeological records of sites and finds, the new perspectives derived from the analysis of ancient DNA and stable isotopes, and the opportunities for fresh reinterpretation of older scholarship. Price does an especially good job of combining different contemporary accounts (and making excellent use of Islamic sources) with art and archaeological evidence.

Price also offers fresh insights into Viking burial rituals, combining findings derived from dendrochronology, ancient DNA, and isotopic sourcing of human remains. We now know, he reveals, that one of the women buried in state in the Oseberg ship—a well-preserved Viking ship discovered in a large burial mound outside Tønsberg in Norway—was from the Caspian and that the ship mound was left half-opened so that mourners could offer an animal sacrifice before it was hurriedly nailed shut.

Price posits that Viking burial rituals represented a form of performance art. His argument is compelling and makes sense of the extreme variability that we see in excavated burials across the Viking world. His exploration of the concept of hydrarchy—the notion that one might gain power over lands by dominating the seas—and of raiding brotherhoods, drawn in part from studies of 17th- to 19th-century pirate bands, is also well developed and demonstrates the value of a wider historical view.

Its scholastic merits aside, *Children of Ash and Elm* is also a wonderful read, with prose that flows like poetry in places and modern analogs that inspire creative thinking. The only thing I found disappointing was the editorial decision to forgo references and footnotes. While the book includes an excellent, long bibliographic essay, this does not help the reader who is interested in tracking down the science and scholarship behind the stories. That fault aside, this volume would make an excellent textbook and a splendid introduction to the world of the Vikings for any reader. ■



Children of Ash and Elm: A History of the Vikings
Neil Price
Basic Books, 2020.
624 pp.

The reviewer is at the Department of Anthropology, Hunter College, New York, NY 10065, USA, and the Graduate Center, City University of New York, New York, NY 10016, USA.
Email: thomas.h.mcgovern@gmail.com

10.1126/science.abb3580

Published by AAAS

SCIENCE LIVES

A history of insatiable intellectuals

Proficient in a range of topics, polymaths are overlooked as often as they are celebrated

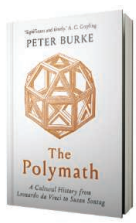
By **Andrew Robinson**

The development of science from the mid-19th century abounds in specialists, observes veteran cultural historian Peter Burke in his new book, *The Polymath*. Think of Louis Pasteur in medicine—the field that first embraced specialization—or Gregor Mendel in genetics, Marie Curie in physics, Edwin Hubble in astronomy, or Dorothy Hodgkin in chemistry.

Yet despite the trend toward specialization since that time, polymaths, whose expertise spans a range of subjects and who flourished in previous centuries, have remained vital to science, if fewer in number. Charles Darwin and Alan Turing come to mind. As does Linus Pauling, who contributed not only to physical chemistry and mathematical physics but also to biology, medicine, and international peace—for which he received two Nobel Prizes. Indeed, Pauling came close to discovering the structure of DNA in the 1950s but failed to achieve his goal, "perhaps because he was distracted by his other interests."

Burke names "500 western polymaths" in the appendix to his mind-stretching and thought-provoking history, which spans the half-millennium that followed the life of the most wide-ranging polymath of all, Leonardo da Vinci. Like da Vinci, their collective work embraces both the sciences and the humanities. Some notable scientific figures among them include (in order of birth) Nikolaus Copernicus, Christopher Wren, Alexander von Humboldt, Mary Somerville, Ernst Haeckel, Henri Poincaré, Alfred Wegener, Erwin Schrödinger, Richard Feynman, Benoit Mandelbrot, and Oliver Sacks. Oddly omitted is the founder of modern nursing, Florence Nightingale, who was also distinguished for her work as a social reformer and statistician.

The first half of the book sketches the lives and work of many of these individuals, with longer sections dedicated to key figures. Here we learn of Thomas Young (1773–1829)—now regarded by many as the greatest polymath since da Vinci—whose tombstone in London's Westminster Abbey describes him as "eminent in almost every department of human learning." Formally trained as a physician in the 1790s, Young taught himself physics and philology. He discovered the phenomenon of astigmatism and first proposed the three-color theory of how the retina responds to light. He showed, in his famous double-slit experiment, that light could behave as both a particle and a wave, a remarkable observation that Richard Feynman declared "the heart of quantum mechanics" and its "only mystery." Young also named the Indo-European family of languages and took the first crucial steps in deciphering the Rosetta stone and the Egyptian hieroglyphs.



The Polymath: A Cultural History from Leonardo da Vinci to Susan Sontag
Peter Burke
Yale University Press, 2020. 352 pp.

Chapter 6 aims to analyze the personal characteristics of polymaths as a group, singling out their unusual powers of concentration, capacious memory, speed, imagination, restlessness, industriousness, and obsession with not wasting time. "[An] overdose of curiosity, long known as the *libido sciendi* and described by the polymath Francis Bacon as 'inquisitive appetite,' is surely the most general as well as the most obvious characteristic of the species," notes Burke.

Two further chapters consider the careers led by polymaths, fleshing out their familial relationships, their educations and working environments, and their reception by professional institutions and by specialists. For all its considerable insight into polymathy, however, at no point does the book address what distinguishes a polymathic genius from a specialist of comparable standing. Should Albert Einstein, for example, be considered a specialist or a polymath? Einstein receives just one glancing reference in *The Polymath*, yet it could be argued that the physicist showed some polymathic tendencies, particularly later in life when he began to engage more deeply in political affairs (to say nothing of his well-known musical ability).

There has always been a tension between specialization and polymathy. Universities and professions are chiefly organized for the benefit of specialists, not polymaths. Moreover, in addition to greater funding, specialists typically receive more recognition than polymaths do, as evidenced by the Nobel Prizes' emphasis on domain-specific advances. And yet, some of the greatest scientific discoveries and works of art have benefited from interdisciplinarity and even polymathy. As Thomas Young privately concluded, near the end of his underappreciated life, "It is probably best for mankind that the researches of some investigators should be conceived within a narrow compass, while others pass more rapidly through a more extensive sphere of research." ■

10.1126/science.abb7546

The reviewer is the author of *The Last Man Who Knew Everything* (Oneworld Publications, 2007) and *Genius: A Very Short Introduction* (Oxford Univ. Press, 2011). Email: andrew@andrew-robinson.org



A healthy dose of skepticism can help one spot misleading data.

PHOTO: ISTOCK.COM/WUTWHANFOTO

PODCAST

Calling Bullshit: The Art of Skepticism in a Data-Driven World
Carl T. Bergstrom and Jevin D. West
Random House, 2020. 336 pp.

There are three kinds of lies, the oft-quoted saying goes: lies, damned lies, and statistics. In a world drowning in data, statistical methods and other tools of scientific inquiry are increasingly being used to advance erroneous claims. This week on the *Science* podcast, evolutionary biologist Carl Bergstrom explains how to identify data-driven misinformation and disinformation.

10.1126/science.abd9788

Published by AAAS



The Vulnerable fishing cat (*Prionailurus viverrinus*) has been seen in an area India has cleared for development.

PHOTO: FELIS IMAGES/MINDEN PICTURES

Edited by **Jennifer Sills**

Development projects jeopardize India's forests

As climate change intensifies, India, one of the world's most megadiverse countries, has put its forests and wildlife in jeopardy. Between January and May, India's Ministry of Environment, Forest, and Climate Change gave environmental clearance (i.e., approval to proceed) to 73 projects located within 10 km of a forest, including some planned near forests with protected status (*1*). Protected areas and forests can effectively mitigate the effects of urbanization and anthropogenic pressure (*2*), but human land use along their boundaries reduces their effectiveness (*3*), especially if the activities include permanent infrastructure (*4*). Low- and middle-income countries like India should mitigate environmental damage caused by development by imposing more stringent rules and regulations for environmental clearances.

India's approved projects include industrial construction, roadways, mining, and new infrastructure. Such development could lead to permanent changes in land use, land cover, and topography. Environmental impact assessment reports for 23 of the proposals found species in the vicinity of the project area that were designated as Schedule I (the most vulnerable) in India's Wildlife (Protection) Act (*5*), yet environmental clearances were granted regardless. For example, one approved project [IA/AS/IND2/92824/2007 in (*1*)] outlines a plan to conduct extension drilling and test hydrocarbons 1.34 km from Dibru-Saikhowa National park (*6*). An environmental impact assessment report confirms the presence of species in the area such as the Critically Endangered white-rumped vulture (*Gyps bengalensis*) (*7*), the Vulnerable fishing cat (*Prionailurus viverrinus*) (*8*), and the Endangered western hoolock gibbon (*Hoolock hoolock*) (*9*), all of which are classified under Schedule I (*5*). Oil and fossil fuel extraction in forests drives wildlife mortality, habitat loss, and fragmentation as well as carbon sequestration (*10*), putting these species in further danger.

Development projects that threaten vulnerable species are antithetical to the United Nations' Sustainable Development Goals (*11*). Megadiverse countries like India should focus on sustainable development and formulate environment-centered development policies. Project approval decisions should rely on the results of mandatory and more stringent environmental impact assessments, which consider areas in the immediate vicinity of protected areas and forests in addition to the designated project area. More research is needed to assess the efficacy of the mitigation measures mentioned in environmental impact assessment reports. India could conserve forests and protect the habitats of vulnerable species by restricting development to urban areas and industrial zones.

Suvarna Khadakkar
RTM Nagpur University, Nagpur 440022, India.
Email: suvarnask17@gmail.com

REFERENCES AND NOTES

1. Ministry of Environment, Forest, and Climate Change, Government of India, Parivesh (2020); http://parivesh.nic.in/. To search for approved projects, click on the "Environment clearance" icon at the bottom of the page. Then click "Dashboard" on the blue bar. Select 2020 in the "Year of Submission" dropdown and search. Then click "More info" under the box labeled "EC granted." The resulting table provides details for 200 approved projects, 73 of which are within 10 km of a forest.
2. L. Naughton-Treves, M. Holland, K. Brandon, *Annu. Rev. Environ. Resour.* **30**, 219 (2005).
3. S. Leroux, J. Kerr, *Conserv. Biol.* **27**, 166 (2012).
4. A. Hansen *et al.*, *BioScience* **61**, 363 (2011).
5. Government of India, The Wildlife (Protection) Act, 1972 (1972); www.indiacode.nic.in/handle/123456789/1726.
6. Oil India Limited, "EIA for extension drilling and testing of hydrocarbons at 7 (seven) locations under Dibru-Saikhowa National Park area: NorthWest of Baghjan PML under Tinsukia District" (2018); http://environmentclearance.nic.in/writereaddata/EIA/28022019DSI931GBEIAReport-ERDLocationsText.pdf.
7. BirdLife International. *Gyps bengalensis* (The IUCN Red List of Threatened Species, 2017).
8. S. Mukherjee *et al.*, *Prionailurus viverrinus* (The IUCN Red List of Threatened Species, 2016).
9. W. Brockelman, S. Molur, T. Giessmann, *Hoolock hoolock* (The IUCN Red List of Threatened Species, 2019).
10. N. Jones, L. Pejchar, J. Kiesecker, *BioScience* **65**, 290 (2015).
11. United Nations, Department of Economic and Social Affairs, Sustainable Development (https://sdgs.un.org/#goal_section).

10.1126/science.abd4305

North African forests falling to charcoal

In North Africa, illegal man-made fires are on the rise (*1*), driven in large part by a growing demand for charcoal (*2*), which is

Published by AAAS

produced by burning wood in underground pits. In the past, local people collected dead wood on a small scale, and larger operations were regulated by the government (*3*). However, as the demand for charcoal and the price for which it sells have increased (*2*), illegal efforts to produce it have multiplied (*4*). North African countries must address the resulting fires to protect forest ecosystems, which play an important role in the welfare of both urban and rural people and harbor unique and sensitive biodiversity (*5*).

The religious feast of the sacrifice (Eid al-Adha)—an important celebration in the Islamic religion during which a sheep is sacrificed—increases the demand for charcoal, which is required for traditional meat grilling (*6*). This year in Algeria, the number of fires peaked on 27 July (4 days before Eid al-Adha) with 66 simultaneous fires in 20 provinces (*7*). The next few years could bring substantial environmental damage because Eid al-Adha will take place during the summer, when forest fires peak and become difficult to contain (*8*).

The business model for producing charcoal is ominous for the health of North Africa's forests because the revenue from the charcoal sold increases with the size of the exploited area. The fires are likely to deteriorate the soil, intensify desertification, and exacerbate climate change (*9*). A Mediterranean forest requires several decades to recover and reach its equilibrium (*10*). Climate change has already increased the total area burned by wildfire in the region (*11*), and the increased anthropogenic disturbance related to charcoal production could threaten the resilience of the environment, the economy, and human well-being (*12*).

Local authorities should increase enforcement to curtail illegal charcoal production, which often takes place at night and across wide geographic areas (*2*). Educating the public about the value of these forests and facilitating collaborations between scientists, the government, and rural and urban residents would also be effective strategies. If residents were motivated to report fires and other environmental crimes, authorities could take prompt action for the benefit of the environment.

Rassim Khelifa
Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
Email: rassimkhelifa@gmail.com

REFERENCES AND NOTES

1. O. Meddour-Sahar *et al.*, *Méditerranée* **121**, 33 (2013).
2. H. Faouzi, *Les Cahiers d'Outre-Mer* **262**, 155 (2013).
3. L. Auclair, M. R. Saïdi, *Forêt Méditerranéenne* **23**, 133 (2002) [in French].
4. H. Hadri, M. Guellouz, *Forests and Rangelands in the Near East Region: Facts and Figures* (FAO Office for the Near East, Cairo, 2011); www.fao.org/3/a-i12557e.pdf.
5. L. Croitoru, *For. Pol. Econ.* **9**, 536 (2007).
6. M. Gagaoua, H.-R. Boudechicha, *J. Ethnic Foods* **5**, 83 (2018).
7. "Incendies en Algérie: Plus de 8.000 hectares brûlés, le président ordonne une enquête," *Le Figaro* (2020) [in French].
8. O. Meddour-Sahar, R. Meddour, V. Leone, R. Lovreglio, A. Derridj, *iForest-Biogeosci. Forestry* **6**, 247 (2013).
9. R. Alkama, A. Cescatti, *Science* **351**, 600 (2016).
10. A. Rodrigo, J. Retana, F. X. Picó, *Ecology* **85**, 716 (2004).
11. C. Giannakopoulos *et al.*, *Glob. Planet. Change* **68**, 209 (2009).
12. D. M. Bowman *et al.*, *Science* **324**, 481 (2009).

10.1126/science.abe2315

Protect Iran's Zagros forests from wildfires

Since 21 March, more than 374 wildfires have spread through the oak forests in Iran's Zagros Mountains (*1*), burning more than 50,000 ha (*2*), compared with the average of 15,000 ha of Iranian forests that burn annually in wildfires (*3*). These deciduous forests, which include trees that are more than 400 years old (*4*), cover an area of 6 million hectares and are home to a wide variety of wildlife (*5*), including wolves, Persian squirrels, the endangered Persian leopard (*6*), and the Persian fallow deer, which was once thought to be extinct (*5*). Because the Zagros oak trees grow slowly (*2*), it could take 100 years for the burned areas to recover (*7*). The Zagros forests feed more than 50% of Iran's livestock (*8*) and include some of the country's most important riverine sources of fresh water (*8*). Given their valuable biodiversity and strategic importance, the Iranian government must address the recent wildfires.

The head of Iran's Department of Environment has announced that the magnifying effects of glass and plastic waste caused the wildfires (*9*). However, some environmentalists and researchers suspect that local people may be setting fires, either deliberately (*3*, *7*) or accidentally (*10*). Deforestation and illegal logging also contribute to wildfire ignition and spread (*5*, *7*). The Iranian government should clearly identify the causes of these frequent fires.

The Iranian forests are protected by the Forests, Range, and Watershed Management Organization, which has provincial centers monitoring their local territories (*11*). If deliberate illegal action is the cause of the increased wildfires, policy-makers should introduce national legislation to implement more stringent regulations against starting fires and to increase enforcement in the Zagros Mountains to prevent illegal logging and brush burning. The government should identify and prioritize the trees and endangered species that are at risk, improve fire planning, and streamline fire responses by learning from other countries with similar experiences, such as Australia (*12*). Without these actions, wildfires will continue to destroy trees, increase endangered species' risk of extinction, and deprive the Iranian people of crucial resources.

Mostafa Kheshti
Shandong University, Jinan, China.
Email: kheshti@sdu.edu.cn

REFERENCES AND NOTES

1. "Are the fires intentional?" *Aftab Yazd* (2020); www.pishkhan.com/news/183021/ [in Farsi].
2. "Ashes of the Zagros forests in the flames of negligence," *Islamic Republic News Agency* (2020); *www.irna.ir/news/83850837/* [in Farsi].
3. "Why are Zagros forests burning?" *Tabnak* (2020); www.tabnak.ir/fa/news/982140/ [in Farsi].
4. M. A. Haqshenas, "Zagros oak forests would cry out if they could!" *Tehran Times* (2018).
5. "Saving Zagros forests must become first priority," *Tehran Times* (2019).
6. A. B. Stein, "*Panthera pardus* (amended version of 2019 assessment)" (IUCN Red List of Threatened Species, 2020).
7. "Zagros in the fire of human fires," *Mashregh News* (2020); www.mashreghnews.ir/news/1089354/ [in Farsi].
8. G. Tavangar, "Homelessness of Zagros species with the fire of negligence," *Kayhan* (2020); www.kayhan.ir/fa/news/191035/ [in Farsi].
9. "Kalantari: Less than 1,500 hectares of forest have been set on fire," *Khabar Online* (2020); www.khabaronline.ir/news/1397174/ [in Farsi].
10. "The story of recent fires in Iran; Intentional or unintentional?" *Borna News* (2020); www.borna.news/fa/tiny/news-1006988/ [in Farsi].
11. "Executive policies of Zagros forest ecosystem management," *Zistonline* (2019); www.zistonline.com/news/80630/share/do.
12. R. M. Kooyman *et al.*, *Science* **367**, 1083 (2020).

10.1126/science.abd2967

ERRATA

Erratum for the Research Article: "Global wildlife trade across the tree of life," by B. R. Scheffers *et al.*, *Science* 369, eabd8164 (2020). Published online 24 July 2020; 10.1126/science.abd8164

Erratum for the Report: "Estimating the burden of SARS-CoV-2 in France," by H. Salje *et al.*, *Science* 368, eabd4246 (2020). Published online 26 June 2020; 10.1126/science.abd4246

Erratum for the Report "Forest microclimate dynamics drive plant responses to warming" by F. Zellweger *et al.*, *Science* 368, eabd3881 (2020). Published online 26 June 2020; 10.1126/science.abd3881

Erratum for the Report: "The global tree restoration potential" by J.-F. Bastin *et al.*, *Science* 368, eabc8905 (2020). Published online 29 May 2020; 10.1126/science.abc8905

Published by AAAS



U.S.-funded R&D projects drive innovation and science forward.

PHOTO: CHARLES A. SMITH/JSU

Increases in U.S. federal R&D needed in a global crisis

Support can boost the U.S. economy, innovation, and global standing

By **Anne Q. Hoy**

Imagine a world without the internet, a Google search engine, magnetic-resonance imaging, or the Human Genome Project—a sampling of American innovations that, like many scientific tools and research efforts, evolved from U.S. R&D investments.

Until Congress recently boosted federal R&D funding as a share of the U.S. economy, such investments had been on the decline—sliding from a high just shy of 1.9% of the gross domestic product in 1964 to 0.62% of the nation's gross domestic product (GDP) in 2018—impeding scientific advances, slowing innovation, and clipping the nation's share of global R&D funding at a time when the country faces challenges, economic and human, triggered by coronavirus disease 2019 (COVID-19).

The American Association for the Advancement of Science presented recommendations on how to advance scientific discovery, expand innovation, drive economic advancement, and ensure the scientific community supports opportunities for all at the "Fueling American Innovation and Recovery" hearing held by the U.S. House of Representatives Budget Committee on 8 July.

Sudip Parikh, AAAS CEO and executive publisher of the *Science* family of journals, was among the expert witnesses at the hearing that examined ways to reinvigorate U.S. economic competitiveness, renew federal investment in scientific R&D, and address the global pandemic.

Since 1995, the global ranking of U.S. R&D investment as a percentage of its GDP slipped from 4th to 10th place, said House Budget Committee Chairman John Yarmuth during the hearing.

The United States now lags behind competitors such as South Korea, Taiwan, Japan, and Germany in R&D investments, said Deborah L. Wince-Smith, president and CEO of the Council on Competitiveness, during her testimony at the House Budget Committee hearing.

Yarmuth warned that delays in scientific research projects as well as other economic challenges triggered by COVID-19 could diminish the U.S. position as a leader of global R&D funding.

Updating a long-standing framework that guides federal investment in scientific research, ensuring effective coordination of federal responses to the crises facing the nation, and committing to provide scientific evidence to propel racial equity in science and national policy-making formed the core of Parikh's recommendations to the panel.

Parikh pointed to research that found U.S. R&D as a share of GDP is well below its historic peak and below the current investment levels of nine other countries.

U.S. funded R&D projects drive innovation forward. The crises at hand require a federal investment of 1.9% of GDP, a level that represents an annual funding increase of 11% in scientific R&D, said Parikh.

Support for the full spectrum of innovation is needed, including "fundamental science, mission-driven technology, useful knowledge programs to meet local, national and international needs with the federal government as a key partner," said Parikh. New approaches to R&D funding models and networks should be explored, including "project grants, people-centered grants, teams and hubs, and prizes."

"Broadly, federal research is effective in producing discoveries that lead to high-impact, novel inventions, often in technology areas that have not yet received much industry attention," said Parikh.

AAAS tracks administration and congressional appropriations through the R&D Budget and Policy program, an outreach effort that keeps scientists and policy-makers informed through regular analyses, reports and media outreach. AAAS also engages in advocacy efforts to address executive branch and congressional actions that pose negative consequences for U.S. scientific research activities.

AAAS has issued, for instance, statements on the U.S. withdrawal from the World Health Organization, which collaborates across borders to support human health, and an administration proposal to limit the participation of international students at U.S. academic institutions despite a long history of important scientific contributions that foreign national students have made to the U.S. scientific enterprise.

"From the beginning, the Trump administration has taken a hard-fiscal line on most research and development programs, favoring Department of Defense technology development and acquisition at the expense of basic and applied research, even Defense research activities," noted Matthew Hourihan, director of the AAAS R&D Budget and Policy Program.

In fiscal year 2017, Congress rejected steep spending cuts outlined in the administration's budget proposal and instead adopted significant spending increases, particularly for the Department of Energy's Office of Science for basic science and R&D programs.

By fiscal year 2020, congressional R&D increases stood as a rebuttal to the president's consistent budget reduction proposals that sought more than $12 billion in spending to be shed from federal basic and applied research programs, according to a 2019 analysis by the AAAS R&D Budget and Policy Program.

Federal R&D investment supports government laboratories, research activities at federal agencies, academic institutions, and private-sector facilities to drive U.S. scientific advances.

Published by AAAS

AAAS NEWS & NOTES

The 2020 fiscal year spending package Congress approved dedicated the largest funding increases to the life sciences, particularly for the National Institutes of Health's basic research on human health and related topics, and made low-carbon energy and space exploration programs the second-largest funding recipients, as documented by the AAAS R&D Budget and Policy dashboard that tracks congressional appropriations trends across the scientific enterprise.

Despite the rise in the levels of federal R&D funding from 2000 to 2017, "the share of total U.S. R&D funded by the federal government declined from 25% to 22%," according to "The State of U.S. Science and Engineering," which highlights the 2020 "Science and Engineering Indicators," a suite of reports that provide findings on thematic scientific topics. The National Science Board is required to deliver the findings to Congress and the president every 2 years.

The business sector and U.S. academic institutions of higher education have stepped up financial support for R&D programs and activities, aware that R&D investments often spark novel scientific knowledge that, in turn, opens new research avenues, contributes to the training of young scientists, and helps fuel the U.S. economy. The business sector plays the most prominent role, having outpaced federal R&D funding to become primarily responsible for the rise in R&D support since 2000.

Universities and colleges are the second-largest contributors to R&D and play an important role in the progress of the nation's overall R&D activities by funding more than half of both U.S. basic research and the training of incoming scientists and engineers, according to the 2020 "Science and Engineering Indicators."

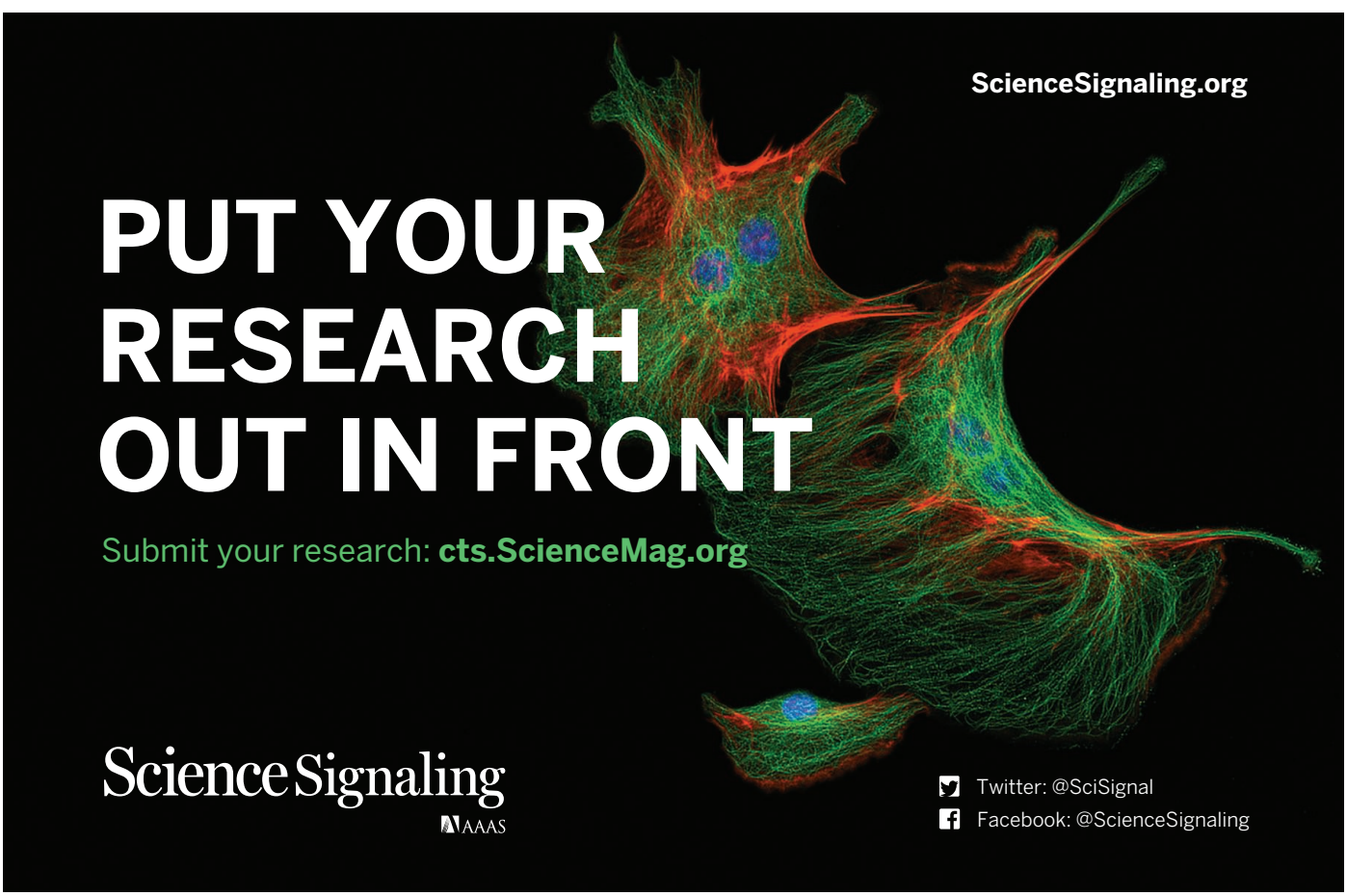
Yet, the combination of an overall decline in state support for public universities and colleges, and a leveling off of federal R&D funding for academic institutions at an annual $30 billion, risks a weakening of the United States' standing in innovation, reports suggest.

More encouraging, federal science and engineering support for historically black colleges and universities delivered HBCUs a 5.4% R&D funding increase for research and experimental development, according to the National Center for Science and Engineering Statistics.

Joining Parikh and Wince-Smith at the House hearing were two other expert witnesses, including Paul Romer, a Nobel-Prize winning economist and professor at New York University. Romer called on U.S. universities to give talented science and technology students a larger voice and leeway to pursue their own innovative ideas, not just those of professors. He said science and engineering degree fellowships should be granted to highly talented undergraduates to help expand U.S. innovation.

Economist Willy Shih, a professor at Harvard Business School, said the COVID-19 pandemic has exposed U.S. reliance on other nations for equipment, devices, and pharmaceuticals, a reality that requires expanded federal investment in basic scientific research and direct stimulus spending for technology investments.

"If the United States does not make needed investments in its future, increase its scope and rate of innovation, its fundamental capacity to grow its economy, create jobs, maintain national security, solve societal challenges and provide a social safety net, it will continue to erode—and its geopolitical leadership will be at increasing risk," said Wince-Smith.



RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



Bacteriophages infecting a bacterial cell, as seen in a colored transmission electron micrograph

MICROBIOLOGY

Prokaryotic enzymes for viral defense

The arms race between prokaryotes and viruses provides a strong evolutionary force to create diverse enzymatic activities that mediate antiviral immune responses. These immune components often cluster together in the host genomes, leading to expanded defense systems. Taking advantage of the evolutionary modularity of defense systems, Gao *et al.* bioinformatically predicted defense genes in most available bacterial and archaeal genomes. In addition, they reconstituted the newly identified systems in Escherichia coli and verified their defense functions against specific bacteriophages. In particular, they characterized defense functions for several predicted nucleoside triphosphatases. —SYM

Science, this issue p. 1077

STRUCTURAL BIOLOGY

Plant cell wall construction crew

Plants produce a complex cell wall in which cellulose, a glucose polymer, is a major component. Cellulose fibers are formed from close-packed single chains of cellulose that have been proposed to be formed by multimeric complexes (18 or more subunits) of the enzyme cellulose synthase, which exists in several isoforms. Purushotham *et al*. determined a cryo–electron microscopy structure of a trimer of a single isoform of cellulose synthase. A large channel forms a path for cellulose chains through the membrane-embedded complex. The structure also reveals oligomeric interfaces and provides a framework for modeling the larger complexes seen in plant membranes. The close arrangement of exit sites for nascent glycan chains is consistent with the enzyme complex playing a role in directing cellulose microfibril formation. —MAF

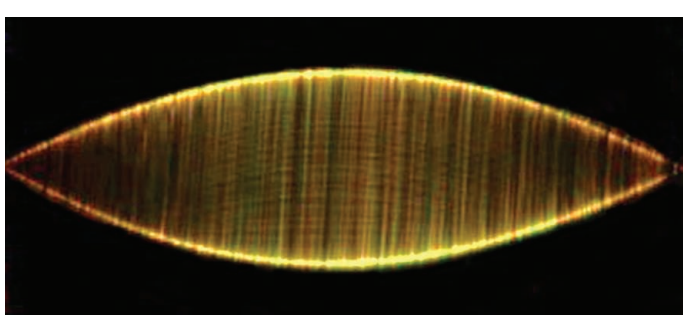
Science, this issue p. 1089

CARBON NANOTUBES

Finding the source of their durability

Many high-strength materials can fail after repeated exposure to subcritical loading; therefore, understanding the causes of fatigue is key to extending the longevity of the material in use. Bai *et al*. used chemical vapor deposition to produce centimeter-long carbon nanotubes. They then decorated these nanotubes with titanium dioxide nanoparticles for optical visualization and studied them using noncontact acoustic resonance tests. The authors determined that fatigue lifetime is dependent on the formation of the first defects on straining and is higher at lower temperatures. —MSL

Science, this issue p. 1104



An array of suspended, vibrating carbon nanotubes

ELECTROCHEMISTRY

Easing water apart

Traditional setups for splitting water into hydrogen and oxygen operate either in acid or in base. A bipolar membrane can potentially enhance efficiency by bridging acidic hydrogen evolution to basic oxygen evolution. Oener *et al*. undertook a systematic study of how catalysts paired with such a membrane might accelerate the preliminary step of water dissociation into protons and hydroxide ions. Using insights from this study for optimal catalyst integration, they were able to substantially lower the overpotential of a bipolar membrane electrolyzer. —JSY

Science, this issue p. 1099

IMAGES (FROM TOP): LEE D. SIMON / SCIENCE SOURCE; BAI *ET AL.*

Published by AAAS

IMMUNOGENOMICS

Teasing apart the mosquito immune system

Hemocytes are key immune cells of insects, playing a major role in how vector species transmit disease. Raddi *et al*. collected and sequenced the RNA of more than 8000 individual hemocytes from the disease-carrying mosquitoes *Anopheles gambiae* and *Aedes aegypti*. These data were then analyzed to determine hemocyte differentiation lineages and population changes in response to noninfected and *Plasmodium*-infected blood meals. From these data, the researchers identified a new hemocyte type, the megacyte, defined by the transcription of specific marker genes. Gene-silencing experiments implied that the megacyte functions in hemocyte differentiation during immune priming and thus may be involved in the immune response in mosquitoes. —LMZ

Science, this issue p. 1128

PROTEIN DESIGN

Exploring the design landscape

Protein design typically selects a protein topology and then identifies the geometries (secondary-structure lengths and orientations) that give the most stable structures. A challenge for this approach is that functional sites in natural proteins often adopt nonideal geometries. Pan *et al*. addressed this issue by exploring the diversity of geometries that can be sampled by a given topology. They developed a computational method called LUCS that systematically samples geometric variation in loop-helix-loop elements and applied it to two different topologies. This method generated families of well-folded proteins that include structures with non-native geometries. The ability to tune protein geometry may enable the custom design of new functions. —VV

Science, this issue p. 1132

BIOMATERIALS

Salubrious solution for the small intestine

Orally delivered drugs are primarily absorbed in the small intestine, which is also the major site of nutrient absorption. Li *et al.* developed a versatile, polydopamine-based solution that can coat and polymerize in situ in the small intestine. The synthetic coating adhered to pig and human tissue ex vivo, remaining stable for up to 24 hours. When administered to pigs endoscopically, polydopamine solution with suspended β-galactosidase increased digestive enzyme activity, whereas solution with embedded polydopamine nano–cross-linkers created a coating with transient barrier function that prevented glucose absorption. Encapsulating the anthelmintic drug praziquantel in polydopamine solution enhanced intestinal retention and absorption in pigs, demonstrating the potential utility of this biomaterial for drug delivery. —CC

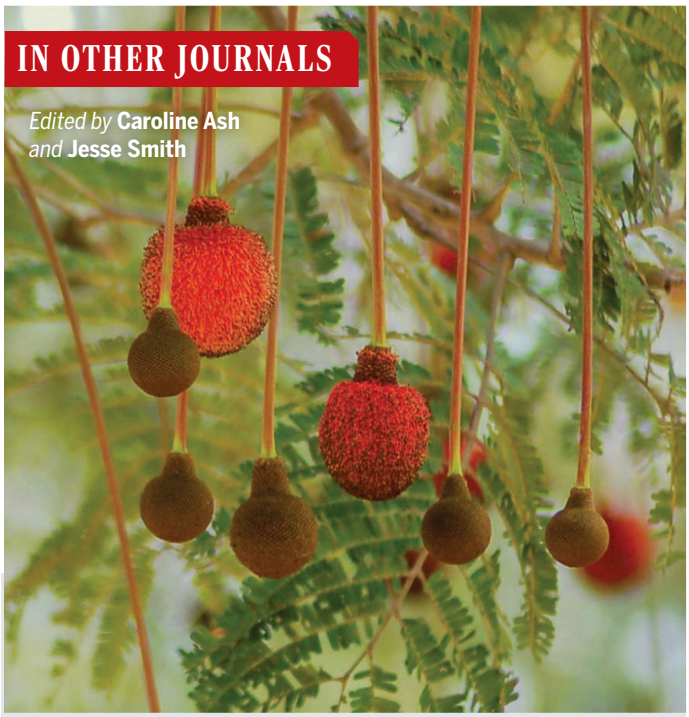
Sci. Transl. Med. **12**, eabc0441 (2020).

IMMUNOLOGY

Neuroimmune link in type 1 diabetes

Type 1 diabetes is an autoimmune disease in which T lymphocytes infiltrate pancreatic islets and destroy insulin-producing beta cells. It has long been known that this immune assault is heterogeneous, affecting different lobes of the endocrine pancreas in a patchy manner. Christoffersson *et al.* speculated that this heterogeneity might relate to the innervation of the islets. Using a mouse model, the authors found that cutting or depleting sympathetic neurons or inhibiting adrenoceptors prevented the T cell–mediated destruction of islets. Neuroimmune interactions with islet-resident macrophages controlled islet destruction and may represent a new target for type 1 diabetes. —DRG

Sci. Adv. 10.1126/sciadv.abb2878 (2020).



IN OTHER JOURNALS

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

PLANT SCIENCE

Population genetics of *Parkia*

In the West African savannah, the melliferous tree *Parkia biglobosa*, also known as the African locust bean, provides rural communities with food, fodder, and fuel, as well as tannins. Its pollinators include bats and honey bees, and its seeds are dispersed by humans, chimpanzees, rodents, and birds. This valuable tree is under threat from pesticide exposure and loss of habitat. Unfortunately, planting cotton, which does not thrive in shade, encourages the removal of trees and increasingly fragments *P. biglobosa*'s range. Lompo *et al*. analyzed polymorphic nuclear microsatellites to study the effects of fragmentation. The authors showed that distance favors genetic isolation and low population density favors inbreeding, but for the time being, cotton agriculture seems not to have diminished the genetic diversity of these long-lived trees. —PJH *Am. J. Bot.* **107**, 1041 (2020).

***Parkia biglobosa* flowers are rich in nectar, and its seeds, leaves, and gum are valuable products, but the tree is suffering from habitat fragmentation.**

CELL BIOLOGY

Tuning membrane tension

The tension of a cell's plasma membrane influences many processes that require membrane remodeling, including cell migration and endocytosis. However, little is known about how organelle membrane tension may regulate cell function. After endocytosis, endosomal membranes are remodeled in an ESCRT-III–dependent process that leads to the formation of intralumenal vesicles (ILVs). Mercier *et al*. investigated this phenomenon using a specific fluorescent membrane tension probe. ESCRT-III subunits were recruited to endosomal membranes as membrane tension reduced. Tension-dependent membrane recruitment was associated with ESCRT-III polymerization and membrane deformation and with increased ILV formation. In cultured cells, endosomal membrane tension decreased when ILV formation was triggered

PHOTO: TREEAID/FLICKR/CC BY

Published by AAAS

by epidermal growth factor. Thus, endosomal membrane tension regulates ILV formation, and reduced tension facilitates intracellular membrane-remodeling processes. —SMH

Nat. Cell Biol. **22**, 947 (2020).

STRUCTURAL BIOLOGY

Phosphorylation remodels amyloid

Abnormal aggregation of the protein α-synuclein (α-syn) is a hallmark of Parkinson's disease (PD) and several other neurodegenerative disorders. Disease-associated α-syn amyloid fibrils accumulate in Lewy bodies and are often phosphorylated. Phosphorylation at tyrosine 39 (Y39) increased neuropathology in a PD mouse model. Zhao *et al*. combined chemical synthesis and bacterial expression to prepare α-syn that could be site-specifically phosphorylated at Y39 (pY39). A structure determined by cryo–electron microscopy showed that in the pY39 fibril, the fibril core extends to include the entire N-terminal domain (the core ordinarily starts at around residue 40). Charged residues at the N terminus, including pY39, form a network of interactions in the center of the fibril core. This extended core may explain why pY39 fibrils are resistant to proteolysis, and the structure highlights how posttranslational modifications can affect the structure of fibrils, which in turn mediates their pathology. —VV

Proc. Natl. Acad. Sci. U.S.A. **117**, 20305 (2020).

ORGANIC CHEMISTRY

Boron needle stitches carbon thread

The Fischer-Tropsch process relies on transition metal catalysis at high temperature to stitch together carbon monoxide and hydrogen into hydrocarbon chains. Phanopoulos *et al*. now report that a pair of boron reagents can do something similar at room temperature, albeit less efficiently. Specifically, the combination of a terminal alkyl borane and triethylborohydride can catenate as many as eight carbon monoxide equivalents, producing a mixture of linear aldehydes after oxidative workup. Theoretical simulations and high-pressure nuclear magnetic resonance spectroscopy elucidated the catalytic role of the terminal boron center. —JSY

J. Am. Chem. Soc. **142**, 14064 (2020).

STELLAR ASTROPHYSICS

Identifying disrupted star clusters

Stars form in clusters, each produced from the gas in a single molecular cloud. Some clusters are massive enough to remain bound together, but most gradually dissipate as their member stars spread through the Milky Way. Price-Jones *et al*. identified groups of stars that are widely separated in space but have very similar composition, suggesting that they formed together in a cluster. Based on the abundances of eight chemical elements, they identified 21 candidate former star clusters. Stellar ages, kinematics, and another seven elements were used to support the identifications. The method could be used to unpack the Milky Way's star formation history. —KTS

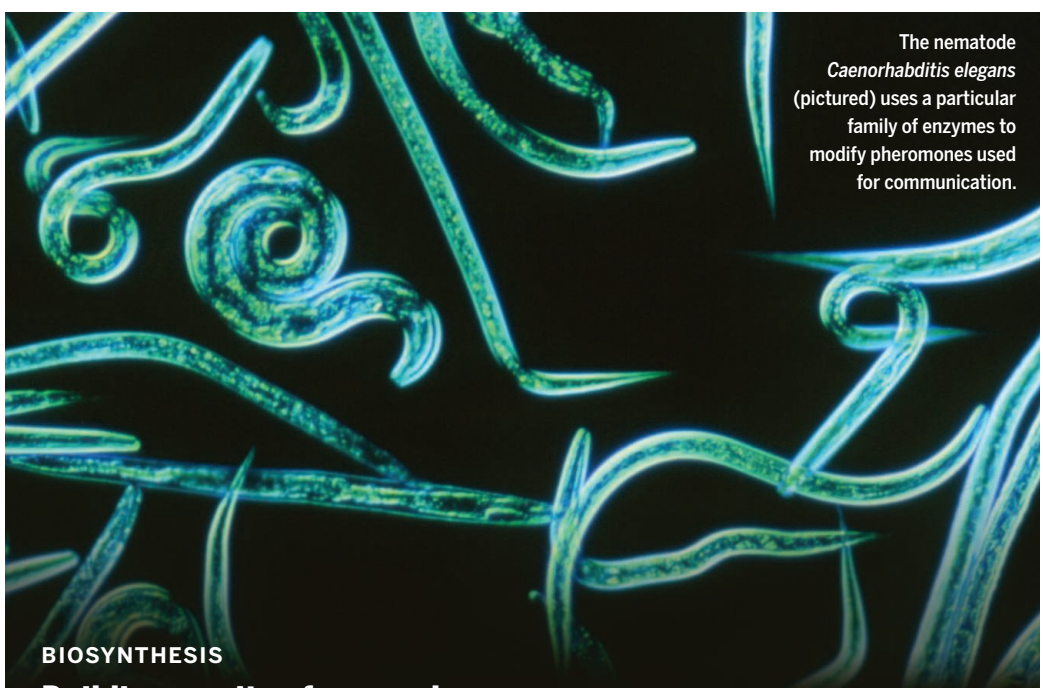
Mon. Not. R. Astron. Soc. **496**, 5101 (2020).

HOST DEFENSE

Immune escape by polyphosphate?

Polyphosphates are inorganic linear polymers with diverse biological functions. In mammals, low–molecular weight short-chain polyphosphates are sequestered within platelets and mast cell secretory dense granules and act as signaling molecules on release. Bacteria such as *Escherichia coli* can also synthesize high–molecular weight long-chain polyphosphates that can serve as, among other things, energy and phosphate reserves. Roewe *et al.* report that polyphosphates in *E. coli* impaired survival in a mouse model of sepsis. In vitro, exposure to bacterial polyphosphates hinders lipopolysaccharide (LPS)–induced polarization of macrophages into a proinflammatory M1 phenotype. Instead, an anti-inflammatory M2 signature is favored, in part caused by suppression of LPS-TLR4–induced interferon β production. Although incompletely understood, *E. coli* polyphosphates antagonize antimicrobial macrophage functions such as phagocytosis, nitric oxide synthesis, and inflammatory cytokine production. —STS

Nat. Commun. **11**, 4035 (2020).



The nematode *Caenorhabditis elegans* (pictured) uses a particular family of enzymes to modify pheromones used for communication.

IMAGE: SINCLAIR STAMMERS/SCIENCE SOURCE

BIOSYNTHESIS

Building a suite of worm pheromones

The nematode *Caenorhabditis elegans* uses small-molecule pheromones to communicate and to influence group behavior. Studying the biosynthesis of one group of pheromones, the ascarosides, Faghih *et al.* found that different members of a group of 34 homologous enzymes attached different functional groups to a sugar component of the pheromone. The resulting molecules were set up to be tailored by β-oxidation, a process that has been shown previously to change the message according to the metabolic state of the worm. —MAF

J. Am. Chem. Soc. **142**, 13645 (2020).

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PALEOECOLOGY

Using the past to inform the future

The late Quaternary paleorecord, within the past ~130,000 years, can help to inform present-day management of the Earth's ecosystems and biota under climate change. Fordham *et al*. review when and where rapid climate transitions can be found in the paleoclimate record. They show how such events in Earth's history can shape our understanding of the consequences of future global warming, including rates of biodiversity loss, changes in ecosystem structure and function, and degradation in the goods and services that these ecosystems provide to humanity. They also highlight how recent developments at the intersection of paleoecology, paleoclimatology, and macroecology can provide opportunities to anticipate and manage the responses of species and ecosystems to changing climates in the Anthropocene. —AMS

Science, this issue p. 1072

MICROBIAL ENGINEERING

Building niches for multiple microbes

Microbial conversion of heterogeneous feedstocks such as lignocellulose into one or more desired products requires assembly of pathways to both break down the inputs and produce the outputs. Shahab *et al*. assembled a collection of microbes that occupy different spatial niches within a bioreactor and have different metabolic capabilities. In the simplest version of their setup, an aerobic fungus adeptly breaks down cellulose into short sugar chains, an oxygen-tolerant bacterium converts these to lactic acid, and an anaerobic bacterium uses the lactic acid to synthesize the short-chain fatty acid butyric acid. Additional microbes handle limitations in the intermediate stream or divert lactic acid to longer-chain fatty acids, which are potentially more valuable. The integrated production and use of lactic acid for biosynthetic reactions could ideally serve as a platform for biosynthesis using robust, heterogeneous microbial consortia. —MAF

Science, this issue p. 1073

CORONAVIRUS

Digital health in India

The coronavirus disease 2019 (COVID-19) pandemic has highlighted the need for effective digital health strategies across the world because of the necessary lockdowns and social distancing. In a Perspective, Agrawal takes the example of digital health transformation in India, which has been reinvigorated by the COVID-19 pandemic to ensure an equitable and broad coverage. Such digital transformation is not without challenges, and it is important to maintain a balance between digital and in-person health care. —GKA

Science, this issue p. 1050

NEUROSCIENCE

Synthetic excitatory synaptic organizer

The human brain contains trillions of synapses within a vast network of neurons. Synapse remodeling is essential to ensure the efficient reception and integration of external stimuli and to store and retrieve information. Building and remodeling of synapses occurs throughout life under the control of synaptic organizer proteins. Errors in this process can lead to neuropsychiatric or neurological disorders. Suzuki *et al*. combined structural elements of natural synaptic organizers to develop an artificial version called CPTX, which has different binding properties (see the Perspective by Salinas). CPTX could act as a molecular bridge to reconnect neurons and restore excitatory synaptic function in animal models of cerebellar ataxia, familial Alzheimer's disease, and spinal cord injury. The findings illustrate how structure-guided approaches can help to repair neuronal circuits. —SMH

Science, this issue p. 1074; see also p. 1052

CELL BIOLOGY

You can find your own way

Migration of cells through tissues and embryos is often steered by gradients of attractive chemicals in a process called chemotaxis. Cells are best at navigating complex routes, for which they use "self-generated chemotaxis" and create their own attractant gradients. An example of this is when neutrophils migrate into tissues to attack infection. Using modeling and live-cell data, Tweedy *et al*. found that self-generated chemotaxis allows cells to obtain surprising amounts of information about their environment. Cells of the slime mold *Dictyostelium discoideum* and mouse pancreatic cancer–derived cells were able to use the diffusion of attractants to identify the best route through complex mazes, even when the correct path was long and twisted, without ever entering incorrect paths. —SMH

Science, this issue p. 1075

NEUROSCIENCE

To continue or to switch strategy?

Successful behavior in an uncertain, changing, and open-ended environment critically relies on the ability to decide between continuing with the ongoing strategy or exploring new options. Neuroimaging studies have shown that the human medial prefrontal cortex (mPFC) is the part of the brain that primarily deals with this dilemma. However, the contribution of the different mPFC regions remains largely unknown. Domenech *et al*. recorded neuronal activity in six epileptic patients with depth electrodes in this brain area (see the Perspective by Steixner-Kumar and Gläscher). The ventral mPFC inferred the reliability of the ongoing action plan according to action outcomes. It proactively flagged outcomes either as learning signals to better exploit this plan or as potential triggers to explore new ones. The dorsal mPFC then evaluated action outcomes and generated an adaptive behavioral strategy. —PRS

Science, this issue p. 1076; see also p. 1056

QUANTUM COMPUTING

Twelve-qubit quantum computing for chemistry

Accurate electronic structure calculations are considered one of the most anticipated applications of quantum computing that will revolutionize theoretical chemistry and other related fields. Using the Google Sycamore quantum processor, Google AI Quantum and collaborators performed a variational quantum eigensolver (VQE) simulation of two intermediate-scale chemistry problems: the binding energy of hydrogen chains (as large as H_{12}) and the isomerization mechanism of diazene (see the Perspective by Yuan). The simulations were performed on up to 12 qubits, involving up to 72 two-qubit gates, and show that it is possible to achieve chemical accuracy when VQE is combined with error mitigation strategies. The key building blocks of the proposed VQE algorithm are potentially scalable to larger systems that cannot be simulated classically. —YS

Science, this issue p. 1084; see also p. 1054

BACTERIAL METABOLISM

Soil sulfur metabolism surprise

Soil bacteria have a range of metabolic pathways that contribute to acquiring and recycling nutrients and carbon. Curiously, some of

Published by AAAS

these organisms give off ethylene gas when starved for sulfur under anaerobic conditions. North *et al*. traced the source of ethylene to a small, sulfur-containing organic molecule produced by certain reactions in cells. Growing cells in sulfur-limiting conditions enabled them to identify the enzymes involved in sulfur salvage, and the concomitant ethylene production, through this pathway. Methane and ethane were also observed as products when appropriate substrates were provided. The key genes involved are distantly related to nitrogenase and several other reductase enzymes found in bacteria and archaea. The involvement of such nitrogenase-like genes in sulfur metabolism highlights the potential of unexplored diversity in this family of enzymes and raises many mechanistic and evolutionary questions that are now ripe for exploration. —MAF

Science, this issue p. 1094

CORONAVIRUS

From cough to splutter

In epidemiology, serial intervals are measured from when one infected person starts to show symptoms to when the next person infected becomes symptomatic. For any specific infection, the serial interval is assumed to be a fixed characteristic. Using valuable transmission pair data for coronavirus disease (COVID-19) in mainland China, Ali *et al*. noticed that the average serial interval changed as nonpharmaceutical interventions were introduced. In mid-January 2020, serial intervals were on average 7.8 days, whereas in early February 2020, they decreased to an average of 2.2 days. The more quickly infected persons were identified and isolated, the shorter the serial interval became and the fewer the opportunities for virus transmission. The change in serial interval may not only measure the effectiveness of infection control interventions but may also indicate rising population immunity. —CA

Science, this issue p. 1106

COSMOCHEMISTRY

An unexpected source of Earth's water

The abundances of Earth's chemical elements and their isotopic ratios can indicate which materials formed Earth. Enstatite chondrite (EC) meteorites provide a good isotopic match for many elements but are expected to contain no water because they formed in the hot inner Solar System. This would require Earth's water to be from a different source, such as comets. Piani *et al*. measured hydrogen contents and deuterium/hydrogen ratios (D/H) in 13 EC meteorites (see the Perspective by Peslier). They found far more hydrogen than is commonly assumed, with D/H close to that of Earth's mantle. Combining these data with cosmochemical models, they show that most of Earth's water could have formed from hydrogen delivered by EC meteorites. —KTS

Science, this issue p. 1110; see also p. 1058

ORGANIC CHEMISTRY

Expanding the kinetic resolution purview

Dynamic kinetic resolution (DKR) is a powerful method to transform a pair of mirror-image reactants into just one of the two possible mirror-image products. The key is to find a means of rapidly interconverting the reactants while one of them is being more efficiently funneled to product. DeHovitz *et al*. report cooperative application of organocatalysis, photoredox catalysis, and enzymatic catalysis to achieve DKR of β-substituted ketones into chiral alcohols. This β position has typically been considered too configurationally stable for a DKR approach. —JSY

Science, this issue p. 1113

CORONAVIRUS

A common theme in antibody responses

In the fight against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), antibodies are a key tool, both as potential therapeutics and to guide vaccine development. Yuan *et al*. focused on finding shared antibody responses, in which multiple individuals develop antibodies against the same antigen using the same genetic elements and modes of recognition. The authors identified the immunoglobulin heavy-chain variable region 3-53 gene as the most frequently used among 294 antibodies that target the receptor-binding domain (RBD) of the viral spike protein. These antibodies have few somatic mutations, and crystal structures of two neutralizing antibodies bound to the RBD show that mostly germline-encoded residues are involved in binding. The minimal affinity maturation and high potency of these antibodies is promising for vaccine design. —VV

Science, this issue p. 1119

FLAVIVRUSES

Double whammy

Dengue and Zika virus epidemics have been lapping each other around the globe. These are closely related mosquito-borne viruses with about 40% homology within the envelope protein. We know that subsequent dengue infections bring a risk of antibody-dependent disease enhancement. Whereas emphasis has been placed on how prior dengue immunity affects Zika infection, little is known about how prior Zika immunity may affect dengue disease. Katzelnick *et al*. have been following a well-characterized and established pediatric cohort in Nicaragua who were serially exposed to both flaviviruses in recent years (see the Perspective by Clapham). This study shows not only that a previous history of just one round of dengue is a problem but also that prior Zika immunity creates an increased risk for severe dengue virus sereotype 2 infection. By contrast, multiple infections raise antibodies to protective levels. —CA

Science, this issue p. 1123; see also p. 1055

PHARMACOLOGY

Drugging an "undruggable" phosphatase

Mitogen-activated protein kinase (MAPK) phosphatases (MKPs) have roles in many diseases but have been considered "undruggable" because of their high degree of conservation. Gannam *et al*. identified and characterized a small molecule that inhibited MKP5 but did not affect the activity of related phosphatases. In muscle cells, the inhibitor blocked MKP5 from inactivating the MAPKs and inhibited a signaling pathway that promotes fibrosis in dystrophic muscle disease. These results demonstrate that it is possible to target MKPs with a high degree of specificity. —WW

Sci. Signal. **13**, eaba3043 (2020).

INFLAMMATORY DISEASE

Profiles of rogue gut lymphocytes

Dysregulated human gut B and T lymphocytes contribute to the immunopathogenesis of ulcerative colitis (UC), a type of inflammatory bowel disease characterized by mucosal damage in the colon. Boland *et al.* probed the immunologic perturbations associated with active UC by sequencing messenger RNA and clonal antigen receptors from individual $CD45^+$ cells isolated from rectal biopsies or blood of patients and healthy controls. The resulting single-cell sequencing resource revealed heterogeneity among tissue-resident memory T cells (T_{RMs}) in UC, including expansion of an inflammatory $CD8^+$ T_{RM} subset expressing the Eomesodermin transcription factor. The identification of this T_{RM} population and other disease-associated T and B cell subsets provides a platform for future functional studies addressing how these subsets conspire to trigger chronic mucosal injury in UC. —AB

Sci. Immunol. **5**, eabb4432 (2020).

Published by AAAS

REVIEW SUMMARY

PALEOECOLOGY

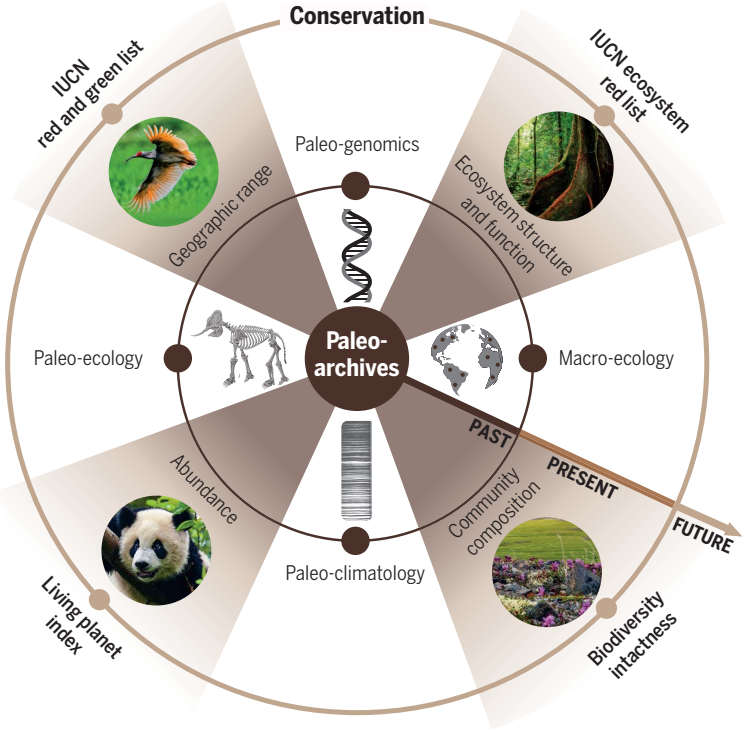
Using paleo-archives to safeguard biodiversity under climate change

Damien A. Fordham*, Stephen T. Jackson, Stuart C. Brown, Brian Huntley, Barry W. Brook, Dorthe Dahl-Jensen, M. Thomas P. Gilbert, Bette L. Otto-Bliesner, Anders Svensson, Spyros Theodoridis, Janet M. Wilmshurst, Jessie C. Buettel, Elisabetta Canteri, Matthew McDowell, Ludovic Orlando, Julia Pilowsky, Carsten Rahbek, David Nogues-Bravo

BACKGROUND: Effects of recent global warming have been documented in every biome on Earth. Safeguarding biodiversity and ecosystem services against future impacts requires reliable predictions of climate-driven biological responses and effective solutions. Integrated research in paleoecology, paleogenomics, paleoclimatology, and macroecology offers new prospects for projecting and managing current biotic responses to climate change. By revealing mechanisms that shaped past and present biodiversity patterns, this interdisciplinary intersection provides an empirical foundation for anticipating responses to accelerated climate change. New insights are coming from developments in high-throughput sequencing, computational technologies, ecological simulation models, and increased spatiotemporal resolution of paleoenvironmental data from late Quaternary paleo-records (the past ~130,000 years). Although these advances reveal biodiversity responses to past global change, benefits for improving forecasting of biodiversity impacts and refining conservation policies are lagging. Abundant opportunities exist for using the late Quaternary paleo-record to inform conservation practices and policies in the context of climate change.

ADVANCES: The threat of anthropogenic climate change demands that conservationists seek more effective ways of improving management of biodiversity and ecosystems. Analytical approaches that combine high-resolution paleoclimate proxy and simulation data, precisely dated fossils, and genetic diversity estimates from ancient DNA are unveiling biotic responses to various rates and magnitudes of natural climate warming, some comparable with 21st century projections. Reference periods in Earth's history provide natural laboratories for testing fundamental ecological theory and offer opportunities to identify ecological processes that influence the likelihood of extinction and ecosystem change, to test efficacy of threatened-species assessments and resilience of biota during periods of abrupt warming, and to locate biogeographic areas that remain stable under shifting climates. Refinement of essential biodiversity variables by using past biodiversity dynamics will improve our understanding of climate-driven shifts in species populations, community composition, and ecosystem structure and function. From this, biodiversity early-warning systems, conservation strategies, and decision-making tools can be tested at fine-grain spatiotemporal scales, providing an evidence base for understanding and improving projections of species- and ecosystem-level collapse.

OUTLOOK: As paleo-archives become more routinely integrated into conservation science, guidelines for the management of nature will benefit from understanding how different spatiotemporal scales of past climate change affected species and ecosystems across the planet. This will require global initiatives to harmonize vast numbers of paleoclimate-proxy and paleoecological records with high-resolution paleoclimate projections from Earth system models. Paleoecological data offer a means to disentangle climate and nonclimate drivers of biodiversity and ecosystem function, particularly in concert with simulation models and integrated analytical techniques that compare biotic change across regions with contrasting histories of human colonization and land use. Moreover, developments in paleogenomics that pinpoint adaptation across and within species will identify microevolutionary processes that lend resilience to biodiversity in shifting climates. Incorporating paleo-archives in conservation policies will equip decision-makers with improved strategies for mitigating biodiversity loss from climate change in the Anthropocene. ■



Paleo-archives offer new prospects for benchmarking and maintaining future biodiversity. Integrated research using paleo-archives provides empirical foundations for contextualizing climate-driven changes in species populations, community composition, and ecosystem structure and function. These observations can inform conservation strategies under anthropogenic climate change.

CREDITS: HUNG CHUNG CHIH, WANG LIQIANG, ANDREI STEPANOV, AUSTRALIANCAMERA, ALHOVIK, BLUERING MEDIA, NAZARI M, ANDERS SVENSSON/ SHUTTERSTOCK

The list of author affiliations is available in the full article online.
*Corresponding author. Email: damien.fordham@adelaide.edu.au
Cite this article as D. A. Fordham *et al.*, *Science* **369**, eabc5654 (2020). DOI: 10.1126/science.abc5654

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc5654

REVIEW

PALEOECOLOGY

Using paleo-archives to safeguard biodiversity under climate change

Damien A. Fordham[1,2]*, Stephen T. Jackson[3,4], Stuart C. Brown[1], Brian Huntley[5], Barry W. Brook[6], Dorthe Dahl-Jensen[7,8], M. Thomas P. Gilbert[9,10], Bette L. Otto-Bliesner[11], Anders Svensson[7], Spyros Theodoridis[2], Janet M. Wilmshurst[12,13], Jessie C. Buettel[6], Elisabetta Canteri[1,2], Matthew McDowell[6], Ludovic Orlando[14,15], Julia Pilowsky[1,2], Carsten Rahbek[2,16,17,18], David Nogues-Bravo[2]

Strategies for 21st-century environmental management and conservation under global change require a strong understanding of the biological mechanisms that mediate responses to climate- and human-driven change to successfully mitigate range contractions, extinctions, and the degradation of ecosystem services. Biodiversity responses to past rapid warming events can be followed in situ and over extended periods, using cross-disciplinary approaches that provide cost-effective and scalable information for species' conservation and the maintenance of resilient ecosystems in many bioregions. Beyond the intrinsic knowledge gain such integrative research will increasingly provide the context, tools, and relevant case studies to assist in mitigating climate-driven biodiversity losses in the 21st century and beyond.

Global mean temperatures are nearing the upper bound of those experienced over the past 1.2 million years (*1*), and their effects are being documented across every biome on Earth (*2*). As climate change intensifies in the coming decades, safeguarding biodiversity and the services that ecosystems provide to nature and people must remain high on the international policy agenda (*1*–*5*). Effective interventions will require conservation actions based on reliable predictions and evidence-based solutions (*3*). However, robust forecasts of species- to ecosystem-level responses to changing climates remain difficult (*4*), adversely affecting conservation planning and outcomes (*5*).

Research at the intersection of paleoecology, paleoclimatology, paleogenomics, macroecology, and conservation biology is offering new approaches to anticipate and manage responses of biodiversity and ecosystems to climate and other environmental change (*6*). By revealing ecological and evolutionary mechanisms that have shaped past and current-day biodiversity patterns, this research provides an empirical foundation for quantifying the broad footprint of accelerated rates of climate change on natural systems and for identifying long-term ecological and evolutionary processes that govern climate-biodiversity dynamics (*7*).

Although human land use, over-exploitation, and movement of non-native species remain primary drivers of biodiversity loss (*8*), climate change will grow in importance in the coming decades (*1*, *9*). Paleo-archives allow biodiversity responses to climate perturbations of varying rates and magnitude, some approximating those predicted for the near future (*10*), to be tracked in situ (places where they occurred) over centuries to many millennia (*11*, *12*). Past warming intervals provide critical reference points in Earth's history that can be used as natural laboratories to identify biotic vulnerability and resilience to rapid climatic change (*7*) and to connect ecological and evolutionary theory to the design and implementation of conservation practices to protect biodiversity (*13*).

Many species on Earth today have existed for hundreds of thousands to millions of years (*14*), having experienced many global glacial-interglacial cycles, each including rapid regional warming events, some spanning only decades (*15*). Although in many biogeographic regions these events are comparable in pace and magnitude with 21st century forecasts (*16*), they do not offer a direct analog for future global warming from recent anthropogenic climate forcing (*17*). Nevertheless, they can reveal actual species- and ecosystem-level responses to previous rapid changes in climate (*11*, *12*, *18*, *19*). One of the most powerful features of the paleo record is its heuristic nature (*20*), providing concrete narratives, scenarios, and other thinking tools to better anticipate and visualize the potential ecological and evolutionary consequences of future climate change, enhancing knowledge of principles and mechanisms for conserving biodiversity and ecosystem services (*6*, *7*).

Increased availability of precisely dated fossil records, genome-scale ancient DNA, and simulations with sufficient temporal resolution for reconstructing rapid climate change events means that the late Quaternary (last ~130,000 years) is providing new and effective opportunities to understand better the effects of climate change on biodiversity dynamics (*7*, *21*), bolstering conservation knowledge (*22*, *23*), particularly in the face of uncertainties on impacts of future climate change (*24*). This includes improved information on the mechanisms by which species have coped with high rates and magnitudes of climate change at a range of spatiotemporal scales, including those directly relevant to vulnerability assessments (*3*), based on sensitivities and adaptive capacities to climate change at human-relevant time scales (*5*).

With a growing emphasis on integrating paleobiology into conservation biology (*13*, *20*, *22*–*25*), and the emergence of conservation paleobiology (*6*), clear guidelines are needed to define when, where, and how scientists can use the late Quaternary paleo-record, spanning the 132,000 years since the start of the Last Interglacial (LIG), to provide insights for conservation policies that address climate change. The late Quaternary represents the origins of extant ecosystems (*21*), providing a suitable geohistorical period for informing responsible management of Earth's ecosystems and diverse biota under trajectories of future climate change. Here, we pinpoint where and when climatic transitions on human-relevant time scales are found in the paleoclimate record and show how these reference points in Earth's history can be used as mensurative experiments to establish likely consequences of future global warming for terrestrial biodiversity loss and ecosystem properties, including goods and services provided to humanity.

Approximately 40% of terrestrial ecosystems are projected to have experienced past shifts in temperature that are similar in pace and magnitude to regional-scale future forecasts (*16*). Thus, there is enormous potential to use geohistorical data to better derive and strengthen

[1]The Environment Institute and School of Biological Sciences, University of Adelaide, South Australia 5005, Australia. [2]Center for Macroecology, Evolution, and Climate, GLOBE Institute, University of Copenhagen, Copenhagen Ø 2100, Denmark. [3]Southwest and South Central Climate Adaptation Science Centers, U.S. Geological Survey, Tucson, AZ 85721, USA. [4]Department of Geosciences and School of Natural Resources and the Environment, University of Arizona, Tucson, AZ 85721, USA. [5]Department of Biosciences, Durham University, Durham, DH1 3LE, UK. [6]School of Natural Sciences and ARC Centre of Excellence for Australian Biodiversity and Heritage, University of Tasmania, Hobart, Tasmania 7001, Australia. [7]Centre for Ice and Climate, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø 2100, Denmark. [8]Centre for Earth Observation Science, University of Manitoba, Winnipeg MB R3T 2N2, Canada. [9]Center for Evolutionary Hologenomics, GLOBE Institute, University of Copenhagen, Copenhagen Ø 2100, Denmark. [10]University Museum, Norwegian University of Science and Technology, Trondheim, Norway. [11]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO 80307-3000, USA. [12]Long-Term Ecology Laboratory, Manaaki Whenua–Landcare Research, Lincoln 7640, New Zealand. [13]School of Environment, The University of Auckland, Auckland 1142, New Zealand. [14]Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse UMR 5288, Université de Toulouse, CNRS, Université Paul Sabatier, France. [15]Section for GeoGenetics, GLOBE Institute, University of Copenhagen, Copenhagen Ø 2100, Denmark. [16]Department of Life Sciences, Imperial College London, Ascot SL5 7PY, UK. [17]Danish Institute for Advanced Study, University of Southern Denmark, Odense, Denmark. [18]Institute of Ecology, Peking University, Beijing 100871, China.
*Corresponding author. Email: damien.fordham@adelaide.edu.au

conservation management and policy through improved knowledge of biotic responses to climatic stressors. Leveraging these paleo "logbooks" of climate-driven biodiversity dynamics requires harmonized measures of ancient, recent, and forecast responses of biodiversity to climatic change. Newly constructed "essential biodiversity variables" (*26*) can quantify biotic change, vulnerability, and resilience to climatic shifts of the late Quaternary, as well as those in recent history.

Ancient warm periods and biodiversity consequences

Earth has both globally and locally experienced conditions warmer than the 20th century during the Quaternary and the Tertiary (*10*), providing numerous glimpses of what a warmer world looks like ecologically (*27*). Although not direct analogs for an enhanced-greenhouse future, because of different processes in the climate system (*17*), these ancient warm periods provide an empirical basis for identifying and understanding ecological and biogeographical implications of a warmer world, at spatiotemporal scales and levels of biological complexity that cannot be captured through deliberate experiments or direct monitoring.

Late Quaternary intervals comparable with near-future climate projections include the Holocene Thermal Maximum (HTM) between 12,000 and 5,000 years B.P. (before the present) (*28*, *29*), in which surface-temperature anomalies relative to preindustrial climates were in some regions +3°C or more (Fig. 1), and the earliest millennia of the LIG (129,000 to 111,000 years B.P.), when the global mean temperature anomaly was >1°C above preindustrial conditions (*10*). Summer warming was prominent in Northern Hemisphere continental interiors during these periods; by contrast, ocean warming was comparatively small (Fig. 1). Deeper-time warm intervals include the Early Eocene (~50 million years ago) and the Mid-Pliocene (3.3 million to 3.0 million years ago), when atmospheric CO_2 concentrations were greater or comparable with that of the present; mean annual surface temperatures were >10°C (Early Eocene) and ~3°C warmer (Mid-Pliocene) than preindustrial temperatures (*10*).

During the HTM, warmer temperatures and different moisture regimes led to climate-driven movements of biome boundaries and elevational shifts in montane vegetation belts. For example, the tundra-forest boundary shifted ~200 km northward in Central Siberia (*30*), whereas the prairie-forest boundary on the North American Great Plains shifted ~200 to 250 km eastward under increased aridity (*31*). Montane treelines were lower than today on islands in the Southern Ocean during HTM, owing to drier conditions under higher temperatures (*29*). In much of northern subtropical Africa, HTM climate was wetter (although cooler) than today, largely because of higher summer temperatures in the northern Sahara (Fig. 1) leading monsoon rains to extend farther north, allowing many savanna plants (and riparian and wetland animals) to extend their ranges northward by ~400 to 500 km in what is today a desert (*32*). In the Central Andes (South America), warm and wet conditions during the HTM resulted in plants moving up to 1000 m downslope (*33*).

During the LIG, boreal forests extended north into Greenland (*34*), whereas the Sahara was largely occupied by savanna (*35*). In Europe, warmer temperatures allowed the hippopotamus (*Hippopotamus amphibius*) to expand its range as far north as Britain (*36*). The occurrence of giant tortoise *Hesperotestudo* (*Geochelone*) *crassiscutata* in the American Midwest in the LIG implies temperatures above freezing throughout the year (*37*). Mid-Pliocene warming

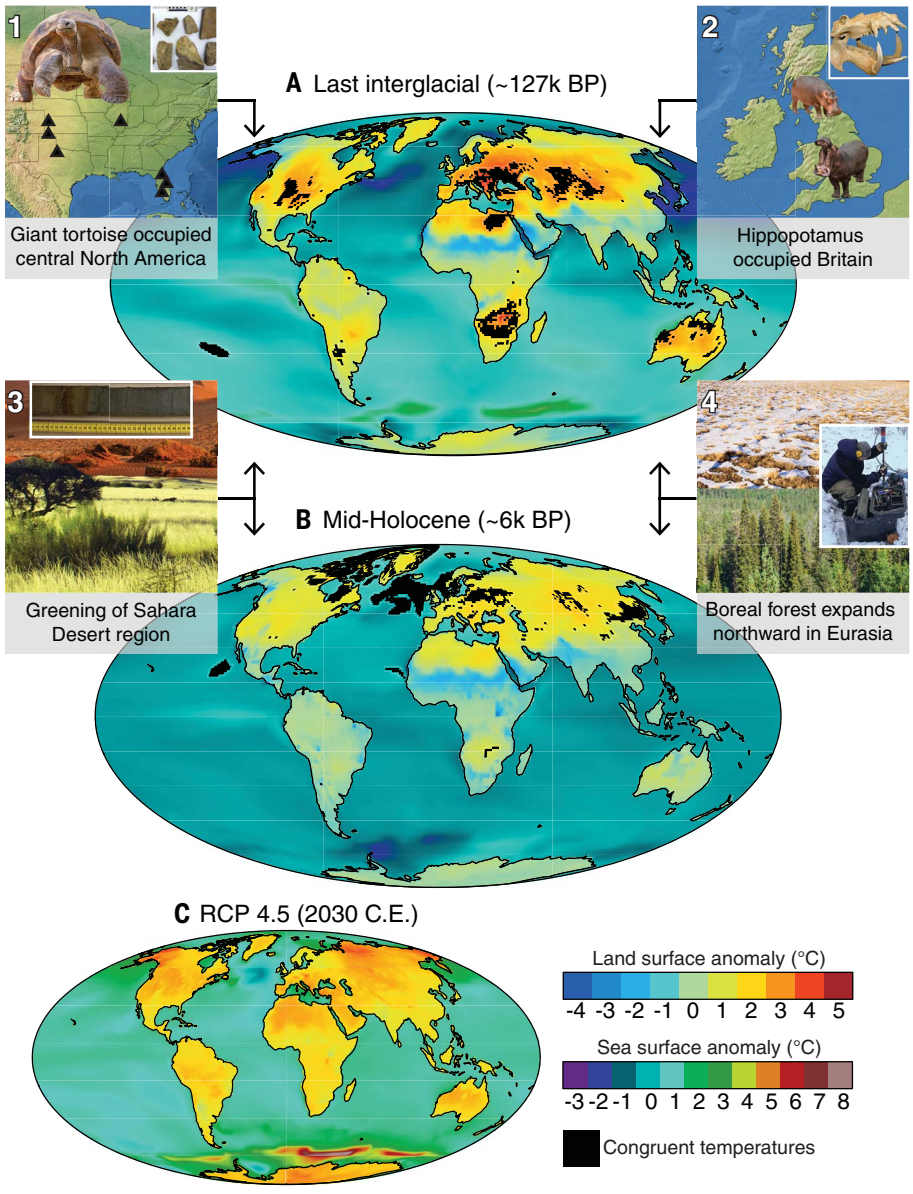


Fig. 1. Similarity in temperatures between past warm periods and the near future. (**A** to **C**) Boreal summer (JJA) land and sea-surface projected temperature changes from a preindustrial baseline for the (A) LIG (~127,000 years B.P.), (B) mid-Holocene (~6000 years B.P.), and (C) 2030 under RCP 4.5. Black areas in (A) and (B) represent locations with temperature anomalies that are congruent with (C). Boxes 1 to 4 show biotic responses inferred from paleo-archives (insets), with arrows pointing to the relevant time periods. Paleo-archives include fossil bones (boxes 1 and 2) and pollen in sediment cores (boxes 3 and 4). Triangles in box 1 indicate fossil sites. Paleoclimate simulations are detailed in appendix S1. [Image credits: Susan Schmitz, Smithsonian, Vincent Gordano Photo, Wlad74, Anan Kwaewkhammul, Janelle Lugge, JaySi, Northwestern, Real Window Creative, Natalia Golubnycha, Audrius Merfeldas]

caused similar latitudinal displacements in vegetation, including poleward contractions of tundra, northward expansions of boreal forest, and greening of the Sahara (*38*). Tundra-like communities with dwarf-shrub species of southern beech (*Nothofagus beardmorensis*) occurred at high altitudes in Antarctica (*39*), where today there is polar desert.

Well-documented geographic displacements of species and biomes across the globe during these past warm periods underscore the need for species' ability to shift their range boundaries by hundreds to thousands of kilometers in response to future warming. They also lead to expectations that many species will respond individualistically to future climatic changes, resulting in communities and ecosystems without modern analog (*40*). These mechanisms of biodiversity and ecosystem change are unlikely to differ, at least qualitatively, in the future (*41*).

Ecological ramifications of rapid warming

Between 21,000 and 18,000 years B.P., Earth began to emerge from a glacial climate—with low CO_2 concentrations, lower sea levels, and large land-based ice sheets—into the Holocene interglacial (*42*). During the global warming of the last deglaciation (LD; 20,000 to 10,000 years B.P.), many terrestrial regions experienced temperature increases of 4° to >10°C (Fig. 2). Some were gradual, but others occurred during episodes that spanned decades to centuries (*43*). In some regions, these rapid warmings were up to half as large as the entire difference between glacial and modern conditions (*44*).

Greenland ice cores reveal rapid large-magnitude warming events at 14,700 and 11,700 years B.P., with warming rates of ~10°C spanning decades (*43*). These are similar to rates forecast for Northern Hemisphere polar regions under Representative Concentration Pathway (RCP) 8.5 (Fig. 2). In Europe, regional temperatures increased 3° to 5°C during the rapid warming event 14,700 years B.P. (*45*). Although regional climates over the North Atlantic region and Europe reorganized quickly during the LD warmings, signatures of these rapid events are found globally at varying amplitudes and lags, including in the tropics (Fig. 2). In the southern hemisphere, more gradual and less pronounced warming events occurred during cold stadial periods in the northern hemisphere (*44*).

Paleoclimate simulations since 21,000 years B.P. suggest that 39% of terrestrial Earth experienced near-centennial climate shifts of similar magnitude to those of regional-scale future forecasts under RCP 8.5 (Fig. 3). These rapid temperature changes occurred largely during the LD (*16*). Zoogeographic regions (*46*) with the largest overlap between past and future paces of temperature change include Arctico-Siberian, Eurasian, Amazonian, and Novozelandic (Fig. 3); many of these regions are rich in paleoecological records, providing natural laboratories for better anticipating biotic responses to climatic changes.

Past rapid warmings profoundly affected biological diversity at multiple levels, including genetic diversity, species abundances and geographical ranges, community composition, and ecosystem structure. In Eurasia, many cold-adapted mammal species experienced reductions in population size of 50 to 90%, causing bottlenecks and consequent losses of genetic diversity (*47*). Population decreases and extirpations were widespread in Eurasia and North America during the LD warming events, with accompanying shifts in species distributions (including that of anatomically modern humans) and reorganization of diversity patterns (*11*). For example, in Eurasia, the saiga antelope (*Saiga tatarica*) retracted its range, surviving the 14,700 years B.P. warming event in temporary cool refugia (*48*), whereas the Arctic Fox (*Alopex lagopus*) was unable to track habitat shifts associated with LD warming events in mid-latitude Europe, leading to regional extinction (*49*). In North America, species evenness of small-mammal communities decreased sharply in response to rapid warming, negatively affecting functional stability and community persistence (*50*), and a widespread and regionally abundant tree species (*Picea critchfieldii*) went extinct (*51*).

These alterations in species distributions and abundances radically changed the structure and function of ecosystems, influencing the physical formation of habitats, geochemical cycles, and primary productivity (*18*, *52*). For example, in Britain and Ireland, rapid warming caused changes in plant-soil and plant-plant interactions, resulting in above- and belowground shifts in ecosystem composition and overall reductions in productivity (*19*). Moreover, decreases in the ranges and abundances of terrestrial megafauna reduced their contribution to nutrient cycling and redistribution (*52*), with consequences for structure and function of ecosystems (*53*). Late-glacial population decline of megaherbivores in the North American interior reduced browsing pressure on broadleaved trees, provoking a rapid transition from coniferous forest to mixed coniferous-deciduous forest, along with a change in fire regime (*18*).

Geohistorical records of biotic responses to rapid warming events of the LD indicate that abrupt climate-driven changes in species distributions and numbers will be far reaching, causing feedbacks on the climate system (such as decreasing albedo and changes in evapotranspiration) that alter the pace of warming (*45*), affecting human well-being through changed ecosystem health and services (*54*). Because many LD warming events are similar in magnitude and pace to 21st-century forecasts (Figs. 2 and 3), LD paleo-archives provide invaluable "logbooks" for establishing how future global warming is likely to redistribute terrestrial biodiversity and change ecosystem functioning.

Biotic responses to extreme climatic events

Conservation management and planning decisions are typically based on species, populations, or ecosystems in specific habitats and locales, involving time horizons of seasons to decades. Paleo-archives offer opportunities, particularly in the mid- to late Holocene, to identify processes underlying climate-change responses at the taxonomic, spatial, and temporal scales used by decision-makers. Genetic and demographic studies of extant populations can often be tied seamlessly to late Holocene records of population and biogeographic dynamics accompanied by independent paleoclimate records.

Holocene paleoclimate reconstructions show that extreme climate anomalies are often clustered, resulting in periods of drought or high temperatures, with variable durations, return frequencies, and magnitudes (*55*). These episodic climatic events govern rates and patterns of range expansion and contraction (*12*). Studies of tree populations, for example, reveal the importance of nonstationary Holocene climate variability and its interactions with long-distance dispersal, local demographic processes, and species life-history traits (*56*–*59*). Moreover, Holocene records show that population expansions and declines are not necessarily accompanied by changes in geographic distribution (*60*, *61*), as is often assumed in conservation assessments (*62*). For example, a rapid population increase of eastern hemlock (*Tsuga canadensis*) near its western range margin was not accompanied by geographic expansion (*60*), whereas its geographic distribution held steady during a dramatic range-wide population decline (*61*). The latter example represents a rapid ecosystem transformation, in which a dominant conifer (hemlock) was replaced by deciduous trees and pines, forming forests with different structural and functional properties, in response to a contingent series of climatic and ecological processes operating at different temporal and spatial scales (*61*, *63*).

As threats of ecosystem transformations increase in response to current global warming (*64*), Holocene records offer opportunities to better understand the mechanisms that underlie these transformations. For example, paleoecological studies confirm that coniferous forest stands can be transformed by a single severe climate-driven fire event (*65*), a phenomenon that appears to be under way in parts of western North America (*66*).

Strengthening conservation policies and strategies

As anthropogenic climate change accelerates, the value of knowledge on past responses will

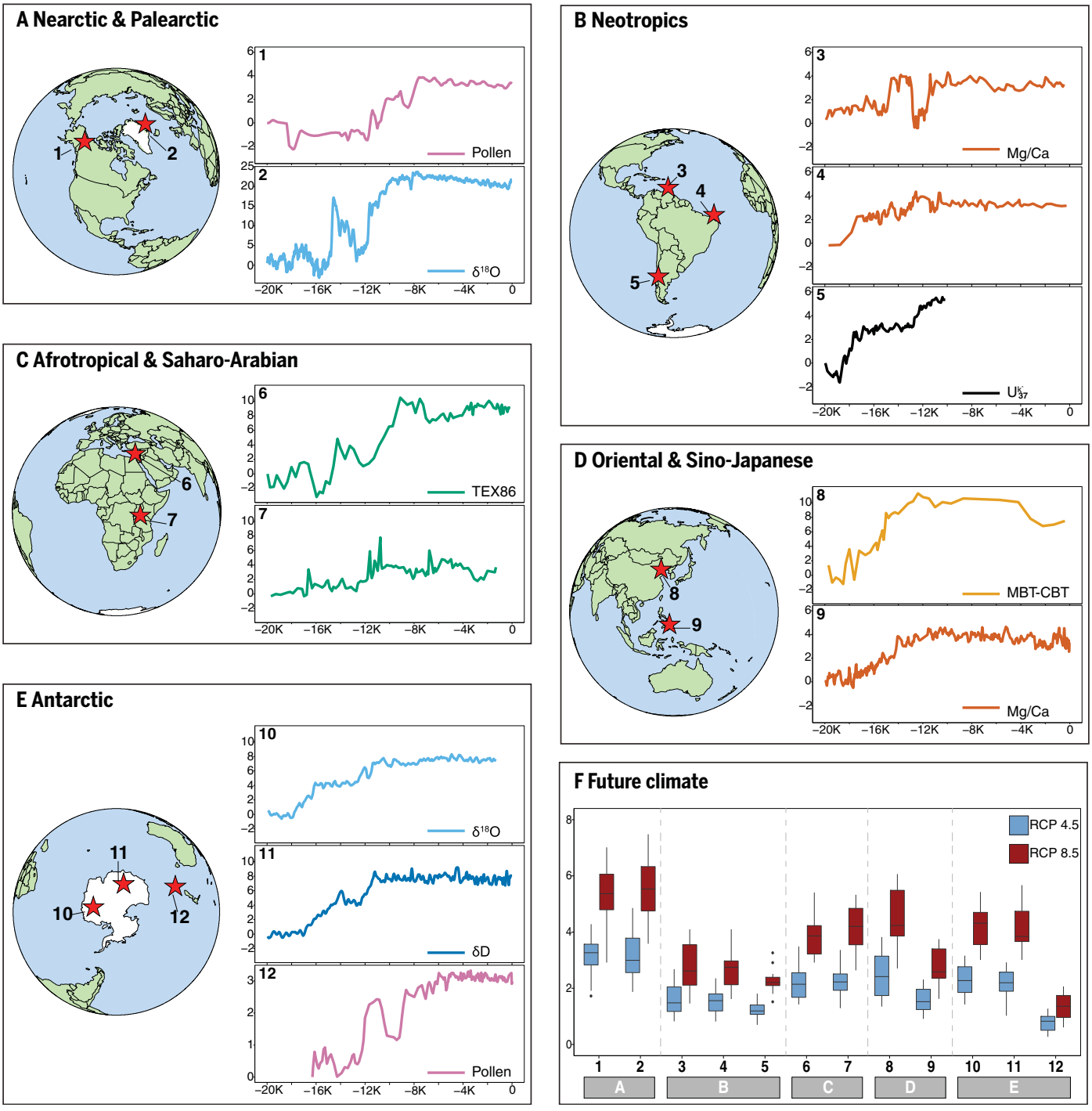


Fig. 2. Abrupt warming during the last deglaciation. (**A** to **E**) Temperature change (degrees Celsius; *y* axes) since 20,000 years B.P. (*x* axes) or the earliest record after 20,000 years ago [(A), 5; (C), 6; (D), 9; and (E), 12] for 12 paleoclimate reconstructions (*29*, *99*) from eight zoogeographic realms (*46*), including Antarctica. (**F**) Boxplots show a multimodel averaged forecast of annual-mean temperature change for 2080 from a baseline focused on 1990 (*16*) under two radiative-forcing scenarios for sites in (A) to (E). The *y* axis scale differs in (A) to (F). The *x* axis shows time before the present in thousands of years. Details of paleoclimate reconstructions in (A) to (E) are provided in appendix S1.

depend on developing harmonized metrics that can be used reliably and consistently to quantify both past and recent trends in extirpation and extinction rates, community composition, and ecosystem structure and function (Fig. 4). In the modern context, ecologists have proposed classes of "essential biodiversity variables" (EBVs) for measuring biodiversity change and deriving conservation policies (*26*). However, to date, their temporal reach has been limited to a few decades (*67*).

Paleo-archives provide opportunities to estimate EBVs under a far wider range of shifts in climatic and environmental conditions, including variables measured at the population, species, community, and ecosystem levels (Fig. 4). The application of paleo-archives to a subset of EBVs (paleo-enabled EBVs) makes the direct translation of past biodiversity responses to future environmental management more straightforward because EBVs comprise

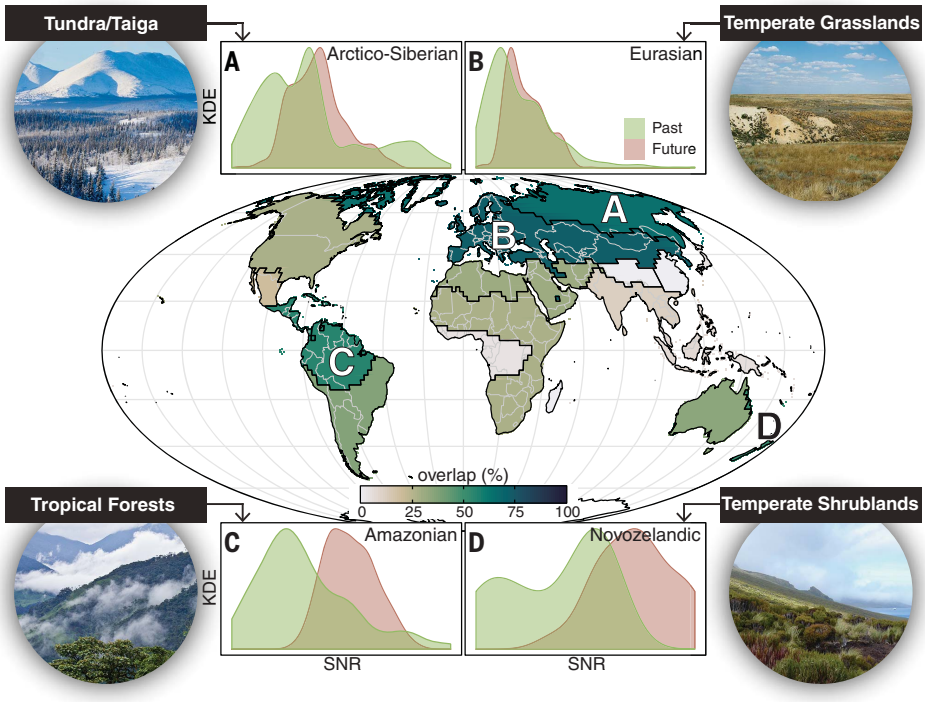


Fig. 3. Similarity in past and future rapid temperature shifts. Map shows overlap (%) in the signal-to-noise ratio (SNR) of rapid climate change events for past (21,000 years B.P. to 1850 CE) and future (21st century, RCP 8.5) temperature change in zoogeographic regions (*46*). Colors represent the percent overlap in SNR. (**A** to **D**) Plots show SNR distributional overlap (green indicates past, and red indicates future) within the (A) Arctico-Siberian (64%), (B) Eurasian (72%), (C) Amazonian (55%), and (D) Novozelandic (52%) regions based on kernal density estimates (KDE). Pictures show examples of ecosystems in regions (A) to (D) with geohistories suitable for assessing and maintaining future biodiversity. Data are based on (*16*). Distributional overlaps for all zoogeographic regions are provided in (*100*). [Image credits: Pi-Lens, Bogdan Kovenkin, Glenn R. Specht-grs photo, Dale Lorna Jacobsen]

a common currency for estimating and discussing biodiversity change among the scientific and policy communities (*67*).

Species distributions and abundances

Species' abundance is an EBV that can be inferred from fossil records (morphological, molecular, or both) by using count data or from genes in extant populations by using coalescent estimates of effective population size (Fig. 4). Probabilistic and bias-correction methods can be used to account for inconsistencies in radiometric dating, sampling rates, fossil preservation, and taphonomy (*25*, *68*), making inferences of biotic change from paleo data more comparable with the abundance-based indicators of biodiversity change used today in conservation policy [Living Planet Index, International Union for Conservation of Nature (IUCN) Red List, and IUCN Green List], and assessment of biodiversity targets [Convention on Biological Diversity (CBD)] used in the face of anthropogenic climate and environmental change (*69*).

Because warning signals of biotic transitions or state shifts are commonly identified by using time-series abundance data (*70*), paleo-archives provide opportunities to test conservation criteria and model efficacy, improving knowledge of critical thresholds for population collapses, possible extinction events, and ecosystem shifts (*71*). Moreover, a longer-term perspective enables natural variability bounds in species abundances and other conservation-relevant variables to be estimated over large-scale climatic shifts, often in settings where there were no substantial direct or indirect human impacts to confound the signal. In this way, paleo-archives allow thresholds of natural variability to be identified and integrated into threat-classification metrics to ensure that declines from greenhouse gas warming and other human-environment interactions are being prioritized (*72*).

Change in spatial distributions of species is another paleo-enabled EBV that can be inferred from the fossil record with ecological models and used to monitor single or aggregated taxonomic units under late Quaternary climate change, allowing improved understanding of natural distributions for recovery assessments and the role that climate had in past range contractions and expansions (Fig. 4, Muskox; *Ovibos moschatus*). Furthermore, paleo-enabled conservation variables allow morphological changes within species populations to be measured across time (Fig. 4) (*22*), providing a deeper understanding of species-level responses to climate change (*73*) and improved capacity to detect ecological regime shifts (*74*).

Assessments of biodiversity values are strongly dependent on the spatial and temporal scale of their evaluation, which in turn influences conservation-planning decisions (*22*). The common method of using the historic record after 1500 CE as a baseline for vulnerability assessments (*8*) can overlook long-term trends in species ranges, population size, and genetic diversity (Fig. 4) (*72*). This can directly affect evaluations of conservation status (IUCN Red List) and measures of conservation success (IUCN Green List) (*75*). Although paleo-archives can address this shortfall, a challenge will be to develop and apply long-term comparative frameworks to biological signals preserved in paleo-archives for local and range-wide extinctions of related species and subspecies.

Paleo-archives are now being integrated into ecological and evolutionary models to improve theories and make generalizations regarding the spatial dynamics of range collapses of species (*13*). For example, a common view in conservation is that densities in peripheral populations are typically lower and less stable than at the center of a species' range (*62*), which has led to the widely adopted goal of avoiding the range periphery in conservation

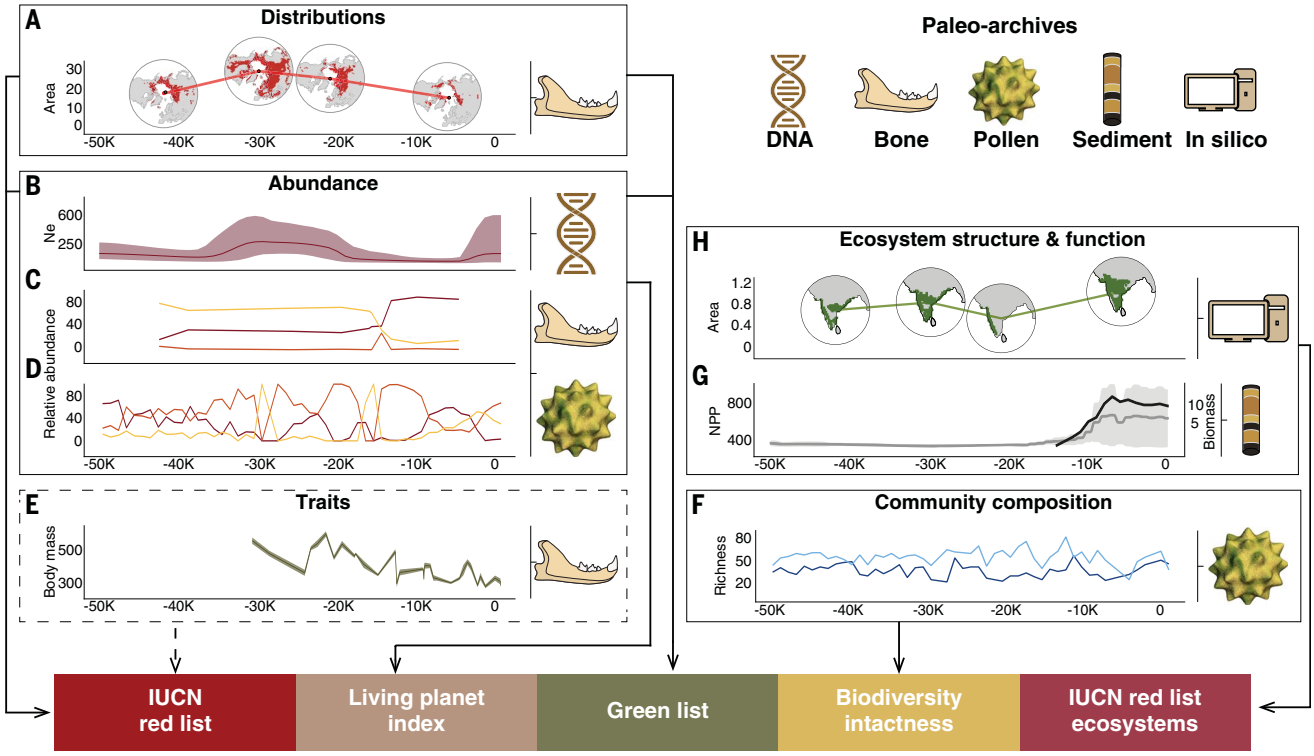


Fig. 4. Paleo-archives for informing conservation under climate change. (**A** to **H**) Paleo-archives (DNA, fossils, and sediments, including in silico computational reconstructions) provide "log books" of changes in species' (A) distributions, [(B) to (D)] abundances, (E) traits, (F) community compositions, and [(G) and (H)] ecosystem structures and functions. Projected (A) geographic range ($km^2 \times 10^6$) and (B) effective population size ($N_e \times 1000$) of muskox (*Ovibos moschatus*); (C) Mammal relative abundance, southern Australia [*Pseudomys australis* (red), *Rattus fuscipes* (orange), and *Pseudomys apodemoides* (yellow)]; (D) Plant relative abundance, southern Italy [*Albies alba* (red), *Fagus sylvatica* (orange), and *Carpinus betulus* (yellow)]; (E) Woodrat (*Neotoma cinerea*) body mass, Great Basin, United States; (F) Taxonomic richness of plants in southern Italy (dark blue) and Guatemala (light blue); (G) Net primary productivity, North America (gray), and plant biomass (black), eastern Canada; (H) Projected area ($km^2 \times 10^6$) of savanna ecosystem, India. Relevant conservation schemes are the IUCN Red List, Living Planet Index, and IUCN Green List (abundance, distribution, and traits); biodiversity intactness (community composition); and IUCN Red List ecosystems (ecosystem structure and function). The *x* axis shows time before the present in thousands of years. Trajectories in (A) to (H) are detailed in appendix S1. [Image credits: LynxVector, Knorre, turtle_stock, flaticon]

strategies. However, if applied indiscriminately, this could result in extinctions of species whose persistence may depend on populations close to the periphery of their historical ranges (according to paleo-archives) (*76*). Inferences of range shifts during the late Quaternary offer distinct opportunities to develop an improved theory of population declines, allowing direct tests of the respective circumstances under which geographic ranges collapse first along the periphery or in the range interior.

Communities and ecosystems

The paleoecological record provides a wealth of data documenting properties of past communities and ecosystems and how they have changed at local to regional scales during the late Quaternary (*33*, *54*, *77*). Environmental DNA can (alongside fossils) document temporal dynamics of communities and ecosystems across hundreds to many thousands of years, helping to assess the resilience of ecosystem services to climate variation and change (*78*). Community reconstructions of taxonomic richness and evenness for plants and animals using paleo-archives (Fig. 4) can be used to calculate "biodiversity intactness" (*79*) and thus guide policies for reducing future biodiversity loss (*69*). This is done by quantifying change in the diversity of a wide assortment of organisms within a given geographical area after a climatic shift or other environmental disturbance. Furthermore, paleo-data can give insight into the relative roles of biotic and abiotic controls on ecosystem properties, including nutrient cycling (*19*), net primary productivity, and plant biomass (Fig. 4) over periods of stability and disturbance. These paleo-enabled measures of ecosystem function can reveal thresholds for ecosystem collapse and inform recent protocols for assessing ecosystem-level threat status under climate change (*80*).

Processes regulating climate-biodiversity dynamics

The challenge of synthesizing disparate evidence from paleo-archives—to assess threats to species and ecosystems from climatic change, and potential consequences of their loss—is being addressed by using process-based (theory- and data-driven) simulation models. These approaches, which run at fine temporal and spatial scales and across large geographical extents, open windows into climate-biodiversity dynamics during the late Quaternary (*81*). By directly capturing spatiotemporal variations in biodiversity at biologically relevant spatiotemporal scales, simulation models provide improved ways to establish ecological baselines and to understand long-term ecological and evolutionary responses to climatic shifts and anthropogenic activities (*82*). They can be used to disentangle multiple drivers of biodiversity change and infer causality, making them particularly suited to guide decisions regarding the pace of change and desired states of ecological systems.

These process-based simulation models are increasingly useful for assessing the relative importance of ecological and evolutionary responses to different spatiotemporal scales of past climatic and environmental change in shaping different levels of biological organization.

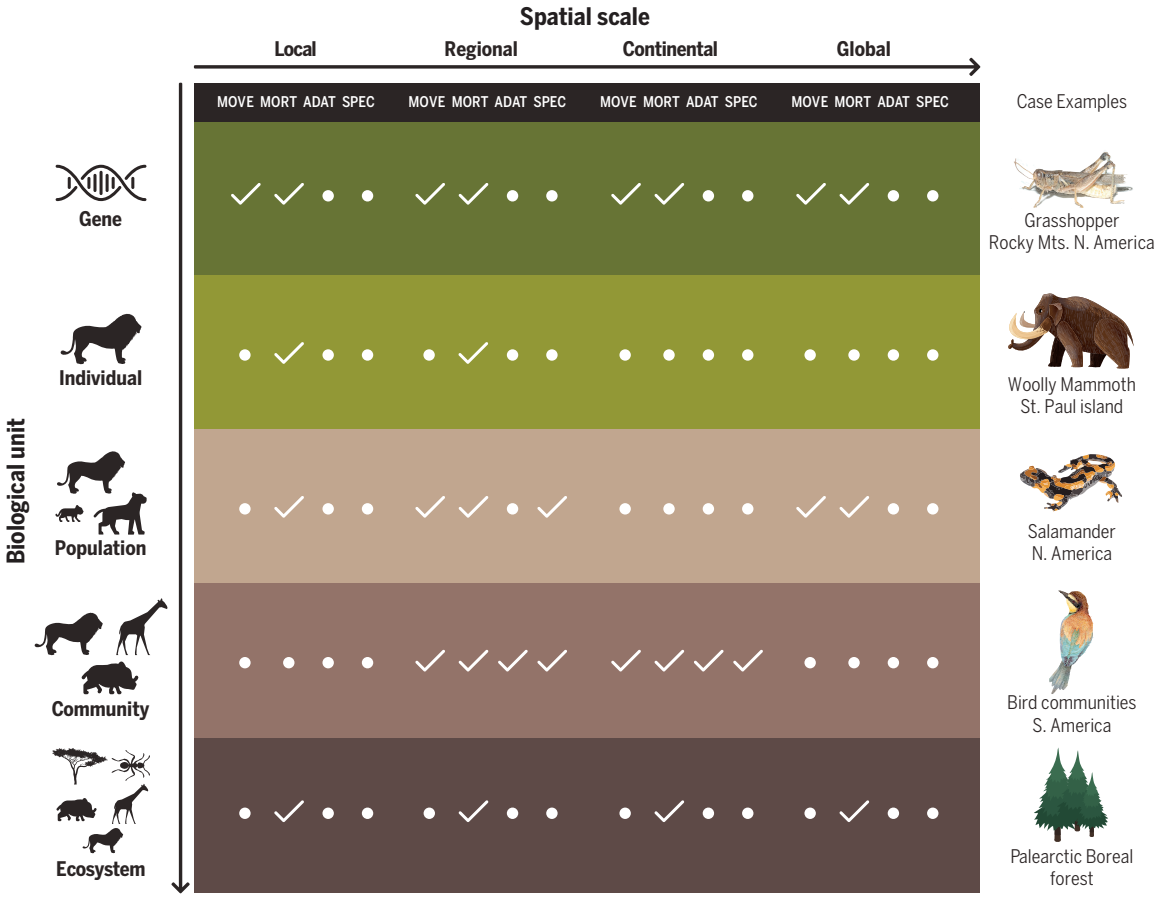


Fig. 5. Simulating mechanisms of past climate-biodiversity dynamics. Process-explicit models built across different levels of biological organization (gene, individual, population, community, and ecosystem) and spatial scales (local, regional, continental, and global). Ticks indicate ecological and evolutionary processes that have been modeled at these different biological and spatial scales: MOVE, movement; MORT, mortality; ADAT adaptation; and SPEC, speciation. Dots indicate an absence of information. Illustrations show exemplar case studies for different biological organizations: grasshopper (*Melanoplus oregonensis*), Rocky Mountains of North America; woolly mammoth (*Mammuthus primigenius*), St. Paul Island; a salamander (*Plethodontid sp.*), North America; bird communities in South America; and boreal forests in the Palearctic. Details of studies are provided in (*100*). [Image credits: Samuel Barrena, tbob, shaineast, HN Works, flaticon]

Driver-state relationships that can now be simulated include effects of climatic change on migration, adaptation, extinction, and speciation (Fig. 5). However, the number and combination of processes simulated depends on the spatial scale and unit of biological organization being modeled (Fig. 5). Such simulations—which unpack complex patterns of biodiversity, driven by multiple biological processes and agents—enable biotic responses to environmental change to be better contextualized and integrated into future biodiversity management (*13*).

Process-based models also permit in silico experiments on the biological consequences of rates of past climatic and environmental changes, allowing counterfactual hypotheses of eco-evolutionary dynamics to be tested (*82*). Model parameters can be held constant to understand the isolated or interacting effects of other dynamical processes on emergent patterns (*81*). Detailed scenario comparisons have been used to investigate competing explanations for ecological regime shifts in the Holocene (*83*), migration patterns of anatomically modern humans since the LIG (*84*), and extirpation patterns during the Pleistocene-Holocene transition (*85*).

Pattern-orientated modeling (POM) of complex paleoecological systems is emerging as a powerful tool (*82*). Spatiotemporal patterns inferred from the paleo-record are used as filters for evaluating whether a model is adequate in its structure and parameterization to simulate the underlying mechanisms. A requirement of POM is that its state variables are expressed in equivalent units to target variables inferred (or measured directly) from paleo-archives. Using paleo-enabled EBVs as targets in POM analysis improves the realism of selected models (those with sufficient structural complexity and adequate parameterization) for conservation management and forecasting.

Looking ahead

Climatologists are taking full advantage of the long-term history of the planet as recorded in paleo-archives to understand mechanisms of long-term climate forcing, quantify trends, and develop scenarios of future climate change (*86*). A wider usage of paleo-archives for quantifying and reporting climate-biodiversity dynamics, particularly in regions where past and future climate change is likely to be similar (Fig. 3), requires an expansion of the taxonomic, spatial, and temporal extent of current paleo-archives. Community-scale genome sequencing of biological remains that are preserved in the permafrost, ice-cores, marine and lake deposits, and coprolites provides solutions for increasing the taxonomic coverage of paleo-samples, often improving knowledge of the geographical distribution of species and ecological communities (*87*). These inferences will be strengthened through robust knowledge of post mortem processes of fossilization (*25*). An expansion of paleo-ecological information in open-access global databases (*88*), including geo-referenced species- and community-level ancient DNA data, will further increase the geographical representation

and temporal coverage of paleo-enabled conservation metrics.

Methods for extending the temporal coverage of past climate-biodiversity dynamics beyond the ~50,000-year age-limit of radiocarbon dating include uranium-thorium dating, optically stimulated luminescence, paleo-magnetism, and infrared stimulated luminescence. However, dates for climatic and biotic events derived from these approaches are generally not resolved at human-relevant time scales (owing to dating limitations and uncertainty), limiting the assignment of ecological shifts to rapid climatic changes during periods earlier than ~30,000 years ago. Tephra layers from volcanic activity, preserved in ice and sediment, have the potential to align paleoenvironmental proxies in space and time for the LD and earlier, making them particularly useful for interpreting biodiversity and ecosystem responses to regional climatic change (*89*).

Individual paleoclimate proxies only approximate broader-scale changes in climatic conditions (*42*); hence, a more thorough understanding of the paleoecology of species and their ecosystems will result from higher spatiotemporal-resolution paleoclimate projections from earth-systems models that use continuous simulations of the transient climate from the LIG to present (*90*). Including improved reconstructions of solar variability, volcanic eruptions, and land-use during the Holocene in these transient simulations will provide a more complete picture of temporal change in regional climates over this period.

Although modern DNA sequencing technologies can generate genetic-based estimates of population change (such as through nuclear single-nucleotide polymorphisms and mitochondrial markers) and adaptive potential to climate change (allelic diversity) (*91*), they have been primarily applied to contemporary or 20th-century samples (*92*). Genomic information extracted from well-preserved fossils is now permitting changes in genetic diversity to be estimated across extended time periods (back at least to the LD) (*93*), informing research on climate change susceptibility and providing preindustrial baseline genomic information for conservation programs, including genetic rescue. Aligning population-level genomic information on inbreeding and connectivity from ancient samples (*94*) with paleoenvironmental proxies presents opportunities to infer magnitudes and rates of change in genetic-based conservation metrics under climate change. Advances that allow changes in allelic diversity to be uncovered from ancient genomes could provide an additional class of paleo-enabled EBV ("genetic composition") relevant to assessing climate-sensitivity and threat status of different taxa.

Decreasing costs of recovering genomic-scale endogenous DNA from fossils are facilitating the compilation of more robust and taxonomically diverse data sets (*95*). This will likely improve understanding of how gene function and gene expression regulate the abilities and speeds at which in situ populations evolve under different rates and magnitudes of climate change, making the inclusion of genomic adaptation in conservation schemes and metrics more feasible (*91*). Indirect paleogenomic evidence suggests a strong potential for fast genomic adaptation to rapid climate change (*96*). However, the generality of these findings (and their potential relevance to rapid warming events) is unknown, given that they are based on a single taxon and small sample sizes.

Simulation and advanced empirical approaches that compare mechanisms of change in nearby regions with different histories of human colonization and land use are likely to further improve knowledge of the synergistic roles of humans and climate in the distributions of taxonomically diverse organisms and communities in space and time. Human-induced landscape transformation on the hundreds of islands in East Polynesia, and potentially Madagascar, did not occur with any severity until within the past millennium (*97*, *98*). Detailed comparisons of long-term biodiversity and ecosystem change on these islands, in comparison and contrast with those on islands of comparable size that have longer human histories (such as Fiji, Cuba, and Tasmania), could prove enlightening.

REFERENCES AND NOTES

1. W. Steffen *et al*., Trajectories of the Earth System in the Anthropocene. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8252–8259 (2018). doi: 10.1073/pnas.1810141115; pmid: 30082409
2. Working Group II, *AR5 Climate Change 2014: Impacts, Adaptation, and Vulnerability* (Intergovernmental Panel on Climate Change, 2014).
3. T. P. Dawson, S. T. Jackson, J. I. House, I. C. Prentice, G. M. Mace, Beyond predictions: Biodiversity conservation in a changing climate. *Science* **332**, 53–58 (2011). doi: 10.1126/science.1200303; pmid: 21454781
4. M. C. Urban *et al*., Improving the forecast for biodiversity under climate change. *Science* **353**, aad8466 (2016). doi: 10.1126/science.aad8466; pmid: 27609898
5. W. B. Foden *et al*., Climate change vulnerability assessment of species. *Wiley Interdiscip. Rev. Clim. Change* **10**, e551 (2018). doi: 10.1002/wcc.551
6. G. P. Dietl, K. W. Flessa, *Conservation Paleobiology: Science and Practice*. (Univ. Chicago Press, 2017).
7. D. Nogués-Bravo *et al*., Cracking the code of biodiversity responses to past climate change. *Trends Ecol. Evol.* **33**, 765–776 (2018). doi: 10.1016/j.tree.2018.07.005; pmid: 30173951
8. C. N. Johnson *et al*., Biodiversity losses and conservation responses in the Anthropocene. *Science* **356**, 270–275 (2017). doi: 10.1126/science.aam9317; pmid: 28428393
9. H. M. Pereira *et al*., Scenarios for global biodiversity in the 21st century. *Science* **330**, 1496–1501 (2010). doi: 10.1126/science.1196624; pmid: 20978282
10. K. D. Burke *et al*., Pliocene and Eocene provide best analogs for near-future climates. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 13288–13293 (2018). doi: 10.1073/pnas.1809600115; pmid: 30530685
11. A. Cooper *et al*., Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. *Science* **349**, 602–606 (2015). doi: 10.1126/science.aac4315; pmid: 26250679
12. S. T. Jackson, J. L. Betancourt, R. K. Booth, S. T. Gray, Ecology and the ratchet of events: Climate variability, niche dimensions, and species distributions. *Proc. Natl. Acad. Sci. U.S.A.* **106** (Suppl 2), 19685–19692 (2009). doi: 10.1073/pnas.0901644106; pmid: 19805104
13. D. A. Fordham *et al*., Predicting and mitigating future biodiversity loss using long-term ecological proxies. *Nat. Clim. Chang.* **6**, 909–916 (2016). doi: 10.1038/nclimate3086
14. O. Hagen, T. Andermann, T. B. Quental, A. Antonelli, D. Silvestro, Estimating age-dependent extinction: Contrasting evidence from fossils and phylogenies. *Syst. Biol.* **67**, 458–474 (2018). doi: 10.1093/sysbio/syx082; pmid: 29069434
15. W. Dansgaard *et al*., Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364**, 218–220 (1993). doi: 10.1038/364218a0
16. S. C. Brown, T. M. L. Wigley, B. L. Otto-Bliesner, C. Rahbek, D. A. Fordham, Persistent Quaternary climate refugia are hospices for biodiversity in the Anthropocene. *Nat. Clim. Chang.* **10**, 244–248 (2020). doi: 10.1038/s41558-019-0682-7
17. T. J. Crowley, Are there any satisfactory geologic analogs for a future greenhouse warming? *J. Clim.* **3**, 1282–1292 (1990). doi: 10.1175/1520-0442(1990)003<1282:ATASGA>2.0.CO;2
18. J. L. Gill, J. W. Williams, S. T. Jackson, K. B. Lininger, G. S. Robinson, Pleistocene megafaunal collapse, novel plant communities, and enhanced fire regimes in North America. *Science* **326**, 1100–1103 (2009). doi: 10.1126/science.1179504; pmid: 19965426
19. E. S. Jeffers *et al*., Plant controls on Late Quaternary whole ecosystem structure and function. *Ecol. Lett.* **21**, 814–825 (2018). doi: 10.1111/ele.12944; pmid: 29601664
20. G. P. Dietl *et al*., Conservation paleobiology: Leveraging knowledge of the past to inform conservation and restoration. *Annu. Rev. Earth Planet. Sci.* **43**, 79–103 (2015). doi: 10.1146/annurev-earth-040610-133349
21. C. Nolan *et al*., Past and future global transformation of terrestrial ecosystems under climate change. *Science* **361**, 920–923 (2018). doi: 10.1126/science.aan5360; pmid: 30166491
22. A. D. Barnosky *et al*., Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science* **355**, eaah4787 (2017). doi: 10.1126/science.aah4787; pmid: 28183912
23. S. T. Turvey, E. E. Saupe, Insights from the past: Unique opportunity or foreign country? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20190208 (2019). doi: 10.1098/rstb.2019.0208; pmid: 31679483
24. D. A. Fordham, B. W. Brook, C. Moritz, D. Nogués-Bravo, Better forecasts of range dynamics using genetic data. *Trends Ecol. Evol.* **29**, 436–443 (2014). doi: 10.1016/j.tree.2014.05.007; pmid: 24951394
25. S. M. Kidwell, Biology in the Anthropocene: Challenges and insights from young fossil records. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4922–4929 (2015). doi: 10.1073/pnas.1403660112; pmid: 25901315
26. H. M. Pereira *et al*., Essential biodiversity variables. *Science* **339**, 277–278 (2013). doi: 10.1126/science.1229931; pmid: 23329036
27. H. Fischer *et al*., Palaeoclimate constraints on the impact of 2 °C anthropogenic warming and beyond. *Nat. Geosci.* **11**, 474–485 (2018). doi: 10.1038/s41561-018-0146-0
28. H. Renssen, H. Seppa, X. Crosta, H. Goosse, D. M. Roche, Global characterization of the Holocene Thermal Maximum. *Quat. Sci. Rev.* **48**, 7–19 (2012). doi: 10.1016/j.quascirev.2012.05.022
29. M. S. McGlone, C. S. M. Turney, J. M. Wilmshurst, J. Renwick, K. Pahnke, Divergent trends in land and ocean temperature in the Southern Ocean over the past 18,000 years. *Nat. Geosci.* **3**, 622–626 (2010). doi: 10.1038/ngeo931
30. N. H. Bigelow, Climate change and Arctic ecosystems: Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. *J. Geophys. Res.* **108** (D19), 8170 (2003). doi: 10.1029/2002JD002558
31. J. W. Williams, B. Shuman, P. J. Bartlein, Rapid responses of the prairie-forest ecotone to early Holocene aridity in mid-continental North America. *Global Planet. Change* **66**, 195–207 (2009). doi: 10.1016/j.gloplacha.2008.10.012
32. J. Watrin, A.-M. Lézine, C. Hély, Plant migration and plant communities at the time of the "green Sahara". *C. R. Geosci.* **341**, 656–670 (2009). doi: 10.1016/j.crte.2009.06.007
33. F. P. Díaz *et al*., Multiscale climate change impacts on plant diversity in the Atacama Desert. *Glob. Change Biol.* **25**, 1733–1745 (2019). doi: 10.1111/gcb.14583; pmid: 30706600

34. A. de Vernal, C. Hillaire-Marcel, Natural variability of Greenland climate, vegetation, and ice volume during the past million years. *Science* **320**, 1622–1625 (2008). doi: 10.1126/science.1153929; pmid: 18566284
35. J. C. Larrasoaña, A. P. Roberts, E. J. Rohling, Dynamics of green Sahara periods and their role in hominin evolution. *PLOS ONE* **8**, e76514 (2013). doi: 10.1371/journal.pone.0076514; pmid: 24146882
36. D. C. Schreve, Differentiation of the British late Middle Pleistocene interglacials: The evidence from mammalian biostratigraphy. *Quat. Sci. Rev.* **20**, 1693–1705 (2001). doi: 10.1016/S0277-3791(01)00033-6
37. J. E. King, J. J. Saunders, Geochelone in Illinois and the Illinoian Sangamonian vegetation of the type region. *Quat. Res.* **25**, 89–99 (1986). doi: 10.1016/0033-5894(86)90046-3
38. U. Salzmann, A. M. Haywood, D. J. Lunt, P. J. Valdes, D. J. Hill, A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Glob. Ecol. Biogeogr.* **17**, 432–447 (2008). doi: 10.1111/j.1466-8238.2008.00381.x
39. A. C. Ashworth, D. J. Cantrill, Neogene vegetation of the Meyer desert formation (Sirius Group) Transantarctic Mountains, Antarctica. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **213**, 65–82 (2004). doi: 10.1016/S0031-0182(04)00359-1
40. S. T. Jackson, J. T. Overpeck, Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* **26** (S4), 194–220 (2000). doi: 10.1017/S0094837300026932
41. G. T. Pecl *et al*., Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017). doi: 10.1126/science.aai9214; pmid: 28360268
42. P. U. Clark *et al*., Global climate evolution during the last deglaciation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1134–E1142 (2012). doi: 10.1073/pnas.1116619109; pmid: 22331892
43. J. P. Steffensen *et al*., High-resolution Greenland ice core data show abrupt climate change happens in few years. *Science* **321**, 680–684 (2008). doi: 10.1126/science.1157707; pmid: 18566247
44. R. B. Alley, Ice-core evidence of abrupt climate changes. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1331–1334 (2000). doi: 10.1073/pnas.97.4.1331; pmid: 10677460
45. H. Renssen, R. F. B. Isarin, The two major warming phases of the last deglaciation at ~14.7 and ~11.5 ka cal BP in Europe: Climate reconstructions and AGCM experiments. *Global Planet. Change* **30**, 117–153 (2001). doi: 10.1016/S0921-8181(01)00082-0
46. B. G. Holt *et al*., An update of Wallace's zoogeographic regions of the world. *Science* **339**, 74–78 (2013). doi: 10.1126/science.1228282; pmid: 23258408
47. E. D. Lorenzen *et al*., Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359–364 (2011). doi: 10.1038/nature10574; pmid: 22048313
48. A. Nadachowski, G. Lipecki, U. Ratajczak, K. Stefaniak, P. Wojtal, Dispersal events of the saiga antelope (*Saiga tatarica*) in Central Europe in response to the climatic fluctuations in MIS 2 and the early part of MIS 1. *Quat. Int.* **420**, 357–362 (2016). doi: 10.1016/j.quaint.2015.11.068
49. L. Dalén *et al*., Ancient DNA reveals lack of postglacial habitat tracking in the arctic fox. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6726–6729 (2007). doi: 10.1073/pnas.0701341104; pmid: 17420452
50. J. L. Blois, J. L. McGuire, E. A. Hadly, Small mammal diversity loss in response to late-Pleistocene climatic change. *Nature* **465**, 771–774 (2010). doi: 10.1038/nature09077; pmid: 20495547
51. S. T. Jackson, C. Weng, Late quaternary extinction of a tree species in eastern North America. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13847–13852 (1999). doi: 10.1073/pnas.96.24.13847; pmid: 10570161
52. C. E. Doughty *et al*., Global nutrient transport in a world of giants. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 868–873 (2016). doi: 10.1073/pnas.1502549112; pmid: 26504209
53. S. A. Zimov *et al*., Steppe-tundra transition—A herbivore-driven biome shift at the end of the Pleistocene. *Am. Nat.* **146**, 765–794 (1995). doi: 10.1086/285824
54. E. S. Jeffers, S. Nogué, K. J. Willis, The role of palaeoecological records in assessing ecosystem services. *Quat. Sci. Rev.* **112**, 17–32 (2015). doi: 10.1016/j.quascirev.2014.12.018
55. F. K. Fye, D. W. Stahle, E. R. Cook, Paleoclimatic analogs to twentieth-century moisture regimes across the United States. *Bull. Am. Meteorol. Soc.* **84**, 901–910 (2003). doi: 10.1175/BAMS-84-7-901
56. M. R. Lesser, S. T. Jackson, Contributions of long-distance dispersal to population growth in colonising *Pinus ponderosa* populations. *Ecol. Lett.* **16**, 380–389 (2013). doi: 10.1111/ele.12053; pmid: 23279647
57. A. Hampe, M.-H. Pemonge, R. J. Petit, Efficient mitigation of founder effects during the establishment of a leading-edge oak population. *Proc. Biol. Sci.* **280**, 20131070 (2013). doi: 10.1098/rspb.2013.1070; pmid: 23782887
58. E. M. Herring, D. G. Gavin, S. Z. Dobrowski, M. Fernandez, F. S. Hu, Ecological history of a long-lived conifer in a disjunct population. *J. Ecol.* **106**, 319–332 (2018). doi: 10.1111/1365-2745.12826
59. J. S. Elleouet, S. N. Aitken, Long-distance pollen dispersal during recent colonization favors a rapid but partial recovery of genetic diversity in *Picea sitchensis*. *New Phytol.* **222**, 1088–1100 (2019). doi: 10.1111/nph.15615; pmid: 30485444
60. A. Dawson *et al*., Quantifying trends and uncertainty in prehistoric forest composition in the upper Midwestern United States. *Ecology* **100**, e02856 (2019). doi: 10.1002/ecy.2856; pmid: 31381148
61. R. K. Booth, S. Brewer, M. Blaauw, T. A. Minckley, S. T. Jackson, Decomposing the mid-Holocene Tsuga decline in eastern North America. *Ecology* **93**, 1841–1852 (2012). doi: 10.1890/11-2062.1; pmid: 22928413
62. J. H. Lawton, Range, population abundance and conservation. *Trends Ecol. Evol.* **8**, 409–413 (1993). doi: 10.1016/0169-5347(93)90043-O; pmid: 21236213
63. T. Ramiadantsoa, M. A. Stegner, J. W. Williams, A. R. Ives, The potential role of intrinsic processes in generating abrupt and quasi-synchronous tree declines during the Holocene. *Ecology* **100**, e02579 (2019). doi: 10.1002/ecy.2579; pmid: 30707453
64. M. G. Turner *et al*., Climate change, ecosystems and abrupt change: Science priorities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190105 (2020). doi: 10.1098/rstb.2019.0105; pmid: 31983326
65. S. D. Crausbay, P. E. Higuera, D. G. Sprugel, L. B. Brubaker, Fire catalyzed rapid ecological change in lowland coniferous forests of the Pacific Northwest over the past 14,000 years. *Ecology* **98**, 2356–2369 (2017). doi: 10.1002/ecy.1897; pmid: 28500791
66. C. I. Millar, N. L. Stephenson, Temperate forest health in an era of emerging megadisturbance. *Science* **349**, 823–826 (2015). doi: 10.1126/science.aaa9933; pmid: 26293954
67. W. D. Kissling *et al*., Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev. Camb. Philos. Soc.* **93**, 600–625 (2018). doi: 10.1111/brv.12359; pmid: 28766908
68. F. Saltré *et al*., Climate change not to blame for late Quaternary megafauna extinctions in Australia. *Nat. Commun.* **7**, 10511 (2016). doi: 10.1038/ncomms10511; pmid: 26821754
69. G. M. Mace *et al*., Aiming higher to bend the curve of biodiversity loss. *Nat. Sustain.* **1**, 448–451 (2018). doi: 10.1038/s41893-018-0130-0
70. C. F. Clements, A. Ozgul, Indicators of transitions in biological systems. *Ecol. Lett.* **21**, 905–919 (2018). doi: 10.1111/ele.12948; pmid: 29601665
71. V. Dakos *et al*., Slowing down as an early warning signal for abrupt climate change. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14308–14312 (2008). doi: 10.1073/pnas.0802430105 pmid: 18787119
72. K. J. Willis *et al*., How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 175–186 (2007). doi: 10.1098/rstb.2006.1977; pmid: 17255027
73. W. B. Foden *et al*., Identifying the world's most climate change vulnerable species: A systematic trait-based assessment of all birds, amphibians and corals. *PLOS ONE* **8**, e65427 (2013). doi: 10.1371/journal.pone.0065427; pmid: 23950785
74. T. L. Spanbauer *et al*., Body size distributions signal a regime shift in a lake ecosystem. *Proc. Biol. Sci.* **283**, 20160249 (2016). doi: 10.1098/rspb.2016.0249; pmid: 27335415
75. M. Grace *et al*., Using historical and palaeoecological data to inform ambitious species recovery targets. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20190297 (2019). doi: 10.1098/rstb.2019.0297; pmid: 31679497
76. R. Channell, M. V. Lomolino, Dynamic biogeography and conservation of endangered species. *Nature* **403**, 84–86 (2000). doi: 10.1038/47487; pmid: 10638757
77. S. T. Jackson, J. L. Blois, Community ecology in a changing environment: Perspectives from the Quaternary. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4915–4921 (2015). doi: 10.1073/pnas.1403664111; pmid: 25901314
78. M. Bálint *et al*., Environmental DNA Time Series in Ecology. *Trends Ecol. Evol.* **33**, 945–957 (2018). doi: 10.1016/j.tree.2018.09.003; pmid: 30314916
79. R. J. Scholes, R. Biggs, A biodiversity intactness index. *Nature* **434**, 45–49 (2005). doi: 10.1038/nature03289; pmid: 15744293
80. D. A. Keith *et al*., The IUCN Red List of Ecosystems: Motivations, challenges, and applications. *Conserv. Lett.* **8**, 214–226 (2015). doi: 10.1111/conl.12167
81. T. F. Rangel *et al*., Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science* **361**, eaar5452 (2018). doi: 10.1126/science.aar5452; pmid: 30026200
82. G. L. W. Perry, J. Wainwright, T. R. Etherington, J. M. Wilmshurst, Experimental simulation: Using generative modeling and palaeoecological data to understand human-environment interactions. *Front. Ecol. Evol.* **4**, (2016). doi: 10.3389/fevo.2016.00109
83. T. A. Prowse, C. N. Johnson, C. J. Bradshaw, B. W. Brook, An ecological regime shift resulting from disrupted predator-prey interactions in Holocene Australia. *Ecology* **95**, 693–702 (2014). doi: 10.1890/13-0746.1; pmid: 24804453
84. A. Timmermann, T. Friedrich, Late Pleistocene climate drivers of early human migration. *Nature* **538**, 92–95 (2016). doi: 10.1038/nature19365; pmid: 27654920
85. Y. Wang *et al*., Mechanistic modeling of environmental drivers of woolly mammoth carrying capacity declines on St. Paul Island. *Ecology* **99**, 2721–2730 (2018). doi: 10.1002/ecy.2524; pmid: 30365160
86. S. P. Harrison *et al*., Evaluation of CMIP5 palaeo-simulations to improve climate projections. *Nat. Clim. Chang.* **5**, 735–743 (2015). doi: 10.1038/nclimate2649
87. E. Willerslev *et al*., Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**, 47–51 (2014). doi: 10.1038/nature12921; pmid: 24499916
88. D. Kaufman *et al*., A global database of Holocene paleotemperature records. *Sci. Data* **7**, 115 (2020). doi: 10.1038/s41597-020-0445-3; pmid: 32286335
89. D. J. Lowe, Tephrochronology and its application: A review. *Quat. Geochronol.* **6**, 107–153 (2011). doi: 10.1016/j.quageo.2010.08.003
90. B. L. Otto-Bliesner *et al*., The PMIP4 contribution to CMIP6 – Part 2: Two interglacials, scientific objective and experimental design for Holocene and Last Interglacial simulations. *Geosci. Model Dev.* **10**, 3979–4003 (2017). doi: 10.5194/gmd-10-3979-2017
91. D. Díez-Del-Molino, F. Sánchez-Barreiro, I. Barnes, M. T. P. Gilbert, L. Dalén, Quantifying temporal genomic erosion in endangered species. *Trends Ecol. Evol.* **33**, 176–185 (2018). doi: 10.1016/j.tree.2017.12.002; pmid: 29289355
92. S. Feng *et al*., The genomic footprints of the fall and recovery of the crested ibis. *Curr. Biol.* **29**, 340–349.e7 (2019). doi: 10.1016/j.cub.2018.12.008; pmid: 30639104
93. E. Palkopoulou *et al*., Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Curr. Biol.* **25**, 1395–1400 (2015). doi: 10.1016/j.cub.2015.04.007; pmid: 25913407
94. A. Fages *et al*., Tracking five millennia of horse management with extensive ancient genome time series. *Cell* **177**, 1419–1435.e31 (2019). doi: 10.1016/j.cell.2019.03.049; pmid: 31056281
95. B. Shapiro, M. Hofreiter, A paleogenomic perspective on evolution and gene function: New insights from ancient DNA. *Science* **343**, 1236573 (2014). doi: 10.1126/science.1236573; pmid: 24458647
96. P. Librado *et al*., Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6889–E6897 (2015). doi: 10.1073/pnas.1513696112; pmid: 26598656
97. A. Anderson *et al*., New evidence of megafaunal bone damage indicates late colonization of Madagascar. *PLOS ONE* **13**, e0204368 (2018). doi: 10.1371/journal.pone.0204368; pmid: 30303989

98. J. M. Wilmshurst, T. L. Hunt, C. P. Lipo, A. J. Anderson, High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1815–1820 (2011). doi: 10.1073/pnas.1015876108; pmid: 21187404
99. J. D. Shakun *et al.*, Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484**, 49–54 (2012). doi: 10.1038/nature10915; pmid: 22481357
100. D. A. Fordham *et al.*, Paleo-archives for informing conservation under climate change. *Figshare* (2020); doi: 10.25909/5eaa200e1b2e2

ACKNOWLEDGMENTS

We thank the IUCN SSC Climate Change Specialist Group for sponsoring a scientific discussion meeting. Full image credits and sources are listed in the supplementary materials. **Funding:** D.A.F. acknowledges funding from the University of Adelaide's Environment Institute, the Australian Research Council (FT140101192, DP180102392), and a residency fellowship from Danmarks Nationalbank; D.N-B. acknowledges funding from DFF2-DEMOCHANGE and DNRF-CMEC (DNRF96); and D.D-J and A.S. acknowledge funding from Villum Investigator (IceFlow; 16572). **Author contributions:** D.A.F., D.N-B., and S.T.J. conceived the idea for the meeting and paper. All authors contributed to writing the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data are available in the main text, references, and supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/eabc5654/suppl/DC1
Appendix S1: Enhanced Figure Legends
Full Image Sources and Credits
References (*101*–*110*)

10.1126/science.abc5654

RESEARCH ARTICLE SUMMARY

MICROBIAL ENGINEERING

A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose

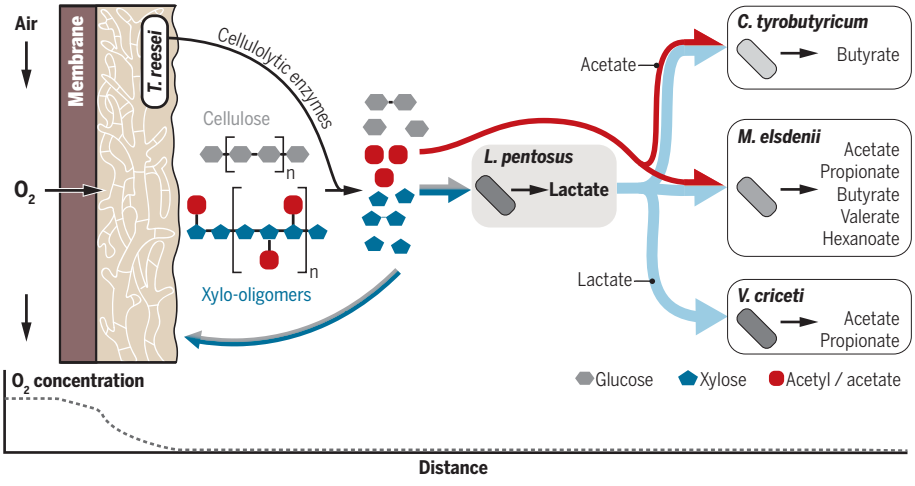
Robert L. Shahab, Simone Brethauer, Matthew P. Davey, Alison G. Smith, Silvia Vignolini, Jeremy S. Luterbacher, Michael H. Studer*

INTRODUCTION: Consortia of microorganisms are a promising alternative to monocultures for complex biotransformations because of their inherent advantages, such as the distribution of the metabolic burden by division of labor, the ability to convert complex substrates more efficiently, and their modularity. Consortia are also inherently more stable and less susceptible to contamination than monocultures. However, achieving stable and reproducible ratios of the different strains and their performance is challenging, especially for microbial consortia where mutual beneficial interactions are missing—for example, if one strain secretes cellulolytic enzymes to provide carbohydrates as a public good. Moreover, the need to find matching fermentation conditions in well-mixed reactors limits the selection of possible consortium members and prevents the harnessing of the full power of artificial consortia.

RATIONALE: To tackle these challenges and to advance the understanding and engineering of synthetic microbial consortia, we attempted to mimic one of the well-known natural mechanisms for consortium stabilization: the spatial organization, along several different gradients, of abiotic factors within a biofilm. To this end, we established a membrane-aerated biofilm reactor, which allows for the creation of an oxygen-replete niche in an otherwise anoxic environment, to harbor artificial cross-kingdom microbial consortia for the consolidated bioprocessing of lignocellulose to short-chain fatty acids (SCFAs). Because many SCFA producers have a limited carbohydrate substrate spectrum, the heterogeneous lignocellulosic sugar mixture is funneled to lactate as the central intermediate.

RESULTS: The lactate platform consists of the aerobic fungus *Trichoderma reesei*, which secretes cellulolytic enzymes and thus provides soluble sugars; facultative anaerobic lactic acid bacteria (LAB); and the obligate anaerobic lactate-fermenting strain, which provides the product. To set up the consortia and to form the necessary oxygen gradient that allows the growth of all members, we applied a sequential inoculation scheme starting with the fungus, followed by LAB, and lastly the lactate fermenter. We found that the metabolic activity of *T. reesei* alone is not sufficient to create an environment with a redox potential suitable for obligate anaerobes; rather, it requires the additional presence of LAB. We tested the feasibility of the lactate platform to produce butyric acid using *Clostridium tyrobutyricum*, which co-consumes lactic acid and acetic acid in a 3:1 molar ratio, thereby valorizing this often unused component of lignocellulose and allowing for a higher carbon yield relative to direct carbohydrate conversion. We were able to produce a stable cocultivation of the three strains, and, with a mixture of cellulose and xylose as model substrate, we obtained 0.35 g of butyric acid per gram of carbohydrates—a value slightly exceeding yields achieved with soluble substrates or by addition of commercial cellulolytic enzymes. With beechwood as substrate, a yield equivalent to 196 kg of butyric acid per metric ton of feedstock was reached. Switching to *Veillonella criceti*, an obligate anaerobic bacterium unable to utilize carbohydrates, a mixture of propionic and acetic acid was produced from lignocellulose, underpinning the suitability of the lactate platform approach to compensate for missing metabolic capabilities of product-forming strains. Longer SCFAs such as valeric and caproic acid, which are more valuable than the C_2 to C_4 SCFAs, could be produced by integrating *Megasphaera elsdenii*, a strict anaerobic rumen bacterium able to perform chain elongation, into the lactate platform. By extending this configuration with *V. criceti*, the ratio of odd- and even-numbered SCFAs could be tuned, and the valeric acid concentration was higher than with the three-member community by a factor of nearly 3. To demonstrate the platform's wide applicability, we identified several other chemicals—including alcohols (ethanol, propanol, butanol, 1,2-propanediol), polyhydroxybutyrate, and lipids—that could be produced through the lactate platform.

CONCLUSION: The lactate platform demonstrates several inherent benefits of synthetic consortia, such as their modularity and their ability to convert complex mixed substrates into valuable platform chemicals. The use of spatial niches in a membrane reactor—not limited to oxygen, but also for other abiotic factors such as temperature, light, or pH—expands the available toolbox for successfully engineering stable and controllable synthetic communities for numerous novel applications. ■



Principle of the lactate platform. A membrane-aerated bioreactor featuring an oxygen gradient (depicted at the bottom) allows for the consolidated bioprocessing of lignocellulose to short-chain fatty acids by a heterogeneous consortium consisting of an aerobic, cellulolytic fungus growing as a biofilm in the oxygen-replete spatial niche, a facultative anaerobic lactic acid bacterium to funnel the carbohydrate mixture to lactate, and a lactate-consuming obligate anaerobic bacterium for product formation.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: michael.studer1@bfh.ch
Cite this article as R. L. Shahab *et al.*, *Science* **369**, eabb1214 (2020). DOI: 10.1126/science.abb1214

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb1214

RESEARCH ARTICLE

MICROBIAL ENGINEERING

A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose

Robert L. Shahab[1,2], Simone Brethauer[2], Matthew P. Davey[3], Alison G. Smith[3], Silvia Vignolini[4], Jeremy S. Luterbacher[1], Michael H. Studer[2]*

Microbial consortia are a promising alternative to monocultures of genetically modified microorganisms for complex biotransformations. We developed a versatile consortium-based strategy for the direct conversion of lignocellulose to short-chain fatty acids, which included the funneling of the lignocellulosic carbohydrates to lactate as a central intermediate in engineered food chains. A spatial niche enabled in situ cellulolytic enzyme production by an aerobic fungus next to facultative anaerobic lactic acid bacteria and the product-forming anaerobes. *Clostridium tyrobutyricum*, *Veillonella criceti*, or *Megasphaera elsdenii* were integrated into the lactate platform to produce 196 kilograms of butyric acid per metric ton of beechwood. The lactate platform demonstrates the benefits of mixed cultures, such as their modularity and their ability to convert complex substrates into valuable biochemicals.

Lignocellulosic biomass is an abundant source of fixed, renewable carbon and is a promising alternative to fossil petroleum as feedstock for a range of chemicals. Short-chain fatty acids (SCFAs) with two to six carbon atoms are currently produced from fossil resources and have been suggested as potential bio-derived platform molecules (*1*). However, the combination of the heterogeneous carbohydrate composition of lignocellulose and the limited metabolic flexibility of many SCFA producers leads to low product yields and undesirable by-product formation (fig. S1A). A more homogeneous carbon source for SCFA producers can be provided by metabolically funneling the heterogeneous mixture of carbohydrates in the biomass to a central platform molecule in a primary fermentation, followed by a secondary conversion to the desired product (fig. S1B). An ideal system would combine the primary and secondary fermentation, as well as the depolymerization of the carbohydrates by cellulolytic enzymes, in a single processing step termed consolidated bioprocessing (CBP) (*2*).

The engineering of such complex reaction cascades into a single microbial host can create a substantial metabolic burden that limits the overall productivity of the system (*3*). Consequently, low carbon fluxes remain a major challenge for CBP approaches based on genetically engineered microorganisms (*4*, *5*), which are often characterized by metabolic imbalances (*6*–*8*). This limitation could be overcome by a metabolic division of labor (*3*), which is also observed in natural consortia, where complex reaction cascades are common and biochemically difficult tasks are catalyzed by multiple species according to the principle of metabolic compartmentalization (*9*). Each member of the community occupies its ecological niche and is specialized in a subtask or exchanges different beneficial molecules such as vitamins or amino acids. However, the high complexity of natural consortia makes the targeted engineering of the desired trait difficult and limits their industrial application for the production of biochemicals. Alternatively, less complex and more defined synthetic consortia can be engineered (*10*) and have been applied for CBP of lignocellulose to produce isobutanol (*11*), ethanol (*12*, *13*), or acetone, butanol, and ethanol (*14*). A major challenge in the design of such bottom-up consortia is the requirement of overlapping tolerance ranges of several abiotic process parameters such as oxygen concentration, pH, and temperature, which limits the selection of possible consortium members (*15*). Thus, the majority of coculture studies reported to date were carried out using only two species (*10*) with similar abiotic requirements, and few examples of artificial consortia with several populations have been reported (*16*–*19*).

Furthermore, under homogeneous growth conditions, population control and stable reproducible performance are challenging to achieve, especially for microbial consortia where mutual beneficial interactions are missing (*9*, *20*, *21*)—for example, if one strain secretes cellulolytic enzymes to provide carbohydrates as a public good (*22*). In such cooperator/cheater communities, different growth rates of consortium members may lead to overgrowth of one population and extinction of the other (*23*). To tackle these challenges, we proposed to switch from a homogeneous to a heterogeneous reactor system that provides defined spatial niches where different abiotic conditions fitted to each consortium member (e.g., oxygen, temperature, light) prevail in the otherwise homogeneous reactor environment (*24*). This enforced niche partitioning can stabilize microbial communities of, for example, an aerobe and facultative anaerobes (*24*–*27*) and enables the cocultivation of microorganisms with nonoverlapping abiotic requirements.

Here, we demonstrate the feasibility of engineering spatial and metabolic niches as a strategy to build and control artificial cross-kingdom microbial consortia and introduce a CBP strategy termed the "lactate platform" where lactic acid is used as a central and versatile intermediate for the direct production of different SCFAs from lignocellulose.

Many SCFA producers are noncellulolytic strains of bacteria with limited flexibility to utilize different carbohydrates, but with the ability to metabolize lactic acid. This ability makes lactic acid a potential central intermediate in a microbial consortium–based CBP strategy aiming at the conversion of the complete carbohydrate fraction of lignocellulose to the target product (fig. S1B). The lactate platform consists of (i) the aerobic fungus *Trichoderma reesei* for the cooperative secretion of cellulolytic enzymes, which provide soluble sugars; (ii) facultative anaerobic lactic acid bacteria (LAB) that are able to convert all types of solubilized sugars to lactic acid; and (iii) the lactate-fermenting anaerobic product-forming strain (Fig. 1A). In contrast to other abiotic factors such as temperature or pH (Fig. 1B), the oxygen requirements of the members of the artificial consortium do not overlap (Fig. 1C). By using a membrane-aerated reactor in combination with the facultative anaerobic LAB, an oxygen gradient is generated in the reactor that allows the redox potential requirements of all strains to be fulfilled (Fig. 1D). To build the consortia, we followed a three-step inoculation scheme: First, the reactor was inoculated with *T. reesei* to allow the formation of a fungal biofilm on top of the membrane (Fig. 1E and fig. S2). After 48 hours, the LAB was inoculated (designated as phase I with the time set to t = 0 hours), followed by the subsequent inoculation of the lactate-fermenting strain(s) after waiting another 48 hours (beginning of phase II). All insoluble substrates were added to the reactor at the time of the fungal inoculation, and soluble substrates were fed to the reactor after inoculation with LAB.

Butyric acid production

To produce butyric acid, we set up a synthetic consortium consisting of *T. reesei*, *Lactobacillus*

[1]Laboratory of Sustainable and Catalytic Processing, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. [2]Laboratory of Biofuels and Biochemicals, School of Agricultural, Forest and Food Sciences, Bern University of Applied Sciences (BFH), CH-3052 Zollikofen, Switzerland. [3]Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK. [4]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK.
*Corresponding author. Email: michael.studer1@bfh.ch

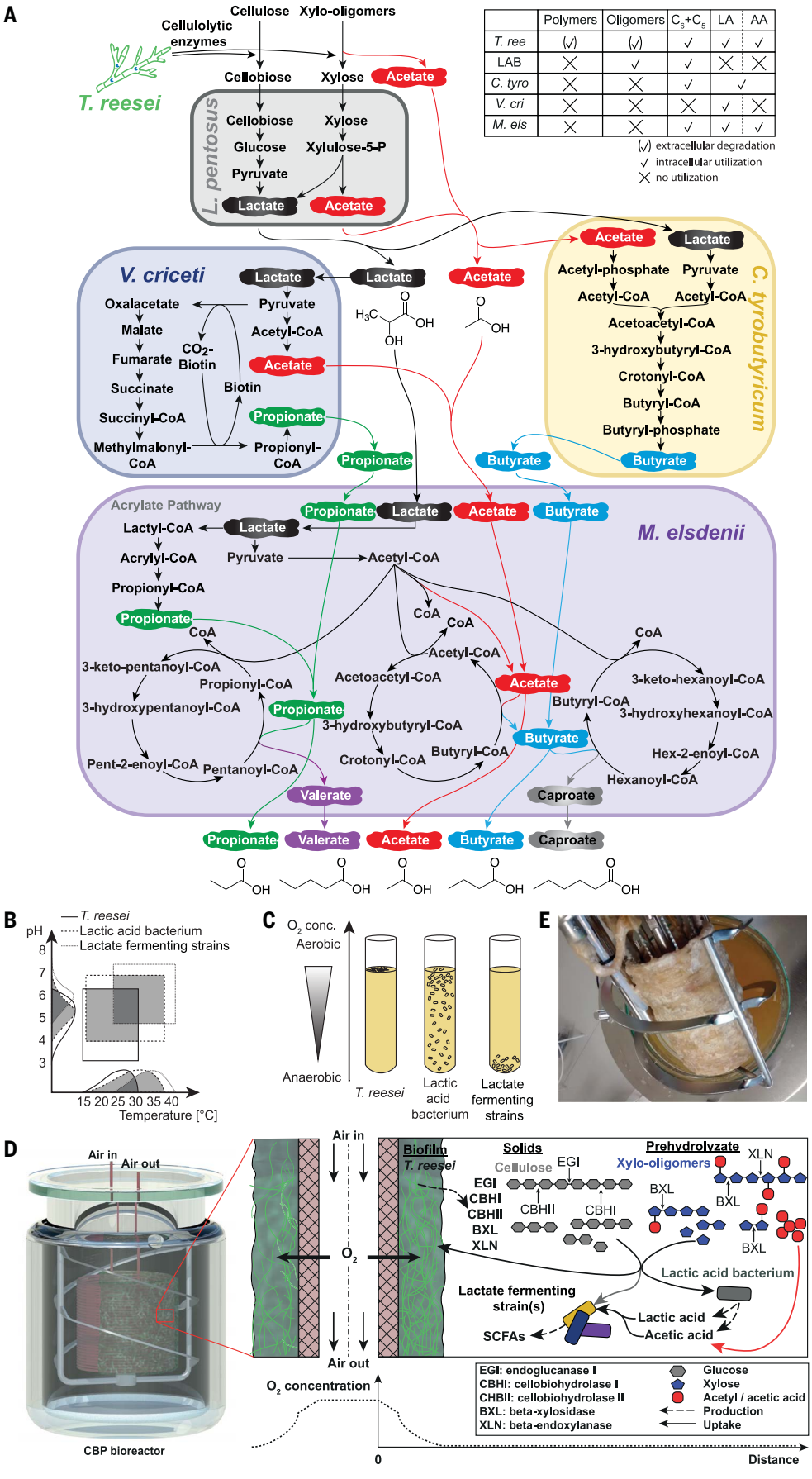


Fig. 1. Setup of the lactate platform. (**A**) Selected metabolic pathways in the community-based CBP using lactate and acetate as central intermediates. The table (upper right) shows the capability of strains to metabolize polymers (cellulose and hemicellulose), oligomers (gluco- and xylo-oligosaccharides), C_5/C_6 sugars, lactate (LA), and acetate (AA). (**B**) Approximate tolerance ranges for pH and temperature of strains (*42*). (**C**) Requirements of strains for dissolved oxygen concentration. (**D**) Schematic representation of the membrane-aerated bioreactor for CBP of lignocellulosic biomass to SCFAs with dissolved oxygen concentration profile. (**E**) Photograph of mature fungal biofilm on the membrane.

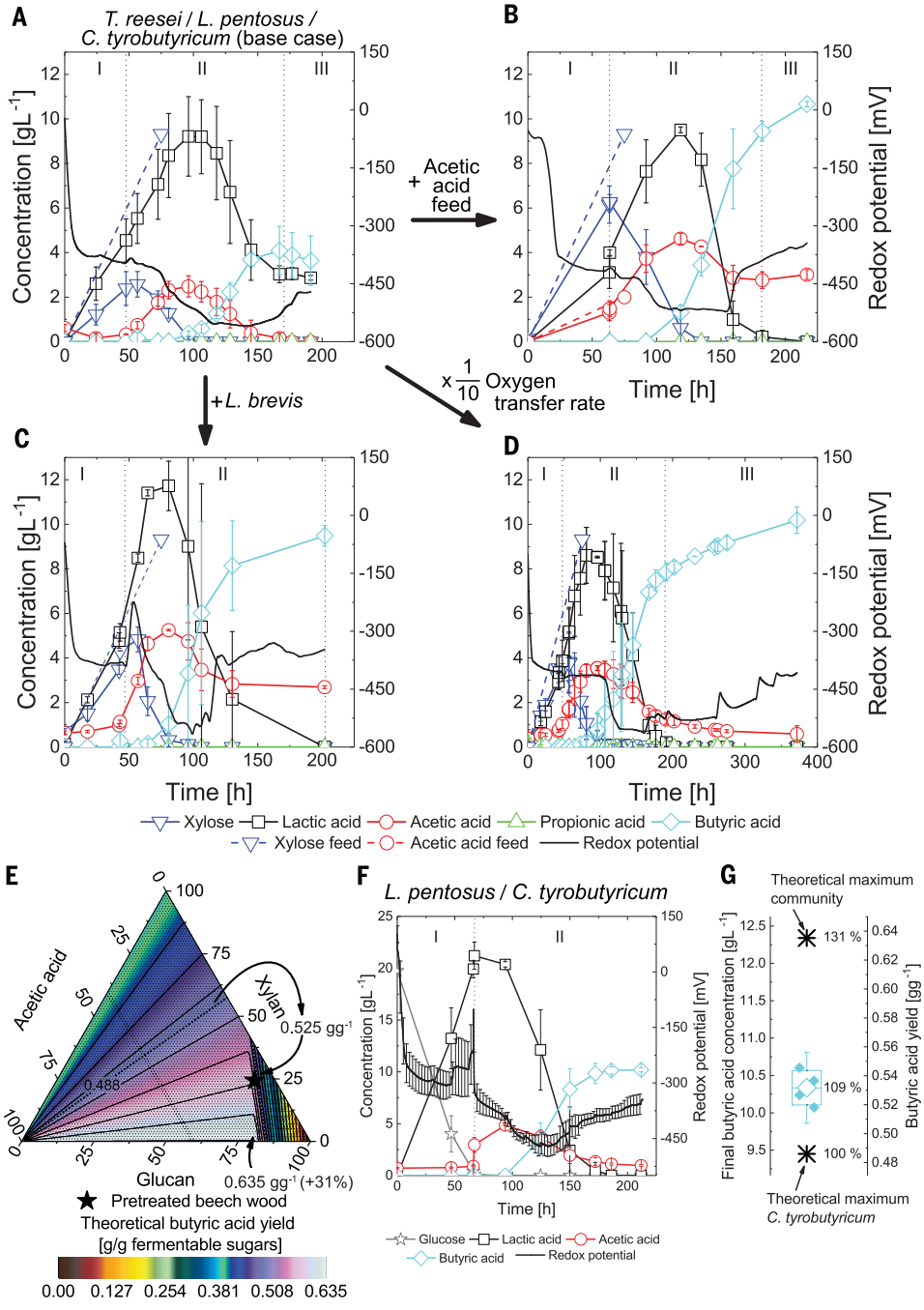


Fig. 2. CBP of cellulose and xylose to butyric acid. (**A**) In the base case, 17.5 g liter^{-1} Avicel and 9.32 g liter^{-1} xylose (fed over a time period of 75 hours; dashed blue line) were converted to butyric acid by a consortium of *T. reesei*, *L. pentosus*, and *C. tyrobutyricum* at an oxygen transfer rate (OTR) of 3.24 hour^{-1}. (**B** to **D**) Differences from the base case: (B) Acetic acid feed to 2 g liter^{-1} over a time period of 75 hours (dashed red line), (C) addition of *L. brevis* to the consortium, (D) reduction of OTR to 0.34 hour^{-1}. Stages are indicated at the top; stage III depicts the condition in which *C. tyrobutyricum* lacks acetic or lactic acid. (**E**) Ternary plot of theoretical butyric acid yield for feedstocks containing different fractions of glucan, xylan, and acetic acid (see supplementary materials). (**F**) Two-stage fermentation of the yield-optimized glucose/acetate mixture. In stage I, glucose was fermented by *L. pentosus* to lactic acid, followed by acetic acid addition and inoculation with *C. tyrobutyricum* in phase II. (**G**) Observed and theoretical butyric acid concentration and yield for the lactate platform. Data in (A) to (D) are means ± SD of two independent experiments; data in (F) and (G) are means ± SD of four independent experiments. The box in (G) denotes SD of the mean; error bars denote the 99% confidence interval.

pentosus (for lactic acid production from glucose and xylose), and *Clostridium tyrobutyricum* (for producing the target product). *C. tyrobutyricum* is able to convert glucose, xylose, or a mixture of lactic and acetic acid to butyric acid (Fig. 1A and table S1). We first used a mixture of cellulose (17.5 g liter^{-1}) and xylose (9.32 g liter^{-1}), which represented a model substrate for lignocellulose (Fig. 2A). The metabolic activity of *T. reesei* reduced the dissolved oxygen content below the detection limit within 8 to 12 hours after fungal inoculation and lowered the redox potential to values between +20 and −20 mV (fig. S3), which is not suitable for obligate anaerobes. Consequently, the cocultivation of *T. reesei* and the obligate anaerobe *C. tyrobutyricum* was not possible (fig. S4). Upon inoculation of *L. pentosus*, the redox potential dropped to less than −300 mV by reduction of oxygen to either hydrogen peroxide or water for the regeneration of nicotinamide adenine dinucleotide (NAD^+) by NADH oxidases (table S1) (*28*), thereby creating suitable growth conditions for *C. tyrobutyricum*. Butyric acid production began after 100 hours and was accompanied by a decrease of lactic and acetic acid. A maximum concentration of 4.1 g liter^{-1} (0.14 g g^{-1} fermentable sugars, 93.1% selectivity, 0.02 g total SCFA liter^{-1} hour^{-1}) was reached at the end of phase II (table S2). At the beginning of phase III, lactic acid utilization by *C. tyrobutyricum* stopped, presumably because lactic acid cannot serve as a sole carbon source but must be co-consumed with either glucose or acetic acid (fig. S5, A to E) (*29*). To verify this hypothesis, we fed acetic acid (2 g liter^{-1}) as an additional carbon source. Indeed, this enabled the complete utilization of lactic acid to produce a butyric

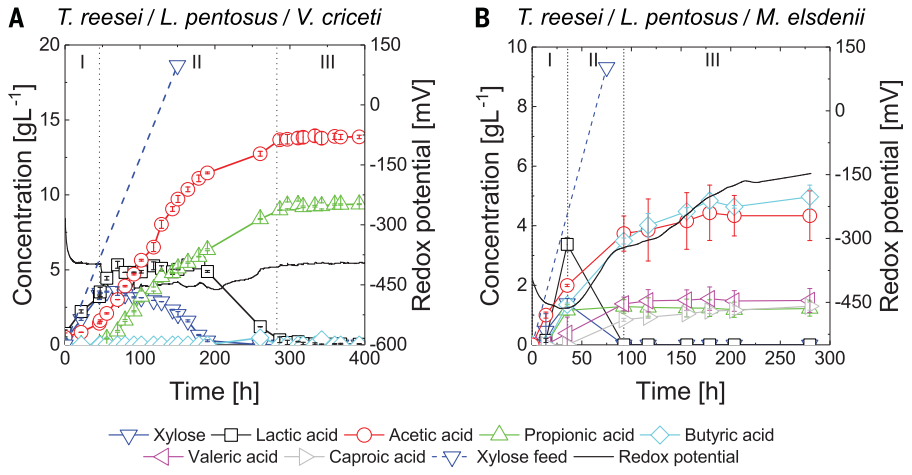


Fig. 3. CBP of cellulose and xylose to propionic and acetic acid or mixed SCFAs. (**A**) CBP of 35 g liter^{-1} Avicel and 18.64 g liter^{-1} xylose (fed over a time period of 150 hours) to propionic and acetic acid with *V. criceti* as the lactate-fermenting strain. (**B**) CBP of 17.5 g liter^{-1} Avicel and 9.32 g liter^{-1} xylose (fed over a time period of 75 hours) to mixed SCFAs with *M. elsdenii* as the lactate-fermenting strain. Stage III depicts the condition in which *V. criceti* or *M. elsdenii* lack lactic acid. Data are means ± SD of two independent experiments.

acid concentration of 10.7 g liter^{-1} (0.37 g g^{-1}, 77.6% selectivity, 0.05 g total SCFAs liter^{-1} hour^{-1}; Fig. 2B and table S2).

As an alternative to the addition of acetic acid, we aimed to enhance the metabolic flux from carbohydrates to acetic acid by extending the microbial consortium with the obligate heterofermentative strain *L. brevis*, which produces lactic and acetic acid from both hexoses and pentoses via the phosphoketolase (PK) pathway (table S1). In comparison with the base-case consortium (Fig. 2A), the extended microbial consortium produced a higher amount of acetic acid in phases I and II (Fig. 2C). This enabled the complete reutilization of lactic acid, leading to a butyric acid concentration of 9.5 g liter^{-1} (0.33 g g^{-1}, 77.9% selectivity, 0.06 g total SCFAs liter^{-1} hour^{-1}; table S2). The lower butyric acid selectivity relative to the base case is due to the slight overproduction of acetic acid.

Finally, we tested elevating the acetic acid concentration in the system by minimizing the previously reported degradation by *T. reesei* (*26*), which could be limited by the amount of oxygen entering the reactor. Indeed, decreasing the oxygen transfer rate by a factor of ~10 by using a siloxane membrane with a thicker wall led to a higher concentration of accumulated acetic acid in phase II (Fig. 2D). Then, lactic acid could be completely converted to butyric acid at a concentration of 10.2 g liter^{-1} after 372 hours (0.35 g g^{-1}, 96.7% selectivity, 0.03 g total SCFAs liter^{-1} hour^{-1}; table S2).

To obtain a benchmark for conventional butyric acid production, we also performed simultaneous saccharification and fermentation experiments with the same substrates, using only *C. tyrobutyricum* and a cellulolytic enzyme cocktail. A yield of 0.34 g g^{-1} butyric acid with a selectivity of 82% (fig. S7) was achieved, which is slightly lower than the corresponding results of the lactate platform (table S2).

Influence of substrate composition on theoretical butyric acid yields

To evaluate and compare the different options for producing butyric acid from lignocellulosic biomass and model substrates, we calculated the theoretical yields as a function of the substrate composition and the chosen fermentation route. To determine the theoretical yields of the lactate platform, we assumed that all hexoses are fermented by *L. pentosus* through the Embden-Meyerhof-Parnas pathway to lactic acid, although *C. tyrobutyricum* is also able to consume glucose. This assumption is justified, because *L. pentosus*—in contrast to *C. tyrobutyricum*—can use cellobiose intracellularly, which favors the metabolic flux of hexoses to lactic acid (Fig. 1A) (*30*). Furthermore, the cessation of lactic acid consumption by *C. tyrobutyricum* upon acetic acid depletion observed in the experiments described above (Fig. 2A) is a strong indication that no glucose is available for this strain, as lactic acid can also be metabolized in the presence of glucose (fig. S5, F to J). We also assume that xylose is metabolized through the PK pathway to lactic and acetic acid solely by *L. pentosus* (table S1). These products, as well as the acetic acid present in the feedstock, are then converted by *C. tyrobutyricum* to butyric acid. We experimentally determined the required ratio of lactic to acetic acid and found that 3 mol of lactic acid and 1 mol of acetic acid were converted to 2 mol of butyric acid (fig. S5, A to E), which is in line with literature data (*29*). In fig. S6, we propose the biochemical basis of the lactate/acetate-cofermenting pathway to butyric acid and compare it to the glucose fermentation to butyric acid in *C. tyrobutyricum*.

The yield based on the mass of fermentable carbohydrates for direct sugar conversion varies only slightly between 0.486 g g^{-1} for glucose and 0.493 g g^{-1} for xylose; however, it ranges from 0 g g^{-1} (for pure glucan) up to 0.635 g g^{-1} (81% glucan, 19% acetic acid) for the lactate platform (Fig. 2E). Because acetic acid is typically an unused resource in real biomass, only the carbohydrates were counted as carbon source. If acetic acid is counted as part of the substrate, the maximal theoretical butyric acid yield of the lactate platform is 0.52 g g^{-1} carbon source, which is 7% higher than the theoretical yield of the direct sugar conversion. For a feedstock composition corresponding to that of pretreated beechwood (68.5% glucan, 24.3% xylan, 7.2% acetic acid), the theoretical yield of butyric acid produced via the lactate platform is 0.525 g g^{-1} fermentable carbohydrates compared to 0.488 g g^{-1} for the direct fermentation of the carbohydrates. To underpin this theoretical framework with experimental data, we determined the butyric acid yield of the lactate platform using the optimal mixture of glucose and acetic acid. To this end, glucose was first metabolized to lactic acid using *L. pentosus*, followed by the addition of *C. tyrobutyricum* and acetic acid in the second phase of the fermentation. A butyric acid yield of 0.53 g g^{-1} glucose or 0.43 g g^{-1} carbon source, respectively, could be achieved, which corresponds to 109.4 ± 2.7% of the theoretical yield for direct glucose conversion by *C. tyrobutyricum* (Fig. 2, F and G). This result also outperforms published yields for fermentations of the same substrates using natural or genetically engineered strains (table S4).

Propionic and acetic acid production

To demonstrate the modularity of the lactate platform, we replaced *C. tyrobutyricum* with *Veillonella criceti* to produce propionic acid. *V. criceti* is an obligate anaerobic bacterium that is unable to utilize carbohydrates but converts lactic acid to propionic and acetic acid (Fig. 1A and table S3). Thus, the lactate platform approach compensates for the missing metabolic flexibility of the strain and enables the production of propionic acid from carbohydrates. CBP of 17.5 g liter^{-1} microcrystalline cellulose and 9.32 g liter^{-1} xylose yielded 8.9 g liter^{-1} total carboxylic acids (fig. S8 and table S5). When doubling both carbon sources, 13.7 g liter^{-1} acetic acid and 9.3 g liter^{-1} propionic acid were produced (Fig. 3A and table S5).

Mixed C_2 to C_6 SCFA production

Higher SCFAs such as valeric (C_5) and caproic (C_6) acid are more valuable than the lower SCFAs (C_2 to C_4). However, few microbial species are known to produce these compounds in

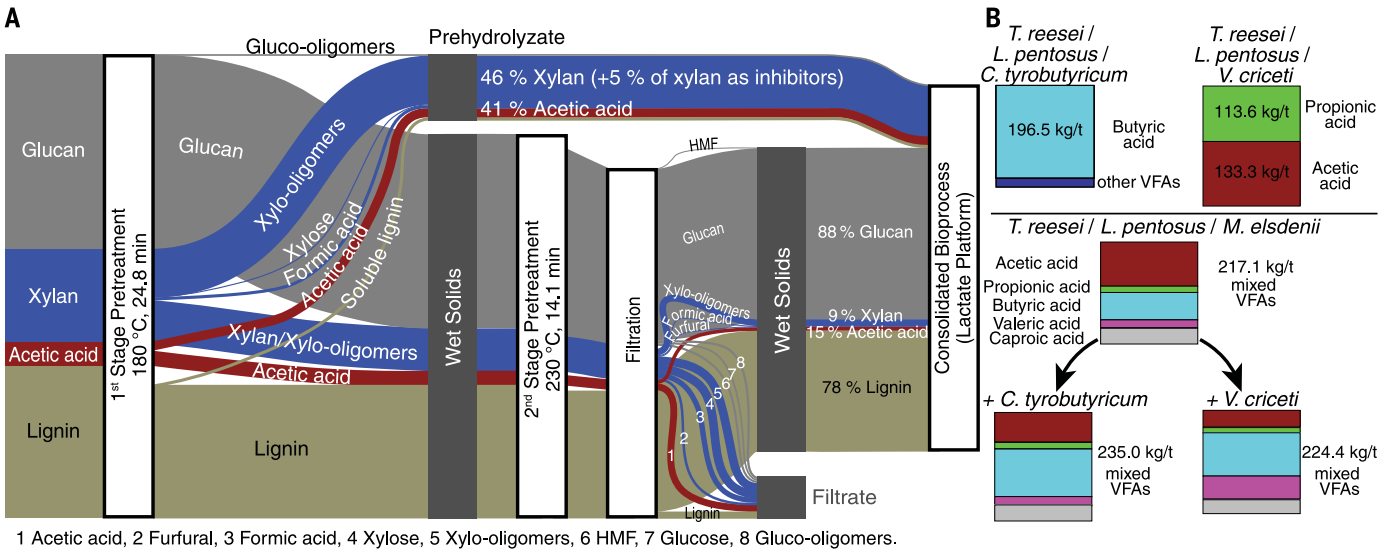


Fig. 4. Steam pretreatment of beechwood and CBP to SCFAs. (**A**) Mass fluxes of glucan, xylan, acetic acid, and lignin in the two-stage steam pretreatment (further details in fig. S9). (**B**) Experimental yields of CBP with different microbial consortia to various SCFAs (fig. S10).

substantial amounts (*31*). One of the exceptions is *Megasphaera elsdenii*, a strict anaerobic rumen bacterium. *M. elsdenii* favors lactic acid over glucose as carbon source and is not able to metabolize xylose or cellulose (*32*, *33*), thereby making it a suitable candidate for the lactate platform (Fig. 1A). We tested the cocultivation of *T. reesei*, *L. pentosus*, and *M. elsdenii* in the membrane-aerated reactor following the subsequent inoculation strategy described above. With 17.5 g liter^{-1} cellulose and 9.32 g liter^{-1} xylose as substrates, we recorded the simultaneous accumulation of 4.3 g liter^{-1} acetic acid, 1.2 g liter^{-1} propionic acid, 5.0 g liter^{-1} butyric acid, 1.5 g liter^{-1} valeric acid, and 1.3 g liter^{-1} caproic acid in 280 hours. The obtained SCFA yields were 0.46 ± 0.01 g SCFA per g fermentable sugars (Fig. 3B and table S6), showing the successful cocultivation of the three strains. However, after inoculation of *M. elsdenii* the redox potential increased (~1.5 mV hour^{-1} from t = 50 hours to t = 200 hours), indicating that the metabolic activity of *T. reesei* and *L. pentosus* was not sufficient to maintain a sufficiently low redox potential in longer fermentation runs, likely due to product toxicity.

CBP of pretreated beechwood

After successfully converting model substrates, we targeted the direct conversion of a lignocellulosic substrate—pretreated beechwood—to the different target products. We applied a two-stage steam pretreatment procedure to recover temperature-sensitive solubilized xylo-oligomers and acetic acid in a moderately inhibitory liquid stream termed prehydrolyzate after a mild first pretreatment and a well-digestible solid substrate after the second, more severe pretreatment step (Fig. 4 and fig. S9). For CBP of beechwood, we added 3% (w/w) pretreated solids at the beginning of the fermentation and started the first-stage prehydrolyzate feed with the inoculation of *L. pentosus*. With *C. tyrobutyricum* as lactic acid–fermenting strain, the consortium produced 9.5 g liter^{-1} butyric acid (0.38 g g^{-1} fermentable carbohydrates, 91.5% selectivity, 0.014 g total SCFA liter^{-1} hour^{-1}; fig. S10A and table S2), which corresponds to an overall yield of 196.5 kg of butyric acid per metric ton of dry raw beechwood (Fig. 4B). This yield was approximately twice as high as the reported yields of a natural consortium (*34*), an anaerobic coculture (*35*), or a natural single cellulolytic strain (*36*) and was even comparable to yields achieved on lignocellulosic hydrolysates (*37*, *38*), although we did not add external enzymes (table S4). When *V. criceti* was used as fermenting strain in the lactate platform, 6.7 g liter^{-1} acetic acid and 5.6 g liter^{-1} propionic acid (0.49 g g^{-1} fermentable carbohydrates, 0.031 g total SCFA liter^{-1} hour^{-1}; fig. S10B and table S5) were produced. This corresponded to 133.3 kg of acetic acid and 113.6 kg of propionic acid per metric ton of raw biomass (Fig. 4B). Finally, when *M. elsdenii* was integrated into the lactate platform, a mixture of 6.8 g liter^{-1} acetic acid, 4.2 g liter^{-1} butyric acid, 2.4 g liter^{-1} caproic acid, and small amounts of propionic and valeric acid (each ~1 g liter^{-1}) (fig. S10C and table S6) could be produced from pretreated beechwood. Overall, the process yielded 0.42 g of total SCFAs per gram of total fermentable sugars, corresponding to 217.1 kg of total SCFAs per metric ton of raw beechwood with a productivity of 0.012 g liter^{-1} hour^{-1}. Even more than 600 hours after inoculation of *M. elsdenii*, the accumulation of metabolic end products of the anaerobic pathways such as caproic acid was recorded, indicating that the reduced oxygen transfer rate achieved by using a thicker membrane prolonged the metabolic activity of the strict anaerobic strain (fig. S10C and table S6). This is in line with the measured smaller slope of the redox potential (~0.23 mV hour^{-1} from t = 400 hours to t = 1094 hours) in comparison to the approach with the thinner membrane (Fig. 3B).

Finally, we aimed to influence the ratio of odd- or even-numbered SCFAs produced by *M. elsdenii* (Fig. 1A). *M. elsdenii* can take up extracellular propionic acid, which is activated intracellularly to propionyl–coenzyme A (CoA) by a propionyl-CoA transferase (PCT) (Fig. 1A). Because PCT has a broad substrate spectrum, the functional group CoA can be swapped between acetyl-CoA and SCFA (*39*, *40*), which is essential for the chain elongation process. On this basis, we hypothesized that an increased availability of propionic acid increases the production of higher, odd-numbered SCFAs. Because *V. criceti* was shown to convert lactic acid to propionic and acetic acid, thereby increasing the concentration of propionic acid, we set up a lactate platform containing both *V. criceti* and *M. elsdenii* as lactate-fermenting strains. Using pretreated beechwood as substrate, this artificial consortium produced 2.6 g liter^{-1} acetic acid, 0.8 g liter^{-1} propionic acid, 6.7 g liter^{-1} butyric acid, 3.5 g liter^{-1} valeric acid, and 2.1 g liter^{-1} caproic acid (Fig. 4B and fig. S10E). Thus, the valeric acid concentration was higher by a factor of nearly 3 relative to the mixture produced by the community without *V. criceti*. The process yield was 0.42 g of total SCFAs per gram of total fermentable

sugars and 0.22 g of total SCFAs per gram of raw beechwood, similar to the community without *V. criceti*.

The metabolism of *M. elsdenii* also allows the reverse β-oxidation of butyric acid (electron acceptor) taken from the environment to caproic acid if an electron donor such as lactate is available (Fig. 1A). Thus, by extending the community with *C. tyrobutyricum* instead of *V. criceti*, we sought to increase the share of even-numbered longer SCFAs. By doing so, we measured a SCFA mixture consisting of 4.6 g liter^{-1} acetic acid, 1.0 g liter^{-1} propionic acid, 7.4 g liter^{-1} butyric acid, 1.4 g liter^{-1} valeric acid, and 2.5 g liter^{-1} caproic acid (Fig. 4B and fig. S10D). Relative to the base case, the butyric acid concentration increased by a factor of 1.75 and the acetic acid concentration decreased by a factor of 0.66. However, the presence of *C. tyrobutyricum* did not increase the caproic acid concentration, which indicates the lack of electron donors that are required for chain elongation by *M. elsdenii* (*41*). Overall, the yield achieved with the community extended with *C. tyrobutyricum* increased to 0.46 g of total SCFAs per gram of total fermentable sugars and 0.24 g of SCFAs per gram of raw beechwood.

CBP of lignocellulosic substrates to mixed SCFAs can also be performed by natural mixed cultures such as the rumen microbiome or different kinds of anaerobic digestion consortia, where methane formation is inhibited. These processes are summarized as the "carboxylate platform" (*1*) and are typically performed with undefined waste substrates, making a comparison of the yields with those of the lactate platform difficult. However, the yield that we achieved with the lactate platform is comparable to reported *M. elsdenii* fermentations on glucose or corn stover hydrolysate, and we even exceeded the reported titers (table S4). Furthermore, in the lactate platform, xylose was also converted, which is not possible in axenic *M. elsdenii* cultures.

In the batches run on beechwood, we also quantified hydrolytic enzyme expression over time to determine whether *T. reesei* was metabolically active and was producing enzymes during all phases of the fermentations. We consistently found increasing activities of cellobiohydrolases, endoglucanases, β-glucosidases, and xylanases in the supernatants after inoculation with the lactic acid and product-forming strains (fig. S11). Together with further experiments described in figs. S18 to S20, we conclude that *T. reesei* was metabolically active throughout the experiments.

Discussion and outlook

In this study, we merged synthetic microbial ecology with process engineering to realize and control artificial food chains distributed over as many as four highly diverse microorganisms. Our goal was to provide a platform that minimizes the required metabolic capabilities of product-forming microorganisms to facilitate their integration into an artificial community utilizing all biomass fractions. To demonstrate the platform's wide applicability, we identified several other bulk or fine chemicals, including acetic acid, alcohols (ethanol, propanol, butanol, 1,2-propanediol), polyhydroxybutyrate, and lipids, that could be produced through the lactate platform (fig. S12). We analyzed the stoichiometries of the required pathways and found that, relative to the alternative direct sugar fermentation routes for acetic acid, polyhydroxybutyrate, ethanol, and butanol, the lactate platform theoretically boosts the carbon flux from lignocellulose to the target product (fig. S12 and tables S7 to S9). A switch to another target product may require the exchange of the lactate-fermenting strain(s), and as a consequence, different abiotic requirements such as pH, temperature, or presence of co-substrates will likely have to be fulfilled. Furthermore, for some products the artificial food chain must be extended; that is, consortia containing at least four different microorganisms would be necessary. Thus, the creation of additional spatial and/or metabolic niches is desirable to fully harness the potential of the lactate platform. To this end, we propose several possibilities to adapt the biofilm membrane reactor (*24*). For instance, temperature gradients could be created by installing local heat sinks or sources in the reactor, or pH gradients could be formed by introducing carbon dioxide or ammonia through a membrane to the reactor (fig. S13). In the same way, co-substrates such as hydrogen or carbon monoxide could be fed to the consortium. Furthermore, light management in the reactor could facilitate the mixotrophic growth of microalgae on lignocellulosic feedstocks, which paves the way to a range of valuable products such as lipids for biodiesel production or omega-3 fatty acids as aquatic feed. As a first step in the direction of the latter, we showed that *T. reesei* and the microalgae *Chlamydomonas reinhardtii* were able to grow together in an illuminated aerobic biofilm using pure cellulose as a substrate (figs. S14 to S17). Prior to a potential industrial application, further optimization of the lactate platform is necessary; in particular, the product concentrations and the corresponding productivities need to be enhanced. At this proof-of-concept stage, the obtained productivities in the range of 0.01 to 0.06 g liter^{-1} hour^{-1} (tables S2, S5, and S6) compare well with literature data (table S4). One reason for the moderate productivities is the gap between the optimal cellulase production temperature (28° to 30°C) and the optimal temperature for enzyme activity (40° to 50°C). Besides using an advanced industrial cellulase-producing strain, the use of a temperature gradient reactor could also increase the productivity. The bulk phase could be operated at elevated temperatures while the enzyme-producing biofilm could be cooled via the membrane.

Taken together, the results achieved with the lactate platform demonstrate the inherent benefits of mixed artificial cultures, such as their modularity and their ability to convert complex substrates to valuable platform chemicals. This reactor engineering complements the available toolbox for successfully engineering stable and controllable synthetic communities for numerous novel applications.

Materials and methods

Fungal, bacterial, and algal strains and culturing methods

L. pentosus (DSM-20314), *L. brevis* (DSM-20054), *C. tyrobutyricum* (DSM-2637), *V. criceti* (DSM-20734), and *M. elsdenii* (DSM-20460) were purchased from Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures (Germany).

One ml of glycerol stock was inoculated to MRS medium for precultures of *Lactobacillus* in 500-ml Erlenmeyer flasks and incubated for 48 hours at 150 rpm at 33°C. The MRS medium contained (g liter^{-1}): peptone from casein, 10; meat extract, 10; yeast extract, 5; glucose, 20; Tween 80, 1; K_2HPO_4, 2; sodium acetate, 5; ammonium citrate, 2; and $MgSO_4 \cdot 7H_2O$, 0.2. The pH was adjusted to 6.2–6.5 with hydrochloric acid.

C. tyrobutyricum and *M. elsdenii* precultures were cultivated in a reinforced clostridial medium (RCM). Composition of the RCM medium (in g liter^{-1}): yeast extract, 13; peptone, 10; glucose, 5; soluble starch, 1; sodium chloride, 5; sodium acetate, 3; L-cysteine-HCl, 0.5; agar, 0.5.

V. criceti precultures were cultivated in medium containing (in g liter^{-1}): trypticase (BBL), 5; yeast extract, 3.0; Na-(DL)-lactate, 7.5; Na-thioglycolate, 0.75; Tween 80, 1.0; glucose, 1.0; Na-resazurin solution [0.1% (w/v)], 0.5 ml. The medium for *C. tyrobutyricum* and *V. criceti* was boiled under continuous flushing with nitrogen and the pH was adjusted to 6.8 before it was aliquoted to 100-ml serum bottles (50 ml in each) and autoclaved at 121°C for 20 min.

V. criceti is known for its high nutrient requirements (*43*, *44*), but we could show that it is still active on Mandels medium containing only 1.0 g liter^{-1} complex media components. We found that the fermentation time was prolonged by a factor of ~3 without further supplementation of complex nutrients, while the yields remained the same (table S3). Therefore, we used Mandel medium in the CBP experiments with *V. criceti*.

All serum bottles were inoculated from a 2-day liquid culture stored at 4°C and incubated on an orbital shaker at 140 rpm at 37°C.

T. reesei Rut-C30 (ATCC 56765) was purchased from the VTT, Finland. Spores from a

7-day-old potato dextrose agar slant culture were incubated at 28°C and stored at 4°C. For precultures, Mandels medium containing 7.5 g liter^{-1} microcrystalline cellulose (Avicel PH-101, Sigma Aldrich) was inoculated with spores from the agar slant in an Erlenmeyer flask for 4 days at 28°C and 150 rpm. Mandels medium contained the following ingredients (g liter^{-1}): KH_2PO_4, 2; $(NH_4)_2SO_4$, 1.4; $MgSO_4{\cdot}7H_2O$, 0.3; $CaCl_2{\cdot}6H_2O$, 0.4; urea, 0.3; peptone, 0.75; yeast extract, 0.25; and trace element stock, 1 ml liter^{-1}. The trace element stock contained (g liter^{-1}): $FeSO_4{\cdot}7H_2O$, 5; $MnSO_4{\cdot}H_2O$, 1.6; $ZnSO_4{\cdot}7H_2O$, 1.4; $CoCl_2{\cdot}6H_2O$, 3.7; and 10 ml liter^{-1} concentrated hydrochloric acid, and was sterile filtered. To avoid precipitation, 100× $CaCl_2$ and $MgSO_4$ solutions were autoclaved separately before merging with the remaining ingredients.

Stocks of *C. reinhardtii* wild-type strain (WT12) were stored at room temperature in TAP medium. TAP medium contained per liter: TRIS, 2.42 g; TAP salt stock, 25 ml; phosphate solution, 0.375 ml; six Kropat's trace element stock solutions, 1 ml each; and glacial acetic acid, 1 ml. The medium was autoclaved for 20 min at 121°C. TAP salt stock solution contained (in g liter^{-1}): NH_4Cl, 15.0; $MgSO_4{\cdot}7H_2O$, 4.0; $CaCl_2{\cdot}2H_2O$, 2.0. The phosphate solution contained 28.8 g K_2HPO_4 and 14.4 g KH_2PO_4 and water was added to a volume of 100 ml. Kropat's trace element stock solutions contained: (1) 25 mM EDTA-Na_2; (2) 28.5 mM $(NH_4)_6MO_7O_{24}$; (3) 2 mM $CuCl_2{\cdot}2H_2O$ and 2 mM EDTA; (4) 2.5 mM $ZnSO_4{\cdot}7H_2O$ and 2.75 mM EDTA; (5) 6 mM $MnCl_2{\cdot}4H_2O$ and 6 mM EDTA; (6) 20 mM $FeCl_3{\cdot}6H_2O$, 22 mM EDTA, and 22 mM Na_2CO_3. For liquid precultures of *C. reinhardtii*, Erlenmeyer flasks were filled with fresh TAP medium and inoculated with 5% (v/v) stock solution and incubated at 25°C with an 18 hours light to 6 hours dark cycle with 120 mol m^{-2} s^{-1} for 5 days under shaking at 120 rpm.

Membrane bioreactor

The membrane bioreactors were based on Multifors 2 and Labfors 5 bioreactors (Infors HT, Switzerland) with 0.5 and 2.7 liter working volume, respectively. Both reactors were modified with a tubular polydimethylsiloxane (PDMS) membrane. The membrane area to volume ratio was 0.3 cm^2 ml^{-1} at 0.5 liter scale. Mono-Lumen Tubing 0.64 × 1.19 × 0.28 mm (Dow Corning) gave a k_La value with standard error of 3.24 ± 0.05 hour^{-1} in the Multifors reactors. The membrane area to volume ratio was 0.3 cm^2 ml^{-1} in the Labfors. Mono-Lumen Tubing 1.58 × 3.18 × 0.80 mm (Dow Corning) membrane gave a k_La value with standard error of 0.34 ± 0.01 hour^{-1}. The membrane was continuously flushed with air at 140 ml min^{-1} per liter liquid phase. The temperature was set to 30°C and the pH to 6.0 using 4 N phosphoric acid and 4 M sodium hydroxide. The redox potential and the pH were measured using EasyFerm Plus ORP Arc and EasyFerm Plus PHI Arc probes (Hamilton, Switzerland). Liquid samples were collected using a UV lamp as the contamination barrier. No external cellulolytic enzyme mixture or nitrogen purging were used. The reactor was autoclaved for 20 min at 121°C with microcrystalline cellulose or pretreated beechwood. The remaining medium components were added when the reactor was inoculated with 5% (v/v) *T. reesei*. After 48 hours, the lactic acid bacteria were inoculated to an optical density (OD_{600}) of 0.5 by centrifuging the cells at 3000 rpm for 10 min and suspending them in fresh Mandels medium. This time was set as time zero. Another 48 hours after inoculation of the LAB, the *C. tyrobutyricum* or *V. criceti* was inoculated to 5% (v/v) from a 2-day-old pre-culture. For fed-batch experiments, the feed solution was sterile filtered and fed at a constant feeding rate.

Experiments with the green alga *C. reinhardtii* (WT12) were performed in a custom made 1-liter glass stirred tank reactor equipped with the same membrane as in the Labfors fermenters described above and a membrane area to volume ratio of 0.3 cm^2 ml^{-1}. In order to enable mixotrophic growth of the microalgae, a radially emitting light fiber (side glow fiber, 1.5 mm diameter, Starscape, UK) connected to an external light source (Cree XLAMP XHP 70.2 – 3000 K 80 CRI) was mounted next to the membrane (figs. S14 and S15). The reactor was autoclaved for 20 min at 121°C with water and microcrystalline cellulose. The remaining medium components were added when the reactor was inoculated at the same time with 5% (v/v) *T. reesei* and 5% (v/v) *C. reinhardtii*. For coculture experiments, Mandel's medium was used and the temperature was set to 26°C.

Steam pretreatment and beechwood analysis

Beechwood chips (*Fagua sylvatica*) from a local forest in Bern, Switzerland, were air-dried (dry matter 94%) and milled to <1.5 mm. The raw biomass was composed of [% (w/w)] glucan 40.0 ± 1.3, xylan 19.0 ± 0.5, acetic acid 7.3 ± 0.2, acid-insoluble lignin 25.6 ± 0.7, acid-soluble lignin 5.3 ± 0.1, ash 0.5, and extractives 0.9. A two-stage steam pretreatment was applied with a custom-built steam gun (Industrieanlagen Planungsgesellschaft m.b. H., Austria) to target the partial recovery of soluble xylo-oligosaccharides from hemicellulose in the prehydrolyzate and simultaneously maximize the glucose yield in the enzymatic hydrolysis of the solids. Beechwood was heated to 180°C by the injection of saturated steam and pretreated for a residence time of 24.8 min (severity of 3.75). At constant pressure and temperature, the liquid phase was removed through a circular nozzle located slightly above the lower ball valve. Subsequently, the pressure was slowly released to 2.0 bar, which is below the known pressure needed to see an effect of the explosion (*45*). At 2.0 bar, the lower ball valve opened and the solids were removed from the reactor. The solids were composed of [in % (w/w)] 49.0 ± 0.6 glucan, 7.6 ± 0.1 xylan, 4.2 ± 0.1 acetic acid, 29.9 ± 0.7 acid-insoluble lignin, and 3.7 ± 0.4 acid-soluble lignin. The recovered solids were treated a second time at 230°C for 14.1 min and the pressure was abruptly released. The solids contained [in % (w/w)] 55.0 ± 1.4 glucan, 2.1 ± 0.1 xylan, 1.2 ± 0.2 acetic acid, 35.4 ± 0.8 acid-insoluble lignin, and 2.5 ± 0.1 acid-soluble lignin. The 180°C liquid fraction contained (in g liter^{-1}): acetic acid, 2.1; formic acid, 0.4; xylose, 1.1; xylo-oligosaccharides calculated as xylan, 22.4. During acid hydrolysis, an additional 5.57 g liter^{-1} acetic acid was released from deacetylation of xylo-oligomers. The severity was calculated according to Overend *et al.* (*46*). Mass fluxes of glucan, xylan, acetic acid, and lignin are shown in Fig. 4A; see fig. S9 for further details of recovery percentages.

Enzyme assays

EG, BG, CBH, and xylanase activity assays were performed at pH 5.0 in the presence of 50 mM citrate buffer, 0.02 (w/v) % sodium azide at 50°C in duplicates. The enzymatic activity was expressed in international units (IU). One unit is defined as the amount of enzyme that catalyzes the conversion of one micromole of substrate per minute. EG, BG, and CBH assays were performed as described (*47*). Ultralow-viscosity carboxymethylcellulose [2 (w/w) %] was used as substrate for EG assays, Avicel [2 (w/w) %] for CBH, and 4-nitrophenyl α-D-glucopyranoside (1 mM) for BG. Birch wood xylan [1 (w/w) %] was used as substrate for xylanase activity measurements (12.5 µl sample in 125 µl total reaction volume). Reducing sugars were quantified using the 3,5-dinitrosalicylic acid (DNS) method (*48*).

Analytical methods

Acetic acid, lactic acid, propionic acid, butyric acid, formic acid, glucose, cellobiose, xylose, hydroxymethylfurfural, and furfural were quantified by high-performance liquid chromatography (Waters 2695 Separation Module, Waters Corp., Milford, MA) using an Aminex HPX-87H column (Bio-Rad) at 65°C with 5 mM H_2SO_4 as the mobile phase, flow at 0.6 ml min^{-1}, a refractive index detector (Waters 410) at 40°C, and a photodiode array detector (Waters 2998) at 210 nm. The detection limit was 0.05 g liter^{-1}. Structural carbohydrates and lignin were measured according to standard methods (*49*).

Chlorophyll a, chlorophyll b, total chlorophyll, and total carotenoids in biofilms of *T. reesei* and *C. reinhardtii* were quantified by first extracting pigments with dimethylformamide (DMF) and measuring the absorbance at 480,

647, and 664.5 nm (Thermo Spectronic UV-1, Thermo Fisher). Prior to the measurement, samples with a specific volume or mass were centrifuged in a 1.5-ml Eppendorf vial, 1 ml DMF was added to the pellet, and the vials were agitated using a vortex at room temperature for 15 min. The vials were centrifuged at 10,000 rpm on a benchtop centrifuge. The chlorophyll a, chlorophyll b, and total chlorophyll contents were calculated using the equations from (*50*) and the amount of carotenoids using the equation from (*51*).

Autofluorescence from photosynthetic pigments of the stained biofilm samples was also measured using a confocal laser scanning microscope (Leica SP8, Leica Microsystems). The setup was equipped with a white light laser set to 30% laser power to excite the sample. The scan speed was 400 Hz and the objective HC PL APO CS2 40×/1.30 oil with a magnification of 40 was installed. Autofluorescence from photosynthetic pigments was detected in the red channel (excitation at 650 nm and emission at 680 to 700 nm). Fluorescein diacetate was used as viability stain. Fluorescein was detected in the green channel (excitation 495 nm and emission at 514 to 521 nm). The fluorescence projections of the biofilm were generated using Leica LAS software.

REFERENCES AND NOTES

1. M. T. Agler, B. A. Wrenn, S. H. Zinder, L. T. Angenent, Waste to bioproduct conversion with undefined mixed cultures: The carboxylate platform. *Trends Biotechnol.* **29**, 70–78 (2011). doi: 10.1016/j.tibtech.2010.11.006; pmid: 21190748
2. L. R. Lynd, P. J. Weimer, W. H. van Zyl, I. S. Pretorius, Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506–577 (2002). doi: 10.1128/MMBR.66.3.506-577.2002; pmid: 12209002
3. R. Tsoi *et al*., Metabolic division of labor in microbial systems. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2526–2531 (2018). doi: 10.1073/pnas.1716888115; pmid: 29463749
4. D. Dugar, G. Stephanopoulos, Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nat. Biotechnol.* **29**, 1074–1078 (2011). doi: 10.1038/nbt.2055; pmid: 22158355
5. J. C. Liao, L. Mi, S. Pontrelli, S. Luo, Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* **14**, 288–304 (2016). doi: 10.1038/nrmicro.2016.32; pmid: 27026253
6. B. W. Biggs, B. De Paepe, C. N. S. Santos, M. De Mey, P. Kumaran Ajikumar, Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* **29**, 156–162 (2014). doi: 10.1016/j.copbio.2014.05.005; pmid: 24927371
7. A. Gupta, I. M. B. Reizman, C. R. Reisch, K. L. J. Prather, Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat. Biotechnol.* **35**, 273–279 (2017). doi: 10.1038/nbt.3796; pmid: 28191902
8. P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, J. D. Keasling, Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320–328 (2012). doi: 10.1038/nature11478; pmid: 22895337
9. N. I. Johns, T. Blazejewski, A. L. Gomes, H. H. Wang, Principles for designing synthetic microbial communities. *Curr. Opin. Microbiol.* **31**, 146–153 (2016). doi: 10.1016/j.mib.2016.03.010; pmid: 27084981
10. L. Goers, P. Freemont, K. M. Polizzi, Co-culture systems and technologies: Taking synthetic biology to the next level. *J. R. Soc. Interface* **11**, 20140065 (2014). doi: 10.1098/rsif.2014.0065; pmid: 24829281
11. J. J. Minty *et al*., Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14592–14597 (2013). doi: 10.1073/pnas.1218447110; pmid: 23959872
12. L. Xu, U. Tschirner, Improved ethanol production from various carbohydrates through anaerobic thermophilic co-culture. *Bioresour. Technol.* **102**, 10065–10071 (2011). doi: 10.1016/j.biortech.2011.08.067; pmid: 21908184
13. T. R. Zuroff, S. B. Xiques, W. R. Curtis, Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium phytofermentans*/yeast co-culture. *Biotechnol. Biofuels* **6**, 59 (2013). doi: 10.1186/1754-6834-6-59; pmid: 23628342
14. Z. Wen *et al*., Artificial symbiosis for acetone-butanol-ethanol (ABE) fermentation from alkali extracted deshelled corn cobs by co-culture of Clostridium beijerinckii and Clostridium cellulovorans. *Microb. Cell Fact.* **13**, 92 (2014). doi: 10.1186/s12934-014-0092-5; pmid: 25023325
15. S. Ben Said, R. Tecon, B. Borer, D. Or, The engineering of spatially linked microbial consortia - potential and perspectives. *Curr. Opin. Biotechnol.* **62**, 137–145 (2020). doi: 10.1016/j.copbio.2019.09.015; pmid: 31678714
16. Y. Liu *et al*., A three-species microbial consortium for power generation. *Energy Environ. Sci.* **10**, 1600–1609 (2017). doi: 10.1039/C6EE03705D
17. H. J. Kim, J. Q. Boedicker, J. W. Choi, R. F. Ismagilov, Defined spatial structure stabilizes a synthetic multispecies bacterial community. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18188–18193 (2008). doi: 10.1073/pnas.0807935105; pmid: 19011107
18. S. Kato, S. Haruta, Z. J. Cui, M. Ishii, Y. Igarashi, Stable coexistence of five bacterial strains as a cellulose-degrading community. *Appl. Environ. Microbiol.* **71**, 7099–7106 (2005). doi: 10.1128/AEM.71.11.7099-7106.2005; pmid: 16269746
19. R. B. Payne, U. Ghosh, H. D. May, C. W. Marshall, K. R. Sowers, A Pilot-Scale Field Study: In Situ Treatment of PCB-Impacted Sediments with Bioamended Activated Carbon. *Environ. Sci. Technol.* **53**, 2626–2634 (2019). doi: 10.1021/acs.est.8b05019; pmid: 30698958
20. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663–666 (2015). doi: 10.1126/science.aad2602; pmid: 26542567
21. X. N. Peng, S. P. Gilmore, M. A. O'Malley, Microbial communities for bioprocessing: Lessons learned from nature. *Curr. Opin. Chem. Eng.* **14**, 103–109 (2016). doi: 10.1016/j.coche.2016.09.003
22. C. D. Nadell, K. Drescher, K. R. Foster, Spatial structure, cooperation and competition in biofilms. *Nat. Rev. Microbiol.* **14**, 589–600 (2016). doi: 10.1038/nrmicro.2016.84; pmid: 27452230
23. Y. Jiang *et al*., Recent advances of biofuels and biochemicals production from sustainable resources using co-cultivation systems. *Biotechnol. Biofuels* **12**, 155 (2019). doi: 10.1186/s13068-019-1495-7; pmid: 31285755
24. R. L. Shahab, S. Brethauer, J. S. Luterbacher, M. H. Studer, Engineering of ecological niches to create stable artificial consortia for complex biotransformations. *Curr. Opin. Biotechnol.* **62**, 129–136 (2020). doi: 10.1016/j.copbio.2019.09.008; pmid: 31671322
25. S. Brethauer, M. H. Studer, Consolidated bioprocessing of lignocellulose by a microbial consortium. *Energy Environ. Sci.* **7**, 1446–1453 (2014). doi: 10.1039/c3ee41753k
26. R. L. Shahab, J. S. Luterbacher, S. Brethauer, M. H. Studer, Consolidated bioprocessing of lignocellulosic biomass to lactic acid by a synthetic fungal-bacterial consortium. *Biotechnol. Bioeng.* **115**, 1207–1215 (2018). doi: 10.1002/bit.26541; pmid: 29315476
27. J. B. Xavier, K. R. Foster, Cooperation and conflict in microbial biofilms. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 876–881 (2007). doi: 10.1073/pnas.0607651104; pmid: 17210916
28. C. Nowak *et al*., A water-forming NADH oxidase from Lactobacillus pentosus suitable for the regeneration of synthetic biomimetic cofactors. *Front. Microbiol.* **6**, 957 (2015). doi: 10.3389/fmicb.2015.00957; pmid: 26441891
29. Y. Tao *et al*., Production of Butyrate from Lactate by a Newly Isolated *Clostridium* sp. BPY5. *Appl. Biochem. Biotechnol.* **179**, 361–374 (2016). doi: 10.1007/s12010-016-1999-6; pmid: 26842597
30. A. Endo, L. M. Dicks, in *Lactic Acid Bacteria: Biodiversity and Taxonomy*, W. H. Holzapfel, B. J. B. Wood, Eds. (Wiley, 2014), pp. 13–30.
31. P. J. Weimer, G. N. Moen, Quantitative analysis of growth and volatile fatty acid production by the anaerobic ruminal bacterium *Megasphaera elsdenii* T81. *Appl. Microbiol. Biotechnol.* **97**, 4075–4081 (2013). doi: 10.1007/s00253-012-4645-4; pmid: 23271673
32. S. R. Elsden, B. E. Volcani, F. M. Gilchrist, D. Lewis, Properties of a fatty acid forming organism isolated from the rumen of sheep. *J. Bacteriol.* **72**, 681–689 (1956). doi: 10.1128/JB.72.5.681-689.1956; pmid: 13376515
33. A. Ohnishi, in *Microbial Factories* (Springer, 2015), pp. 47–71.
34. B. Ai *et al*., Consolidated Bioprocessing for Butyric Acid Production from Rice Straw with Undefined Mixed Culture. *Front. Microbiol.* **7**, 1648 (2016). doi: 10.3389/fmicb.2016.01648; pmid: 27822203
35. X. Chi, J. Li, X. Wang, Y. Zhang, P. Antwi, Hyper-production of butyric acid from delignified rice straw by a novel consolidated bioprocess. *Bioresour. Technol.* **254**, 115–120 (2018). doi: 10.1016/j.biortech.2018.01.042; pmid: 29413911
36. K. Merklein, S. S. Fong, Y. Deng, Production of butyric acid by a cellulolytic actinobacterium *Thermobifida fusca* on cellulose. *Biochem. Eng. J.* **90**, 239–244 (2014). doi: 10.1016/j.bej.2014.06.012
37. Y. Suo *et al*., Butyric acid production from lignocellulosic biomass hydrolysates by engineered *Clostridium tyrobutyricum* overexpressing Class I heat shock protein GroESL. *Bioresour. Technol.* **250**, 691–698 (2018). doi: 10.1016/j.biortech.2017.11.059; pmid: 29220814
38. M. Kim *et al*., Butyric acid production from softwood hydrolysate by acetate-consuming *Clostridium* sp. S1 with high butyric acid yield and selectivity. *Bioresour. Technol.* **218**, 1208–1214 (2016). doi: 10.1016/j.biortech.2016.07.073; pmid: 27474955
39. C. H. Martin *et al*., A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-γ-butyrolactone. *Nat. Commun.* **4**, 1414–1419 (2013). doi: 10.1038/ncomms2418; pmid: 23361005
40. B. S. Jeon, O. Choi, Y. Um, B.-I. Sang, Production of medium-chain carboxylic acids by *Megasphaera* sp. MH with supplemental electron acceptors. *Biotechnol. Biofuels* **9**, 129 (2016). doi: 10.1186/s13068-016-0549-3; pmid: 27340431
41. T. Hino, K. Miyazaki, S. Kuroda, Role of extracellular acetate in the fermentation of glucose by a ruminal bacterium, *Megasphaera elsdenii*. *J. Gen. Appl. Microbiol.* **37**, 121–129 (1991). doi: 10.2323/jgam.37.121
42. P. Vos *et al*., *Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes* (Springer, 2011).
43. T. D. Mays, L. V. Holdeman, W. E. C. Moore, M. Rogosa, J. L. Johnson, Taxonomy of the Genus *Veillonella* Prévot. *Int. J. Syst. Bacteriol.* **32**, 28–36 (1982). doi: 10.1099/00207713-32-1-28
44. D. Dietz, W. Sabra, A. P. Zeng, Co-cultivation of *Lactobacillus zeae* and *Veillonella criceti* for the production of propionic acid. *AMB Express* **3**, 29 (2013). doi: 10.1186/2191-0855-3-29; pmid: 23705662
45. T. Pielhop, J. Amgarten, P. R. von Rohr, M. H. Studer, Steam explosion pretreatment of softwood: The effect of the explosive decompression on enzymatic digestibility. *Biotechnol. Biofuels* **9**, 152 (2016). doi: 10.1186/s13068-016-0567-1; pmid: 27453727
46. R. P. Overend, E. Chornet, J. A. Gascoigne, Fractionation of Lignocellulosics by Steam-Aqueous Pretreatments. *Philos. Trans. R. Soc. A* **321**, 523–536 (1987).
47. C. Xiros, M. H. Studer, A Multispecies Fungal Biofilm Approach to Enhance the Celluloyltic Efficiency of Membrane Reactors for Consolidated Bioprocessing of Plant Biomass. *Front. Microbiol.* **8**, 1930 (2017). doi: 10.3389/fmicb.2017.01930; pmid: 29067006
48. G. L. Miller, Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* **31**, 426–428 (1959). doi: 10.1021/ac60147a030
49. A. Sluiter *et al*., *Determination of Structural Carbohydrates and Lignin in Biomass* (National Renewable Energy Laboratory, rev. ed., 2012); www.nrel.gov/docs/gen/fy13/42618.pdf.
50. W. P. Inskeep, P. R. Bloom, Extinction coefficients of chlorophyll a and b in *n,n*-dimethylformamide and 80% acetone. *Plant Physiol.* **77**, 483–485 (1985). doi: 10.1104/pp.77.2.483; pmid: 16664080
51. A. R. Wellburn, The Spectral Determination of Chlorophylls a and b, as well as Total Carotenoids, Using Various Solvents with Spectrophotometers of Different Resolution. *J. Plant Physiol.* **144**, 307–313 (1994). doi: 10.1016/S0176-1617(11)81192-2

ACKNOWLEDGMENTS

We thank Genencor-DuPont for providing the enzymes used in this work and S. G. Oliver for generously providing us with *S. cerevisiae* FYdelta pet191::KanMX4. **Funding:** Supported by the Swiss National Science Foundation in the framework of National Research Program NRP 70 "Energy Turnaround" grant 407040-153868 (M.H.S.); Biotechnology and Biological Sciences Research Council (BBSRC) Innovate UK Industrial Biotechnology Catalyst

grant BB/M018210/01 "Developing Platforms for the Production of Diterpenoids" (A.G.S.); BBSRC Global Challenges Research Fund grant BB/S0014020/1 "Enhancing cobalamin (vitamin B12) bioavailability in culturally appropriate foods in India" (A.G.S.); and Leverhulme Trust grant RPG-2017-077 "Snow Algae—are they the most abundant photosynthetic organisms in terrestrial Antarctica?" (A.G.S.). The study has been conducted in the framework of the SCCER BIOSWEET, supported by the Swiss Confederation through Innosuisse–Swiss Innovation Agency (M.H.S.).
Author contributions: R.L.S., S.B., J.S.L., A.G.S., S.V., and M.H.S. conceived the study and designed the experiments; R.L.S. conducted the experiments; and R.L.S., S.B., M.P.D., J.S.L., A.G.S., S.V., and M.H.S. analyzed the experimental results; R.L.S. and S.B. drafted the manuscript. All authors critically revised the article and approved the final manuscript. **Competing interests:** R.L.S. and M.H.S. submitted an international patent application (EP 2019/061402) on a method for the microbial production of fatty acids and lipids from lignocellulose with a consortium of at least three microorganisms. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/eabb1214/suppl/DC1
Figs. S1 to S20
Tables S1 to S9
References (*52*–*70*)

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

14 February 2020; accepted 26 June 2020
10.1126/science.abb1214

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

A synthetic synaptic organizer protein restores glutamatergic neuronal circuits

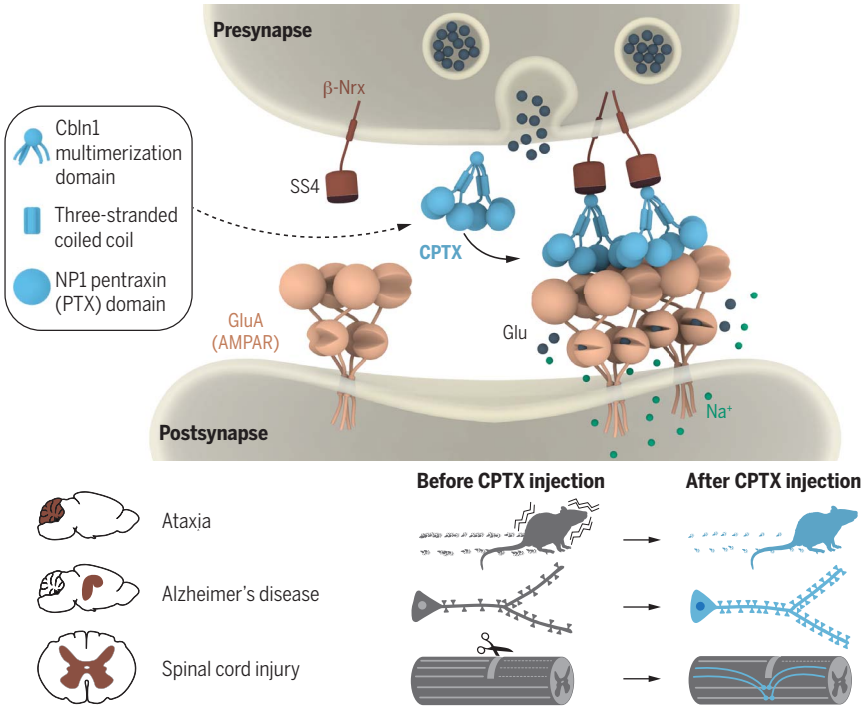
Kunimichi Suzuki*, Jonathan Elegheert*, Inseon Song*, Hiroyuki Sasakura*, Oleg Senkov, Keiko Matsuda, Wataru Kakegawa, Amber J. Clayton, Veronica T. Chang, Maura Ferrer-Ferrer, Eriko Miura, Rahul Kaushik, Masashi Ikeno, Yuki Morioka, Yuka Takeuchi, Tatsuya Shimada, Shintaro Otsuka, Stoyan Stoyanov, Masahiko Watanabe, Kosei Takeuchi, Alexander Dityatev†, A. Radu Aricescu†, Michisuke Yuzaki†

INTRODUCTION: Synapses are fundamental structural and functional units within neural circuits, where they define the connectivity between neurons and provide avenues for communication. At a molecular level, synapses are highly dynamic and their remodeling is essential for all aspects of brain physiology. However, errors in this process can happen and often lead to an imbalance of excitatory and inhibitory signaling. This is thought to be a major cause of neuropsychiatric or neurological disorders, including autism spectrum disorders, epilepsy, schizophrenia, and Alzheimer's disease. Thus, molecular tools to control the number and/or function of synapses would be highly desirable. Physiologically, synapse formation is driven by synaptic organizer proteins. Among these, extracellular scaffolding proteins (ESPs) such as cerebellin-1 (Cbln1) and neuronal pentraxin-1 (NP1) are distinctive in that they could rapidly induce synapse differentiation by binding pre- and/or postsynaptic cell surface proteins at the synaptic cleft. We hypothesized that synthetic molecules that would combine structural features of Cbln1 and NP1 could be used to efficiently reverse the loss of excitatory synapses and promote the structural and functional recovery of damaged neuronal circuits in animal models of neurological disease.

RATIONALE: NP1 recruits postsynaptic AMPA-subtype ionotropic glutamate receptors (AMPARs), responsible for excitatory neurotransmission, through its pentraxin domain. However, NP1 does not seem to induce presynaptic specializations in vivo. By contrast, Cbln1 promotes presynaptic differentiation by interacting with the cell adhesion molecule neurexin (Nrx) through its N-terminal multimerization domain, but cannot bind AMPARs. Guided by structural information, we developed a hexameric synthetic soluble ESP, termed CPTX, which includes the multimerization domain of Cbln1 and the pentraxin domain of NP1. We hypothesized that CPTX should induce Nrx–CPTX–AMPAR transsynaptic molecular bridges, and thus accumulate and align presynaptic vesicle release machinery and postsynaptic neurotransmitter receptors.

RESULTS: Recombinant CPTX selectively bound presynaptic Nrx containing the spliced sequence 4 [Nrx(+4)] with nanomolar affinity and most AMPAR subtypes with micromolar affinity. When administered to cerebellar granule cells and hippocampal neurons in vitro, CPTX acted as a bidirectional synapse organizer and induced excitatory pre- and postsynaptic sites. In vivo, CPTX increased the number of functional excitatory synapses and improved gait performance upon injection into the cerebellum of the ataxic *Cbln1*-null and *GluD2*-null mice. Furthermore, when injected into the hippocampus of 5xFAD mice, a model of familial Alzheimer's disease, CPTX restored dendritic spine numbers, excitatory synaptic transmission, and long-term potentiation and improved hippocampus-dependent learning. Finally, in mouse models of spinal cord injury, single injections of CPTX into the damaged tissue were sufficient to reorganize excitatory circuits and restore locomotion for more than 7 to 8 weeks.

CONCLUSION: We developed a synthetic, structure-guided, synaptic organizer termed CPTX, which induced functional and structural excitatory synapses in the cerebellar, hippocampal, and spinal cord neuronal circuits in vivo. Molecular components involved in excitatory synapses are considerably different among neuronal circuits. Rationally designed ESPs targeting distinct pre- and postsynaptic molecules may be useful to modulate neural circuit connectivity. This approach may inspire the development of a variety of innovative molecular tools for basic neuroscience as well as the treatment of neurological disorders. ■



Structure-guided design and applications of the synthetic synapse organizer CPTX. CPTX combines structural elements from the endogenous synaptic organizers Cbln1 and NP1 to bridge presynaptic neurexins (Nrxs) and postsynaptic AMPA receptors (AMPARs) across the synaptic cleft. CPTX can rapidly induce the formation of excitatory synapses and efficiently reorganize and functionally restore neuronal circuits in mouse models of ataxia, Alzheimer's disease, and spinal cord injury.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: alexander.dityatev@dzne.de (A.D.); radu@mrc-lmb.cam.ac.uk (A.R.A.); myuzaki@keio.jp (M.Y.)
Cite this article as K. Suzuki *et al.*, *Science* **369**, eabb4853 (2020). DOI: 10.1126/science.abb4853

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb4853

RESEARCH ARTICLE

NEUROSCIENCE

A synthetic synaptic organizer protein restores glutamatergic neuronal circuits

Kunimichi Suzuki[1]*†, Jonathan Elegheert[2]†‡, Inseon Song[3]†, Hiroyuki Sasakura[4]†, Oleg Senkov[3], Keiko Matsuda[1], Wataru Kakegawa[1], Amber J. Clayton[2]§, Veronica T. Chang[2,5], Maura Ferrer-Ferrer[3], Eriko Miura[1], Rahul Kaushik[3,6], Masashi Ikeno[4], Yuki Morioka[4], Yuka Takeuchi[4], Tatsuya Shimada[1], Shintaro Otsuka[1], Stoyan Stoyanov[3], Masahiko Watanabe[7], Kosei Takeuchi[4], Alexander Dityatev[3,6,8]||, A. Radu Aricescu[2,5]||, Michisuke Yuzaki[1]||

Neuronal synapses undergo structural and functional changes throughout life, which are essential for nervous system physiology. However, these changes may also perturb the excitatory–inhibitory neurotransmission balance and trigger neuropsychiatric and neurological disorders. Molecular tools to restore this balance are highly desirable. Here, we designed and characterized CPTX, a synthetic synaptic organizer combining structural elements from cerebellin-1 and neuronal pentraxin-1. CPTX can interact with presynaptic neurexins and postsynaptic AMPA-type ionotropic glutamate receptors and induced the formation of excitatory synapses both in vitro and in vivo. CPTX restored synaptic functions, motor coordination, spatial and contextual memories, and locomotion in mouse models for cerebellar ataxia, Alzheimer's disease, and spinal cord injury, respectively. Thus, CPTX represents a prototype for structure-guided biologics that can efficiently repair or remodel neuronal circuits.

A broad range of neuropsychiatric and neurological disorders, including autism spectrum disorders, epilepsy, schizophrenia, and Alzheimer's disease (AD), are thought to be caused by an imbalance between excitatory and inhibitory (E/I) synaptic functions (*1*–*5*). During organism development, but also throughout life, the formation and remodeling of complex yet precise neuronal circuits rely on specific synaptic organizing proteins (*6*, *7*). These include cell adhesion molecules, such as neurexins (Nrxs), neuroligins (*8*, *9*), and receptor protein tyrosine phosphatases (*10*), as well as secreted proteins, such as fibroblast growth factors, semaphorins, Wnt, and extracellular scaffolding proteins (ESPs) (*7*, *11*, *12*). ESPs directly connect pre- and postsynaptic membrane proteins to form molecular bridges that span the synaptic cleft and mediate bidirectional signaling. For example, cerebellin-1 (Cbln1) is released from parallel fibers (PFs; axons of cerebellar granule cells) and contributes to the synapse-spanning tripartite complex Nrx–Cbln1–GluD2 (the ionotropic glutamate receptor family member delta-2) (Fig. 1A, left panel). Cbln1 can simultaneously bind Nrx isoforms containing the 30-residue "spliced sequence 4" (SS4) insert [Nrx(+4)] expressed at PF terminals and the amino-terminal domains (ATDs) of GluD2 on Purkinje cells (PCs) (*13*–*15*). Notably, a single injection of recombinant Cbln1 into the cerebellum can restore ~75% of PF–PC synapses and normal motor coordination within 1 day in adult *Cbln1*-null mice in vivo (*16*). Cerebellins (Cbln1–4) are expressed in nearly all brain regions with distinct patterns and developmental dynamics (*17*, *18*). Cbln1 promotes the formation or maintenance of excitatory synapses in the nucleus accumbens (*19*), whereas Cbln2 is necessary for excitatory synapses in the interpeduncular nucleus (*20*) and the hippocampus (*21*, *22*). Cbln4 mediates the formation of inhibitory synapses between pyramidal neurons and cortical interneurons (*23*) by interacting postsynaptically with GluD1 (*24*), a receptor closely related to GluD2 (*25*). Other ESPs, such as the C1q-related C1ql1, promote excitatory synapse formation and maintenance in the cerebellum by binding to the adhesion G protein–coupled receptor BAI3 (*26*, *27*). C1ql3 also mediates synapse formation and maintenance in the medial prefrontal cortex (*28*). C1ql2 and C1ql3 are produced by mossy fibers and recruit kainate-subtype ionotropic glutamate receptors (KARs) by directly binding to the ATDs of GluK2/4 KAR subunits in CA3 hippocampal neurons (*29*, *30*). Molecular components of excitatory synapses differ considerably among neuronal circuits (*31*, *32*), and other ESPs likely remain to be discovered. Nevertheless, the Cbln1–Cbln4 and C1ql1–C1ql3 examples prompted us to hypothesize that synthetic molecules with defined pre- and postsynaptic binding specificities could be designed and employed to reverse the loss of synapses and promote the structural and functional recovery of damaged neuronal circuits.

[1]Department of Physiology, Keio University School of Medicine, Tokyo 160-8582, Japan. [2]Division of Structural Biology, University of Oxford, Oxford OX3 7BN, UK. [3]Molecular Neuroplasticity, German Center for Neurodegenerative Diseases (DZNE), 39120 Magdeburg, Germany. [4]Department of Medical Cell Biology, School of Medicine, Aichi Medical University, Aichi, Japan. [5]Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. [6]Center for Behavioral Brain Sciences (CBBS), 39106 Magdeburg, Germany. [7]Department of Anatomy, Hokkaido University Graduate School of Medicine, Sapporo 060-8638, Japan. [8]Medical Faculty, Otto von Guericke University, 39120 Magdeburg, Germany.
*Present address: Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK.
†These authors contributed equally to this work.
‡Present address: Interdisciplinary Institute for Neuroscience (IINS), UMR5297 CNRS/UB, 33076 Bordeaux, France.
§Present address: Adaptimmune, 60 Jubilee Avenue, Milton Park, Abingdon OX14 4RX, UK.
||Corresponding author. Email: alexander.dityatev@dzne.de (A.D.); radu@mrc-lmb.cam.ac.uk (A.R.A.); myuzaki@keio.jp (M.Y.)

Design of a synthetic ESP-type synaptic organizer

Neuronal pentraxins (NPs) are a family of oligomeric secreted (NP1 and NP2) or membrane-attached (NPR) proteins that induce clustering of postsynaptic AMPA-type ionotropic glutamate receptors (AMPARs) through direct interactions between pentraxin (PTX) domains and the AMPAR ATDs (*33*–*36*). Astrocyte-secreted glypican 4 indirectly leads to AMPAR recruitment by enhancing the release of NP1 from presynaptic axons (*37*). However, unlike Cbln1 (*13*, *16*) and Cbln4 (*23*, *24*), NPs do not seem to induce presynaptic specializations in vivo (*38*, *39*) (Fig. 1A, middle panel). Thus, we hypothesized that we could exploit the modular architecture of NP1 and Cbln1 to develop a synthetic hexameric ESP, termed CPTX, which would include the N-terminal cysteine-rich region (CRR) of Cbln1 that binds Nrx(+4) and the PTX domain of NP1 (Fig. 1A, right panel). CPTX should have the potential to induce the formation of excitatory synapses by organizing Nrx nanoclusters (*40*) and to form a transsynaptic complex by recruiting postsynaptic AMPARs (*41*).

To define the correct domain boundaries for CPTX assembly, we first solved the crystal structure of the PTX domain of NP1 ($NP1_{PTX}$) at 1.45 Å resolution (Fig. 1B and table S1). Reminiscent of short pentraxins (*42*), such as the C-reactive protein and serum amyloid P component, $NP1_{PTX}$ formed a two-layered β sheet with a flattened jellyroll topology containing two Ca^{2+} ions (Fig. 1B and fig. S1). However, unlike short pentraxins, $NP1_{PTX}$ is monomeric in the crystal as well as in solution, as confirmed by multi-angle light scattering (MALS; Fig. 1C). Surface plasmon resonance (SPR) assays, performed to identify the minimal domain requirements for NP1–GluA interactions, revealed that $NP1_{PTX}$ bound the ATDs of GluA1, GluA3, and GluA4 AMPARs with affinities in the high micromolar range (Fig. 1D). A recent analysis of the Nrx–Cbln1–GluD2 transsynaptic complex showed a similarly weak interaction between the globular domain of Cbln1 and the GluD2 ATD; however, avidity effects arising from the oligomeric nature of Cbln1 (hexamer) and GluD2 (tetramer) increase the apparent affinity between the full-length partners considerably (*15*). To mimic the structural organization of Cbln1 in CPTX, we first

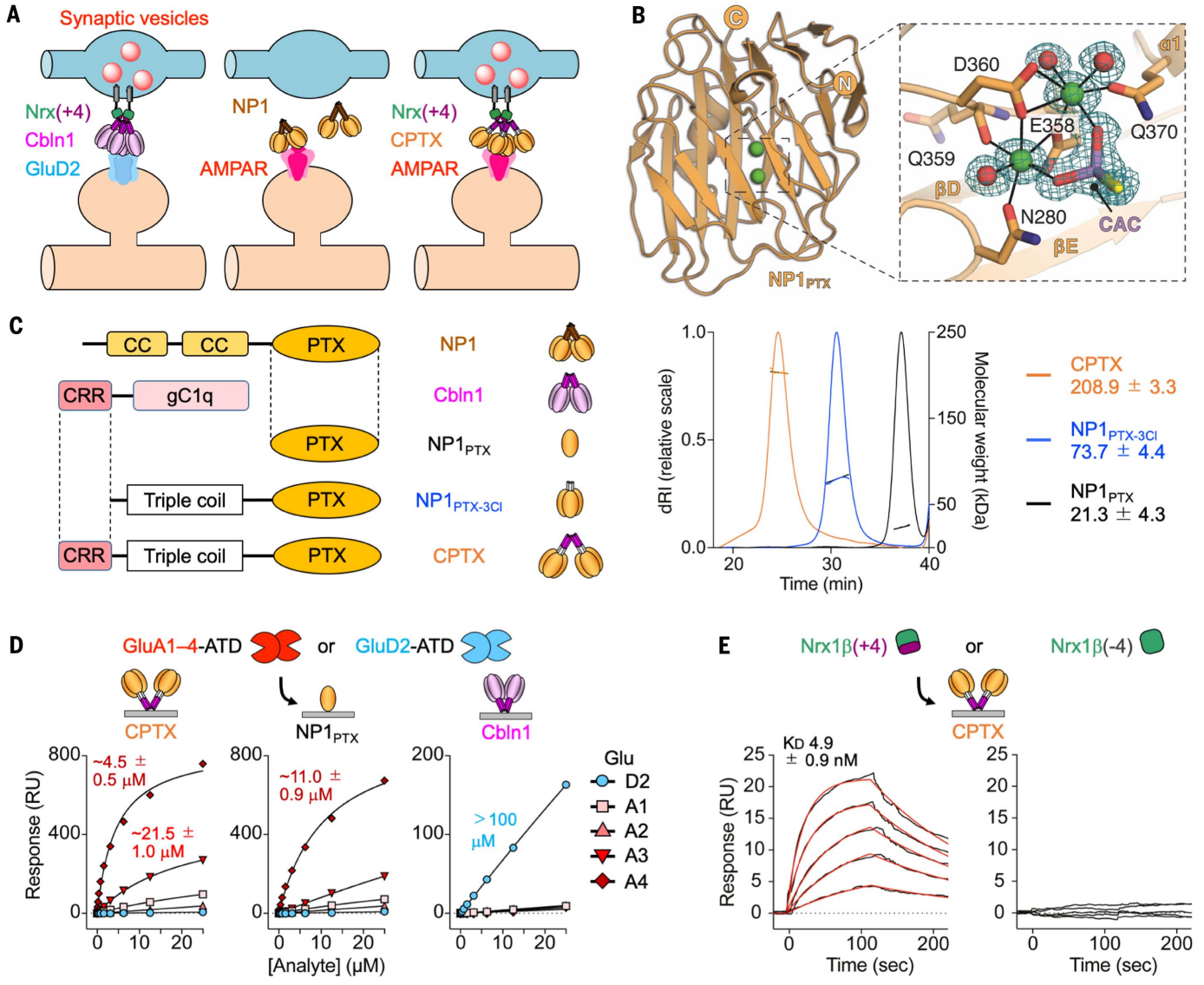


Fig. 1. Structure-guided design and biophysical characterization of CPTX. (**A**) The CPTX concept. Whereas Cbln1 induces synapse formation by binding to postsynaptic GluD2 and presynaptic Nrx(+4), NP1 induces clustering of postsynaptic AMPARs without inducing presynaptic elements. CPTX is a chimeric ESP consisting of Cbln1 and NP1 structural elements. (**B**) Crystal structure of the pentraxin domain of human NP1 ($NP1_{PTX}$). The N and C termini are annotated. Calcium ions (Ca^{2+}) and water molecules are shown as green and red spheres, respectively. 2mFo-DFc electron density is contoured at 1.0 σ. The Ca^{2+} coordination shell, which includes a cacodylate buffer molecule (CAC), is indicated by black lines. (**C**) Diagrams of CPTX and related constructs. CC, coiled-coil domain; PTX, pentraxin domain; CRR, cysteine-rich region; gC1q, globular C1q domain; 3Cl, triple coil. The graph on the right shows molecular masses of monomeric $NP1_{PTX}$, trimeric $NP1_{PTX\text{-}3Cl}$, and hexameric CPTX. dRI, differential refractive index. (**D**) Binding isotherms for the interactions between the GluA1–4 or GluD2 ATDs and immobilized CPTX, $NP1_{PTX}$ and Cbln1. The GluA4 ATD interacted strongest with CPTX (K_D ~ 4.5 ± 0.5 μM) and $NP1_{PTX}$ (K_D ~ 11.0 ± 0.9 μM), but not with Cbln1. By contrast, the ATD of GluD2 only interacted with Cbln1 (K_D > 100 μM). (**E**) Single-cycle kinetic sensorgrams for the interaction between the ectodomain of Nrx1β(±4) and immobilized CPTX. CPTX directly bound to Nrx1β(+4) (K_D ~ 4.90 ± 0.90 nM), but not to Nrx1β(−4).

added a triple-coil–forming mutant GCN4 peptide to the N terminus of $NP1_{PTX}$, resulting in the $NP1_{PTX\text{-}3Cl}$ construct, and confirmed its trimeric stoichiometry (Fig. 1C). The Cbln1 CRR (*15*) was then added to the $NP1_{PTX\text{-}3Cl}$ N terminus, leading to the hexameric CPTX molecule (Fig. 1C and fig. S1).

We next examined whether CPTX was equipped with the intended dual binding capacity in vitro. SPR assays showed that CPTX and $NP1_{PTX}$ bound the ATD of GluA4 with comparable apparent affinities (dissociation constant K_D = 4.5 ± 0.5 μM and 11.0 ± 0.9 μM, respectively; Fig. 1D and fig. S2). Like $NP1_{PTX}$, CPTX also bound to the other GluA ATDs with a GluA3 > GluA1 > GluA2 preference. However, these interactions were weaker, and accurate K_D values could not be determined reliably (Fig. 1D and fig. S2). In addition, CPTX bound specifically to the Nrx1β(+4) ectodomain (K_D = 4.9 ± 0.9 nM), but not Nrx1β(−4) (Fig. 1E), thus retaining the strict isoform recognition specificity of Cbln1. Cell-based binding assays also showed that CPTX bound specifically to human embryonic kidney (HEK) cells displaying GluA1–GluA4 ATDs on their surface, whereas Cbln1 only bound

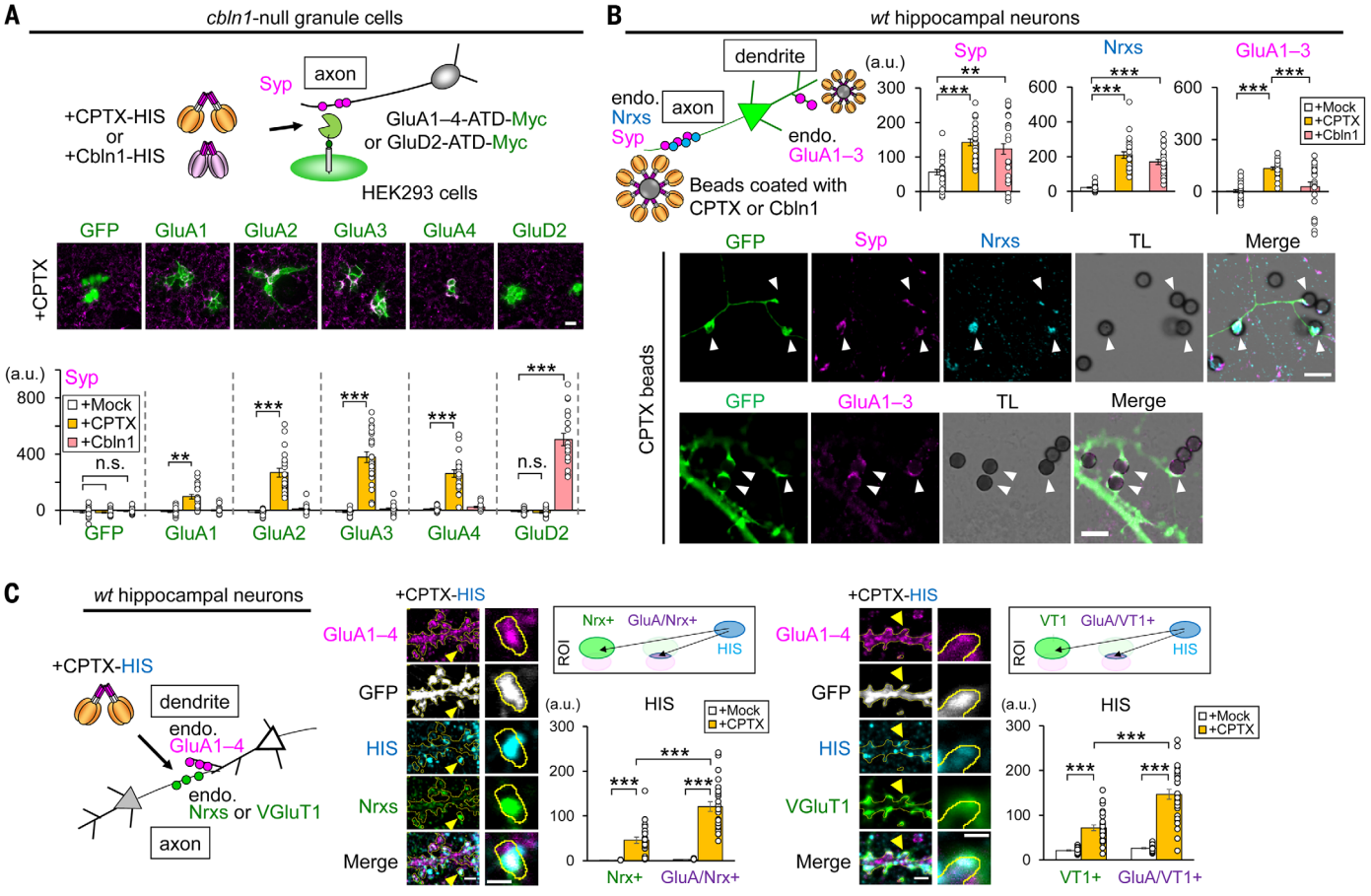


Fig. 2. CPTX directly induces excitatory pre- and postsynaptic sites in vitro. (**A**) CPTX induced the accumulation of presynaptic terminals of cerebellar granule cells onto co-cultured HEK293 cells displaying AMPAR ATDs. The intensities of the synaptophysin immunoreactivity (Syp; magenta) onto HEK293 cells (green) are quantified in the lower graph. Mock, vehicle [HEPES buffered saline (HBS)] controls. The error bars represent the mean ± SEM. ***P < 0.001, **P < 0.01; n.s., not significant; n = 16 to 22 fields from seven independent experiments, one-way analysis of variance (ANOVA) followed by Tukey's test. Scale bar, 20 μm. (**B**) Beads coated with CPTX induced the formation of presynaptic boutons positive for endogenous synaptophysin (Syp) and neurexins (Nrxs), and postsynaptic boutons positive for GluA1–3, in co-cultured hippocampal neurons transfected with GFP. Arrowheads indicate beads immunopositive for Syp (magenta, middle panel), Nrxs (cyan), or GluA1–3 (magenta, lower panel). Mock, beads coated with anti-HIS antibody. TL, transmitted light. Scale bars, 5 μm. The intensities of the Syp, Nrxs and GluA1–3 immunoreactivity onto beads are quantified in the upper graphs. The error bars represent the mean ± SEM. ***P < 0.001, **P < 0.01, n = 21 fields from five experiments, one-way ANOVA followed by Tukey's test. (**C**) CPTX localizes at excitatory synapses. Representative immunocytochemical staining images show Nrxs or VGluT1 (green), GluA1–4 (magenta), and HIS-tagged CPTX (cyan) in hippocampal neurons. Dendritic spines indicated by yellow arrowheads are magnified. Mean intensities of HIS signals in the Nrx-positive (Nrx+) or VGluT1-positive (VT1+) areas were measured and compared to those in the areas overlapping with GluA1–3 signals (GluA/Nrx+ or GluA/VT1+). The error bars represent the mean ± SEM. ***P < 0.001, n = 24 to 26 fields from two to three independent experiments, Student's t test with Bonferroni correction. Scale bars, 2 μm (left column of each panel), 1 μm (magnifications).

those displaying the GluD2 ATD (fig. S3A). By contrast, both Cbln1 and CPTX bound to HEK293 cells expressing full-length Nrx1α, Nrx1β, Nrx2β, and Nrx3β only when the SS4 insert was present (fig. S3B). The ectodomain of Nrx1β(+4) specifically bound to HEK293 cells expressing GluA4 ATD under CPTX application (fig. S3C). Thus, CPTX can simultaneously bind GluAs and neurexin isoforms containing the SS4 insert.

CPTX induces excitatory pre- and postsynaptic sites in vitro

We examined whether CPTX could serve as a synaptic organizer with the designed GluA versus GluD specificity in neuronal cell cultures. Application of CPTX to *Cbln1*-null cerebellar granule cells induced punctate accumulation of endogenous synaptophysin, a synaptic vesicle marker, in axons that contact HEK293 cells displaying GluA ATDs on their surface, but not in those that contact cells displaying the GluD2 ATD (Fig. 2A and fig. S4A). Similarly, CPTX induced accumulation of synaptophysin (fig. S4B) or vesicular glutamate transporter 1 (VGluT1) (fig. S4C) in wild-type hippocampal neurons that contacted HEK293 cells displaying GluA ATDs. By contrast, NP1 could bind to the surface of HEK293 cells displaying the GluA4 ATD but failed to accumulate synaptophysin (fig. S4D). Conversely, application of Cbln1 induced presynaptic synaptophysin accumulation in axons that contacted co-cultured HEK293 cells expressing GluD2, but not GluA ATDs (Fig. 2A and fig. S4, A and B). To examine the direct effect of CPTX, we applied beads coated with CPTX to hippocampal neurons in culture. Cbln1- or CPTX-coated beads accumulated presynaptic sites positive for endogenous synaptophysin, Nrx, and VGluT1 (Fig. 2B and fig. S5, A and B), indicating that CPTX can directly induce presynapses in vitro.

We further asked whether CPTX could induce postsynaptic sites in vitro. Coexpression of Nrx1β(+4) and CPTX, but not Cbln1, on the

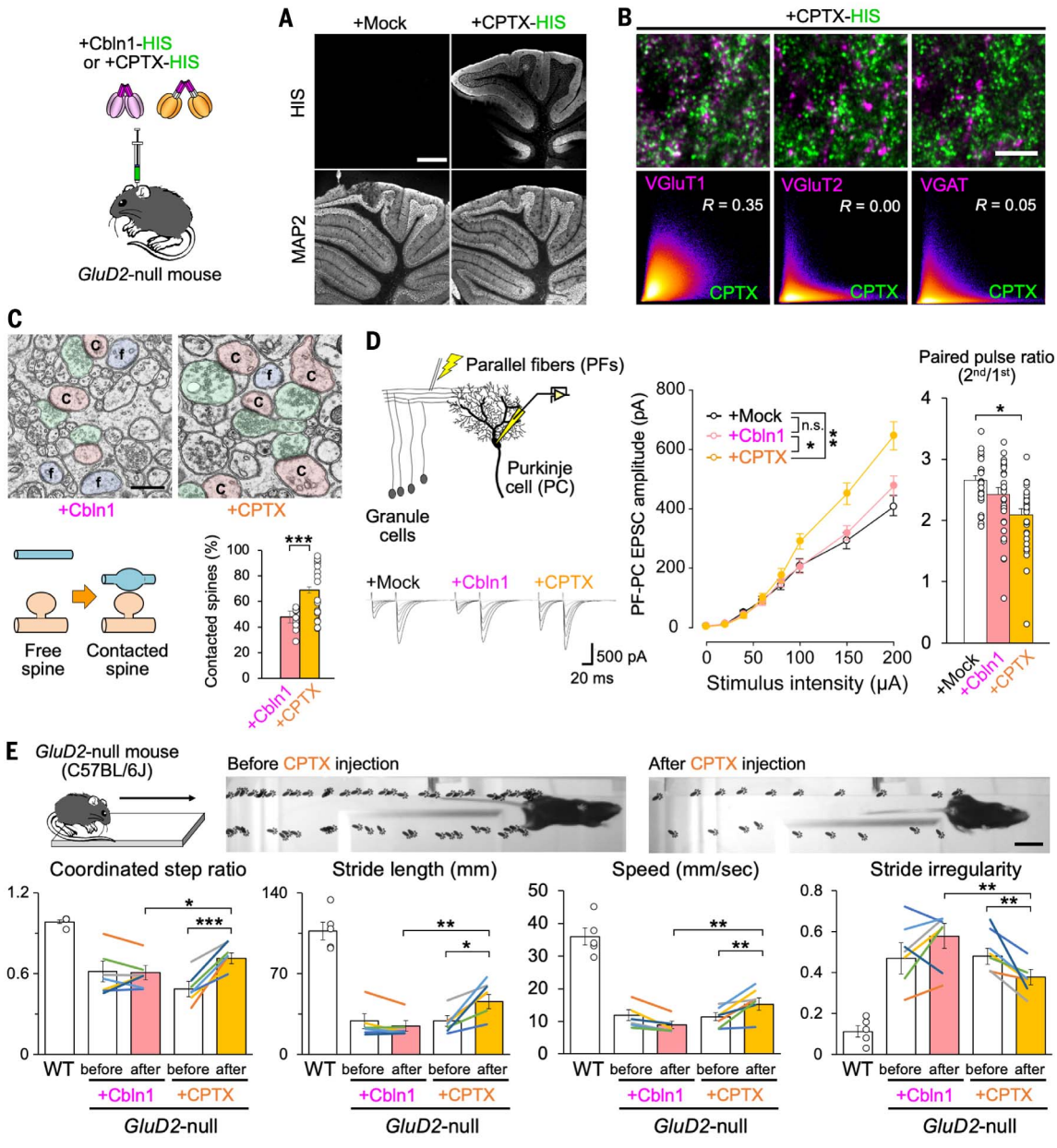


Fig. 3. CPTX restores PF-PC synapses and motor coordination in *GluD2*-null mice. (**A**) HIS-tagged CPTX injected into the *GluD2*-null cerebellum localizes at PF–PC synapses. Mock, vehicle. Scale bar, 500 μm. (**B**) HIS immunoreactivity in the molecular layer of the injected area (A) colocalized with VGluT1 (a PF marker), but not VGluT2 (a climbing fiber marker) or VGAT (an inhibitory input marker). The scatter plots show the 2D pixel intensity histograms for CPTX (green) and each synapse marker (magenta). Pearson's correlation coefficient *R* values are indicated. Scale bar, 2 μm. (**C**) Representative electron microscopic images show free dendritic spines (f, blue) and contacted spines (c, red) innervated by PFs (green) in the *GluD2*-null cerebellum 3 days after injection of Cbln1 or CPTX. Scale bar, 500 nm. The fractions of contacted PC spines are quantified in the lower graph. ***$P < 0.001$, n = 10 to 20 sections from one or two mice, χ^2 test. (**D**) CPTX restores functional PF–EPSCs in *GluD2*-null mice 3 days after injection. Representative traces are shown. The middle graph shows averaged input-output relationships of PF–EPSCs for each treatment. **$P < 0.01$, *$P < 0.05$; n.s., not significant; n = 30 cells each, two-way ANOVA followed by Scheffe post-hoc test. The right graph shows the paired-pulse ratio of the second to first PF–EPSC amplitudes. *$P < 0.05$, n = 30 cells each, Kruskal-Wallis test followed by Scheffe *post-hoc* test. (**E**) CPTX improves the gait of *GluD2*-null mice 3 days after the injection. Representative footprints before and after CPTX injection are shown near the top. The lower graphs show the quantification of gait parameters. Each line represents an individual score before and after injection with either Cbln1 or CPTX. ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, n = 5 or 6 mice, Student's paired t test. Scale bar, 50 mm. The error bars in (C) to (E) represent the mean ± SEM.

surface of HEK293 cells facilitated accumulation of endogenous GluA1–3 in dendrites of contacted hippocampal neurons (fig. S6A). Similarly, beads coated with CPTX, but not Cbln1, accumulated endogenous GluA1–GluA3 in dendrites of most hippocampal neurons (Fig. 2B and fig. S6B) and GluA4 in parvalbumin-positive (PV^+) interneurons (fig. S6C). Immunocytochemical analyses revealed that soluble CPTX added to the culture medium accumulated between Homer- and Bassoon-positive puncta (fig. S7A) and at puncta double immunopositive for GluA1–4 and Nrx or VGluT1 (Fig. 2C and fig. S7B). CPTX also increased Nrx and GluA intensities in GluA/Nrx- and GluA/VGluT1–double immunopositive areas, respectively. Thus, CPTX serves as a bidirectional synaptic organizer by binding presynaptic Nrx(+4) and postsynaptic AMPARs in vitro.

Rescue of synapse formation and motor coordination by CPTX in cerebellar ataxia mice

To test the impact of CPTX in vivo, we injected it into the cerebellum (lobules VI and VII) of adult *Cbln1*-null mice. Immunohistochemical analyses 1 day after the injection showed that CPTX was preferentially confined to the injected molecular layer (fig. S8A) and colocalized with VGluT1 (a PF terminal marker), but not VGluT2 (a climbing fiber terminal marker)

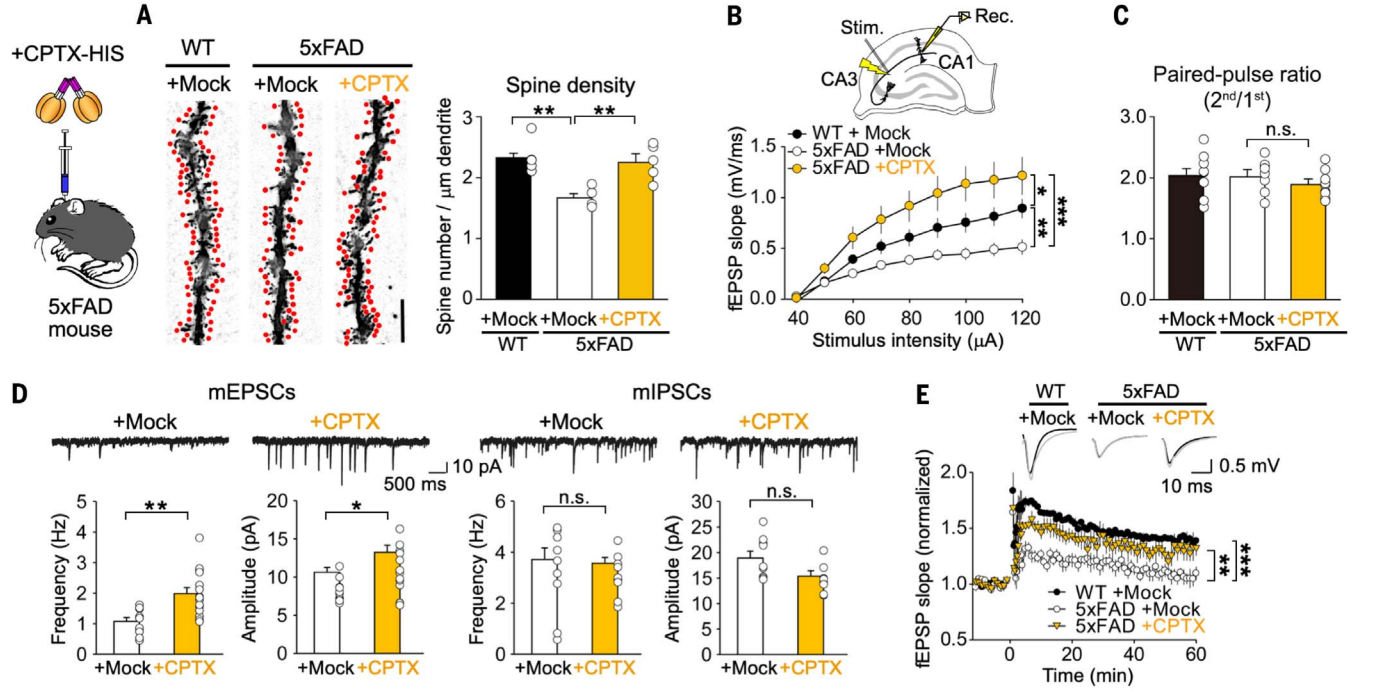


Fig. 4. CPTX restores spines and LTP in the hippocampus of an Alzheimer's disease model. (**A**) CPTX restores dendritic spine density in 5xFAD mice. Representative Golgi-Cox staining of apical dendrites of CA1 pyramidal neurons in the dorsal hippocampus. Red marks indicate counted spines. Mock, vehicle (HBS buffer) controls. Scale bar, 5 µm. The graph shows the averaged spine density for seven or eight secondary apical dendrites for each animal. The bars represent the mean ± SEM. $^{**}P < 0.01$, $n = 4$ or 5 mice, one-way ANOVA followed by Holm-Sidak post-hoc test. (**B**) CPTX restores Schaffer collateral (SC)–evoked field excitatory postsynaptic potentials (fEPSPs) in 5xFAD mice. The graph shows averaged input-output relationships of SC–fEPSPs for each treatment. Mock, vehicle (HBS buffer) controls. The bars represent the mean ± SEM. $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, $n = 6$ to 9 slices, repeated two-way ANOVA. (**C**) CPTX does not affect the presynaptic release probability. The graph shows the paired-pulse ratio of the second to first SC-fEPSP amplitudes at 50-ms intervals. The bars represent the mean ± SEM. n.s., not significant, $n = 8$ or 9 slices, Student's t test. (**D**) CPTX increases the frequency and the amplitude of mEPSCs, but not mIPSCs, in the 5xFAD hippocampus. Representative traces are shown near the top. $^{**}P < 0.01$, $^{*}P < 0.05$, $n = 9$–15 cells, Student's t test. (**E**) CPTX restores LTP in 5xFAD mice. Application of theta-burst stimulation to SC three times induced a robust LTP at SC-CA1 synapses in the mock-treated WT, but not in the mock-treated 5xFAD hippocampus. Representative SC-fEPSP traces are shown. The bars represent the mean ± SEM. $^{***}P < 0.001$, $^{**}P < 0.01$, $n = 8$ or 9 slices, one-way ANOVA on ranks followed by Holm-Sidak post-hoc test for the mean values from the last 15 min of the recording.

or the vesicular γ-aminobutyric acid (GABA) transporter (VGAT, a marker for inhibitory inputs) (fig. S8B). Thus, CPTX was mainly located along PC dendrites where PFs that specifically express Nrx(+4) (*43*) make synapses. Electron microscopy analysis revealed that single injections of CPTX partially restored PF–PC synapses in *Cbln1*-null mice (fig. S8C). The effect was weaker than for Cbln1 (*13*, *16*), probably because there is a large excess of GluD2 relative to GluAs at PF–PC synapses (*44*). Whole-cell patch-clamp recordings in acute slice preparations showed that CPTX, but not $NP1_{PTX\text{-}3Cl}$, increased the amplitudes of PF-evoked excitatory postsynaptic currents (PF-EPSCs) (fig. S8D). Thus, CPTX most likely regulates synaptic functions by bridging pre- and postsynaptic sites. Furthermore, CPTX was as effective as Cbln1 in restoring normal gait patterns (figs. S8E and S9A and movies S1 and S2). We also observed recovery of motor coordination as measured by the rotor-rod test in the ICR mouse genetic background (fig. S9B), but not in the C57BL/6 one where the loss of the skilled motor coordination phenotype was too severe. Similar to Cbln1 (*16*), the effect of CPTX on the rotor-rod performance recovery was maximal at 3 days after injection but decayed afterward (fig. S9C), suggesting that continued presence of CPTX is necessary to maintain PF–PC synapses in the longer term. Thus, CPTX can induce functional PF–PC synapses that contribute to the restoration of the motor phenotypes in adult *Cbln1*-null mice.

We next investigated whether CPTX could serve as a synaptic organizer in the absence of GluD2 in vivo. As observed in *Cbln1*-null mice, injected CPTX was localized at PF–PC synapses in *GluD2*-null mice (Fig. 3, A and B). Although Cbln1 and GluD2 form a transsynaptic complex, *GluD2*-null mice are reported to show more severe motor phenotypes with less reduced PF–PC synapses (~40 to 60% of wild-type) than *Cbln1*-null mice (~20 to 30%) (*13*, *16*). Whereas Cbln1 was completely ineffective in *GluD2*-null mice in restoring morphological (Fig. 3C) and functional (Fig. 3D) PF–PC synapses, CPTX partially restored these phenotypes. CPTX, but not Cbln1, reset the increased paired-pulse ratio (PPR) of PF-EPSCs to wild-type values (*16*) (Fig. 3D). Furthermore, unlike Cbln1, CPTX improved the gait of *GluD2*-null mice (Fig. 3E and movie S3). Thus, CPTX can induce PF–PC synapses that contribute to cerebellar functions independent of GluD2 in vivo.

CPTX induces synaptic AMPAR accumulation in the wild-type hippocampus

To test if CPTX could act in brain regions outside the cerebellum, we injected CPTX into the dorsal hippocampus of Thy1–green fluorescent protein (GFP) and wild-type mice. Injected CPTX diffused throughout the entire CA1 regions of dorsal hippocampi and was detected between Bassoon-positive presynaptic terminals and GFP-positive postsynaptic sites of the CA1 pyramidal neurons in Thy1-GFP mice at 3 days after injection (fig. S10A). Super-resolution microscopy revealed that CPTX was localized between VGluT1 and AMPAR puncta in the CA1 stratum radiatum at 1 day after injection (fig. S10, B and C). By contrast, CPTX

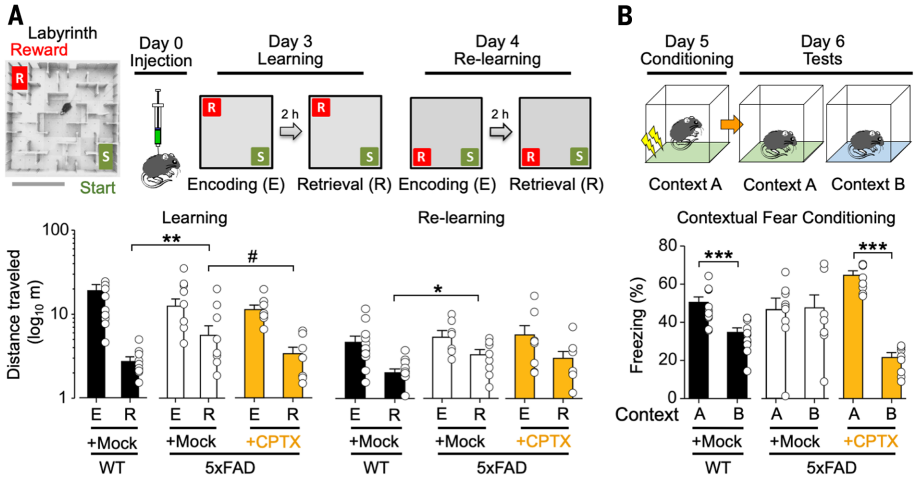


Fig. 5. CPTX restores hippocampus-dependent behaviors in an Alzheimer's disease model. (**A**) CPTX improves spatial memory in 5xFAD mice. CPTX or Mock (vehicle) was injected in 5xFAD mice on day 0. On day 3, mice were placed at the start point (S) of a 3D-printed maze and the pellet was placed at the reward point (R). Two hours after the initial encoding (E) session, mice were returned to the same start point to examine memory integrity in the retrieval (R) session. On the next day, the position of reward was changed, and reversal learning (relearning) was evaluated. Scale bar, 30 cm. The averaged total distances that mice traveled to reach the goal during the encoding and the retrieval sessions are shown in the lower graph. Log_{10} scaling of the *y* axis facilitates comparison of distances before and after training. The bars represent the mean ± SEM. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{\#}P < 0.1$, $n = 8$ to 11 mice, two-way repeated measures ANOVA followed by Fisher's post-hoc least significant difference (LSD) test. (**B**) CPTX restores context discrimination in 5xFAD mice. CPTX or Mock (vehicle) was injected in 5xFAD mice on day 0. Electrical shock was applied to wild-type and 5xFAD mice in context A on day 5. Freezing time was measured in context A and context B on day 6. The lower graph shows the mean (± SEM) freezing time of each group. $^{***}P < 0.001$, $n = 8$ to 11 mice, Student's paired *t* test.

did not show any colocalization with inhibitory synaptic markers (fig. S11). Immunohistochemistry of CPTX and each AMPAR subunit (fig. S12A) revealed that CPTX shifted the intensity histogram of GluA1 to higher values, while causing redistribution of GluA2/3 signals among puncta (fig. S12, B to D). As a result, CPTX increased the size of high-intensity (> mean + 3 SD) puncta of GluA1 and GluA2/3 (fig. S12, E and F). By contrast, although CPTX also shifted the intensity histogram of GluA4 to higher values (fig. S12D), it did not affect the size of high-intensity GluA4 puncta (fig. S12F). Thus, CPTX selectively binds to excitatory synapses in the hippocampus and modifies the distribution of AMPARs in a subunit-dependent manner.

Functionally, the CPTX injection increased the frequency of miniature EPSCs (mEPSCs) in CA1 pyramidal neurons as early as 4 hours after application to acute hippocampal slices (fig. S13A). CPTX also up-regulated both the amplitude and frequency of mEPSCs in hippocampal slices prepared 3 days after injection (fig. S13B) but did not enhance the frequency and amplitude of mEPSCs in CA1 PV^+ interneurons (fig. S14), which express high amounts of GluA4 and NP1/NPR (figs. S15 and S16, A, B, and E). In the CA1 lacunosum moleculare where NP1/NPR immunoreactivities were much weaker (fig. S16C), CPTX injections showed stronger effects in increasing the GluA1–3 intensity (fig. S17) than in the stratum radiatum (fig. S12). Thus, the impact of CPTX depends on the endogenous amounts of NPs: In CA1 PV^+ interneurons, the ATD of GluA4, where CPTX binds, would most likely be occupied by endogenous NPs whereas in other regions, where NPs are less abundant, CPTX appears more likely to bind and cluster GluAs.

CPTX restores spines, LTP, and cognition in an Alzheimer's disease model

Prompted by these observations, we next examined whether CPTX could modify neuronal circuits in the hippocampus under pathological conditions in vivo. We used 5xFAD mice, a model of early familial AD, which allowed us to test the impact of CPTX in the context of severe loss of neurons, synapses, synaptic plasticity, and cognitive abilities (*45*, *46*). We focused on Schaffer collateral (SC)–CA1 synapses, which are most relevant to contextual fear conditioning. Golgi staining showed that the reduced spine densities in 5xFAD mice (11 to 12 months old) were restored by injection of CPTX into the hippocampus of 5xFAD mice (Fig. 4A and movie S4). We next immunostained PSD95, the key excitatory synaptic scaffold protein required for synaptic stabilization, together with synaptophysin, GluA1–3, and GluN1, an essential *N*-methyl-D-aspartate (NMDA) receptor subunit in the hippocampus (fig. S18A). CPTX caused the redistribution of PSD95 as well as GluA1–3 signal intensities without affecting synaptophysin or GluN1 signals (fig. S18, B and C). CPTX also increased the proportion of large puncta positive for PSD95 and GluA1–3, but not GluN1 or synaptophysin (fig. S18D). Extracellular field recordings confirmed that excitatory postsynaptic potentials (fEPSPs) induced by SC stimulation were reduced in acute slices prepared from 5xFAD mice (Fig. 4B), as reported earlier (*47*). CPTX restored the amplitude of SC-fEPSPs, which became even larger than in the wild-type hippocampus (Fig. 4B). By contrast, $NP1_{PTX\text{-}3Cl}$ failed to increase SC-fEPSPs (fig. S19), a result consistent with the lack of effect of NP1 in inducing synapses in vitro (fig. S4D). CPTX did not affect the PPR of SC-fEPSPs, probably because the PPR was normal in 5xFAD hippocampus (Fig. 4C), a phenotype different from *GluD2*-null cerebellum (Fig. 3D). Moreover, whole-cell patch-clamp recordings from CA1 pyramidal neurons revealed that the frequency and amplitude of mEPSCs were increased in 5xFAD mice injected with CPTX (Fig. 4D). By contrast, miniature inhibitory synaptic currents (mIPSCs) were unaffected (Fig. 4D). Thus, CPTX specifically improved excitatory transmission between principal cells without changing GABAergic transmission in 5xFAD mice.

Impaired long-term potentiation at SC-CA1 synapses (SC-LTP) is believed to underlie certain cognitive abnormalities in 5xFAD mice (*47*). CPTX injection rescued impaired SC-LTP in middle-aged 5xFAD mice (Fig. 4E). Because CPTX did not increase GluN1 immunoreactivities, it is plausible that newly formed spines (Fig. 4A) and redistribution of PSD95 and AMPARs (fig. S18) may contribute to restoration of LTP. To evaluate spatial memory in a paradigm with minimal stress exposure, we placed three groups of mice in a labyrinth with food pellets as a reward and measured the total distance that mice traveled to reach the goal during the encoding and retrieval sessions (Fig. 5A). Although there was no difference in the travel distance between each group during the encoding session, 5xFAD mice traveled longer distances than wild-type mice during the retrieval session. CPTX injection 3 days before the learning test decreased the distance that 5xFAD mice traveled in the retrieval session, suggesting an improvement of spatial learning after the CPTX treatment. After reversal learning, when mice had to find a reward in a new position, the performance of 5xFAD mice was impaired in the mock-, but not in the CPTX-injected group, as compared to wild-type mice (Fig. 5A). The modest effect of CPTX in this test may be caused by the insufficient spread of CPTX because spatial memory requires

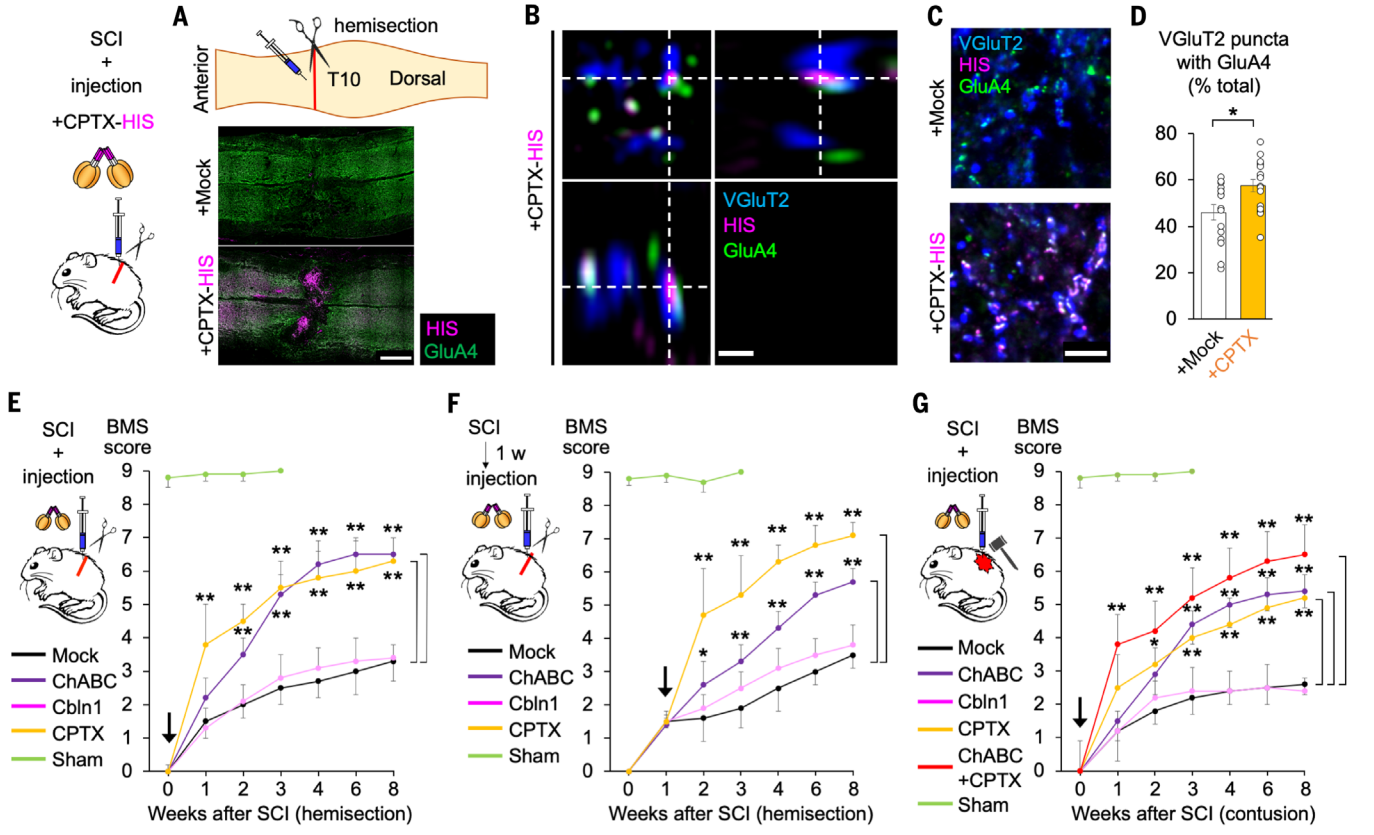


Fig. 6. CPTX promotes restoration of synapses and motor function in spinal cord injury. (**A**) Schematic depiction of SCI caused by hemisection at the 10th thoracic vertebra (T10). The lower panels show representative horizontal sections of the spinal cord immunostained for VGluT2, HIS (CPTX), and GluA4 upon Mock (vehicle) or CPTX injection. Scale bar, 0.5 mm. (**B**) Representative orthogonal images obtained by Airyscan super-resolution microscopy indicating the localization of CPTX in proximity of VGluT2- and GluA4-immunopositive puncta. Scale bar, 0.5 μm. (**C**) Representative immunohistochemical staining images of coronal sections stained for VGluT2 (blue), GluA4 (green), and HIS (CPTX; magenta) from mock- or CPTX-treated spinal cords. Scale bar, 5 μm. (**D**) Quantification of the fraction of GluA4+/VGluT2+–double-positive puncta. The bars represent the mean ± SEM. *P < 0.05, n = 16 slices from eight mice, Student's t test. (**E** to **G**) Time-course analyses of locomotion [Basso mouse scale (BMS) score] in SCI mice after injections. Mock (vehicle), chondroitinase ABC (ChABC), Cbln1, or CPTX were injected into the spinal cord immediately (E) or 1 week after (F) hemisection, or immediately after contusion by impactors (G) (70 kdyn impact force). For the sham controls, the spinal cord was surgically exposed without imposing injury or injection. Mice that showed a BMS score of 1.5 at 1 week after hemisection were selected for (F). **P < 0.01, *P < 0.05, n = 9 mice for each treatment, repeated two-way ANOVA with post-hoc Bonferroni-Dunn test (comparing the various treatments with Mock for each time point).

integration of information from head direction cells, grid cells, and place cells with landmarks (context). To more directly assess the function of the CA1 region, we next evaluated contextual fear conditioning by exposing the same three groups of mice to an electrical shock in context A and measuring the freezing time in the conditioned context A and a neutral context B 1 day later. 5xFAD mice failed to discriminate context A from B, i.e., showed generalized rather than context-specific learned fear responses. CPTX injection 5 days before fear conditioning improved context discrimination in 5xFAD mice to be even better than in mock-treated wild-type controls (Fig. 5B), indicating that CPTX could rescue the context generalization in 5xFAD mice. Thus, CPTX injection rescues the impaired basal excitatory transmission, SC-LTP, and hippocampus-dependent learning in middle-aged 5xFAD mice.

CPTX restores synapses and locomotion in spinal cord injury models

Reorganization and the E/I balance of spared intraspinal networks have recently been shown to contribute to functional recovery after spinal cord injury (SCI) (*48*, *49*). To test whether CPTX could restore glutamatergic circuits in the spinal cord, we performed dorsal hemisections at the thoracic vertebra 10 (T10) in mice and injected CPTX rostral to the injured region (Fig. 6A). Spinal cords were dissected 2 to 5 days after injections, and coronal serial sections were immunostained for VGluT2 and CPTX (fig. S20A). CPTX was distributed in the gray matter and was largely colocalized with VGluT2 and GluA4 immunopositive (VGluT2+ and GluA4+) puncta in a region 1.4 to 1.6 mm rostral to the epicenter of the injured site (fig. S20A–C). Super-resolution microscopy revealed that CPTX was located between VGluT2+ and GluA4+ puncta (Fig. 6B). Most excitatory spinal cord interneurons that are involved in locomotion express high amounts of VGluT2, GluA4 (*50*) and Nrx(+4) (*51*). By contrast, NP1/NPR immunoreactivities were very weak in most of the spinal cord (fig. S16, D and E), indicating that the ATDs of GluA4 are not occupied by endogenous NPs. Thus, CPTX may connect excitatory pre- and postsynaptic sites of spared neuronal circuits.

We next assessed whether CPTX induced excitatory synapses by quantifying the intensity and the size of VGluT2+, GluA4+, and CPTX+ puncta (Fig. 6C). CPTX signal was higher at GluA4+/VGluT2+-double-positive puncta than at single GluA4+ or VGluT2+ puncta, suggesting preferential localization of CPTX at synapses (fig. S20D). Furthermore, CPTX increased the percentage of VGluT2+ puncta that were also GluA4+ (Fig. 6D). Thus,

CPTX induced GluA4+ excitatory synapses within regions spared by the injury (fig. S20, A and B).

Finally, we assessed the recovery of locomotion after spinal cord hemisections. For comparison, we separately injected chondroitinase ABC (ChABC), an enzyme that promotes robust nerve regeneration (*52*), as well as Cbln1 to the lesioned spinal cord. During the acute stage, single injections of CPTX as well as ChABC, but not Cbln1, significantly restored the Basso mouse scale (BMS) score (*53*) (Fig. 6E and movie S5) and other locomotive parameters examined until 6 weeks after the injury (fig. S21A and movie S6). ChABC delivered 1 week after the hemisection was less effective (Fig. 6F), as reported previously (*54*). By contrast, when CPTX was delivered 1 week after injury, it was most effective (Fig. 6F, fig. S21B, and movies S7 and S8). CPTX also efficiently restored locomotion in a contusion SCI model, which is considered closer to pathophysiology that occurs commonly in humans (Fig. 6G and fig. S21C). In this paradigm, a combination of CPTX and ChABC treatment showed the most prominent recovery of locomotion, especially in the early stages after contusion. Immunoblot analyses revealed that by 7 days after injection, CPTX became undetectable in the spinal cord as well as in the hippocampus and the cerebellum (fig. S22). Thus, CPTX may directly enhance excitatory connectivity, which may be further stabilized by endogenous synaptic organizers, to promote sustained functional recovery in spinal cord injury models. Mechanistically, this process appears to be distinct from the one induced by ChABC treatment.

Discussion

Here, we developed a synthetic, structure-guided, synaptic organizer termed CPTX that can induce functional and structural synapses in the cerebellar, hippocampal. and spinal cord neuronal circuits in vivo. CPTX induced accumulation of AMPARs, as revealed by increased GluA1 intensity in all puncta and redistribution of GluA2/3 signals in the hippocampus 1 day after injection. CPTX application to the wild-type hippocampus led to increased mEPSC frequency in ~4 hours, and then both the frequency and the amplitude of mEPSCs ~3 days later. Thus, we speculate that CPTX first accumulates preexisting surface AMPARs at immature synapses with loose perisynaptic extracellular matrix (*55*), leading to redistribution of AMPARs and the increased mEPSC frequency. Considering that GluA1-containing AMPARs traffic to the surface in an activity-dependent manner (*56*), CPTX may stabilize newly inserted GluA1-containing AMPARs in mature synapses, leading to increased GluA1 surface amounts and mEPSC amplitudes. NP1 or $NP1_{PTX\text{-}3Cl}$, which could bind to AMPARs, failed to induce excitatory synapses in the hippocampus. Moreover, $NP1_{PTX\text{-}3Cl}$ could not restore PF-EPSCs in *Cbln1*-null mice. Thus, we propose that CPTX induces functional synapses by connecting presynaptic neurexin with postsynaptic AMPARs and facilitating transsynaptic alignment of AMPARs with the synaptic vesicle release sites. Such "molecular bridges" may initiate presynaptic maturation by inducing Nrx nanoclusters (*40*) and may eventually integrate within transsynaptic nanocolumns at excitatory synapses (*41*). Because molecular components are considerably different among neuronal synapse types and circuits, further studies are warranted to test this hypothesis and clarify the detailed molecular mechanisms by which CPTX restores functional and structural excitatory synapses under pathological conditions.

Potential adverse effects of CPTX could include inhibition of synapses where CPTX binds GluAs but does not have a presynaptic Nrx(+4) partner. However, injected CPTX did not replace endogenous NPs that target the ATD of GluA4. CPTX could also cause artificial hyper-connectivity, but we did not notice obvious abnormal behaviors, such as seizures or hyperactivities, in mice injected with CPTX into the hippocampus, the spinal cord or the cerebellum. Most likely this is because CPTX selectively strengthens synapses that contain Nrx(+4) at presynaptic sites. However, constitutive expression of Nrx3(+4) in genetically modified mice reduces mEPSC frequency and decreases LTP (*51*). Thus, the dose and the affinity of CPTX-like molecules should be optimized for future therapeutic use.

The catalog of ESPs continues to grow. For example, LGI1, an ESP produced by the hippocampal neurons, is particularly interesting because it indirectly recruits AMPARs by binding to pre- and postsynaptic ADAM23 and ADAM22, respectively (*57*, *58*). Owing to their modular structure, we anticipate that the toolkit of synthetic ESPs can be expanded to include a variety of pre- and postsynaptic specificities, with a range of affinities and enhanced resistance to proteolytic degradation in vivo. Such ESPs could be used to restore or modify synaptic connectivity in various neuronal circuits. The structure-guided design of ESPs that target specific $GABA_A$ receptor subtypes (*59*), as well as specific AMPAR or KAR subunits, will be useful to restore E/I balance in specific neuronal circuits affected by autism spectrum disorders and schizophrenia (*1*, *5*, *60*), epilepsy (*3*), and AD (*2*). Notably, although the effect of CPTX was transient in *Cbln1*-null ataxia mice, a single injection of CPTX greatly enhanced locomotion for at least 7 to 8 weeks in SCI models, long after injected CPTX was degraded. In the cerebellum, PF-PC synapses rely heavily on Cbln1 and GluD2 as synaptic organizers. Thus, once the injected CPTX is degraded, cerebellar synapses are rapidly lost and ataxic phenotypes return quickly. By contrast, in 5xFAD mice and SCI models, the new synapses formed upon CPTX injection might survive longer because other (endogenous) synaptic organizers could be recruited by increased neuronal activities. Thus, synthetic ESPs built upon the principles described here may become powerful tools to investigate and potentially cure disorders associated with impaired neuronal connectivity.

Materials and methods

Mice

Wild-type (Charles River Laboratories and Japan SLC) and *Cbln1*-null mice were used for primary culture. In experiments assessing cerebellar functions, we used male and female *Cbln1*-null mice with the C57BL/6J or ICR genetic backgrounds and *GluD2*-null mice (*13*) with the C57BL/6J genetic background. They were housed on a conventional 12 hours light/dark cycle (light on at 8:30 a.m.) and their behavior was tested in the light phase of the cycle. Mice expressing enhanced green fluorescent protein (EGFP) under the control of a modified Thy1 promoter region (The Jackson Laboratory, cat. no. 007788) were used for immunohistochemistry. In experiments assessing hippocampal functions, we used male C57BL/6J, PV-*Cre*/Ai9 double transgenic mice (The Jackson Laboratory, cat. no. 017320 and 007909), 5xFAD mice (The Jackson Laboratory, cat. no. 006554) which carry five familial AD mutations, and their wild-type littermates with C57BL/6J genetic background (Charles River Laboratories). At least one week before starting the experiments, mice were transferred from the main animal facility of DZNE (Magdeburg, Germany) to a small vivarium, where they were housed individually with food and water ad libitum before food restriction diet (FRD) was given, on a reversed 12 hours light/dark cycle (light on at 9:00 p.m.). All behavioral experiments assessing hippocampal functions were performed in the late afternoons during the dark phase of the cycle when mice were active, under constant temperature (22° ± 1°C) and humidity (55 ± 5%). For the spinal cord injury (SCI) model, wild-type mice with the ICR genetic background (Japan SLC) were used. They were housed on a conventional 12 hours light/dark cycle. Locomotor functions were assessed at a fixed time in the evening. All treatments and behavioral procedures were conducted in accordance with ethical animal research standards defined by German and Japanese law and approved by the Ethical Committee on Animal Health and Care of the State of Saxony-Anhalt, Germany, with license numbers 42502-2-1343 DZNE and 42502-2-1322 DZNE, by the Animal Resource Committee of Keio University (no. 09050) and by the Institutional Animal Care and Use Committee of Aichi Medical University (no. 1559).

Plasmids

For surface plasmon resonance (SPR) experiments, cDNAs encoding the extracellular human glutamate receptor D2 amino-terminal domain (ATD; UniProt ID O43424; GluD2 ATD: Asp24-Gly440), the mouse glutamate receptor A1 ATD (UniProt ID P23818; GluA1 ATD: Ala19-Thr412), the mouse glutamate receptor A2 ATD (UniProt ID P23819; GluA2 ATD: Val22-Thr419), the mouse glutamate receptor A3 ATD (UniProt ID Q9Z2W9; GluA3 ATD: Gly23-Thr422), the mouse glutamate receptor A4 ATD (UniProt ID Q9Z2W8; GluA4 ATD: Gly21-Thr420), the human β-neurexin-1 LNS6 domain including sliced sequence no. 4 (UniProt ID P58400; β-Nrx1 LNS6(+SS4): His86-Val295) and the pentraxin domain of human neuronal pentraxin-1 (UniProt ID Q15818; $NP1_{PTX}$: Pro224-Ile431) were cloned into the pHLsec-Avitag3 expression vector (Addgene cat. no. 99847) (*61*). To produce a trimeric form of the pentraxin domain, $NP1_{PTX}$ was fused to the C terminus of a three-stranded GCN4 leucine zipper coiled-coil sequence (*62*) to yield $NP1_{PTX\text{-}3Cl}$ (fig. S1, A and B). To produce CPTX, $NP1_{PTX\text{-}3Cl}$ was fused to the C terminus of the cysteine-rich region (CRR) (*15*) of human Cbln1 (UniProt ID P23435; Gln22-Ile53) (fig. S1, A and B). For the in vitro binding and synapse formation assays, cDNA encoding the (i) full-length mouse NP1 (UniProt ID Q62443) was cloned into the pCAGGS vector (a gift from J. Miyazaki, Osaka University, Osaka, Japan) with a C-terminal hemagglutinin (HA)–tag (fig. S4D), (ii) full-length mouse β-Nrx1 (UniProt ID P0DI97; ±SS4 –SS5), full-length rat β-Nrx2 (UniProt ID Q63376; ±SS4 –SS5) with the signal sequence of mouse neuroligin-1 (UniProt ID Q99K10; Met1-Lys47), full-length mouse β-Nrx3 (UniProt ID Q8C985; ±SS4 +SS5), and full-length rat α-Nrx1 (UniProt ID Q63372; –SS1,2,4,6 +SS3,5 or –SS1,2,5,6 +SS3,4; a gift from P. Scheiffele, University of Basel, Switzerland) was cloned into pCAGGS with a C-terminal FLAG-tag (figs. S3B and S6A), (iii) mouse β-Nrx1 extracellular domain (UniProt ID P0DI97; Met1-Thr392, +SS4 –SS5) was cloned into the pCAGGS vector with a C-terminal human Fc tag (fig. S3C), and (iv) Myc-tagged GluA1–4 ATDs were cloned into a modified pDisplay vector (Invitrogen) as described previously (*29*) (figs. S3, A and C, and S4, A to D). The 21-nucleotide microRNA RNA interference sequences targeting mouse *Nptx1* and *Nptxr* mRNA were designed by BLOCK-iT (Invitrogen) and inserted upstream of IRES-GFP in the pCAGGS vector as described previously (*63*) (sequences: 5′-AGA CAA GTT TCA GCT GAC ATT-3′ for NP1 and 5′-TGC TCA GTC GCT TCC TCT GTA-3′ for NPR). For immunocytochemical analyses to check the selectivity of anti-NP1 and anti-NPR antibodies (fig. S15, C and D), full-length mouse NP1 (UniProt ID Q62443), full-length mouse NP2 (GenBank ID O70340) and full-length mouse NPR (UniProt ID Q99J85) were cloned into the pCAGGS vector with a C-terminal HA-tag. For the immunoblotting analyses to check the selectivity of anti-NP1 and anti-NPR antibodies (fig. S15, A and B), mouse NP1 (without the signal sequence: ΔMet1-Ala22), NP2 (without signal sequence: ΔMet1-Ala14) and NPR (without the transmembrane domain, ΔMet1-Ile23) were cloned into pCAGGS downstream of the Igκ secretion signal sequence, followed by a 2xHA tag.

Production of recombinant proteins

For large-scale production, proteins were expressed using established procedures for transient transfection (*61*) or lentiviral transduction (*64*) of HEK293 cells. Briefly, for large-scale transient transfections, the pHLsec expression plasmids encoding the proteins of interest ($NP1_{PTX}$, Nrx1β(+4), Nrx1β(–4), $GluA1_{ATD}$, $GluA2_{ATD}$, $GluA3_{ATD}$, $GluA4_{ATD}$, and $GluD2_{ATD}$) were transfected into 3-liter adherent HEK293T (ATCC cat. no. CRL-1573) expression cell cultures. The supernatants were collected 5 days after transfection and filtered (0.22 μm). For the lentiviral infections, the pHR-CMV-$TetO_2$ transfer plasmid (Addgene cat. no. 113887) encoding the protein of interest (CPTX), the psPAX2 packaging plasmid and the pMD2.G envelope plasmid were cotransfected into the HEK293T Lenti-X producer cell line (Takara Bio cat. no. 632180). The supernatant containing the viral particles was harvested after 3 days, filtered (0.45 μm) and directly used for the infection of HEK293T expression cells. The infected cells were grown and expanded to 3-L adherent cell cultures. The supernatants were collected 5 days after infection and filtered (0.22 μm). All supernatants were concentrated and buffer-exchanged using the AKTA flux system (GE Healthcare) or a QuixStand benchtop diafiltration system (GE Healthcare). Proteins were first purified by immobilized metal-affinity chromatography (IMAC) using pre-packed nickel Sepharose columns (GE Healthcare), or in batch mode on TALON beads (Takara Bio) collected on an EconoColumn (Bio-Rad). Then, elution fractions were pooled after SDS–polyacrylamide gel electrophoresis (PAGE) analysis and concentrated, followed by further purification by size exclusion chromatography (SEC) using a Superdex 200 16/600 column (GE Healthcare) in 10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (pH 7.4), 150 mM sodium chloride and 3 mM calcium chloride (HBS-C) for structural studies ($NP1_{PTX}$), or in 10 mM HEPES (pH 7.4), 150 mM sodium chloride (HBS; vehicle) for biological studies (CPTX), or in 10 mM Tris (pH 7.4), 150 mM sodium chloride, 3 mM calcium chloride and 0.005% (v/v) Tween-20 (TBS-CT) for interaction studies using SPR (Nrx1β(+4), Nrx1β(–4), $GluA1_{ATD}$, $GluA2_{ATD}$, $GluA3_{ATD}$, $GluA4_{ATD}$, and $GluD2_{ATD}$).

Protein crystallization

Crystallization trials, using 100 nl protein solution plus 100 nl reservoir solution in sitting drop vapor diffusion format, were set up in 96-well Greiner plates using a Cartesian Technologies robot (*65*). Purified human $NP1_{PTX}$ (Pro224-Ile431), produced from HEK293T cells and concentrated to 6.5 g/liter, crystallized in 15% (v/v) glycerol, 25.5% (w/v) polyethylene glycol (PEG) 8000, 0.17M ammonium sulfate and 0.085 M sodium cacodylate (pH 6.5). Crystals were cryoprotected using a reservoir solution containing 30% (v/v) ethylene glycol.

Crystallographic data collection and structure determination

Diffraction data for $NP1_{PTX}$ were collected at Diamond Light Source (DLS) beamline I03 to a nominal resolution of 1.45 Å in space group (SG) *C*2. All data were indexed, integrated, and scaled using the automated XIA2 expert system (*66*), using the Labelit (*67*), POINTLESS and AIMLESS (*68*, *69*), and XDS (*70*) software. The structure of $NP1_{PTX}$ was solved by molecular replacement using Phaser (*71*) and using the human SAP (Serum Amyloid P Component; PDB code 1SAC) crystal structure (*42*) as a search model. Crystallographic data collection and refinement statistics are presented in table S1.

Structure refinement and model analysis

Maximum-likelihood refinement of $NP1_{PTX}$ was initially performed with Refmac5 using "jelly body" restraints (*72*), and finally with phenix.refine (*73*) using automated X-ray and atomic displacement parameter (ADP) weight optimization applied throughout. Automated model building was performed using ARP/wARP (*74*) and further manual model building was performed using Coot (*75*). Structure validation was performed using the MolProbity routines within the PHENIX software suite (*73*, *76*). Molecular representations were made using PyMOL (*77*).

Multi-angle light scattering (MALS)

Protein samples concentrated to ~1.0 g/liter were injected into an HPLC-driven SEC column (Superdex 200 10/30 column, GE Healthcare) equilibrated with HBS-C buffer. The SEC column was coupled to an online UV detector (Shimadzu), an 18-angle light scattering detector (DAWN HELEOS; Wyatt Technology), and a refractive index detector (Optilab T-rEX; Wyatt Technology). CPTX contains an N-linked sugar on the Cbln1 N-terminal cysteine-rich region (CRR), and molecular mass determination was performed using an adapted RI increment value (*dn/dc* standard value; 0.185 ml/g) to account for the glycosylation state. $NP1_{PTX}$ and $NP1_{PTX\text{-}3Cl}$ are not glycosylated hence the standard *dn/dc* value was used for molecular mass determination. Data

analysis was carried out using the ASTRA V software (Wyatt Technology).

Surface plasmon resonance (SPR)

cDNA for the immobilized proteins (CPTX, $NP1_{PTX}$, Cbln1) was cloned into the pHLsec-Avitag3 vector (*61*), resulting in proteins carrying a C-terminal biotin ligase (BirA) recognition sequence (Avitag). Constructs were cotransfected with pDisplay-BirA-ER (Addgene cat. no. 20856; coding for an ER-resident biotin ligase) (*78*) for in vivo biotinylation in HEK293T cells in small-scale six-well plates in a 3:1 pHLsec:pDisplay stoichiometric ratio. A concentration of 100 μM D-biotin was maintained in the expression medium to ensure biotinylation of the Avitag. After 48 hours of expression, conditioned medium was collected and dialyzed against 10 mM Tris (pH 7.4), 150 mM sodium chloride, 3 mM calcium chloride and 0.005% (v/v) Tween-20 (TBS-CT). SPR experiments were performed on a Biacore T200 (GE Healthcare) operated at a data collection frequency of 10 Hz. Streptavidin (Sigma-Aldrich cat. no. S4762) was chemically coupled via amine coupling chemistry onto CM5 chips to a response unit (RU) level of 5000 RU. Then, biotinylated proteins were captured to the desired RU level. In each instance, for every two analyte binding cycles, a buffer injection was performed, allowing for double referencing of the binding responses (*79*).

Antibodies

The origin, dilution, company and catalog number are as follows: Anti-calbindin (Goat, 1:500, Frontier Institute, Af1040), anti-HIS (Mouse, 1:1000, MBL, D291-3 or Rabbit, 1:1000, CST, 2365), anti-Myc (Rabbit, 1:1000, MBL, 562), anti-FLAG (Rabbit, 1:1000, Sigma-Aldrich, F7425), anti-HA (Mouse, 1:1000, BAbCo, MMS-101P), anti-synaptophysin (Mouse, 1:500, Sigma-Aldrich, S5768 or Guinea pig, 1:500, Frontier Institute, Af300), anti-neurexin (Chicken, 1:500, a gift from P. Scheiffele, University of Basel, Basel, Switzerland), anti-GFP (Rabbit 1:1000, Frontier Institute, Af-2020), anti-MAP2 (Goat, 1:500, Frontier Science, Af860), anti-PSD95 (Mouse, 1:500, Invitrogen, MA1-046), anti-pan-AMPAR (Guinea pig, 1:500, Frontier Institute, Af580), anti-GluA1 (Rabbit, 1:100, Calbiochem, PC246 or Guinea Pig, 1:500, Frontier Institute, Af380), anti-GluA2/3 (Rabbit, 1:1000, Chemicon, AB1506), anti-GluA4 (Rabbit, 1:1000, Pharmingen, 60666N or Guinea Pig, 1:500, Frontier Institute, Af640), anti-GluN1 (Mouse, 1:500, BD Biosciences, 556308 or Mouse, 1:500, Chemicon, MAB363), anti-VGluT1 (Rabbit, 1:500, Frontier Institute, Af570 or Goat, 1:500, Frontier Institute, Af310), anti-VGluT2 (Guinea Pig, 1:500, Frontier Institute, Af720 or Goat, 1:500, Frontier Institute, Af310), anti-VGAT (Goat, 1:500, Frontier Institute, Af620 or Guinea pig, 1:500, Frontier Institute, Af1000), anti-Parvalbumin (Goat, 1:500, Frontier Institute, Af460), anti-Homer 1 (Guinea pig, 1:1000, Synaptic Systems, 160-004) and anti-Bassoon (Rabbit, 1:500, Synaptic Systems, 141-003 or Mouse, 1:500, Enzo Life Sciences, SAP7F407). Anti-NP1 and NPR antibodies were newly raised by guinea pigs and rabbits immunized with peptides Glu185-Lys227 (NP1) or Ser226-Glu263 (NPR). For the simultaneous detection of various GluAs (GluA1–3 or GluA1–4), the mixture of the primary antibodies described above was used. Secondary antibodies conjugated with DyLight 405, Alexa 488, 546, 647 and Cy3 (Invitrogen or Jackson ImmunoResearch Laboratories) against the respective primary antibody were used in the dilution 1:1000 for immunocytochemistry or 1:200 for immunohistochemistry. Isospecific secondary antibodies against mouse IgG1 or IgG2a were used for the simultaneous staining of two types of mouse primary antibodies.

Immunoblotting

For fig. S15, A and B, the hippocampus and conditioned medium were solubilized in laminae buffer (2% SDS, 80 mM Tris-HCl (pH 6.8), 10% glycerol, 0.00625% Coomassie blue G 250) and the proteins were reduced by boiling for 3 min using 2% 2-mercaptoethanol. For fig. S22A, brain tissue from the hippocampus, cerebellum (vermis) and spinal cord were solubilized by 190 to 250 μl 1% SDS buffer, sonicated, and quantified for protein concentration. Samples were diluted in laminae buffer and reduced by boiling for 20 min using 2% 2-mercaptoethanol. Samples were subjected to gradient gel SDS-PAGE (Fujifilm Wako Chemicals) and blotted onto the membrane (Immobilon-P PVDF membranes, Merck Millipore). Membranes were blocked with TS-tween (0.1% Tween-20, 50 mM Tris-HCl pH7.6, 150 mM NaCl) containing 5% skimmed milk (CP-Meiji) and subjected to the primary antibody for 2 hours and HRP-conjugated secondary antibody (GE Healthcare) for 30 min. Chemiluminescence was generated by the ImmunoStar (Fujifilm Wako Chemicals) or Immobilon (Merck Millipore) detection kits and detected by the LAS-3000 mini (Fujifilm) or iBright systems (Thermo Fisher Scientific).

Cell-based in vitro binding assays

HEK293 tsA201 cells (a gift from R. Horn, Thomas Jefferson University, Philadelphia, USA) were cultured in high glucose DMEM (Sigma-Aldrich) containing 10% FBS (HyClone), 50 U/ml Penicillin, 50 mg/ml Streptomycin (Invitrogen), 2 mM L-Glutamine in 10% CO_2 at 37°C. Cells were transfected with pDisplay encoding myc-tagged GluA1–4 ATD or with pCAGGS encoding GFP or FLAG-tagged Nrx, using Lipofectamin 2000 (Invitrogen). On the following day, transfected cells were detached by phosphate-buffered saline (PBS) containing 5 mM EDTA and seeded on 12-mm coverslips coated with poly-L-lysine (PLL; Sigma-Aldrich) at 2×10^4 cells/well. One hour after seeding, cells were treated with Mock (vehicle; HBS buffer), recombinant Cbln1-HIS or CPTX-HIS (23.6 nM, calculated for the hexamers) for 4 hours to overnight. For the in vitro tripartite binding assay (fig. S3C), cells were further incubated with the conditioned medium containing Nrx1β(+4)-hFc for 4 hours. Cells were fixed with 4% PFA/PBS for 15 min and washed with PBS three times. After blocking with 3% BSA/PBS (BSA: Sigma-Aldrich) for 30 min without permeabilization, cells were further incubated with primary antibodies against the ligand tag (HIS, HA and/or hFc) in 3% BSA/PBS for 2 hours at room temperature (RT) or 24 hours at 4°C. Following three washes with PBS, cells were permeabilized with PBS containing 0.1% Triton-X 100 (Sigma-Aldrich) and 3% BSA. Cells were stained with primary antibodies against the receptor tag (FLAG or Myc) for 2 hours at RT or 24 hours at 4°C, followed by washing with PBS and incubation with the respective secondary antibodies for 30 min. After washing with PBS, coverslips were mounted on a glass slide with Fluoromount-G. Samples in which HEK293 cells expressed GFP were not stained by primary antibodies against GFP but treated with the same primary antibodies against the receptor tag (FLAG or Myc) and the fluorescence of GFP was detected by microscopy.

In vitro synapse formation assays

Hippocampal culture

Preparation from embryonic day 16-17 (E16-17) wild-type ICR mice was performed as described (*13*) with some modifications. Neurons (2×10^5) were plated on 12-mm cover glasses coated with PLL and maintained in Neurobasal cell culture medium (Invitrogen) containing NS21 supplement (*80*), 50 U/ml Penicillin, 50 mg/ml Streptomycin (Invitrogen), 2 mM L-Glutamine and 2% FBS (HyClone) in 5% CO_2 at 37°C. One to 2 hours after the initial incubation, the culture medium was changed to fresh medium without FBS. For induction of presynapses by HEK293 cells (fig. S4), neurons were co-cultured with 2×10^4 HEK293 cells expressing GFP or Myc-tagged GluA1–4 or GluD2 ATDs at day-in-vitro (DIV) 4, 5 or 7 for 1 hour, followed by treatment with vehicle or recombinant Cbln1-HIS or CPTX-HIS (23.6 to 30 nM, calculated for the hexamers), or with conditioned medium containing NP1-HA or CPTX-HIS secreted from HEK293 cells transfected with those plasmids for 2 days. For induction of AMPARs by HEK293 cells (fig. S6A), neurons were co-cultured with HEK293 cells expressing Nrx-3xFLAG and GFP, Cbln1-HIS or CPTX-HIS at DIV9 for 2 days. For synapse induction by beads, neurons were treated with streptavidin Dynabeads M-280 (Thermo Fisher

Scientific) coated with Cbln1-HIS or CPTX-HIS through anti-HIS mAb-biotin (MBL) for 1 to 2 days (DIV20–22 for Syp, Nrx and GluA1–3 in Fig. 2B and figs. S5A and S6B, DIV4–6 for VGluT1 in fig. S5B, and DIV30–32 for GluA4 in fig. S6C). To label axons and dendrites, GFP was transfected with Lipofectamine 2000 at DIV6. Following the incubations, the cells were fixed with 4% PFA/PBS for 5 min at RT. Cells were washed with PBS and fixed with chilled methanol for 8 min on ice, followed by washing with PBS. Cells were permeabilized and blocked with PBS containing 0.1% Triton-X 100, 3% BSA and stained with primary antibodies diluted in blocking buffer for 2 hours at RT or 24 hours at 4°C, followed by washing with PBS and incubation with secondary antibodies diluted in blocking buffer for 30 min. Following washing with PBS, coverslips were mounted on a glass slide with Fluoromount-G (SouthernBiotech).

Cerebellar culture

Cerebellar cultures were prepared from postnatal day 0 to 3 (P0 to P3) *Cbln1*-null mice (on the ICR genetic background) as described (*13*) with minor modifications. Neurons (2×10^5) were plated on 12 mm cover glasses coated with PLL and were maintained in DMEM/F12 (Sigma-Aldrich) containing 100 μM putrescine, 30 nM sodium selenite, 0.5 g/ml tri-iodothyronine, 0.25 mg/ml BSA, 3.9 mM glutamate and N3 supplement (100 μg/ml apo-transferrin, 10 μg/ml insulin and 20 nM progesterone) and 10% FBS in 5% CO_2 at 37°C. One to 2 hours after incubation, the culture medium was changed to fresh medium without FBS. Co-culturing with HEK293 cells (at DIV 4 or 5 for 2 days), fixation, permeabilization and staining were performed as described above for the hippocampal culture.

Quantification of signal intensities of in vitro assays

Fluorescence images of cultured samples were captured using a CCD camera attached to a conventional fluorescence microscope (BX63, Olympus) or a confocal laser scanning microscope (FV1000, Olympus) using a 40× objective for the detection of HEK293 cells or using a 60× objective with 5× digital zoom for the detection of beads. The same parameters, such as exposure time and excitation power, were used for the sample set to be compared, except for GFP signals. Analyses were performed semi-automatically by a customized macro using Image J (Fiji). Background subtraction using the rolling ball method (radius = 20 to 100 pixels (px)) was applied to images to reduce the background noise. Regions of interest (ROIs) of beads or HEK293 cells expressing receptors were determined by the signal intensity and area size fulfilling the threshold (Otsu or Triangle) and particle definition. For quantification of signals on HEK293 cells, the signal intensity within the ROI was determined by subtracting the mean signal intensity outside the ROI. For quantification of signals on beads and to minimize non-specific signals associated with the cross-reactivity of beads with antibodies, the signal intensity was determined by subtracting the mean intensity of beads located in the field where no cells were present (Fig. 2B and figs. S5, A and B, and S6, B and C). For fig. S6C, ROIs of beads were manually classified into two groups, those contacting and non-contacting with PV^+ interneurons, which were identified by PV and GluA4 immunoreactivities. The signal intensity was calculated by subtracting the intensity on non-contacting beads from that on contacting beads.

Immunocytochemistry and immunohistochemistry

Hippocampal culture

Dissociated hippocampal neurons were prepared from E17 embryos, infected by AAV CAG-GFP at DIV12 to express GFP (Fig. 2C and fig. S7B) and cultured for 21 DIV. They were incubated overnight (between DIV20 and 21) with 7 nM CPTX added directly to the culture medium. Neurons were either fixed with 4% PFA, permeabilized with 0.1% Triton-X 100 in PBS for 10 min, washed three times and blocked (0.1% Glycine + 0.1% Tween-20 + 10% normal goat serum in PBS for fig. S7A) for 45 min, or fixed with 4% PFA, fixed with chilled methanol, permeabilized and blocked (2% Goat serum 2% BSA + 0.1% Triton-X 100 in PBS for Fig. 2C and fig. S7B) for 30 min at RT. Cells were incubated 1 hour or overnight with primary antibodies diluted in the blocking buffer and washed with PBS three times. Cells were incubated with respective secondary antibodies for 1 hour at RT. Coverslips were washed with PBS, incubated with or without DAPI (Invitrogen, cat. no. D1306) for 10 min and mounted using Fluoromount aqueous mounting media (Sigma Aldrich, cat. no. F4680) or Fluoromount-G (SouthernBiotech). For the acquisition of the images in fig. S7A, a complete Z stack spanning primary hippocampal cultures was acquired on a Leica TCS SP8 3X microscope equipped with a 405 nm diode laser and a white light laser (WLL) for excitation. The format for all images was set to 1024 × 1024 pixels with an optical zoom of 2× resulting in a pixel size of 90 nm in the x-y plane. Images were taken at 400 lines/s and line averaging of 2. To improve image quality for representation, raw confocal images were deconvoluted using the Huygens Professional software (SVI, v15.10). The theoretical point spread function (PSF) parameters were automatically acquired from confocal LIF-files. Within the deconvolution wizard, images were subjected to manual background correction and signal-to-noise ratio optimization. The "optimized iteration" mode of the Good's roughness Maximum Likelihood Estimation algorithm (GMLE) was applied until it reached a quality threshold of 0.01. For the representation of the images, the contrast was enhanced for the complete image by linear methods using Fiji for each channel separately. For the acquisition of the images in Fig. 2C and fig. S7B, a confocal laser scanning microscope (FV1000, Olympus) was used to acquire single-plane images around GFP-positive dendrites without the cell body, using a 100× objectives with 4× optical zoom.

Hippocampal sections

For fig. S10A, Thy1-GFP mice were transcardially perfused 3 days after CPTX injection with 4% PFA in 0.1 M sodium phosphate buffer (PB; pH 7.2) for 10 min under deep isoflurane anesthesia. The dissected brains were incubated in PBS containing 4% PFA, cryoprotected in PBS containing 30% sucrose for 48 hours, frozen at −80°C and cut into 50 μm thick sections by cryostat (Leica). Membrane permeabilization was performed with 0.5% Triton-X 100 in PBS for 30 min. Sections were blocked with blocking buffer (0.1% Glycine + 0.1% Tween-20 + 10% normal goat serum in PBS) for 90 min. Above mentioned antibodies were used for the detection of Bassoon (1:500) and HIS (1:500), diluted in blocking buffer and applied for 2 days at 4°C. Sections were washed three times with PBS followed by incubation with the respective secondary antibodies (1:1000) for 3 hours at RT. Sections were mounted using Fluoromount aqueous mounting media (Sigma Aldrich, cat. no. F4680). Complete Z stacks spanning 50 μm thin hippocampal slices were acquired on a Leica TCS SP8 3X microscope. For acquiring images, the Leica objective HC PL APO CS2 63×/1.40 oil was used. Bassoon (405 nm), Thy1-GFP (488 nm), CPTX (546 nm) images were acquired using the confocal scanning mode. The format for all images was set to 512 × 512 pixels with an optical zoom of 3.04 resulting in a pixel size of 119 nm in the x-y plane. Images were taken at 400 lines/s and line averaging of 4.

For figs. S10B, S11, S12, and S17, 1 day after the unilateral hippocampal injection of CPTX into WT mice, and for fig. S18, 3 days after the unilateral hippocampal injection of CPTX into 5xFAD mice (11 to 12 months old), the brain was freshly frozen with liquid nitrogen after decapitation under deep isoflurane anesthesia. Sections (10 to 20 μm thick) were cut by cryostat (Leica), mounted on glass slides or coverslips and postfixed with 4% PFA in sodium phosphate buffer (PB; pH 7.2) or 3% glyoxal for 15 min to 2 hours at RT. Following washing with PBS + 0.1% Triton X-100, sections were treated with 10% donkey serum for 30 min at RT, incubated with a mixture of primary antibodies overnight and with a mixture of respective secondary antibodies for 2 hours. Finally, the sections were attached

to the glass slides and mounted with Fluoromount-G, VECTASHIELD (Vector Laboratories) or ProLong glass (Invitrogen). Fluorescence images were captured using confocal microscopy (SD-OSR, Olympus or LSM880, ZEISS) using the same parameters such as objectives, zoom, laser power, exposure time, gain and offset for the comparison of samples. Super-resolution images were obtained by structural illumination with SD-OSR (Objective: 100×/1.49 oil with 3× zoom lens, pixel size: 40 nm/px, configuration: XY < 120 nm, Z < 300 nm) or by Airyscan with LSM880 (63×/1.40 oil with 3× digital zoom, 35 nm/px, XY < 120 nm, Z < 350 nm). Images obtained by Airyscan were 3D-reconstructed by Imaris (Bitplane).

For the quantification of pre- and postsynaptic molecule clustering (figs. S12, S17, and S18), images were obtained using a 100× objective around the area of the injected side of the hippocampus (stratum radiatum or stratum lacunosum moleculare) and the corresponding area of the contralateral uninjected or control-injected side. Injected areas were confirmed by the HIS immunoreactivity. The dynamic range of signals was first determined by maximal and minimal signal intensities in pair images of the injected and the contralateral area and then each signal intensity was normalized to this value (max-min). Normalized intensity histograms were calculated by dividing the normalized signal intensity into 40 bins on a logarithmic scale. Thus, the signal intensity in the n^{th} bin equals $10^{(n \times (\text{max-min}) / 40)}$. To analyze the bright pixels, the mean intensity + 3 SD of the contralateral area was used as the threshold. Selected pixels were analyzed as particles using the "Analyze Particles" command within Image J without any constraint on size or circularity, leading to the calculation of particle size, number, area fraction and intensity.

Cerebellar sections

One day after the cerebellar injection of vehicle or CPTX into *Cbln1*-null mice or *GluD2*-null mice (3 or 7 months old), they were transcardially perfused with 4% PFA in 0.1 M sodium phosphate buffer (PB; pH 7.2) for 10 min under deep pentobarbital anesthesia, followed by post-fixation of dissected brain samples with 4% PFA in PB for 2 hours at 4°C. Alternatively, the brain was freshly frozen with liquid nitrogen after decapitation under deep isoflurane anesthesia. Free-floating sections (50 μm thickness) from 4% PFA fixed brains were prepared with a microslicer (DTK-1000, Dosaka EM), followed by treatment with 1 mg/ml pepsin (Dako) in PBS containing 0.1% Triton X-100 and 0.2 N HCl for 10 min at 37°C, washing of the sections with PBS containing 0.1% Triton X-100 and blocking with 10% donkey serum for 30 min at RT. Sections of freshly frozen brains were prepared by cryostat (Leica) at 20 μm thickness, mounted on glass slides and fixed with 3% glyoxal for 15 min, followed by blocking with 10% donkey serum for 30 min at RT and permeabilization with PBS containing 0.1% Triton X-100 (wash buffer). Sections were incubated with a mixture of primary antibodies diluted in wash buffer overnight, washed three times, incubated with a mixture of respective secondary antibodies in wash buffer and again washed three times. Finally, the sections were mounted with VECTASHIELD (Vector Laboratories). Fluorescence images were captured using a confocal laser scanning microscope (FV1000, Olympus) using the same parameters such as objectives, zoom, laser power, gain and offset for the comparison of samples. Colocalization of signals was analyzed by the plugin "Coloc 2" in Image J (*81*).

Cortical sections

One-month-old mice, into which miR vectors were transfected by in utero electroporation, were transcardially perfused with 3% glyoxal for 10 min under deep pentobarbital anesthesia, followed by post-fixation of the dissected brain with 3% glyoxal for 2 hours at 4°C. Free-floating sections (100 μm thickness) were prepared with a microslicer (DTK-1000, Dosaka EM). Following washing of the sections with PBS containing 0.1% Triton X-100 (wash buffer), sections were subsequently treated with 10% donkey serum for 30 min at RT, incubated with a mixture of primary antibodies diluted in wash buffer overnight, washed three times, incubated with a mixture of respective secondary antibodies in wash buffer and again washed three times. Finally, the sections were attached to glass slides and mounted with VECTASHIELD (Vector Laboratories). Z-stacked fluorescence images of NP1 and NPR were obtained throughout the brain slice in the transfected cortical area at every 2 μm using a confocal microscope (SD-OSR, Olympus) using 20× objectives. Soma area regions of interest (ROIs) in layers 2/3 with abundant expression of NP1/NPR were found in a large population of neurons and were manually selected in an unbiased manner according to the GFP signal. Signal intensities of NP1/NPR were quantified using Image J. Background signals were subtracted using the rolling ball method (radius = 50 px) and the signals were normalized to the surrounding signals enlarged to 5 px because the absolute signal intensity was strongly affected by the z-position of the ROI in the brain section at 100 μm thickness.

Spinal cord sections

Two to 5 days after the injection the mice were transcardially perfused with 3% glyoxal for 10 min under deep pentobarbital anesthesia, followed by post-fixation of the dissected spinal cord with 3% glyoxal overnight at 4°C and by cryoprotection in 30% sucrose/PBS for several days. Spinal cords embedded in Tissue-Tek O.C.T. compound (Sakura Finetek, cat. no. 4583) were horizontally or coronally sectioned by cryostat (Leica) at 20 to 40 μm thickness in 4 to 5 mm around the epicenter of the injury and the slices were mounted on APS-coated glass slides (Matsunami) at 200 to 400 μm intervals. Following washing with PBS + 0.1% Triton X-100, sections were treated with 10% donkey serum for 30 min at RT and incubated with a mixture of primary and respective secondary antibodies overnight. Finally, the sections were attached to glass slides and mounted with Fluoromount-G (SouthernBiotech). Fluorescence images of all sections were captured and digitized as a virtual slide (OlyVIA, Olympus) using a conventional fluorescence microscope (BX63, Olympus) with a 4× objective to determine the section at 1.4 to 1.6 mm upstream of the precise epicenter of the injury. Fluorescence images of the section were subsequently captured using a confocal microscope (SD-OSR, Olympus; 63× objective) around the gray matter ventral root using the same parameters such as objectives, laser power, exposure time, gain and offset for the comparison of samples. The super-resolution images were obtained by Airyscan 2 with a confocal laser scanning microscope (LSM980, ZEISS; 63×/1.40 oil objective with 2.5x digital zoom, 35 nm/px, XY < 120 nm, Z < 350 nm). To define the particles of GluA4, VGluT2 and their intersections, the fluorescent images were subjected to background subtraction (50 px), Laplace filtering (9 × 9), box filtering (5 × 5), auto thresholding (Otsu method), particle extraction (> 5 px), binarization and segmentation by watershed. The defined particles were used as ROI to quantify the size and mean intensity of each channel. The percentage of GluA4+/VGluT2+-double-positive puncta was calculated as the number of VGluT2 particles accompanied by more than 1 px of defined GluA4 particles in the ROI, divided by the total number VGluT2 particles. All image processing was performed by ImageJ. For fig. S16, D and E, slices of the spinal cord and hippocampus from the same mouse were mounted on glass slides, stained using the same procedures and the images were taken using the same microscope settings.

Golgi-Cox staining and dendritic spine analysis

WT and 5xFAD mice (11 to 12 months old) were sacrificed and decapitated 3 days after CPTX or Mock injection. Brains were quickly removed from the skull and washed with bi-distilled water to remove blood from the surface. Golgi-Cox impregnation of neurons was performed using the FD Rapid GolgiStain kit (FD NeuroTechnologies, cat. no. PK401). Dye-impregnated brains were rapidly frozen on methyl butane and dry ice, embedded in Tissue-Tek O.C.T. compound (Sakura Finetek, cat. no. 4583), cryo-sectioned coronally at 100 μm

thickness and directly mounted on gelatin-coated slides (FD NeuroTechnologies, cat. no. PO101) with the help of solution C provided in the kit. Sections were stained according to the manufacturer's protocol and mounted using the ROTI Histokitt embedding medium (Carl Roth, cat. no. 6638). Secondary apical dendrites from CA1 pyramidal neurons with cell bodies located in the dorsal part of the hippocampus were imaged for spine analysis. Sections were imaged by an experimenter blinded to experimental conditions using a Leitz DMRXE microscope (Leica) as Z stacks with 0.25 μm intervals and at a magnification of 100×. 7-8 dendrites of 60 to 80 μm length were imaged for each animal. The analysis was done by an experimenter blinded to experimental conditions using the Neurolucida software v11 (MBF Bioscience). Dendritic spines were defined as the protrusions greater than 0.4 μm that emerged from the dendrites and were identified going through the focal planes of the Z stacks. To facilitate visual inspection of the Z-minimal projections in Fig. 4A and of the 3D images in Movie S4, the background was removed using the rolling ball method (radius = 30 px) and images were corrected using an unsharp mask filter (radius = 2 px, mask weight = 0.7) in Fiji (*81*). Then, 3D projections (made using the 3D projection tool in Fiji) were additionally processed using the rolling ball method (radius = 15 px), an unsharp mask filter (radius = 1 px, mask weight = 0.6) and contrasted to remove a remaining blur around dendritic shafts and enhance the visibility of small protrusions.

Electron microscopy

For the in vivo Cbln1 and CPTX administration experiments, *Cbln1*-null or *GluD2*-null mice (C57BL/6J genetic background, 3 to 6 months old) were perfused transcardially with 2% PFA + 2% glutaraldehyde in 0.1 M PB (pH 7.2) under deep pentobarbital anesthesia 3 days after the injection into the cerebellum. Parasagittal microslicer sections of the cerebellum (300 μm thickness) were postfixed for 2 hours with 1% OsO_4 in 0.1 M PB. After block staining in 1% aqueous uranyl acetate solution and dehydration with graded alcohols, the sections were embedded in Epon 812. Serial ultrathin sections (70 nm) were made using an ultramicrotome and stained with 2% uranyl acetate for 5 min and mixed lead solution for 2 min. Electron micrographs of the molecular layer were taken using an H-7100 electron microscope (Hitachi) at a magnification of 4,000× and printed at a magnification of 16,000×. For the quantitative analysis of PF–Purkinje cell synapses, the asymmetrical postsynaptic density in the cerebellar cortex was randomly photographed and the frequency of contacted and naked/free synapses was analyzed using the MetaMorph software as described previously (*16*). Three series of serial sections, each of which consisted of ten electron micrographs, were analyzed (n = 1 to 2 mice).

Ex vivo electrophysiology

Hippocampus (Fig. 4 and figs. S13 and S14)

All experiments were performed in 4- to 6-week-old C57BL/6J WT mice, 4- to 6-week-old PV-*Cre*/tdTomato-FLEX double transgenic mice (for miniature excitatory postsynaptic current (mEPSC) recordings from CA1 PV+ interneurons), 5-month-old 5xFAD mice (for mEPSC recordings), or 15- to 18-month-old 5xFAD transgenic mice (for long-term potentiation (LTP) recordings), as described previously (*82*). Animals were sacrificed and then decapitated. Hippocampi from both hemispheres were isolated and placed on an agar block to cut transverse slices (350 μm thick) using a vibrating microtome (VT1200S, Leica) in an ice-cold solution containing 2 mM KCl, 1 mM $MgCl_2$, 2 mM $MgSO_4$, 1.25 mM NaH_2PO_4, 26 mM $NaHCO_3$, 1 mM $CaCl_2$, 10 mM D-glucose, and 230 mM sucrose for mice younger than 6 months or 2.5 mM KCl, 10 mM MgSO4, 1.2 mM NaH_2PO_4, 30 mM $NaHCO_3$, 0.5 mM $CaCl_2$, 20 mM HEPES, 25 mM glucose, 93 mM NMDG, 5 mM sodium ascorbate, and 3 mM sodium pyruvate for 15- to 18-month-old mice (pH 7.4 with HCl). All solutions were saturated with 95% O_2 and 5% CO_2 and the osmolality was adjusted to 300 ± 5 mOsm. For in vitro CPTX treatments, slices were incubated in a 2-ml chamber for 4 hours at room temperature in the artificial cerebrospinal fluid (ACSF) solution containing 124 mM NaCl, 2.5 mM KCl, 1.3 mM $MgSO_4$, 1 mM NaH_2PO_4, 26.2 mM $NaHCO_3$, 2.5 mM $CaCl_2$, and 11 mM D-glucose. CPTX was added at 20 μg/ml to slices in the chamber, whereas control slices were incubated in ACSF without CPTX. Next, the slices were transferred to the recording chamber and were continuously perfused with the same ACSF for mEPSC measurements. Picrotoxin (50 μM, Tocris Bioscience), CGP 55845 (3 μM, Tocris), and tetrodotoxin (Tocris) were added to ACSF to block $GABA_A$, $GABA_B$ receptor and Na^+ channels, respectively, and hence to isolate mEPSCs. Whole-cell patch recordings from CA1 pyramidal neurons were obtained using glass electrodes (4 to 6 MΩ, outer diameter of 1.5 mm, wall thickness of 0.315 mm; Hilgenberg). The glass electrodes were filled with a solution containing 140 mM K-gluconate, 8 mM NaCl, 0.2 mM $CaCl_2$, 10 mM HEPES, 5 mM lidocaine *N*-ethyl bromide (QX-314 Br), 0.5 mM Na_2GTP, and 2 mM Mg_2ATP (pH 7.2, 290 mOsm). The membrane potential was clamped at −70 mV. For miniature inhibitory postsynaptic current (mIPSC) recordings, NBQX (25 μM, AMPAR antagonist, Tocris), D-AP5 (50 μM, NMDAR antagonist, Tocris) and CGP 55845 were added to the ACSF solution. The glass electrodes were filled with a solution containing 120 mM CsCl, 8 mM NaCl, 0.2 mM $MgCl_2$, 10 mM HEPES, 2 mM EGTA, 5 mM QX-314 Br, 0.3 mM Na_2GTP, and 2 mM Mg_2ATP (pH 7.2, 290 ± 3 mOsm). mEPSC and mIPSC events were analyzed offline using the Mini Analysis software (v6.0.3, Synaptosoft).

ACSF solution containing 120 mM NaCl, 2.5 mM KCl, 1.5 mM $MgCl_2$, 1.25 mM NaH_2PO_4, 24 mM $NaHCO_3$, 2 mM $CaCl_2$, and 25 mM D-glucose was used for field excitatory postsynaptic potential (fEPSP) recording. Thin glass electrodes (outer diameter 1.5 mm, wall thickness 0.188 mm, ~2 MΩ) filled with ACSF were used for stimulation and recording of fEPSPs. The Shaffer collaterals pathway was stimulated using a glass electrode and theta-burst stimulation (TBS) trains were applied three times to induce LTP (*83*). The stimulation intensity was determined based on the input-output curve and was set to give fEPSPs with a slope of ~50% of the supramaximal fEPSP. Single stimuli were repeated every 20 s for at least 10 min for baseline recording before and for 60 min after LTP induction. The paired-pulse ratio (PPR) was evaluated at 50-ms intervals under the same conditions. All recordings were obtained at 30 ± 1°C using an EPC-10 amplifier (HEKA Elektronik). The recordings were filtered at 1 to 3 kHz and digitized at 10-20 kHz.

Cerebellum

Parasagittal cerebellar slices (200 μm thick) were prepared from 4- to 6-week-old *GluD2*-null or *Cbln1*-null mice at 3 days after vehicle, Cbln1, $NP1_{PTX\text{-}3Cl}$ or CPTX injection, as described previously (*84*)). Whole-cell patch-clamp recordings were made from PCs using a 60× water-immersion objective attached to an upright microscope (BX51WI, Olympus) at RT. The resistance of the patch pipettes was 1.5–3 MΩ when filled with an intracellular solution composed of 150 mM Cs-gluconate, 10 mM HEPES, 4 mM $MgCl_2$, 4 mM Na_2ATP, 1 mM Na_2GTP, 0.4 mM EGTA, and 5 mM QX-314 Br (pH 7.3, 298 mOsm/kg) for measurements of the input-output relationship and paired-pulse ratio of PF-EPSC amplitudes, and of 65 mM Cs-methanesulfonate, 65 mM K-gluconate, 20 mM HEPES, 10 mM KCl, 1 mM $MgCl_2$, 4 mM Na_2ATP, 1 mM Na_2GTP, 5 mM sucrose, and 0.4 mM EGTA (pH 7.25, 295 mOsm/kg) for long-term depression (LTD) recordings. The solution used for the slice storage and recording consisted of 125 mM NaCl, 2.5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 1.25 mM NaH_2PO_4, 26 mM $NaHCO_3$ and 10 mM D-glucose, bubbled continuously with a mixture of 95% O_2 and 5% CO_2. Picrotoxin (100 μM, Sigma Aldrich) was always present in the saline to block the inhibitory inputs. To evoke the PF-EPSCs, square pulses were applied at various stimulus intensities (10 μs, 0 to 200 μA) through a stimulating electrode placed on the molecular layer (~50 μm

away from the pial surface). Selective stimulation of the PFs was confirmed by paired-pulse facilitation (PPF) of EPSC amplitudes at a 50-ms inter-stimulus interval. Current responses were recorded with an Axopatch 200B amplifier (Molecular Devices), and pClamp software (v9.2, Molecular Devices) was used for data acquisition and analysis. Signals were filtered at 1 kHz and digitized at 4 kHz.

CPTX injection

Acute intrahippocampal injection

We used a digitally controlled infusion system (UltraMicroPump III with Micro4 Controller, World Precision Instruments) fed with a 10 μl Hamilton syringe and a NanoFil (35 gauge) beveled needle. The mice were first anesthetized with 1 to 3% isoflurane and placed into the stereotaxic frame (Narishige). 1 μl CPTX (1.49 mg/ml) or vehicle (HBS buffer) was bilaterally injected at a rate of 3 nl/s. We used the following coordinates for bilateral injection: AP = −2.0 mm from Bregma and L = ±1.5 mm; DV = 2.0 mm from the brain surface according to the mouse brain atlas (*85*).

Acute intracerebellar injection

Mice were anesthetized with an intraperitoneal injection of ketamine/xylazine (80/20 mg/kg body weight, Daiichi-Sankyo or Sigma Aldrich). A small hole in the occipital bone was made with a dental drill, and the dura matter was ablated as described previously (*16*). A glass needle was inserted into the vermis of lobule VI/VII at ± 0.7 mm from the midline at a depth of ~200 μm. Approximately 7 μl of vehicle, Cbln1-HIS or CPTX-HIS (7.7 μM) was bilaterally injected at a rate of 10 to 40 μl/hour.

Acute intrahippocampal injection through guide cannulas

Intrahippocampal guide cannulas were implanted as described previously (*86*–*88*) with minor changes. In brief, mice were anesthetized with 1 to 3% isoflurane mixed with O_2 through a vaporizer (Matrx VIP 3000, Midmark). Prior to any surgical manipulation, the mice were given the analgesic carprofen (5 mg/kg body weight (b.w.), subcutaneous injection (s.c.), Rimadyl, Pfizer Pharma) and the skin was cleaned by 75% ethanol, followed by 10% povidone iodine (Dynarex) and additional analgesic lidocaine (Xylocaine spray, 10 mg/metered dose, AstraZeneca). The mice were placed in a stereotaxic frame (Narishige), and all procedures were performed under a surgical binocular microscope (Labomed Prima DNT, Labo America) and on a heating pad (DC Temperature Controller, World Precision Instruments) to maintain mouse body temperature at 34° to 36°C. The mouse scalp skin was circularly incised (ϕ 10 mm) and removed. The edges of the skin were processed with 75% ethanol and xylocaine. The scalp bone was carefully cleaned with 75% ethanol and 3% H_2O_2 and dried under 40°C (HG 2310 LCD, Steinel). After marking coordinates for implantation, 4 small holes for anchoring screws were drilled in the frontal and parietal bones by using a dental micro motor (Eickemeyer). Then all four screws were covered with acrylic dental cement Paladur (Heraeus Kulzer) leaving the marked areas for cannulas free of acrylic. Implantation began with drilling of a hole for the left cannula. The cannula was stereotaxically implanted by using a self-made universal holder for cannulas (small curved tweezers with a metal ring from Fine Science Tools). During implantation, the cannula gently touched the surface of the brain and only then it was secured with a small amount of acrylic. Next, the right cannula was similarly implanted and secured with acrylic. Coordinates for bilateral cannulas (AP: 2.0 mm; L: ±1.5 mm from bregma, the bottom end of cannulas should touch the surface of the brain), were set according to the mouse brain atlas (*85*). After implantation of cannulas, more dental cement was added to secure the whole system on the bone. After the surgery which took ~60 min, mice were returned to their home cages. Carprofen (5 mg/kg b.w. s.c., Rimadyl, Pfizer Pharma) was used as a postoperative analgesic. The intrahippocampal injections and behavioral tests were performed after the mice had fully recovered at least 5 to 7 days after surgery. Injections were done using a 10 μl Hamilton syringe as described in the *Acute intrahippocampal injection* section but being inserted in the guide cannula under brief sedation with 1% isoflurane.

Acute intraspinal injection to SCI mice

A volume of 0.5 μl of CPTX (1 μg/μl), Cbln1 (1 μg/μl), Chondroitinase ABC (0.5 U/μl; Sigma Aldrich cat. no. C2905) or vehicle (HBS buffer) were injected in the proximity of the injury site by electrical micro-injector (BJ-110, BEX Co.) through glass capillaries (Drummond Scientific Company, cat. no. 3-000-203-G/X). Following injection, the muscle layers and skin were again closed by suturing.

In utero electroporation

In utero electroporation (IUE) was performed at E14.0 in the cortex of ICR mice as described previously (*89*). Briefly, plasmid DNA was injected into the third ventricle by a glass pipette and a set of electrical pulses was applied 5 times. Positive electrodes were placed onto the cortical side. Plasmid DNA was dissolved in 21 mM HEPES, 137 mM NaCl, 5 mM KCl and 0.7 mM Na_2HPO_4 at a concentration of 1 mg/ml.

Surgical procedure for spinal cord injury (SCI)

Mice (9 to 11 weeks old) were subjected to hemisection- or compression-induced SCI. After the anesthesia, the spinal cord was surgically exposed and the dorsal column at the 10^{th} thoracic vertebra (T10) was dissected by micro-scissors and a scalpel. The muscle layers and skin were closed by sutures. Animals were recovered from the anesthetic by the administration of the antagonist. Mice were subjected to compression-induced damage using an SCI impactor device (Infinite Horizon Impactor; Precision Systems and Instrumentation) (*90*) and using a 70-kdyn impact force. We monitored the strength and the duration of the impact of compression through the automated recording system in the SCI impactor device.

Behavioral Analysis

Spatial navigation in a labyrinth

The 3D-printed dry maze was developed at the DZNE. In short, before implantation of cannulas, mice were put on the light food restriction diet (LFRD) scheme where ad libitum food was removed from mice and was given daily manually at the same time in amounts comparable with normal consumption (4 to 5 g/day/mouse). The weight of all mice was monitored daily and kept within a 10 to 15% reduction. Before the cannulation of mice, labyrinth habituation and training was performed for 2-3 weeks. Once all mice were familiar with the labyrinth, the room was changed to introduce new landmarks and cues, and the configuration of the labyrinth was changed to be more complex (Fig. 5A). Testing in the labyrinth began on day 3 after CPTX or Mock injection. Mice were placed into a starting zone in the labyrinth and given 20 min to find a hidden colorless and odorless round pellet reward (AIN-76A 10 mg tablet, TestDiet, cat. no. 1811213). Two hours later the mice were placed again in the same labyrinth with the same room landmarks and cues for 10 min to find the reward. On day 4, the mice were placed again into the same labyrinth in which the reward zone was changed, and the mice had to relearn the new position of the reward within 10 min. Two hours later, during the retrieval session (5 min), the mouse had to find the new location of the reward. The labyrinth was cleaned with 60% ethanol between trials.

The position of mice in the labyrinth was recorded using a USB video camera (Microsoft) and special behavioral video acquisition and analysis software (ANY-maze). All recorded movies were analyzed using ANY-maze by a trained observer blind to the mouse genotype and treatments. ANY-maze was run in tracking mode to trace coordinates of mice (10 frames/s) and compute distance traveled from the starting zone to the hidden reward zone. The criteria for successful performance were that during encoding and retrieval sessions, the hidden rewards (7 to 8 pellets) were found and eaten. Travel distances in meters were computed as follows: mice had to find the reward and take at least 1 pellet. This time point, as a beginning of the association of the reward location in the labyrinth and all

landmarks and cues in the room, was used to calculate travel distances.

Contextual fear conditioning

To validate the action of CPTX in another hippocampus-dependent cognitive task, we chose contextual fear conditioning (CFC). Shortly after spatial navigation in the labyrinth, at day 5 all mice were first recorded in a neutral context (context B, 5 min) and then 1 hour later they were put into a conditioned context (context A, 5 min) and given three foot shocks, each with 0.5 mA and 1 s in duration and separated by 1 min. Contexts A and B were different in terms of walls (chess-like black-and-white pattern versus grey color), smell (different cleaning solutions) and floor (context A; metal grid to deliver foot shock, context B: white plastic floor. At day 6, 1 day after CFC, mice were first put in the context A for 5 min, then at least 1 hour later in the context B for 5 min for retrieval sessions. CFC memory in retrieval sessions was measured as the fraction of the time a mouse was freezing in contexts A and B. Freezing of mice was defined as total immobility of animals (except for breathing) with a characteristic tense fearful posture. The analysis of data was done by using ANY-maze software. The fear conditioning system consisted of a touch-pad controller and a conditional sound-attenuated cabinet and chambers (Ugo Basile).

Rotor-rod test and gait analyses

To examine the skilled motor coordination in the mice, the rotor-rod test was performed. As this phenotype is clearly observed on the ICR genetic background (*13*, *16*, *84*), we used in 4- to 6-week-old *Cbln1*-null mice with this background. The rotor-rod test was performed 1 day before and 3 days after the injection. Six trials were performed at 20 rpm with a 30-s interval between trials, and the latency to fall from the rotor-rod was measured (maximum score; 120 s). For gait analyses, we used 4- to 6-month-old *Cbln1*-null and *GluD2*-null (C57BL/6J) mice. In addition, 4- to 6-week-old *Cbln1*-null mice (ICR) were used. Mice were habituated to the experimenter and the behavior room from 2 to 3 days before the test. Then, the mice were habituated to the gait analysis apparatus made of transparent plexiglass (6 cm × 85 cm) and trained to walk straight to their home cage. Each step during the walking was recorded by a video camera across the transparent floor at 30 frames per second. The time points and locations of the hind limb to attach and detach on the floor were manually determined by an experimenter blinded to experimental conditions, using custom software written in HSP (Hot Soup Processor, ONION software). These analyses were performed > 4 times for each mouse 1 day before and 3 days after Mock, Cbln1 or CPTX injection. Gait parameters were calculated as follows: (i) the coordinated step ratio was calculated as the number of alternate left/right steps divided by the number of total steps; (ii) the stride length of each step along the x-axis (plus direction to home cage); (iii) the stride speed and (iv) stride irregularity (the standard deviation of stride length during the single journey divided by the average stride length).

Locomotor recovery after SCI

The locomotor recovery assessment was performed using video recording as previously reported (*90*). BMS (Basso Mouse Scale) open-field scoring and footfall tests were performed weekly during the 6 to 8 weeks following SCI. The BMS score was evaluated by at least two independent investigators who were blinded to the experimental groups. Mice were excluded if they had an incomplete injury (BMS score > 0) on day 2 after SCI. To quantify the recovery effect, the increase in BMS score between the 1st and 2nd week after SCI was calculated as ΔBMS/week = BMS score (2nd week) − BMS score (1st week).

For footfall tests, mice were placed on a wire-mesh grid and videotaped for 5 min while on the grid. Mice that walked longer than 3 min with more than 70 steps were subjected to scoring by three independent examiners. The total number of footfalls from the bars during the total walking time was counted.

REFERENCES AND NOTES

1. S. B. Nelson, V. Valakh, Excitatory/inhibitory balance and circuit homeostasis in autism spectrum disorders. *Neuron* **87**, 684–698 (2015). doi: 10.1016/j.neuron.2015.07.033; pmid: 26291155
2. J. J. Palop, L. Mucke, Network abnormalities and interneuron dysfunction in Alzheimer disease. *Nat. Rev. Neurosci.* **17**, 777–792 (2016). doi: 10.1038/nrn.2016.141; pmid: 27829687
3. Y. Bozzi, G. Provenzano, S. Casarosa, Neurobiological bases of autism-epilepsy comorbidity: A focus on excitation/inhibition imbalance. *Eur. J. Neurosci.* **47**, 534–548 (2018). doi: 10.1111/ejn.13595; pmid: 28452083
4. J. H. Foss-Feig *et al*., Searching for cross-diagnostic convergence: Neural mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum disorders. *Biol. Psychiatry* **81**, 848–861 (2017). doi: 10.1016/j.biopsych.2017.03.005; pmid: 28434615
5. M. P. Forrest, E. Parnell, P. Penzes, Dendritic structural plasticity and neuropsychiatric disease. *Nat. Rev. Neurosci.* **19**, 215–234 (2018). doi: 10.1038/nrn.2018.16; pmid: 29545546
6. T. C. Südhof, Towards an understanding of synapse formation. *Neuron* **100**, 276–293 (2018). doi: 10.1016/j.neuron.2018.09.040; pmid: 30359597
7. M. Yuzaki, Two classes of secreted synaptic organizers in the central nervous system. *Annu. Rev. Physiol.* **80**, 243–262 (2018). doi: 10.1146/annurev-physiol-021317-121322; pmid: 29166241
8. S. Jang, H. Lee, E. Kim, Synaptic adhesion molecules and excitatory synaptic transmission. *Curr. Opin. Neurobiol.* **45**, 45–50 (2017). doi: 10.1016/j.conb.2017.03.005; pmid: 28390263
9. T. C. Südhof, Synaptic neurexin complexes: A molecular code for the logic of neural circuits. *Cell* **171**, 745–769 (2017). doi: 10.1016/j.cell.2017.10.024; pmid: 29100073
10. H. Takahashi, A. M. Craig, Protein tyrosine phosphatases PTPδ, PTPσ, and LAR: Presynaptic hubs for synapse organization. *Trends Neurosci.* **36**, 522–534 (2013). doi: 10.1016/j.tins.2013.06.002; pmid: 23835198
11. E. M. Johnson-Venkatesh, H. Umemori, Secreted factors as synaptic organizers. *Eur. J. Neurosci.* **32**, 181–190 (2010). doi: 10.1111/j.1460-9568.2010.07338.x; pmid: 20646052
12. T. J. Siddiqui, A. M. Craig, Synaptic organizing complexes. *Curr. Opin. Neurobiol.* **21**, 132–143 (2011). doi: 10.1016/j.conb.2010.08.016; pmid: 20832286
13. K. Matsuda *et al*., Cbln1 is a ligand for an orphan glutamate receptor delta2, a bidirectional synapse organizer. *Science* **328**, 363–368 (2010). doi: 10.1126/science.1185152; pmid: 20395510
14. T. Uemura *et al*., Trans-synaptic interaction of GluRdelta2 and Neurexin through Cbln1 mediates synapse formation in the cerebellum. *Cell* **141**, 1068–1079 (2010). doi: 10.1016/j.cell.2010.04.035; pmid: 20537373
15. J. Elegheert *et al*., Structural basis for integration of GluD receptors within synaptic organizer complexes. *Science* **353**, 295–299 (2016). doi: 10.1126/science.aae0104; pmid: 27418511
16. A. Ito-Ishida *et al*., Cbln1 regulates rapid formation and maintenance of excitatory synapses in mature cerebellar Purkinje cells in vitro and in vitro. *J. Neurosci.* **28**, 5920–5930 (2008). doi: 10.1523/JNEUROSCI.1030-08.2008; pmid: 18524896
17. E. Miura, T. Iijima, M. Yuzaki, M. Watanabe, Distinct expression of Cbln family mRNAs in developing and adult mouse brains. *Eur. J. Neurosci.* **24**, 750–760 (2006). doi: 10.1111/j.1460-9568.2006.04950.x; pmid: 16930405
18. E. Seigneur, T. C. Südhof, Cerebellins are differentially expressed in selective subsets of neurons throughout the brain. *J. Comp. Neurol.* **525**, 3286–3311 (2017). doi: 10.1002/cne.24278; pmid: 28714144
19. V. Krishnan *et al*., Autism gene Ube3a and seizures impair sociability by repressing VTA Cbln1. *Nature* **543**, 507–512 (2017). doi: 10.1038/nature21678; pmid: 28297715
20. E. Seigneur, J. S. Polepalli, T. C. Südhof, Cbln2 and Cbln4 are expressed in distinct medial habenula-interpeduncular projections and contribute to different behavioral outputs. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10235–E10244 (2018). doi: 10.1073/pnas.1811086115; pmid: 30287486
21. E. Seigneur, T. C. Südhof, Genetic ablation of all cerebellins reveals synapse organizer functions in multiple regions throughout the brain. *J. Neurosci.* **38**, 4774–4790 (2018). doi: 10.1523/JNEUROSCI.0360-18.2018; pmid: 29691328
22. W. Tao, J. Díaz-Alonso, N. Sheng, R. A. Nicoll, Postsynaptic δ1 glutamate receptor assembles and maintains hippocampal synapses via Cbln2 and neurexin. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5373–E5381 (2018). doi: 10.1073/pnas.1802737115; pmid: 29784783
23. E. Favuzzi *et al*., Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science* **363**, 413–417 (2019). doi: 10.1126/science.aau8977; pmid: 30679375
24. M. Fossati *et al*., Trans-synaptic signaling through the glutamate receptor Delta-1 mediates inhibitory synapse formation in cortical pyramidal neurons. *Neuron* **104**, 1081–1094.e7 (2019). doi: 10.1016/j.neuron.2019.09.027; pmid: 31704028
25. M. Yuzaki, A. R. Aricescu, A GluD coming-of-age story. *Trends Neurosci.* **40**, 138–150 (2017). doi: 10.1016/j.tins.2016.12.004; pmid: 28110935
26. W. Kakegawa *et al*., Anterograde C1ql1 signaling is required in order to determine and maintain a single-winner climbing fiber in the mouse cerebellum. *Neuron* **85**, 316–329 (2015). doi: 10.1016/j.neuron.2014.12.020; pmid: 25611509
27. S. M. Sigoillot *et al*., The secreted protein C1ql1 and its receptor BAI3 control the synaptic connectivity of excitatory inputs converging on cerebellar purkinje cells. *Cell Rep.* **10**, 820–832 (2015). doi: 10.1016/j.celrep.2015.01.034; pmid: 25660030
28. D. C. Martinelli *et al*., Expression of C1ql3 in discrete neuronal populations controls efferent synapse numbers and diverse behaviors. *Neuron* **91**, 1034–1051 (2016). doi: 10.1016/j.neuron.2016.07.002; pmid: 27478018
29. K. Matsuda *et al*., Transsynaptic modulation of kainate receptor functions by C1q-like proteins. *Neuron* **90**, 752–767 (2016). doi: 10.1016/j.neuron.2016.04.001; pmid: 27133466
30. C. Straub *et al*., Distinct subunit domains govern synaptic stability and specificity of the kainate receptor. *Cell Rep.* **16**, 531–544 (2016). doi: 10.1016/j.celrep.2016.05.093; pmid: 27346345
31. J. Schwenk *et al*., High-resolution proteomics unravel architecture and molecular diversity of native AMPA receptor complexes. *Neuron* **74**, 621–633 (2012). doi: 10.1016/j.neuron.2012.03.034; pmid: 22632720

32. K. H. Loh *et al*., Proteomic analysis of unbounded cellular compartments: Synaptic clefts. *Cell* **166**, 1295–1307.e21 (2016). doi: 10.1016/j.cell.2016.07.041; pmid: 27565350
33. D. Xu *et al*., Narp and NP1 form heterocomplexes that function in developmental and activity-dependent synaptic plasticity. *Neuron* **39**, 513–528 (2003). doi: 10.1016/S0896-6273(03)00463-X; pmid: 12895424
34. R. J. O'Brien *et al*., Synaptic clustering of AMPA receptors by the extracellular immediate-early gene product Narp. *Neuron* **23**, 309–323 (1999). doi: 10.1016/S0896-6273(00)80782-5; pmid: 10399937
35. G. M. Sia *et al*., Interaction of the N-terminal domain of the AMPA receptor GluR4 subunit with the neuronal pentraxin NP1 mediates GluR4 synaptic recruitment. *Neuron* **55**, 87–102 (2007). doi: 10.1016/j.neuron.2007.06.020; pmid: 17610819
36. S. J. Lee *et al*., Presynaptic neuronal pentraxin receptor organizes excitatory and inhibitory synapses. *J. Neurosci.* **37**, 1062–1080 (2017). doi: 10.1523/JNEUROSCI.2768-16.2016; pmid: 27986928
37. I. Farhy-Tselnicker *et al*., Astrocyte-secreted glypican 4 regulates release of neuronal pentraxin 1 from axons to induce functional synapse formation. *Neuron* **96**, 428–445.e13 (2017). doi: 10.1016/j.neuron.2017.09.053; pmid: 29024665
38. M. C. Chang *et al*., Narp regulates homeostatic scaling of excitatory synapses on parvalbumin-expressing interneurons. *Nat. Neurosci.* **13**, 1090–1097 (2010). doi: 10.1038/nn.2621; pmid: 20729843
39. K. A. Pelkey *et al*., Pentraxins coordinate excitatory synapse maturation and circuit integration of parvalbumin interneurons. *Neuron* **85**, 1257–1272 (2015). doi: 10.1016/j.neuron.2015.02.020; pmid: 25754824
40. J. H. Trotter *et al*., Synaptic neurexin-1 assembles into dynamically regulated active zone nanoclusters. *J. Cell Biol.* **218**, 2677–2698 (2019). doi: 10.1083/jcb.201812076; pmid: 31262725
41. T. Biederer, P. S. Kaeser, T. A. Blanpied, Transcellular nanoalignment of synaptic function. *Neuron* **96**, 680–696 (2017). doi: 10.1016/j.neuron.2017.10.006; pmid: 29096080
42. N. Srinivasan *et al*., Comparative analyses of pentraxins: Implications for protomer assembly and ligand binding. *Structure* **2**, 1017–1027 (1994). doi: 10.1016/S0969-2126(94)00105-7; pmid: 7881902
43. T. Iijima, Y. Iijima, H. Witte, P. Scheiffele, Neuronal cell type-specific alternative splicing is regulated by the KH domain protein SLM1. *J. Cell Biol.* **204**, 331–342 (2014). doi: 10.1083/jcb.201310136; pmid: 24469635
44. M. Masugi-Tokita *et al*., Number and density of AMPA receptors in individual synapses in the rat cerebellum as revealed by SDS-digested freeze-fracture replica labeling. *J. Neurosci.* **27**, 2135–2144 (2007). doi: 10.1523/JNEUROSCI.2861-06.2007; pmid: 17314308
45. H. Oakley *et al*., Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: Potential factors in amyloid plaque formation. *J. Neurosci.* **26**, 10129–10140 (2006). doi: 10.1523/JNEUROSCI.1202-06.2006; pmid: 17021169
46. J. Götz, L. G. Bodea, M. Goedert, Rodent models for Alzheimer disease. *Nat. Rev. Neurosci.* **19**, 583–598 (2018). doi: 10.1038/s41583-018-0054-8; pmid: 30194347
47. N. Crouzin *et al*., Area-specific alterations of synaptic plasticity in the 5XFAD mouse model of Alzheimer's disease: Dissociation between somatosensory cortex and hippocampus. *PLOS ONE* **8**, e74667 (2013). doi: 10.1371/journal.pone.0074667; pmid: 24069328
48. L. Asboth *et al*., Cortico-reticulo-spinal circuit reorganization enables functional recovery after severe spinal cord contusion. *Nat. Neurosci.* **21**, 576–588 (2018). doi: 10.1038/s41593-018-0093-5; pmid: 29556028
49. B. Chen *et al*., Reactivation of dormant relay pathways in injured spinal cord by KCC2 manipulations. *Cell* **174**, 521–535.e13 (2018). doi: 10.1016/j.cell.2018.06.005; pmid: 30033363
50. A. Sathyamurthy *et al*., Massively parallel single nucleus transcriptional profiling defines spinal cord neurons and their activity during behavior. *Cell Rep.* **22**, 2216–2225 (2018). doi: 10.1016/j.celrep.2018.02.003; pmid: 29466745
51. J. Aoto, D. C. Martinelli, R. C. Malenka, K. Tabuchi, T. C. Südhof, Presynaptic neurexin-3 alternative splicing trans-synaptically controls postsynaptic AMPA receptor trafficking. *Cell* **154**, 75–88 (2013). doi: 10.1016/j.cell.2013.05.060; pmid: 23827676
52. E. J. Bradbury *et al*., Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature* **416**, 636–640 (2002). doi: 10.1038/416636a; pmid: 11948352
53. D. M. Basso *et al*., Basso Mouse Scale for locomotion detects differences in recovery after spinal cord injury in five common mouse strains. *J. Neurotrauma* **23**, 635–659 (2006). doi: 10.1089/neu.2006.23.635; pmid: 16689667
54. C. H. Cheng *et al*., Local delivery of high-dose chondroitinase ABC in the sub-acute stage promotes axonal outgrowth and functional recovery after complete spinal cord transection. *PLOS ONE* **10**, e0138705 (2015). doi: 10.1371/journal.pone.0138705; pmid: 26393921
55. M. Ferrer-Ferrer, A. Dityatev, Shaping synapses by the neural extracellular matrix. *Front. Neuroanat.* **12**, 40 (2018). doi: 10.3389/fnana.2018.00040; pmid: 29867379
56. G. H. Diering, R. L. Huganir, The AMPA receptor code of synaptic plasticity. *Neuron* **100**, 314–329 (2018). doi: 10.1016/j.neuron.2018.10.018; pmid: 30359599
57. Y. Fukata *et al*., Epilepsy-related ligand/receptor complex LGI1 and ADAM22 regulate synaptic transmission. *Science* **313**, 1792–1795 (2006). doi: 10.1126/science.1129947; pmid: 16990550
58. K. L. Lovero, Y. Fukata, A. J. Granger, M. Fukata, R. A. Nicoll, The LGI1-ADAM22 protein complex directs synapse maturation through regulation of PSD-95 function. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4129–E4137 (2015). doi: 10.1073/pnas.1511910112; pmid: 26178195
59. D. Laverty *et al*., Cryo-EM structure of the human α1β3γ2 $GABA_A$ receptor in a lipid bilayer. *Nature* **565**, 516–520 (2019). doi: 10.1038/s41586-018-0833-4; pmid: 30602789
60. R. Canitano, M. Pallagrosi, Autism spectrum disorders and schizophrenia spectrum disorders: Excitation/inhibition imbalance and developmental trajectories. *Front. Psychiatry* **8**, 69 (2017). doi: 10.3389/fpsyt.2017.00069; pmid: 28507523
61. A. R. Aricescu, W. Lu, E. Y. Jones, A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243–1250 (2006). doi: 10.1107/S0907444906029799; pmid: 17001101
62. P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401–1407 (1993). doi: 10.1126/science.8248779; pmid: 8248779
63. Y. H. Takeo, W. Kakegawa, E. Miura, M. Yuzaki, RORα regulates multiple aspects of dendrite development in cerebellar Purkinje cells in vitro. *J. Neurosci.* **35**, 12518–12534 (2015). doi: 10.1523/JNEUROSCI.0075-15.2015; pmid: 26354918
64. J. Elegheert *et al*., Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat. Protoc.* **13**, 2991–3017 (2018). doi: 10.1038/s41596-018-0075-9; pmid: 30455477
65. T. S. Walter *et al*., A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 651–657 (2005). doi: 10.1107/S0907444905007808; pmid: 15930615
66. G. Winter, C. M. Lobley, S. M. Prince, Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260–1273 (2013). doi: 10.1107/S0907444913015308; pmid: 23793152
67. N. K. Sauter, R. W. Grosse-Kunstleve, P. D. Adams, Robust indexing for automatic data collection. *J. Appl. Crystallogr.* **37**, 399–409 (2004). doi: 10.1107/S0021889804005874; pmid: 20090869
68. P. Evans, Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006). doi: 10.1107/S0907444905036693; pmid: 16369096
69. P. R. Evans, An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011). doi: 10.1107/S090744491003982X; pmid: 21460446
70. W. Kabsch, Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010). doi: 10.1107/S0907444909047374; pmid: 20124693
71. A. J. McCoy *et al*., Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007). doi: 10.1107/S0021889807021206; pmid: 19461840
72. G. N. Murshudov *et al*., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011). doi: 10.1107/S0907444911001314; pmid: 21460454
73. P. D. Adams *et al*., PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi: 10.1107/S0907444909052925; pmid: 20124702
74. G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008). doi: 10.1038/nprot.2008.91; pmid: 18600222
75. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi: 10.1107/S0907444910007493; pmid: 20383002
76. V. B. Chen *et al*., MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi: 10.1107/S0907444909042073; pmid: 20057044
77. The PyMOL Molecular Graphics System, Version 1.7 Schrödinger, LLC.
78. M. Howarth *et al*., Monovalent, reduced-size quantum dots for imaging receptors on living cells. *Nat. Methods* **5**, 397–399 (2008). doi: 10.1038/nmeth.1206; pmid: 18425138
79. D. G. Myszka, Improving biosensor analysis. *J. Mol. Recognit.* **12**, 279–284 (1999). doi: 10.1002/(SICI)1099-1352(199909/10)12:5<279::AID-JMR473>3.0.CO;2-3; pmid: 10556875
80. Y. Chen *et al*., NS21: Re-defined and modified supplement B27 for neuronal cultures. *J. Neurosci. Methods* **171**, 239–247 (2008). doi: 10.1016/j.jneumeth.2008.03.013; pmid: 18471889
81. J. Schindelin *et al*., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019; pmid: 22743772
82. I. Pavlov *et al*., Tonic GABAA conductance bidirectionally controls interneuron firing pattern and synchronization in the CA3 hippocampal network. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 504–509 (2014). doi: 10.1073/pnas.1308388110; pmid: 24344272
83. A. G. Nikonenko *et al*., Enhanced perisomatic inhibition and impaired long-term potentiation in the CA1 region of juvenile CHL1-deficient mice. *Eur. J. Neurosci.* **23**, 1839–1852 (2006). doi: 10.1111/j.1460-9568.2006.04710.x; pmid: 16623841
84. W. Kakegawa *et al*., Differential regulation of synaptic plasticity and cerebellar motor learning by the C-terminal PDZ-binding motif of GluRdelta2. *J. Neurosci.* **28**, 1460–1468 (2008). doi: 10.1523/JNEUROSCI.2553-07.2008; pmid: 18256267
85. S. L. Franklin, S. Love, J. R. Greene, S. Betmouni, Loss of perineuronal net in ME7 prion disease. *J. Neuropathol. Exp. Neurol.* **67**, 189–199 (2008). doi: 10.1097/NEN.0b013e3181654386; pmid: 18344910
86. O. Senkov *et al*., Polysialylated neural cell adhesion molecule is involved in induction of long-term potentiation and memory acquisition and consolidation in a fear-conditioning paradigm. *J. Neurosci.* **26**, 10888–109898 (2006). doi: 10.1523/JNEUROSCI.0878-06.2006; pmid: 17050727
87. O. Senkov, A. Mironov, A. Dityatev, A novel versatile hybrid infusion-multielectrode recording (HIME) system for acute drug delivery and multisite acquisition of neuronal activity in freely moving mice. *Front. Neurosci.* **9**, 425 (2015). doi: 10.3389/fnins.2015.00425; pmid: 26594144
88. D. Minge *et al*., Heparan sulfates support pyramidal cell excitability, synaptic plasticity, and context discrimination. *Cereb. Cortex* **27**, 903–918 (2017). doi: 10.1093/cercor/bhx003; pmid: 28119345
89. T. Saito, In vivo electroporation in the embryonic mouse central nervous system. *Nat. Protoc.* **1**, 1552–1558 (2006). doi: 10.1038/nprot.2006.276; pmid: 17406448
90. K. Takeuchi *et al*., Chondroitin sulphate N-acetylgalactosaminyl-transferase-1 inhibits recovery from neural injury. *Nat. Commun.* **4**, 2740 (2013). doi: 10.1038/ncomms3740; pmid: 24220492

ACKNOWLEDGMENTS

We thank staff at the Diamond Light Source beamline I03; T. Walter and K. Harlos for support with protein crystallization; E. Budinger, O. Kobler, and J. Schneeberg for histology advice and assistance; and H. Sakuma and A. Sekigawa for the Airyscan super-resolution imaging. **Funding:** This work was funded by the Human Frontier Science Program (grant RGP0065/2014 to M.Y., A.D. and A.R.A.), the UK Medical Research Council (MRC; grants L009609 and MC_UP_1201/15 to A.R.A.), the Japan Society for the Promotion of Science (JSPS; grant 15H05772 and 16H06461 to M.Y.; 17K10949 and 1705584 to K.T.; 25893233, 26860148, 14J07587 and 18K19380 to K.S.; 18H04563 to W.K.; JP16H06280 to K.M. and K.S.), the Keio University Grant-in-Aid for Encouragement of Young Medical Scientists (to K.S.), the Keio Association Grant-in-Aid (to K.S.), the Astellas Foundation for Research on Metabolic Disorders (to K.S.), the Daiichi Sankyo Foundation of Life Science (to K.S.), the AMED (iD3 16nk0101302h0002 to K.T.; JP18dm0107124 to W.K.), the JST CREST (JPMJCR1854 to M.Y.), the JST ERATO (JPMJER1802 to

W.K.), the Takeda Science Foundation (to M.Y. and W.K.), the Marie-Curie Actions postdoctoral fellowship (grant 328531 to J.E.), the University of Bordeaux Initiative of Excellence (IdEx) fellowship (to J.E.), the European Research Council (ERC Starting Grant 850820 SynLink to J.E.), the European Union 7th Framework Programme Initial Training Network (EU FP7 ITN; grant 606950 EXTRABRAIN to A.D.), and the Bundesministerium für Bildung und Forschung (BMBF; Energl project, TP5, 01GQ1421A to A.D.). **Author contributions:** K.S. performed cell biological experiments, immunohistochemical experiments of hippocampus, cerebellum, and spinal cord tissues, and behavioral studies of cerebellar function; A.J.C., J.E. and A.R.A. collected x-ray data and solved the $NP1_{PTX}$ crystal structure; A.R.A. conceived the CPTX organizer; J.E. performed SPR and MALS experiments; W.K. performed electrophysiological experiments in the cerebellum; I.S. performed electrophysiological experiments in the hippocampus; J.E., V.T.C., and A.J.C. prepared recombinant proteins; O.S. developed a 3D-printed labyrinth and performed behavioral studies of hippocampal function; M.F.-F. and S.S. performed spine analysis; E.M. performed electron microscopic analyses and immunohistochemical experiments of the spinal cord; R.K. performed immunocytochemical and immunohistochemical analyses of hippocampal neurons; K.M. performed immunocytochemical, immunohistochemical and biochemical analyses; K.S. and T.S. performed cell-based binding assays; S.O. developed analytical tools for gait performance; H.S., M. I., Y.M., Y.T. and K.T. generated the experimental models of spinal cord injury and performed behavioral analyses. M.W. raised new antibodies against neuronal pentraxins; A.D., A.R.A., and M.Y. designed and supervised the project; K.S., J.E., A.D., A.R.A., and M.Y. wrote the paper. **Competing interests:** K.S., K.T., and M. Y. are coinventors on Japan patent application 2020-2879 (10 January 2020) describing the use of synthetic synapse connector for spinal cord injury. **Data and materials availability:** All data are available in the main text or the supplementary materials. All the plasmids and antibodies generated in this study are available from the authors upon request. Structure factors and coordinates of $NP1_{PTX}$ are deposited in the Protein Data Bank (PDB code 6YPE).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/eabb4853/suppl/DC1
Figs. S1 to S22
Table S1
Movies S1 to S8

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

27 February 2020; accepted 24 July 2020
10.1126/science.abb4853

RESEARCH ARTICLE SUMMARY

CELL BIOLOGY

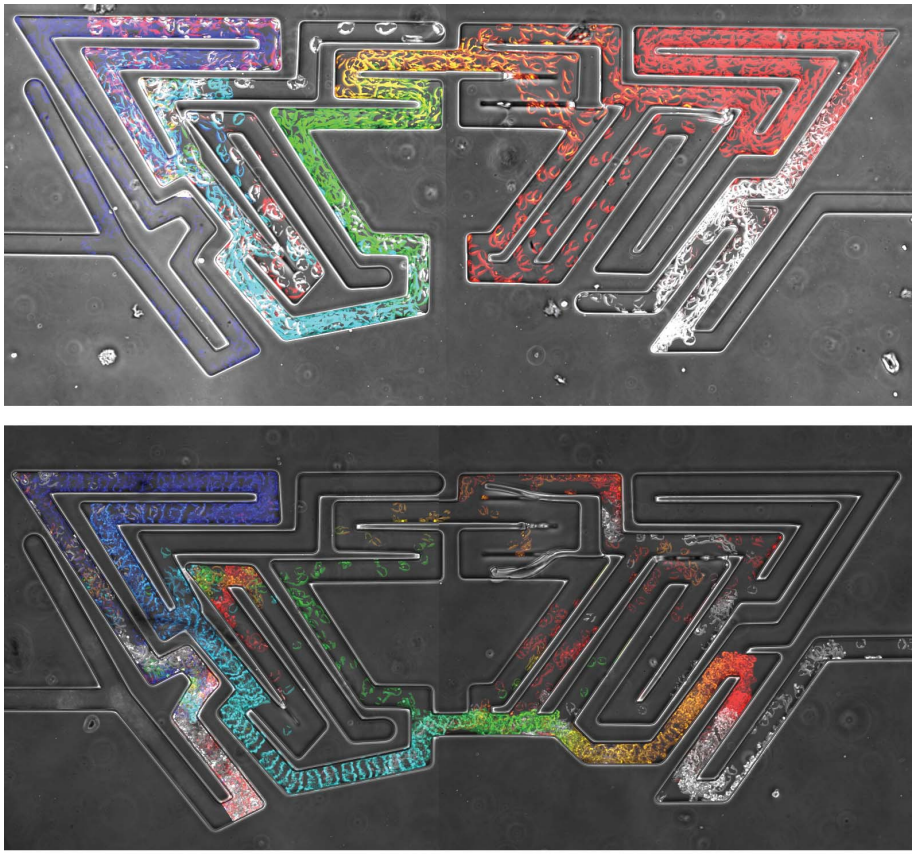
Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown

Luke Tweedy, Peter A. Thomason, Peggy I. Paschke, Kirsty Martin, Laura M. Machesky, Michele Zagnoni, Robert H. Insall*

INTRODUCTION: Cells migrate over long distances during normal embryonic development and in disease, and they navigate through complex, branched paths. Chemotaxis, cell migration steered by gradients of attractive chemicals, is a central regulator of this behavior. Its sensitivity is limited by the responsiveness of attractant receptors. Simple chemotaxis is unable to guide cells over long distances and cannot resolve complexities and branches.

Self-generated chemotaxis, a process in which cells create their own local, dynamic gradients by breaking down an attractant in their environment, is different. Gradients can be steeper because they are only made in the vicinity of the cells. They also work over long distances because they are remade locally as the cells migrate. Furthermore, self-generated gradients can follow complex paths because they are constantly generated by interactions between the cells and their environment.

RATIONALE: The migration of cells through complex structures such as developing embryos or vascularized tumors involves both long distances and numerous decisions about which path to follow. Therefore, we examined chemotaxis in artificial complex environments, which are represented here by mazes. We predicted the pathfinding of cells in mazes using computational and mathematical modeling of self-generated gradients and tested the outcomes with real cells in microfluidic mazes.



Cells solving mazes by using self-generated chemotaxis and identifying a shortcut. (Top) *D. discoideum* solving a representation of the maze from Hampton Court Palace. (Bottom) Cells sensing a shortcut using the self-generated gradient and selecting a new minimal route. Cells enter from the left and exit at the right. Time is represented by cell color; blue shows early times and red later times.

RESULTS: We modeled cells as they navigated through various junctions between a connection to an attractant well and a dead end. We then tested the model's predictions with cells that typically use long-range navigation: *Dictyostelium discoideum* cells (which find each other over large distances in the environment) and metastatic cancer cells (which spread around the human body). Both successfully solved a range of mazes, even remarkably complex ones (see figure), and could identify optimum paths. The results confirm that a model of self-generated chemotaxis captures all essential decision-making in these mazes: Cells broke down attractants locally, which were replenished at different, predictable rates by each possible path. The resulting local attractant gradients allowed cells to sense upcoming maze junctions before reaching them, even around corners, and to make correct decisions about paths they had not yet encountered. Attractant breakdown thus allowed cells to create a detailed picture of their surroundings.

Real cell behavior closely agreed with the models for both rapidly moving *D. discoideum* and slower-moving pancreatic cancer cells, implying that the underlying conceptual model is an effective explanation of general cell behavior. In particular, the diffusivity of attractant, which is ignored by simple models, is crucial to cells' responses. We simulated and then constructed "easy" and "hard" mazes, which appear similar but caused radically different pathfinding by cells. Finally, we designed maze geometries that passively created a temporary peak of attractant flux, deliberately misdirecting cells into a dead end and creating a chemotactic "mirage."

CONCLUSION: Self-generated chemoattractant gradients allow cells to navigate complex paths with great efficiency. Diffusion and attractant breakdown allow cells to obtain detailed information about their surroundings that could not be provided by simple attractant gradients. Cell migration in vivo cannot be understood without considering the interplay among cells, attractants, and the structure of the local environment. Microfluidic mazes are an excellent experimental system for studying these interactions. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: Robert.Insall@glasgow.ac.uk
Cite this article as L. Tweedy *et al.*, *Science* **369**, eaay9792 (2020). DOI: 10.1126/science.aay9792

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aay9792

RESEARCH ARTICLE

CELL BIOLOGY

Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown

Luke Tweedy[1], Peter A. Thomason[1], Peggy I. Paschke[1], Kirsty Martin[1]*, Laura M. Machesky[1,2], Michele Zagnoni[3], Robert H. Insall[1,2]†

During development and metastasis, cells migrate large distances through complex environments. Migration is often guided by chemotaxis, but simple chemoattractant gradients between a source and sink cannot direct cells over such ranges. We describe how self-generated gradients, created by cells locally degrading attractant, allow single cells to navigate long, tortuous paths and make accurate choices between live channels and dead ends. This allows cells to solve complex mazes efficiently. Cells' accuracy at finding live channels was determined by attractant diffusivity, cell speed, and path complexity. Manipulating these parameters directed cells in mathematically predictable ways; specific combinations can even actively misdirect them. We propose that the length and complexity of many long-range migratory processes, including inflammation and germ cell migration, means that self-generated gradients are needed for successful navigation.

Cells migrating in embryogenesis (*1*–*3*), immune responses (*4*, *5*), and neural pathfinding (*6*, *7*) steer using chemotaxis, migrating up gradients of attractants such as chemokines and netrins. Simple chemotaxis, in which gradients are established between a localized attractant source and an external sink, only provides short-range guidance (*8*). It becomes inefficient at distances above 500 μm (*9*), cannot resolve complex paths, and is only effective over a narrow range of attractant concentrations. These restrictions have confounded our understanding of how chemotaxis drives longer-range phenomena such as neural crest migration (*10*) and cancer metastasis (*11*). However, when cell groups locally break down an attractant found throughout their surroundings, they create their own local, dynamic gradients (*1*, *12*–*14*). These typically direct migration away from areas with a high density of cells, promoting metastasis (*15*). Such self-generated gradients work over arbitrarily long distances and work equally robustly with a wide range of attractant concentrations (*16*). Here, we set out to test their role in resolving complex paths such as, for example, one that a cell migrating through an embryo would follow. We found that cells using self-generated gradients could make accurate choices about paths they had not yet encountered. This enabled them to solve complex mazes even when the initial environment was homogeneous and the correct destination was distant. Computational models, combined with live observations using microfluidic devices (*17*), have revealed that the accuracy of decisions in complex environments is determined by the complexity and lengths of the paths and the migration speeds of the cells. This mechanism explains how cells can be guided over far greater distances than can be explained by simple chemotaxis and how they can interpret environmental features in a way that would be impossible with simple attractant sources.

[1]CRUK Beatson Institute, Switchback Road, Glasgow G61 1BD, UK. [2]Institute for Cancer Sciences, University of Glasgow, Glasgow G61 1QH, UK. [3]Centre for Microsystems and Photonics, Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, UK.
*Present address: CRUK AstraZeneca Antibody Alliance Laboratory, Granta Park, Cambridge CB21 6GS, UK.
†Corresponding author. Email: Robert.Insall@glasgow.ac.uk

Self-generated gradients promote long-range chemotaxis

Chemotactic cells detect attractant gradients by comparing receptor occupancy at different places. Cells can resolve 1% differences (*18*) between the occupancy at their fronts and rears, but this is only enough to navigate short distances. Beyond 0.5 to 1 mm, gradients contain zones that are either too shallow or too saturating to cause a 1% occupancy difference and thus become undetectable (*9*). However, cells can make sharp, local gradients by breaking down attractants. The resulting dynamic, self-generated gradients are usually impossible to measure directly. They are best analyzed using computational models, with the predictions of the models tested experimentally in quantitative experiments on living cells. Figure 1A and movie S1 show modeled cells responding to a passive 1-mm gradient or to a self-generated gradient in which the attractant is broken down by a cell surface enzyme. Passive gradients of this distance are either too shallow or, if steeper, rapidly saturate the receptors, so cells always steer poorly in simple 1-mm gradients. By contrast, the self-generated gradient gives robust chemotaxis throughout because the gradient is always sharp and local to the group of cells that makes it, resulting in a nonsaturating attractant concentration around the cells (Fig. 1A and movie S1). Secreted enzymes give similar results (*9*). Real cells behaved the way the model predicted (Fig. 1B and movie S1).

Self-generated gradients allow cells to make long-range route decisions

Chemotaxis studies often ignore diffusion because they consider the steady state of imposed, linear gradients (*19*, *20*). For dynamic gradients such as self-generated gradients, diffusion is a key determinant. Interactions between the depletion of attractant by cells, the diffusion of fresh attractant toward cells, and the movement of cells in the resulting gradients can lead to counterintuitive results. We therefore modeled the way that cells make decisions at junctions (Fig. 1, C to E). Even in the simple case of two equivalent routes to an attractant reservoir (Fig. 1C), self-generated and static gradients elicited different behaviors. In the static gradient, each cell chose a route randomly. The self-generated gradient, on the other hand, robustly directed equal numbers of cells into each branch. Stochastic variations were balanced out; branches containing more cells evolved shallower attractant gradients, so newly arriving cells were directed into the other branch (shown quantitatively in Fig. 1F). If one of the branches was closed off from the reservoir (Fig. 1D), then the number of cells entering the closed branch was small because the first arrivals rapidly depleted the attractant and prevented further recruitment.

When the closed branch was shorter (Fig. 1E), an unexpected behavior emerged from the models. Before the migrating cells reached the junction, they cleared the attractant from the short branch by diffusion and therefore never made the decision to enter the short branch. In essence, the cells sensed the space ahead of themselves at a distance, making accurate predictions about the outcome of pathways that they had never visited.

The effect depended strongly on the length of the dead end. Fewer than 10% of cells committed to dead ends shorter than 250 μm (Fig. 1G). The time taken for attractant to diffuse out of the dead end increased quadratically with the length of the channel, so cell commitment increased substantially with the length of the dead end. Dead ends longer than 600 μm attracted 40% of exploring cells.

To investigate how cells sense closed paths without visiting them, we modeled cells following self-generated gradients through complex routes, analogous to physiological problems such as finding a path from a tumor into a blood vessel. In our models, a single correct path, 1050 μm long, connected each starting well to a reservoir of attractant; the number and nature of dead ends varied. Parameters were taken from measurements of *Dictyostelium discoideum*

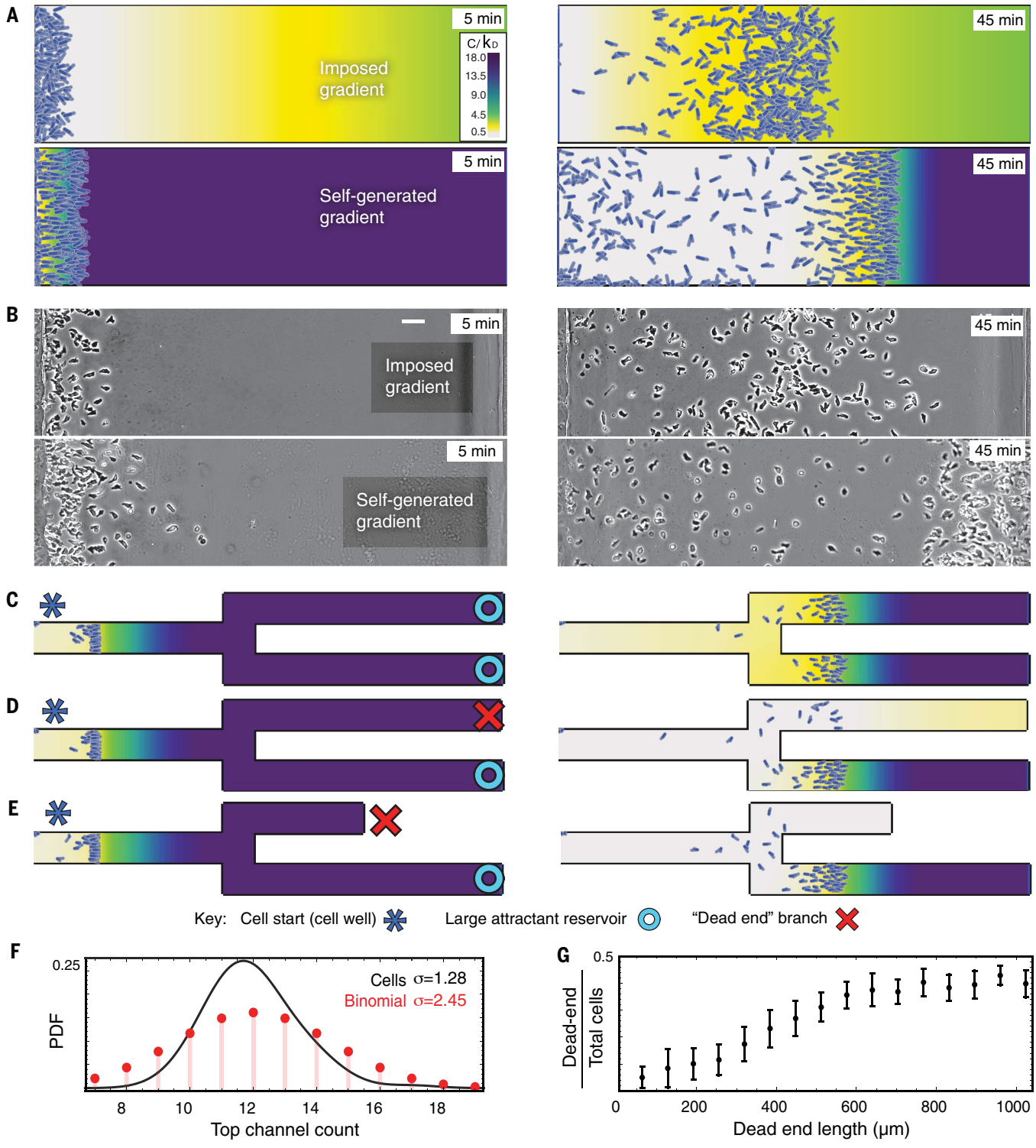


Fig. 1. Self-generated gradients allow cells to explore remote features. (**A**) Comparison of imposed and self-generated gradients guiding cells across 1 mm to a full attractant well. (**B**) Experimental verification of (A) using *D. discoideum*. The imposed gradient uses the nondegradable attractant Sp-cAMPS. The self-generated gradient uses uniform cAMP. Cells following the imposed gradient performed worse, especially after the halfway point. Scale bar, 50 μm. See movie S1. (**C** to **E**) Simulated navigation past a junction. In (C), both branches are identical and connect to an attractant reservoir. Each recruits the same number of cells. In (D), one branch is a dead end. Some cells do still commit because of residual attractant in the channel. In (E), the dead end is much shorter and is almost entirely free of attractant as the cells reach the junction. (**F**) Number of cells selecting the top channel through repeated simulations of (C). Concentration is tuned so that an average of 24 cells commit. The self-generated gradient has a smaller standard deviation than a random choice, revealing active sorting. (**G**) The fraction of cells committing to a dead end as a function of its length. Few cells commit to short dead ends, but there is an apparently linear increase from 250 μm to 650 μm; above this, the fraction plateaus at about 0.4.

cells (*21*, *22*), which exhibit chemotaxis toward 3′,5′- cyclic adenosine monophosphate (cAMP) while breaking it down using a cell surface phosphodiesterase (*23*). We verified the model's predictions by fabricating identical microfluidic mazes in polydimethylsiloxane (PDMS; fig. S2). Cells were able to find the correct route through all designs despite starting with uniform chemoattractant, so there were initially no directional cues. As shown below, the simulations accurately predicted the behavior of real cells (Figs. 2, 3, 5, and 6).

Self-generated gradients allow cells to navigate mazes

We found that *D. discoideum* cells were able to navigate environments of surprising length and complexity by combining the long range of self-generated chemotaxis with the ability to detect dead ends ahead (Fig. 2). To ensure that the cells were not convolving the results by generating their own cAMP signals, we generated a new adenylyl cyclase mutant (*aca*A) in a wild-type (NC4) parent, which is healthier and migrates better than more widely used axenic strains. These cells recapitulated the model's predictions with considerable accuracy (movies S2 to S4; compare Fig. 2, A and B, D and E, and G and H). We confirmed that the gradients were self-generated using a nondegradable ligand. Cells made almost no progress when the breakdown-resistant cAMP analog Sp-cAMPS (*24*) was used instead of cAMP (fig. S1, bottom panels). To minimize confounding hydraulic effects, we used 40 μm × 25 μm maze channels that are much wider and taller than *D. discoideum* cells (<10 μm wide and ~5 μm tall).

These results show that self-generated gradients allowed cells to steer themselves through long, complex mazes that would be impossible to navigate using simple gradients.

Accuracy of decisions is controlled by length and complexity of paths

We designed three different mazes to test the accuracy of cells' decision-making strategies (Fig. 2, A, D, and G). Each had the same correct path, but "simple" mazes had three short dead ends (Fig. 2, A to C), "short" mazes had dead ends half the length of their live channels (Fig. 2, D to F), and "long" mazes had symmetrical live and dead channels, with the only difference being the connection to the attractant reservoir (Fig. 2, G to I). We tested the model's predictions using two very different cell types: *D. discoideum* cells and mouse pancreatic cancer cells that self-generate gradients of lysophosphatidic acid (LPA) in serum (*25*).

The model predicted that most cells would avoid dead ends in the simple mazes (Fig. 2A and movie S2). Both cell types fulfilled this prediction, with a leading group of cells migrating accurately through the maze and a few cells deviating into the dead ends (Fig. 2, B and C, and movies S3 and S4). Decisions were poorer in short mazes (Fig. 2, D to F), and in long mazes, many cells erroneously chose the dead ends (Fig. 2, G to I).

To measure the dynamic fidelity of decisions, we recorded the cumulative number of cells committing to the dead ends and live channels. We observed that the error rate was highest in the first decision in each maze in both simulations and experiments and therefore focused on the behavior at this branch point (Fig. 3, A and B). The results confirmed that longer dead ends lead to poorer decision fidelity: Cells performed consistently better in simple versus long mazes and better in general in short versus long mazes (Fig. 3, C to E; the blue line shows the numbers of cells choosing the correct path and the red line the number committing to the dead end). Long mazes were significantly less biased over the first hour than the other designs (Fig. 3F; one-way ANOVA and Tukey's test, $\alpha = 0.05$).

The behavior of cells in long mazes changed over time. Initially, there was little difference between the numbers of cells choosing correctly and incorrectly, but as cells depleted the attractant in the dead ends, their decisions became more accurate (Fig. 3, C, D, and G). The cancer cells (in homogeneous medium, where the attractant derived from the 10% serum) moved much more slowly than the *D. discoideum* cells but showed the same trends in decision-making (Fig. 3E and movie S4). The similarity in behavior between these different cell types indicates that these are general features of the chemotactic response to self-generated gradients.

Cell speed and attractant diffusivity

We supplemented our modeling and experimental approaches with a mathematical analysis that considered the effects of the lengths of the live channel and dead end on a decision at a T junction (described in detail in the supplementary materials, section 3.1). Because real cells are not static when they decide a direction at junctions, we developed a mathematical mapping connecting the cell speed to a static waiting time; this was validated by comparing it with the simulations. The analysis yielded three key predictions, which are shown in Fig. 4A: (i) shorter dead ends gave more accurate decisions, (ii) shorter live ends also gave more accurate decisions, and (iii) decisions were more accurate if cells took longer to make them.

We had already observed the first prediction experimentally, giving confidence in the others. The second prediction resulted from two factors: For short live channels, equilibrium was quickly reached and the resulting well-to-junction gradient was steeper. The third prediction resulted from cells having a longer time to clear attractant diffusing out of dead ends before deciding which path to take. This explains the cells' greater accuracy in the second and third decisions in the mazes (e.g., Fig. 3G and movies S2 and S3). The same pattern was observed when cell speed in the simulations was altered because slower cells effectively took longer to make decisions (Fig. 4B; note similarity to Fig. 4A).

These findings raised an apparent contradiction. We predicted that slower cells would make better decisions than faster ones, yet *D. discoideum*, which solved mazes in 2 hours, performed similarly to cancer cells, which took ~2 days. The attractants for these cells, cAMP (*26*) and LPA (*15*) respectively, have similar molecular masses, so we expected them to have similar diffusivities. However, lipids are often carried by proteins such as albumin, slowing their effective diffusion. We predicted (Fig. 4C) the relationship between the diffusivities of the two attractants that would lead to equal decision fidelity and then performed a photobleaching assay on fluorescently labeled cAMP and LPA (Fig. 4, D and E). As expected, the measured effective diffusivity of LPA was much lower than that of cAMP and close to the value predicted by our model (10% of the cAMP value versus ~5% in our prediction; see the supplementary materials, section 3.2).

This result highlights how much the accuracy of decisions in a maze depends on the rate of diffusion of the attractant. We therefore simulated the three maze designs using attractants with a range of diffusivities. The mathematical analysis predicted that high diffusivity would yield excellent decisions, and that fidelity would decrease with diffusivity to a limit where decisions were made with no information, so 50% of cells head in each direction. The models followed this prediction, with one surprising exception: Cells in the short maze, but not the long maze, were predicted to do worse than 50:50 for slow-diffusing attractants, with diffusivities ~1/10 that of a small molecule such as LPA or cAMP (Fig. 4, F and G). However, on detailed inspection, we found that in the short maze, two dead-end channels were close enough to the junction to supply diffusing attractant molecules, whereas in the live channel, molecules traveling a similar distance came from a single source. The quadratic relationship between diffusive flux and distance meant that branches further up the live channel had a minimal effect. This explains the behavior reversal between the short and long mazes and again emphasizes the importance of modeling in interpreting the complex behaviors of self-generated gradients.

Complex topologies drive cells into incorrect decisions

Analysis of the counterintuitive short maze result leads to a surprising prediction: A dead

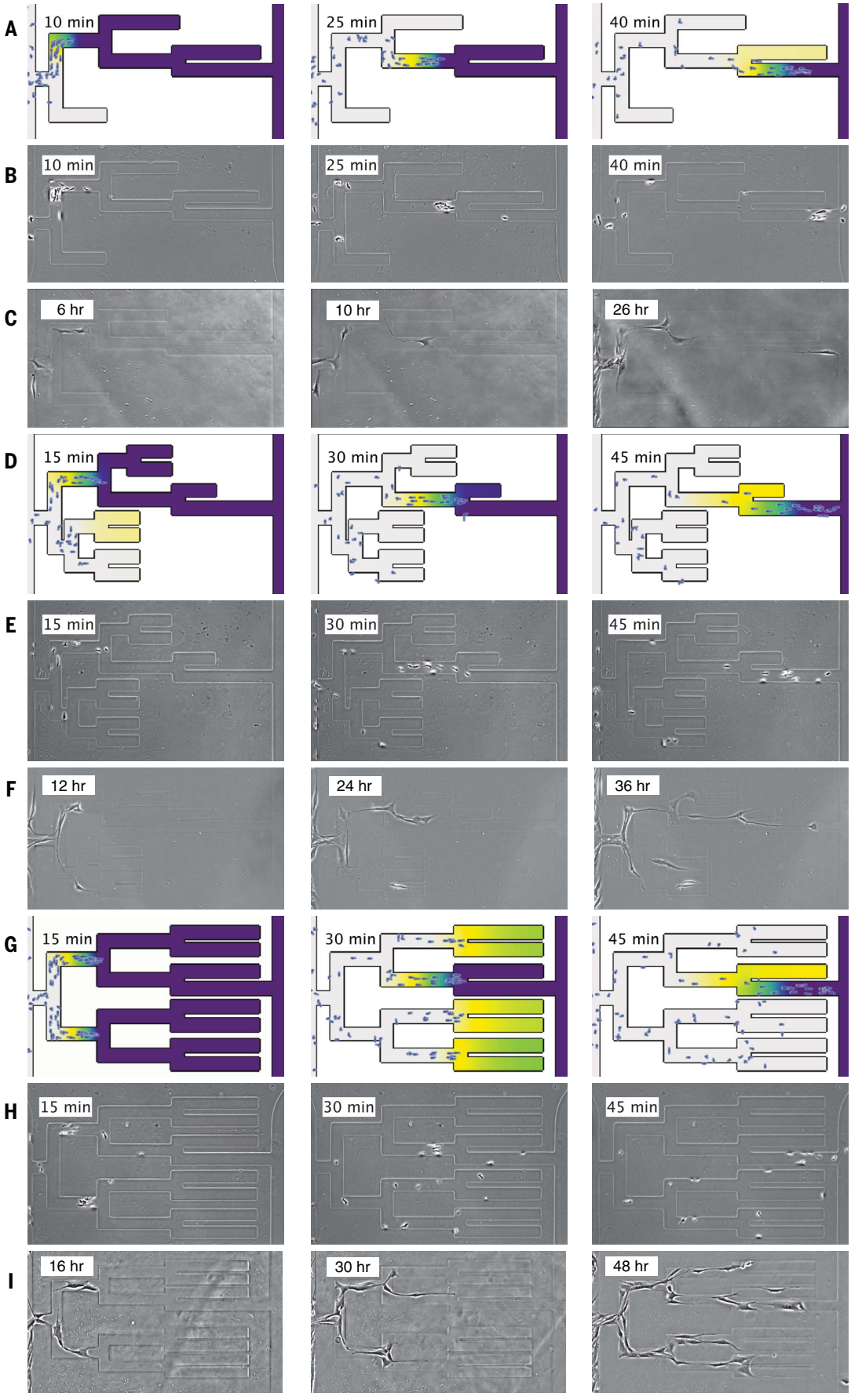


Fig. 2. Real cells can solve mazes. (**A**) Simulation of cells navigating the maze with short dead ends at time points past the first, second, and third decisions. In this design, cells are predicted to almost always commit to the correct path to lead them to the attractant well. (**B**) *D. discoideum* cells migrating through the same maze design initially filled uniformly with the attractant cAMP. (**C**) Pancreatic cancer cells in the same maze design with an initial background of 10% FCS. (**D** to **F**) Simulations of the short-branched maze (D) compared with *D. discoideum* cells (E) and pancreatic cancer cells (F) navigating the same design. (**G** to **I**) Simulations of the long-branched maze design (G) compared with *D. discoideum* cells (H) and pancreatic cancer cells (I). See movie S2 for (A), (D), and (G); movie S3 for (B), (E), and (H); and movie S4 for (C), (F), and (I). Device width = 850 μm; channel width = 40 μm; channel height = 25 μm.

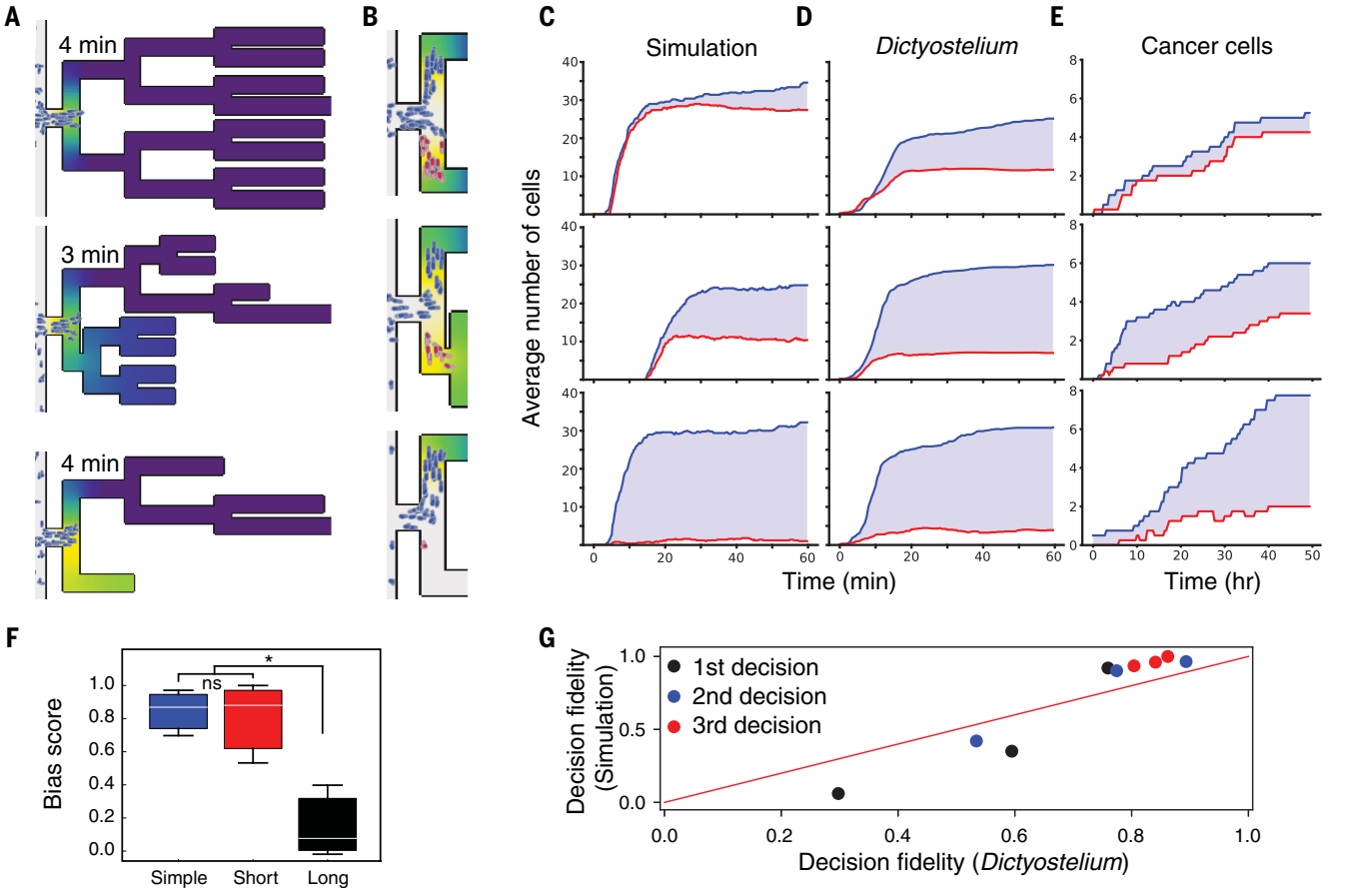


Fig. 3. More distant and complex dead ends are harder to resolve. (**A**) Simulations of the long-, short-, and simple-branched mazes as the cells reach the junction between the live path and the first dead end. (**B**) The same simulations at a later time. Cell color has been altered for those that have committed to the live end (deep blue) or dead end (red). Uncommitted cells are shown in the original gray blue. (**C** to **E**) Running totals of cells committed to the live end (blue) and dead end (red), as pictured in (B). Results are shown for simulations (C), *D. discoideum* experiments (D), and pancreatic cancer experiments (E), with t = 0 when cells first reach the entrance to the maze. Light blue shading highlights the difference between these values. (**F**) Decision fidelity scores for *D. discoideum* cells at the first junction of each maze. The simple- and short-branched mazes both differ significantly in their fidelity to the long-branched maze (one-way ANOVA, α = 0.05). (**G**) Overall decision fidelity scores for the simulations compared with their experimental counterparts. Decision fidelity is $t - f/t + f$ over a 1-hour window, where t and f are the number of cells currently committed to the correct and incorrect paths. Later decisions generally have higher fidelity.

end can be more attractive than a live channel if the dead end branches or widens. This can create chemotactic mirages that lead cells away from the source of attractant.

We built new mazes to test whether mirages actually occur. Each maze connected the cell well and large attractant reservoir with a path of identical length (~800 μm). After a variable approach, cells encountered a T junction, at which they could either migrate toward the large reservoir of attractant or down a dead end, again of variable length, toward a smaller attractant well (Fig. 5, A to C). We created and tested 16 designs in total, with four approach lengths (150, 300, 450, and 600 μm) and four dead-end lengths (150, 300, 450, and 600 μm).

We expected shorter approach lengths to yield more severe mirages for two reasons. First, a shorter approach time gives cells less time to clear attractant from the dead end. Second, the attractant well down the live path would be farther away and would therefore signal more weakly. We also predicted that short dead ends would generate stronger mirages because a shorter distance between the side well and the junction would give a greater supply of attractant molecules from the side well. Finally, we predicted that if the live path to the large reservoir was shorter than the path to the dead end, then no mirage would occur.

In simulations and experiments, short approaches and short dead ends produced a profound mirage, steering cells toward the trap (Fig. 5, A and D, top panels). In intermediate designs, signals from the true reservoir were strong enough to cause some cells to reverse their initial, incorrect decision (Fig. 5, B and D, middle panels, and movie S5). For long approaches and long dead ends, the stronger signal of the large reservoir dominated decision-making, so there was no mirage (Fig. 5, C and D, bottom panels). A new prediction emerged from simulations of these designs: Because these easy-to-avoid mirages moved the trap a large distance away from the cell well, attractant flux through the entrance of the maze was reduced, so fewer cells were recruited (Fig. 5D, compare top and bottom panels). We measured the decision bias through the first hour for each design and found that the relationship between the lengths and decision-making held true across all conditions and was accurately predicted by the model (Fig. 5E). The dead-end mazes confirm that most phosphodiesterase was bound to the cell surface; models in which the enzyme diffuses reproduced the data far less accurately (fig. S7)

These results emphasize a key result that underpins chemotaxis using self-generated gradients: The most important external determinant for decision-making is attractant flux, not attractant quantity. In the mirage

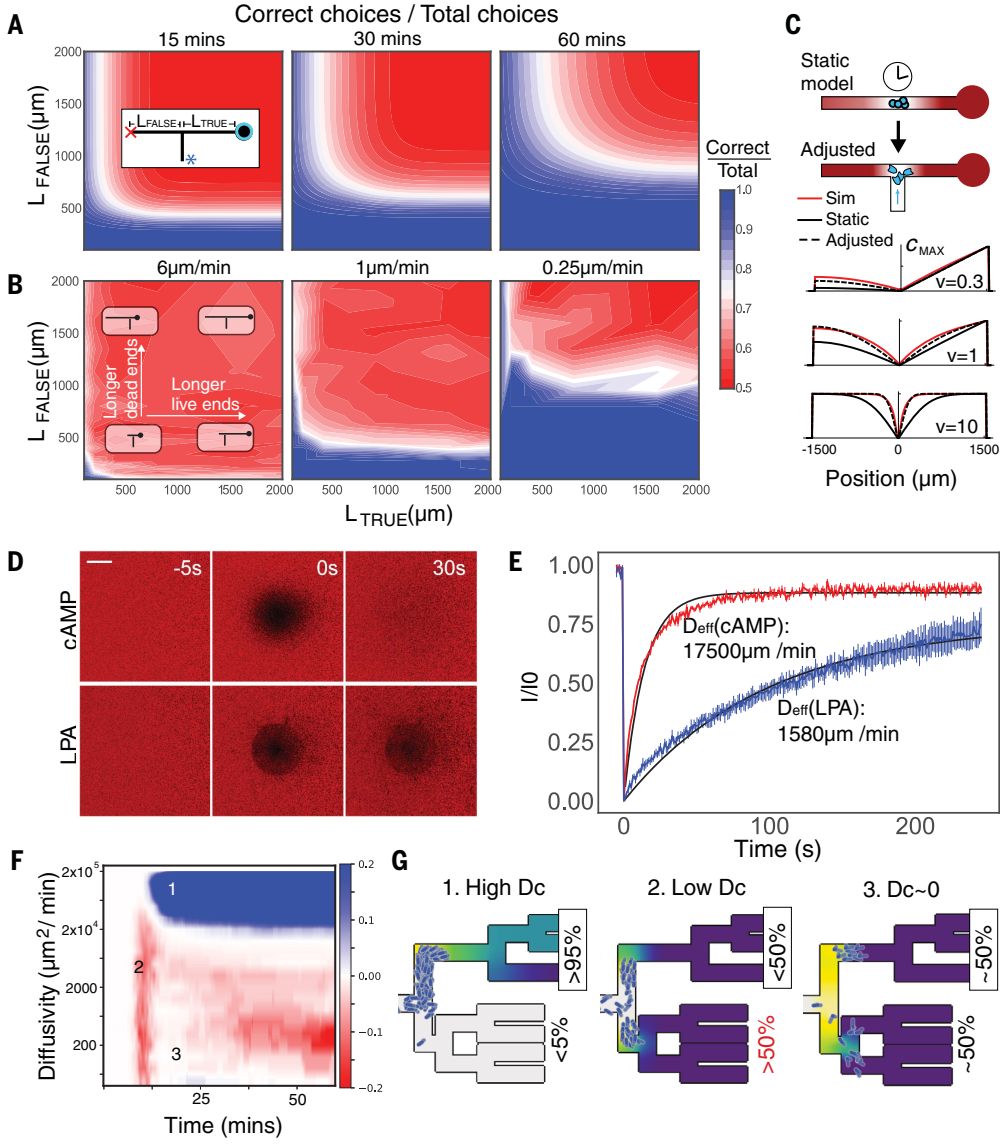


Fig. 4. Slower cells and faster diffusion lead to more accurate decisions. (**A**) Mathematical model of decision fidelity at a T junction. Each panel shows a snapshot after a different waiting time at the junction before making a decision. Lengths of the live and dead ends are shown on the *x*- and *y*-axes, respectively. (**B**) Decisions made by simulated cells moving at different speeds and therefore taking different times to reach the junction so as to correspond with the panels above. (**C**) To correct for cell movement sharpening the gradient, we created a mapping from the solvable static model to an adjusted model. The static and adjusted models are shown for three cell speeds against the true gradient seen in the simulations. (**D**) Photobleaching of 8-fluo-cAMP and TopFluor Lyso PA. (**E**) Recovery curves and fits for the photobleached areas suggest that the effective diffusivity of LPA is ~1/10 that of cAMP. Scale bar, 50 μm. (**F**) Expected decision fidelities for the short-branched maze across time simulated over a variety of attractant diffusivities. A phase change in behavior occurs for lower diffusivities, in which most cells choose the dead end. This does not happen in the long-branched maze, revealing that in branching topologies, shorter dead ends may in fact lead to worse decisions. (**G**) Snapshots of three simulations at the three numbered points in (F).

maze designs, long dead ends contain more attractant than short dead ends but signal more weakly. This has obvious relevance to migration in vivo; cancer cell metastasis, for example, favors tracks with low resistance (*27*), which allow greater attractant flux.

Targeting specific outcomes in topology design

As a final test of our models, we made two pairs of closely related mazes. For each pair, the paths to the source and dead-end space were similar, but one was designed to be "easy" (models predict accurate decisions) and one "hard" (movies S6 and S7). Easy mazes had short, weakly branched dead ends. Hard mazes had long, highly branched dead ends, with branching beginning near the entry to maximize the mirage effect (Fig. 6).

The progress of cells was scored by counting how many passed a checkpoint after each major decision (in Fig. 6A, yellow, blue, and purple spots mark checkpoints). Cell behavior in "labyrinth" designs was very similar to that predicted by the simulations (Fig. 6, B and C, show the first checkpoint for the easy and hard designs, respectively). Cells scored extremely divergently in the two designs despite the similarity of the paths (Fig. 6D). The second pair of mazes, easy and hard "trident" mazes (Fig. 6, E to G), yielded very similar results (Fig. 6, F to H). In both design pairs, large parts of the easy mazes were not visited, whereas the hard maze was thoroughly explored (Fig. 6, I and J). Therefore, understanding the principles of chemotactic mazes can accurately inform how real cells respond to complex environments.

Discussion

Our results have implications for many areas of biology. We found that the detailed structure of the environments that cells encounter substantially affects their ability to steer accurately. In many cases, self-generated gradients allow cells to steer with an accuracy that seems intuitively impossible. Changes in maze design that seem minor make a substantial—and predictable—difference in the accuracy and eventual destination of the cells. Because our experimental results were accurately predicted by mathematical and computational methods, we believe that our findings are general and apply to any system in which an attractant is degraded by the cells that respond to it. This mechanism is entirely different from that used by *Physarum polycephalum* to solve mazes (*28*), which relies on the plasmodium migrating down all branches simultaneously before pruning useless paths (*29*).

Many situations in which chemotaxis occurs in vivo, for example, neutrophils extravasating to an infection in tissue (*30*) or germ cells migrating through an embryo (*31*), have

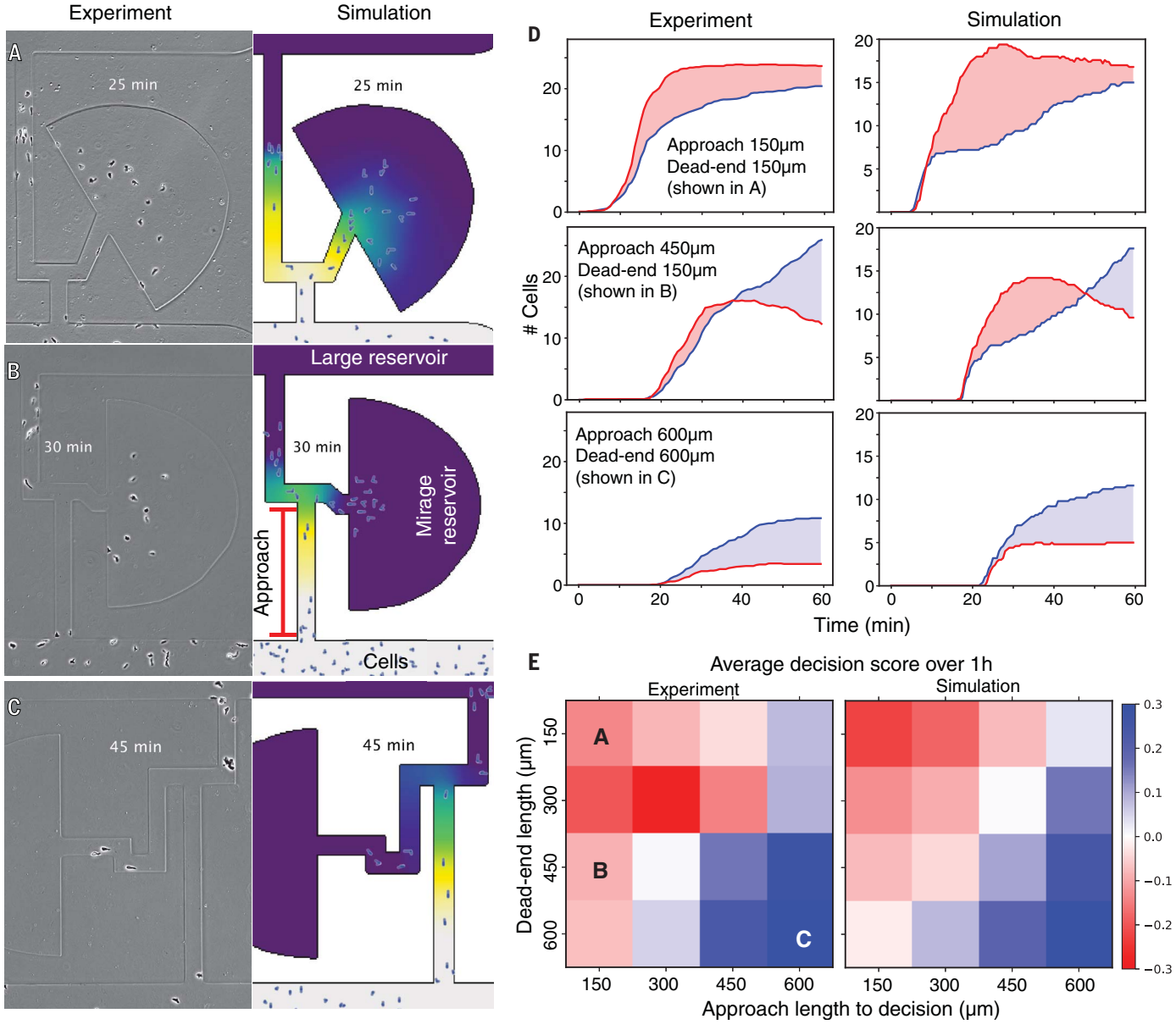


Fig. 5. Cell decisions depend on the rate of attractant transport. (**A** to **C**) Using four dead-end lengths and four junction approach lengths [labeled in (B)], we generated 16 mazes in total with varying predictions of cell behavior. We show real cells (left) and simulation predictions (right) for three of the 16 designs. (A) shows short approach length and dead-end length (both 150 µm). (B) shows a 150-µm dead end and a 450-µm approach, with cells faring better than in (A). (C) shows an approach length and a dead-end length of 600 µm. In this case, cells overwhelmingly steer correctly toward the large reservoir. Dynamics of (A) to (C) can be seen in movie S5. Distance between reservoirs = 800 µm. (**D**) Total number of cells committing to the live and dead ends on average for the designs shown in (A) to (C). Blue shading shows a positive bias (favoring the live end) and red shading shows a negative bias (favoring the dead end). For a short approach and a short dead end, real cells have a strongly negative bias that is not overcome within the hour-long observation (top left); this was predicted by the simulation (top right). An intermediate approach length and a short dead end causes a brief chemotactic mirage, but the dead end is short enough that the misdirected cells then receive attractant flux from the large reservoir and rectify their mistake (middle left). This, again, is as predicted (middle right). A long approach and long dead end (bottom left) has two features. First, the bias is consistently positive and no mirage forms. Second, this design recruits fewer cells overall. Both features were predicted by the simulations. (**E**) Average bias score over a 45-min observation for all 16 designs. The general trend of short approach and dead ends both causing negative bias, which was predicted by the model (right), is seen clearly in the experimental data (left).

equivalently complex migration paths. Similarly, attractant degradation is widespread, with examples known in immunity (*32*), development (*31*), and cancer (*33*). Chemotactic cells have a variety of mechanisms for depleting attractants, including receptor-ligand endocytosis (*34*), decoy receptors (*1*, *12*, *35*), and cell surface enzymes that degrade attractants (*33*, *36*). Ligand breakdown is rarely considered when interpreting spatial patterns of chemotaxis data because its effects can be complex, counterintuitive, and difficult to measure, but our results show that it needs to be analyzed and understood. Additionally, many cells create attractants as well as degrading them, and additional attractants may influence behavior independently. The need to analyze the interactions between these complex and

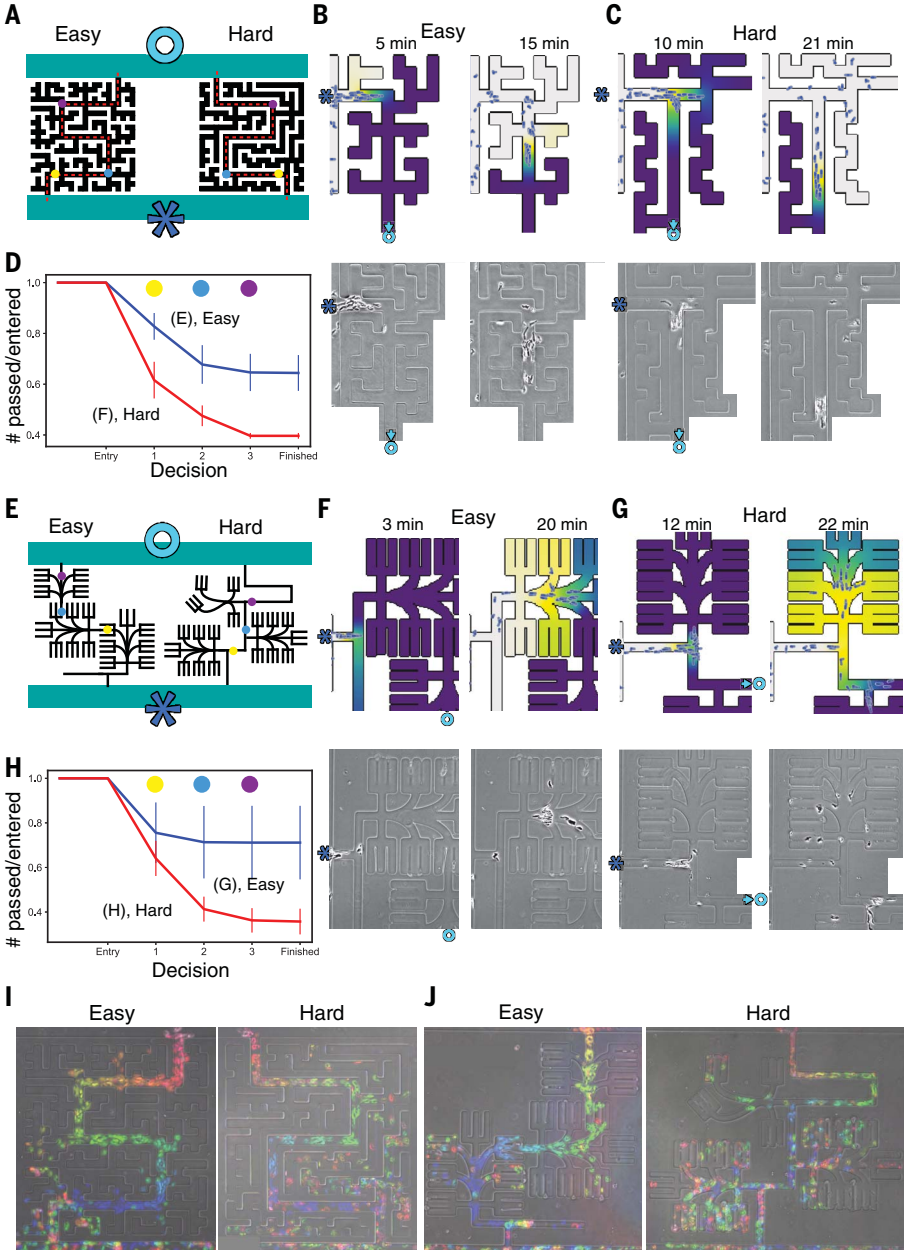


Fig. 6. Deliberate misdirection of live cells. (**A**) Designs for easy and hard labyrinth mazes. Although both have the same amount of dead-end space, the hard labyrinth is designed with fewer, longer, and more branched dead ends. We note that the easy labyrinth is harder to solve visually. Checkpoints are marked (colored dots) and were used to score the navigational success of cells. Devices are 800 μm wide. (**B** and **C**) Example images of cells passing the first checkpoint in both labyrinths alongside the behavior predicted by simulations. See movie S6. (**D**) Number of *D. discoideum* cells passing each checkpoint in (A) as a fraction of maze entrants. The checkpoint color tag is shown above each value. (**E**) Designs for easy (left) and hard (right) trident mazes along similar principles to (A). Devices are 810 μm wide. (**F** to **H**) Similar images of cells navigating trident designs. See movie S7. Pairwise *t* testing of the final decision yielded $P < 0.05$ for both design pairs. (**I** and **J**) Cell images taken at 5-min intervals color coded by time and superimposed on the unmoving, median image of the mazes for all four designs show the paths taken by cells.

counterintuitive processes points to the pivotal role of computational modeling combined with quantitative experimental measurements in the future of biology.

We have described decisions within our mazes as correct or incorrect, but these loaded terms need not apply to in vivo contexts. Our aim is to understand how complex topology draws cells to locations that intuition would not predict. The chemotactic mirage, in which expanding and branching topologies provide more powerful chemoattraction than a direct source of chemoattractant, is particularly counterintuitive. This could be crucial to understanding migratory behavior in complex in vivo environments (*11*); for example, neutrophil extravasation into tissues, migration of melanoblasts through the embryonic dermis, or metastasis of glioblastoma through white matter tracts of the brain (*37*).

In summary, the interaction between the degradation of chemoattractants by cells and the structure of the landscape explains how cells navigate over long distances in complex environments. Self-generated chemoattractant gradients represent an understudied phenomenon that could offer answers to many unexplained physiological behaviors.

Materials and Methods

Cell lines

Mutant *aca*A *D. discoideum* cells were generated by homologous recombination in an NC4 parental background using resuspended bacteria according to (*38*), and were grown nonaxenically on bacterial lawns. To sensitize cells to attractant, they were collected at 1.5 to 3 × 10^7 cells/ml in development buffer (DB), shaken for 1 hour, and then pulsed at 6-min intervals for 3.5 to 4 hours with 300 nM cAMP. Cells were then pelleted and resuspended in maze medium [DB + 2.5 μM cAMP + 0.05% bovine serum albumin (BSA)] and inoculated into the cell well at high density (~90% confluency). The cancer line used was KPC model murine pancreatic cancer (-kras -p53). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) + 20% fetal calf serum (FCS), trypsinized, resuspended in DMEM + 10% fresh FCS, and placed in the cell well of the maze. In these experiments, mazes were filled with DMEM + 10% fresh FCS.

Basic chemotaxis experiments

D. discoideum cells were harvested, collected at 1.5 to 3 × 10^7 cells/ml in DB, and shaken for 1.5 hour. Cells were then pulsed with 300 nM cAMP for 4.5 hours at 6-min intervals, bringing them into an aggregation competent state. For the results shown in Fig. 1, an Insall chamber was used to create a stable attractant gradient in a viewing bridge between two connected reservoirs of attractant. For an imposed gradient, the inner attractant reservoir was filled with DB containing 1 μM Sp-cAMPS, and the outer reservoir was filled with attractant-free DB. For a self-generated gradient, both the inner and outer reservoir were filled with 10 μM cAMP in DB. Cells were mixed into the outer well medium at a density of 2.5 × 10^6/ml.

Maze design and fabrication

Design schematics for mazes are given in the supplementary materials. Microfluidic mazes were fabricated in PDMS (Sylgard 184, Dow

Corning, US) using standard soft lithography techniques. Briefly, silicon masters were produced using an SU8 3005 photoresist (3000 series, MicroChem, US) on silicon wafers following the manufacturer's protocol to achieve a final resist thickness of 25 μm. The resist was exposed through a photomask (JD Photo-Tools, UK) to collimated UV light and developed in MicroPosit EC solvent (Rohm and Haas, US). To prevent PDMS adhesion to the silicon master, it was salinized by vapor deposition of 1H,1H,2H,2H perfluorooctyl-trichlorosilane for 1 hour. PDMS was poured onto the silicon master at a 10:1 ratio of base to curing agent, degassed in a vacuum desiccator chamber, and cured at 70°C for at least 3 hours. PDMS devices were then peeled from the mold, cut to the desired size, and 2-mm holes were punched to obtain inlet and outlet ports. PDMS devices were then cleaned and irreversibly bonded to glass-bottomed petri dishes (MatTek Life Sciences, U.S.) using oxygen plasma.

Maze use

Mazes were filled uniformly with medium by filling all inlet ports with the medium of choice (typically ~6 μl per well) and then placing into a vacuum desiccator for ~20 min, degassing the PDMS. When the vacuum is released, the pressure difference draws medium into all parts of the maze, including dead ends, although this functions best if additional medium is pipetted up and down in each well because this dislodges residual gas bubbles. Mazes used for cancer cells were prefilled with 0.05% BSA in sterile, deionized water to block the PDMS and prevent any attractants from adhering. The prefill was dried out by first draining the wells thoroughly with a pipette, and then placing them in a tissue culture hood for ~2 hours. As soon as a maze was observed to be dry, it was refilled with experimental medium, which for cancer cells was DMEM + 10% FCS, freshly added.

Simulations

Simulations were written in Java. Diffusion in a complex environment was simulated using the semi-implicit DuFort-Frankel method. Agent-based model cells then made decisions using a persistent, biased random walk. The persistent, random element comes from drawing a new direction at each step from wrapped-normal distribution centered on the current direction of motion. The attractant gradient direction was estimated from grid points that overlapped the cell (all those grid points within 6 μm of the cell centroid) and used to generate a bias vector. The bias vector was added to the persistent, random vector to choose a final direction of motion, and the cell moved in this direction at its current speed unless it collided with a wall, in which case its movement distance was reduced.

Cells degrade attractant at a rate r determined by Michaelis-Menten kinetics as follows:

$$r = v_{max}\frac{c}{c + k_m}$$

with attractant removed evenly from all grid points overlapped by the cell.

Analysis

All decisions in mazes were binary, with cells committing to a live or a dead end. *D. discoideum* cells were counted as having committed once the whole cell body passed out of the junction into one or the other channel. Cancer cells, which are much larger, were counted as having committed once the nucleus left the junction. Decision reversals were tracked, with a cell reentering the junction lowering the score of its channel. To account for some stochastic variation from random movement, all mazes were timed from the arrival of the second cell.

All maze figures involve comparisons between designs. In all cases, these different designs were tested against one another with the same cells on the same day. For the long, short, and simple mazes shown in Figs. 2 and 3, $n = 3$, with 14, 14, and 12 technical replicates used, respectively. For the mirage mazes shown in Fig. 5, $n = 4$, with 9 to 12 technical replicates performed in total for each design. For the easy and hard mazes in Fig. 6A, $n = 3$, with 11 technical replicates of each in total, and for those in Fig. 6C, $n = 3$, with 12 technical replicates of each in total.

REFERENCES AND NOTES

1. E. Donà *et al*., Directional tissue migration through a self-generated chemokine gradient. *Nature* **503**, 285–289 (2013). doi: 10.1038/nature12635; pmid: 24067609
2. M. Chuai, C. J. Weijer, Regulation of cell migration during chick gastrulation. *Curr. Opin. Genet. Dev.* **19**, 343–349 (2009). doi: 10.1016/j.gde.2009.06.007; pmid: 19647425
3. X. Yang, D. Dormann, A. E. Münsterberg, C. J. Weijer, Cell movement patterns during gastrulation in the chick are controlled by positive and negative chemotaxis mediated by FGF4 and FGF8. *Dev. Cell* **3**, 425–437 (2002). doi: 10.1016/S1534-5807(02)00256-3; pmid: 12361604
4. B. A. David, P. Kubes, Exploring the complex role of chemokines and chemoattractants in vivo on leukocyte dynamics. *Immunol. Rev.* **289**, 9–30 (2019). doi: 10.1111/imr.12757; pmid: 30977202
5. I. R. Evans, W. Wood, *Drosophila* blood cell chemotaxis. *Curr. Opin. Cell Biol.* **30**, 1–8 (2014). doi: 10.1016/j.ceb.2014.04.002; pmid: 24799191
6. C. E. Adler, R. D. Fetter, C. I. Bargmann, UNC-6/Netrin induces neuronal asymmetry and defines the site of axon formation. *Nat. Neurosci.* **9**, 511–518 (2006). doi: 10.1038/nn1666; pmid: 16520734
7. T. F. Sloan, M. A. Qasaimeh, D. Juncker, P. T. Yam, F. Charron, Integration of shallow gradients of Shh and Netrin-1 guides commissural axons. *PLOS Biol.* **13**, e1002119 (2015). doi: 10.1371/journal.pbio.1002119; pmid: 25826604
8. L. Tweedy, O. Susanto, R. H. Insall, Self-generated chemotactic gradients-cells steering themselves. *Curr. Opin. Cell Biol.* **42**, 46–51 (2016). doi: 10.1016/j.ceb.2016.04.003; pmid: 27105308
9. L. Tweedy, D. A. Knecht, G. M. Mackay, R. H. Insall, Self-generated chemoattractant gradients: Attractant depletion extends the range and robustness of chemotaxis. *PLOS Biol.* **14**, e1002404 (2016). doi: 10.1371/journal.pbio.1002404; pmid: 26981861
10. R. McLennan, J. M. Teddy, J. C. Kasemeier-Kulesa, M. H. Romine, P. M. Kulesa, Vascular endothelial growth factor (VEGF) regulates cranial neural crest migration in vivo. *Dev. Biol.* **339**, 114–125 (2010). doi: 10.1016/j.ydbio.2009.12.022; pmid: 20036652
11. E. T. Roussos, J. S. Condeelis, A. Patsialou, Chemotaxis in cancer. *Nat. Rev. Cancer* **11**, 573–587 (2011). doi: 10.1038/nrc3078; pmid: 21779009
12. G. Venkiteswaran *et al*., Generation and dynamics of an endogenous, self-generated signaling gradient across a migrating tissue. *Cell* **155**, 674–687 (2013). doi: 10.1016/j.cell.2013.09.046; pmid: 24119842
13. S. R. Schwab, J. G. Cyster, Finding a way out: Lymphocyte egress from lymphoid organs. *Nat. Immunol.* **8**, 1295–1301 (2007). doi: 10.1038/ni1545; pmid: 18026082
14. C. Scherber *et al*., Epithelial cell guidance by self-generated EGF gradients. *Integr. Biol.* **4**, 259–269 (2012). doi: 10.1039/c2ib00106c; pmid: 22314635
15. A. J. Muinonen-Martin *et al*., Melanoma cells break down LPA to establish local gradients that drive chemotactic dispersal. *PLOS Biol.* **12**, e1001966 (2014). doi: 10.1371/journal.pbio.1001966; pmid: 25313567
16. L. Tweedy, R. H. Insall, Self-generated gradients yield exceptionally robust steering cues. *Front. Cell Dev. Biol.* **8**, 133 (2020). doi: 10.3389/fcell.2020.00133; pmid: 32195256
17. D. Irimia, Microfluidic technologies for temporal perturbations of chemotaxis. *Annu. Rev. Biomed. Eng.* **12**, 259–284 (2010). doi: 10.1146/annurev-bioeng-070909-105241; pmid: 20450351
18. S. H. Zigmond, Mechanisms of sensing chemical gradients by polymorphonuclear leukocytes. *Nature* **249**, 450–452 (1974). doi: 10.1038/249450a0; pmid: 4834231
19. S. H. Zigmond, J. G. Hirsch, Leukocyte locomotion and chemotaxis. New methods for evaluation, and demonstration of a cell-derived chemotactic factor. *J. Exp. Med.* **137**, 387–410 (1973). doi: 10.1084/jem.137.2.387; pmid: 4568301
20. D. Zicha, G. A. Dunn, A. F. Brown, A new direct-viewing chemotaxis chamber. *J. Cell Sci.* **99**, 769–775 (1991). pmid: 1770004
21. P. J. Van Haastert, Binding of cAMP and adenosine derivatives to *Dictyostelium discoideum* cells. Relationships of binding, chemotactic, and antagonistic activities. *J. Biol. Chem.* **258**, 9643–9648 (1983). pmid: 6309779
22. L. Wu, J. Franke, R. L. Blanton, G. J. Podgorski, R. H. Kessin, The phosphodiesterase secreted by prestalk cells is necessary for *Dictyostelium* morphogenesis. *Dev. Biol.* **167**, 1–8 (1995). doi: 10.1006/dbio.1995.1001; pmid: 7851634
23. G. Gerisch, D. Hülser, D. Malchow, U. Wick, Cell communication by periodic cyclic-AMP pulses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **272**, 181–192 (1975). doi: 10.1098/rstb.1975.0080; pmid: 1814
24. C. Rossier, G. Gerisch, D. Malchow, Action of a slowly hydrolysable cyclic AMP analogue on developing cells of *Dictyostelium discoideum*. *J. Cell Sci.* **35**, 321–338 (1979). pmid: 217886
25. A. Juin *et al*., N-WASP control of LPAR1 trafficking establishes response to self-generated LPA gradients to promote pancreatic cancer cell metastasis. *Dev. Cell* **51**, 431–445.e7 (2019). doi: 10.1016/j.devcel.2019.09.018; pmid: 31668663
26. Y. Artemenko, T. J. Lampert, P. N. Devreotes, Moving towards a paradigm: Common mechanisms of chemotactic signaling in *Dictyostelium* and mammalian leukocytes. *Cell. Mol. Life Sci.* **71**, 3711–3747 (2014). doi: 10.1007/s00018-014-1638-8; pmid: 24846395
27. P. G. Gritsenko, O. Ilina, P. Friedl, Interstitial guidance of cancer invasion. *J. Pathol.* **226**, 185–199 (2012). pmid: 22006671
28. T. Nakagaki, H. Yamada, A. Tóth, Maze-solving by an amoeboid organism. *Nature* **407**, 470 (2000). doi: 10.1038/35035159; pmid: 11028990
29. T. Nakagaki, H. Yamada, A. Tóth, Path finding by tube morphogenesis in an amoeboid organism. *Biophys. Chem.* **92**, 47–52 (2001). doi: 10.1016/S0301-4622(01)00179-X; pmid: 11527578
30. E. Kolaczkowska, P. Kubes, Neutrophil recruitment and function in health and inflammation. *Nat. Rev. Immunol.* **13**, 159–175 (2013). doi: 10.1038/nri3399; pmid: 23435331
31. A. D. Renault, P. S. Kunwar, R. Lehmann, Lipid phosphate phosphatase activity regulates dispersal and bilateral sorting of embryonic germ cells in *Drosophila*. *Development* **137**, 1815–1823 (2010). doi: 10.1242/dev.046110; pmid: 20431117
32. B. Bréart *et al*., Lipid phosphate phosphatase 3 enables efficient thymic egress. *J. Exp. Med.* **208**, 1267–1278 (2011). doi: 10.1084/jem.20102551; pmid: 21576386
33. O. Susanto *et al*., LPP3 mediates self-generation of chemotactic LPA gradients by melanoma cells. *J. Cell Sci.* **130**, 3455–3466 (2017). doi: 10.1242/jcs.207514; pmid: 28871044

34. M. A. Puthenveedu *et al*., Sequence-dependent sorting of recycling proteins by actin-stabilized endosomal microdomains. *Cell* **143**, 761–773 (2010). doi: 10.1016/j.cell.2010.10.003; pmid: 21111236
35. A. Rot *et al*., Cell-autonomous regulation of neutrophil migration by the D6 chemokine decoy receptor. *J. Immunol.* **190**, 6450–6456 (2013). doi: 10.4049/jimmunol.1201429; pmid: 23670187
36. G. L. Garcia, E. C. Rericha, C. D. Heger, P. K. Goldsmith, C. A. Parent, The group migration of *Dictyostelium* cells is regulated by extracellular chemoattractant degradation. *Mol. Biol. Cell* **20**, 3295–3304 (2009). doi: 10.1091/mbc.e09-03-0223; pmid: 19477920
37. C. Denier, A. F. Carpentier, K. Mokhtari, J. Y. Delattre, Cerebral glioblastomas and systemic metastases [Article in French]. *Rev. Neurol.* **157**, 1531–1534 (2001). pmid: 11924450
38. P. Paschke *et al*., Rapid and efficient genetic engineering of both wild type and axenic strains of *Dictyostelium discoideum*. *PLOS ONE* **13**, e0196809 (2018). doi: 10.1371/journal.pone.0196809; pmid: 29847546
39. L. Tweedy, P. A. Thomason, P. I. Paschke, K. Martin, L. M. Machesky, M. Zagnoni, R. H. Insall, Time-lapse images for: Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown, Enlighten: Research Data (2020); https://dx.doi.org/10.5525/gla.researchdata.1029.

ACKNOWLEDGMENTS

We are very grateful to R. Ward for several rounds of helpful comments on the manuscript and presentation, H. Spence for assistance with tissue culture, and the members of Beatson Advanced Imaging Resource (BAIR) facility for help with microscopy. **Funding:** This work was funded by core grants (A24450 and A15673) from Cancer Research UK to R.I. and L.M., a multidisciplinary project award (A20017) from Cancer Research UK to R.I., CRUK Beatson Institute Core Services and Advanced Technologies (C596/A17196), and internal Strathclyde University Funds to M.Z. **Author contributions:** L.T.: Conceptualization, Formal analysis, Investigation, Software, Visualization, Writing – original draft; P.T.: Investigation, Writing – review and editing; P.P.: Methodology; K.M.: Methodology; L.M.: Conceptualization, Writing – review and editing; M.Z.: Methodology, Funding acquisition; R.I.: Conceptualization, Funding acquisition, Supervision, Visualization. **Competing interests:** M.Z. is cofounder and CSO of ScreenIn3D Limited. The authors declare no other competing interests. **Data and materials availability:** Data are available in the main text or the supplementary materials. Time-lapse images are archived at Enlighten: Research Data (39). Source code is available on GitHub at https://github.com/ltweedy/MazeNavigation.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/eaay9792/suppl/DC1
Supplementary Text
Figs. S1 to S10
Movies S1 to S7
MDAR Reproducibility Checklist

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

2 August 2019; resubmitted 12 May 2020
Accepted 2 July 2020
10.1126/science.aay9792

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

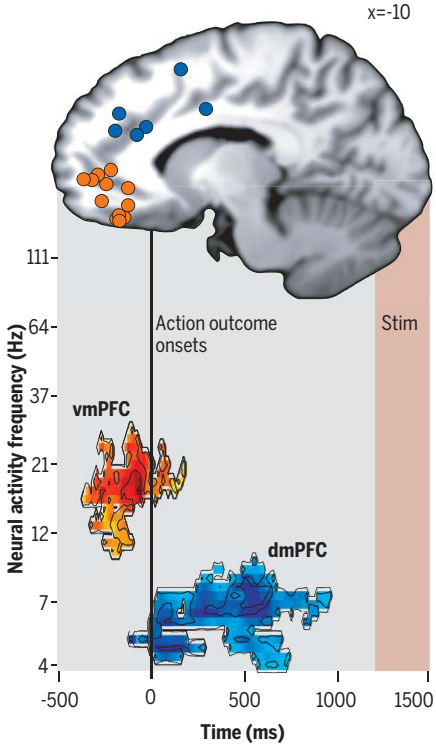
Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex

Philippe Domenech[1,3,4], Sylvain Rheims[5,6], Etienne Koechlin[1,2,7]*

INTRODUCTION: Everyday life often requires arbitrating between pursuing an ongoing action plan by possibly adjusting it versus exploring new action plans instead. Resolving this so-called exploitation-exploration dilemma is critical to gradually build a repertoire of action plans for efficient adaptive behavior in uncertain, changing, and open-ended everyday environments. Previous studies have shown that its resolution primarily involves the medial prefrontal cortex (mPFC). Human functional magnetic resonance imaging shows that activations in the ventromedial PFC (vmPFC) reflect the subjective value of the ongoing plan according to action outcomes, whereas the dorsomedial PFC (dmPFC) exhibits activations when this value drops and the plan is abandoned for exploring new ones. However, the neural mechanisms that resolve the dilemma and make the decision to exploit versus explore remain largely unknown.

RATIONALE: We addressed this issue by recording neuronal activity in participants using intracranial electroencephalography while they were performing a task that induced systematic exploitation-exploration dilemmas in an uncertain, changing, and open-ended environment. Participants were six epileptic patients with electrodes implanted in the vmPFC and dmPFC (see the figure), who were eventually diagnosed with temporal or parietal lobe epilepsy with no impacts in the PFC. Using computational modeling, we identified from participants' behavior the so-called stay trials, when participants adjusted and exploited their ongoing action plan through reinforcement learning, and the switch trials, when action outcomes instead led participants to covertly switch away from this plan and explore new ones in the following trials. We then analyzed vmPFC and dmPFC neural activity in both stay and switch trials.

RESULTS: vmPFC neural activity in the high-gamma frequency band (>50 Hz) that reflects local processing was found to encode outcome expectations after action selection. This vmPFC high-gamma activity further encoded the prior and posterior reliability of the ongoing action plan relative to action outcomes, which, according to the computational model, subserved the arbitration between exploiting and exploring. Notably, this reliability encoding yielded vmPFC activity to proactively flag forthcoming action outcomes as potential triggers to explore rather than as learning signals to exploit. Preceding the occurrence of action outcomes, switch trials—unlike stay trials—witnessed an increased neural activity in the beta frequency band (13 to 30 Hz) that reflects top-down neural processing (see the figure). Following action outcomes in switch compared with stay trials, dmPFC neural activity then decreased in the theta frequency band (4 to 8 Hz), which indicates that the dmPFC was then configured to respond to action outcomes according to this vmPFC proactive construct. In stay trials, outcome expectations encoded in the vmPFC were transmitted to the dmPFC, so that from 300 ms after action outcomes, dmPFC neural activity in the high-gamma frequency band encoded the reward prediction error (i.e., the discrepancy between expected and actual outcomes that scales reinforcement learning). In switch trials, by contrast, this encoding was disrupted through reconfiguring dmPFC activity in the alpha frequency band (8 to 12 Hz) to release the inhibition bearing upon alternative action plans from 250 ms after action outcomes.



Action outcomes triggering exploration. Neural activity around outcome onsets in switch compared with stay trials recorded in ventromedial (orange, vmPFC) and dorsomedial (blue, dmPFC) prefrontal electrodes implanted in the six patients. Electrode localizations are shown on a canonical sagittal brain slice [Montreal Neurological Institute (MNI) coordinate: x = −10], and neural activity is shown against time according to its spectral decomposition. vmPFC activity reflecting top-down neural processing increased and proactively flagged action outcomes as potential triggers to explore rather than as learning signals to exploit. dmPFC activity followed action outcomes triggering exploration through reconfiguring neural processing. Stim, stimulus.

CONCLUSION: The medial PFC resolves exploitation-exploration dilemmas through a predictive coding mechanism that was originally proposed for perception. The vmPFC monitors the reliability of the ongoing action plan to proactively set the functional signification of forthcoming action outcomes as either learning signals to exploit or potential triggers to explore. The dmPFC responds to action outcomes according to this functional construct, yielding to either stay and adjust the ongoing plan through reinforcement learning or switch away from this plan to explore new ones. This predictive coding mechanism has the advantage of speeding up the abandonment of ongoing action plans and preventing action outcomes that trigger exploration from inappropriately acting as learning signals. These findings support the idea that predictive coding also operates within the prefrontal executive system and constitutes a general mechanism that underlies information processing across the cerebral cortex. In perceptual neural systems, predictive coding operates so that observers' prior beliefs about a scene alter how they perceive the scene. Our findings suggest that within the prefrontal executive system, predictive coding operates by proactively altering the functional signification of behavioral events according to the agents' beliefs about their own behavior. ■

[1]Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France. [2]Université Paris Sciences et Lettres (PSL) Research University, Ecole Normale Supérieure, Paris, France. [3]Paris Brain Institute, Paris, France. [4]APHP, Groupe Hospitalier Henri Mondor, DMU Psychiatry, Department of Neurosurgery, Université Paris Est Créteil, Créteil, France. [5]Department of Functional Neurology and Epileptology, Hospices Civils de Lyon, University of Lyon, Lyon, France. [6]Lyon's Neuroscience Research Center, INSERM-U1028, CNRS-UMR 5292, Lyon, France. [7]Université Pierre et Marie Curie, Paris, France.
*Corresponding author. Email: etienne.koechlin@upmc.fr
Cite this article as P. Domenech *et al.*, *Science* **369**, eabb0184 (2020). DOI: 10.1126/science.abb0184

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb0184

RESEARCH ARTICLE

NEUROSCIENCE

Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex

Philippe Domenech[1,3,4], Sylvain Rheims[5,6], Etienne Koechlin[1,2,7]*

Everyday life often requires arbitrating between pursuing an ongoing action plan by possibly adjusting it versus exploring a new action plan instead. Resolving this so-called exploitation-exploration dilemma involves the medial prefrontal cortex (mPFC). Using human intracranial electrophysiological recordings, we discovered that neural activity in the ventral mPFC infers and tracks the reliability of the ongoing plan to proactively encode upcoming action outcomes as either learning signals or potential triggers to explore new plans. By contrast, the dorsal mPFC exhibits neural responses to action outcomes, which results in either improving or abandoning the ongoing plan. Thus, the mPFC resolves the exploitation-exploration dilemma through a two-stage, predictive coding process: a proactive ventromedial stage that constructs the functional signification of upcoming action outcomes and a reactive dorsomedial stage that guides behavior in response to action outcomes.

Everyday life frequently necessitates arbitration between the pursuit of an ongoing action plan with possible adjustments versus the exploring of a new plan. Resolving this so-called exploitation-exploration dilemma is critical for efficient adaptive behavior in uncertain, changing, and open-ended everyday environments (*1*, *2*) (supplementary note), and it primarily involves the medial prefrontal cortex (mPFC) (*3*–*8*). Human neuroimaging shows that activation in the ventromedial PFC (vmPFC) reflects the subjective value of the ongoing plan according to action outcomes, whereas the dorsomedial PFC (dmPFC) exhibits activation when this value drops and the plan is abandoned for exploring new ones (*6*). However, the neural mechanisms that resolve the dilemma and make the decision to exploit versus explore remain largely unknown. According to the classical view, the brain processes information in a feed-forward manner from stimuli to responses; namely, the vmPFC and dmPFC implement a bottom-up, reactive process evaluating the ongoing plan only after experiencing action outcomes to decide between further exploiting versus exploring. Alternatively, however, predictive coding that has been proposed for perception (*9*–*11*) could extend to the prefrontal executive function. Accordingly, the vmPFC and dmPFC might implement a top-down, proactive process wherein the ongoing plan is evaluated before experiencing action outcomes to prospectively code upcoming action outcomes as either learning signals to improve the ongoing plan or as potential triggers to explore new plans rather than exploit the ongoing plan. We addressed this issue by recording neuronal activity in six participants while they were performing a task that induced systematic exploitation-exploration dilemmas in an uncertain, changing, and open-ended environment (*2*, *6*) (Fig. 1).

[1]Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France. [2]Université Paris Sciences et Lettres (PSL) Research University, Ecole Normale Supérieure, Paris, France. [3]Paris Brain Institute, Paris, France. [4]APHP, Groupe Hospitalier Henri Mondor, DMU Psychiatry, Department of Neurosurgery, Université Paris Est Créteil, Créteil, France. [5]Department of Functional Neurology and Epileptology, Hospices Civils de Lyon, University of Lyon, Lyon, France. [6]Lyon's Neuroscience Research Center, INSERM-U1028, CNRS-UMR 5292, Lyon, France. [7]Université Pierre et Marie Curie, Paris, France.
*Corresponding author. Email: etienne.koechlin@upmc.fr

Experimental paradigm

Participants were patients with drug-resistant focal epilepsy who had been stereotactically implanted with multilead electroencephalography (EEG) depth electrodes (*12*, *13*), which passed notably through the vmPFC and dmPFC (Fig. 1A and table S1). Participants were all eventually diagnosed with temporal or parietal lobe epilepsy with no impact observed on the PFC (see the Materials and methods section). Participants responded to visually, successively presented digits and searched for digit-response combinations by trial and error. Feedback was binary (positive versus negative) and stochastic (10% noise). Combinations changed episodically and unpredictably (every 33 to 57 trials), thereby dividing trial series into successive latent episodes associated with distinct combinations. Thus, the task induced participants to constantly arbitrate according to feedback between two options: (i) staying with the same presumed combination by possibly adjusting it or (ii) abandoning it to explore new combinations. In every trial, participants' responses could be correct (chance level, 25%), perseverating (incorrect in the current episode but correct in the preceding episode; chance level, 25%), or ancillary (neither correct nor perseverative; chance level, 50%) (Materials and methods). Overall, participants performed notably below the statistical optimum but similarly to healthy participants who had been previously tested in the same task (*2*, *6*). Correct response rates increased from ~5% at episode onsets to ~80%, perseverative response rates decreased from ~80% at episode onsets to ~5% after 25 trials, and ancillary response rates increased from ~15% at episode onsets, peaked to ~40% about 6 trials later, and returned to ~15% about 25 trials later (Fig. 2A).

Identifying covert switches into exploration

To determine when participants switched from exploiting presumed combinations to exploring new ones, we leveraged previous studies that have shown that these switches derive from an online algorithm approximating the optimal adaptive process (*2*, *6*). According to this model, presumed combinations form action plans associating digits, responses, and expected feedback. These plans are monitored online: The algorithm probabilistically infers the reliability value of presumed plans—i.e., the belief λ_i that plan i matches the current true combination—given observed action outcomes and the possibility that no monitored plans match. When plan i_0 is more likely matching than not matching the current combination—i.e., $\lambda_{i_0} > 1 - \lambda_{i_0}$ (or equivalently, $\lambda_{i_0} > 0.5$)—the plan is said to be reliable and the others are necessarily unreliable (because $\Sigma_i \lambda_i < 1$). The algorithm is then in the exploitation state: The reliable plan constitutes the actor that guides ongoing behavior and learns from feedback through reinforcement learning (RL) processes. When the actor becomes unreliable and all other monitored plans remain unreliable (all $\lambda_i < 0.5$), the algorithm switches into the exploration state: A new presumed plan is formed from mixing previously learned plans stored in long-term memory, and this new plan is used as a provisional actor guiding behavior and learning from feedback. The algorithm eventually returns to the exploitation state, when this provisional actor or another monitored plan becomes reliable. In the former case, the provisional actor is then consolidated in long-term memory, whereas it is disbanded in the latter case (see supplementary text for the full model description).

This model closely fitted and reproduced participants' performances (Fig. 2A), as has been explained and shown in previous studies (*2*, *6*, *14*). The model especially reveals when participants covertly switched from exploitation to exploration. After episode changes, the model switched, as expected, from exploitation to exploration at variable time points ranging from two to eight trials after episode onsets [95% confidence interval (CI)]. In these switch trials, feedback induced posterior actor reliability to decrease and drop below the reliability

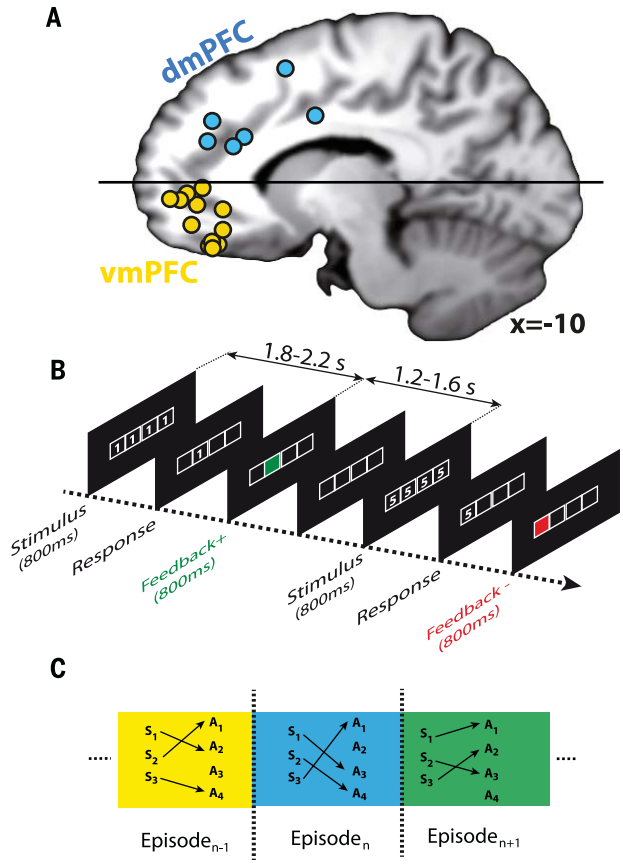


Fig. 1. Experimental protocol. (**A**) Localization of transverse electrodes within the medial PFC across participants, reconstructed on a canonical T1-weighted sagittal MRI brain slice [Montreal Neurological Institute (MNI) coordinate: x = −10]. Subgenual (yellow) and dorsogenual (blue) electrodes were ascribed to the vmPFC and dmPFC, respectively. The vmPFC and dmPFC comprised *N* = 13 and *N* = 12 electrode contacts, respectively. See fig. S1 for exact localizations in every participant. (**B**) Behavioral protocol. Trials started with the display of one out of three possible stimuli (digits; duration: 800 ms). Participants responded by pressing one out of four response buttons. Participants then received a positive or negative feedback (digits turned green or red, respectively; duration: 800 ms), depending on the current correct digit-button combination. Stimuli-feedback asynchrony was jittered (range: 400 ms). (**C**) Episode structure of the behavioral protocol. Current correct digit-button combinations episodically changed after an unpredictable number of trials, thereby defining successive latent episodes. There were no overlaps in digit-button associations ($S_i \rightarrow A_j$) between two successive episodes.

threshold ($\lambda = 0.5$), which led the model to start exploring in the subsequent trials. Realigning both the model's and participants' performances on switch trials rather than episode onsets (Fig. 2B) shows that both the model's and participants' responses were unaffected by episode changes up to these switch trials: Most responses remained perseverative (~85%), whereas residual responses were randomly distributed across ancillary and correct responses (~10 and ~5%, respectively). By contrast, in trials immediately after switch trials, both the model's and participants' perseverative responses abruptly dropped off to ~40%, whereas ancillary and correct responses abruptly increased close to their chance level (~40 and ~20%, respectively) (Fig. 2B). This abrupt algorithmic transition was also independently observed in participants' reaction times, which suddenly increased by ~100 ms in the trials immediately after switch trials (Fig. 2B)—an effect that could neither be ascribed to increases of response shifts after switch trials (fig. S5) nor to the occurrence of unexpected feedback (fig. S6). Thus, in switch trials, feedback induced participants to switch away from their ongoing plan to explore new plans in the next trials. In stay trials, by contrast, the model reveals that participants' behavior instead derived from RL processes—i.e., participants stayed with the same actor plan that adapted to external contingencies through RL processes—particularly when eventually returning from exploration to exploitation. Overall, switch trials occurred in the model contingent upon posterior rather than prior actor reliability relative to the actual feedback that participants observed. Switch trials predicted that participants behaviorally switched into exploration in the trials that immediately followed. Accordingly, participants covertly switched into exploration posterior to this feedback and before participants' responses that immediately followed.

Neural mechanisms inferring actor reliability

Using the standard time-frequency decomposition (*15*), we investigated whether local neural processing in the vmPFC and/or the dmPFC probabilistically infers and tracks actor reliability predicted by the model. We extracted neural activity in the high-gamma frequency band (50 to 150 Hz) reflecting local neuronal spiking (*16*) at each time point within the trials (Materials and methods). We entered this activity into a single multiple regression analysis performed over all trials that included model variables as within-subject regressors, including the following: (i) prior and posterior actor reliability relative to feedback, orthogonalized in that order so that the second regressor captures only reliability updates from feedback; (ii) the same reliability regressors for alternative plans monitored along the actor according to the model; and (iii) chosen values, i.e., the RL value of chosen actions, which, as feedback were binary, also measured positive feedback likelihoods involved in computing posterior from prior actor reliability. Additional regressors were included to remove potential confounding factors (Materials and methods).

vmPFC high-gamma activity linearly encoded prior actor reliability along virtually the whole trial epoch including intertrial intervals (Fig. 3). Moreover, ~300 ms after participants' responses, this activity further started encoding chosen values linearly. This encoding gradually increased up to feedback occurrences and then decreased afterward, which indicates that the vmPFC also anticipated when feedback occurred. Finally, ~350 ms after feedback occurrences, the activity further started encoding posterior actor reliability linearly. We detected no significant correlations with the reliability of alternative strategies (fig. S7), confirming the vmPFC specific role in monitoring the action plan driving ongoing behavior (*6*). In the dmPFC by contrast, high-gamma activity was uncorrelated with any preceding regressors (Fig. 3 and fig. S7). In both regions, finally, we found no significant correlations with neural activity in other frequency bands.

vmPFC and the prospect of exploration

To understand how monitoring of actor reliability leads to switch from exploitation to exploration, we compared neural activity locked on feedback onset between switch and stay trials. In the vmPFC, switch compared with neighboring stay trials elicited only a significant increase of neural activity in beta-band frequencies (13 to 30 Hz), which started at ~350 ms, peaked ~70 ms before feedback onset, and

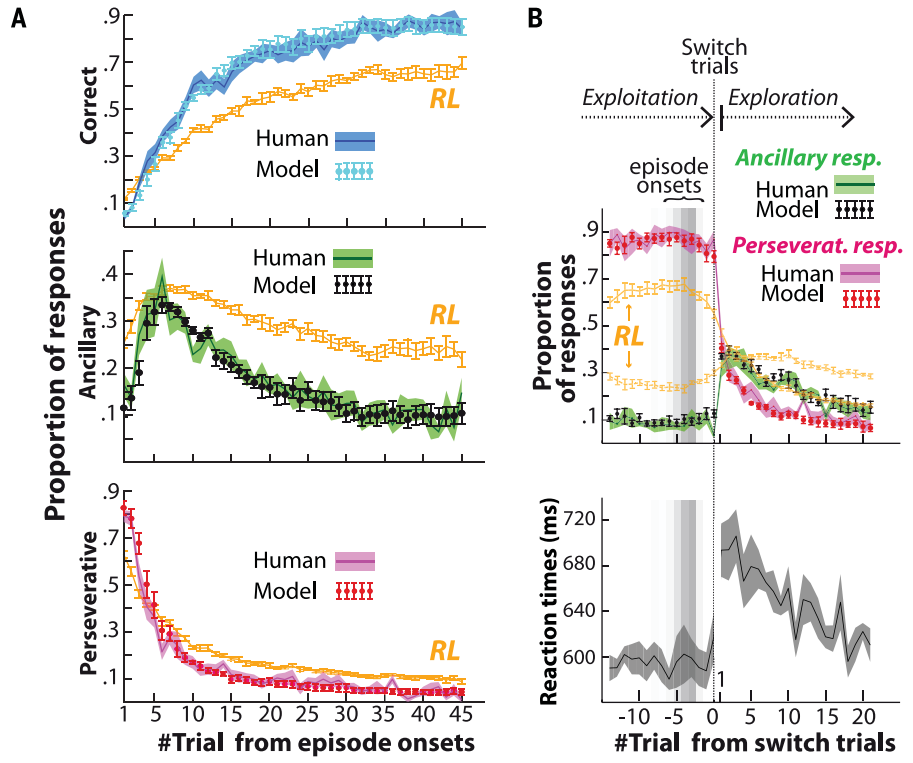


Fig. 2. Behavioral performances. (**A**) Proportion of correct, ancillary, and perseverative responses (summing up to 1) from participants and the model, according to the number of trials after episode onsets (i.e., combination changes). Ancillary and perseverative responses are both incorrect, but perseverative responses further correspond to correct responses in the preceding episode. Both the model and participants adapted much faster to combination changes than a Rescorla and Wagner's RL model that was fitted on participants' data (orange). (**B**) Participants' responses and model predictions (top) along with participants' reaction times (bottom) realigned on switch trials occurring in the model rather than on episode onsets. Orange lines show RL predictions. All model predictions are computed in every trial given participants' responses in previous trials. Error bars represent SEMs across participants. See Materials and methods for modeling details and tables S2 and S3 for model parameters.

vanished ~150 ms after feedback onset (Fig. 4A). This prefeedback effect is consistent with the documented role of beta-band activity in anticipating upcoming stimuli (*17*–*19*). The effect abruptly appeared in switch trials and was neither present in previous and subsequent neighboring trials (Fig. 4B) nor in stay trials that lead to either negative or positive feedback (Fig. 5). Thus, the effect was unlikely to reflect differences in reward or feedback expectations between stay and switch trials: In switch and immediately preceding stay trials, positive feedbacks were equally frequent and chosen values (or feedback likelihoods) were virtually identical (Fig. 4D). The effect was also unlikely to reflect actor reliability, which gradually decreased, whereas beta-band activity remained constant along stay trials preceding switch trials (Fig. 4D). We also dismissed the possibility that this prefeedback effect reflects the commitment to switch into exploration. The analysis of behavioral performances reported above indicates that participants covertly switched into exploration contingent upon and, consequently, posterior to these feedbacks. However, this analysis indicates that in switch compared with stay trials, the following prefeedback event occurred: Prior actor reliability monitored in the vmPFC (see Fig. 3) approached the 0.5 reliability threshold closely enough that upcoming feedback could cause posterior actor reliability to cross the threshold and trigger exploration. We therefore concluded that the prefeedback effect observed in the vmPFC reflects this event—namely, the sudden possibility that upcoming feedback may cause posterior actor reliability to cross the threshold or, equivalently, the fact that prior actor reliability is close enough to the threshold. Accordingly, the vmPFC appears to evaluate actor reliability relative to the threshold before feedback occurrences. This prospectively flags upcoming feedback in switch trials as potential triggers committing to exploration rather than as regular learning signals serving to adjust the ongoing actor plan.

dmPFC responses to exploration triggers versus learning signals

In the dmPFC, by contrast, switch compared with neighboring stay trials exhibited a significant differential neural activity only after feedback onset. The activity started at feedback onset and lasted up to ~200 ms after feedback offset (i.e., lasting ~1000 ms). This postfeedback effect occurred in theta-band frequencies (4 to 8 Hz) and slightly extended to alpha-band frequencies (8 to 12 Hz) (Fig. 4C). Although this postfeedback activity remained unchanged in stay trials preceding and after switch trials, it abruptly decreased in switch trials (Fig. 4B) (*20*, *21*). For the same reasons as above, this abrupt postfeedback effect could neither be ascribed to any variations in reward-feedback expectations or in actor reliability across stay and switch trials (Fig. 4D) nor to differences in reward prediction errors (RPEs), because these errors were virtually identical in switch and immediately preceding stay trials (Fig. 4D). Moreover, actor reliability was unrelated to dmPFC neural activity and was updated in the vmPFC only ~350 ms after feedback onset. Consequently, the effect emerging at feedback onset was unlikely to reflect a bottom-up, reactive process comparing posterior actor reliability with the 0.5 reliability threshold and leading to the decision to explore in switch trials. However, when this dmPFC postfeedback effect emerged at feedback onset, the vmPFC prefeedback effect started declining. This suggests that the vmPFC proactively configured the dmPFC to process feedback differently in stay and switch trials—i.e., as learning signals versus exploration triggers, respectively. Consistent with this interpretation, the vmPFC prefeedback effect arose in beta-band frequencies known to convey top-down, predictive neural processing (*22*–*24*), whereas the dmPFC postfeedback effect arose in theta-band frequencies considered as reflecting the configuration of PFC neural networks underpinning behavioral control (*25*–*28*).

In stay trials, dmPFC neural responses to positive and negative feedback exhibited a single common feature starting at feedback onset—namely, a strong increased activity in alpha-band frequencies that vanished ~600 ms later (Fig. 6A). dmPFC-centered alpha-band activities are thought to drive the inhibition of neural representations that are irrelevant to ongoing behavior (*29*–*33*), thereby favoring the maintenance of the ongoing actor plan and its adjustment in response to feedback. Consistently, dmPFC neural responses to positive and negative feedback exhibited a second common feature from ~200 to ~600 ms after feedback onset, which corresponded to the signature of RL processes—namely, a strong increase of high-gamma activity that correlated with unsigned RPEs (i.e., the discrepancy between chosen values and actual feedback) (*34*) (Fig. 6, A, C, and D). Thus, dmPFC neuronal responses to feedback in stay trials transiently encoded unsigned RPEs scaling RL processes. As expected, this postfeedback RPE encoding

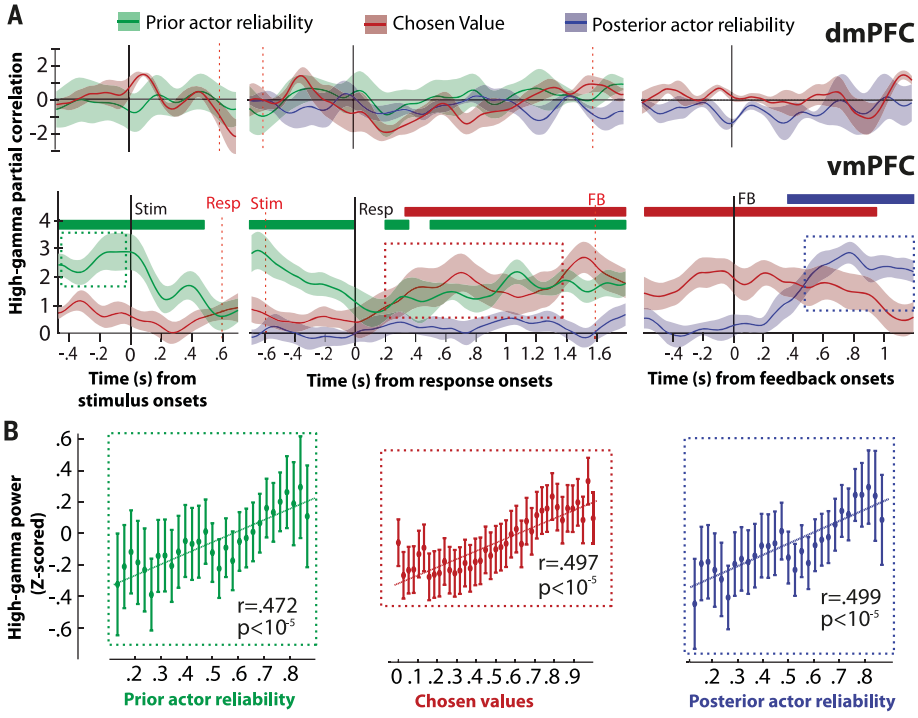


Fig. 3. Neural encoding of model variables in the medial PFC. (**A**) Time courses of partial correlation coefficients (betas) at each time point within trials between high-gamma neural activity (local field potentials >40 Hz), averaged over electrode contacts within the dmPFC and vmPFC and model variables (prior and posterior actor reliability, chosen values) derived from multiple regression analyses across all trials. Time courses are locked on stimulus onsets (left), participants' responses (middle), and feedback onset (right). Red vertical dashed lines show the average onsets of trial events. Shaded areas represent SEMs across contacts. Thick horizontal colored bars indicate statistical significance at $P < 0.05$, corrected for multiple comparisons at cluster level. The dmPFC exhibited no significant correlations. Resp, response; Stim, stimulus; FB, feedback. (**B**) vmPFC high-gamma activities (Z-scored for each electrode contact across trials before averaging) averaged over specific time windows [dashed boxes in (A)] and plotted against prior actor reliability (left), chosen values (middle), and posterior actor reliability (right). Error bars are SEMs across electrode contacts. Linear regression coefficients, r, are shown with P values. See fig. S2 for data from one vmPFC individual contact.

in the dmPFC was functionally connected with the prefeedback encoding of chosen values that we observed in the vmPFC (see the Psychophysiological interactions section in the Materials and methods; Fig. 7). vmPFC prefeedback and dmPFC postfeedback high-gamma activities were indeed correlated, and this cross-temporal correlation increased with chosen values (Fig. 7). This indicates that the chosen value representation in the vmPFC is transmitted to the dmPFC to subserve RPE computations. Consistently, the cross-temporal correlation further decreased when unsigned RPEs increased [i.e., decreased with the discrepancy between expected feedbacks (chosen values) encoded in the vmPFC and actual feedbacks processed in the dmPFC] (Fig. 7). By contrast, the cross-temporal correlation was unrelated to actor reliability (Fig. 7), which is in agreement with the absence of any dmPFC activity related to action plan reliability (Fig. 3 and fig. S7).

In switch trials, this cascade of dmPFC responses to feedback was disrupted. First, dmPFC theta-band activity in switch compared with stay trials decreased from feedback onset; second, the dmPFC alpha-band activity was notably reduced and rapidly vanished ~250 ms later (Fig. 6, A and B); third, this suppression was accompanied by a drop-off of dmPFC high-gamma activity observed in stay trials from ~200 to ~600 ms after feedback onset, which further ceased coding for any RPEs (Fig. 6, A and D). Thus, the dmPFC processed feedback in switch trials unlike learning signals, with the alpha-band response suppression presumably releasing the inhibition bearing upon neural representations that are irrelevant to the ongoing actor plan and preventing RL processes from adjusting this action plan. The dmPFC thus appears to process feedback in switch trials as favoring the covert emergence of neural representations forming new action plans to explore from ~250 ms after feedback onset.

Resolving exploitation-exploration dilemmas through predictive coding

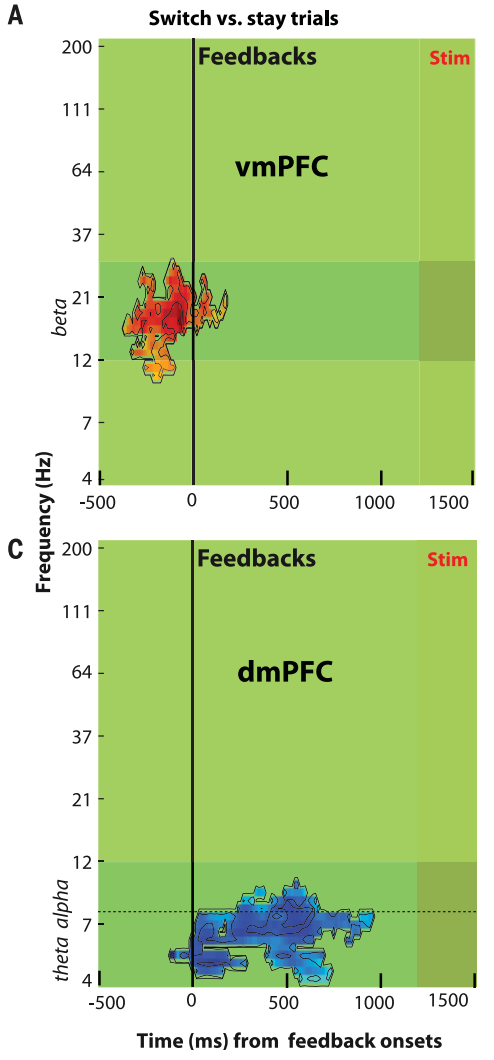
vmPFC neural activity in gamma-band frequencies infers and tracks the reliability of the ongoing action plan according to action outcomes. After the action, it proactively flags (through beta-band frequencies) upcoming outcomes as either learning signals to better exploit this plan or potential triggers to explore new ones. According to this functional construct, dmPFC activity in theta-band frequencies appears to reflect the dmPFC configuration to respond to action outcomes. The dmPFC response to outcomes that are flagged as actual triggers then appears to realize the switch into exploration through the suppression of neural activity in alpha-band frequencies. This favors the emergence of neural representations that form new action plans and, through inhibiting dmPFC high-gamma activity–scaling RL processes, prevents the ongoing plan from adjusting through RL. Thus, the medial PFC resolves exploitation-exploration dilemmas through a top-down, predictive coding process from the vmPFC to the dmPFC. This predictive coding process has the advantage of speeding up the abandonment of the ongoing action plan and preventing action outcomes that trigger exploration from inappropriately acting as learning signals.

Predictive coding, which was originally developed to describe perceptual cortical processes (*9*–*11*, *35*), may also play a role in prefrontal executive processes. In perceptual predictive coding, observers' prior beliefs about a scene alter how they perceive the scene. Our findings suggest that within the prefrontal executive system, predictive coding proactively alters the functional signification of behavioral events according to the agents' beliefs about their own behavior.

Materials and methods

Participants

Six patients with drug-resistant focal epilepsy (one female; age range: 25 to 49 years old; see table S1) from the Department of Functional Neurology and Epileptology at University Hospital of Lyon participated in the present study. The participants belonged to a group of patients, which were stereotactically implanted with intracranial EEG depth electrodes to locate epileptic foci because noninvasive methods were unsuccessful (*13*, *36*). Implantation sites were selected according to clinical requirements, independent of the present study. We recruited the patients with at least one electrode implanted in the vmPFC and who were eventually diagnosed with temporal lobe ($n = 5$) or parietal lobe ($n = 1$) epilepsy with no electrophysiological impacts observed in the PFC. The study (DSI-SEEG protocol, NCT02869698) was approved by the Institutional Review Board (ANSM no. 2009-A00239-48) and National French



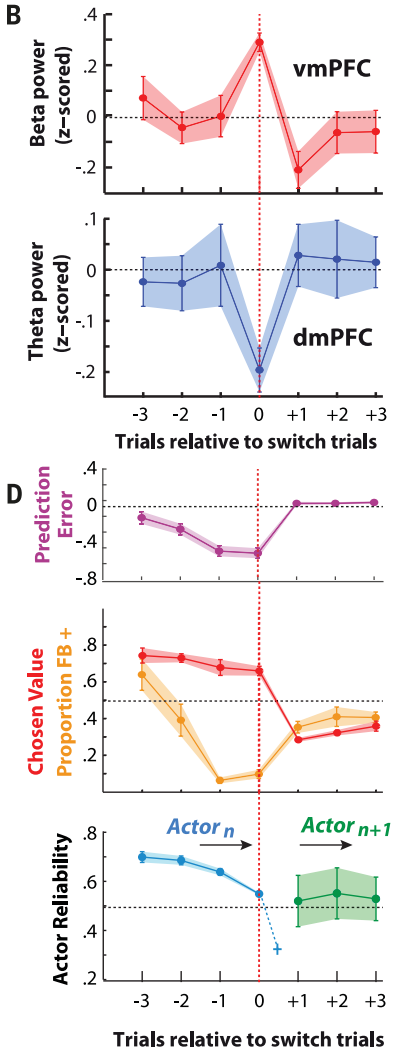


Fig. 4. Medial PFC neural activity associated with switch compared with stay trial feedback. (**A** and **C**) Time-frequency analyses (*T*-value maps) of neural local field potentials in switch compared with neighboring stay trials (from −2 to +2 trials relative to switch trials) locked on feedback and averaged over vmPFC (A) and dmPFC (C) electrode contacts. Power increases (positive *T* values) and decreases (negative *T* values) are shown in red and blue, respectively. *T*-value maps are thresholded at $P < 0.05$, corrected for multiple comparisons (cluster-level, family-wise error corrections). Black contours delimit statistical thresholds from $P < 0.05$ to $P < 5.0 \times 10^{-6}$. Vertical shaded areas indicate onset windows of stimuli from the next trial. See fig. S3 for the unthresholded maps. (**B**) Power amplitudes in switch and neighboring stay trials averaged over the vmPFC beta-band and dmPFC theta-band clusters shown in (A) and (C), respectively. Power amplitudes were *Z*-scored in each electrode contact before averaging. Error bars are SEMs over trials. (**D**) Signed prediction errors, chosen values, proportion of positive feedback (FB+), and prior actor reliability in switch and neighboring stay trials. In switch trials, posterior actor reliability dropped below the 0.5 reliability threshold (blue cross), and a new actor is formed to guide exploration in subsequent trials. Error bars are SEMs over participants.

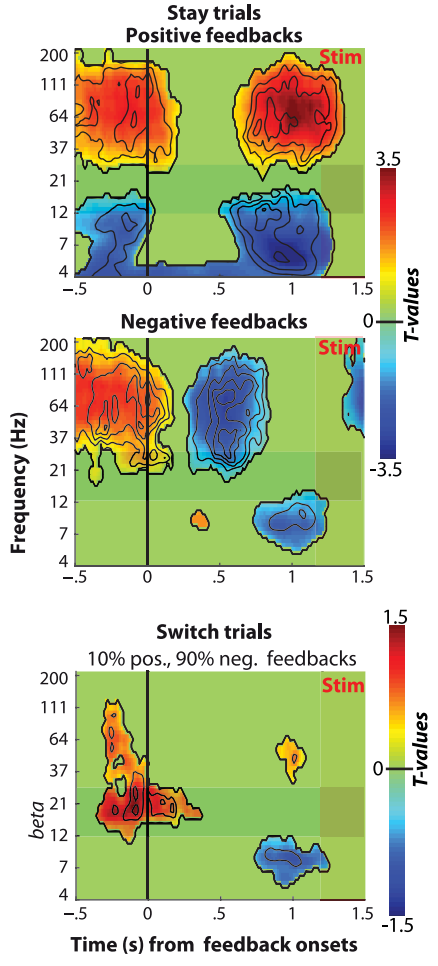


Fig. 5. vmPFC neural activity associated with response feedback. Time-frequency analyses (*T*-value maps) of vmPFC local field potentials relative to trial grand averages (from −2 s to +5 s relative to stimulus onsets), averaged across electrode contacts and locked on feedback onset in stay (top: positive feedback; middle: negative feedback) and switch (bottom) trials. Switch trials comprised 10% positive and 90% negative feedback. Maps are thresholded at $P < 0.05$, corrected for multiple comparisons (cluster-wise, family-wise error corrections). Increases (positive *T* values) and decreases (negative *T* values) are shown in red and blue, respectively. Black contours delimit statistical thresholds from $P < 0.05$ to $P < 0.000005$. Vertical shaded areas indicate onset windows of stimuli from the next trial. See fig. S4 for the unthresholded maps.

science ethic committee (CPP 09-CHUG-12, no. 0907). All participants volunteered to participate and provided written informed consent before participation.

Intracranial electroencephalography

We collected intracranial electroencephalography (iEEG) recordings from the six patients. They were chronically implanted with 12 to 15 stereotactic multilead depth electrodes for 2 to 3 weeks. These semirigid electrodes (DIXI Medical Instrument) have a diameter of 0.8 mm and, depending on the target structure, consisted of 5 to 15 linearly arranged contact leads (2 mm wide), with a 1.5-mm gap between two consecutive leads. Overall, we recorded from 929 contacts distributed across 81 depth electrodes, among which 21 depth electrodes (185 contacts) were located in the PFC. All these electrodes were implanted orthogonal to the interhemispheric plane with the deepest contacts located in the medial PFC. Over the six patients, 13 electrode contacts were localized in the vmPFC and 12 in the dmPFC (see Fig. 1A and fig. S1). T1-weighted anatomical magnetic resonance imaging (MRI) [three-dimensional gradient-recalled echo (3D GRE); resolution, 1 mm^3; matrix size, 256 voxels by 256 voxels] were acquired before and after surgery. Exact contact locations were manually determined

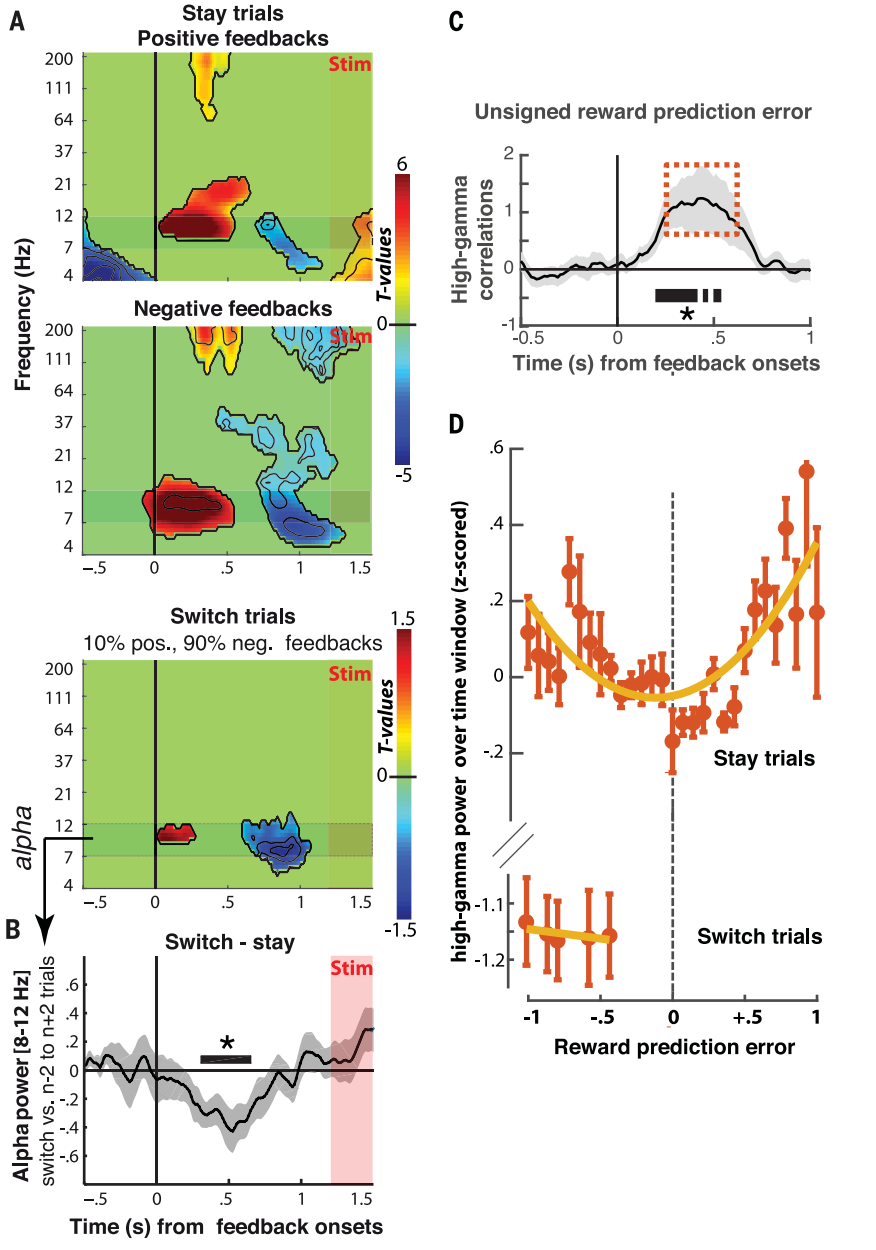


Fig. 6. dmPFC neural activity associated with response feedback. (**A**) Same as in Fig. 5 but for local field potentials averaged across dmPFC electrode contacts. See fig. S4 for the unthresholded maps. (**B**) Time courses of dmPFC alpha-band power in switch compared with neighboring stay trials (from −2 to +2 trials relative to switch trials) locked on feedback onset. (**C**) Correlation between dmPFC high-gamma neural activity and unsigned RPEs plotted against time from feedback onset. Shaded areas in (B) and (C) are SEMs across dmPFC electrode contacts. Horizontal black bars in (B) and (C) indicate statistical significance at $P < 0.05$, corrected for multiple comparisons (cluster-wise, family-wise error correction). (**D**) High-gamma activity over time window shown in (C) (orange) plotted against RPEs in stay and switch trials. Error bars are SEMs across binned trials. Lines show second-order polynomial regressions in stay trials (df = 340; linear $P = 0.34$; quadratic $P < 0.0001$) and in switch trials (df = 212; both linear and quadratic $P > 0.89$).

from the position of the corresponding artifact relative to the main anatomical landmarks on the postsurgery MRI. We recorded iEEG ~8 days after surgery (8.6 days ±1.4 SEM; see table S1) using a 128-channels video-EEG monitoring system (Micromed; sampling rate, 512 Hz). iEEG data were bandpass filtered online (0.1 to 200 Hz).

Experimental paradigm

The participants performed a variant of the Wisconsin Card Sorting Test, in which they learned combinations between digits and response buttons by trial and error. Combinations changed episodically and unpredictably. The experimental paradigm is identical to that used in previous studies testing healthy participants within and outside MRI scanners (*2*, *6*). In the present study, we only adjusted the event timing and jittering to iEEG constraints.

Four white boxes representing four response buttons were displayed on a black background at the center of a screen (Fig. 1B). Each trial started with the display of a white digit (out of three possible digits) in every box during 800 ms. The patients responded to this stimulus by pressing one of four buttons (response box: Cedrus Lumina, LU444-RH). Patients had to use the same finger to press the same button throughout the experiment. If a response occurred within 1500 ms from the stimulus onset, all displayed digits disappeared between 1800 and 2200 ms after the stimulus onset, except the digit displayed in the box related to the pressed button: If the participant pressed the correct button, this digit instead turned green with 90% probability (positive feedback) and turned red with 10% probability (negative feedback). If the participant pressed another button, positive and negative feedback probabilities were reversed. Response feedbacks were thus stochastic. If no responses occurred within 1500 ms from the stimulus onset, all digits disappeared and were replaced by an uninformative, neutral feedback (dashes in every box). Response feedbacks were presented during 800 ms. After 800 ms, the feedback disappeared, leaving the four boxes empty. The next trial started after a delay of 1200 to 1600 ms from the feedback onset. Participants were explicitly instructed that in every trial, each digit was associated with only one correct response button, and that distinct digits were associated with distinct responses (see Fig. 1C). Participants were also told that response feedbacks were not fully reliable (they were however not informed about the exact feedback probabilities). Finally, participants were informed that digit-response combinations could change episodically and unpredictably. No additional instructions were provided to participants. We refer to series of trials with no combination changes as episodes. Episode lengths pseudorandomly ranged between 33 and 57 trials. When combinations changed, every digit-response association was changed (Fig. 1C)

Overall, the experiment included two behavioral sessions administered on successive days. Each session included 1011 trials comprising 24 episodes. Each session included five short breaks occurring within episodes. After each break, the last digit-response combination was used again during six to nine trials, so that breaks were unrelated to combination changes (participants were explicitly instructed that breaks were unrelated to combination changes). Stimuli were pseudorandomly drawn from the set {1,3,5} for one session and {2,4,6} for the other session, counterbalanced across participants. In one session, three distinct digit-response combinations were pseudorandomly

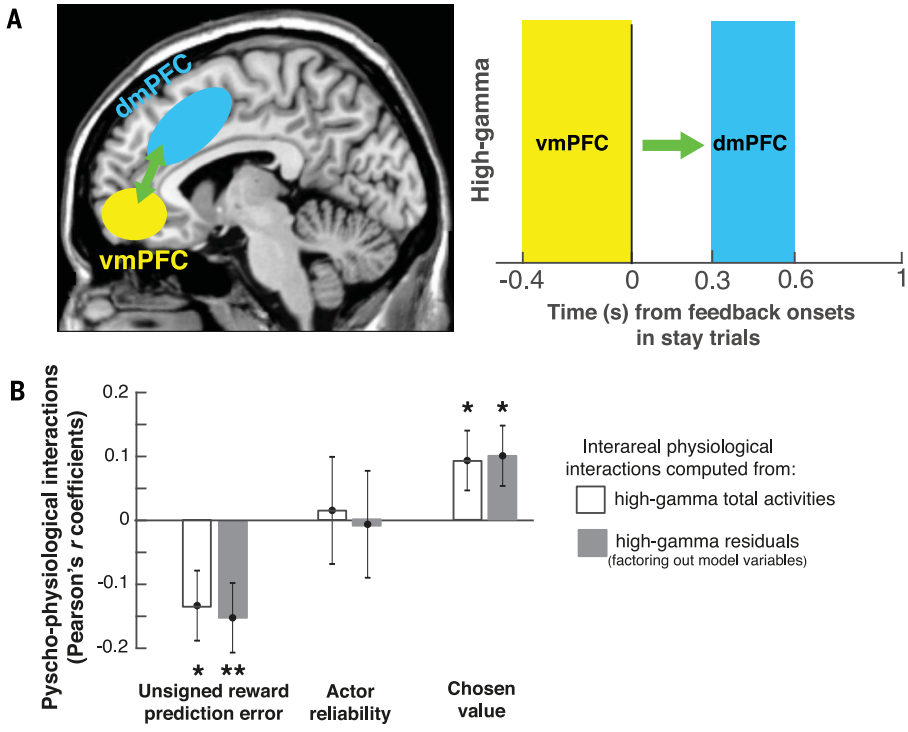


Fig. 7. Functional connectivity analysis from vmPFC to dmPFC high-gamma activities in stay trials. (**A**) The analysis investigated the physiological correlations between vmPFC high-gamma activity recorded from −400 ms to feedback onsets when this activity maximally encoded chosen values (Fig. 3A, right) and dmPFC high-gamma activity recorded from 300 to 600 ms after feedback onsets when this activity maximally encoded unsigned RPEs (Fig. 6C). Note that Fig. 3A shows vmPFC high-gamma activity to still encode chosen values and actor reliability over this postfeedback 300- to 600-ms time window. However, the analysis was restricted to prefeedback vmPFC activity to estimate the directional connectivity from vmPFC to dmPFC. Only the three participants with electrodes implanted in both the vmPFC and dmPFC were included in the analysis (see fig. S1). Physiological correlations were computed within every pair comprising one vmPFC and one dmPFC electrode contact. (**B**) Psychophysiological interactions (PPIs)—i.e., variations of physiological correlations with the model variables of interest, namely actor reliability, chosen value, and (unsigned) RPE. PPIs were computed for physiological correlations between high-gamma total activities as well as for correlations between high-gamma residual activities factoring out the influence of model variables on local neural activities. The graph shows that these physiological correlations increased with chosen values (*P < 0.039), decreased when unsigned RPEs increased (*P < 0.016 and **P < 0.006), and were unrelated to actor reliability (P < 0.85). Error bars are SEMs across contact pairs. As only the vmPFC encoded chosen values and actor reliability (Fig. 3A), the results indicate that vmPFC high-gamma activity conveys chosen value rather than actor reliability information to the dmPFC, with this chosen value information serving to compute RPEs that dmPFC high-gamma activity encoded (Fig. 6C).

repeated over episodes. In the other session, 24 distinct digit-response combinations were used across the 24 episodes (no combination repetitions). Order of sessions were counterbalanced across participants.

The day before the first session, the participants performed a short training session under supervision of the experimenter with a simplified protocol, using only two shapes (instead of three numbers) and two buttons (instead of four). This was done to ensure that participants correctly understood the task and experienced stochastic feedbacks and episode changes.

Computational model

In two previous studies (*2*, *6*), we proposed a computational model describing how the human prefrontal function guides adaptive behavior in uncertain, changing, and open-ended environments. Collins and Koechlin (*2*) have described how the proposed model derives and approximates optimal adaptive processes (Dirichlet processes mixture) and, using experimental protocols comprising the present one, accounts for human adaptive performance better than various alternative models. Donoso *et al.* (*6*) have provided additional behavioral and neuroimaging evidence supporting the model and have revealed the central role of the medial PFC in arbitrating between exploiting the actor action plan versus exploring alternative plans. The present study relies on the exact same model, which is described in the supplementary text.

Model fitting procedures

The model has seven continuous parameters and one discrete parameter (buffer size N comprising monitored action plans, see model description in the supplementary text). To fit the model to each participant choice behavior, we used the slice-sampling Markov chain Monte Carlo method (*37*). Accordingly, we drew 7 million samples from the parameter posterior (with uniform priors over parameter ranges; four independent sets of 400,000 samples for each buffer size ranging from 1 to 6; 50,000 burn-in). We computed the model log-likelihood for each participant by summing the log-likelihoods provided by the model's epsilon-softmax function over all trials. To obtain robust individual parameter estimates, we computed the log-likelihood weighted average over all the samples drawn using the buffer size with maximal posterior probability (table S2). Finally, we computed trial-by-trial Monte Carlo estimates of model variables (prior and posterior reliabilities, choice values, etc.) by randomly resampling from our full sample set (70,000 resamples) and computing their log-likelihood weighted average.

Preprocessing of iEEG data

In the present study, we only report data from electrode contacts located in the mPFC. We performed all preprocessing steps using EEGLAB (*38*). First, we removed electrical artifacts from power lines and medical equipment using frequency-domain regressions and a threshold based on a Thompson F statistic for artifact detection (Cleanline, http://chronux.org; bandwidth 2 Hz, sliding window length of 2- and 1.5-s steps). Then, iEEG data were epoched from −2000 ms before stimulus onset to +5000 ms after stimulus onset, and detrended using third degree polynomial fits (overall participants performed 2022 trials over two sessions leading to 2022 epochs). Each contact trace was subsequently re-referenced with respect to its nearest neighbor along the same electrode (bipolar derivation). We used bipolar rather than unipolar derivations because it allows for better signal artifact removal and achieves high spatial resolution (~3 mm^3) by canceling out contributions of distant sources which spread equally to recording sites (*39*). Epochs with possible epileptic spikes, electrical

artefacts along with those corresponding to missed trials and to the first six to nine trials after breaks were removed. Overall, we excluded ~5% of epochs.

Time-frequency analyses were carried out using the FieldTrip toolbox for MATLAB (*15*). Spectral powers were estimated using an adaptive multi-tapering time-frequency transform (*40*) (Slepian tapers; lower frequency range: 4 to 32 Hz, six cycles, and three tapers per window; higher frequency range: 32 to 200 Hz, fixed time windows of 240 ms, 4 to 31 tapers per window). This approach uses a constant number of cycles across frequencies <32 Hz (time window durations decrease when frequencies increase) and above 32 Hz, a fixed time window with an increasing number of tapers to obtain more precise power estimates (smoothing adaptively increases with frequencies).

iEEG data analyses

Model-based analyses of gamma-band activity

We first computed local activity reflecting neural spiking by averaging spectral powers (in decibels) across gamma and high-gamma frequencies at each time point in each trial. For each contact, we characterized the lower and upper boundaries of the gamma band on the full time-frequency map averaged across trials. On average, the lower bound was 41.9 ± 5.6 Hz (mean ± SEM; minimum, 40 Hz; maximum, 60 Hz), and the upper bound was 140.4 ± 30.8 Hz (mean ± SEM; minimum, 110 Hz; maximum, 200 Hz). Extracting the gamma-band activity using fixed boundaries (50 to 150 Hz) instead of contact specific gamma bands changed neither the pattern nor the overall significance of the reported results (Fig. 3 and fig. S7).

Second, we regressed across trials the gamma-band amplitude computed as above against key parametric variables from the model. More specifically, we estimated a general linear model predicting the trial-to-trial variability in gamma-band amplitude independently at each time step (13.8 ms) within time windows locked on events of interest (from −700 ms to +700 ms relative to stimulus onsets, from −700 ms to +1800 ms relative to responses, from −600 ms to +1200 ms relative to feedback onsets). The general linear model included the following parametric regressors: (i) the posterior reliability of the actor plan in the previous trial, its prior and posterior reliability in the current trial, orthogonalized in that order so that the current prior and posterior reliability regressor properly reflects the sequences of reliability updates; (ii) the same regressors for the most reliable alternative plan in the monitoring buffer; (iii) the RL chosen value $Q^t_{\text{actor}}(s_t, a_{\text{chosen}})$ in the previous and current trial, orthogonalized in that order; and (iv) two potentially confounding factors, namely choice uncertainty [entropy over current actor Q values $Q^t_{\text{actor}}(s_t, a_t)$] and reaction time. For model variable regressors, we used the Monte Carlo trial-by-trial estimates of model variables described in section model fitting procedure above. These regressions were carried out separately for each of contact, then averaged across contacts and participants to generate group-level averages. Fig. 3A shows the time courses of betas associated with the regressors of interest—i.e., the partial correlation slope between gamma-band activity and model variables at every time step. Statistical inferences were performed with a statistical threshold of $P < 0.05$ corrected for multiple comparisons [Bonferroni-Holmes Family-Wise-Error corrections at the cluster level (*41*); $n = 150$ permutations within contacts, $n = 50{,}000$ permutations between contacts].

Analyses of switch compared with stay trials

For each contact, we computed the paired difference in frequency powers at each time step (from −500 ms to +1500 ms relative to feedback onsets) for every frequency (4 to 200 Hz) between switch trials and the two neighboring stay trials (from −2 to +2 trials relative to switch trials). We then computed the corresponding T values, which we averaged across contacts and participants (Fig. 4A). Statistical inferences were performed with a statistical threshold of $P < 0.05$ corrected for multiple comparisons [Bonferroni-Holmes Family-Wise-Error corrections at the cluster level (*41*); $n = 150$ permutations within contacts, $n = 50{,}000$ permutations between contacts].

Neural activity associated with response feedbacks

To analyze the vmPFC and dmPFC local field potentials associated with response feedbacks (Fig. 5 and Fig. 6), we computed in every trial the frequency power at each time step (from −500 ms to +1500 ms relative to feedback onset) relative to the trial grand average (from −2 s to +5 s around stimulus onset) and we averaged across trials. We then computed the mean across contacts and participants. Statistical inferences were performed with a statistical threshold of $P < 0.05$ corrected for multiple comparisons [Bonferroni-Holmes Family-Wise-Error corrections at the cluster level (*41*); $n = 150$ permutations within contacts, $n = 50{,}000$ permutations between contacts].

Psychophysiological interactions

We performed psychophysiological interaction (PPI) analyses (Fig. 7) in stay trials (*42*) between prefeedback high-gamma activities in the vmPFC (averaged from −400 ms to feedback onsets), which encoded chosen values and prior actor reliability (Fig. 3A) and post-feedback high-gamma activities in the dmPFC (averaged from +300 ms to +600 ms after feedback onsets), which encoded unsigned prediction errors (Fig. 6C). We thus tested whether in stay trials, the correlation between these vmPFC and dmPFC activities varied with each of these three model variables (chosen values, actor reliability, and unsigned prediction errors). We performed the three corresponding PPI analyses across all pairs of vmPFC-dmPFC contacts in participants with electrodes implanted in both the vmPFC and dmPFC (three participants, fig. S1, for a total of $n = 18$ contact pairs). For each analysis, we first sorted the stay trials into 30 bins according to the model variable of interest. Across the trials within each bin, we then computed for each contact pair the correlation between vmPFC and dmPFC contact activities, namely the Fisher Z-transformed correlation coefficient. We then estimated the PPI for every contact pair by computing the Pearson's r correlation between these Fisher Z-transformed correlation coefficients and the model variable of interest. PPI statistical significances were assessed by entering the Pearson's r correlations across contact pairs into one-sample two-sided t tests. The results are reported in Fig. 7 (white bars). For completeness, we also performed the same three analyses from residual rather than total high-gamma activities, i.e., after regressing out the joint contribution of the three model variables to every electrode contact activity (Fig. 7, gray bars). PPI results were robust to changes in the number of bins and in the exact boundaries of time windows used for extracting high-gamma activities.

REFERENCES AND NOTES

1. J. D. Cohen, S. M. McClure, A. J. Yu, Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**, 933–942 (2007). doi: 10.1098/rstb.2007.2098; pmid: 17395573
2. A. Collins, E. Koechlin, Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLOS Biol.* **10**, e1001293 (2012). doi: 10.1371/journal.pbio.1001293; pmid: 22479152
3. B. Y. Hayden, J. M. Pearson, M. L. Platt, Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011). doi: 10.1038/nn.2856; pmid: 21642973
4. N. Kolling, T. E. Behrens, R. B. Mars, M. F. Rushworth, Neural mechanisms of foraging. *Science* **336**, 95–98 (2012). doi: 10.1126/science.1216930; pmid: 22491854
5. H. Seo, X. Cai, C. H. Donahue, D. Lee, Neural correlates of strategic reasoning during competitive games. *Science* **346**, 340–343 (2014). doi: 10.1126/science.1256254; pmid: 25236468
6. M. Donoso, A. G. Collins, E. Koechlin, Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science* **344**, 1481–1486 (2014). doi: 10.1126/science.1252254; pmid: 24876345
7. T. C. Blanchard, S. J. Gershman, Pure correlates of exploration and exploitation in the human brain. *Cogn. Affect. Behav. Neurosci.* **18**, 117–126 (2018). doi: 10.3758/s13415-017-0556-2; pmid: 29218570
8. M. Sarafyazd, M. Jazayeri, Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911 (2019). doi: 10.1126/science.aav8911; pmid: 31097640
9. R. P. Rao, D. H. Ballard, Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999). doi: 10.1038/4580; pmid: 10195184

10. K. Friston, A theory of cortical responses. *Phil. Trans. R. Soc. B* **360**, 815–836 (2005). doi: 10.1098/rstb.2005.1622; pmid: 15937014
11. T. Lochmann, S. Deneve, Neural processing as causal inference. *Curr. Opin. Neurobiol.* **21**, 774–781 (2011). doi: 10.1016/j.conb.2011.05.018; pmid: 21742484
12. J. Parvizi, S. Kastner, Promises and limitations of human intracranial electroencephalography. *Nat. Neurosci.* **21**, 474–483 (2018). doi: 10.1038/s41593-018-0108-2; pmid: 29507407
13. P. Ryvlin, J. H. Cross, S. Rheims, Epilepsy surgery in children and adults. *Lancet Neurol.* **13**, 1114–1126 (2014). doi: 10.1016/S1474-4422(14)70156-5; pmid: 25316018
14. T. Hare, Exploiting and exploring the options. *Science* **344**, 1446–1447 (2014). doi: 10.1126/science.1256862; pmid: 24970062
15. R. Oostenveld, P. Fries, E. Maris, J. M. Schoffelen, FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011). doi: 10.1155/2011/156869; pmid: 21253357
16. S. Ray, N. E. Crone, E. Niebur, P. J. Franaszczuk, S. S. Hsiao, Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J. Neurosci.* **28**, 11526–11536 (2008). doi: 10.1523/JNEUROSCI.2848-08.2008; pmid: 18987189
17. H. Liang, S. L. Bressler, M. Ding, W. A. Truccolo, R. Nakamura, Synchronized activity in prefrontal cortex during anticipation of visuomotor processing. *Neuroreport* **13**, 2011–2015 (2002). doi: 10.1097/00001756-200211150-00004; pmid: 12438916
18. J. Gross *et al*., Anticipatory control of long-range phase synchronization. *Eur. J. Neurosci.* **24**, 2057–2060 (2006). doi: 10.1111/j.1460-9568.2006.05082.x; pmid: 17067302
19. Y. Zhang, X. Wang, S. L. Bressler, Y. Chen, M. Ding, Prestimulus cortical activity is correlated with speed of visuomotor processing. *J. Cogn. Neurosci.* **20**, 1915–1925 (2008). doi: 10.1162/jocn.2008.20132; pmid: 18370597
20. J. F. Cavanagh, Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *Neuroimage* **110**, 205–216 (2015). doi: 10.1016/j.neuroimage.2015.02.007; pmid: 25676913
21. J. F. Cavanagh, M. J. Frank, T. J. Klein, J. J. Allen, Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* **49**, 3198–3209 (2010). doi: 10.1016/j.neuroimage.2009.11.080; pmid: 19969093
22. X. J. Wang, Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* **90**, 1195–1268 (2010). doi: 10.1152/physrev.00035.2008; pmid: 20664082
23. A. K. Engel, P. Fries, Beta-band oscillations—Signalling the status quo? *Curr. Opin. Neurobiol.* **20**, 156–165 (2010). doi: 10.1016/j.conb.2010.02.015; pmid: 20359884
24. A. Oswal, V. Litvak, P. Sauleau, P. Brown, Beta reactivity, prospective facilitation of executive processing, and its dependence on dopaminergic therapy in Parkinson's disease. *J. Neurosci.* **32**, 9909–9916 (2012). doi: 10.1523/JNEUROSCI.0275-12.2012; pmid: 22815506
25. G. Buzsáki, A. Draguhn, Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004). doi: 10.1126/science.1099745; pmid: 15218136
26. J. F. Cavanagh, M. J. Frank, Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414–421 (2014). doi: 10.1016/j.tics.2014.04.012; pmid: 24835663
27. B. Voytek *et al*., Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. *Nat. Neurosci.* **18**, 1318–1324 (2015). doi: 10.1038/nn.4071; pmid: 26214371
28. R. F. Helfrich, R. T. Knight, Oscillatory Dynamics of Prefrontal Cognitive Control. *Trends Cogn. Sci.* **20**, 916–930 (2016). doi: 10.1016/j.tics.2016.09.007; pmid: 27743685
29. W. Klimesch, P. Sauseng, S. Hanslmayr, EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Res. Rev.* **53**, 63–88 (2007). doi: 10.1016/j.brainresrev.2006.06.003; pmid: 16887192
30. K. E. Mathewson *et al*., Pulsed out of awareness: EEG alpha oscillations represent a pulsed-inhibition of ongoing cortical processing. *Front. Psychol.* **2**, 99 (2011). doi: 10.3389/fpsyg.2011.00099; pmid: 21779257
31. S. Sadaghiani, A. Kleinschmidt, Brain Networks and α-Oscillations: Structural and Functional Foundations of Cognitive Control. *Trends Cogn. Sci.* **20**, 805–817 (2016). doi: 10.1016/j.tics.2016.09.004; pmid: 27707588
32. T. J. Buschman, E. L. Denovellis, C. Diogo, D. Bullock, E. K. Miller, Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* **76**, 838–846 (2012). doi: 10.1016/j.neuron.2012.09.029; pmid: 23177967
33. O. Jensen, M. Bonnefond, Prefrontal α- and β-band oscillations are involved in rule selection. *Trends Cogn. Sci.* **17**, 10–12 (2013). doi: 10.1016/j.tics.2012.11.002; pmid: 23176827
34. J. M. Pearce, G. Hall, A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980). doi: 10.1037/0033-295X.87.6.532; pmid: 7443916
35. L. H. Arnal, A. L. Giraud, Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **16**, 390–398 (2012). doi: 10.1016/j.tics.2012.05.003; pmid: 22682813
36. J. Isnard, M. Guénot, K. Ostrowsky, M. Sindou, F. Mauguière, The role of the insular cortex in temporal lobe epilepsy. *Ann. Neurol.* **48**, 614–623 (2000). doi: 10.1002/1531-8249(200010)48:4<614::AID-ANA8>3.0.CO;2-S; pmid: 11026445
37. R. M. Neal, Slice sampling. *Ann. Stat.* **31**, 705–767 (2003). doi: 10.1214/aos/1056562461
38. A. Delorme, S. Makeig, EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004). doi: 10.1016/j.jneumeth.2003.10.009; pmid: 15102499
39. J.-P. Lachaux, D. Rudrauf, P. Kahane, Intracranial EEG and human brain mapping. *J. Physiol. Paris* **97**, 613–628 (2003). doi: 10.1016/j.jphysparis.2004.01.018; pmid: 15242670
40. P. P. Mitra, B. Pesaran, Analysis of dynamic brain imaging data. *Biophys. J.* **76**, 691–708 (1999). doi: 10.1016/S0006-3495(99)77236-X; pmid: 9929474
41. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007). doi: 10.1016/j.jneumeth.2007.03.024; pmid: 17517438
42. K. J. Friston *et al*., Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* **6**, 218–229 (1997). doi: 10.1006/nimg.1997.0291; pmid: 9344826
43. P. Domenech, S. Rheims, E. Koechlin, PROBE iEEG, version 1.0.0, OpenNeuro (2020); https://openneuro.org/datasets/ds003078/versions/1.0.0.

ACKNOWLEDGMENTS

We thank V. Wyart and J. Drugowitsch for their help. **Funding:** This work was supported by a European Research Council grant (ERC-AdG no. 250106) to E.K. and the Institut Hospitalo-Universitaire CESAME (ANR-10-IBHU-0003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. **Author contributions:** E.K. and P.D. conceived the study and designed the experimental protocol. P.D. and S.R. collected the data. P.D. and E.K. analyzed the data. E.K. and P.D. wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Intracranial EEG and MRI data are available in the OpenNeuro public repository database (*43*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/eabb0184/suppl/DC1
Supplementary Text
Figs. S1 to S7
Tables S1 to S3
References (*44*–*49*)

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

23 January 2020; accepted 4 June 2020
10.1126/science.abb0184

RESEARCH ARTICLE

MICROBIOLOGY

Diverse enzymatic activities mediate antiviral immunity in prokaryotes

Linyi Gao[1,2,3], Han Altae-Tran[1,2,3], Francisca Böhning[1,2], Kira S. Makarova[4], Michael Segel[1,2,3,5,6], Jonathan L. Schmid-Burgk[1,2,3,5,6], Jeremy Koob[1,2], Yuri I. Wolf[4], Eugene V. Koonin[4], Feng Zhang[1,2,3,5,6]*

Bacteria and archaea are frequently attacked by viruses and other mobile genetic elements and rely on dedicated antiviral defense systems, such as restriction endonucleases and CRISPR, to survive. The enormous diversity of viruses suggests that more types of defense systems exist than are currently known. By systematic defense gene prediction and heterologous reconstitution, here we discover 29 widespread antiviral gene cassettes, collectively present in 32% of all sequenced bacterial and archaeal genomes, that mediate protection against specific bacteriophages. These systems incorporate enzymatic activities not previously implicated in antiviral defense, including RNA editing and retron satellite DNA synthesis. In addition, we computationally predict a diverse set of other putative defense genes that remain to be characterized. These results highlight an immense array of molecular functions that microbes use against viruses.

Bacterial and archaeal viruses are the most abundant, and possibly the most diverse, biological entities on Earth (*1*, *2*). To resist frequent and varied attacks by viruses, prokaryotes possess multiple antiviral defense systems. These include the adaptive immune system CRISPR-Cas, which provides immunity by memorizing past infection events (*3*), and a variety of innate immune systems, such as restriction-modification (RM) systems that target specific, predefined sequences in the viral DNA; abortive infection (Abi) systems that induce cell dormancy or death upon viral infection; and additional systems with mechanisms that have not yet been elucidated (*4*). Antiviral defense systems range in complexity from a single small protein (e.g., certain types of Abi systems) to 10 or more proteins acting in concert (e.g., type I and type III CRISPR-Cas systems). Conversely, viruses have evolved strategies to counteract many of these defense systems, including anti-CRISPR and antirestriction proteins (*5*, *6*). Given the vast diversity of viruses and their complex patterns of coevolution with defense systems (*7*–*9*), more types of defense systems with diverse mechanisms can be expected to exist than are currently known.

[1]Howard Hughes Medical Institute, Cambridge, MA 02139, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA. [5]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [6]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. Email: zhang@broadinstitute.org

Domain-independent prediction of uncharacterized defense systems

Many antiviral defense genes in bacterial and archaeal genomes show a distinctive tendency to cluster together in defense "islands" (*7*, *10*). As a consequence, an uncharacterized gene whose homologs consistently occur next to, for instance, RM genes has an increased likelihood of being involved in defense (*11*, *12*). Using this principle, a recent analysis (*4*) identified and validated 10 previously unknown defense systems, based on the requirement that each (putative) system contain at least one annotated protein domain that is enriched in defense islands.

We hypothesized that additional, unknown defense systems exist that either lack annotated domains or only contain domains that are not typically associated with defense but have been co-opted in specific instances to perform defense functions. To test this hypothesis, we developed an expanded computational approach in which putative defense systems are predicted independent of domain annotations (Fig. 1A). We analyzed all bacterial and archaeal genomes available in GenBank as of November 2018, collectively encoding 620 million proteins. To identify candidate defense genes, we first compiled a list of all genes within 10 kb or 10 open reading frames away from known defense systems (materials and methods). This initial list (n = 8.7 × 10^6), which evidently contained both defense genes and nondefense ones, was clustered to yield 6 × 10^5 representative sequences ("seeds"). To distinguish between defense and nondefense seeds, we identified all homologs of each seed present in GenBank and analyzed their gene neighborhoods. The seed was predicted to be a defense gene if these neighborhoods resembled those of known defense genes—in particular, if a high percentage of homologs were located in proximity to known defense genes and displayed context diversity (Fig. 1B, fig. S1, and materials and methods). All clustering and homolog detection steps were performed on the basis of amino acid sequences, without invoking existing domain annotations and thus allowing the prediction of previously unknown types of defense genes.

After all filtering and curation steps, we identified a total of 7472 seeds (table S1) that represented putative defense genes, along with 4555 seeds for known defense genes under the same analysis parameters (Fig. 1C and table S2). These seeds were analyzed with additional, more sensitive analysis of their domain content (table S3). Of the uncharacterized genes, 1687 (23%) had either no annotated domains or contained only domains of unknown function (DUFs), and an additional 2756 (37%) contained only domains that are different from the characteristic domains of known defense genes. These results suggest the existence of a diverse set of defense genes with mechanisms that remain to be investigated.

Candidate defense systems exhibit antiviral activity in a heterologous system

To characterize the functional diversity among the predicted defense genes, we selected 48 candidate systems to test experimentally for defense activity. Candidate systems were prioritized on the basis of the presence of predicted molecular functions not previously implicated in defense; broad phylogenetic distribution; the presence of at least one protein larger than 300 amino acids (to increase the likelihood of the presence of enzymes); and, for multigene systems, conservation of the component genes. Because wild-type bacterial strains are likely to harbor multiple active defense systems, thereby maintaining phage resistance even if one of the systems were knocked out (*13*), we elected to assay activity by heterologous reconstitution. For each system, one to four homologs were selected, cloned from the source organism into the low-copy vector pACYC, and transformed into *Escherichia coli* (Fig. 2A), comprising a total of 395 kb of exogenous DNA (see tables S4 to S11 for sequence, accession, and source organism information). Three previously identified defense systems, BREX type I (*13*, *14*), Druantia type I (*4*), and the Abi reverse transcriptase RT-Abi-P2 (*15*) were included as positive controls. Each system was then challenged with a diverse panel of coliphages with double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), or single-stranded RNA (ssRNA) genomes, and phage sensitivity of the bacteria was compared to that observed with the empty vector control (Fig. 2, B and C).

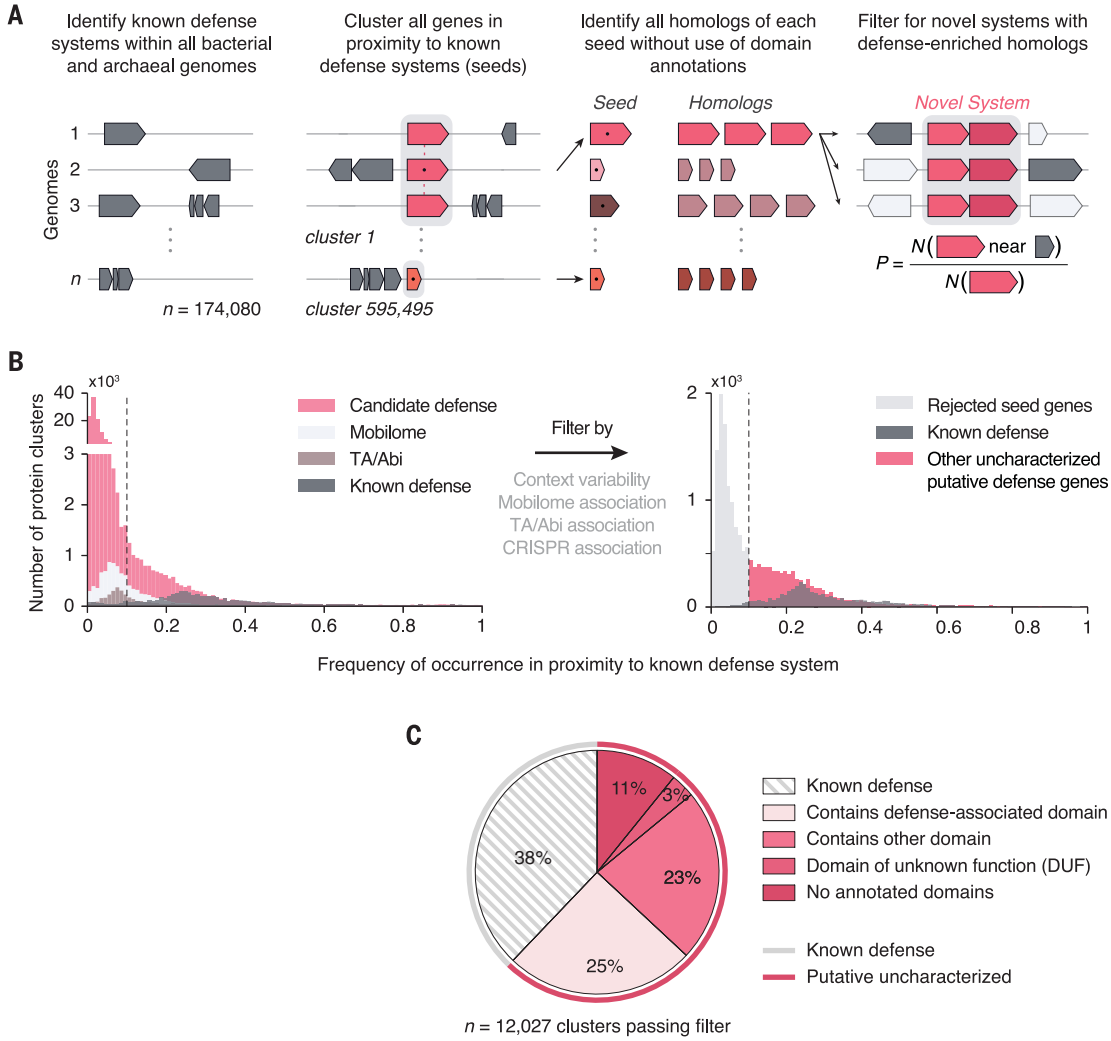


Fig. 1. Domain-independent prediction of putative antiviral defense systems. (**A**) Computational pipeline to identify uncharacterized putative defense systems across all sequenced bacterial and archaeal genomes. Defense systems were predicted on the basis of analysis of amino acid sequences, independent of domain annotations. (**B**) Histograms of defense association frequencies before filtering and after neighborhood context–based filtering (minimum 50 homologs). Seeds to the right of the dashed line (0.1) were selected for further analysis. TA, toxin-antitoxin. (**C**) Pie chart of the domain diversity among predicted defense genes, based on additional analysis using HHpred against pfam domains.

We observed antiphage activity for 29 of the 48 tested candidates (60%) (fig. S2). Systems from source organisms outside the Enterobacteriaceae family, which consists of *Escherichia* and closely related genera including *Salmonella* and *Klebsiella*, had little to no activity, suggesting the importance of host compatibility. The most active representative in each of these 29 systems (representing 4% of the uncharacterized defense seeds) was further tested with an expanded panel of phages in two *E. coli* strains (Fig. 2D and fig. S3). All 29 systems were active against at least one dsDNA phage, and four were active against ssDNA phages (M13 or φX174). Phage specificity was typically narrow and varied widely across systems. The abundance of these defense systems among the sequenced bacterial and archaeal genomes spans two orders of magnitude, ranging from ~0.1 to ~10% of the genomes (Fig. 2D). Overall, 32% of all sequenced bacterial and archaeal genomes contain at least one of these defense systems, which are broadly distributed across bacterial and archaeal phyla (fig. S4).

RADAR contains a divergent adenosine deaminase that edits RNA in response to phage infection

We identified a two-gene cassette consisting of an adenosine triphosphatase (ATPase) (~900 residues) and a divergent adenosine deaminase (~900 residues) that was active against dsDNA phages T2, T3, T4, and T5. Because deaminase activity had not been previously implicated in antiviral defense, we focused on this system for further investigation. The system appears in diverse defense contexts and forms three subtypes (Fig. 3A and fig. S5A). In most cases, it consists of the ATPase and deaminase only, but some variants also include a small membrane protein, either a SLATT domain (*16*) or the type VI-B CRISPR ancillary protein Csx27 (*17*). Mutations in the ATPase Walker B motif or in the putative divalent metal cation-binding HxH motif of the deaminase abolished defense activity, whereas the SLATT domain membrane protein was required for resistance against phage T5 but not against phage T2 (Fig. 3B).

Given the large size of the deaminase compared with typical metabolic adenosine deaminases and its sequence divergence due to large insertions in the deaminase domain (fig. S5B), we hypothesized that it acts on nucleic acids rather than on free nucleosides or nucleotides. To test this hypothesis, we performed whole-transcriptome sequencing and found an enrichment of A-to-G substitutions in sequencing reads at specific sites in the presence of phage, whereas C, G, or U bases were not affected (Fig. 3C and fig. S6A), consistent with RNA editing of adenosine to inosine. Furthermore, the overall expression of phage genes, including early genes, was reduced by ~100-fold even at a multiplicity of infection (MOI) of 2 (Fig. 3D). Because most of the cells in the culture were expected to be infected, this suggested that defense activity occurs early in the infection cycle, which was not evident from efficiency of plating alone.

RNA editing occurred only when both the defense system and the phage were present; expression of the defense system without the phage resulted in a near-baseline level of

editing, and no editing was detected in the absence of the system. Mutations in the ATPase or deaminase active sites abolished editing, and no DNA editing was detected (fig. S6B). Editing sites were broadly distributed throughout the *E. coli* transcriptome (Fig. 3E, figs. S6A and S7, and table S12), and editing could also be induced by coexpressing specific phage proteins with the system (fig. S8 and table S13). RNA secondary-structure predictions indicated a characteristic stem-loop structure at strong editing sites; specific adenosines in loops were edited with up to ~90% frequency, whereas adenosines in the stem were not edited within the limit of detection (Fig. 3E and fig. S7). Finally, some of the editing sites are likely to be deleterious to the host cell, resulting in nonsynonymous mutations such as at the UAA stop codon of the transfer messenger RNA (tmRNA) (fig. S8B), which rescues ribosomes stalled during translation (*18*).

On the basis of these results, we named this system phage restriction by an adenosine deaminase acting on RNA (RADAR). Growth kinetics at varying phage MOI revealed a threshold MOI above which RADAR-expressing cells had a lower optical density at 600 nm

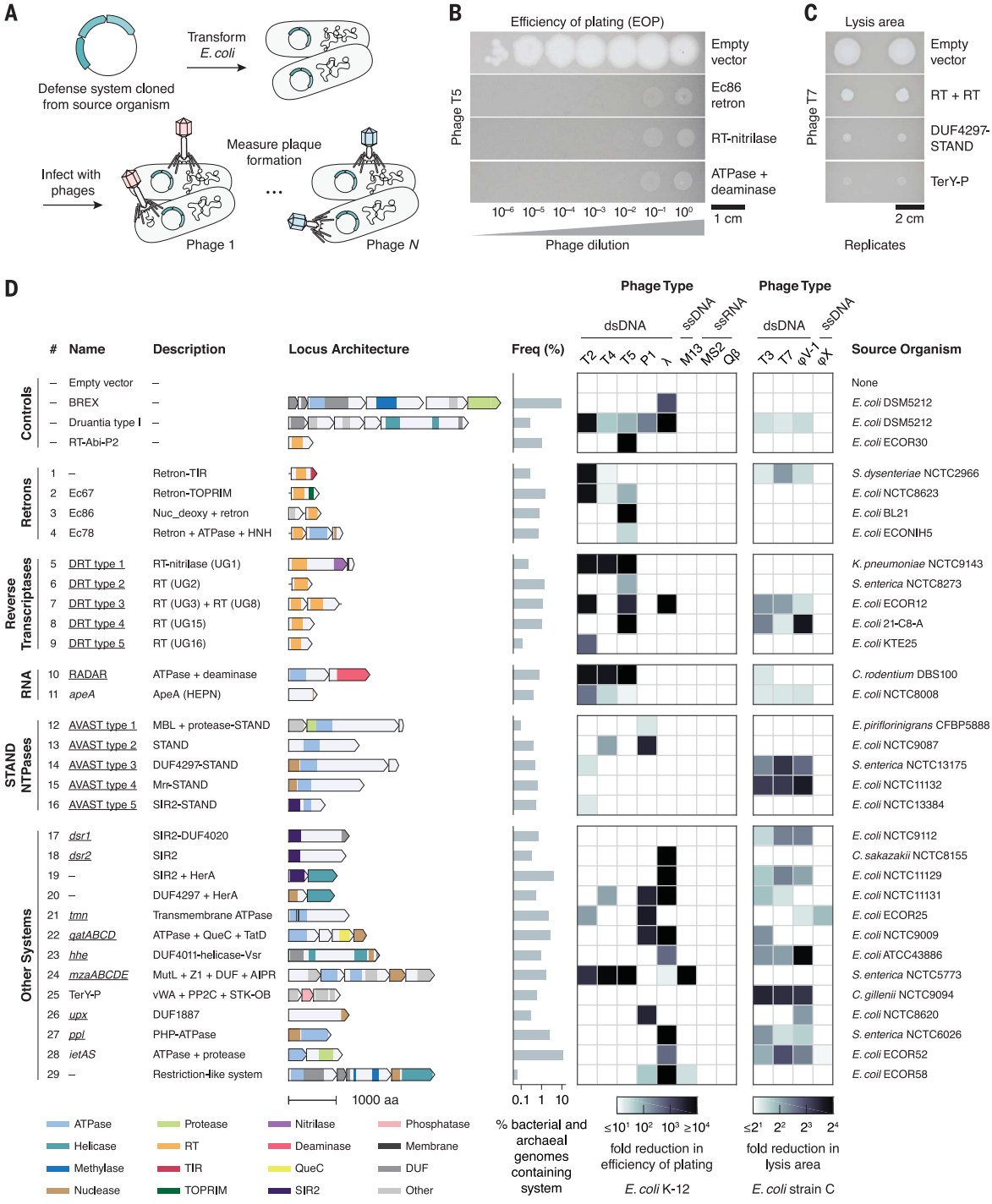


Fig. 2. Candidate defense systems exhibit antiviral activity in a heterologous system. (**A**) Experimental validation pipeline using phage plaque assays on *E. coli* heterologously expressing a cloned candidate defense system. (**B**) Example plaques and (**C**) zones of lysis for six candidate defense systems. (**D**) Antiphage activity across a panel of 12 coliphages with dsDNA, ssDNA, or ssRNA genomes (mean of two replicates). The bar graph shows the abundance of each system in sequenced bacterial and archaeal genomes. Domains: RT, reverse transcriptase; TIR, Toll/interleukin-1 receptor homology domain; TOPRIM, topoisomerase-primase domain; QueC, 7-cyano-7-deazaguanine synthase-like domain; SIR2, sirtuin; mem, membrane, transmembrane helix; DUF, domain of unknown function. Proposed gene names: DRT, defense-associated reverse transcriptase; RADAR, phage restriction by an adenosine deaminase acting on RNA; AVAST, antiviral ATPase/NTPase of the STAND superfamily; *dsr*, defense-associated sirtuin; *tmn*, transmembrane NTPase; *qat*, QueC-like associated with ATPase and TatD DNase; *hhe*, HEPN, helicase, and Vsr endonuclease; *mza*, MutL, Z1, and AIPR; *upx*, uncharacterized (P)D-(D/E)-XK defense protein; *ppl*, polymerase/histidinol phosphatase-like. aa, amino acids; HerA, helicase; MBL, metallo β-lactamase.

(OD600) compared with the empty vector control, suggestive of RADAR-mediated growth arrest (Fig. 3F). Together with the abundance and broad distribution of editing sites in the host transcriptome (figs. S6 and S7), these results are consistent with an editing-dependent Abi mechanism that is activated by phage.

A widespread family of defense systems containing reverse transcriptases

We discovered that a family of uncharacterized reverse transcriptases (RTs) are active defense systems. Although most RTs in prokaryotes are components of mobile retroelements, distinct clades of RTs that lack the hallmarks of mobility also exist, including 16 unknown groups (UGs) (*19*–*22*). We independently identified many of these uncharacterized RTs through our pipeline, suggesting that they might be defense genes (Fig. 4A). Indeed, six of these candidates (UG1, UG2, UG3, UG8, UG15, and UG16) provided robust protection

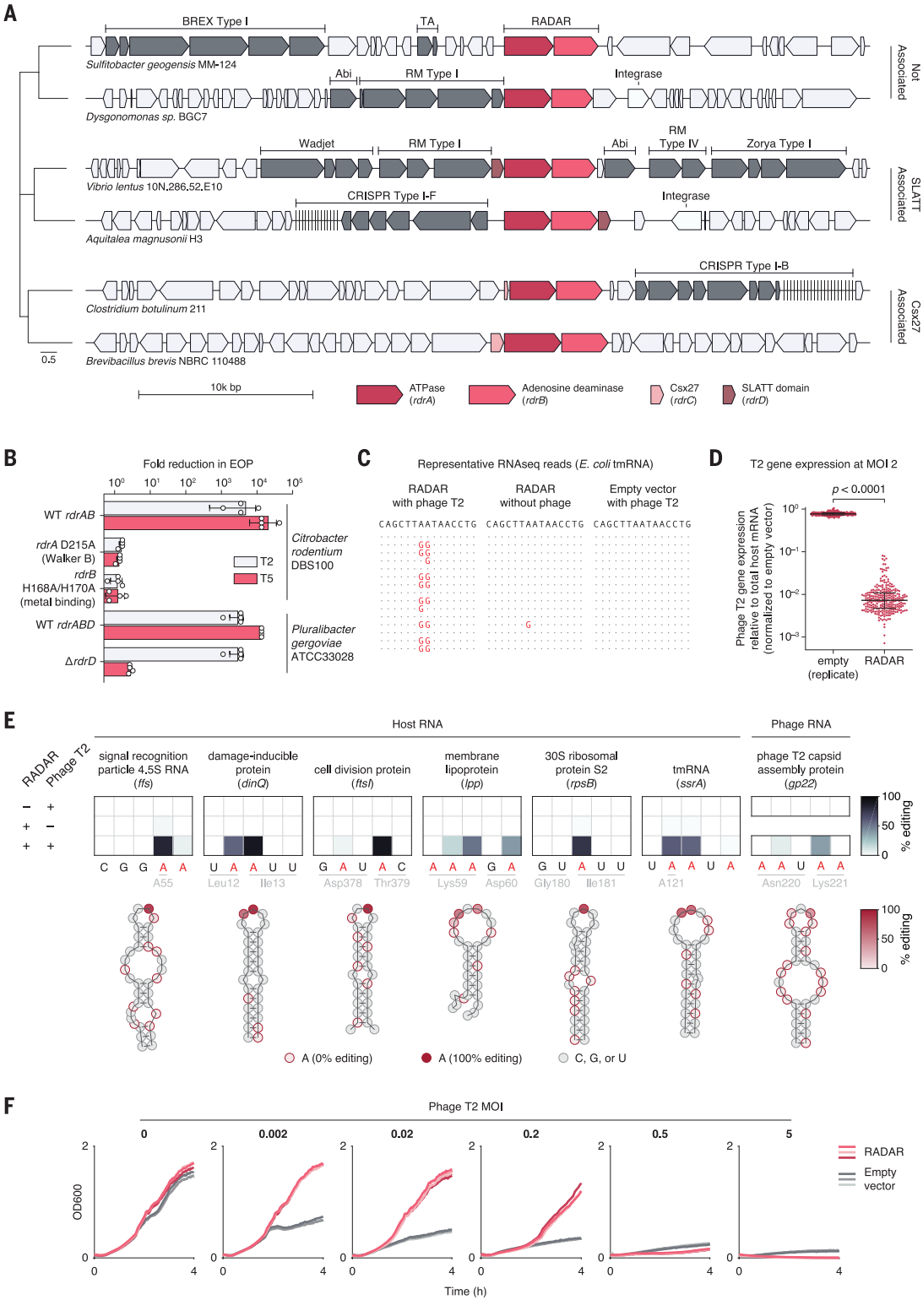


Fig. 3. RADAR mediates RNA editing in response to phage infection. (**A**) Examples of genomic loci containing three subtypes of RADAR (standalone, Csx27-associated, and SLATT-associated). (**B**) Essentiality of the core RADAR genes *rdrAB* and the accessory gene *rdrD* against phages T2 and T5. D215A, Asp^{215}→Ala; H168A, His^{168}→Ala; H170A, His^{170}→Ala; WT, wild type. (**C**) Representative RNA sequencing (RNAseq) reads from *E. coli* expressing either RADAR or an empty vector control. (**D**) Expression of phage T2 RNA relative to total host RNA in *E. coli* containing RADAR. Each dot represents a phage gene. Cells were infected at a MOI of 2. The *p* value was determined by a Wilcoxon signed-rank test. (**E**) Representative editing sites in the host and phage transcriptomes, with corresponding predicted RNA secondary structures. (**F**) Growth kinetics of RADAR-containing *E. coli* in comparison with an empty vector control under varying MOI by phage T2.

against dsDNA phages. In all cases, mutations in the RT active site [(Y/F)xDD to (Y/F)xAA, where x is any amino acid] abolished activity (Fig. 4B and fig. S9, A and B). We named these genes defense-associated RTs (DRTs).

Each of these RT systems displayed a distinct pattern of phage resistance (Fig. 2D). Moreover, whereas UG2 (*drt2*), UG15 (*drt4*), and UG16 (*drt5*) act as individual genes, the UG3 (*drt3a*) and UG8 (*drt3b*) RTs are components of the same defense system (DRT type 3), with both RTs required for defense activity. Like RADAR, some subtypes of the UG1 (DRT type 1) and DRT type 3 systems are also associated with small membrane proteins (Fig. 4A). Moreover, DRT type 1 encompasses a much larger protein (~1200 residues) than the other five RTs and also contains a C-terminal nitrilase domain. Mutation of the catalytic cysteine of the nitrilase to alanine (C1119A) abolished activity (Fig. 4B). Nitrilases typically function in processes unrelated to defense, such as nucleotide metabolism and small-molecule biosynthesis (*23*). Thus, DRT type 1, which is divergent from typical nitrilases and forms a distinct clade in the phylogenetic tree of the nitrilase family (fig. S10), exemplifies a non-defense domain that was apparently co-opted for a defense function.

To further characterize these RTs, we performed whole-transcriptome sequencing of RT-expressing *E. coli* during phage infection. These experiments revealed substantial differences in phage gene expression across the different RTs (Fig. 4C). For instance, DRT type 1 strongly suppressed the expression of phage late genes, such as capsid proteins, whereas early and middle genes were not substantially affected, suggesting that it is active before the late stage of infection

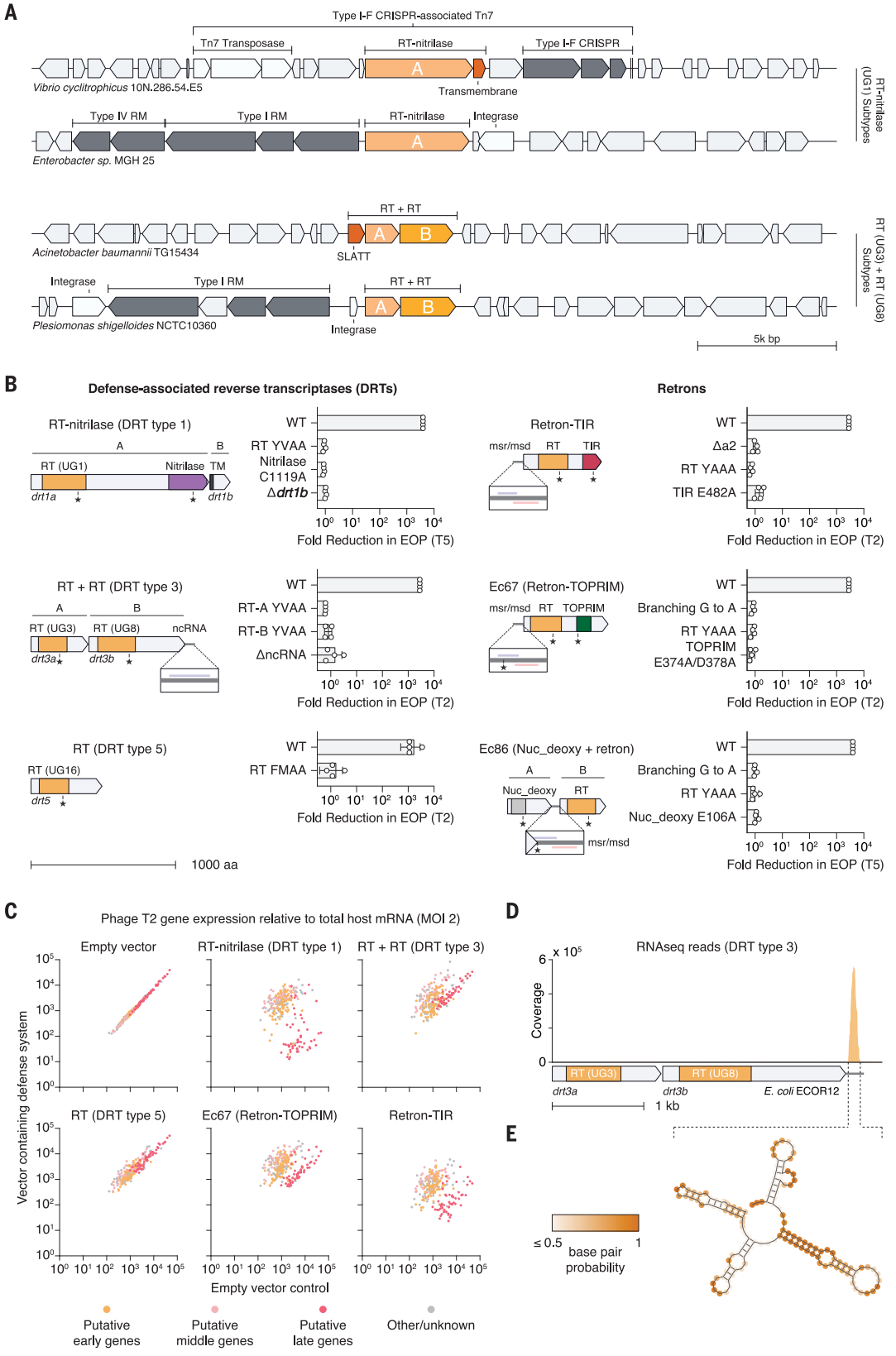


Fig. 4. Diverse families of RTs mediate antiviral defense. (**A**) Examples of genomic loci containing two RT-based defense systems (DRT type 1 and type 3), with two representative subtypes shown for each system. (**B**) Essential components of non-retron RTs (left panel) and retrons (right panel). TM, transmembrane; ncRNA, noncoding RNA; msr/msd: genes encoding msRNA and msDNA, respectively; a2, retron 5′ inverted repeat. (**C**) Effect of defense RTs on the expression of phage T2 genes in *E. coli* infected at an MOI of 2. (**D**) RNAseq reads mapping to the DRT type 3 system. (**E**) Predicted secondary structure of the highly expressed noncoding RNA identified in (D).

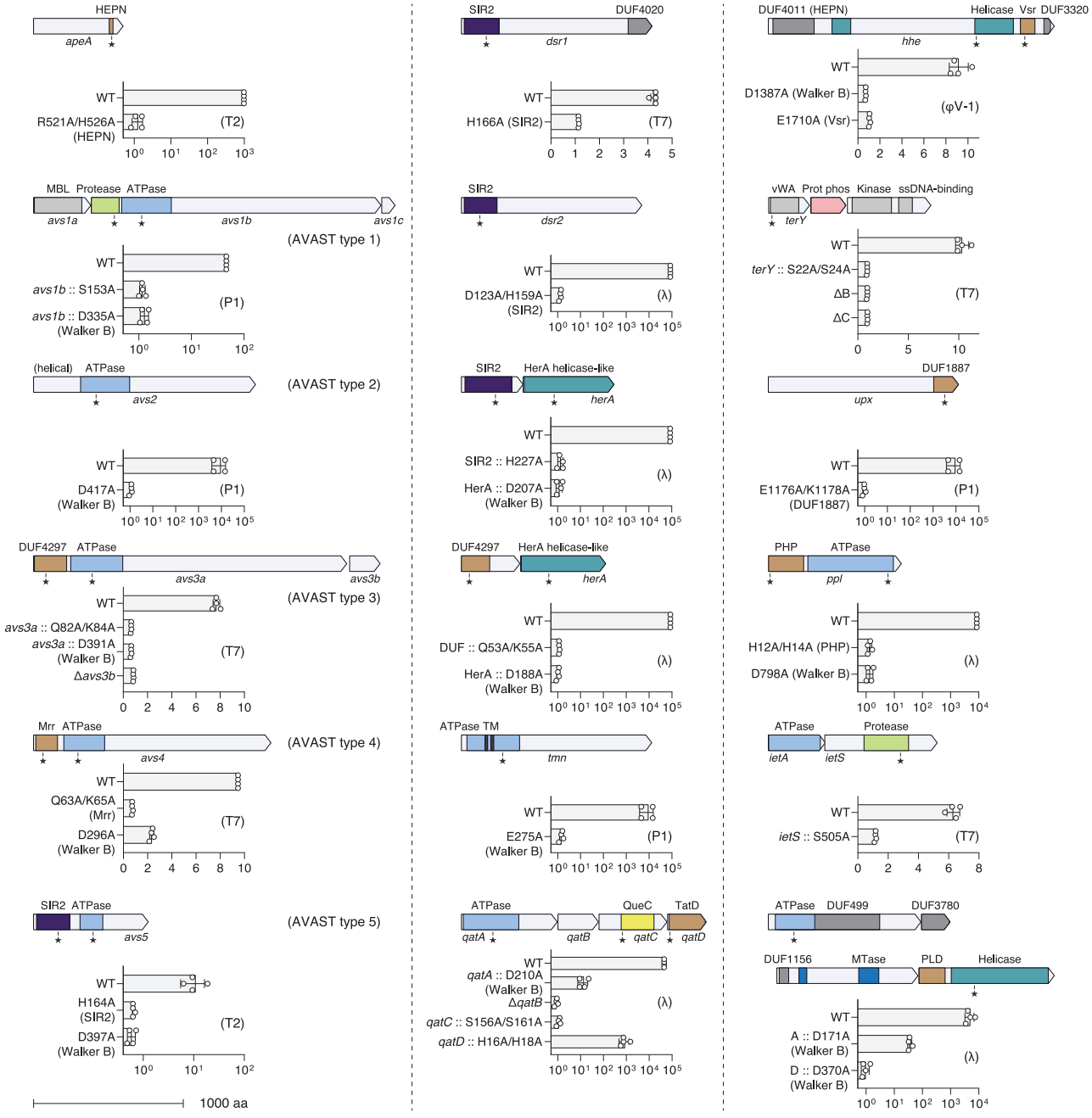


Fig. 5. Domain architectures and mutational analysis of additional defense systems. Graphics show domains identified by using HHpred, and stars indicate locations of active site mutations. Bar graphs (n = 4 replicates per bar) show either $\log_{10}$ fold change of efficiency of plating (for phages T2, P1, and λ) or $\log_2$ fold change in the area of the zone of lysis (for phages T7 and φV-1) relative to the empty vector control. MBL, metallo β-lactamase; SIR2, sirtuin; HerA, helicase; QueC, 7-cyano-7-deazaguanine synthase-like domain; Vsr, very short patch repair endonuclease; TatD, DNase; vWA, von Willebrand factor type A; Prot phos, serine/threonine protein phosphatase; PHP, polymerase/histidinol phosphatase; MTase, methyltransferase; PLD, phospholipase D; DUF, domain of unknown function.

but does not prevent the injection of phage DNA into the host cell. By contrast, DRT type 3 did not strongly suppress expression of any of the phage genes, despite growing at a rate similar to DRT type 1 during phage infection (fig. S11A). Transcriptome sequencing also identified a highly expressed, structured noncoding RNA at the 3′ end of the DRT type 3 system that is required for activity (Fig. 4, B, D, and E).

Retrons mediate antiphage defense

We also found that retrons, a distinct class of RTs that produce extrachromosomal satellite DNA [multicopy ssDNA (msDNA)], are active antiphage defense systems. The retron msDNA is produced from the 5′ untranslated region of its own mRNA and is covalently linked to an internal guanosine of the RNA through a 2′-5′ phosphodiester bond (*24*). First identified more than 30 years ago, retrons have been harnessed for bacterial genome engineering (*25*), but their native biological function has remained unknown. We found that the original *E. coli* retrons Ec67 (*26*) and Ec86 (*27*), as well as a homolog of the Ec78 retron (*28*) and a previously uncharacterized TIR (Toll/interleukin-1 receptor) domain–associated retron, mediate defense against dsDNA phages. The Ec86

retron is natively present in the widely used laboratory *E. coli* strain BL21. Mutations in the (Y/F)xDD active site motif of the RT, as well as at the branching guanosine, abolished activity, indicating that the defense function depends on msDNA synthesis (Fig. 4B and fig. S9C). Furthermore, perturbations to the msDNA also abolished activity (fig. S11), suggesting that its structure, and not simply formation, is essential for the defense activity. Indeed, a single nucleotide mismatch in the msDNA hairpin reduced activity by 100- to 1000-fold, but introducing a second mutation on the complementary strand to restore the structure of the msDNA also restored wild-type activity (fig. S11). Notably, these retrons are associated with other domains, including TOPRIM (topoisomerase-primase) (*29*); TIR (*30*), a nucleoside deoxyribosyltransferase–like enzyme; and the Septu defense system (*4*), all of which are required for activity (Fig. 4B).

Additional molecular functions of defense systems

We investigated several additional systems with diverse components (Fig. 5 and fig. S12). These include a three-gene system containing a von Willebrand factor A (vWA) metal ion binding protein, a serine/threonine protein phosphatase, and a serine/threonine protein kinase that provided strong protection against T7-like phages (T3, T7, and φV-1). This system, called the TerY-phosphorylation triad (TerY-P), has been previously analyzed computationally in the context of tellurite resistance–associated stress response and might operate as a phosphorylation switch that couples the activities of the kinase and the phosphatase (*31*).

Additional systems include proteins containing a SIR2 (sirtuin) deacetylase domain that is also present in the recently discovered Thoeris system (*4*) and has also been detected in the same neighborhoods with prokaryotic Argonaute proteins (*32*); ApeA, a predicted HEPN-family Abi protein (*33*) and a putative ancestor of the type VI CRISPR effector Cas13; a ~1300-residue P-loop ATPase containing an unusual insertion of two transmembrane helices into the ATPase domain, similar to the KAP ATPases (*34*); and a four-gene cassette containing a 7-cyano-7-deazaguanine synthase-like protein (QueC), suggestive of small-molecule biosynthesis. All of these components are essential for defense activity (Fig. 5). Further investigation is required to understand the mechanisms by which these systems sense and respond to phage infection.

Finally, we also demonstrate defense functions for several predicted nucleoside triphosphatases (NTPases) of the STAND (signal transduction ATPases with numerous associated domains) superfamily (Fig. 5). This expansive superfamily consists of multidomain proteins that include eukaryotic ATPases and GTPases involved in programmed cell death and various forms of signal transduction (*35*, *36*). Typically, STAND NTPases contain a C-terminal helical sensor domain that, on target recognition, induces oligomerization by means of ATP or GTP hydrolysis, leading to activation of the N-terminal effector domain. The role of the STAND NTPases in prokaryotes has long remained enigmatic (*35*, *37*); the few for which experimental data are available contain a helix-turn-helix domain and have been shown to regulate transcription (*36*). Several STAND NTPases were active against dsDNA phages (Fig. 2D); these proteins contained different putative effector domains, including DUF4297 [a putative PD-(D/E)xK-family nuclease], an Mrr-like nuclease, SIR2, a trypsin-like serine protease, and an uncharacterized helical domain. We named these systems antiviral ATPases/NTPases of the STAND superfamily (AVAST). As homologs of essential eukaryotic programmed cell death effectors, AVAST systems are likely to function through an Abi mechanism, i.e., by causing growth arrest or programmed cell death in infected hosts.

Discussion

These findings substantially expand the space of protein domains, molecular functions, and interactions that are used by bacteria and archaea in antiviral defense. Some of these functions, including RNA editing, have not been previously implicated in defense mechanisms. The high success rate of defense system prediction based on the evolutionary conservation of their proximity to previously identified defense genes supports the defense island concept (*4*, *7*, *10*) and demonstrates its growing utility at the time of rapid expansion of sequence databases. Furthermore, the computational approach implemented in this work provided for a substantial expansion of the range of the identified putative defense systems. Many of these previously unknown systems contain enzymatic activities as well as predicted sensor components that potentially could be engineered for new biotechnology applications.

Despite similarities in domain architectures among some of the identified defense systems, their phage specificities differ significantly, emphasizing the importance of multiple defense mechanisms for the survival of prokaryotes in the arms race against viruses. These observations are compatible with the concept of distributed microbial immunity, according to which defense systems encoded in different genomes collectively protect microbial communities from the diverse viromes they confront (*38*). Additionally, several of the identified defense systems incorporate molecular functions from typically nondefense sources, highlighting the versatility of activities that are recruited for antiviral defense. These include the RADAR deaminase, nitrilases, RTs, and retrons. The demonstration of defense functions for multiple RTs, which are generally associated with mobile genetic elements (MGEs), is consistent with the "guns for hire" paradigm whereby enzymes are shuttled between MGEs and defense systems during microbial evolution (*8*). Finally, most of these defense systems do not appear to be substantially enriched within prophages, suggesting that they are dedicated host defense genes, rather than virus superinfection exclusion modules (fig. S13 and materials and methods).

The overall patchy pattern of phage specificity observed for the different defense systems was unexpected. In some cases, the same system exhibited widely varying levels of protection against similar phages; for instance, retron Ec67 and DRT type 3 offered full protection against phage T2 but poor protection against phage T4, which is ~82% identical to T2. We hypothesize that phage-encoded inhibitors or antidefense genes that have yet to be discovered may play an important role in determining the specificity of defense systems. Indeed, many phage-encoded proteins have no known function and remain to be investigated for antidefense activity.

With the exception of RADAR, we do not yet know the mechanisms of most of the identified defense systems. The range of domains contained in these systems indicates that they use diverse biochemical activities. Additional experimental characterization is required to elucidate the effector functions for these systems and the molecular basis of antiphage action and specificity. The identification of these defense systems, as well as others we have predicted computationally, provides a foundation for further mechanistic investigation.

The results described here have broad implications for understanding antiviral resistance and host-virus interactions in natural populations of microbes, as well as for technological applications such as the development of antibacterial therapeutics, nucleic acid editing, molecular detection, and targeted cell destruction.

REFERENCES AND NOTES

1. C. A. Suttle, *Genome* **56**, 542–544 (2013).
2. A. G. Cobián Güemes *et al.*, *Annu. Rev. Virol.* **3**, 197–214 (2016).
3. F. Hille *et al.*, *Cell* **172**, 1239–1259 (2018).
4. S. Doron *et al.*, *Science* **359**, eaar4120 (2018).
5. J. E. Samson, A. H. Magadán, M. Sabri, S. Moineau, *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
6. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, *Nature* **493**, 429–432 (2013).
7. K. S. Makarova, Y. I. Wolf, E. V. Koonin, *Nucleic Acids Res.* **41**, 4360–4377 (2013).
8. E. V. Koonin, K. S. Makarova, Y. I. Wolf, M. Krupovic, *Nat. Rev. Genet.* **21**, 119–131 (2020).
9. G. Faure *et al.*, *Nat. Rev. Microbiol.* **17**, 513–525 (2019).
10. K. S. Makarova, Y. I. Wolf, S. Snir, E. V. Koonin, *J. Bacteriol.* **193**, 6039–6056 (2011).

11. S. A. Shmakov, K. S. Makarova, Y. I. Wolf, K. V. Severinov, E. V. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5307–E5316 (2018).
12. S. A. Shmakov *et al.*, *Nat. Protoc.* **14**, 3013–3031 (2019).
13. J. Gordeeva *et al.*, *Nucleic Acids Res.* **47**, 253–265 (2019).
14. T. Goldfarb *et al.*, *EMBO J.* **34**, 169–183 (2015).
15. R. Odegrip, A. S. Nilsson, E. Haggård-Ljungquist, *J. Bacteriol.* **188**, 1643–1647 (2006).
16. A. M. Burroughs, D. Zhang, D. E. Schäffer, L. M. Iyer, L. Aravind, *Nucleic Acids Res.* **43**, 10633–10654 (2015).
17. K. S. Makarova, L. Gao, F. Zhang, E. V. Koonin, *FEMS Microbiol. Lett.* **366**, fnz088 (2019).
18. C. D. Rae, Y. Gordiyenko, V. Ramakrishnan, *Science* **363**, 740–744 (2019).
19. S. Zimmerly, L. Wu, *Microbiol. Spectr.* **3**, A3–A0058, 2014 (2015).
20. N. Toro, R. Nisa-Martínez, *PLOS ONE* **9**, e114083 (2014).
21. K. K. Kojima, M. Kanehisa, *Mol. Biol. Evol.* **25**, 1395–1404 (2008).
22. D. M. Simon, S. Zimmerly, *Nucleic Acids Res.* **36**, 7219–7229 (2008).
23. H. C. Pace, C. Brenner, *Genome Biol.* **2**, S0001 (2001).
24. A. J. Simon, A. D. Ellington, I. J. Finkelstein, *Nucleic Acids Res.* **47**, 11007–11019 (2019).
25. F. Farzadfard, T. K. Lu, *Science* **346**, 1256272 (2014).
26. B. C. Lampson *et al.*, *Science* **243**, 1033–1038 (1989).
27. D. Lim, W. K. Maas, *Cell* **56**, 891–904 (1989).
28. T. M. Lima, D. Lim, *Plasmid* **38**, 25–33 (1997).
29. L. Aravind, D. D. Leipe, E. V. Koonin, *Nucleic Acids Res.* **26**, 4205–4213 (1998).
30. S. Horsefield *et al.*, *Science* **365**, 793–799 (2019).
31. V. Anantharaman, L. M. Iyer, L. Aravind, *Mol. Biosyst.* **8**, 3142–3165 (2012).
32. K. S. Makarova, Y. I. Wolf, J. van der Oost, E. V. Koonin, *Biol. Direct* **4**, 29 (2009).
33. V. Anantharaman, K. S. Makarova, A. M. Burroughs, E. V. Koonin, L. Aravind, *Biol. Direct* **8**, 15 (2013).
34. L. Aravind, L. M. Iyer, D. D. Leipe, E. V. Koonin, *Genome Biol.* **5**, R30 (2004).
35. D. D. Leipe, E. V. Koonin, L. Aravind, *J. Mol. Biol.* **343**, 1–28 (2004).
36. O. Danot, E. Marquenet, D. Vidal-Ingigliardi, E. Richet, *Structure* **17**, 172–182 (2009).
37. E. V. Koonin, L. Aravind, *Cell Death Differ.* **9**, 394–404 (2002).
38. A. Bernheim, R. Sorek, *Nat. Rev. Microbiol.* **18**, 113–119 (2020).

ACKNOWLEDGMENTS

We thank A. Ladha, J. Strecker, S. Kannan, J. Kreitz, and G. Faure for valuable discussions and experimental assistance; R. Macrae for a critical reading of the manuscript; S. Richards and R. Belliveau for assistance with ordering bacterial strains; and the entire Zhang lab for support and advice. **Funding:** K.S.M. and E.V.K. are supported by intramural funds of the U.S. Department of Health and Human Services (to National Library of Medicine). F.Z. is supported by the National Institutes of Health (grants 1R01-HG009761, 1R01-MH110049, and 1DP1-HL141201); the Howard Hughes Medical Institute; the Open Philanthropy Project, the Harold G. and Leila Mathers and Edward Mallinckrodt, Jr. Foundations; the Poitras Center for Psychiatric Disorders Research at MIT; the Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT; and by the Phillips family and J. and P. Poitras.
Author contributions: L.G. and F.Z. conceived of the project. L.G. and H.A.-T. performed defense island analysis. L.G. cloned defense systems and performed transcriptome sequencing. L.G. performed plaque assays with assistance from F.B., M.S., and J.K. L.G. and F.B. performed bacterial density and phage fragmentation assays. H.A.-T. performed phylogenetic distribution and prophage analysis. L.G., H.A.-T., F.B., K.S.M., J.L.S.-B., Y.I.W., E.V.K., and F.Z. analyzed data. F.Z. supervised the research and experimental design. L.G., H.A.-T., E.V.K., and F.Z. wrote and revised the manuscript with input from all authors. **Competing interests:** F.Z. is a scientific advisor and cofounder of Editas Medicine, Beam Therapeutics, Pairwise Plants, Arbor Biotechnologies, and Sherlock Biosciences. L.G, J.L.S.-B., and F.Z. are co-inventors on US provisional application no. 62/928,269, which includes bacterial defense systems described in this manuscript. **Data and materials availability:** Expression plasmids are available from Addgene (nos. 157879 to 157912) under the Uniform Biological Material Transfer Agreement. The genome assembly of *E. coli* phage ϕV-1 has been deposited in GenBank (accession number MT542512). All other data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1077/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S13
References (*39*–*60*)
MDAR Reproducibility Checklist

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

29 October 2019; accepted 6 July 2020
10.1126/science.aba0372

QUANTUM COMPUTING

Hartree-Fock on a superconducting qubit quantum computer

Google AI Quantum and Collaborators*†

The simulation of fermionic systems is among the most anticipated applications of quantum computing. We performed several quantum simulations of chemistry with up to one dozen qubits, including modeling the isomerization mechanism of diazene. We also demonstrated error-mitigation strategies based on *N*-representability that dramatically improve the effective fidelity of our experiments. Our parameterized ansatz circuits realized the Givens rotation approach to noninteracting fermion evolution, which we variationally optimized to prepare the Hartree-Fock wave function. This ubiquitous algorithmic primitive is classically tractable to simulate yet still generates highly entangled states over the computational basis, which allowed us to assess the performance of our hardware and establish a foundation for scaling up correlated quantum chemistry simulations.

The prediction of molecular properties and chemical reactions from ab initio quantum mechanics has emerged as one of the most promising applications of quantum computing (*1*). This fact is due both to the commercial value of accurate simulations as well as the relatively modest number of qubits required to represent interesting instances. However, as the age of "quantum supremacy" dawns (*2*), so has a more complete appreciation of the challenges required to scale such computations to the classically intractable regime using near-term intermediate scale quantum (NISQ) devices. Achieving that objective will require further algorithmic innovations, hardware with more qubits and low error rates, and more effective error-mitigation strategies. Here, we report a variational quantum eigensolver (VQE) (*3*) simulation of molecular systems with progress in all three directions.

We used the Google Sycamore quantum processor to simulate the binding energy of hydrogen chains as large as H_{12}, as well as a chemical reaction mechanism (the isomerization of diazene). The Sycamore quantum processor consists of a two-dimensional array of 54 transmon qubits (*2*). Each qubit is tunably coupled to four nearest neighbors in a rectangular lattice. Our largest simulations used a dozen qubits—twice the size as the largest prior quantum simulations of chemistry (*4*)—and required only nearest-neighbor coupling (Fig. 1). Prior simulations of chemistry on superconducting qubit devices and trapped ion systems demonstrated the possibility of error mitigation through VQE (*4*–*10*), albeit on a small scale. We demonstrated that within the model, achieving chemical accuracy through VQE is possible for intermediate-scale problems when combined with effective error mitigation strategies. Furthermore, we argue that the circuit ansatz we used for VQE is especially appealing as a benchmark for chemistry.

We simulated quantum chemistry in a second-quantized representation in which the state of each of N qubits encoded the occupancy of an orbital basis function. We used what are commonly referred to as "core orbitals" as the initial orbitals (shown for H_{12} in Fig. 1A, left), which are the eigenfunctions of the molecular Hamiltonian without the electron-electron interaction term. The goal of this experiment was to use a quantum computer to implement the Hartree-Fock procedure, which is a method for obtaining the best single-particle orbital functions assuming each electron feels the average potential generated from all the other electrons. This assumption is enforced by constraining the wave function to be a product of one-particle functions, which has been appropriately antisymmetrized to satisfy the Pauli exclusion principle. An initial guess for the Hartree-Fock state, from which we can optimize the orbitals, was obtained by filling the lowest-energy $\eta/2$ orbitals, each with a spin-up electron and a spin-down electron, where η is the number of electrons. Because we simulated the singlet ground state for all molecules considered here, there is no spin component to the mean-field approximation; thus, we only needed to explicitly simulate the $\eta/2$ spin-up electrons.

By performing a unitary rotation of the initial (core) orbital basis $\varphi_p(r)$, one can obtain a new valid set of orbitals $\tilde{\varphi}_p(r)$ as a linear combination of the initial ones:

$$\tilde{\varphi}_p(r) = \sum_{q=1}^{N} [e^{\kappa}]_{pq} \varphi_q(r) \qquad (1)$$

where κ is an $N \times N$ anti-Hermitian matrix and $[e^{\kappa}]_{pq}$ is the p, q element of the matrix exponential of κ. A result attributable to Thouless (*11*) is that one can express the unitary that applies this basis rotation to the quantum state as time-evolution under a noninteracting fermion Hamiltonian. Specifically, if we take $a_p^\dagger$ and a_p to be fermionic creation and annihilation operators, respectively, for the core orbital $\varphi_p(r)$, then we can parameterize $|\psi_\kappa\rangle$, an antisymmetric product state in the new basis $\tilde{\varphi}_p(r)$, as noninteracting fermion dynamics from a computational basis state $|\eta\rangle = a_\eta^\dagger \cdots a_1^\dagger |0\rangle$ in the core orbital basis:

$$|\psi_\kappa\rangle = U_\kappa |\eta\rangle,$$
$$U_\kappa = \exp\left(\sum_{p,q=1}^{N} \kappa_{pq} a_p^\dagger a_q\right) \qquad (2)$$

Such states are referred to as Slater determinants.

To complete the accurate preparation of Hartree-Fock states, we implemented variational relaxation of the κ parameters to minimize the energy of $|\psi_\kappa\rangle$ starting from the optimal κ determined by solving the Hartree-Fock equations classically. This is an idealized implementation of VQE that allowed us to demonstrate error mitigation of coherent errors through variational relaxation. We defined the Hartree-Fock state $|\psi_{HF}\rangle$ to be the lowest-energy Slater determinant for the molecular Hamiltonian H

$$|\psi_{HF}\rangle = |\psi_{\kappa^*}\rangle, \kappa^* = \arg\min_\kappa \langle\psi_\kappa|H|\psi_\kappa\rangle \qquad (3)$$

We applied U_κ to $|\eta\rangle$ using our quantum computer and then performed the optimization over κ through feedback from a classical optimization routine. The energy decreased because the initial core orbitals were obtained by ignoring the electron-electron interaction and variational relaxation compensates for coherent errors. Because the generator for U_κ corresponds to a noninteracting fermion Hamiltonian, its action on a product state in second quantization can be classically simulated in $O(N^3)$ by diagonalizing the one-body operator, and in some cases, the Hartree-Fock procedure can be made to converge with even lower complexity. Despite this, we argue that this procedure is still a compelling experiment for a quantum computer.

The Hartree-Fock state is usually the initial state for classical correlated electronic structure calculations such as coupled cluster and configuration interaction methods, as well as for many quantum algorithms for chemistry. Thus, often one chooses to work in the molecular orbital basis, which is defined so that the Hartree-Fock state is a computational basis state. However, the molecular orbital basis Hamiltonian has a large number of terms that can be challenging to simulate and measure with low complexity. Accordingly, the most efficient quantum algorithms for chemistry (*12*–*15*) require that one perform the simulation in more structured bases with asymptotically fewer terms (*16*–*18*), necessitating that U_{κ^*} is applied explicitly at the beginning of the computation. Even when simulating chemistry

*Corresponding author. Email: nickrubin@google.com (N.C.R.); babbush@google.com (R.J.B.) †Google AI Quantum authors, collaborators, and affiliations are listed in the supplementary materials.

in an arbitrary basis, the most efficient strategies are based on a tensor factorization of the Hamiltonian that requires many applications of U_κ to simulate (*19*, *20*). Exploiting this tensor factorization with basis rotations is also key to the most efficient strategy for measuring $\langle H \rangle$ in variational algorithms and requires implementing U_κ before measurement (*21*).

We used this variational ansatz based on basis rotations to benchmark the Sycamore processor for linear hydrogen chains of length 6, 8, 10, and 12 and two pathways for diazene bond isomerization. We modeled hydrogen chains of length N with N qubits. Our simulations required N qubits to simulate $2N$ spin orbitals because of the constraint that the α–spin orbitals have the same spatial wave function as that of the β–spin orbitals. For diazene, we required 10 qubits after preprocessing. The hydrogen chains are a common benchmark in electronic structure (*22*–*24*), and the diazene bond isomerization provides a system in which the required accuracy is more representative of typical electronic structure problems and has been used as a benchmark for coupled-cluster methods (*25*). For the diazene isomerization, our goal was to resolve the energetic difference between the transition states of two competing mechanisms, requiring accuracy of about 40 millihartree. This objective differs from prior quantum simulations of chemistry, which have focused on bond dissociation curves (*4*–*7*).

One motivation for this work was to calibrate and validate the performance of our device in realizing an important algorithmic primitive for quantum chemistry and lattice model simulation. Our experiment was also appealing for benchmarking purposes because the circuits we explored generated highly entangled states but with a special structure that enabled the efficient measurement of fidelity and the determination of systematic errors. Further motivation was to implement the largest variational quantum simulation of chemistry so that it would be possible to better quantify the current gap between the capabilities of NISQ devices and real applications. Even though the Hartree-Fock ansatz is efficient to simulate classically, the circuits in our experiment are far more complex than prior experimental quantum simulations of chemistry. Last, the structure of the Hartree-Fock state enabled us to sample the energy and gradients of the variational ansatz with fewer measurements than typically would be required, allowing us to focus on other aspects of simulating chemistry at scale, such as the effectiveness of various types of error mitigation. Thus, our choice to focus on Hartree-Fock for this experiment embraces the notion that we should work toward valuable quantum simulations of chemistry by first scaling up important components of the exact solution (such as error-mitigation strategies and basis rotations) in a fashion that enables us to completely understand and perfect those primitives.

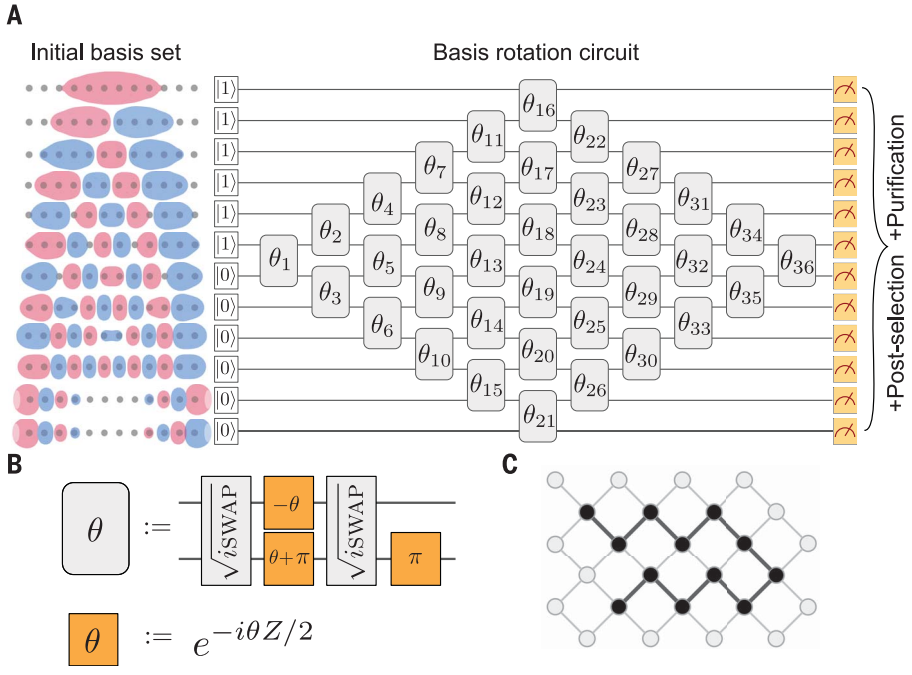


Fig. 1. Basis rotation circuit and compilation. (**A**) To the left of the circuit diagram are the initial orbitals for the H_{12} chain with atom spacings of 1.3 Å, obtained by diagonalizing the Hamiltonian, ignoring electron-electron interactions. The circuit diagram depicts the basis rotation ansatz for a linear chain of 12 hydrogen atoms. Each gray box with a rotation angle θ represents a Givens rotation gate. (**B**) Compilation of the Givens rotation gate to $\sqrt{i\text{SWAP}}$ gates and single-qubit gates that can be realized directly in hardware. The H_{12} circuit involves 72 $\sqrt{i\text{SWAP}}$ gates and 108 single-qubit Z rotation gates with a total of 36 variational parameters. (**C**) Depiction of a 12-qubit line on a subgrid of the entire 54-qubit Sycamore device. All circuits only require gates between pairs of qubits that are adjacent in a linear topology.

Variational algorithms are specified in the form of a functional minimization. This minimization has three main components: ansatz specification in the form of a parameterized quantum circuit (the function), observable estimation (the functional), and outer-loop optimization (the minimization). Each component is distinctively affected by our choice to simulate a model corresponding to noninteracting fermion wave functions. Symmetries built into this ansatz allowed for reduction of the number of qubits required to simulate molecular systems, a reduction in the number of measurements needed to estimate the energy, and access to the gradient without additional measurements beyond those required for energy estimation. Details on how we realized Hartree-Fock with VQE are available in the supplementary materials, section A.

The unitary in Eq. 2 can be compiled exactly (without Trotterization) by using a procedure based on Givens rotations. This strategy was first suggested for quantum computing in work on linear optics in (*26*) and later in the context of fermionic simulations in (*27*). We implemented these basis rotations using the optimal compilation of (*28*), which has gate depth $N/2$ and requires only $\eta(N - \eta)$ two-qubit "Givens rotation" gates on a linearly connected architecture, giving one rotation for each element in the unitary basis change. These Givens rotation gates were implemented by decomposition into two $\sqrt{i\text{SWAP}}$ gates and three Rz gates. In Fig. 1, we depict the basis change circuit for the H_{12} chain, which has a diamond-shaped structure. We further review the compilation of these circuits in the supplementary materials, section B.

The average energy of any molecular system can be evaluated with knowledge of the one-particle reduced-density matrix (1-RDM), $\langle a_p^\dagger a_q \rangle$, and the two-particle reduced-density matrix (2-RDM), $\langle a_p^\dagger a_q^\dagger a_r a_s \rangle$. In general, it is not possible to exactly reconstruct the 2-RDM from knowledge of just the 1-RDM. However, for single-Slater determinants (as in our Hartree-Fock experiment), the 2-RDM is completely determined by the 1-RDM (*29*):

$$\left\langle a_p^\dagger a_q^\dagger a_r a_s \right\rangle = \left\langle a_p^\dagger a_s \right\rangle \left\langle a_q^\dagger a_r \right\rangle - \left\langle a_q^\dagger a_s \right\rangle \left\langle a_p^\dagger a_r \right\rangle \quad (4)$$

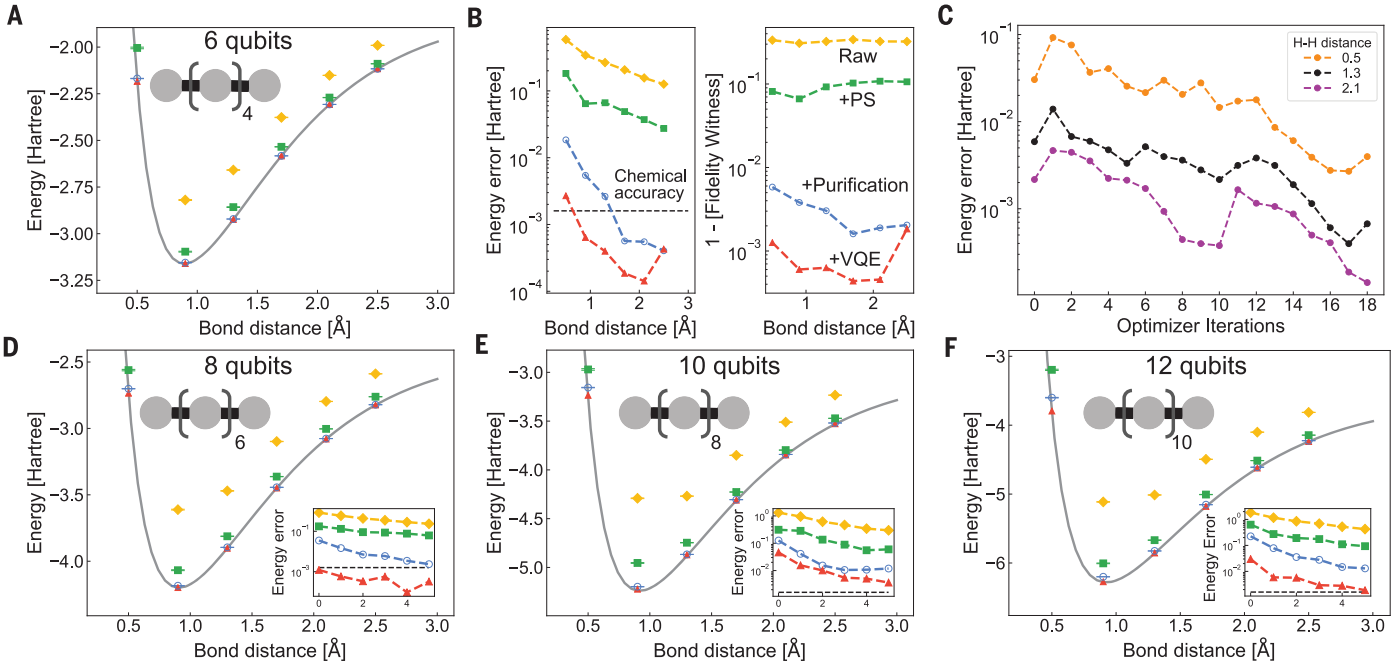


Fig. 2. Static and VQE performance on hydrogen chains. Binding curve simulations for H_6, H_8, H_{10}, and H_{12} with various forms of error mitigation. (**A** and **D** to **F**) Comparison of Sycamore's raw performance (yellow diamonds) with postselection (green squares), purification (blue circles), and error-mitigated combined with variational relaxation (red triangles). For all hydrogen systems, the raw data at 0.5-Å bond length is off the top of the plot. The yellow, green, and blue points were calculated by using the optimal basis rotation angles computed from a classical simulation; thus, the variational optimization shown here is only used to correct systematic errors in the circuit realization. (**B**) The absolute error and infidelity for the H_6 system. For all points, we calculated a fidelity witness described in the supplementary materials, section D. The error bars for all points were computed by estimating the covariance between simultaneously measured sets of 1-RDM elements and resampling those elements under a multivariate Gaussian model. Energies from each sample were tabulated, and the standard deviation is used as the error bar. The "+PS" means applying postselection to the raw data, "+Purification" means applying postselection and McWeeny purification, and "+VQE" means postselection, McWeeny purification, and variational relaxation. (**C**) Optimization traces for three H_6 geometries (bond distances of 0.5, 1.3, and 2.1 Å). All optimization runs used between 18 and 30 optimization. The lowest-energy solution from the optimization trace was reported.

Thus, in our experiment we only needed to sample the 1-RDM to estimate the energy. Because the 2-RDM has quadratically more elements than the 1-RDM, this approach is a substantial simplification. We measured the 1-RDM using a protocol described in the supplementary materials, section C. This protocol enabled us to optimally parallelize measurement of all N^2 1-RDM elements with $N + 1$ distinct circuits. For each distinct circuit, we made 250,000 measurements.

We performed two types of error mitigation on our measured data: postselection on particle number (conserved in basis rotations) and pure-state projection. To apply postselection, we modified our circuits by first rotating into a basis that diagonalizes $a_p^\dagger a_q + a_q^\dagger a_p$ for N different pairs of p and q so that these elements could be sampled at the same time as the total particle-number operator. Following the strategy in the supplementary materials, section C, this measurement was accomplished at the cost of two T gates and one $\sqrt{i\mathrm{SWAP}}$ gate per pair of qubits. We then postselected to discard measurements where the total number of excitations changed from $\eta/2$.

For pure-state purification, we leveraged the fact that the 1-RDM of any single-Slater determinant wave function ψ_κ has eigenvalues restricted to be 0 and 1 (*30*). We performed projection back to the pure set of 1-RDMs using a technique known as McWeeny purification (*29*). Details on the procedure and sampling bounds for guaranteeing that the procedure has a fixed-point 1-RDM corresponding to a Slater determinant can be found in the supplementary materials, section E. Although McWeeny purification only works for Slater determinant wave functions, pure-state N-representability conditions are known for more general systems (*31*), and we expect that a computational procedure similar to enforcing ensemble constraints could be used (*32*, *33*).

A variety of circuit optimization techniques based on gradient and gradient-free methods have been proposed in the context of NISQ algorithms. We developed an optimization technique that exploits local gradient and Hessian information in a fashion that is distinctive to the Hartree-Fock model. It is based on a proposal for iterative construction of a wave function to satisfy the Brillouin condition for a single-particle model (*34*). Our optimization protocol used the property that at a local optima, the commutator of the Hamiltonian H with respect to any generator of rotation G is zero—$\langle\psi|[H, G]|\psi\rangle = 0$—and sequential basis change circuits can be concatenated into a single basis change circuit ($U_a U_b = U_{ab}$). Using these relations and taking $G = \sum_{pq} \kappa_{pq} a_p^\dagger a_q$, as in our experiment, the double commutator $\langle\psi|[H, G], G|\psi\rangle$ determined an augmented Hessian (matrix of derivatives) that we could use to iteratively update the wave function so that the first-order condition was approximately satisfied. Regularization was added by limiting the size of update parameters (*35*). Details are provided in the supplementary materials, section H.

As a benchmark, we studied symmetrically stretched hydrogen chains of lengths 6, 8, 10, and 12 atoms (Fig. 2). The initial parameters were set to the parameters obtained by solving the Hartree-Fock equations on a classical computer. The data from the quantum computer were plotted along with classical Hartree-Fock results, showing better and better agreement as we added postselection, postselection and purification, and then error-mitigated variational relaxation. The 6- and 8-qubit data achieved chemical accuracy after VQE, and even the 12-qubit data followed the expected energy closely. The error data in Fig. 2B and

the other inserts show a large and consistent decrease, about a factor of 100, when using these protocols. The substantial decrease in error when using a modest number of VQE iterations is shown in Fig. 2C.

A fidelity witness can be efficiently computed from the experimental data (supplementary materials, section D.2) (*36*). This value is a lower bound to the true fidelity and thus potentially loose when fidelity is small. However, this fidelity generally tracked the measured errors (Fig. 2B). As shown in Table 1, fidelity increased as we added various forms of error mitigation, starting in the left column, where the optimal angles were computed classically. Uncertainties in the last digit, indicated in parentheses, were calculated by means of the procedure described in the supplementary materials, section C.5. The first column of Table 1 is an estimate of the fidelity based on multiplying the fidelity for all the gates and readout assuming 99.5% fidelity for single qubit gates, 99% fidelity for two-qubit gates, and 97% fidelity for readout. We see that this estimate qualitatively follows the "raw" fidelity witness estimates, except when the witness value is very small. For all hydrogen systems studied, we observed drastic fidelity improvements, with combined error mitigation.

We simulated two isomerization pathways for diazene, marking the first time that a chemical reaction mechanism has been modeled by using a quantum computer. Hartree-Fock theory reverses the order of the transition states; however, we focused on the accuracy of the computation with respect to the simulated model. Correctly identifying this pathway requires resolving the energy gap of 40 millihartree between the two transition states. The pathways correspond to the motion of the hydrogen in the process of converting *cis*-diazene to *trans*-diazene. One mechanism is in-plane rotation of a hydrogen, and the other is an out-of-plane rotation corresponding to rotation of the HNNH dihedral angle. VQE-optimized data simulating nine points along the reaction coordinates for in-plane and out-of-plane rotation of hydrogen are shown in Fig. 3. For all points along the reaction coordinate, the initial parameter setting was the solution to the Hartree-Fock equations. VQE produced 1-RDMs with an average fidelity greater than 0.98 after error mitigation. Our full error-mitigation procedure again substantially improved the accuracy of our calculation.

Our VQE calculations on diazene predicted the correct ordering of the transition states to within the chemical model with an energy gap of 41 ± 6 millihartree, and the true gap is 40.2 millihartree. We provide a more detailed analysis of the error mitigation performance on the diazene circuits in the supplementary materials, section F, considering that the $\sqrt{i\mathrm{SWAP}}$ gates we used had a residual CPHASE ($\pi/24$) and Rz gates had stochastic control angles. This simulation reinforced VQE's ability to mitigate systematic errors at the scale of 50 $\sqrt{i\mathrm{SWAP}}$ gates and over 80 Rz gates.

In this work, we took a step toward answering the question of whether NISQ computers can offer quantum advantage for chemical simulation by studying VQE performance on basis rotation circuits that are widely used in quantum algorithms for fermionic simulation. The considered ansatz afforded ways to minimize the resource requirements for VQE and study device performance for circuits that are similar to those needed for full Hamiltonian simulation. These basis rotation circuits also made an attractive benchmark because of their prevalence, optimal known compilation, and ability to extract fidelity and fidelity witness

Table 1. Average-fidelity lower bounds for hydrogen chain calculations. We report values of the fidelity witness from (*36*), averaged across H–H separations of {0.5, 0.9, 1.3, 1.7, 2.1, 2.5} Å, starting from circuits with the theoretically optimal variational parameters (κ). "Estimate" corresponds to an estimate of the fidelity derived by multiplying gate errors assuming 0.5% single-qubit gate error, 1% two-qubit gate error, and 3% readout error. "Raw" corresponds to fidelities from constructing the 1-RDM without any error mitigation. "+PS" corresponds to fidelities from constructing the 1-RDM with postselection on particle number. "+Pure" corresponds to fidelities from constructing the 1-RDM with postselection and applying purification as postprocessing. "+VQE" corresponds to fidelities from using all previously mentioned error mitigation techniques in conjunction with variational relaxation. Numbers in parentheses are the uncertainties of the last, preceding digit. For small values (such as the "raw" value for H_{12}), we expect that the fidelity lower-bound is more likely to be loose.

System	Estimate	Raw	+PS	+Pure	+VQE
H_6	0.571	0.674(2)	0.906(2)	0.9969(1)	0.99910(9)
H_8	0.412	0.464(2)	0.827(2)	0.9879(3)	0.99911(8)
H_{10}	0.277	0.316(2)	0.784(3)	0.9704(5)	0.9834(4)
H_{12}	0.174	0.010(2)	0.654(3)	0.9424(9)	0.9913(3)

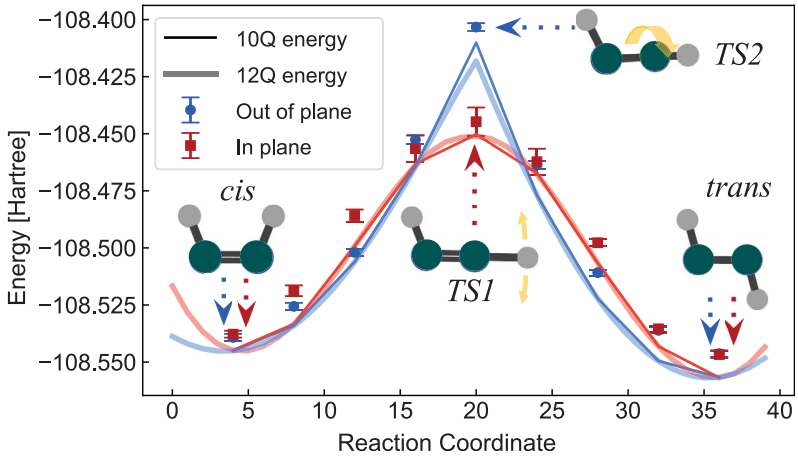


Fig. 3. VQE performance on distinguishing the mechanism of diazene isomerization. Hartree-Fock curves for diazene isomerization between cis and trans configurations. TS1 and TS2 are the transition states for the in-plane and out-of-plane rotation of the hydrogen, respectively. The yellow arrows on TS1 and TS2 indicate the corresponding reaction coordinate. The solid curve is the energy obtained from optimizing a 10-qubit problem generated by freezing the core orbitals generated from two self-consistent field cycles. The transparent lines of the same color are the full 12-qubit system, indicating that freezing the lowest two levels does not change the characteristics of the model chemistry. Nine points along the reaction paths are simulated on Sycamore by using VQE. We allowed the optimizer 30 iterations for all points, except for fifth and sixth points from the left of the in-plane rotation curve, for which we allowed 60 steps. The error bars for all points were computed by estimating the covariance between simultaneously measured sets of 1-RDM elements and resampling those elements under a multivariate Gaussian model. Energies from each sample were tabulated, and the standard deviation is used as the error bar. No purification was applied for the computation of the error bar. If purification is applied, the error bars become smaller than the markers. Each basis rotation for diazene contains 50 $\sqrt{i\mathrm{SWAP}}$ gates and 80 Rz gates.

values and because they parameterize a continuous family of analytically solvable circuits that demonstrate a high degree of entanglement. The circuits also serve as a natural progression toward more correlated ansatz such as a generalized swap network (*28*) or a non–particle-conserving circuit ansatze followed by particle-number projection.

We demonstrated the performance of two error-mitigation techniques on basis rotation circuit fidelity. The first is postselection on total occupation number when measuring all elements of the 1-RDM. This step was accomplished by permuting the basis rotation circuit so that all measurements involved estimating nearest-neighbor observables and measuring each pair of observables so that the total occupation number is preserved. The second is the application of McWeeny purification as a postprocessing step. The energy improvements from projecting back to the pure-state N-representable manifold was evidence that generalized pure-state N-representability conditions would be instrumental in making NISQ chemistry computations feasible. This underscores the importance of developing procedures for applying pure-state N-representability conditions in a more general context. The postselection and RDM measurement techniques can be generalized to measuring all 1-RDM and 2-RDM elements when considering a less restrictive circuit ansatz by permuting the labels of the fermionic modes. For ansatz such as the generalized swap network (*28*), the circuit structure would not change, only the rotation angles. Thus, the measurement schemes presented here are applicable in the more general case. Furthermore, it is important to understand the performance of these error-mitigation techniques when combined with alternatives such as noise extrapolation (*37*).

Last, we were able to show further evidence that variational relaxation effectively mitigates coherent errors that arise in the implementation of physical gates. The performance of our problem-specific optimization strategy motivates the study of iterative wave function constructions (*38*) in a more general setting. The combination of these error-mitigation techniques with VQE unambiguously resolved a chemical mechanism to within the model chemistry by using a quantum computation. It is still an open question whether NISQ devices will be able to simulate challenging quantum chemistry systems, and it is likely that major innovations will be required. However, we find the accuracy of these experiments and the effectiveness of these error-mitigation procedures to be an encouraging signal of progress in that direction.

REFERENCES AND NOTES

1. A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, M. Head-Gordon, *Science* **309**, 1704–1707 (2005).
2. F. Arute *et al*., *Nature* **574**, 505–510 (2019).
3. A. Peruzzo *et al*., *Nat. Commun.* **5**, 4213 (2014).
4. A. Kandala *et al*., *Nature* **549**, 242–246 (2017).
5. P. J. J. O'Malley *et al*., *Phys. Rev. X* **6**, 031007 (2016).
6. C. Hempel *et al*., *Phys. Rev. X* **8**, 031022 (2018).
7. A. Kandala *et al*., *Nature* **567**, 491–495 (2019).
8. J. I. Colless *et al*., *Phys. Rev. X* **8**, 011021 (2018).
9. S. E. Smart, D. A. Mazziotti, *Phys. Rev. A* **100**, 022517 (2019).
10. R. Sagastizabal *et al*., *Phys. Rev. A* **100**, 010302 (2019).
11. D. J. Thouless, *Nucl. Phys.* **21**, 225–232 (1960).
12. G. H. Low, N. Wiebe, arXiv:1805.00675 [quant-ph] (2018).
13. R. Babbush, D. W. Berry, J. R. McClean, H. Neven, *npj Quantum Information* **5**, 92 (2019).
14. A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, S. Zhu, arXiv:1912.08854 [quant-ph] (2019).
15. R. Babbush *et al*., *Phys. Rev. X* **8**, 041015 (2018).
16. R. Babbush *et al*., *Phys. Rev. X* **8**, 011044 (2018).
17. S. R. White, *J. Chem. Phys.* **147**, 244102 (2017).
18. J. R. McClean *et al*., arXiv:1909.00028 [quant-ph] (2019).
19. M. Motta *et al*., arXiv:1808.02625 [physics.comp-ph] (2018).
20. D. Berry, C. Gidney, M. Motta, J. McClean, R. Babbush, *Quantum* **3**, 208 (2019).
21. W. J. Huggins *et al*., arXiv:1907.13117 [quant-ph] (2019).
22. M. Motta *et al*., *Phys. Rev. X* **7**, 031059 (2017).
23. P. A. Limacher *et al*., *J. Chem. Theory Comput.* **9**, 1394–1401 (2013).
24. J. Hachmann, W. Cardoen, G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).
25. R. K. Chaudhuri, K. F. Freed, S. Chattopadhyay, U. Sinha Mahapatra, *J. Chem. Phys.* **128**, 144304 (2008).
26. M. Reck, A. Zeilinger, H. J. Bernstein, P. Bertani, *Phys. Rev. Lett.* **73**, 58–61 (1994).
27. D. Wecker *et al*., *Phys. Rev. A* **92**, 062318 (2015).
28. I. D. Kivlichan *et al*., *Phys. Rev. Lett.* **120**, 110501 (2018).
29. R. McWeeny, *Rev. Mod. Phys.* **32**, 335–369 (1960).
30. A. J. Coleman, *Rev. Mod. Phys.* **35**, 668–686 (1963).
31. D. A. Mazziotti, *Phys. Rev. A* **94**, 032516 (2016).
32. N. C. Rubin, R. Babbush, J. McClean, *New J. Phys.* **20**, 053020 (2018).
33. A. A. Klyachko, *J. Phys.* **36**, 72 (2006).
34. W. Kutzelnigg, *Chem. Phys. Lett.* **64**, 383–387 (1979).
35. Q. Sun, arXiv:1610.08423 [physics.chem-ph] (2016).
36. M. Gluza, M. Kliesch, J. Eisert, L. Aolita, *Phys. Rev. Lett.* **120**, 190501 (2018).
37. K. Temme, S. Bravyi, J. M. Gambetta, *Phys. Rev. Lett.* **119**, 180509 (2017).
38. H. R. Grimsley, S. E. Economou, E. Barnes, N. J. Mayhall, *Nat. Commun.* **10**, 3007 (2019).
39. J. R. McClean *et al*., arXiv:1710.07629 [quant-ph] (2017).
40. F. Arute *et al.*, Hartree-fock on a superconducting qubit quantum computer. Dryad, Dataset (2020); doi:10.5061/dryad.2jm63xsm2.

ACKNOWLEDGMENTS

D.B. is a CIFAR Associate Fellow in the Quantum Information Science Program. **Funding:** This work was supported by Google. **Competing interests:** The authors declare no competing interests. **Author contributions:** N.C.R. designed the experiment. C.N. assisted with data collection. Z.J., V.S., and N.W. assisted with analytical calculations and gate synthesis. N.C.R. and R.J.B. wrote the paper. Experiments were performed by using a quantum processor that was recently developed and fabricated by a large effort involving the entire Google Quantum team. **Data and materials availability:** The code used for this experiment and a tutorial for running it can be found in the open-source library ReCirq, located at https://github.com/quantumlib/ReCirq (*39*). All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. Data presented in the figures can be found in the Dryad repository (*40*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1084/suppl/DC1
Supplementary Text
Figs. S1 to S11
Tables S1 to S11
References (*41*–*50*)

1 April 2020; accepted 18 June 2020
10.1126/science.abb9811

STRUCTURAL BIOLOGY

Architecture of a catalytically active homotrimeric plant cellulose synthase complex

Pallinti Purushotham*, Ruoya Ho*, Jochen Zimmer†

Cellulose is an essential plant cell wall component and represents the most abundant biopolymer on Earth. Supramolecular plant cellulose synthase complexes organize multiple linear glucose polymers into microfibrils as load-bearing wall components. We determined the structure of a poplar cellulose synthase CesA homotrimer that suggests a molecular basis for cellulose microfibril formation. This complex, stabilized by cytosolic plant-conserved regions and helical exchange within the transmembrane segments, forms three channels occupied by nascent cellulose polymers. Secretion steers the polymers toward a common exit point, which could facilitate protofibril formation. CesA's N-terminal domains assemble into a cytosolic stalk that interacts with a microtubule-tethering protein and may thus be involved in CesA localization. Our data suggest how cellulose synthase complexes assemble and provide the molecular basis for plant cell wall engineering.

A main product of plant photosynthesis is glucose, which is a precursor of many naturally occurring carbohydrates and a building block for cell wall materials in most photosynthetic organisms (*1*). Cellulose, produced by cellulose synthase (CesA), is a linear β-1,4–linked glucose polymer and a fundamental cell wall component in land plants (*2*). With an estimated production exceeding 10 billion metric tons annually (*3*), it represents the most abundant biopolymer on Earth as well as a vast carbon dioxide sink and energy resource. CesA is a membrane-embedded family-2 glycosyltransferase (GT) related to chitin, alginate, and hyaluronan synthases (*4*, *5*). The enzyme catalyzes glucose transfer from uridine diphosphate–glucose (UDP-Glc, the donor) to the terminal C4 hydroxyl of the nascent cellulose polymer (the acceptor) (fig. S1A).

CesA combines two functions: It synthesizes cellulose and secretes the polymer through a transmembrane (TM) channel formed by its membrane-spanning segment (*6*). The translocation mechanism has been described in detail for the bacterial homolog BcsA. Cellulose synthase uses an evolutionarily conserved gating loop and finger helix to couple the binding of a new substrate molecule after cellulose elongation with the translocation of the polymer into its TM channel (*7*). Despite similar biosynthesis mechanisms, bacterial and plant cellulose differ substantially. Whereas most bacteria [with some exceptions (*8*)] produce amorphous cellulose (*9*), plant CesAs organize multiple cellulose chains into micro- and macrofibrils (CMFs) that are spun around the cell to form the wall's load-bearing components (*10*).

Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, 480 Ray C. Hunt Dr., Charlottesville, VA 22903, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: jochen_zimmer@virginia.edu

Most plants express several CesA isoforms with different subsets required for primary and secondary cell wall formation (*10*). Reconstitutions of single CesA isoforms from *Physcomitrella patens* and *Populus tremula x tremuloides* (Ptt) demonstrated that, in vitro, single isoforms devoid of any additional plant-derived factors are sufficient to form cellulose fibers resembling authentic wall CMFs (*11*, *12*). CMFs are likely synthesized by supramolecular CesA complexes (CSCs) in the plasma membrane. Accordingly, plant CesAs contain three distinct regions likely mediating oligomerization. First, their N-terminal domains (NTDs) form extended cytosolic structures ~140 to 160 residues long with a RING-like domain near the extreme N terminus followed by hypervariable regions leading into the first TM helix. The isolated NTD from cotton CesA1 has been shown to dimerize in vitro (*13*). Second, a plant-conserved region (PCR; ~118 residues) is inserted into the catalytic GT domain at the membrane-distal side, and third, a class-specific region (CSR; ~124 residues) follows an evolutionarily conserved substrate-binding motif near the active site (*14*, *15*). In the plasma membrane, CesAs assemble into pseudo–sixfold symmetric CSCs, termed rosettes (*16*). CSCs were proposed to consist of hexamers of either trimers, tetramers, or hexamers, thereby giving rise to 18, 24, or 36 CesAs per CSC, respectively, and synthesizing CMFs with a corresponding maximum number of cellulose chains (*17*).

To reveal the mechanism of plant CMF formation, we determined the cryo–electron microscopy (cryo-EM) structure of a homo-oligomeric PttCesA8 complex, implicated in secondary cell wall formation in vivo (*10*), at an intermediate state during cellulose biosynthesis at ~3.5-Å resolution. The structure reveals a trimeric PttCesA8 complex in which each enzyme is associated with a nascent polysaccharide. Our data reveal how CesA trimers are stabilized to steer cellulose chains into protofibrils that could align into CMFs when synthesized by a rosette consisting of six CesA trimers.

Results

Homo-oligomerization of PttCesA8

PttCesA8 was recombinantly expressed in SF9 insect cells with an N-terminal polyhistidine tag facilitating affinity purification (materials and methods). The protein was solubilized from the membrane fraction in a combination of the detergents lauryl maltose neopentyl glycol and cholesteryl-hemisuccinate and exchanged into glyco-diosgenin during purification (fig. S1B). Under these conditions, the purified enzyme is catalytically active and synthesizes a water-insoluble glucan in the presence of UDP-Glc and Mg^{2+} or Mn^{2+} (fig. S1C). The polymer is completely degraded by a β-1,4–specific glucanase, consistent with the synthesis of authentic cellulose, as also previously described for proteoliposome-reconstituted PttCesA8 (*12*). Although most of the in vitro–synthesized cellulose is amorphous and not detectable by negative-stain EM analyses, some thin fibers were observed ~10 to 15 Å in diameter, perhaps representing cellulose protofibrils containing a small number of cellulose chains (fig. S1D).

To test the ability of PttCesA8 to self-assemble into homo-oligomers, we coexpressed three variants carrying either an N-terminal tandem His-Strep–, FLAG-, or Myc-tag in *Pichia pastoris*. Purification of the recombinant proteins by sequential binding to four affinity matrices (TALON resin, followed by anti-Strep, anti-FLAG, and anti-Myc matrices) revealed the copurification of all three PttCesA8 species in a functional form (fig. S1, E and F). Combined with our structural analyses described below, these interaction studies suggest that PttCesA8 can assemble into catalytically active homo-oligomers.

Architecture of the PttCesA8 homotrimer

Structural insights into PttCesA8 homo-oligomers were obtained by cryo-EM (table S1). To minimize sample dissociation during purification, we designed a fast purification scheme consisting of immobilized metal affinity and gel filtration chromatography of the N-terminally His-tagged PttCesA8. Cryo grids were prepared immediately after purification without addition of UDP-Glc, as explained in detail in the materials and methods. The two-dimensional classification of ~3,000,000 particles revealed the presence of small (~85%) and large (~15%) particles, in agreement with monomers and trimers, with the monomers likely arising from detergent-dissociated trimers (Fig. 1A). In both species, the enzyme's prominent cytosolic GT domain protrudes from the micelle-embedded TM segment. Yet only the trimeric complexes contain an additional stalk extending from the

GT domains, which is formed by PttCesA8's NTD (see below) (Fig. 1, A and B).

We focused our structural analysis on the trimeric PttCesA8 particles. Classification in two and three dimensions, ab initio model building, and nonuniform and local refinement in cryoSPARC (*18*) with applied threefold symmetry generated an excellent density map allowing model building of the trimer excluding its N-terminal stalk (residues 1 to 156) (Fig. 1C and fig. S2). The GT domain map significantly exceeds the overall estimated resolution of 3.5 Å as essentially all side chains are resolved (fig. S3A). In the TM segment, aromatic residues as well as comparisons with the bacterial BcsA structure (fig. S4) (*19*) guided model building. Unstructured regions include residues 548 to 609 of the CSR domain, 847 to 868 as part of the gating loop, and 959 to 978 at the extreme C terminus.

PttCesA8's large cytosolic domain is flanked N- and C-terminally by two and five TM helices, respectively (Fig. 1D). TM helices 1 to 6 cluster around three cytosolic interface helices (IF1 to IF3), similar to BcsA from *Rhodobacter sphaeroides* (*19*) (Fig. 1D and fig. S4B). TM topology prediction algorithms frequently predict IF3 of the bacterial and plant enzymes as a proper TM helix, thereby giving rise to eight predicted TM helices in plant CesAs with inverted topologies of TM helices 5 to 7 (fig. S4A). Yet IF3 sits on top of IF1 and IF2 on the cytosolic membrane side and rests against TM helices 3, 5, and 6 in the bilayer, which stabilizes the helix at the cytosolic water-lipid interface (fig. S4, A and B), as observed in the bacterial enzyme (*19*). The C-terminal TM helix 7 (residues 938 to 958) packs into a groove formed by TM helices 4 and 6, as well as IF3 of a neighboring subunit, thereby completing this subunit's TM architecture (Fig. 1C). The extracellular C-terminal residues 959 to 978 are poorly ordered and not included in the model. Most CesAs contain a Cys residue at or close to their C termini (Cys^{978} in PttCesA8), which could be involved in disulfide bond formation or posttranslationally modified, as recently suggested (*20*). The channel generated by PttCesA8's TM helices accommodates a nascent cellulose chain at least five sugar units long (Fig. 1D and fig. S3B). Its nonreducing terminal sugar stacks against Trp^{718} of the conserved QxxRW motif near the active site (see below and fig. S4, C and D).

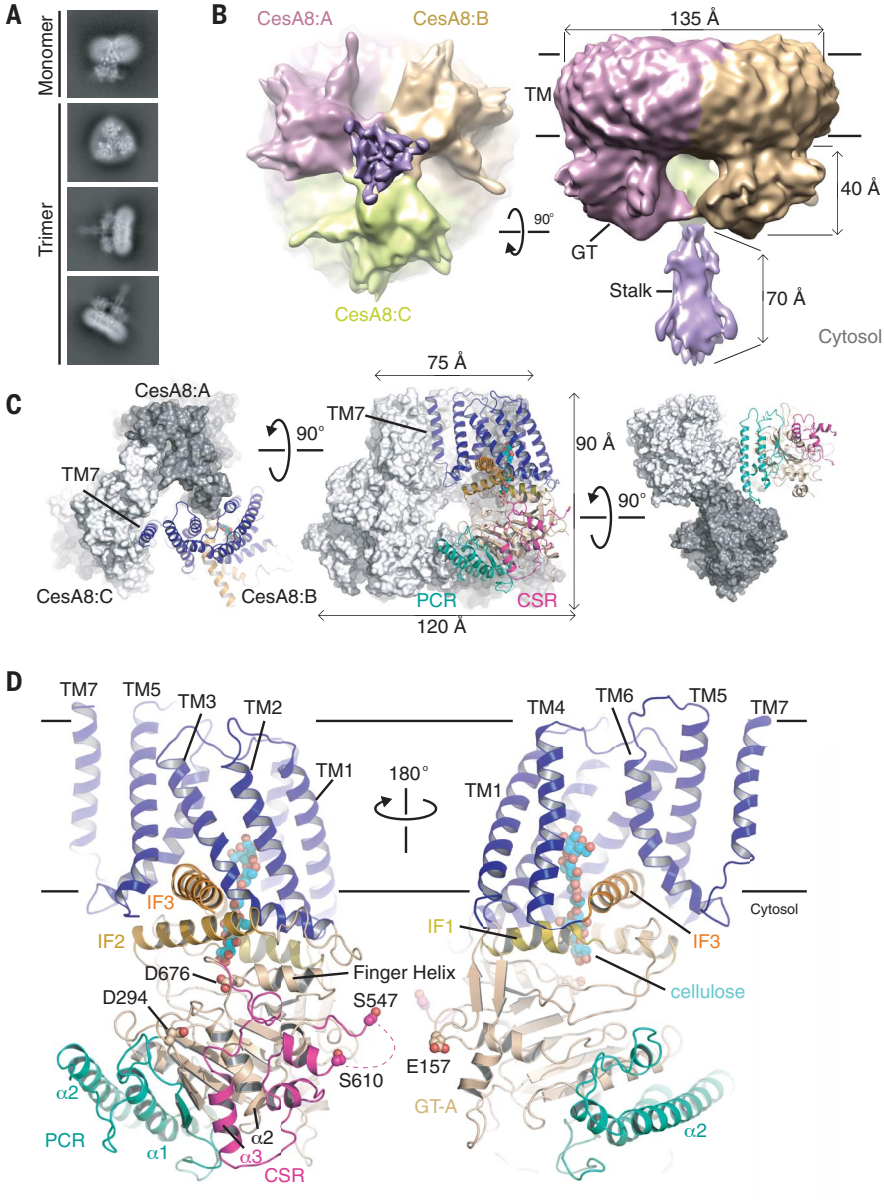


Fig. 1. Structure of PttCesA8. (**A**) Representative two-dimensional class averages of monomeric and trimeric PttCesA8 particles. (**B**) Low-resolution reconstruction of a full-length PttCesA8 trimer shown at a map contour level of 1.8σ. (**C**) Trimeric assembly of the PttCesA8 core. Two copies are shown as surfaces colored light and dark gray, respectively. One subunit is shown as a cartoon colored blue, wheat, blue-green, and magenta for the TM, GT, PCR, and CSR domains, respectively. Amphipathic interface helices are colored in shades of orange. (**D**) Detailed organization of a PttCesA8 monomer. Residues 548 to 609 of the CSR are disordered. Cellulose is shown as cyan and red spheres. D, Asp; E, Glu; S, Ser.

The intracellular catalytic and plant-specific domains

PttCesA8's intracellular catalytic domain adopts a classical GT-A fold consisting of a seven-stranded β sheet surrounded by six α helices that create an active site cleft for substrate and acceptor binding (Fig. 1D and fig. S4C). The plant-specific regions, however, protrude from this core, in agreement with earlier modeling studies (*21*). Whereas the PCR belongs to the region of highest map quality, the CSR is partially disordered, with a central region of 62 residues not included in the model.

The PCR (residues 301 to 418) is inserted into the GT-A fold after Asp^{294} of the conserved substrate-binding "DDG" motif (*19*, *22*) (Fig. 1D and fig. S5). It contains two antiparallel α helices connected by a ~20 residue long linker. PttCesA8's PCR is structurally identical to the corresponding region of rice CesA8, previously crystallized as an isolated domain (*23*), with which it shares 81% sequence identity. Past the last α helix, the PCR is connected to the GT-A core by means of a loop and a short β strand that extends GT-A's central β sheet (Fig. 1D).

The CSR (residues 542 to 665) is inserted into a peripheral loop of the GT-A core, forming helical and disordered regions (fig. S5). The helical subdomain, residues 618 to 665, contains three short helices, of which the C-terminal

helix (residues 640 to 653) packs into a groove formed by the C-terminal β strand of the PCR and α helix 2 of the GT-A core (Fig. 1D). The CSR harbors four cysteine residues in its disordered region close to its N-terminal connection with the GT-A core. Acylation of these residues has recently been implicated in CesA trafficking to the plasma membrane (*20*). The last resolved residue at the beginning of the CSR, Ser^{547}, is only ~20 Å away from the putative cytosolic water-lipid interface, consistent with membrane-association of the acylated CSR under physiological conditions.

The PCR is a trimerization domain

In the trimer, the PttCesA8 protomers are related by a threefold symmetry axis that runs perpendicular to the plane of the membrane. The complex extends ~90 Å along the membrane normal and has a broad intracellular layer ~120 Å in diameter, followed by a narrower TM segment ~75 Å wide (Fig. 1C). Excluding the N-terminal stalk, intermolecular contacts between the subunits are limited to interactions of the C-terminal TM-7 with a neighboring subunit as well as the trimerization of the cytosolic PCRs (Fig. 2A).

PttCesA8 trimerization critically depends on the PCR. In particular, specific intermolecular interactions of the PCR's second α helix (PCR-H2) together with the preceding loop stabilize the trimer interface (Fig. 2A and fig. S5). First, the loop region of subunit A (residues 339 to 348) interacts with a hydrophobic C-terminal segment of PCR-H2 in subunit B (residues 359 to 363). Second, the loop of subunit B contacts the PCR-H2 hydrophobic segment in subunit C, and, third, the loop of subunit C interacts with the hydrophobic PCR-H2 segment of subunit A, thereby forming a triangular PCR interface (Fig. 2B). A similar trimeric PCR arrangement is observed in crystals of the isolated rice CesA8 PCR, in which three PCR copies are related by crystallographic symmetry (*23*) (fig. S6A).

Although the PCR interactions are primarily based on hydrophobic contacts and shape complementarity, the domains create a profound electropositive trimer interface (Fig. 2C). From each PCR-H2, Arg^{352} and Arg^{356} point directly toward the threefold symmetry axis and Arg^{351} and Lys^{355} are similarly positioned on the opposite side of PCR-H2. Ser^{333}, right above the Lys^{355}-Arg^{356} pair, has been predicted to be phosphorylated in *Arabidopsis thaliana* (At) CesA4, which could complement the electropositive environment (Fig. 2C) (*24*). We observe unidentified cone-shaped density ~10 Å long and 6 Å wide on the membrane-proximal side and small spherical density on the membrane-distal side of the PCR trimer (Fig. 2B). The apparent pseudo–threefold symmetric shape of the density is maintained even after refining the data without imposed symmetry, suggesting that it could be a small molecule bound in three different orientations (discussed below).

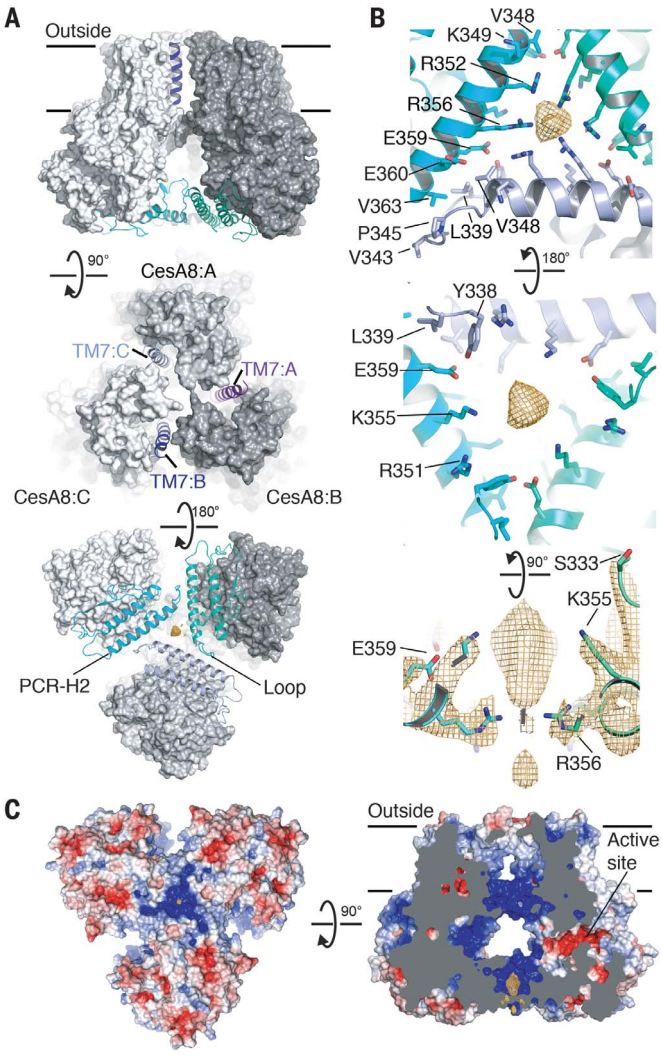


Fig. 2. Trimerization of PttCesA8. (**A**) Surface representation of the CesA8 trimer with all C-terminal TM helices (TM7) and PCR domains shown as cartoons. (**B**) Detailed interactions of the PCR trimer interface. Interface residues are shown as sticks and labeled. Unidentified density is shown as an orange mesh [B-factor sharpened (−93) and contoured at 8 σ]. A larger section of the density map is shown in the bottom panel. K, Lys; L, Leu; P, Pro; R, Arg; V, Val. (**C**) Surface electrostatic potential of the PttCesA8 trimer calculated with the APBS plugin in PyMol and colored red to blue (−5 to +5 kT, respectively). The unidentified density is shown as in (B).

The CesA8 trimer forms three cellulose channels

TM helices 1 to 6, IF3, and TM-7 of a neighboring subunit form a continuous channel across the membrane that accommodates the nascent cellulose chain (Fig. 3, A and B). The polymer is best resolved extending from the active site, yet disconnected density near the extracellular channel exit indicates the presence of a longer, yet flexible, chain (Fig. 3A and fig. S3B). The polymer must have been synthesized in the expression host and copurified with the enzyme, as also observed in bacterial BcsA (*19*).

The TM channel begins right above the active site, with Trp^{718} of IF2, Phe^{513} of IF1, and Val^{529} and Gln^{494} forming the portal (Fig. 3C). The channel is lined with aromatic and hydrophilic residues that likely facilitate glucan translocation. All resolved glucosyl units are coordinated by aromatic residues through CH–π-stacking interactions as well as contacts to hydrophilic residues through their equatorial hydroxyl groups (Fig. 3C and fig. S5).

The extracellular channel exit is formed by the TM-3 and TM-4 loop and TM-6 of one protomer as well as TM-7 from the neighboring subunit (Figs. 1B and 3B). The TM channel is curved toward the trimer's threefold symmetry axis and forms a lateral exit in the extracellular membrane leaflet toward the trimer center (Fig. 3B). Therefore, cellulose chains simultaneously synthesized by a PttCesA8 trimer are extracellularly released within ~30 Å from each other, which may facilitate their alignment into a trimeric cellulose protofibril (Fig. 3D), perhaps represented by the fibrillar reaction product produced by detergent-solubilized PttCesA8 (fig. S1D).

Architecture of the full-length PttCesA8 trimer

Two-dimensional classification and our low-resolution shape reconstruction of the PttCesA8 trimer revealed the presence of a stalk protruding from the cytosolic domains (Fig. 1, A and B). The stalk is ~70 Å long and roughly

aligns with the threefold symmetry axis. It represents CesA's NTD, which, in land plants, contains a zinc-coordinating RING-like domain at the extreme N terminus, followed by a hypervariable region 90 to 150 amino acids long (*10*), which is significantly shorter in CesA8 enzymes compared with other isoforms.

Local refinement of the stalk region (fig. S2) improved the map resolution to ~5 Å, revealing a central helical coiled coil of three α helices together with an extended region at its tip (Fig. 4A). Additional short, likely helical, segments are discernable near the interface with the electropositive PCR trimer where both domains interact. The conserved PCR residues Asp^{357} and Lys^{349} are suitably positioned to mediate these interactions (Fig. 4A).

The function of the NTD is currently unknown, yet dimerization and zinc ion coordination of the isolated N terminus from cotton CesA1 have been suggested (*13*). PttCesA8's isolated NTD (residues 1 to 160) also purifies in a monomer-dimer equilibrium, but we can also detect a small fraction of a trimeric species by gel filtration chromatography and western blotting, along with additional species of a higher molecular weight (fig. S6, B and C). These results suggest a propensity of the NTD to self-associate, which is consistent with the trimeric form observed in the EM structure.

The NTD is connected to the well-ordered GT-A core through residues 157 to 165 that form a two-stranded antiparallel β sheet with the TM-2–GT-A connection (Fig. 1D). Considering that the NTD's RING-like domain is formed by the most N-terminal ~50 residues and is followed by a helical region ~30 residues long, the unresolved linker would be ~70 to 90 residues long (depending on the isoform) (Fig. 4A).

Cellulose synthase interactive (CSI) proteins have been shown to interact with CesAs involved in primary cell wall formation as well as microtubules, thereby likely facilitating the directional deposition of cellulose fibrils on the cell surface (*25*–*27*). We observe binding of full-length PttCesA8 as well as its soluble NTD to CSI-1–coated beads (fig. S6D), suggesting the N-terminal stalk could mediate protein-protein interactions implicated in CSC microtubule tethering.

Implications of CesA mutations and phosphorylation

Several CesA mutations have been reported to either alter the overall amount of cellulose produced or its crystallinity (i.e., CMF organization) or to confer resistances to herbicides. Localizing these substitutions in the PttCesA8 trimer structure provides insights into their mode of action. For example, mutations conferring resistance to the herbicides quinoxyphen and isoxaben in AtCesA1, -3, and -6 [CesA1: A903V (Ala^{903}→Val) for quinoxyphen and CesA3:G998D (Gly^{998}→Asp) and CesA6:R1064W (Arg^{1064}→Trp) for isoxaben resistance] map to PttCesA8's Ile^{800}, Gly^{909}, and Lys^{956}, respectively, of which Gly^{909} and Lys^{956} frame the extracellular channel exit (*28*, *29*) (fig. S7A). Ile^{800} is at the center of TM-4 and packs against TM-7 from the neighboring subunit. An additional isoxaben resistance mutation, T942I (Thr^{942}→Ile) in AtCesA3, localizes to the conserved FxVTxK motif of the gating loop,

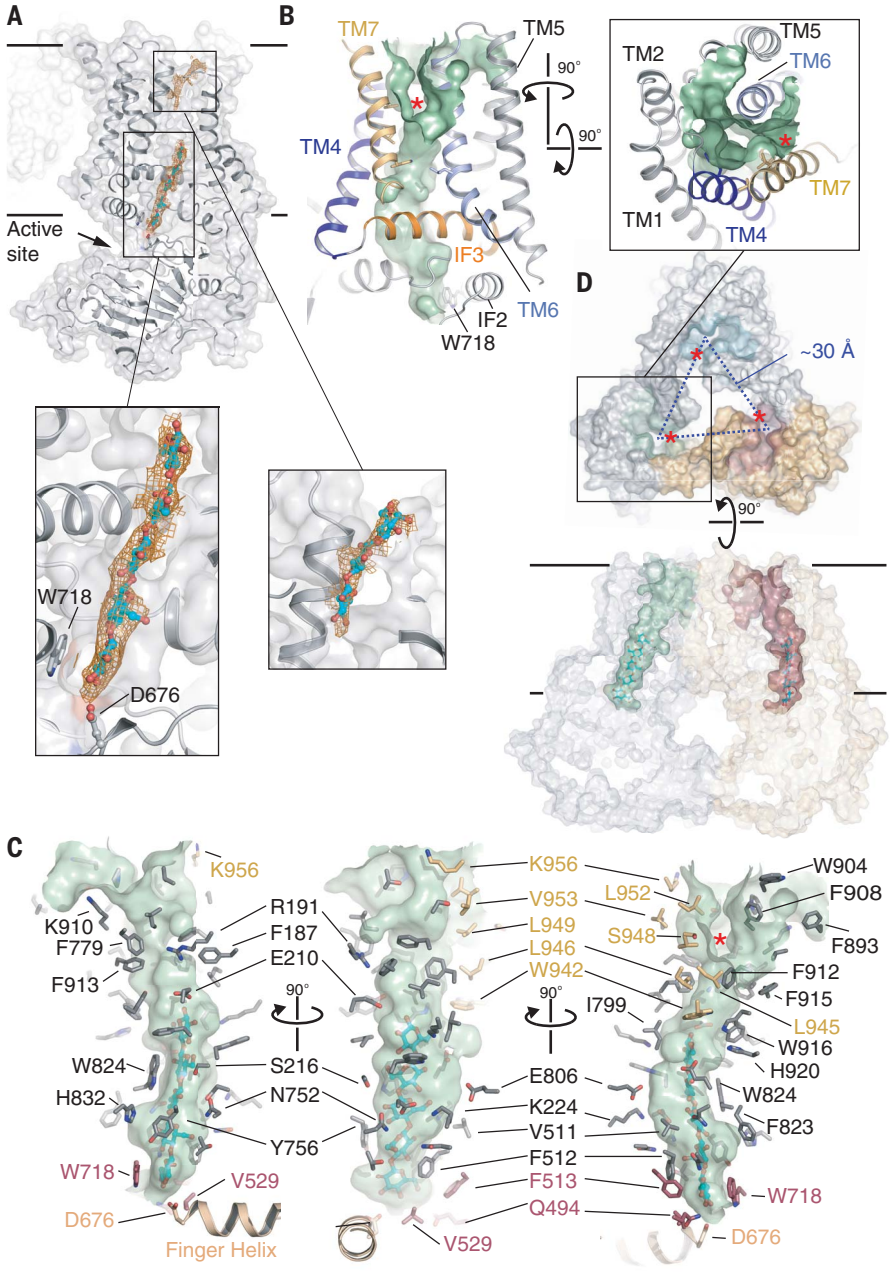


Fig. 3. PttCesA8's channel accommodates a nascent cellulose polymer. (**A**) Each CesA8 subunit contains a glucan chain at least 5 glucose units long. A PttCesA8 subunit is shown as a semitransparent surface and cartoon. Density representing cellulose is shown as a brown mesh. The density near the extracellular channel exit is significantly weaker (shown at 4.5σ contour level) compared with the one near the active site (contoured at 5.5σ) and may represent three glucosyl units (shown as cyan and red ball-and-sticks). (**B**) PttCesA8 forms a continuous TM channel. Channel dimensions were calculated in Hollow (*38*) with a probe radius of 1.8 Å after removing cellulose. (**C**) Cellulose coordination by channel-lining residues from two PttCesA8 subunits. Residues contributed by TM7 of another subunit and those lining the cytosolic channel entrance are labeled in wheat and magenta, respectively. Red asterisk: lateral channel exit. F, Phe; H, His; I, Ile; N, Asn; Q, Gln; W, Trp; Y, Tyr. (**D**) Path of the TM channels in the PttCesA8 trimer. One subunit is colored pale orange; the others are shown in gray.

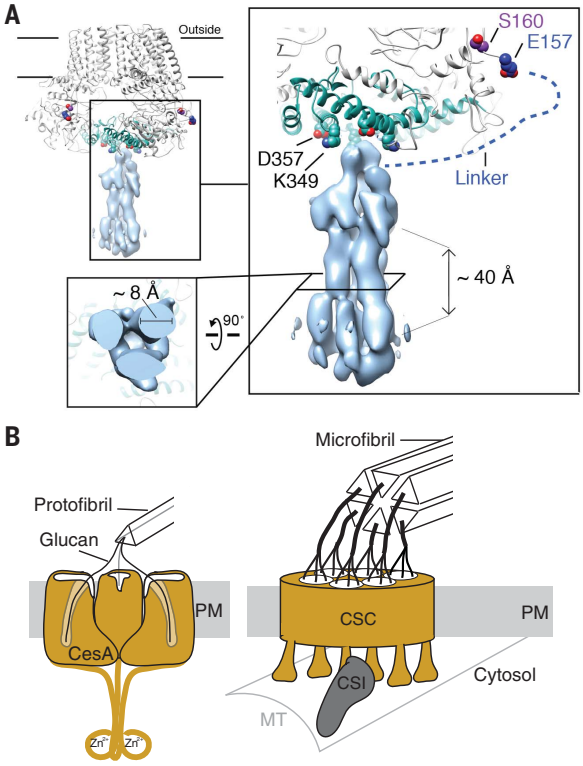


Fig. 4. Structural organization of PttCesA8's stalk and proposed rosette assembly. (**A**) CesA8 core structure aligned with the locally refined stalk density [blue surface, contoured at 1.2σ after B-factor sharpening (−200)]. The central stalk region contains three helical segments. (**B**) Model of cellulose microfibril formation in which trimeric CesA assemblies produce protofibrils that align into a CMF in the context of a rosette. Microtubule (MT) and CesA interacting proteins may facilitate cellular localization and movement of the CSCs in the plasma membrane (PM).

a region that is disordered in our substrate-free structure, similar to observations with bacterial BcsA (*28*, *30*) (figs. S4A and S7A).

Further, several phosphorylation sites have been identified in AtCesAs, potentially affecting CesA activity and degradation (*24*, *31*). With the exception of Ser^{160} near the connection of the N-terminal linker with TM-1 (Fig. 4A) and Ser^{333} at the PCR trimer interface discussed above (Fig. 2B), most sites localize to poorly ordered NTD and CSR regions, where they could affect interactions with other binding partners (fig. S7B).

Discussion

Most plants express several different CesA isoforms that range in sequence identity from ~55 to >68% and are not functionally equivalent. Among the secondary cell wall CesAs (CesA4, -7, and -8), CesA8 is more highly expressed in aspen tension wood compared with isoforms 4 and 7, suggesting that CesA8 may play a particular role in secondary cell wall formation (*32*). Yet the isoforms interact, perhaps by forming heterotrimeric complexes similar to the PttCesA8 homotrimer (*33*). PttCesA8 differs from PttCesA4 and -7 mostly in a shortened hypervariable region of the NTD that forms part of the stalk as well as in the CSR (Fig. 4A and fig. S8). Mapping all conserved residues between the three isoforms onto our PttCesA8 trimer structure reveals that the complex interfaces formed by the PCR and TM region are nearly identical (fig. S8). This may explain why recombinantly expressed PttCesA8 assembles into homotrimers, yet, in vivo, coexpression of different CesA isoforms may favor hetero-oligomer formation.

Our cryo-EM structure suggests that CesA catalytic activity requires its assembly into at least trimeric species. The TM regions of CesA subunits may assemble into trimers upon insertion in the ER membrane, and the tight interaction of the intracellular PCR domains could require posttranslational modifications or binding of a ligand. The electropositive character of the PCR trimer interface (Fig. 2C) almost certainly requires a stabilizing counter charge, which, in turn, could control PCR assembly and, potentially, activity of the trimeric complex. The trimeric structure of the isolated rice PCR domain is stabilized by a phosphate ion (fig. S6A) (*23*). We also observe additional density at this interface significantly larger than a phosphate ion that readily accommodates a dinucleotide, which could provide the necessary counter charge (Fig. 2B and fig. S9).

Cellulose microfibrils have been observed to align with cortical microtubules, and tracks of CSCs during cellulose biosynthesis also align with this network (*34*, *35*). These connections are established by CSI proteins that tether CesAs to the microtubule network (*34*). We reveal that PttCesA8's NTD binds CSI-1 in vitro, and its protrusion into the cytosol would be ideal to also mediate these interactions in vivo. Because the NTD's RING-like domain is shared by all CesA isoforms, an interaction of CSI-1 with this domain seems most likely. However, because CSI proteins have been shown to interact with primary cell wall CesAs in yeast two-hybrid screens using the GT domains (including CSR and PCR) as bait, we cannot exclude multiple interactions with the NTD and GT domain.

The membrane-embedded region of the PttCesA8 trimer superimposes neatly with the rosette's repeat unit structure (*16*) (fig. S10A). Therefore, our data support earlier proposals of rosettes consisting of hexamers of CesA trimers (*16*, *36*). We propose that trimeric CesA complexes produce cellulose protofibrils containing three glucan chains that, in the context of a rosette, assemble into microfibrils containing 18 cellulose chains (Fig. 4B and fig. S10). Recent modeling studies support this hypothesis (*37*).

REFERENCES AND NOTES

1. D. J. Nevins, *Am. J. Clin. Nutr.* **61**, 915S–921S (1995).
2. K. Keegstra, *Plant Physiol.* **154**, 483–486 (2010).
3. J. C. Duchesne, D. W. Larson, *BioScience* **39**, 238–241 (1989).
4. L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.* **77**, 521–555 (2008).
5. B. L. Cantarel *et al.*, *Nucleic Acids Res.* **37** (Suppl. 1), D233–D238 (2009).
6. Y. Bi, C. Hubbard, P. Purushotham, J. Zimmer, *Curr. Opin. Struct. Biol.* **34**, 78–86 (2015).
7. J. L. Morgan *et al.*, *Nature* **531**, 329–334 (2016).
8. M. Iguchi, S. Yamanaka, A. Budhiono, *J. Mater. Sci.* **35**, 261–270 (2000).
9. O. A. McCrate, X. Zhou, C. Reichhardt, L. Cegelski, *J. Mol. Biol.* **425**, 4286–4294 (2013).
10. H. E. McFarlane, A. Döring, S. Persson, *Annu. Rev. Plant Biol.* **65**, 69–94 (2014).
11. S. H. Cho *et al.*, *Plant Physiol.* **175**, 146–156 (2017).
12. P. Purushotham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11360–11365 (2016).
13. I. Kurek, Y. Kawagoe, D. Jacob-Wilk, M. Doblin, D. Delmer, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11109–11114 (2002).
14. C. Somerville, *Annu. Rev. Cell Dev. Biol.* **22**, 53–78 (2006).
15. C. E. Vergara, N. C. Carpita, *Plant Mol. Biol.* **47**, 145–160 (2001).
16. B. T. Nixon *et al.*, *Sci. Rep.* **6**, 28696 (2016).
17. M. C. Jarvis, *Plant Physiol.* **163**, 1485–1486 (2013).
18. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, *Nat. Methods* **14**, 290–296 (2017).
19. J. L. Morgan, J. Strumillo, J. Zimmer, *Nature* **493**, 181–186 (2013).
20. M. Kumar *et al.*, *Science* **353**, 166–169 (2016).
21. L. Sethaphong *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7512–7517 (2013).
22. I. M. Saxena, R. M. Brown Jr., T. Dandekar, *Phytochemistry* **57**, 1135–1148 (2001).
23. P. S. Rushton *et al.*, *Plant Physiol.* **173**, 482–494 (2017).
24. T. L. Speicher, P. Z. Li, I. S. Wallace, *Plants* **7**, 52 (2018).
25. S. Li, L. Lei, C. R. Somerville, Y. Gu, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 185–190 (2012).
26. L. Lei, S. Li, J. Du, L. Bashline, Y. Gu, *Plant Cell* **25**, 4912–4923 (2013).
27. Y. Gu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12866–12871 (2010).
28. W. R. Scheible, R. Eshed, T. Richmond, D. Delmer, C. Somerville, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10079–10084 (2001).
29. D. M. Harris *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4098–4103 (2012).
30. J. L. W. Morgan, J. T. McNamara, J. Zimmer, *Nat. Struct. Mol. Biol.* **21**, 489–496 (2014).
31. S. Chen *et al.*, *Plant Physiol.* **171**, 242–250 (2016).
32. X. Zhang *et al.*, *Plant Physiol.* **177**, 1096–1107 (2018).
33. J. Timmers *et al.*, *FEBS Lett.* **583**, 978–982 (2009).

34. M. Bringmann *et al*., *Trends Plant Sci.* **17**, 666–674 (2012).
35. M. Woodley, A. Mulvihill, M. Fujita, G. O. Wasteneys, *Plants* **7**, 53 (2018).
36. V. G. Vandavasi *et al*., *Plant Physiol.* **170**, 123–135 (2016).
37. J. D. Kubicki *et al*., *Sci. Rep.* **8**, 13983 (2018).
38. B. K. Ho, F. Gruswitz, *BMC Struct. Biol.* **8**, 49 (2008).

ACKNOWLEDGMENTS

We thank S. H. Cho, T. Nixon, and C. Bator for support using the Huck Institute cryo-EM facility at the Pennsylvania State University. We thank G. Skiniotis for discussions and are grateful for support from K. Dryden and M. Purdy from the MEMC at the University of Virginia. We are grateful to D. Cosgrove and B. Chittoor for comments on the manuscript. **Funding:** This work was supported as part of the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences (award DESC0001090). MEMC is funded by National Institutes of Health (NIH) Recovery grant (1G20RR31199). The MEMC Titan Krios and Gatan K3/GIF detector were funded through the NIH (grants SIG S10-RR025067 and U24-GM116790, respectively). **Author contributions:** P.P. established all expression, purification, and activity assay protocols and prepared materials for cryo grid preparation. R.H. prepared and screened cryo grids and collected all data sets. J.Z. processed the data and built and refined the model. All authors contributed to preparing the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** EM maps have been deposited at the Electron Microscopy Data Bank (EMD-21820), and model coordinates have been deposited at the Protein Data Bank under ID 6WLB. All other data are provided in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1089/suppl/DC1
Materials and Methods
Figs. S1 to S10
Table S1
References (*39*–*49*)
MDAR Reproducibility Checklist

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

13 February 2020; accepted 26 June 2020
Published online 9 July 2020
10.1126/science.abb2978

REPORT

BACTERIAL METABOLISM

A nitrogenase-like enzyme system catalyzes methionine, ethylene, and methane biogenesis

Justin A. North[1], Adrienne B. Narrowe[2], Weili Xiong[3]*, Kathryn M. Byerly[1], Guanqi Zhao[1], Sarah J. Young[1], Srividya Murali[1], John A. Wildenthal[1]†, William R. Cannon[4,5], Kelly C. Wrighton[2], Robert L. Hettich[3], F. Robert Tabita[1]‡

Bacterial production of gaseous hydrocarbons such as ethylene and methane affects soil environments and atmospheric climate. We demonstrate that biogenic methane and ethylene from terrestrial and freshwater bacteria are directly produced by a previously unknown methionine biosynthesis pathway. This pathway, present in numerous species, uses a nitrogenase-like reductase that is distinct from known nitrogenases and nitrogenase-like reductases and specifically functions in C–S bond breakage to reduce ubiquitous and appreciable volatile organic sulfur compounds such as dimethyl sulfide and (2-methylthio)ethanol. Liberated methanethiol serves as the immediate precursor to methionine, while ethylene or methane is released into the environment. Anaerobic ethylene production by this pathway apparently explains the long-standing observation of ethylene accumulation in oxygen-depleted soils. Methane production reveals an additional bacterial pathway distinct from archaeal methanogenesis.

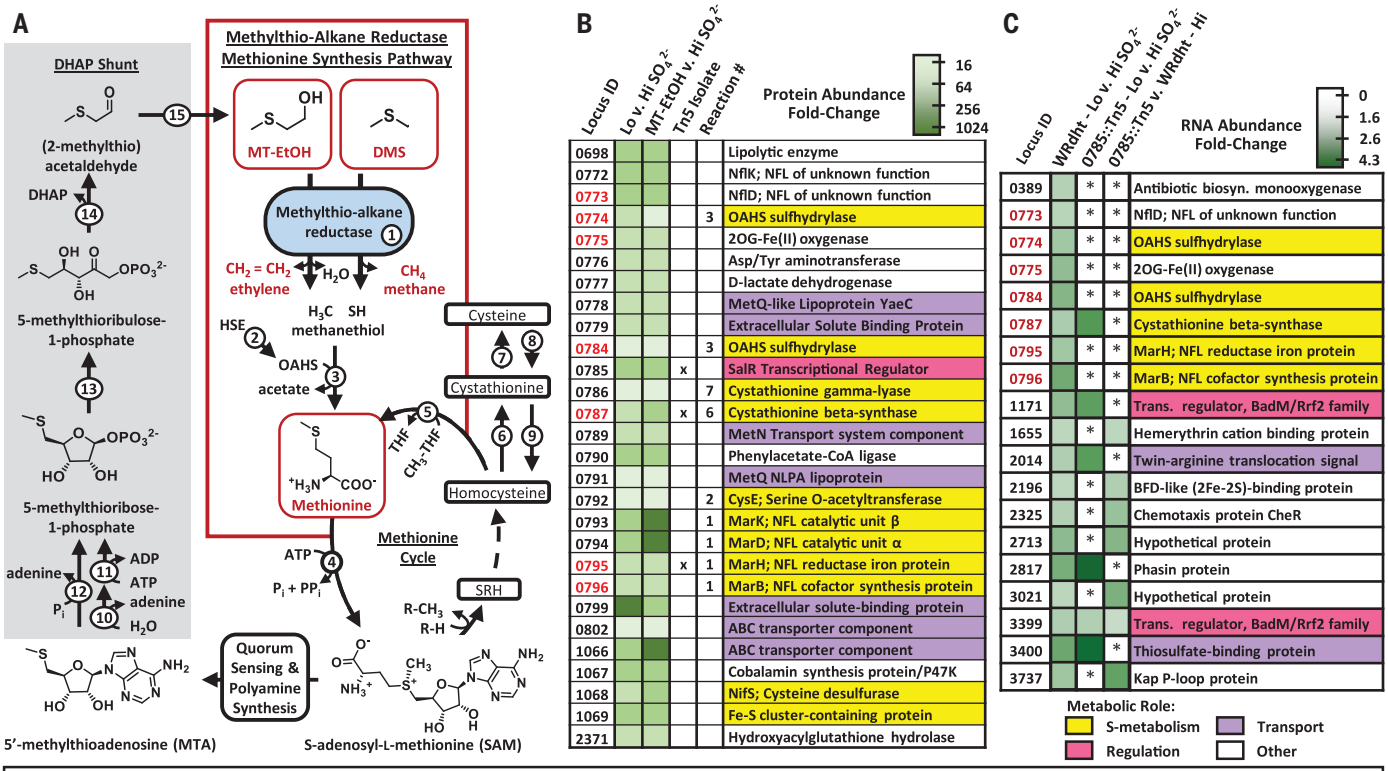
Nitrogenases are an ancient group of enzymes, existing ~3.2 billion years ago, before the evolution of oxygenic photosynthesis and subsequent widespread oxygenation (*1*, *2*). Their essential function is reduction of dinitrogen gas into ammonia, which contributes over half of the annual global nitrogen fixation required for the synthesis of nucleic and amino acids by all life on Earth (*3*). Ancestors to nitrogenase in anaerobic prokaryotes also gave rise to distinct nitrogenase-like reductases for bacterial photosynthesis and archaeal methanogenesis cofactor metabolism (*4*–*7*). These include the dark-operative protochlorophyllide oxidoreductase (DPOR) and chlorophyllide a oxidoreductase (COR) of bacteriochlorophyll biosynthesis, as well as Ni^{2+}-sirohydrochlorin a,c-diamide reductive cyclase for biosynthesis of the archaeal methyl coenzyme-M reductase cofactor F430 (*4*–*7*). However, the evolutionary history of nitrogen fixation revealed overlooked nitrogen fixation–like (NFL) sequences in the genomes of anaerobic bacteria with entirely unknown function. Some of these sequences were associated with sulfur metabolism and transport genes (*8*, *9*), suggesting that the nitrogenase family evolved to tackle elemental processes beyond nitrogen and carbon assimilation.

We previously observed production of ethylene gas (>1 μmol h^{-1} g^{-1} dry cell weight) by photosynthetic Alphaproteobacteria such as *Rhodospirillum rubrum* and *Rhodopseudomonas palustris* when growing anaerobically under the low sulfate concentrations (<200 μM) commonly encountered in their freshwater and soil habitats (fig. S1) (*10*). We traced the precursor of ethylene to (2-methylthio)ethanol (MT-EtOH). This volatile organic sulfur compound (VOSC) was produced from by-products of *S*-adenosyl-L-methionine (SAM) utilization to regenerate methionine [Fig. 1A, dihydroxyacetone phosphate (DHAP) shunt] (*10*). SAM is a key cellular cofactor synthesized directly from methionine and is required by all organisms for diverse processes including DNA, RNA, and protein methylations; polyamine and neurotransmitter synthesis; quorum sensing; and 5′-deoxyadenosyl radical generation by radical SAM enzymes (*11*). However, the enzymes responsible for the liberation of sulfur from MT-EtOH for methionine regeneration and concomitant ethylene formation were unresolved (*10*). Therefore, we grew *R. rubrum* under conditions for ethylene induction (50 μM limiting sulfate or 1 mM MT-EtOH as the sole sulfur source) and ethylene repression (1 mM sulfate) (fig. S1) (*10*). Differential abundance analysis identified multiple proteins that increased by more than 20-fold in proteomes from induced cells compared with those from repressed cells (Fig. 1B). Among these were enzymes involved in cysteine and methionine metabolism: homoserine/serine O-acetyltransferase, O-acetylhomoserine sulfhydrylase, cystathionine beta-synthase, and cystathionine gamma-lyase (Fig. 1, reactions 2, 3, 6, and 7, respectively).

Unexpectedly, several proteins previously identified as NFL sequences of unknown function (*8*, *9*) showed some of the highest increases in abundance under ethylene-inducing conditions (Fig. 1B, Rru_A0772-Rru_A0773 and Rru_A0793–Rru_A0796; see fig. S2 and data S1 for gene organization and sequences). In addition, there was also a large increase in abundance of two proteins likely involved in iron-sulfur cluster metabolism, NifS cysteine desulfurase and a putative [4Fe-4S] scaffold protein (Fig. 1B, Rru_A1068-Rru_A1069). Rru_A1068-Rru_A1069 appear analogous to the *Azotobacter vinelandii* NifUS system for synthesis of nitrogenase-destined iron-sulfur clusters from cysteine (*12*). However, the precise iron-sulfur cluster assembly pathway in *R. rubrum* is unknown. The involvement of the nitrogenase-like system in ethylene production was further bolstered by the *R. rubrum* transposon mutant strain WRdht-66B3, having an inactivated gene encoding a putative nitrogenase reductase–like iron protein (Fig. 1B, Rru_A0795). This and other mutants identified in a random mutagenesis screen were unable to grow anaerobically with MT-EtOH as the sole sulfur source but could still grow using sulfate, indicating a defect in the ethylene-producing pathway (fig. S3). Consistent with the transposon Tn5 mutagenesis results, specific deletion of NFL gene cluster Rru_A0793–Rru_A0796 rendered *R. rubrum* incapable of growth or production of ethylene above basal levels with MT-EtOH as the sole sulfur source (Fig. 2, A and B, and fig. S4C). This result confirmed that the putative nitrogenase-like system encoded by NFL gene cluster Rru_A0793–Rru_A0796 was essential for assimilating sulfur from MT-EtOH to produce ethylene and methionine.

We then tested whether other biologically relevant VOSCs could be used by this putative nitrogenase-like enzyme system (Fig. 2, A and B, and fig. S5A). In addition to MT-EtOH, VOSC utilization with concomitant hydrocarbon production was specific to dimethyl sulfide (DMS), the most abundant environmental VOSC, and ethyl methyl sulfide (EMS) (Fig. 2, A and B). Analogous to MT-EtOH (*10*), use of DMS or EMS resulted in methane or ethane production, respectively, in a 1:1 stoichiometry (Fig. 3, A and B). Specific deletion of the other two NFL genes, Rru_A0772-Rru_A0773, did not affect growth or hydrocarbon production (Fig. 2, A and B, and fig. S4B). Thus, we designate *R. rubrum* genes Rru_A0793–Rru_A0796, previously identified as NFL genes *nflBHDK* of unknown function (*8*, *9*), as methylthio-alkane reductase genes *marBHDK*. This designation is based on corresponding amino acid similarity to *R. rubrum* molybdenum nitrogenase

[1]Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA. [2]Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80523, USA. [3]Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA. [4]Pacific Northwest National Laboratory, Richland, WA 99352, USA. [5]Department of Mathematics, University of California, Riverside, Riverside, CA 92507, USA.
*Present address: Center for Food and Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, MD 20740, USA.
†Present address: Washington University in St. Louis School of Medicine, St. Louis, MO 63110, USA.
‡Corresponding author. Email: tabita.1@osu.edu



Enzyme and Compound Key:

1) methylthio-alkane reductase
2) serine/homoserine O-acetyltransferase (2.3.1.31)
3) O-acetylhomoserine sulfhydrylase (2.5.1.49)
4) S-adenosylmethionine synthetase (2.1.1.13, 2.1.1.14)
5) methionine synthase (2.5.1.6)
6) cystathionine beta-synthase (4.2.1.22)
7) cystathionine gamma-lyase (4.4.1.1)
8) cystathionine gamma-synthase (2.5.1.48)
9) cystathionine beta-lyase (4.4.1.8)
10) MTA nucleosidase (3.2.2.16)
11) 5-methylthioribose kinase (2.7.1.100)
12) MTA phosphorylase (2.4.2.28)
13) 5-methylthioribose-1-phosphate isomerase (5.3.1.33)
14) 5-methylthioribulose-1-phosphate aldolase (4.1.2.n)
15) alcohol dehydrogenase (1.1.1.1)

DHAP) dihydroxyacetone phosphate
HSE) homoserine
OAHS) O-acetyl-L-homoserine
SRH) S-ribosyl-L-homocysteine
R-H) methyl acceptor
THF) tetrahydrofolate
ADP) adenosine diphosphate
ATP) adenosine triphosphate

Fig. 1. Nitrogenase-like proteins linked to VOSC utilization. (**A**) Methylthio-alkane reductase (**1**, blue), the gene product of *marBHDK* (proposed), converts VOSCs to ethylene, methane, and methanethiol for methionine biosynthesis (red box). MT-EtOH is produced by the widespread DHAP shunt (gray shaded box) (*10*, *25*, *26*). (**B**) *R. rubrum* proteins with increased abundance when methylthio-alkane reductase activity is induced ("MT-EtOH" or "Lo" 50 μM sulfate) versus repressed ("Hi" 1 mM sulfate). Isolated Tn5 transposon mutants (fig. S3), which could not utilize MT-EtOH for growth, are marked with an X. (**C**) Changes in gene transcript abundance of *R. rubrum* parent strain (WRdht) and SalR deletion strain (0785::Tn5) under "Hi" and "Lo" sulfate. Asterisk symbol indicates no significant change. $P > 0.25$, two-tailed.

gene products NifB (synthesis of the NifB-cofactor precursor to the nitrogenase catalytic cofactor), NifH (nitrogenase-reductase iron protein), NifD (nitrogenase catalytic subunit α), and NifK (nitrogenase catalytic subunit β) (figs. S6 to S9). NFL genes Rru_A0772-Rru_A0773 remain designated *nflDK* genes of unknown function (*8*, *9*).

When all *R. rubrum* NFL genes were deleted (strain Δ0772:3/Δ0793:6) and specific gene combinations were reintroduced via expression from a plasmid, expression of *marBHDK* was necessary and sufficient to restore growth and hydrocarbon metabolism from VOSCs (Fig. 2, B and C, and fig. S5, B and C). The NFL genes of unknown function, *nflDK*, could not replace *marDK* in complementing for growth. Upon feeding cells expressing *marBH* and *nflDK* with VOSCs, ethylene and ethane production was poorly catalyzed at three- to fourfold above basal levels, and no methane enhancement was observed (Fig. 2, B and C, and fig. S5, B and C). This revealed that *R. rubrum* NflDK could only weakly catalyze methylthio-alkane reduction, indicating a different primary function. Given that *nflDK* is expressed not just in the presence of MT-EtOH but also in response to general sulfate limitation (Fig. 1, B and C), NflDK may catalyze sulfur liberation from alternate, albeit unknown, compounds. Alternately, given gene proximity and amino acid similarity (40%) to MarDK, NflDK may serve as accessory proteins for MarDK assembly analogous to NifEN (*13*), as detailed in the supplementary text section of the supplementary materials. These results demonstrated the requirement of the MarBHDK nitrogenase-like system for the anaerobic assimilation of sulfur from common environmental VOSCs such as DMS and MT-EtOH to support growth and methionine metabolism. Moreover, these observations revealed a previously unknown mechanism for the bacterial production of methane and ethylene.

We metabolically observed the link between VOSC utilization and methionine synthesis via the *marBHDK* gene products by feeding experiments with (2-methyl[^{14}C]thio)ethanol, which enabled us to follow the methylthio group of MT-EtOH. Upon feeding the wild-type strain, MT-EtOH was consumed. Labeled methanethiol ($^{14}CH_3$-SH) and methionine (methyl-^{14}C) were concomitantly produced and observed at low levels (~2% of MT-ETOH concentration) until MT-EtOH was depleted (Fig. 2D). These low levels, as previously observed for methanethiol metabolism from 5′-methylthioadenosine in *R. rubrum* (*12*), are likely due to the flux of methanethiol to methionine and subsequent utilization thereof for protein synthesis and SAM-dependent processes (*11*). This flux is substantiated by ^{14}C incorporation from MT-EtOH into insoluble cell

material (fig. S10). Conversely, in the *marBHDK* deletion strain there was no detectable metabolism of MT-EtOH and thus no methanethiol or methionine production (Fig. 2E and fig. S10). Given that ethylene, ethane, and methane are produced from MT-EtOH, EMS, and DMS, respectively, the observed methanethiol is consistent with C–S single bond reduction and release of the methylthio group by the methylthio-alkane reductase (Figs. 1A, reaction 1, and 2F). Each process is thermodynamically favored for the substrates and products observed (Fig. 2F and fig. S11). The methanethiol, along with O-acetyl-L-homoserine, then serves as a substrate for O-acetylhomoserine sulfhydrylase, which catalyzes the synthesis of methionine (Fig. 1A, reaction 3) (*14*). This defines an anaerobic methylthio-alkane reductase methionine synthesis pathway and establishes the role of a nitrogenase-like enzyme system in sulfur metabolism (Fig. 1A, red box).

Sulfur metabolism is evidently the primary function of these nitrogenase-like methylthio-alkane reductases, as opposed to nitrogen fixation by nitrogenase. *R. rubrum* has molybdenum nitrogenase (NifHDK), which is the default nitrogenase, and iron-only nitrogenase (AnfHDGK), which is synthesized in the absence of molybdenum (*9*). In in vivo activity assays, the *R. rubrum* molybdenum nitrogenase could not perform methylthio-alkane reduction, and methylthio-alkane reductase did not exhibit nitrogenase activity (Fig. 3D; glutamate as nitrogen source and 50 μM sulfate). Indeed, nitrogenase and methylthio-alkane reductase activities were independent and separately regulated, and both systems could be expressed simultaneously (Fig. 3D). *R. rubrum* nitrogenase gene expression (*nifHDK*) is regulated by the transcriptional regulator NifA in response to nitrogen availability (*13*). Methylthio-alkane reductase activity in the presence of 1 mM MT-EtOH or DMS was regulated by sulfate availability, with a median effective concentration (EC_{50}) of ~150 μM sulfate for 50% repression of activity (Fig. 3C). Our random mutagenesis screen identified the specific regulatory gene in the vicinity of *marBHDK* (Rru_A0785; Fig. 1B and figs. S2 and S3). We designate this LysR family regulator as SalR (sulfur salvage regulator) (data S1). Inactivation of *salR* rendered strains incapable of growth or hydrocarbon production using MT-EtOH, DMS, or EMS as the sole sulfur source (Fig. 2, A and B, and fig. S4E; strain 0785::Tn5). Transcriptomics and differential expression analysis of the parent (WRdht) and *salR* deletion strain (0785::Tn5) growing under *marBHDK* inducing and repressing conditions revealed that *marBHDK* and the rest of the methylthio-alkane reductase methionine synthesis pathway are under transcriptional control of SalR (Fig. 1C). Thus, when sufficient sulfur is available (>150 μM), expression appears repressed, but when sulfate becomes limiting, *marBHDK* and O-acetylhomoserine sulfhydrylase gene transcription is specifically up-regulated via SalR to use VOSCs for methionine metabolism (Fig. 1A, reactions 1 and 3).

The nitrogenase superfamily is composed of the bona fide nitrogenase sequences (groups I to III) and nitrogen fixation–like sequences (NFL; groups IV to VI) (Fig. 4) (*9*). Phylogenetic

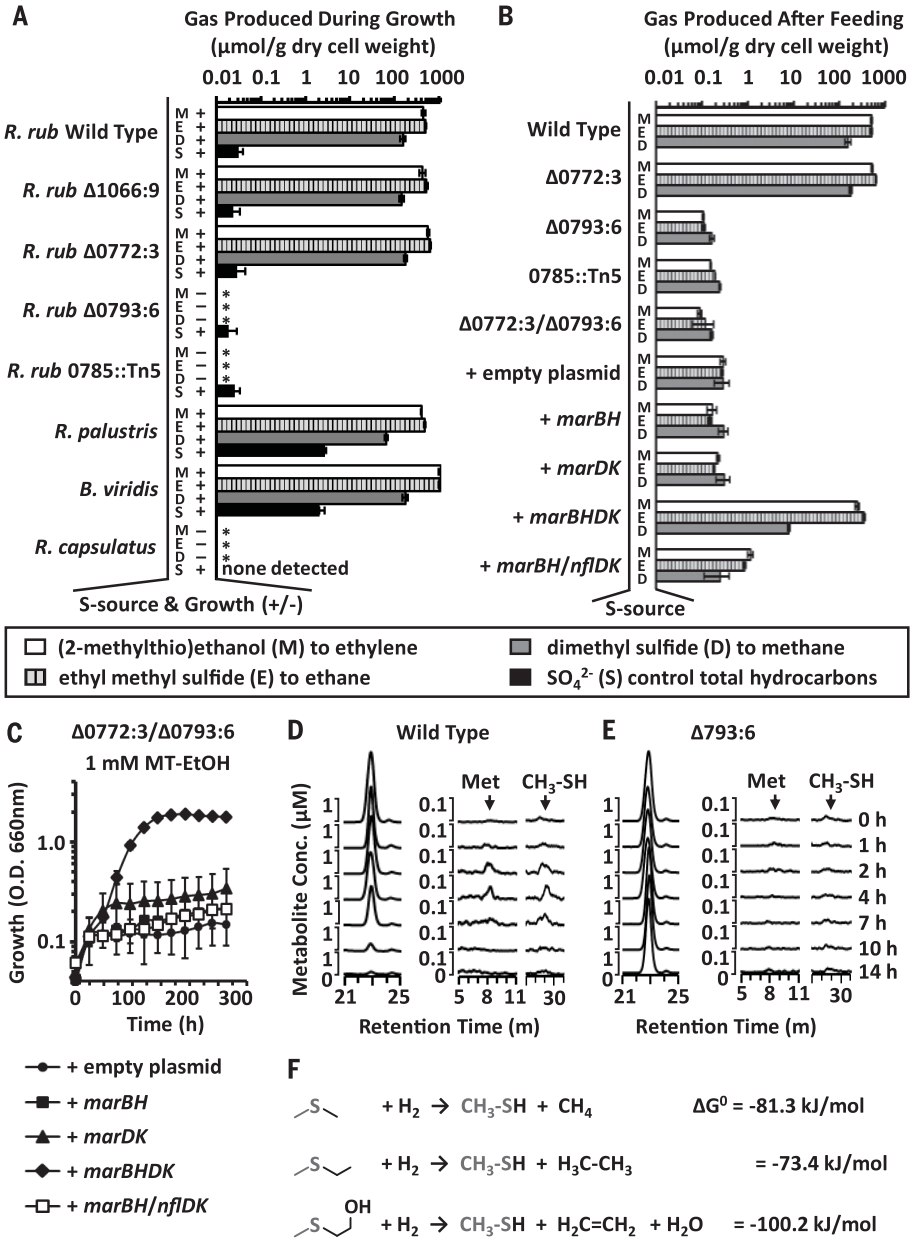
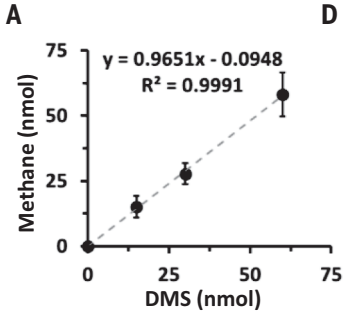
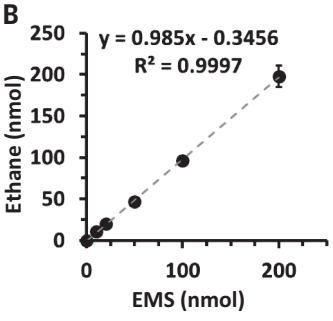


Fig. 2. Genes *marBHDK* are required for anaerobic methionine metabolism from VOSCs. (**A**) Growth and average total hydrocarbon production of strains using sulfate or VOSCs (see figs. S4 and S15). Asterisk symbol, not applicable. (**B**) Total amount of hydrocarbons produced when cells were fed with the indicated VOSC. (**C**) Plasmid-based complementation studies of NFL genes for growth of *R. rubrum* NFL gene deletion strain. O.D. 660 nm, optical density measured at 660 nm. Error bars in (A) to (C) are standard deviation for $N = 3$ independent biological replicates. (**D** and **E**) Identification of methionine [retention time (RT) = 8.5 min] and methanethiol (RT = 28.3 min) upon feeding *R. rubrum* strains with (2-methyl[^{14}C]thio)ethanol (RT = 22.8 min). (**F**) Change in Gibbs free energy under standard conditions for the conversion of VOSCs to methanethiol and the corresponding hydrocarbon. H_2 represents $2H^+$ and $2e^-$ from an electron donor, as detailed in the materials and methods section of the supplementary materials.





D

Strain	Sulfate (μM)	Mar Activity (nmol h^{-1} g^{-1})	N-Source (15 mM)	Nitrogenase Activity (μmol h^{-1} g^{-1})
WT	1000	92	NH_4^+	7
WT	1000	100	Glu	1135
WT	50	478	NH_4^+	5
WT	50	684	Glu	440
ΔΔ	1000	69	NH_4^+	4
ΔΔ	1000	102	Glu	1081
ΔΔ	50	92	NH_4^+	9
ΔΔ	50	88	Glu	520

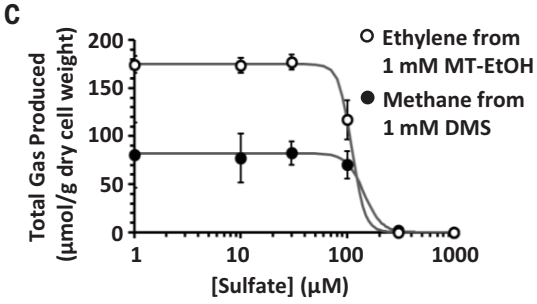


Fig. 3. Methylthio-alkane reductase and nitrogenase are independent. (**A** and **B**) Stoichiometric production of methane and ethane by cells fed with DMS and EMS. R^2, coefficient of determination. (**C**) Competition assays for methylthio-alkane reductase repression in cells grown with 1 mM MT-EtOH or DMS plus the indicated amount of sulfate. Nonlinear fit to the Hill equation gives $EC50_{DMS/sulfate}$ = 140 μM sulfate and $EC50_{MT\text{-}EtOH/sulfate}$ = 110 μM sulfate for 50% activity with DMS and MT-EtOH as substrate, respectively. (**D**) Whole-cell methylthio-alkane reductase (Mar) and molybdenum nitrogenase (NifHDK) activities for wild type (WT) and Δ0772:3/Δ0793:6 deletion (ΔΔ) strains under methylthio-alkane reductase inducing (50 μM sulfate) or repressing (1 mM sulfate) and NifHDK inducing (Glu, glutamate) or repressing (NH_4^+, ammonium) conditions. Standard deviations are [(A) to (C)] the error bars or (D) are <10% for *N* = 3 biological replicates.

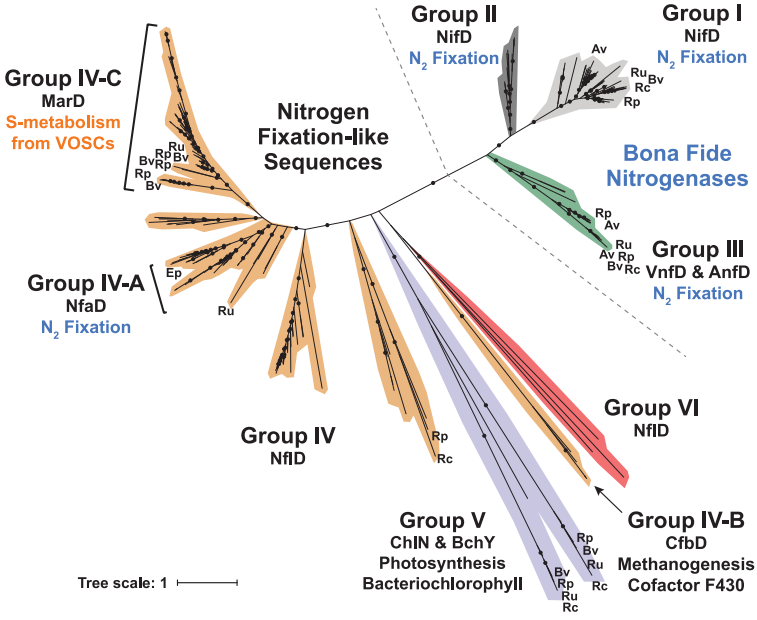


Fig. 4. Methylthio-alkane reductases are phylogenetically distinct. Phylogenetic tree of NifD superfamily homologs. The scale bar represents the number of substitutions per site. Nodes with UFBoot support values ≥ 95% are indicated with black circles. Clade labeling: Group IV-A (NfaD, nitrogen fixation IV-A) (*27*); Group IV-B (CfbD, Ni^{2+}-sirohydrochlorin a,c-diamide reductive cyclase) (*4*, *5*); Group IV-C (MarD, putative methylthio-alkane reductase); Groups IV and VI (NflD, nitrogen fixation–like of unknown function); Group V (ChlN, DPOR; and BchY, COR). Clade labels and colors are per Raymond *et al.* (*9*) and Méheust *et al.* (*28*). Av, *Azotobacter vinelandii*; Bv, *Blastochloris viridis*; Ep, *Endomicrobium proavitum*; Rc, *Rhodobacter capsulatus*; Rp, *Rhodopseudomonas palustris*; Ru, *Rhodospirillum rubrum*.

analysis places methylthio-alkane reductase homologs in their own clade within group IV, which we denote as group IV-C (Fig. 4 and fig. S12). In contrast, the *R. rubrum* NflD protein resides in a separate clade with other NflD sequences of unknown function (Fig. 4), consistent with the poor methylthio-alkane reductase activity exhibited by NflDK (Fig. 2B). Bacteria having MarBHDK sequence homologs of this previously uncharacterized group IV-C clade include members of the Fibrobacter and Bacteroidetes phyla, Rhodospirillales and Rhizobiales within the Proteobacteria phylum, and Selenomonadales and *Clostridium* species within the Firmicutes phylum (fig. S13). To verify our phylogeny results for the Proteobacteria, we tested *R. palustris* and *Blastochloris viridis*, which have group IV-C *marBHDK* homologs. We also tested closely related species *Rhodobacter capsulatus*, which have nitrogenase and *nflBHDK* but no *marBHDK* (Fig. 4 and figs. S12 and S14; Rp, Bv, Rc). Both *R. palustris and B. viridis* were able to grow with MT-EtOH, EMS, or DMS as the sole sulfur source and correspondingly produced ethylene, ethane, or methane (Fig. 2A and fig. S15), demonstrating that methylthio-alkane reductase homologs from these organisms catalyze the same process. Conversely, *R. capsulatus* could not use any of these VOSCs as the sole sulfur source for growth (Fig. 2A and fig. S15), like *R. rubrum* expressing *nflDK* but not *marDK* (Fig. 2, B and C), indicating that group IV NFL proteins of unknown function catalyze different processes than methylthio-alkane reductase.

There are several noteworthy similarities and differences between MarBHDK and other nitrogenase and nitrogenase-like sequences (figs. S6 to S9), as detailed in the supplementary

text. MarDK have all the requisite cysteines for P-cluster coordination found in the bona fide nitrogenases but not in the nitrogenase-like systems for bacteriochlorophyll and F430 biosynthesis (figs. S6 and S7) (*9*). The nitrogenase P-cluster shuttles electrons originating from the NifH [4Fe-4S] cluster to the catalytic FeMo cofactor (*13*). In addition, MarH has the conserved NifH [4Fe-4S] cluster coordination and MgATP hydrolysis domains for transferring electrons to a P-cluster (fig. S8) (*9*). Methylthio-alkane reductases also appear similar to nitrogenases owing to the presence of NifB homologs. MarB has all of the conserved motifs found in nitrogenase NifB sequences, particularly the radical SAM domain for carbide insertion into the NifB-cofactor precursor of the FeMo cofactor (fig. S9) (*12*). However, the conserved NifD histidine responsible for coordinating the FeMo-cofactor molybdenum and homocitrate in nitrogenase is replaced with acidic amino acids in MarD (fig. S6) (*9*, *15*, *16*), which is suggestive of an analogous but distinct catalytic cofactor used by methylthio-alkane reductase. Together, this indicates that methylthio-alkane reductase proceeds via a mechanism, albeit currently unknown, similar to that of nitrogenase (*17*).

Methane release from DMS by the methylthio-alkane reductases is separate and distinct from the other known nonarchaeal methanogenic processes, including photosynthesis-linked methane production by cyanobacteria (*18*), methane release from methylphosphonates by marine bacteria (*19*), and direct reduction of carbon dioxide to methane by iron-only nitrogenase (AnfDHGK) (*20*). In waterlogged soils, strictly anaerobic microbial processes produce ethylene that can accumulate to levels inhibitory to plant root growth, causing crop damage (*21*, *22*). Early attempts at identifying ethylene-producing organisms unexpectedly isolated oxygen-dependent soil bacteria and fungi (*23*, *24*). The organisms and methylthio-alkane reductases we have identified here function anaerobically and could contribute to this soil-ethylene paradox (*10*).

REFERENCES AND NOTES

1. E. E. Stüeken, R. Buick, B. M. Guy, M. C. Koehler, *Nature* **520**, 666–669 (2015).
2. M. C. Weiss *et al*., *Nat. Microbiol.* **1**, 16116 (2016).
3. E. S. Boyd, J. W. Peters, *Front. Microbiol.* **4**, 201 (2013).
4. K. Zheng, P. D. Ngo, V. L. Owens, X. P. Yang, S. O. Mansoorabadi, *Science* **354**, 339–342 (2016).
5. S. J. Moore *et al*., *Nature* **543**, 78–82 (2017).
6. N. Muraki *et al*., *Nature* **465**, 110–114 (2010).
7. J. Nomata, T. Mizoguchi, H. Tamiaki, Y. Fujita, *J. Biol. Chem.* **281**, 15021–15028 (2006).
8. P. C. Dos Santos, Z. Fang, S. W. Mason, J. C. Setubal, R. Dixon, *BMC Genomics* **13**, 162 (2012).
9. J. Raymond, J. L. Siefert, C. R. Staples, R. E. Blankenship, *Mol. Biol. Evol.* **21**, 541–554 (2004).
10. J. A. North, A. R. Miller, J. A. Wildenthal, S. J. Young, F. R. Tabita, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10455–E10464 (2017).
11. N. Parveen, K. A. Cornell, *Mol. Microbiol.* **79**, 7–20 (2011).
12. S. Burén, E. Jiménez-Vicente, C. Echavarri-Erasun, L. M. Rubio, *Chem. Rev.* **120**, 4921–4968 (2020).
13. Y. Zhang, E. L. Pohlmann, P. W. Ludden, G. P. Roberts, *J. Bacteriol.* **182**, 983–992 (2000).
14. T. J. Erb *et al*., *Nat. Chem. Biol.* **8**, 926–932 (2012).
15. D. Sippel, O. Einsle, *Nat. Chem. Biol.* **13**, 956–960 (2017).
16. L.-M. Zhang, C. N. Morrison, J. T. Kaiser, D. C. Rees, *Acta Crystallogr. D* **71**, 274–282 (2015).
17. D. Sippel *et al*., *Science* **359**, 1484–1489 (2018).
18. M. Bižić *et al*., *Sci. Adv.* **6**, eaax5343 (2020).
19. D. Repeta *et al*., *Nat. Geosci.* **9**, 884–887 (2016).
20. Y. Zheng *et al*., *Nat. Microbiol.* **3**, 281–286 (2018).
21. K. A. Smith, R. S. Russell, *Nature* **222**, 769–771 (1969).
22. S. M. N. Manik *et al*., *Front. Plant. Sci.* **10**, 140 (2019).
23. J. M. Lynch, *Nature* **240**, 45–46 (1972).
24. J. M. Lynch, *Nature* **256**, 576–577 (1975).
25. J. A. North *et al*., *Mol. Microbiol.* **113**, 923–937 (2020).
26. G. A. W. Beaudoin *et al*., *Nat. Commun.* **9**, 3105 (2018).
27. H. Zheng, C. Dietrich, R. Radek, A. Brune, *Environ. Microbiol.* **18**, 191–204 (2016).
28. R. Méheust *et al*., *ISME J.* 10.1038/s41396-020-0716-1 (2020).

ACKNOWLEDGMENTS

We thank D. Canniffe (University of Liverpool) for providing *Blastochloris viridis*. **Funding:** Electronic structure calculations were performed at the Environmental Molecular Sciences Laboratory (EMSL), a National Scientific User Facility sponsored by the U.S. Department of Energy (DOE), Office of Science, Biological and Environmental Research (BER) and located at Pacific Northwest National Laboratory, which is operated by Battelle, for the U.S. DOE under contract DE-AC05-76RLO 1830. Transcriptomics work was supported in part by the University of Colorado Cancer Center's Genomics and Microarray Shared Resource (NCI grant P30CA046934) with RNA sequencing services performed by K. Diener. Computational aspects used resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. Proteomics work at Oak Ridge National Laboratory (ORNL) was supported by the Genomic Science Program, U.S. DOE, Office of Science, BER as part of the Plant Microbe Interfaces Scientific Focus Area (http://pmiweb.ornl.gov). ORNL is managed by UT-Battelle, LLC, for the DOE under contract number DE-AC05-00OR22725. This work was supported by an OSU Center for Applied Plant Sciences Grant (to F.R.T.) and the Genomic Science Program, U.S. DOE, Office of Science, BER under award number DE-SC0019338 (to F.R.T., K.C.W., and W.R.C.). **Author contributions:** F.R.T., R.L.H., K.C.W., W.R.C., and J.A.N. designed the experiments. W.X. and R.L.H. performed proteomics. J.A.N. and A.B.N. performed transcriptomics, bioinformatics, and comparative genomics analysis. J.A.N., K.M.B., S.J.Y., G.Z., J.A.W., and S.M. performed the growth, genetics, and metabolite analysis assays. W.R.C. performed thermodynamic simulations. J.A.N., W.X., and A.B.N. analyzed the data. J.A.N. and F.R.T wrote the manuscript, with contributions and approval from all other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials, except as follows: All raw mass spectra and searched data files for the proteome measurements have been deposited into the ProteomeXchange repository: MassIVE accession number MSV000084455 (FTP link to files: ftp://massive.ucsd.edu/MSV000084455/) and ProteomeXchange accession number PXD015818. Raw transcriptomics reads are available in the NCBI Sequence Read Archive under accession numbers SRR10887327 through SRR10887338 and BioProject accession number PRJNA601207.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1094/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
Tables S1 to S6
References (*29*–*65*)
MDAR Reproducibility Checklist
Data S1 to S3

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

7 March 2020; accepted 12 June 2020
10.1126/science.abb6310

ELECTROCHEMISTRY

Accelerating water dissociation in bipolar membranes and for electrocatalysis

Sebastian Z. Oener*, Marc J. Foster, Shannon W. Boettcher*

Catalyzing water dissociation (WD) into protons and hydroxide ions is important both for fabricating bipolar membranes (BPMs) that can couple different pH environments into a single electrochemical device and for accelerating electrocatalytic reactions that consume protons in neutral to alkaline media. We designed a BPM electrolyzer to quantitatively measure WD kinetics and show that, for metal nanoparticles, WD activity correlates with alkaline hydrogen evolution reaction activity. By combining metal-oxide WD catalysts that are efficient near the acidic proton-exchange layer with those efficient near the alkaline hydroxide-exchange layer, we demonstrate a BPM driving WD with overpotentials of <10 mV at 20 mA·cm^{-2} and pure water BPM electrolyzers that operate with an alkaline anode and acidic cathode at 500 mA·cm^{-2} with a total electrolysis voltage of ~2.2 V.

Bipolar membranes (BPMs) consist of a polymeric cation-exchange layer (CEL) with fixed anionic groups and mobile cations in contact with an anion-exchange layer (AEL) with fixed cationic groups and mobile anions (*1*). When a sufficient bias is applied across a BPM, heterolytic water dissociation (WD), $H_2O \rightarrow H^+ + OH^-$, occurs at the AEL/CEL interface, with the H^+ driven through the CEL and the OH^- driven through the AEL (Fig. 1) (*2*). BPMs are used in electrodialysis, desalination, acid-base synthesis, or other electrochemical applications in which a steady-state difference in pH across the membrane is needed (*3*). Emerging applications include the use of BPMs in water (photo)electrolysis ($2H_2O \rightarrow 2H_2 + O_2$), in which the anode and cathode electrocatalyst can operate at a different local pH at steady state (*4*); in CO_2 electrolyzers, in which BPMs attenuate dissolved (bi)carbonate crossover because of the outward flux of H^+/OH^- (*5*–*9*); or in regenerative fuel cells that provide ideal pH conditions for bidirectional electrocatalysts (*10*).

State-of-the-art research and commercial BPMs typically show substantial overpotentials driving water dissociation (η_{wd}) of $\gtrapprox$ 100 mV even at current densities of only ~20 mA·cm^{-2} (table S1; see the supplementary text and figs. S1 and S2 for a precise definition of η_{wd}, including theory and simulation of BPMs). The best-reported BPM uses a three-dimensional (3D) electrospun AEL/CEL junction that passes ~0.5 A·cm^{-2} at η_{wd} ~0.6 V, with $Al(OH)_3$ nanoparticles catalyzing WD at the AEL/CEL interface and neutral salt solutions on both sides of the membranes (*11*). These η_{wd} values are too high for energy-conversion applications that probably require η_{wd} <100 mV at 1 to 2 A·cm^{-2}.

To improve BPM performance, the WD reaction must be accelerated. The addition of polymers such as sulfonated polyether ether ketone (*12*) or metal (hydr)oxide WD catalysts in the BPM junction, e.g., Al_2O_3, $Fe(OH)_3$, and $Cr(OH)_3$, can reduce η_{wd} compared with BPMs without catalyst (*13*–*17*). WD catalysts have also been introduced by the precipitation of metal salts (*17*), by sol-gel chemistry (*14*), or by the addition of nanoparticles (*15*). Previous studies investigated few WD catalysts at a time or focused on the membrane end groups at the AEL/CEL interface (*18*–*22*) and usually treated WD as occurring in a homogeneous, pH-neutral layer. Other polymers are often added in the junction to physically link the constituent layers, and a variety of ionomer membranes (with different chemical end groups and thus different inherent WD kinetics) have been used across different studies. Sometimes, membranes are physically roughened before assembly (*23*). So far, no systematic and comprehensive studies of WD as a function of catalyst composition and location inside the BPM junction have been reported, no consistent mechanistic picture has emerged, and the resulting BPM performance is insufficient for emerging applications (*24*, *25*).

Beyond BPMs, WD is important for any chemical reaction that uses water as a reactant. Understanding fundamental WD mechanisms in pure water, and the reverse H^+/OH^- recombination reaction, has been a focus of experiment and theory for decades (*26*–*28*). The related dissociative adsorption of water, $H_2O \rightarrow H_{ad} + OH_{ad}$, occurs on surfaces and is important when water reacts in thermochemical processes, such as for water-gas-shift catalysis (*29*). The interaction of water with surfaces has thus been of longstanding interest (*30*), with broad importance to corrosion, passivation, and geological processes (*31*). In biology, metalloenzymes activate water. Carbonic anhydrase, for example, binds water to

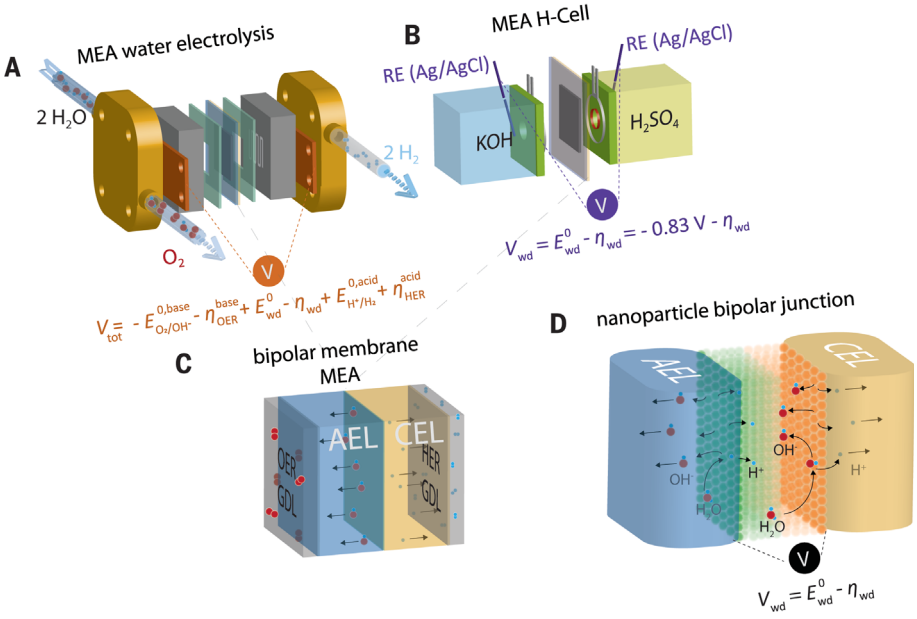


Fig. 1. BPM water-dissociation measurements. (**A**) WD is studied in an electrolyzer fed only pure water. The applied V_{tot} is negative, associated with the positive ΔG_{rxn} of water electrolysis. We define η_{wd} as positive. (**B**) A custom H-cell allows for the direct measurement of η_{wd} to calibrate the electrolysis setup. (**C**) The MEA is kept under constant mechanical pressure, preventing delamination. The MEA is composed of the anode and cathode gas-diffusion layers (GDLs) with IrO_2 OER catalyst and Pt HER catalyst, the AEL and CEL, and the WD catalyst. (**D**) Different nanoparticle WD catalysts can be placed on both the acidic CEL and basic AEL surfaces.

Department of Chemistry and Biochemistry, the Materials Science Institute, and the Oregon Center for Electrochemistry, University of Oregon, Eugene, OR 97403, USA.
*Corresponding author. Email: szo@uoregon.edu (S.Z.O.); swb@uoregon.edu (S.W.B.)

Zn^{2+} and releases H^+ to generate nucleophilic OH^- that reacts with CO_2 to form HCO_3^- (*32*). During the alkaline hydrogen-evolution reaction (HER, $2H_2O + 2e^- \rightarrow H_2 + 2OH^-$), important for low-cost alkaline water electrolysis (*33*), the first step involves WD coupled with the formation of a surface metal hydride ($H_2O + M^* + e^- \rightarrow M\text{-}H + OH^-$) and is thought to be rate limiting. Modification of metals with hydroxides, proposed to catalyze WD, substantially increases HER activity through a possible bifunctional mechanism (*34*, *35*), but alternative explanations have been proposed (*36*–*38*). No direct measurements of the isolated WD step, independent from the other reaction steps, have been conducted on electrocatalysts. WD was also hypothesized to affect the rate of electrocatalytic CO_2 reduction—a reaction that occurs in neutral to basic pH and therefore must use water as a proton source (*39*).

Here, we report the use of a BPM electrolyzer (Fig. 1) to measure WD kinetics on ~40 metal and metal-oxide nanoparticle assemblies composed of ~30 different materials. The electrolyzer was fed only 18.2 MΩ · cm water (without supporting electrolyte that might complicate results; Fig. 1A). The BPMs were encompassed in a membrane-electrode assembly (MEA) held under pressure by two rigid current collectors (Ti/Pt-frit) to prevent delamination and to ensure constant contact between the gas-diffusion electrodes, membranes, and WD-catalyst layers. As different WD catalysts were studied, the materials and preparation parameters used for the MEA were kept constant.

The voltage needed for BPM water electrolysis (V_{tot}) can be conceptually decomposed into the following: (i) the thermodynamic potential of generating O_2 in alkaline media and H_2 in acidic media ($E^{0,acid}_{H^+/H_2} - E^{0,base}_{O_2/OH^-} = -0.40$ V, although, rigorously, the reference half-reactions that define the E^0 on each side of the BPM are at different electrostatic potentials); (ii) the OER and HER overpotentials at a given current density (η_{OER} and η_{HER}, respectively); (iii) the thermodynamic potential needed to generate acid and base from a neutral aqueous solution ($E^0_{wd} = -\Delta G^0_{wd}/F = -0.83$ V at standard conditions); (iv) η_{wd} (defined to be positive here); and (v) the voltage loss caused by ohmic resistance, which is negligible at 20 mA·cm^{-2} for the MEA setup (*40*). We found that V_{tot} depends sensitively on the WD catalyst.

To quantify η_{wd} independently of the anode and cathode reactions, we built an H-cell that maintains the BPM MEA under pressure while contacting the CEL with 1.0 M acid and the AEL with 1.0 M base (Fig. 1B). The electrostatic potential across the BPM was measured using two pH-independent Ag/AgCl reference electrodes in the basic and acidic compartments while a constant current was passed between the anode and cathode (*2*). The η_{wd} values are taken as the increase in the magnitude of the electrostatic potential over the expected equilibrium value as a function of the current density (see the supplementary materials). Because we used high-conductivity electrolytes, and the anode and cathode were in direct contact with the BPM, resistive voltage losses between the BPM and reference electrodes were negligible (fig. S3). By measuring a BPM made with the same IrO_2 WD catalyst in both the H-cell and pure water electrolyzer, we calculated a theoretical electrolyzer baseline for $\eta_{wd} = 0$ V (figs. S4 to S7). Voltages measured for the BPM electrolyzer above this baseline are taken as η_{wd}.

The use of an electrolyzer to study WD kinetics inside the BPM instead of a conventional H-cell is important. We fed pure water, thereby eliminating co-ion (e.g., K^+, Cl^-) crossover in the conventional H-cell measurements that could carry current and potentially influence the WD reaction at the AEL/CEL junction through ion adsorption (*36*, *41*, *42*). In the pure water BPM electrolyzer, current is only carried by H^+ and OH^-. We used IrO_2 WD catalysts to calibrate the electrolyzer with the H-cell measurements because IrO_2 is stable under acidic and alkaline conditions. Co-ions in the H-cell setup are responsible for <2% of the current at 20 mA·cm^{-2} (fig. S35).

The calibrated BPM electrolyzer was used to study WD kinetics on metal-oxide nanoparticles deposited in the BPM junction (table S2). We found that η_{wd} is relatively insensitive to loading once the membrane is covered with ~100 to 200 nm of WD catalyst (see figs. S8 and S9). Because metal-oxide surfaces are polyacids and/or polybases, their surface chemistry and protonation state depend on pH, which could affect WD kinetics (*43*, *44*). The CEL surface is an acidic environment, whereas the AEL surface is alkaline. To study the WD catalysis under a defined local pH, we used a chemically stable WD catalyst at one surface (e.g., NiO at the AEL and Sb:SnO_2 at the CEL) while systematically varying the WD catalyst at the other membrane surface (polarization curves are shown in figs. S10 to S15). We separately measured each oxide's point of zero charge (PZC; fig. S16 and table S2) to assess acid-base properties; WD catalyst activity versus PZC is plotted in Fig. 2. With Sb:SnO_2 on the acidic CEL, the best WD catalysts on the basic AEL tend to be those with basic PZCs, such as NiO. With NiO on the basic AEL, the best WD catalysts on the acidic CEL are those with acidic PZCs, such as IrO_2 or SnO_2. At higher current densities, and thus larger η_{wd}, these weak correlations appear more pronounced (fig. S17).

A proton-transfer mechanism, similar to the one proposed for WD catalysis by ionomer-membrane end groups at the AEL/CEL interface (*18*, *19*, *22*), can be used to tentatively rationalize the data in Fig. 2. Under basic conditions at the AEL, most oxide surfaces are negatively charged and consist of structurally diverse –OH and –O^- species. These sites might act as proton donors or acceptors,

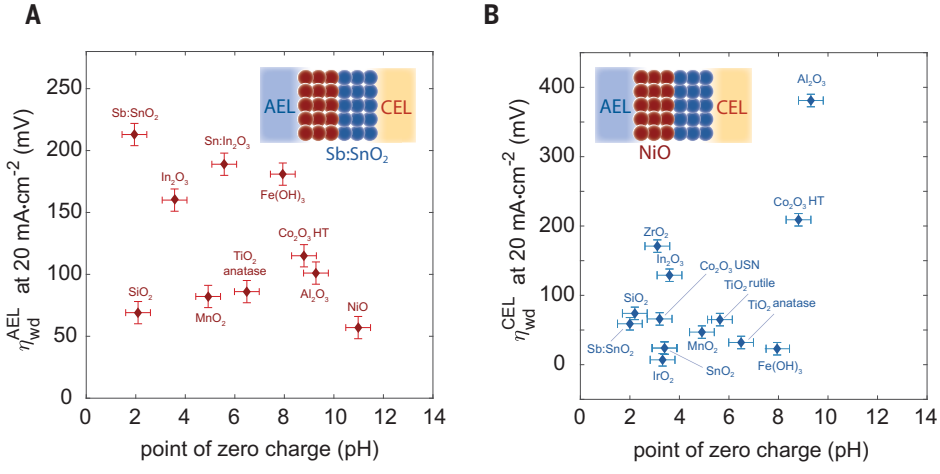


Fig. 2. WD overpotentials and PZC for oxide nanoparticles. (**A**) Values for η_{wd} measured with catalysts placed under locally alkaline conditions in the BPM (η^{AEL}_{wd})using acid-stable Sb:SnO_2 to cover the acidic CEL. (**B**) Values for η_{wd} measured with catalysts placed under locally acidic conditions (η^{CEL}_{wd}) using base-stable NiO to coat the alkaline AEL. The PZCs were measured by finding the pH where the zeta potential (at the hydrodynamic shear plane) tended to zero. The uncertainties in the PZC data are based on measurements made in triplicate (see the supplementary materials and fig. S16), whereas those for η_{wd} reflect the uncertainty of the calibration of the MEA with the H-cell setup (see the supplementary materials and figs. S4 to S6).

respectively, catalyzing WD through a two-step mechanism as follows:

$$MO^- + H_2O \underset{k_1'}{\overset{k_1}{\rightleftarrows}} MOH + OH^- \quad (1)$$

$$MOH + H_2O \underset{k_2'}{\overset{k_2}{\rightleftarrows}} MO^- + H_3O^+ \quad (2)$$

In the above case, MO^- would be a catalyst active site, with MOH an intermediate. Under acidic conditions at the CEL, most oxide surfaces will be protonated, consisting of $-OH_2^+$ and -OH species, and can similarly catalyze WD as follows:

$$MOH_2^+ + H_2O \underset{k_3'}{\overset{k_3}{\rightleftarrows}} MOH + H_3O^+ (3)$$

$$MOH + H_2O \underset{k_4'}{\overset{k_4}{\rightleftarrows}} MOH_2^+ + OH^- (4)$$

A large net rate for either of the above two-step mechanisms requires sufficient concentrations of both the protonated and unprotonated surface species (assuming similarity of the forward rate constants k_1 and k_2 or k_3 and k_4), so that both steps are simultaneously fast. At local pH values near the PZC, a diversity of surface protonation states capable of driving both steps is needed for fast WD. Although this is consistent with the observation that oxide WD catalysts tend to work best when the local pH is near the PZC, the weak trends shown in Fig. 2 illustrate that other factors, such as accessible surface area, the diversity and range of surface pK_a, and the density of surface hydroxyl groups, are also likely to affect WD activity. For materials without a substantial density of ionizable surface groups (e.g., noble metals) PZC data are unlikely to be useful.

Although alternative mechanisms involving the dissociative adsorption of water to form M–OH intermediates are also possible (see the supplementary materials), we hypothesize that these pathways are slow on typical oxides because of the strength of the M–O bonds. Such mechanisms may be prevalent on the precious metals discussed below. A few oxides, e.g., TiO_2 and SiO_2, catalyze WD reasonably in both acidic and basic environments. This may be because of the broad surface pK_a range for TiO_2 and SiO_2 (*45*, *46*), which are better described as "spectra" than discrete pK_a values (*44*). Some oxides are also not stable in acid or base. NiO, for example, would partially dissolve on the CEL surface generating soluble Ni^{2+} that can ion exchange for H^+ in the CEL. Improved mechanistic insight will require a better understanding of the WD-catalyst surface speciation. Raman and Stark effect spectroscopy, for example, might be suitable to detect chemical changes on the catalyst surface during operation and help to elucidate the role of the electric field in the junction, respectively.

These findings suggest a new principle for the design of WD catalysts that accounts for the surface speciation and chemical stability

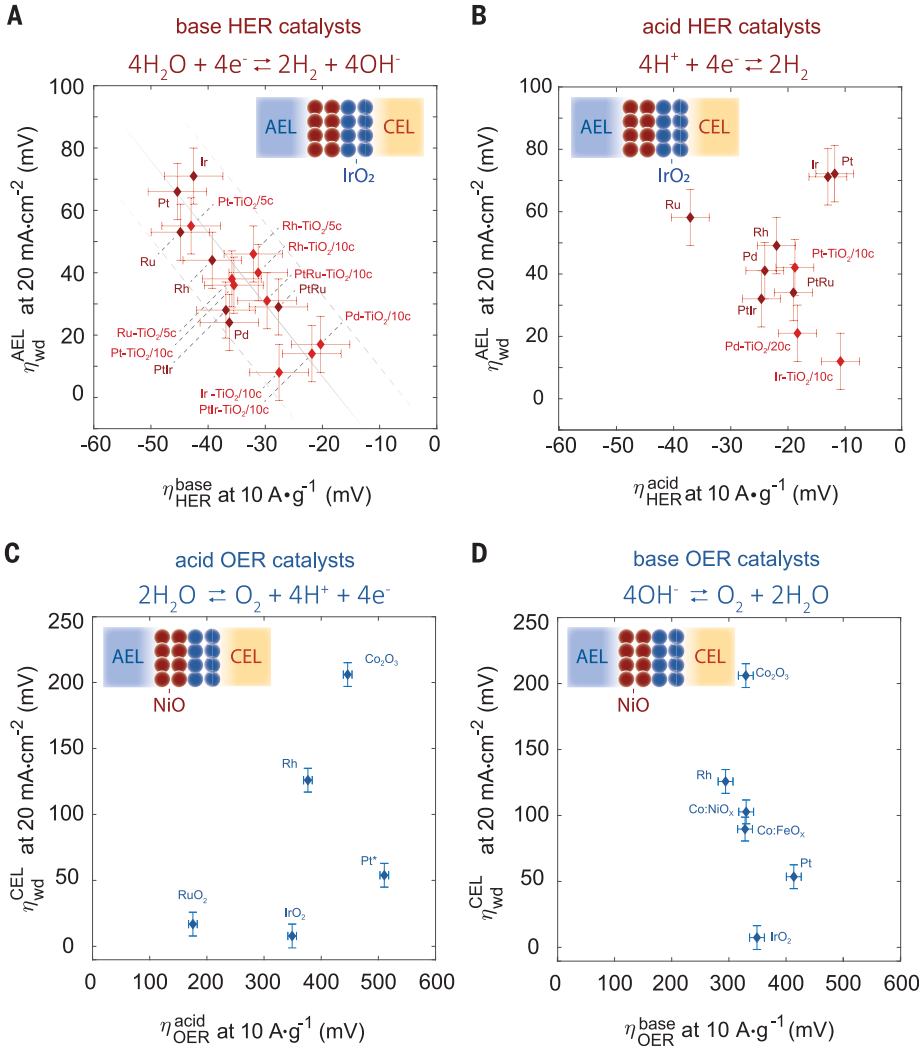


Fig. 3. Comparison between WD and electrocatalytic overpotentials measured on the same nanoparticles. (**A**) The overpotentials for WD in a locally basic environment (with IrO_2 used as the stable WD catalyst on the acidic CEL surface) correlate with the overpotentials measured for HER in 1.0 M KOH [$\rho(\eta_{HER}^{base}, \eta_{wd}^{AEL})$ = −0.81], consistent with WD rate limiting the alkaline HER. (**B**) No correlation is found between η_{wd} and the HER performance in 1.0 M H_2SO_4 [$\rho(\eta_{HER}^{acid}, \eta_{wd}^{AEL})$ = −0.13], consistent with WD not being involved in the acidic HER mechanism. Dark-red markers in (A) and (B) show the as-received nanoparticles and light-red markers show the samples after TiO_2 decoration. (**C**) The η_{wd} in a locally acidic environment (with NiO used as the stable WD catalyst on the basic AEL surface) do not strongly correlate with the overpotentials measured for OER in 1.0 M H_2SO_4 [$\rho(\eta_{OER}^{acid}, \eta_{wd}^{CEL})$ = 0.49], consistent with WD not being rate limiting. (**D**) No strong correlation is present between η_{wd} and the OER overpotentials in 1.0 M KOH [$\rho(\eta_{OER}^{base}, \eta_{wd}^{CEL})$ = −0.45], consistent with the large OH^- concentration obviating the need for WD. The loadings for the electrocatalysis studies were 0.1 to 1 mg · cm^{-2} and the current densities were 1 to 10 mA·cm^{-2}, similar to the WD current densities studied in the BPM. The currents for the electrocatalysts are normalized to mass because higher loadings lead to larger current. The WD current was not normalized to mass because a simple correlation between activity and loading does not exist (see fig. S8). The uncertainties for the OER and HER overpotentials are based on the average standard deviation obtained for each electrolyte condition, whereas those for η_{wd} reflect the uncertainty of the MEA calibration with the H-cell setup (see the supplementary materials and figs. S4 to S6). The Pt OER activity degraded with subsequent cycles because of oxidation (see the supplementary materials).

of the catalyst at the relevant local pH. In the context of BPMs, the data show the superiority of using two WD catalyst layers selected for the appropriate local pH over a single material as in prior work. We envision designing multilayered WD catalysts that provide a gradient in surface acid-base behavior matched to the pH gradient present at the AEL/CEL interface at the operational current density.

The new data also suggest that catalysis of WD is more important than electric-field enhancement (Onsager's theory of the second Wien effect) in the BPM junction, a topic of a long-standing debate (*18*, *24*, *47*, *48*). The electric field across the AEL/CEL junction is ~ $3 \cdot 10^4$ V · cm^{-1}, assuming the electrostatic potential drops linearly across an ~300-nm-thick WD-catalyst layer. This field is not sufficient to substantially increase the rate of WD based on Onsager's theory (see the supplementary materials) (*47*), consistent with previous calculations (*49*). We observed no apparent correlation between η_{wd} and the electronic properties of the oxide. Electrical conductors (IrO_2, RuO_2), *n*- and *p*-type semiconductors (SnO_2, NiO), and insulators (SiO_2) all modify and/or screen the electric field distribution inside the WD-catalyst layer differently yet are all efficient WD catalysts. The η_{wd} values are also temperature dependent (fig. S18). The BPM made without added WD catalyst has the narrowest junction, sharpest electrochemical potential gradient, and thus the largest electric field, but the performance is the worst (see below). These findings suggest the rates of WD are influenced primarily by surface catalysis in the presence of gradients in $\bar{\mu}_{H^+}$ and $\bar{\mu}_{OH^-}$ that separate H^+ from OH^-, with electric field enhancement playing a secondary role.

WD has further been proposed to be rate determining in some electrocatalytic reactions. During the HER in alkaline media ($2H_2O + 2e^- \rightarrow H_2 + 2OH^-$), the first step ($H_2O + e^- \rightarrow H_{ad} + OH^-$) involves WD and metal hydride formation and is thought to be the reason HER is slower in base than in acid (*34*, *35*). The OER under acidic or neutral conditions, $2H_2O \rightarrow O_2 + 4H^+ + 4e^-$, likewise requires WD to generate -OH_{ad} species that can be further oxidized. WD, however, is not thought to be rate limiting for the OER (*50*). A fundamental challenge is the inability to isolate and measure WD kinetics for a particular electrocatalyst independent of other electrochemical processes. Traditionally, the driving force for both WD and electron transfer is simultaneously provided as an applied electrode potential. Here, we isolate WD kinetics by locally generating a gradient in the electrochemical potentials $\bar{\mu}_{H^+}$ and $\bar{\mu}_{OH^-}$ inside the electrically disconnected BPM junction (see the supplementary materials for simulations).

We used the BPM electrolyzer to measure η_{wd} of common HER and OER electrocatalysts (Fig. 3) and analyzed them to ascertain whether η_{wd} correlates with electrocatalytic overpotentials (at similar current density) measured for the same particles in a three-electrode cell (figs. S19 to S23). We found that η_{HER} of precious metal nanoparticles in 1.0 M KOH linearly correlates with η_{wd} measured at the basic AEL surface (Pearson correlation coefficient $\rho(\eta_{HER}^{base}, \eta_{wd}^{AEL}) = -0.81$). As η_{wd} tends to zero, so does η_{HER} at a similar current density. When η_{HER} for those metal nanoparticles is measured in 1.0 M H_2SO_4, where WD is not needed for the HER, there is no significant correlation and $\rho(\eta_{HER}^{acid}, \eta_{wd}^{AEL}) = -0.13$.

The above correlation is consistent with the hypothesis that modifying metals such as Pt with metal hydroxides such as $Ni(OH)_2$ improves alkaline HER activity because of increased WD kinetics (*34*, *35*). Our results show that NiO (which is hydroxylated in water) is an excellent alkaline WD catalyst and will further enable the design of new electrocatalysts. TiO_2 is an active WD catalyst in both acid and base and is chemically stable. We deposited TiO_2 by atomic-layer deposition (ALD) onto Pt, Ir, Ru, Rh, PtRu, and PtIr nanoparticles at 250°C. The addition of TiO_2 simultaneously decreased both η_{wd} and η_{HER}^{base}, providing direct support to a bifunctional alkaline HER mechanism in which the oxide facilitates WD and the metal stabilizes M–H intermediates.

A water-ordering model has been alternatively proposed to explain the increased alkaline HER kinetics of metal-hydroxide-functionalized Pt (*37*). When water is dissociated during HER in base, OH^- is produced and stabilized (i.e. solvated) by orienting nearby water dipoles. Pt, under alkaline HER conditions, carries a substantial surface charge, which also orients nearby water molecules and generates an energy barrier to water reorganization. $Ni(OH)_2$ lowers the surface charge of Pt under alkaline HER conditions, which could reduce the water reorganization energy and thus η_{HER}^{base} (*37*). Although the general trends of PZC with WD rate (Fig. 2) are consistent with this hypothesis, some of our data are not. SiO_2 has an acidic PZC yet works well for WD in base, where it presents charged surfaces that would, in the above model, be expected to retard WD. TiO_2 and $Ni(OH)_2$ have very different PZCs, and thus are expected to have very different net surface charge densities in base. However, both enhance alkaline HER kinetics on Pt and other metal surfaces. This evidence, along with the proposed proton-transfer WD mechanism (Eqs. 1 to 4), favors a bifunctional route for alkaline HER on (hydr)oxide-functionalized metal surfaces. Further work is needed, however, to understand the effects of water ordering on WD activity for the oxides studied here, which are likely to have a diversity of hydroxylated surface sites with a range of pK_a values.

In contrast to alkaline HER, we did not find a strong correlation between η_{wd} and the OER activity in acid [$\rho(\eta_{OER}^{acid}, \eta_{wd}^{CEL}) = 0.49$], where H_2O is also a reagent (Fig. 3C), nor between η_{wd} and the OER activity in base [$\rho(\eta_{OER}^{base}, \eta_{wd}^{CEL}) = -0.45$]. This is consistent with computational models of OER indicating that nonoptimal M–O binding strengths provide rate-limiting barriers (*50*). It is, however, interesting that fast acid-OER catalysts, such as RuO_2 and IrO_2, are also fast WD catalysts.

Although the durability of the WD catalysts in the BPM have not been studied extensively, the observed trends (Figs. 2 and 3) cannot be explained by differences in WD catalyst stability or dissolution (which might provide current-carrying ions or be coupled to WD). The electrolyzer measurements used pure water without soluble electrolyte, and therefore crossover of acid or base into the BPM junction is avoided. The best WD catalysts (IrO_2 in acid and NiO in base) should not substantially

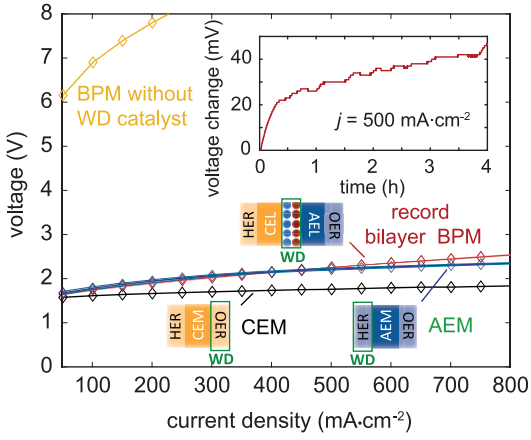


Fig. 4. Alkaline, acidic, and BPM electrolyzers. BPM electrolyzers require >6 V to drive electrolysis at 50 mA·cm^{-2} without additional WD catalysts (yellow). When the new bilayer WD catalysts are added, the performance substantially improves and is similar to a reference AEM electrolyzer. The PEM electrolyzer performs better, which is likely because of the superior properties of Nafion compared with Sustainion. For all electrolyzers, the cathode and anode catalyst composition and loading were identical. The AEM and bilayer BPM electrolyzer polarization curves are averages over three devices and/or datasets, with the error indicated by the thickness of the line (fig. S32). The inset shows that the BPM electrolyzer performance is relatively stable over 4 hours of operation at 500 mA·cm^{-2} (see the supplementary materials for discussion of degradation modes and mitigation strategies).

dissolve based on thermodynamic data (*51*). The rate of voltage increase for the BPM electrolyzer is also not correlated with η_{wd} (figs. S24 to S30), and continuous operation for >4 hours at 500 mA·cm^{-2} leads to an increase of only 40 mV in the BPM electrolysis voltage (Fig. 4).

Taken together, these results help to unify concepts of WD catalysis across the membrane and electrocatalysis sciences. The BPM electrolyzer architecture allows WD reaction rates to be directly probed in relevant acidic or basic local-pH environments and independently of electron-transfer processes. Bifunctional HER catalysts, as demonstrated by others (*34*, *35*) and shown here for Pt/TiO_2, can now be assembled based on independently measured WD activity. This approach should facilitate the design of electrocatalysts for any electrochemical process that uses water as a reactant (including the electrochemical reduction of CO_2 and N_2).

The BPMs fabricated here also have faster WD kinetics than those previously reported (see table S1 for η_{wd} values). The best BPMs, e.g., those with NiO or IrO_2 coated with 10 layers of TiO_2 by ALD (Ir/TiO_2-10c) at the AEL surface and RuO_2 or IrO_2 at the CEL surface, show η_{wd} that is within measurement error (±9 mV) of 0 mV at 20 mA·cm^{-2} in pure water at 50°C. Fig. S31 shows a comparison of a NiO|IrO_2 bilayer BPM with other BPMs measured under comparable H-cell conditions at 25°C. Commercial membranes, e.g., those from Fumatech, show η_{wd} ranging from 100 to 200 mV at 20 mA·cm^{-2} (depending on pretreatment). Research BPMs generally show similar or substantially higher η_{wd} [see fig. S31, table S1, and (*52*)]. Figure 4 compares the polarization curves of our best BPM pure water electrolyzers to reference alkaline-exchange membrane (AEM) and proton-exchange membrane (PEM) electrolyzers, which are made from the same membranes and OER/HER electrocatalysts. The BPM electrolyzers meet or exceed the performance of the AEM electrolyzers up to 400 mA·cm^{-2}. This is possible because, in the AEM electrolyzer, Pt must drive HER in alkaline media with relatively slow kinetics. In the BPM electrolyzer, Pt drives HER with fast kinetics in a locally acidic environment, whereas protons are provided through catalyzed WD in the AEL/CEL junction. This result proves that η_{wd} is relatively small, even at high current densities, for the new BPMs reported here. State-of-the-art 3D junction BPMs, similarly measured with acid and base on each side of the BPM, show η_{wd} of ~1 V at 0.2 A cm^{-2} (*24*). Here, we drove water electrolysis at 0.2 A cm^{-2} with a total voltage of ~1.9 V and an estimated η_{wd} of <0.25 V. For current densities beyond 400 mA·cm^{-2}, the BPM electrolyzer voltage is higher than that of the AEM electrolyzer. We ascribe this difference to water-transport limitations and ionic series resistance inside the BPM junction (see the supplementary materials).

Further innovations are possible to improve performance and achieve small η_{wd} at the ~2 A cm^{-2} relevant to many technologies. Examples include engineering the surface area and surface chemistry of the WD catalyst, the number of different WD-catalyst layers, the interfacial area with a 3D junction (*11*, *24*), the ionomer thickness and chemistry to improve water transport to the AEL/CEL junction, and the ionic conductivity in the WD catalyst layer. The performance decay (Fig. 4, inset), which may be related to junction dehydration at high currents, must also be mitigated before commercial application (see the supplementary materials).

The BPM enables electrochemical devices in which ideal pH conditions can be selected independently for each half-reaction, in contrast to traditional architectures, in which electrode stability, catalyst activity, and cost must be simultaneously optimized at both cathode and anode for a single pH. This system design flexibility could enable high-performance, nonprecious metal BPM electrolyzers. Earth-abundant metal oxides can be used to catalyze the WD reaction inside the BPM junction, with optimal earth-abundant OER catalysts (e.g., Ni_xFe_yO) at the basic anode and HER catalysts at the acidic cathode. Flow batteries that use anolyte and catholyte at different pH values (*53*, *54*), (regenerative) fuel cells (*10*), and direct CO_2 electrolyzers that prevent soluble co-ion crossover (*7*) are other emerging applications facilitated by high-performance BPMs.

REFERENCES AND NOTES

1. T. Luo, S. Abdu, M. Wessling, *J. Membr. Sci.* **555**, 429–454 (2018).
2. R. S. Reiter, W. White, S. Ardo, *J. Electrochem. Soc.* **163**, H3132–H3134 (2016).
3. Y. Tanaka, *Ion Exchange Membranes: Fundamentals and Applications* (Elesevier, 2007).
4. D. A. Vermaas, M. Sassenburg, W. A. Smith, *J. Mater. Chem. A Mater. Energy Sustain.* **3**, 19556–19562 (2015).
5. Y. C. Li *et al*., *ACS Energy Lett.* **1**, 1149–1153 (2016).
6. D. A. Vermaas, W. A. Smith, *ACS Energy Lett.* **1**, 1143–1148 (2016).
7. Y. C. Li *et al*., *Adv. Sustainable Syst.* **2**, 1700187 (2018).
8. M. Lin, L. Han, M. R. Singh, C. Xiang, *ACS Appl. Energy Mater.* **2**, 5843–5850 (2019).
9. Y. C. Li *et al*., *ACS Energy Lett.* **4**, 1427–1431 (2019).
10. J. M. Ahlfield, L. Liu, P. A. Kohl, *J. Electrochem. Soc.* **164**, F1165–F1171 (2017).
11. C. H. Shen, R. Wycisk, P. N. Pintauro, *Energy Environ. Sci.* **10**, 1435–1442 (2017).
12. J. Balster *et al*., *J. Membr. Sci.* **365**, 389–398 (2010).
13. S. S. Mel'nikov, O. V. Shapovalova, N. V. Shel'deshov, V. I. Zabolotskii, *Petrol. Chem.* **51**, 577–584 (2011).
14. A. M. Rajesh, T. Chakrabarty, S. Prakash, V. K. Shahi, *Electrochim. Acta* **66**, 325–331 (2012).
15. F. Hanada, K. Hiraya, N. Ohmura, S. Tanaka, Bipolar membrane and method for its production. European Patent EP0459820B1 (1991).
16. Y. Oda, T. Yawataya, *Desalination* **5**, 129–138 (1968).
17. M.-S. Kang, Y.-J. Choi, H.-J. Lee, S.-H. Moon, *J. Colloid Interface Sci.* **273**, 523–532 (2004).
18. R. Simons, *Nature* **280**, 824–826 (1979).
19. R. Simons, G. Khanarian, *J. Membr. Biol.* **38**, 11–30 (1978).
20. H. Strathmann, H. J. Rapp, B. Bauer, C. M. Bell, *Desalination* **90**, 303–323 (1993).
21. A. B. Yaroslavtsev, V. V. Nikonenko, V. I. Zabolotsky, *Russ. Chem. Rev.* **72**, 393–421 (2003).
22. V. I. Zabolotskii, N. V. Shel'deshov, N. P. Gnusin, *Russ. Chem. Rev.* **57**, 801–808 (1988).
23. A. J. B. Kemperman, *Handbook on Bipolar Membrane Technology* (Twente Univ. Press, 2000).
24. Z. Yan *et al*., *Energy Environ. Sci.* **11**, 2235–2245 (2018).
25. S. Abdu *et al*., *ACS Appl. Mater. Interfaces* **5**, 10445–10455 (2013).
26. M. Moqadam *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4569–E4576 (2018).
27. M. Eigen, L. d. Maeyer, *Z. Elecktrochem.* **59**, 986 (1955).
28. G. Ertl, H. Gerischer, *Z. Elecktrochem.* **66**, 560 (1962).
29. C. V. Ovesen, P. Stoltze, J. K. Nørskov, C. T. Campbell, *J. Catal.* **134**, 445–468 (1992).
30. M. Sterrer *et al*., *J. Mater. Res.* **34**, 360–378 (2019).
31. P. A. Thiel, T. E. Madey, *Surf. Sci. Rep.* **7**, 211–385 (1987).
32. I. Bertini, C. Luchinat, "The reaction pathways of zinc enzymes and related biological catalysts," in *Bioinorganic Chemistry* (University Science Books, 1994), pp. 37–106.
33. D. Xu *et al*., *ACS Catal.* **9**, 7–15 (2019).
34. N. Danilovic *et al*., *Angew. Chem. Int. Ed.* **51**, 12495–12498 (2012).
35. R. Subbaraman *et al*., *Science* **334**, 1256–1260 (2011).
36. X. Chen, I. T. McCrum, K. A. Schwarz, M. J. Janik, M. T. M. Koper, *Angew. Chem. Int. Ed.* **56**, 15025–15029 (2017).
37. I. Ledezma-Yanez *et al*., *Nat. Energy* **2**, 17031 (2017).
38. J. Zheng, W. Sheng, Z. Zhuang, B. Xu, Y. Yan, *Sci. Adv.* **2**, e1501602 (2016).
39. M. Luo *et al*., *Nat. Commun.* **10**, 5814 (2019).
40. S. Z. Oener, S. Ardo, S. W. Boettcher, *ACS Energy Lett.* **2**, 2625–2634 (2017).
41. J. N. Mills, I. T. McCrum, M. J. Janik, *Phys. Chem. Chem. Phys.* **16**, 13699–13707 (2014).
42. E. Liu *et al*., *J. Am. Chem. Soc.* **141**, 3232–3239 (2019).
43. H. P. Boehm, *Discuss. Faraday Soc.* **52**, 264–275 (1971).
44. D. S. Smith, F. G. Ferris, *Environ. Sci. Technol.* **35**, 4637–4642 (2001).
45. X. Liu, J. Cheng, X. Lu, R. Wang, *Phys. Chem. Chem. Phys.* **16**, 26909–26916 (2014).
46. M. K. Ridley, V. A. Hackley, M. L. Machesky, *Langmuir* **22**, 10972–10982 (2006).
47. L. Onsager, *J. Chem. Phys.* **2**, 599–615 (1934).
48. M. A. Blommaert, D. A. Vermaas, B. Izelaar, B. in 't Veen, W. A. Smith, *J. Mater. Chem. A Mater. Energy Sustain.* **7**, 19060–19069 (2019).
49. H. Strathmann, J. J. Krol, H. J. Rapp, G. Eigenberger, *J. Membr. Sci.* **125**, 123–142 (1997).
50. I. C. Man *et al*., *ChemCatChem* **3**, 1159–1165 (2011).
51. M. Pourbaix, *Atlas of Electrochemical Equilibria in Aqueous Solutions* (National Association of Corrosion, 1974).
52. M. B. McDonald, S. Ardo, N. S. Lewis, M. S. Freund, *ChemSusChem* **7**, 3021–3027 (2014).
53. M. Park *et al*., *Adv. Energy Mater.* **9**, 1900694 (2019).
54. J. Xia, G. Eigenberger, H. Strathmann, U. Nieken, *J. Membr. Sci.* **565**, 157–168 (2018).
55. S. Z. Oener, M. J. Foster, S. W. Boettcher, Tabulated data for: Accelerating water dissociation in bipolar membranes and for electrocatalysis. Zenodo (2020); .

ACKNOWLEDGMENTS

We thank M. Lonergan for assistance in developing the numerical simulations, S. Ardo for insightful discussions regarding BPM physics and ion-transfer reactions, and J. Fehrs for assistance analyzing ion-crossover rates. **Funding:** This work was supported by the National Science Foundation Chemical Catalysis program under grant no. CHE-1566348. S.Z.O acknowledges support from a research fellowship of the German Research Foundation (Deutsche Forschungsgemeinschaft) under project no. 408246589 (OE 710/1-1). We acknowledge use of shared instrumentation in the CAMCOR and Rapid Materials Prototyping facilities, which are supported by grants from the M. J. Murdock Charitable Trust, the W. M. Keck Foundation, ONAMI, and the NSF. **Author contributions:** S.Z.O and S.W.B. conceived of the project, analyzed the data, and wrote the manuscript. S.Z.O. completed the experiments with assistance from M.J.F. **Competing interests:** S.Z.O, S.W.B., and M.J.F have filed a provisional patent application (U.S. Patent 62/984,652) based on the results of this manuscript. **Data and materials availability:** Tabulated data are available in the Zenodo repository (*55*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1099/suppl/DC1
Materials and Methods
Figs. S1 to S36
Tables S1 and S2
References (*56*–*73*)

15 August 2019; resubmitted 20 December 2019
Accepted 22 June 2020
Published online 2 July 2020
10.1126/science.aaz1487

CARBON NANOTUBES

Super-durable ultralong carbon nanotubes

Yunxiang Bai[1,2], Hongjie Yue[1], Jin Wang[2,3], Boyuan Shen[1], Silei Sun[1], Shijun Wang[2,3], Haidong Wang[4], Xide Li[2,3], Zhiping Xu[2,3], Rufan Zhang[1]*, Fei Wei[1,2]*

Fatigue resistance is a key property of the service lifetime of structural materials. Carbon nanotubes (CNTs) are one of the strongest materials ever discovered, but measuring their fatigue resistance is a challenge because of their size and the lack of effective measurement methods for such small samples. We developed a noncontact acoustic resonance test system for investigating the fatigue behavior of centimeter-long individual CNTs. We found that CNTs have excellent fatigue resistance, which is dependent on temperature, and that the time to fatigue fracture of CNTs is dominated by the time to creation of the first defect.

Strong materials with outstanding fatigue resistance are in demand in fields such as artificial muscles, airframes, suspension bridges, body armor, cables, sports equipment, and even space elevators (*1*–*3*). Low-dimensional materials such as carbon nanotubes (CNTs) and graphene generally have excellent strength. Extensive research has been conducted on the intrinsic strength of these nanoscale materials (*4*–*6*). It is known that failure in materials is primarily caused by fatigue, rather than by widespread fracture of chemical bonds (*7*). Under cyclic stress, failure occurs at a stress value lower than the intrinsic strength under static loading (*7*). The current dominant paradigm used for understanding the fatigue process is based on defects causing stress concentration during cyclic loading and subsequent failure once critical-length cracks form.

We fabricated centimeter-long CNT arrays (*8*–*10*). Figure 1, A, B, and D, shows these suspended CNTs aligned across the trenches on a Si/SiO_2 substrate. The structure of the CNTs was characterized by transmission electron microscopy (TEM) (Fig. 1C), resonance Rayleigh scattering (Fig. 1E), and Raman spectroscopy (see below). CNTs with a different chirality show a different color in resonance Rayleigh scattering (*11*), and so a single CNT that has a change in structure or chirality along its axial direction will show a color change in this direction. Thus, resonance Rayleigh scattering of CNTs offers an efficient way to identify their chiral consistency (*12*, *13*). As shown in Fig. 1E, the as-grown CNTs gave a monochromatic characteristic, which showed that their chirality remained consistent along the axial direction (*4*, *11*). A noncontact acoustic resonance test (ART) system equipped with a nanoprobe system was designed to investigate the mechanical behavior of the individual CNTs (Fig. 1G and fig. S1). We deposited TiO_2 nanoparticles onto suspended CNTs to enable their optical visualization (Fig. 1B and fig. S2) (*14*) and to control their resonance frequency, which was done by varying the string linear density. By changing the amount of TiO_2 on the CNTs (fig. S2), the resonance frequency was decreased from MHz to hundreds of Hz (text S1). The resonance vibration was excited by emitting low-frequency acoustic waves from a loudspeaker controlled by a digital signal generator (Fig. 1, F and G). Unlike common nanomaterial testing systems that use an electron microscope (*15*), our ART system under ambient conditions not only avoids the potential for defect formation that can be caused by an electron beam, but also enables the testing of samples that are orders of magnitude longer and overcomes the measurement difficulties for nanomaterials, such as specimen gripping and large cycles of tension loading.

The tensile stress–strain curve for an individual CNT at ambient temperature (300 K) was measured using a cantilever test. A typical curve is shown in Fig. 2A. This showed nonlinear elastic behavior, in agreement with the widely reported elasticity of CNTs (*16*, *17*). A tensile strength of 118.9 ± 4.5 GPa and a breaking strain of 16.41 ± 0.22% were obtained. The toughness was calculated as 8.0 ± 0.2 GJ/m^3 (text S2). We measured the fatigue life (in terms of N, number of cycles) of individual CNTs under different applied strains by changing the voltage and frequency to the loudspeaker. The applied strain here is an upper estimate extracted from the deflection shapes of CNTs in the ART (text S3). CNTs possess striking mechanical resilience (*18*–*20*). During the entire ART process, the CNTs vibrate in a plane (Fig. 1F and movie S1) with constant amplitude and frequency; that is, no change in the vibrational state was observed. The normalized strain/stress gave the ratio of applied strain/stress to the average tensile breaking strain ε_b (16.36%) or average tensile strength σ_b (119.2 GPa) of the CNTs obtained from the cantilever test at ambient temperature (text S4). The fatigue life N of the CNTs and corresponding applied normalized strain at ambient temperature at a loading frequency of 500 Hz are shown in

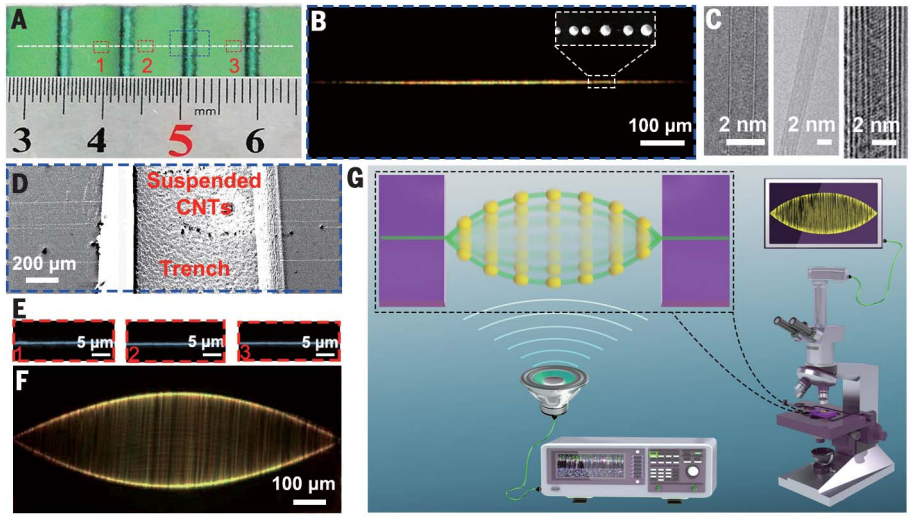


Fig. 1. Structure and fatigue resistance characterization of CNTs. (**A**) Optical image of a Si/SiO_2 substrate with trenches for growing CNTs. (**B**) Optical microscope image of a suspended CNT. Inset: SEM image of decorated TiO_2 nanoparticles. (**C**) TEM micrographs of CNTs. (**D**) SEM image of CNTs growing across a trench. (**E**) Resonance Rayleigh scattering spectra of a CNT. (**F**) Optical image showing the vibration of a suspended CNT; L = 1005.9 μm, h = 143.1 μm, $dy/dx|_{x=L/2}$ = 0.569 (see fig. S1E). (**G**) Schematic illustration of the acoustic resonance test (ART) system.

[1]Beijing Key Laboratory of Green Chemical Reaction Engineering and Technology, Department of Chemical Engineering, Tsinghua University, Beijing 100084, China. [2]Center for Nano and Micro Mechanics, Tsinghua University, Beijing 100084, China. [3]Department of Engineering Mechanics, Tsinghua University, Beijing 100084, China. [4]Key Laboratory for Thermal Science and Power Engineering of Ministry of Education, School of Aerospace, Tsinghua University, Beijing 100084, China.
*Corresponding author. Email: wf-dce@tsinghua.edu.cn (F.W.); zhangrufan@tsinghua.edu.cn (R.Z.)

Fig. 2B. In this work, the loading frequency of the ART was 500 Hz and experiments were conducted at ambient temperature, except for the tests studying the frequency and temperature effect. The applied normalized strain and fatigue life of the CNTs showed a nearly linear relationship on a semi-log plot.

The fatigue limit of a material is the maximum amplitude of cyclic stress/strain that can be applied to it without causing failure. The fatigue limit is often defined as 10^7 cycles (corresponding to ~5.5 hours for a loading frequency of 500 Hz) (*7*). In a design that uses the material, the designed stress levels must remain below the fatigue limit, which is the value of the fatigue limit as a material property. In Fig. 2B, the circles represent CNTs fractured during the loading cycles and the triangles represent those that survived after 10^7 cycles of loading. The normalized fatigue strain limit ε_f was obtained from this to be >0.90 (at the condition of $N = 10^7$ where we consider the fatigue limit was reached). Furthermore, after 10^7 loading cycles at an applied normalized stress of 0.95 (estimated from the applied strain and typical stress-strain curves shown in Fig. 2A), the residual strength (static tensile strength of the CNTs after releasing cyclic loading) did not show obvious change (Fig. 2C). This indicated that the fatigue ratio [normalized fatigue strength or the ratio (σ_f/σ_b) of fatigue strength σ_f to tensile strength σ_b] was >0.95.

Figure S3 shows the tension-relaxation stress-strain curves for a pristine CNT and one that has been loaded for 10^7 cycles at a normalized stress of 0.95. The curves largely overlap. It was also reported that even a single defect can reduce the tensile strength of CNTs by as much as 50% (*21*); however, the almost unchanged residual strength at the different testing stages (Fig. 2C) and the large overlap of curves in fig. S3 indicated that no detectable defects were generated in the whole ART process. To further verify this, we used Raman spectroscopy to monitor possible defect generation at the different testing stages. Figure 3A shows the Raman spectra of a CNT sample measured at different stages of the ART process. There was no detectable D band at ~1350 cm^{-1} (used to indicate defects in CNTs) up to the end of testing, although a shift of the G band occurred because the initial strain was released after the failure of the sample (*4*).

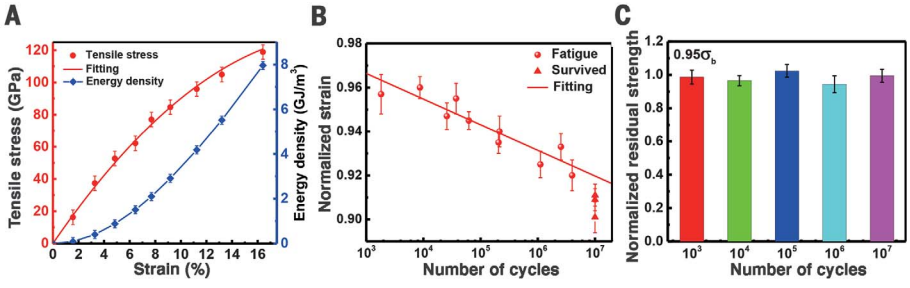


Fig. 2. Mechanical behavior of CNTs at ambient temperature. (**A**) Typical tensile stress–strain curve and energy density–strain curve for CNTs from a cantilever test. (**B**) Fatigue behavior of CNTs at ambient temperature. (**C**) Residual strength of CNTs after specific loading cycles at a normalized stress of 0.95. Error bars in (A) were calculated from the uncertainties in load measurements and cross section (diameter). Error bars in (B) denote SE of five measurements of the strain. Error bars in (C) denote SD (n = 5).

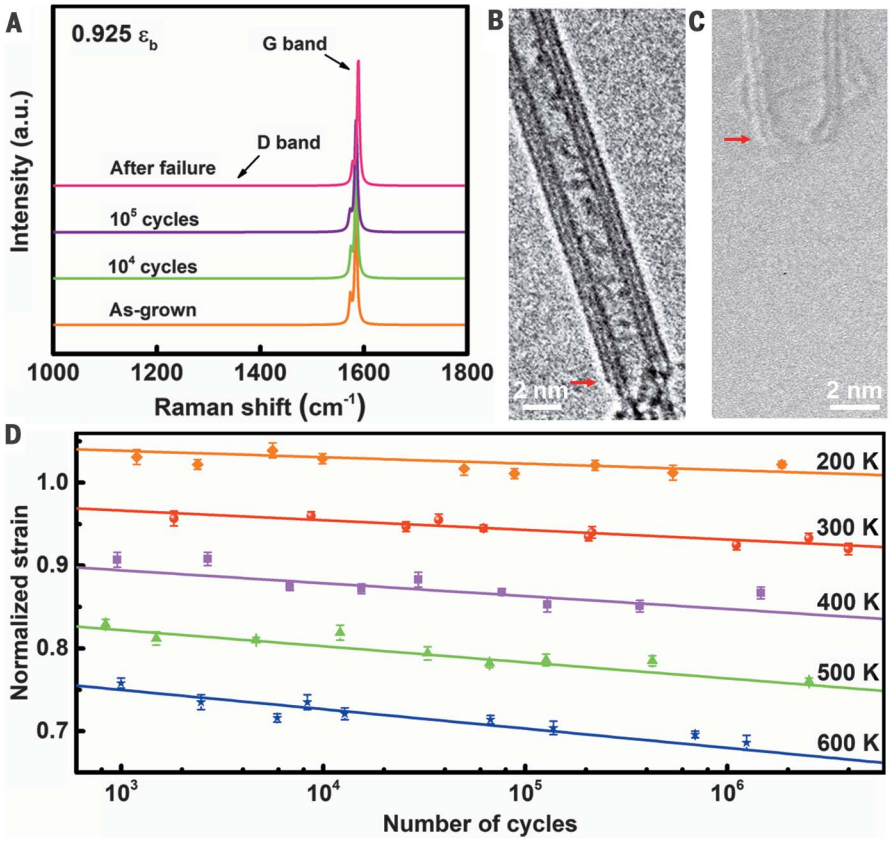


Fig. 3. Fatigue behavior of CNTs at different temperatures. (**A**) Raman spectra of an individual CNT sample after specific loading cycles. (**B** and **C**) TEM images of the flat fracture surfaces (marked by arrows) of the CNTs: (B) static tensile fracture surface; (C) fracture surface after cyclic loading. (**D**) Fatigue behavior of CNTs at different temperatures. Error bars denote SD of five measurements of strain.

All of the above evidence revealed that no detectable defects were present before and during the ART process. This showed that the time for a defect to be created tends to occupy the whole lifetime of the CNTs, which strongly suggested a clean brittle failure. The fragments after the fracture of the CNTs under static tension and cyclic loading were characterized by TEM. The flat fracture surfaces shown in Fig. 3, B and C, indicated a brittle-mode fracture (*22*–*25*).

The dependence of the fatigue behavior on the testing parameters provided further insights into the underlying mechanism. Fatigue experiments were conducted at different loading frequencies and temperatures, which showed that the lifetime of the CNTs was almost independent of the loading frequency (fig. S4) but depended on the temperature. A higher temperature resulted in a lower fatigue resistance (Fig. 3D). It was also observed that at a low temperature of 200 K, the samples exhibited a higher fatigue resistance than at ambient temperature. The same trend was also observed for the breaking strain of CNTs (fig. S5), which is consistent with previous theoretical work (*26*, *27*). This also implied a higher toughness of the CNTs at low temperatures because the modulus was demonstrated to be almost unchanged at low temperature (*26*, *27*).

The observation that the fatigue fracture of CNTs depends on the applied strain and temperature but not the loading frequency indicates a thermally activated process that was previously predicted (*23*, *28*). Our experimental data (lifetime of a CNT) were well fitted by the proposed model, which uses the time required for the first bond rotation to occur in a CNT [also called Stone-Wales (SW) isomerization or SW defect] (Fig. 2B, Fig. 3D, and text S5) (*23*). This says that the subsequent defect/crack propagation is instantaneous (*22*–*24*); that is, the time needed for it is negligible, or the fatigue failure of CNTs is global and catastrophic without progressive damage. Hence, the lifetime of a CNT is virtually dominated by the lifetime of the first defect creation.

REFERENCES AND NOTES

1. Y. Bai *et al*., *Adv. Mater.* **31**, e1800680 (2019).
2. R. H. Baughman, A. A. Zakhidov, W. A. de Heer, *Science* **297**, 787–792 (2002).
3. M. F. De Volder, S. H. Tawfick, R. H. Baughman, A. J. Hart, *Science* **339**, 535–539 (2013).
4. Y. Bai *et al*., *Nat. Nanotechnol.* **13**, 589–595 (2018).
5. C. Lee, X. Wei, J. W. Kysar, J. Hone, *Science* **321**, 385–388 (2008).
6. G.-H. Lee *et al*., *Science* **340**, 1073–1076 (2013).
7. S. Suresh, *Fatigue of Materials* (Cambridge Univ. Press, 1991).
8. R. Zhang, Y. Zhang, F. Wei, *Acc. Chem. Res.* **50**, 179–189 (2017).
9. R. Zhang *et al*., *ACS Nano* **7**, 6156–6161 (2013).
10. R. Zhang, Y. Zhang, F. Wei, *Chem. Soc. Rev.* **46**, 3661–3715 (2017).
11. W. Wu *et al*., *Nano Res.* **8**, 2721–2732 (2015).
12. Z. Zhu *et al*., *Nat. Commun.* **10**, 4467 (2019).
13. Z. Zhu *et al*., *Sci. Adv.* **2**, e1601572 (2016).
14. R. Zhang *et al*., *Nat. Commun.* **4**, 1727 (2013).
15. J. Y. Huang *et al*., *Nature* **439**, 281 (2006).
16. T. Xiao, K. Liao, *Phys. Rev. B* **66**, 153407 (2002).
17. T. Belytschko, S. P. Xiao, G. C. Schatz, R. S. Ruoff, *Phys. Rev. B* **65**, 235430 (2002).
18. B. I. Yakobson, M. P. Campbell, C. J. Brabec, J. Bernholc, *Comput. Mater. Sci.* **8**, 341–348 (1997).
19. B. I. Yakobson, C. J. Brabec, J. Berhnolc, *Phys. Rev. Lett.* **76**, 2511–2514 (1996).
20. B. Yakobson, C. Brabec, J. Bernholc, *J. Comput. Aided Mater. Des.* **3**, 173–182 (1996).
21. L. Zhu, J. Wang, F. Ding, *ACS Nano* **10**, 6410–6415 (2016).
22. M. B. Nardelli, B. I. Yakobson, J. Bernholc, *Phys. Rev. Lett.* **81**, 4656–4659 (1998).
23. T. Dumitrica, M. Hua, B. I. Yakobson, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6105–6109 (2006).
24. B. Yakobson, *Appl. Phys. Lett.* **72**, 918–920 (1998).
25. T. Dumitrică, T. Belytschko, B. I. Yakobson, *J. Chem. Phys.* **118**, 9485–9488 (2003).
26. M. Wen, B. An, S. Fukuyama, K. Yokogawa, A. H. W. Ngan, *Carbon* **47**, 2070–2076 (2009).
27. C. Wei, K. Cho, D. Srivastava, *Phys. Rev. B* **67**, 115407 (2003).
28. G. G. Samsonidze, G. G. Samsonidze, B. I. Yakobson, *Phys. Rev. Lett.* **88**, 065501 (2002).

ACKNOWLEDGMENTS

We thank C. Chen for advice on the manuscript, D. Wang for English editing, and M. Liu, X. Ye, C. Pei, Y. Hu, E. Gao, H. Wang, B. Li, X. Chen, Y. Wang, C. Li, T. Xiao, and Z. Xu for helpful discussion. **Funding:** Supported by Foundation for the National Basic Research Program of China grant 2016YFA0200102, National Natural Science Foundation of China grant 21636005, Beijing Municipal Science and Technology Commission grant D141100000614001, and NSFC grants 51872156, 11872035, and 11632010. **Author contributions:** F.W., R.Z., and Y.B. conceived the main idea of the project; R.Z. co-directed the project; Y.B. designed and performed the experiments and analyzed data; H.Y. participated in the synthesis, characterization, and fatigue test of CNTs; J.W., S.W., and Z.X. participated in the theoretical analysis; B.S., S.S., H.W., and X.L. participated in data analysis and manuscript preparation; and H.W. participated in equipping the ART system with a heating/cooling stage. All authors participated in the project discussions and production of the final manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data are available in the manuscript and supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1104/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S5
Movie S1
References (*29*–*62*)

25 June 2019; accepted 20 July 2020
10.1126/science.aay5220

CORONAVIRUS

Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions

Sheikh Taslim Ali[1]*, Lin Wang[2,3]*, Eric H. Y. Lau[1]*, Xiao-Ke Xu[4], Zhanwei Du[5], Ye Wu[6,7], Gabriel M. Leung[1], Benjamin J. Cowling[1]†

Studies of novel coronavirus disease 2019 (COVID-19), which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), have reported varying estimates of epidemiological parameters, including serial interval distributions—i.e., the time between illness onset in successive cases in a transmission chain—and reproduction numbers. By compiling a line-list database of transmission pairs in mainland China, we show that mean serial intervals of COVID-19 shortened substantially from 7.8 to 2.6 days within a month (9 January to 13 February 2020). This change was driven by enhanced nonpharmaceutical interventions, particularly case isolation. We also show that using real-time estimation of serial intervals allowing for variation over time provides more accurate estimates of reproduction numbers than using conventionally fixed serial interval distributions. These findings could improve our ability to assess transmission dynamics, forecast future incidence, and estimate the impact of control measures.

In December 2019, a novel coronavirus disease [coronavirus disease 2019 (COVID-19)], caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was first reported in Wuhan, China. It has since spread to more than 212 countries, causing more than 10 million confirmed cases and 500,000 deaths worldwide as of 30 June 2020 (*1*). Recent studies have suggested that several demographic and social factors can influence the transmission of COVID-19, including age- and gender-related differences in infection risk (*2*–*4*), reduced risk of infection as a result of intensive nonpharmaceutical interventions (NPIs) (e.g., isolation and social distancing) (*5*–*7*), and abrupt changes in social mixing patterns because of lockdowns and confinement (*8*–*10*). Serial interval, defined as the duration between the symptom-onset time of the infector and that of the infectee, is an essential metric for estimating many other key epidemiological parameters (e.g., reproduction number, generation time, and attack rate), which are used in turn to predict disease trends and health care demands (*11*). In early studies, before the availability of specific data on COVID-19, the serial interval distribution of COVID-19 was assumed to be similar to those of severe acute respiratory syndrome (SARS) or Middle East respiratory syndrome (MERS), with a mean >8 days (*12*, *13*). Once specific data became available on COVID-19 transmission pairs, several studies examined the serial interval distribution of COVID-19 in different locations, with estimates of the mean serial interval varying from 3.1 to 7.5 days (*6*, *14*–*21*). All of these studies have assumed that the timing of transmission events can be described by a single, stable distribution of serial intervals at different stages of an epidemic.

In fact, the serial interval depends on the incubation period, the profile of infectiousness after infection, and the variation in contact structure of the population (as explained in fig. S1) (*22*). The incubation period describes the biological process of disease progression and tends to follow a more similar distribution from one location to another, with minor variations resulting from social or cultural differences in how symptoms are perceived or reported. However, the profile of infectiousness over time can vary because of human behavior. Changes in contact patterns and the use of public health measures can reshape the timing of infection events by limiting successful contacts overall (e.g., social distancing) or after illness onset (e.g., case isolation). Interventions such as the isolation of confirmed and suspected cases, suspension of intra- and intercity travel, and different forms of social distancing were widely implemented in different Chinese cities. This provides an opportunity to study the temporal changes in the serial interval distribution and its association with NPIs. Here, we show that variation in the serial interval can occur and has important implications for the assessment of transmission dynamics and the impact of control measures.

We compiled a database of 1407 COVID-19 transmission pairs, in which symptom-onset dates and social relationships were available for both the infector and infectee of 677 transmission pairs [see table S1 for entire database (*23*) and supplementary materials for details]. Household and nonhousehold transmissions were identified on the basis of the information on social relationships (e.g., familial members of the same household, nonhousehold relatives, colleagues, classmates, friends, and other face-to-face contacts). The data were reconstructed from the publicly available reports of 9120 confirmed COVID-19 cases reported by 27 provincial and 264 urban health commissions in China outside Hubei province. Data from Hubei province were excluded because there was less reliable information on chains of transmission during the widespread community circulation of COVID-19; outside Hubei province, it was more straightforward to link connected cases and derive serial intervals. We focused on 677 transmission pairs with infectors having developed symptoms from 9 January through 13 February 2020. This 36-day period spans a series of key interventions related to the evolving epidemiology and transmission dynamics of COVID-19 in mainland China (*24*–*26*).

We first calculated the number of transmission pairs in our database by the onset dates of infectors (fig. S3). Because many infectors (339) developed symptoms during 23 to 29 January 2020, we defined this 1-week period as the peak week, the previous 14-day period (9 to 22 January 2020) as the prepeak period, and the following 15-day period (30 January to 13 February 2020) as the postpeak period. We computed the serial interval as the number of days between the symptom-onset date of the infector and that of the infectee for each transmission pair. Empirical serial interval distributions for transmission pairs, counting from symptom onsets of the infectors during each period, indicate that the serial intervals shortened over time (Fig. 1A).

We estimated the serial interval distribution during each nonoverlapping period by fitting a normal distribution to the corresponding serial intervals data (supplementary materials). Analysis of all 677 transmission pairs revealed that the serial interval distribution had a mean of 5.1 [95% credibility interval (CrI): 4.7, 5.5] days and a standard deviation (SD) of 5.3 (95% CrI: 5.0, 5.6) days (table S2) overall, which is consistent with other recent

[1]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China. [2]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK. [3]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, UMR2000, CNRS, Paris 75015, France. [4]College of Information and Communication Engineering, Dalian Minzu University, Dalian 116600, China. [5]Department of Integrative Biology, University of Texas at Austin, Austin, TX 78705, USA. [6]School of Journalism and Communication, Beijing Normal University, Beijing 100875, China. [7]Computational Communication Research Center, Beijing Normal University, Zhuhai 519087, China.
*These authors contributed equally to this work.
†Corresponding author. Email: bcowling@hku.hk

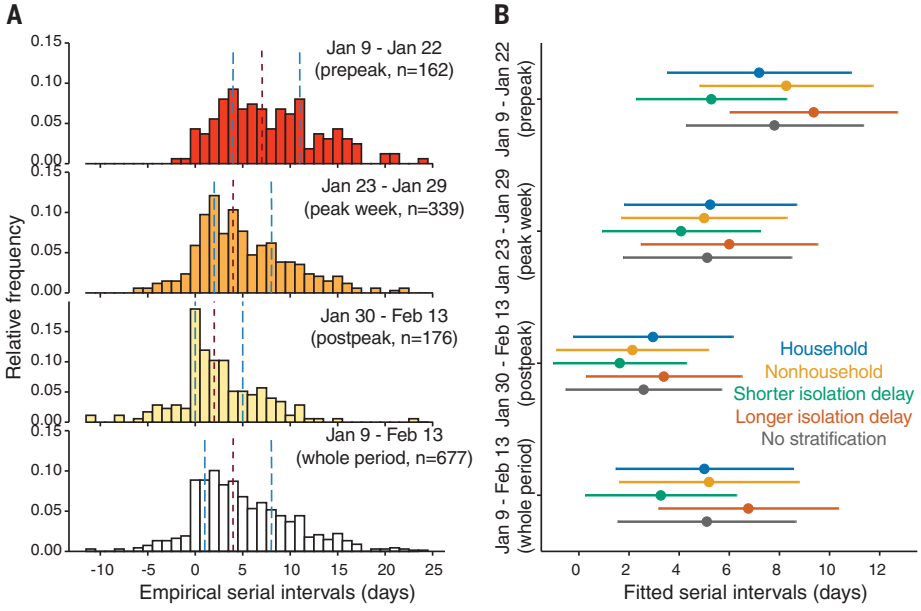


Fig. 1. Serial intervals of SARS-CoV-2 substantially shortened over time in mainland China. (**A**) Empirical serial interval distributions. From top to bottom, transmission pairs were analyzed by selecting infectors who developed symptoms during 9 to 22 January 2020 (prepeak); 23 to 29 January 2020 (peak week); 30 January to 13 February 2020 (postpeak); and 9 January to 13 February 2020 (whole period), respectively. In each panel, vertical dashed lines in red and blue colors indicate the median and interquartile range (IQR), respectively. (**B**) Estimated serial interval distributions by fitting a normal distribution using MCMC. From top to bottom, each group of bars corresponds to the transmission pairs with infectors who developed symptoms during the prepeak (162 pairs), peak week (339 pairs), postpeak (176 pairs), and whole 3 6-day period (677 pairs), respectively. Colored dots and bars correspond to the transmission pairs within households (blue), outside households (yellow), with isolation delays shorter than the median isolation delay of each period (green), and with isolation delays longer than the median isolation delay of each period (orange), respectively. Dark gray bars correspond to transmission pairs with no stratification. Dots and bars indicate the estimated median and IQR, respectively.

studies (*16*, *21*, *27*). However, fitting to data of nonoverlapping periods of time revealed considerable variation in serial interval distributions (Fig. 1B). Before the peak, the mean and SD of serial intervals were estimated to be 7.8 (7.0, 8.6) days and 5.2 (4.7, 5.9) days, respectively. During the peak, the mean and SD reduced to 5.1 (4.6, 5.7) days and 5.0 (4.6, 5.4) days, respectively. After the peak, these estimates further shortened to 2.6 (1.9, 3.2) days and 4.6 (4.2, 5.1) days, respectively (table S2).

Next, we examined the real-time change in serial intervals by using a series of running time windows with fixed lengths of 10, 14, or 18 days (fig. S10). In contrast to the use of a constant distribution of serial intervals, our analysis suggests that serial intervals were gradually shortened over the study period (Fig. 2A), which is robust to alternative specifications of time windows (fig. S10). By fitting the transmission pairs data for each running time window by Markov chain Monte Carlo (MCMC) (Fig. 2A and table S3), we estimated that during the first 14-day period (9 to 22 January 2020), the serial intervals were longer on average [mean: 7.8 (95% CrI: 7.0, 8.6) days; SD: 5.2 (95% CrI: 4.7, 5.9) days]; whereas during the last 14 days (30 January to 13 February 2020), the serial intervals were much shorter on average [mean: 2.2 (1.5, 2.9) days; SD: 4.6 (4.1, 5.1) days]. Notably, the mean serial intervals were shortened by more than a factor of 3 over the 36-day period.

The transmission pair data also contain information for age, sex, household, and isolation delay (i.e., time duration from symptom onset to isolation) for most infectors. This allows for a granular stratification. Using either nonoverlapping or running time windows for data stratified by each of these factors, we find the same pattern of shorter serial intervals over time (Figs. 1B and 2A and tables S2 and S3). Therefore, we termed this changing serial interval the effective serial interval, which accounts for temporal changes caused by its potential driving factors. Notably, the length of effective serial intervals is positively associated with the length of isolation delay (Fig. 2A; figs. S5, S6, and S9; and tables S3 and S4), which accounts for the decreasing isolation delay over time (fig. S2). Therefore, early isolation (shorter than the median isolation delay) translates into shorter serial intervals [mean: 3.3 (2.7, 3.8) days; SD: 4.5 (4.1, 4.9) days], and delayed isolation (longer than the median isolation delay) is associated with longer serial intervals [mean: 6.8 (6.2, 7.3) days; SD: 5.3 (4.9, 5.7) days] (table S2). Stratification by age, gender, or household shows no clear differences in serial interval estimates. Our findings are robust to using alternative distributions (e.g., Gumbel distribution) for model fitting (fig. S11) and the infector-based approach (fig. S14).

Our probabilistic, individual-based simulated and regression models confirm that serial intervals are positively associated with isolation delay (section 5, supplementary materials). We found that the serial intervals become shorter on the basis of how much faster the infectors are isolated, regardless of when an infector starts to be infectious before symptom onset (fig. S5). In an individual-based simulation model with a mean generation time of 7.8 days, the simulated mean serial intervals are reduced from ~8.0 to ~1.2 days when the isolation delay is reduced from 10 to 0 days. We found through regression models that up to 51.5% of the variability in daily empirical serial interval can be explained by isolation delay and further improved by other NPI factors, which explain an additional 15.6 to 16.7% of the variability (table S5).

In practice, the time-varying serial interval may affect the estimation of epidemic parameters, including the transmissibility. The real-time transmissibility of an infectious disease is often characterized by the instantaneous reproduction number (R_t), which is defined as the expected number of secondary infections caused by an infector on day t. The pathogen spreads when $R_t > 1$ and is under control when $R_t < 1$. To examine the effect of serial intervals on R_t, we first obtained the daily number of cases on the basis of the onset dates of infectors and infectees among the 1407 transmission pairs (Fig. 2, B to D). By applying the statistical method developed by Cori *et al.* (*28*), we estimated R_t for each day between 20 January and 13 February 2020. We noticed substantial differences in estimates of R_t between using a single stable serial interval distribution and time-varying effective serial interval distributions. The magnitude of this difference is more prominent during the prepeak and postpeak periods than it is during the peak week when $R_t \approx 1$ (Fig. 2, B to D).

We observed that the serial interval for COVID-19 in mainland China was shortened by more than a factor of 3 in the 36 days between 9 January and 13 February 2020. This reduction was driven by intensive NPIs, particularly the reduction of the isolation delay period. Isolation of an infector 1 day earlier is expected to reduce the mean serial interval by 0.7 days. Thus, the serial interval was shortened by >3 days if infectors were rapidly isolated (Figs. 1B and 2A and tables S2 and S3). This is consistent with advocating isolation of cases

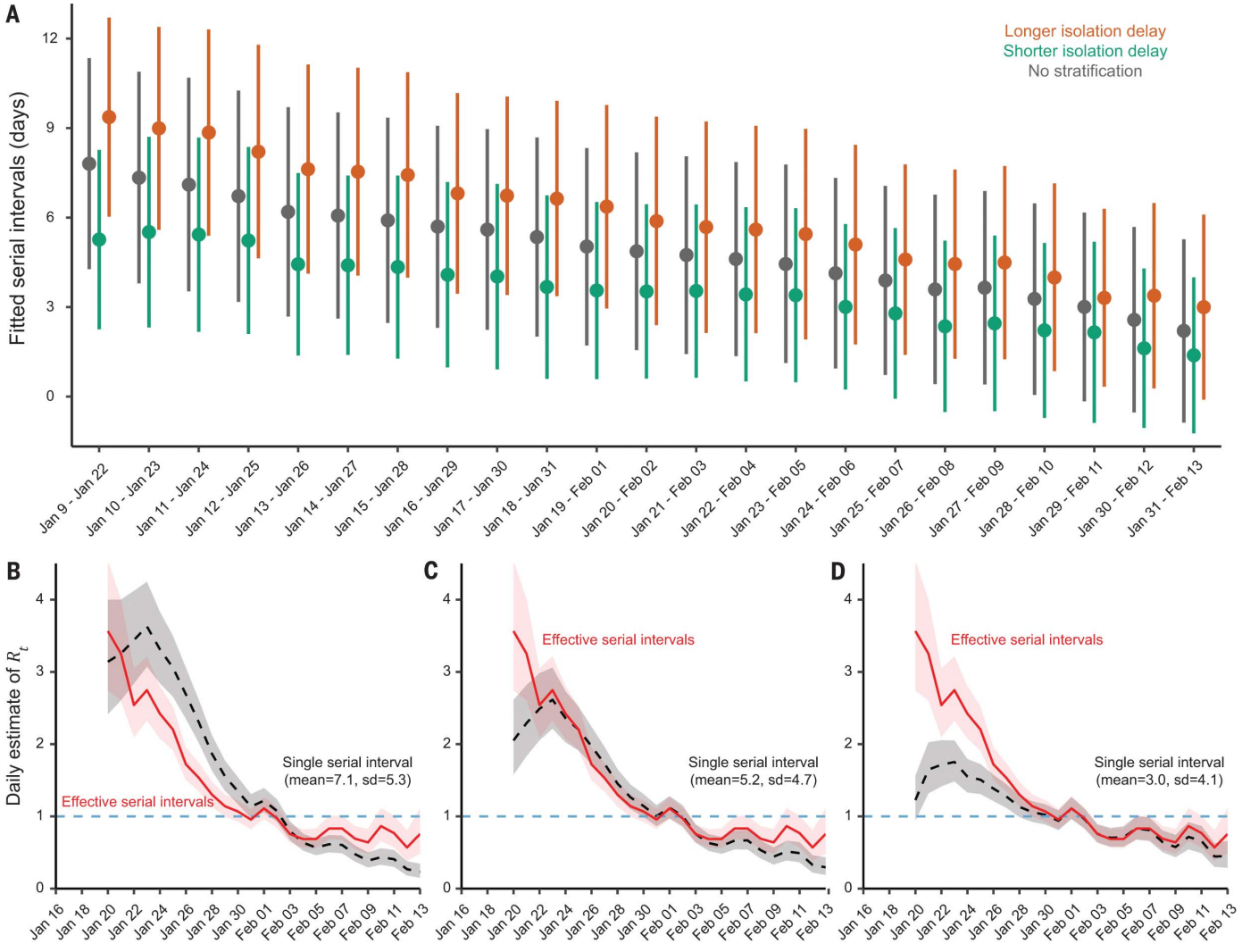


Fig. 2. Real-time effective serial intervals and instantaneous reproduction number R_t. (**A**) Estimated serial interval distribution for each 14-day running time window. Dark gray color indicates fitting data with no stratification, whereas green and orange indicate fitting data with isolation delay shorter and longer, respectively, than the median isolation delay of each running time window. Dots and bars indicate the estimated median and IQR, respectively. (**B** to **D**) Daily estimates of R_t by using real-time effective serial interval distributions [as in (A)] versus using a single fixed serial interval distribution. Red curves and light pink shaded regions indicate the median and 95% CrI, respectively, of daily R_t estimated using real-time effective serial interval distributions. Black dashed curves and light gray shaded regions indicate the median and 95% CrI, respectively, of daily R_t estimated using a single serial interval distribution fixed, with a mean of 7.1 and SD of 5.3 days in (B), a mean of 5.2 and SD of 4.7 days in (C), and a mean of 3.0 and SD of 4.1 days in (D).

and quarantining contacts within 1 day from symptom onset, which has been estimated to reduce COVID-19 transmission by 60% (*8*). We have not identified any substantial effects of gender or age of infectors on serial interval, but the NPIs were found to be significant for the transmission in communities rather than in households (table S5). Other studies (*15*, *20*) have estimated that the infectiousness of COVID-19 is greater at symptom onset. Although a short serial interval indicates that a substantial proportion of transmission events have occurred by the time symptoms are apparent (*14*), because of prolonged viral shedding (*14*, *29*, *30*) case isolation is still likely to reduce further transmission. Changes in the serial interval can therefore indicate effective implementation of specific transmission-reduction measures.

There are some limitations to our work. First, it is possible that there was recall bias on the onset of first symptoms in the line-list data; however, given the centralized pandemic response in mainland China, we expected that recall bias would not affect our main conclusions (figs. S12 and S13). Second, other factors may have influenced the reduction of effective serial intervals, as we can only explain up to 72% of the variance in observed serial intervals. Finally, our current transmission pair data do not contain variables about the potential exposure window of each case, which do not allow further inferences on the transmission potential.

Our results indicate that caution is needed when attempting to generalize estimates of the serial interval distribution to other places or to other periods of time in the same place, for example when estimating instantaneous reproductive numbers (Fig. 2, B to D). The real-time metric of effective serial intervals indicates that transmission models also need to account for the temporal variation in serial intervals as an epidemic proceeds. Effective serial intervals may provide better measurements of instantaneous transmissibility (R_t)—because they include the effects of possible drivers of transmission—and could be helpful

to policy-makers because they offer real-time information on the impact of public health measures.

REFERENCES AND NOTES

1. World Health Organization (WHO), "Coronavirus disease 2019 (COVID-19): Situation report – 162" (WHO, 2020); www.who.int/docs/default-source/coronaviruse/situation-reports/20200629-covid-19-sitrep-161.pdf?sfvrsn=74fde64e_2.
2. M. U. G. Kraemer *et al*., *Science* **368**, 493–497 (2020).
3. C. Wenham, J. Smith, R. Morgan, Gender and COVID-19 Working Group, *Lancet* **395**, 846–848 (2020).
4. J. M. Jin *et al*., *Front. Public Health* **8**, 152 (2020).
5. J. Hellewell *et al*., *Lancet Glob. Health* **8**, e488–e496 (2020).
6. L. Ferretti *et al*., *Science* **368**, eabb6936 (2020).
7. R. Armitage, L. B. Nellums, *Lancet Public Health* **5**, e256 (2020).
8. R. M. Anderson, H. Heesterbeek, D. Klinkenberg, T. D. Hollingsworth, *Lancet* **395**, 931–934 (2020).
9. J. R. Koo *et al*., *Lancet Infect. Dis.* **20**, 678–688 (2020).
10. K. Prem *et al*., *Lancet Public Health* **5**, e261–e270 (2020).
11. M. A. Vink, M. C. Bootsma, J. Wallinga, *Am. J. Epidemiol.* **180**, 865–875 (2014).
12. J. T. Wu, K. Leung, G. M. Leung, *Lancet* **395**, 689–697 (2020).
13. M. Chinazzi *et al*., *Science* **368**, 395–400 (2020).
14. H. Nishiura, N. M. Linton, A. R. Akhmetzhanov, *Int. J. Infect. Dis.* **93**, 284–286 (2020).
15. H. Y. Cheng *et al*., *JAMA Intern. Med.* 10.1001/jamainternmed.2020.2020 (2020).
16. Z. Du *et al*., *Emerg. Infect. Dis.* **26**, 1341–1343 (2020).
17. Q. Li *et al*., *N. Engl. J. Med.* **382**, 1199–1207 (2020).
18. J. M. Griffin *et al*., medRxiv 2020.05.08.20095075 [Preprint]. 11 May 2020. https://doi.org/10.1101/2020.05.08.20095075.
19. S. Ma *et al*., medRxiv 2020.03.21.20040329 [Preprint]. 24 March 2020. https://doi.org/10.1101/2020.03.21.20040329.
20. X. He *et al*., *Nat. Med.* **26**, 672–675 (2020).
21. Q. Bi *et al*., *Lancet Infect. Dis.* **20**, 911–919 (2020).
22. J. Zhang *et al*., *Science* **368**, 1481–1486 (2020).
23. Lin, PDGLin/COVID19_EffSerialInterval_NPI: Serial interval of SARS-CoV-2 was shortened over time by non-pharmaceutical interventions, version v1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.3940300.
24. H. Tian *et al*., *Science* **368**, 638–642 (2020).
25. K. Leung, J. T. Wu, D. Liu, G. M. Leung, *Lancet* **395**, 1382–1393 (2020).
26. A. Pan *et al*., *JAMA* **323**, 1915–1923 (2020).
27. J. Zhang *et al*., *Lancet Infect. Dis.* **20**, 793–802 (2020).
28. A. Cori, N. M. Ferguson, C. Fraser, S. Cauchemez, *Am. J. Epidemiol.* **178**, 1505–1512 (2013).
29. L. Zou *et al*., *N. Engl. J. Med.* **382**, 1177–1179 (2020).
30. Y. Pan, D. Zhang, P. Yang, L. L. M. Poon, Q. Wang, *Lancet Infect. Dis.* **20**, 411–412 (2020).

ACKNOWLEDGMENTS

We thank all the health workers and volunteers who collected data throughout the COVID-19 outbreak. We thank H. Salje, S. Cauchemez, L. A. Meyers, J. Paireau, Q. Bi, B. Yang, X. Liu, and L. Hu for discussions. **Funding:** We acknowledge financial support from the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region, China (grant no. COVID190118); the Investissement d'Avenir program, the Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases program (grant no. ANR-10-LABX-62-IBEID); the European Research Council (grant no. 804744); the European Union's Horizon 2020 research and innovation program under grant agreement no. 101003589 (RECOVER); the National Institutes of Health (no. U01 GM087719); the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Land and Resources, China (no. KF-2019-04-034); the National Natural Science Foundation of China (nos. 61773091, 11875005, 61976025, and 11975025); and a University of Cambridge COVID-19 Rapid Response Grant. S.T.A. acknowledges the research computing facilities and advisory services offered by Information Technology Services, The University of Hong Kong. L.W. acknowledges the computational and storage services provided by the IT department at the Institut Pasteur. **Author contributions:** L.W., S.T.A., E.H.Y.L., and B.J.C. conceived the study, designed the statistical and modeling methods, conducted analyses, interpreted results, and wrote and revised the manuscript. X.-K.X., Z.D., and Y.W. collected and compiled data, interpreted results, and revised the manuscript. G.M.L. supervised the study, interpreted results, and revised the manuscript. **Competing interests:** B.J.C. reports honoraria from Sanofi Pasteur and Roche. The authors report no other potential conflicts of interest. **Data and materials availability:** All data and codes are available in the main text, the supplementary materials, or Zenodo (*23*). This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1106/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S5
References (*31*–*49*)
MDAR Reproducibility Checklist

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

20 May 2020; accepted 13 July 2020
Published online 21 July 2020
10.1126/science.abc9004

COSMOCHEMISTRY

Earth's water may have been inherited from material similar to enstatite chondrite meteorites

Laurette Piani[1]*, Yves Marrocchi[1], Thomas Rigaudier[1], Lionel G. Vacher[1,2], Dorian Thomassin[1], Bernard Marty[1]

The origin of Earth's water remains unknown. Enstatite chondrite (EC) meteorites have similar isotopic composition to terrestrial rocks and thus may be representative of the material that formed Earth. ECs are presumed to be devoid of water because they formed in the inner Solar System. Earth's water is therefore generally attributed to the late addition of a small fraction of hydrated materials, such as carbonaceous chondrite meteorites, which originated in the outer Solar System where water was more abundant. We show that EC meteorites contain sufficient hydrogen to have delivered to Earth at least three times the mass of water in its oceans. EC hydrogen and nitrogen isotopic compositions match those of Earth's mantle, so EC-like asteroids might have contributed these volatile elements to Earth's crust and mantle.

The origin of Earth's water is debated. The isotopic composition of Earth suggests that it is composed of material from the inner Solar System, such as enstatite chondrite (EC) meteorites (*1*–*3*). The inner Solar System was too warm to have retained water ice, so terrestrial water is thought to have been supplied by hydrated materials that originally formed in the outer Solar System before migrating inward (*4*, *5*).

Hydrogen isotopic compositions are conventionally expressed as $\delta D \equiv [(D/H)_{sample}/(D/H)_{SMOW} - 1] \times 1000$, where D/H is the ratio of ^2H (deuterium) to ^1H (protium) and standard mean ocean water (SMOW) is the mean value of Earth's oceans. δD varies greatly among Solar System objects. The bulk protosolar nebula had δD = −865 per mil (‰), estimated from the solar wind and remnant H_2 in the atmospheres of the giant planets (*6*). Compared with this protosolar value, all other materials in the Solar System are enriched in deuterium, with inner Solar System objects having intermediate values [e.g., δD ≡ 0‰ in terrestrial oceans and −165 to +800‰ in asteroids, as recorded in primitive meteorites (*7*)]. Outer Solar System objects, such as comets, are even more enriched in deuterium [e.g., up to δD = 2400‰ in comet 67P/Churyumov-Gerasimenko (*8*)]. Among meteorites, water-rich carbonaceous chondrites (CCs) of Ivuna-type (CI) and Mighei-type (CM) are potential sources of terrestrial water because their δD values are distributed around a value coincident with Earth's oceans—i.e., δD ≡ 0‰ (*9*). Dynamical simulations suggest that C-type (carbonaceous) asteroids, whose optical properties suggest a link to water-rich CCs, could have been scattered into the inner Solar System from beyond the orbit of Jupiter (*10*), potentially delivering water to Earth during or after its main phase of accretion, in the first few million years of Solar System formation.

An alternative possibility is that EC-like materials—which have similar isotopic abundances to those of Earth for elements including oxygen, titanium, calcium, and molybdenum (*3*)—could contain enough hydrogen with the appropriate isotopic composition to provide water to the growing proto-Earth. Efforts to test this scenario have been hampered by difficulties in measuring H concentrations and isotopic ratios in ECs, owing to their assumed low H abundances and the potential for terrestrial contamination (*7*).

We measured H abundances and isotopic compositions for a suite of 13 EC meteorites that spans the full range of thermal metamorphism degrees [from petrologic type EH3 (the least metamorphosed) to EH6 (those most altered by thermal metamorphism on their parent asteroid)]. We covered this range to evaluate potential effects of parent-body processing, which could have altered the original water content and H isotopic composition (*11*). We also analyzed one aubrite meteorite (Norton County), which represents an evolved planetary body of enstatite meteorite composition for which metal-silicate segregation (differentiation) had occurred (*11*). Hydrogen measurements were performed at the Centre de Recherches Pétrographiques et Géochimiques (CRPG, Nancy, France), either on bulk samples with an elemental analyzer (EA) coupled to a stable

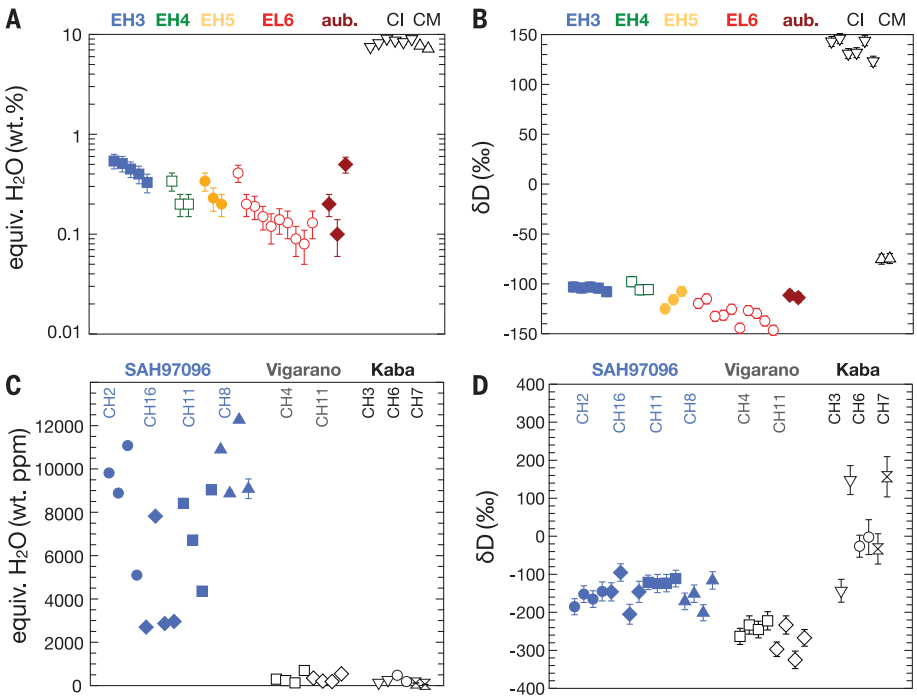


Fig. 1. Hydrogen contents and isotopic compositions of ECs. (**A**) Bulk-rock hydrogen content, reported as water equivalent by weight on a logarithmic scale. (**B**) Bulk-rock deuterium abundance, expressed as an offset from standard mean ocean water. (**C**) Mesostasis hydrogen content. (**D**) Mesostasis deuterium abundance. The CM- and CI-type CCs Orgueil, Alais, and Murchison, as well as the aubrite Norton County (aub.), are included in (A) and (B) for comparison. Data for the primitive EC Sahara 97096 (SAH97096; EH3.1-3.4) and the two CV-type CCs Vigarano and Kaba are shown in (C) and (D) for comparison (*11*). Different symbols in (C) and (D) indicate different chondrules (labeled CH). Error bars are 2σ and are smaller than the symbol size in (B). The data are listed in tables S2 and S4.

[1]Centre de Recherches Pétrographiques et Géochimiques (CRPG), Centre National de Recherche Scientifique (CNRS)–Université de Lorraine, Vandoeuvre-les-Nancy, F-54500, France. [2]Department of Physics, Washington University in St. Louis, St. Louis, MO 63130, USA.
*Corresponding author. Email: laurette.piani@univ-lorraine.fr

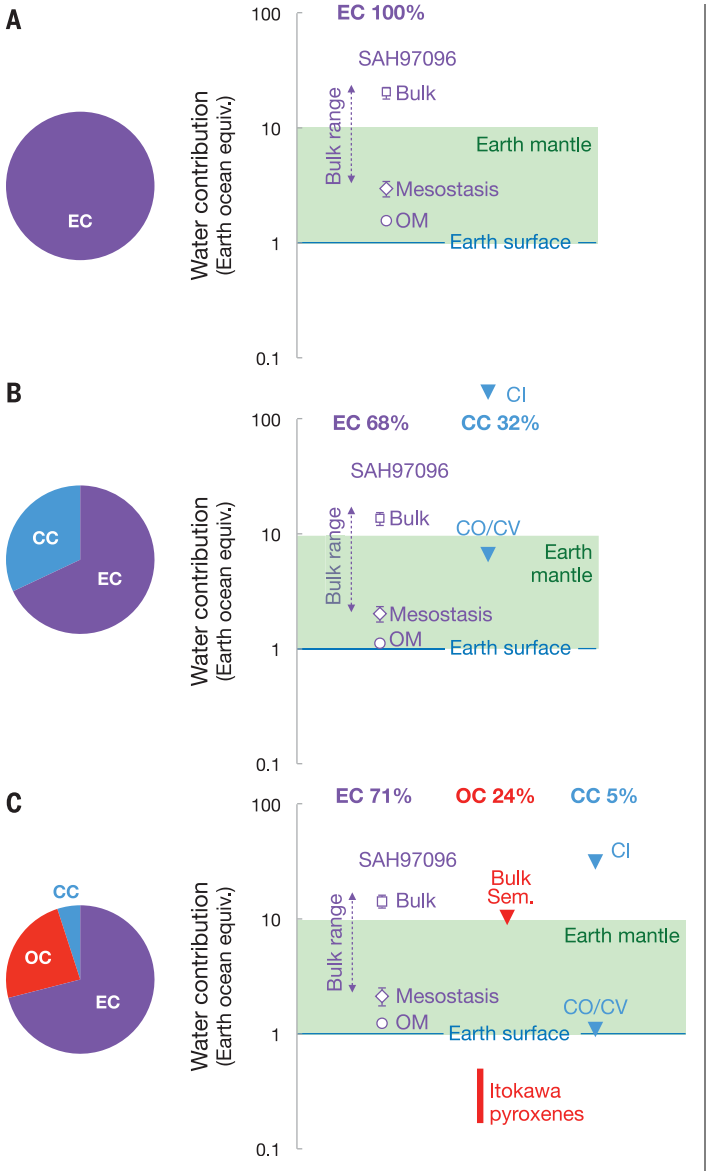


Fig. 2. Contributions of accreting materials to Earth's hydrogen budget in three different models. Values are expressed in equivalent ocean masses, where one ocean = 1.4×10^{21} kg H_2O. (**A**) A model Earth comprising only ECs (*1*). (**B**) A model Earth comprising 68% ECs and 32% CCs (*2*). (**C**) A model Earth comprising 71% ECs, 24% OCs, and 5% CCs [the most extreme case from (*3*)]. The consistency of such models for the isotopic composition of the mantle is discussed in (*11*). SAH97096, Sahara 97096 (EH3); Bulk, bulk rock; OM, organic matter; Sem., Semarkona (LL3.0). CCs are divided into water-rich (CI-type) and water-poor (CV- and CO-type) groups. Water contributions from ECs are shown in purple, from OCs in red, and from CCs in blue. The water content of Earth's mantle (green band) is between 1 and ≥10 ocean masses (*26*). We consider 10 ocean masses in the mantle as a maximum value. Regardless of the model, the water provided by ECs alone lies above the maximum value of 10 ocean masses and thus substantially contributes to Earth's water budget. Data are from (*5*, *7*, *11*, *21*) and this study (tables S2 and S4). Error bars are 2σ.

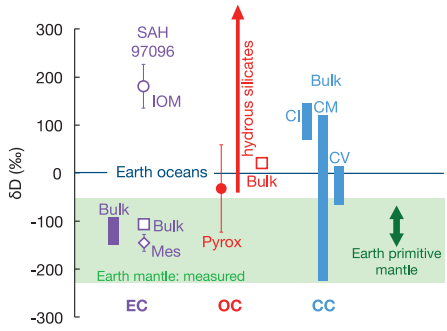


Fig. 3. Hydrogen isotopic compositions of meteorites compared with that of Earth's mantle and oceans. Data are shown for ECs (bulk rock and mesostasis) (purple); OC pyroxenes and hydrous minerals (red); and CCs (bulk rock and CI, CM, and CV types) (blue). Mes, mesostasis; Pyrox, pyroxene. Meteorite data are from (*7*, *12*, *21*, *31*) and this study (tables S2 and S4). Values for Earth's mantle (δD = −220 to −20‰) (green shaded area) are from (*26*), with the most D-depleted values reported in (*28*). Primitive mantle δD values range from−130 to −75‰ (*32*, *33*). Error bars are 2σ.

isotope ratio mass spectrometer (IRMS) or at the micrometer scale by secondary ion mass spectrometry (SIMS) (*11*).

We found that the ECs have bulk hydrogen contents (reported as water equivalents) ranging from 0.08 to 0.54 wt % H_2O (Fig. 1A and table S2). The EH3 meteorites are generally richer in H (0.44 ± 0.04 wt % H_2O; all uncertainties are one standard deviation unless otherwise noted) than the more metamorphosed groups EH4 (0.2 ± 0.1 wt % H_2O), EH5 (0.3 ± 0.1 wt % H_2O), and EL6 (0.3 ± 0.3 wt % H_2O). The Norton County aubrite is 0.3 ± 0.2 wt % H_2O. These EC water contents are well below those measured in the water-rich CCs Orgueil (CI), Alais (CI), and Murchison (CM), which are 7.2 to 9.1 wt % H_2O (Fig. 1A and table S2). However, all analyzed ECs contain hydrogen concentrations above the EA-IRMS detection limit [~0.05 wt % H_2O (*11*)]. The least-altered EH chondrites (EH3 and EH4) have δD values averaging δD = −103 ± 3‰, systematically below the current value of Earth's oceans, whereas metamorphosed ECs (EH5 and EL6) and the aubrite have even lower values, averaging δD = −127 ± 15‰ (Fig. 1B and table S2).

If the hydrogen contents had been entirely due to atmospheric contamination during sample preparation, constant D/H ratios would be observed despite variable H concentrations. Instead, our data follow a negative trend in a diagram of δD versus 1/H (fig. S5), indicating possible loss of D-enriched hydrogen-bearing organic molecules (*12*) during thermal metamorphism within the EC parent asteroid(s). In our least-metamorphosed sample, the pristine EH3 meteorite Sahara 97096 [EH3.1-3.4 (*13*); a fragment of the same meteorite as Sahara 97116, see table S2], only a thin (≈500 μm) surface layer of terrestrial alteration is visible (*11*). Its inner portions remain unaffected, as attested by the presence of unaltered grains of oldhamite (CaS), a highly hydrophilic mineral (*14*, *15*). The fresh interior of Sahara 97096 has lower (and constant) H contents (0.5 ± 0.1 wt % H_2O) than the altered parts (3.0 ± 1.1 wt % H_2O), and their H isotopic compositions are also different (fresh interior δD = −103.6 ± 0.7‰, altered surface δD = −136.3 ± 3.9‰; fig. S4 and table S2). Similar features have been observed in martian meteorites subjected to desert alteration, with the H content and isotopic characteristics being modified only near the surface (i.e., <2 mm) (*16*). In addition, the hydrogen contents and D/H ratios measured for the two meteorite finds (Sahara 97096 and Yamato 791790) are very close to the ones measured for the Indarch chondrite that was collected after its observed fall and are consistent with a metamorphic evolution of the H signatures in ECs (table S2 and fig. S5B). These observations suggest that the H contents of ECs were established during the early Solar System and have experienced minimal modifications on Earth.

To confirm the indigenous nature of H in the interiors of ECs, we used SIMS to analyze the H contents and isotopic compositions of the glass fraction (mesostasis) in four chondrules—submillimeter silicate spherules that correspond to the main mineral assemblages in ECs—of Sahara 97096, the least-metamorphosed EC in our sample (*11*, *12*). For comparison, we performed

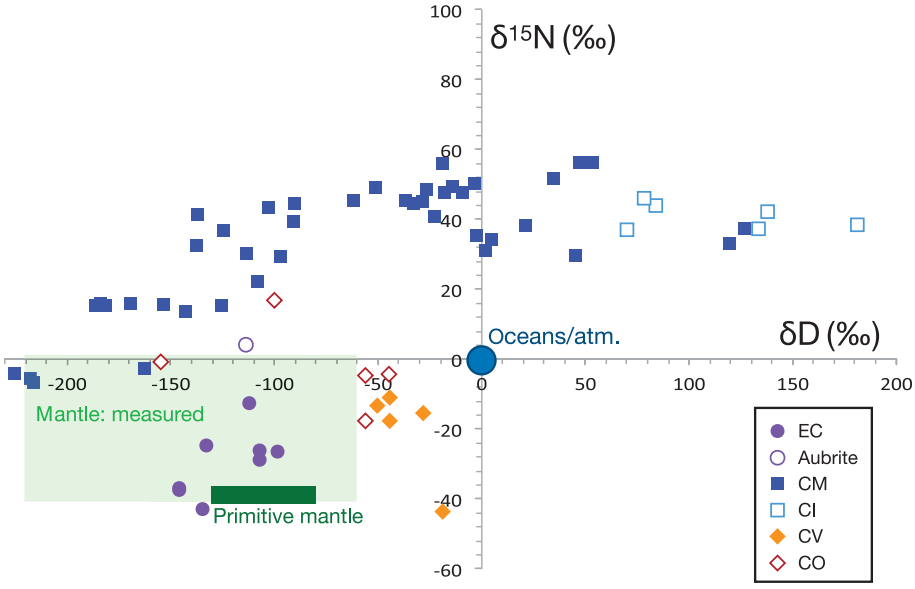


Fig. 4. Hydrogen and nitrogen isotopic compositions of meteorites compared with that of Earth's mantle and surface. Meteorite bulk-rock values are shown for ECs and CM-, CI-, CV-, and CO-type CCs; data are from table S2 and (*7*, *34*, *35*). Mantle δD values are from (*26*, *28*). Primitive mantle δD values are from (*32*) and (*33*) and $\delta^{15}N$ values are from the most negative values (−40‰) in (*35*). atm., atmosphere.

similar measurements in two CV chondrites (Kaba and Vigarano) that have the same degree of metamorphism (*17*). The glassy mesostasis of Sahara 97096 contains between 2700 and 12,300 wt parts per million (ppm) H_2O equivalent, with an average value of 7560 ± 1546 wt ppm H_2O (two standard deviations, 2σ) for 16 measurements (Fig. 1C and table S4). Although their metamorphic grades are similar to that of Sahara 97096, the CV chondrites Vigarano and Kaba have lower H contents of 330 ± 140 and 210 ± 120 wt ppm H_2O, respectively (Fig. 1C and table S4). The H isotopic composition of Sahara 97096 mesostasis is homogeneous, with an average δD value of −147 ± 16‰ (2σ; Fig. 1D and table S4), whereas Vigarano and Kaba have δD = −261 ± 25 and +17 ± 95‰, respectively (2σ; Fig. 1D and table S4). Among these three meteorites, only Kaba underwent aqueous alteration of its chondrules (*18*), which accounts for the variable D/H ratio of its mesostasis. Vigarano contains traces of incipient aqueous alteration restricted to its fine-grained matrix (*19*). Because Sahara 97096 does not show any evidence of aqueous alteration, the abundant H contents of its chondrule mesostases must have been acquired before the EC parent asteroid evolution, possibly during the period of chondrule formation.

Considering the H concentration of Sahara 97096 mesostasis, the modal abundance of mesostasis in Sahara 97096 chondrules [16.2 vol % (*15*)], and the modal abundance of chondrules in ECs [~70 vol % (*20*)], H in chondrule mesostases accounts for 702 ± 117 wt ppm H_2O of the bulk rock, or ~13% of the bulk-rock H content. Another potential source of H in ECs is insoluble organic matter (IOM). However, given the low abundance of IOM in ECs (0.6 wt %) and the H concentration in EC organics [0.7 wt % (*12*)], the contribution of IOM to the bulk-rock H content is only 380 ppm H_2O, or 7.7% of the total H content of the chondrite. Combined, H hosted in chondrule mesostases and IOM can account for only ~20% of the total H measured in the Sahara 97096 bulk rock.

H concentrations have been reported in low-calcium pyroxenes from non-CC materials from (i) the S-type asteroid Itokawa (700 to 1000 wt ppm H_2O) brought back by the Hayabusa sample-return spacecraft and (ii) the metamorphosed ordinary chondrite (OC) Larkman Nunatak 12036 (600 to 1300 wt ppm H_2O) (*21*). We found an even higher H concentration (5300 wt ppm H_2O; table S2) for the pyroxene fraction of the Norton County aubrite. The modal abundance of low-calcium pyroxene in ECs is ~50 vol % (*15*), so enstatite pyroxenes could account for ~15 or ~58% of the bulk H content of ECs, considering the highest H content of OC pyroxenes (1300 wt ppm H_2O) or the aubrite pyroxene fraction, respectively. The remaining ~65 to 20% of H measured in the bulk rock may be contributed by other unknown carriers such as sulfur-rich carbon-bearing porous amorphous silica (*22*), owing to an underestimation of the H content in EC pyroxenes, and/or could be derived from some pervasive terrestrial contamination, although a limited amount of contamination is expected (otherwise, it would have erased the thermal metamorphism effect visible on the diagram of δD versus 1/H; fig. S5). Our SIMS measurements demonstrate that at least 20% of the H in ECs has an identified carrier phase (mesostasis or IOM) and that most of the remaining 80% is probably of indigenous origin. This implies that terrestrial hydrogen could be derived from ECs without invoking additional sources.

We incorporated the H concentrations measured in ECs (tables S2 and S4) into theoretical models of Earth's formation that invoke mixing of chondrite-like materials. These models use various proportions of EC-like, OC-like, and CC-like materials (*1*–*3*, *5*, *23*), so we explored three end-member compositions: (i) 100% EC-like materials (*1*); (ii) 68% EC-like and 32% CC-like materials (*2*, *23*), although these proportions could be inconsistent for the isotopic composition of Earth's mantle (*11*); and (iii) ~70% EC-like, ~25% OC-like, and ~5% CC-like materials (*3*) (Fig. 2). In all cases, EC-like materials contribute substantially to Earth's water budget, supplying 3.4 to 23.1 times the mass of Earth's oceans (1.4×10^{21} kg) once our H content measurements are included (table S2). These estimates are upper limits because loss of volatiles might have occurred during Earth's accretion (*24*, *25*). Considering only identified H-bearing phases (mesostasis and IOM), EC-like materials alone contribute ~3 to 4.5 times the mass of Earth's oceans (Fig. 2). If Sahara 97096 is representative of the most pristine ECs (Fig. 2), its bulk-rock hydrogen content (4700 wt ppm H_2O or ≥14 ocean masses) is sufficient to explain the highest estimates of the water content of Earth's surface and mantle, which are 1000 to 3900 ppm H_2O (*5*, *26*). Mesostasis and IOM H contents alone can account for three to five times the total mass of the oceans, corresponding to intermediate estimates of the water content of Earth's mantle.

The hydrogen isotopic composition of ECs is depleted in deuterium relative to that of Earth's oceans but is still within the range of mantle values (Fig. 3). ECs are not the only group of meteorites with D-poor values consistent with those of Earth's mantle, as CM-type CCs also span the range of mantle H isotopic compositions (Fig. 3). Nitrogen isotopes provide additional constraints (*27*), as the mantle is depleted in both D and ^{15}N relative to the surface reservoirs (atmosphere and oceans). Combining H and N isotopic compositions allows us to identify potential chondritic precursors (Fig. 4). Only ECs are compatible with both the δD and $\delta^{15}N$ values measured in mantle-derived rocks (Fig. 4). D-poor solar hydrogen [δD = −865‰ (*6*)] could also have contributed to the mantle composition (*28*), although it would require extreme enrichments of solar H relative to solar Ne because mantle neon is solar-like. The chondritic H/Ne

ratio [4×10^8 (*29*)] is five orders of magnitude higher than the solar ratio [7.9×10^3 (*30*)]. Therefore, we favor the incorporation of EC-like material in the silicate Earth (*1*).

Earth's surficial (oceans and atmosphere) H and N isotopic compositions ($\delta D \equiv 0‰$ and $\delta^{15}N \equiv 0‰$) cannot be easily derived from EC materials alone because they are richer in D and ^{15}N by ~100 to 200‰ and 5 to 40‰, respectively. Atmospheric escape to space, often invoked for atmospheric species including water vapor, could potentially achieve such enrichments, but the compositions of light noble gases do not support this hypothesis. Ne-Ar isotopic variations in the mantle-atmosphere system are not consistent with the elemental and isotopic variations expected for atmospheric escape; rather, they are best reproduced by mixing between a solar-like component trapped in the solid Earth and a CC-like component with a low $^{20}Ne/^{22}Ne$ ratio (*5*). A contribution of CC-like materials, especially CIs, to the surface inventory would also be consistent with the observed H and N isotopic variations (Fig. 4) and can be computed independently from H and N isotopic ratios (*11*). We find a ~4% contribution from CI-like material to Earth's surficial reservoir based on hydrogen abundances and isotopic compositions and a ~15% contribution based on equivalent data for nitrogen (*11*). Although nitrogen isotopes could also have contributions from cometary materials (*11*), these estimates are similar to those obtained from noble gases (*5*) and molybdenum isotopes (*23*). We cannot determine exactly when the CI-like material was delivered to the surficial reservoir, but it must have been sufficiently late during Earth's formation to prevent a global rehomogenization of the mantle with the surface.

REFERENCES AND NOTES

1. M. Javoy *et al.*, *Earth Planet. Sci. Lett.* **293**, 259–268 (2010).
2. P. H. Warren, *Earth Planet. Sci. Lett.* **311**, 93–100 (2011).
3. N. Dauphas, *Nature* **541**, 521–524 (2017).
4. A. Morbidelli *et al.*, *Meteoritics* **35**, 1309–1320 (2000).
5. B. Marty, *Earth Planet. Sci. Lett.* **313–314**, 56–66 (2012).
6. G. Gloeckler, J. Geiss, in *Primordial Nuclei and Their Galactic Evolution*, N. Prantzos, M. Tosi, R. Von Steiger, Eds. (Springer, 1998), pp. 275–284.
7. C. M. O. D. Alexander *et al.*, *Science* **337**, 721–723 (2012).
8. M. Hässig *et al.*, *Science* **347**, aaa0276 (2015).
9. F. Robert, *Space Sci. Rev.* **106**, 87–101 (2003).
10. S. N. Raymond, A. Izidoro, *Icarus* **297**, 134–148 (2017).
11. Materials and methods are available as supplementary materials.
12. L. Piani *et al.*, *Meteoritics* **47**, 8–29 (2012).
13. E. Quirico, M. Bourot-denise, C. Robin, G. Montagnac, P. Beck, *Geochim. Cosmochim. Acta* **75**, 3088–3102 (2011).
14. M. K. Weisberg, M. Prinz, "Sahara 97096: A highly primitive EH3 chondrite with layered sulfide-metal-rich chondrules," abstr. 1741, 29th Lunar and Planetary Science Conference, Houston, TX, 16 to 20 March 1998; www.lpi.usra.edu/meetings/LPSC98/pdf/1741.pdf.
15. L. Piani, Y. Marrocchi, G. Libourel, L. Tissandier, *Geochim. Cosmochim. Acta* **195**, 84–99 (2016).
16. A. Stephant, L. A. J. Garvie, P. Mane, R. Hervig, M. Wadhwa, *Sci. Rep.* **8**, 12385 (2018).
17. L. Bonal, E. Quirico, M. Bourot-Denise, G. Montagnac, *Geochim. Cosmochim. Acta* **70**, 1849–1863 (2006).
18. L. P. Keller, P. R. Buseck, *Geochim. Cosmochim. Acta* **54**, 2113–2120 (1990).
19. M. R. Lee, R. Hutchison, A. L. Graham, *Meteorit. Planet. Sci.* **31**, 477–483 (1996).
20. E. R. D. Scott, A. N. Krot, in *Meteorites and Cosmochemical Processes*, Volume 1 of *Treatise on Geochemistry*, A. M. Davis, Ed. (Elsevier, ed. 2, 2014), pp. 65–137.
21. Z. Jin, M. Bose, *Sci. Adv.* **5**, eaav8106 (2019).
22. S. W. Lehner, P. Németh, M. I. Petaev, P. R. Buseck, *Meteorit. Planet. Sci.* **52**, 2424–2436 (2017).
23. G. Budde, C. Burkhardt, T. Kleine, *Nat. Astron.* **3**, 736–741 (2019).
24. K. Hamano, Y. Abe, H. Genda, *Nature* **497**, 607–610 (2013).
25. H. Lammer *et al.*, *Mon. Not. R. Astron. Soc.* **439**, 3225–3238 (2014).
26. A. H. Peslier, M. Schönbächler, H. Busemann, S. I. Karato, *Space Sci. Rev.* **212**, 743–810 (2017).
27. M. Javoy, F. Pineau, D. Demaiffe, *Earth Planet. Sci. Lett.* **68**, 399–412 (1984).
28. L. J. Hallis *et al.*, *Science* **350**, 795–797 (2015).
29. B. Marty *et al.*, *Earth Planet. Sci. Lett.* **441**, 91–102 (2016).
30. K. Lodders, *Astrophys. J.* **591**, 1220–1247 (2003).
31. L. Piani, F. Robert, L. Remusat, *Earth Planet. Sci. Lett.* **415**, 154–164 (2015).
32. E. Deloule, F. Albarède, S. M. F. Sheppard, *Earth Planet. Sci. Lett.* **105**, 543–553 (1991).
33. M. W. Loewen, D. W. Graham, I. N. Bindeman, J. E. Lupton, M. O. Garcia, *Earth Planet. Sci. Lett.* **508**, 62–73 (2019).
34. M. M. Grady, I. P. Wright, L. P. Carr, C. T. Pillinger, *Geochim. Cosmochim. Acta* **50**, 2799–2813 (1986).
35. P. Cartigny, B. Marty, *Elements* **9**, 359–366 (2013).

ACKNOWLEDGMENTS

We thank the Field Museum (Chicago, USA), the French National Museum of Natural History (Paris, France), the Japanese National Institute for Polar Research (Tokyo, Japan), the University of New Mexico (Albuquerque, USA), Natural History Museum (Vienna, Austria), and the CEREGE meteoritic collection (Aix en Provence, France) for providing the meteorite samples. We thank N. Bouden and the members of the Ion Probe Team Nancy (IPTN) for help with the SIMS. We thank E. Deloule, A. Gurenko, B. Luais, L. France, M. Broadley, D. Bekaert, and C. Cartier for fruitful discussions. This is CRPG contribution 2997. **Funding:** This work was supported by the French Research National Agency (grant ANR-19-CE31-0027-01 to L.P.) and the European Research Council (grant 695618 to B.M.). **Author contributions:** L.P. designed the study. L.P., T.R., L.G.V., and D.T. performed the bulk and in situ analyses. Y.M. and B.M. contributed to the experimental design. L.P., Y.M., and B.M. discussed the data and wrote the manuscript. All authors provided input to the data analysis and manuscript preparation. **Competing interests:** We declare no competing interests. **Data and materials availability:** All meteorites that we studied are deposited in public museums, as listed in the supplementary materials. Our measured H abundances and D/H ratios, for the meteorite samples and standards, are listed in tables S1 to S4 and data S1.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1110/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S6
References (*36–63*)
Data S1

22 November 2019; accepted 15 July 2020
10.1126/science.aba1948

ORGANIC CHEMISTRY

Static to inducibly dynamic stereocontrol: The convergent use of racemic β-substituted ketones

Jacob S. DeHovitz*, Yong Yao Loh*, Jacob A. Kautzky, Kazunori Nagao, Andrew J. Meichan, Motoshi Yamauchi, David W. C. MacMillan†, Todd K. Hyster†

The synthesis of stereochemically complex molecules in the pharmaceutical and agrochemical industries requires precise control over each distinct stereocenter, a feat that can be challenging and time consuming using traditional asymmetric synthesis. Although stereoconvergent processes have the potential to streamline and simplify synthetic routes, they are currently limited by a narrow scope of inducibly dynamic stereocenters that can be readily epimerized. Here, we report the use of photoredox catalysis to enable the racemization of traditionally static, unreactive stereocenters through the intermediacy of prochiral radical species. This technology was applied in conjunction with biocatalysts such as ketoreductases and aminotransferases to realize stereoconvergent syntheses of stereodefined γ-substituted alcohols and amines from β-substituted ketones.

The concept of static stereocenters is a fundamental principle that underpins the induction, application, and teaching of stereochemistry. The vast majority of stereocenters are considered to be inherently static (i.e., once formed, they are stable toward racemization or epimerization) (Fig. 1A, left). As a result, the configuration of a stereocenter is generally thought to be intrinsically coupled to the step in which it was forged, requiring each bond-forming step to also bear the burden of possessing high levels of kinetic stereocontrol. This approach to structural complexity has been adopted universally (*1*–*6*); however, alternative strategies wherein bond formation and stereochemical induction are decoupled could prove to be both powerful and complementary.

Whereas most of the stereocenters in organic chemistry are classified as static, a small cadre of structural motifs are amenable to stereochemical interconversion (or stereoablation) under prescribed reaction conditions. These inducibly dynamic centers typically incorporate (i) acidic C–H bonds that can undergo reversible deprotonation in the presence of an appropriate base (e.g., α-carbonyl C–H bonds); (ii) α-heteroatom C–H bonds (such as α-hydroxy or α-amino) that can readily participate in hydrogen atom transfer (Fig. 1A, right); and (iii) C–heteroatom bonds that can undergo reversible formation via 1,2- or 1,4-addition (e.g., acetals or β-amino carbonyls) (*7*).

Although dynamic stereochemistry might seem antithetical to the objective of a multistep stereoselective synthesis, there have been a number of well-established strategies that elegantly exploit the capacity for stereogenic racemization or epimerization (Fig. 1A, right). In particular, out-of-equilibrium processes, such as the uphill olefin isomerization shown by Weaver and colleagues (*8*) and the deracemizations detailed by Turner and colleagues (*9*) and Knowles, Miller, and colleagues (*10*), as well as a litany of dynamic kinetic resolutions (DKRs) have successfully used dynamic stereocontrol over the past two decades (*11*–*13*). Among these strategies, DKRs rely on the rapid interconversion of substrate enantiomers, in combination with stereoselective bond formation, to transform racemic substrates into enantio- and diastereomerically pure products (Fig. 1A, right). Such a stereoconvergent approach can substantially simplify the steps required for complex molecule synthesis by effectively decoupling bond formation from stereoinduction, as racemic intermediates can be converted to stereochemically pure final products.

Dynamic kinetic resolutions are nonetheless currently restricted to substrates that incorporate traditionally inducibly dynamic stereocenters, which constitutes a serious limitation given that most stereogenic centers in organic chemistry are static. With this in mind, we sought to determine whether it might be feasible to develop catalytic approaches that would render traditionally static stereocenters dynamic under mild conditions. In particular, we were interested in combining such racemization protocols with the exquisite stereoselectivity of enzymes toward molecules with static distal stereocenters (*14*) (Fig. 1B, right). Here, we describe such an advance, wherein we apply photoredox catalysis to racemize traditionally static β-keto stereocenters while successfully merging this protocol with biocatalysis to achieve a dynamic kinetic resolution.

Photoredox catalysis has emerged as an enabling platform for regioselective C–H functionalization through the generation and harnessing of reactive carbon-centered radical intermediates toward a variety of synthetically useful bond-forming reactions (*15*–*17*). The power of photoredox catalysis is highlighted by its activation of conventionally inert, remote C–H bonds, enabling the functionalization of C–H bonds at β-carbonyl (*18*), β-amino (*19*), and other unactivated aliphatic C–H moieties (*20*) (Fig. 1B, left). On this basis, we questioned whether we could leverage the photoredox-catalyzed generation of these planar, prochiral radicals as a stereoablative tool for traditionally static stereocenters. Subsequent nonstereospecific trapping of these carbon-centered radical species by a suitable hydrogen atom transfer (HAT) catalyst could then enable the racemization of static stereocenters via the reversible cleavage and formation of C–H bonds, allowing the use of substrates with remote, unactivated, static C–H stereocenters in stereoconvergent processes.

Specifically, we postulated that the direct β-functionalization of saturated cyclic ketones via the merger of photoredox and organocatalysis could serve as a framework for a method toward the racemization of β-keto stereocenters through the intermediacy of a key 5π e^- β-enaminyl radical species (Fig. 1C) (*18*). In contrast to the well-precedented DKRs of α-substituted ketones (*11*), DKRs of β-substituted ketones have remained underdeveloped owing to a lack of general methods for the racemization of static β-keto stereocenters. By applying the proposed photoredox racemization protocol in conjunction with a ketoreductase-catalyzed kinetic resolution, we aimed to achieve a broadly applicable dynamic reductive kinetic resolution of β-substituted ketones to furnish stereodefined γ-substituted alcohols as products. A crucial design element of this DKR strategy is the inability of alcohol products to form the key β-enaminyl radical intermediate, thereby precluding the product from undergoing racemization after its formation.

A proposed mechanism for this photoredox- and enzyme-catalyzed DKR of β-substituted ketones is outlined in Fig. 2A. For demonstration purposes, we have depicted the racemization process as the conversion of (*S*)-**1** to the faster reacting active enantiomer (*R*)-**1**. This begins through the condensation of (*S*)-**1** with amine catalyst **2** to form stereodefined enamine (*S*)-**3** in situ. At the same time, photoexcitation of photocatalyst [$Ir(dF(CF_3)ppy)_2(dtbbpy)](PF_6$) (**4**) would generate the highly oxidizing excited state *Ir(III) (**5**) (half-wave reduction potential $E_{1/2}^{red}$ [$*Ir^{III}/Ir^{II}$] = +1.21 V versus the saturated calomel electrode (SCE) in MeCN) (*21*). This species can then oxidize enamine (*S*)-**3** (half-peak potential E_p [$\mathbf{3}/\mathbf{3}^{\bullet+}$] = +0.36 V versus

Merck Center for Catalysis, Princeton University, Princeton, NJ 08544, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: dmacmill@princeton.edu (D.W.C.M.); thyster@princeton.edu (T.K.H.)

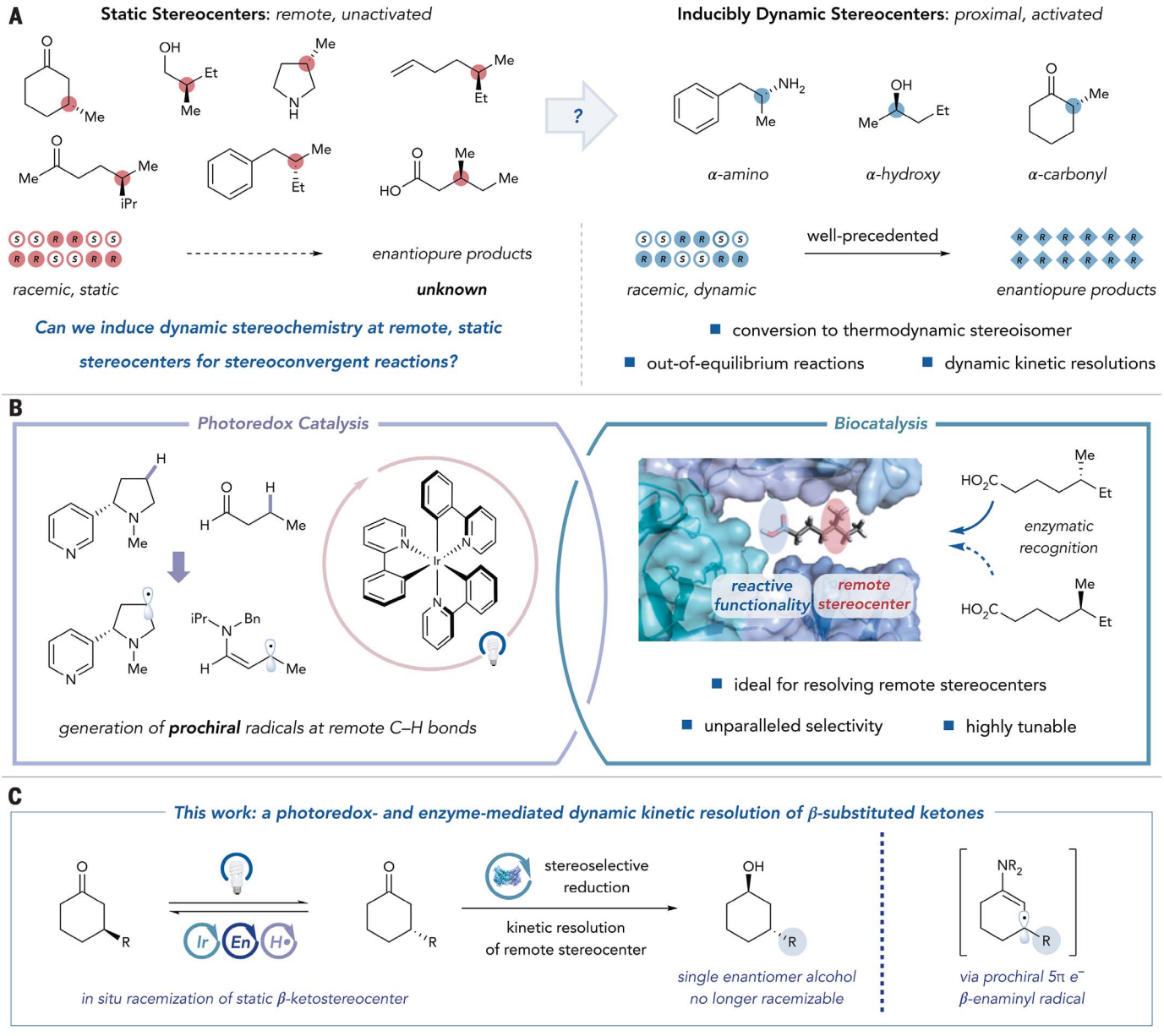


Fig. 1. Inducing dynamic stereochemistry for asymmetric synthesis. (**A**) Static stereocenters versus inducibly dynamic stereocenters. Red circles (filled and unfilled) indicate substrates containing static stereocenters. Blue circles (filled and unfilled) indicate substrates containing inducibly dynamic stereocenters. Blue diamonds indicate products formed from stereoconvergent reactions. *R* and *S* indicate the stereochemistry of compounds. (**B**) Combination of photoredox catalysis and biocatalysis for stereoconvergent reactions. (**C**) Proposed photoredox- and enzyme-catalyzed dynamic kinetic resolution. Me, methyl; Et, ethyl; iPr, isopropyl; Bn, benzyl; En, enamine; H•, hydrogen atom transfer.

SCE in MeCN) (*22*) to give reduced Ir(II) species (**6**) and enaminyl radical cation **7**. Subsequent allylic deprotonation of **7**, due to an increase in acidity of the allylic C–H bond, can then provide the nucleophilic 5π e^- β-enaminyl radical **8** (*23*). The formation of this key prochiral radical intermediate ablates the stereochemistry at the β-carbon, and subsequent polarity-matched HAT can occur between **8** and electrophilic thiol catalyst **9** [cyclohexene allylic C–H bond dissociation energy (BDE) = 83.2 kcal mol^{-1} (*24*), thiophenol S–H BDE = 76.9 kcal mol^{-1} (*25*)]. We expect this HAT event to be nonstereoselective, resulting in the gradual racemization of the stereocenter, with the formation of the opposite enantiomer of the enamine [(*R*)-**3**] occurring only 50% of the time. The photoredox and HAT catalytic cycles can then converge in a single-electron transfer (SET) event between **6** and thiyl radical **10**, regenerating ground-state photocatalyst **4** and forming thiolate **11**, which regenerates thiol catalyst **9** upon protonation [$E_{1/2}^{red}$ [Ir^{III}/Ir^{II}] = −1.37 V versus SCE in MeCN (*21*), $E_{1/2}^{red}$ [$PhS^{\bullet}/PhS^-$] = −0.06 V versus SCE in MeCN] (*26*). Finally, enamine (*R*)-**3** undergoes hydrolysis to furnish the active enantiomer of the ketone (*R*)-**1**. Concurrent with this racemization process, the ketoreductase-catalyzed stereoselective reduction of (*R*)-**1** is constantly occurring to furnish enantiopure alcohol **12** as product, using cofactor NADPH (the reduced form of nicotinamide adenine dinucleotide phosphate) as the hydride source. The reaction can then be rendered catalytic in NADPH via the use of an exogenous terminal reductant. Depending on

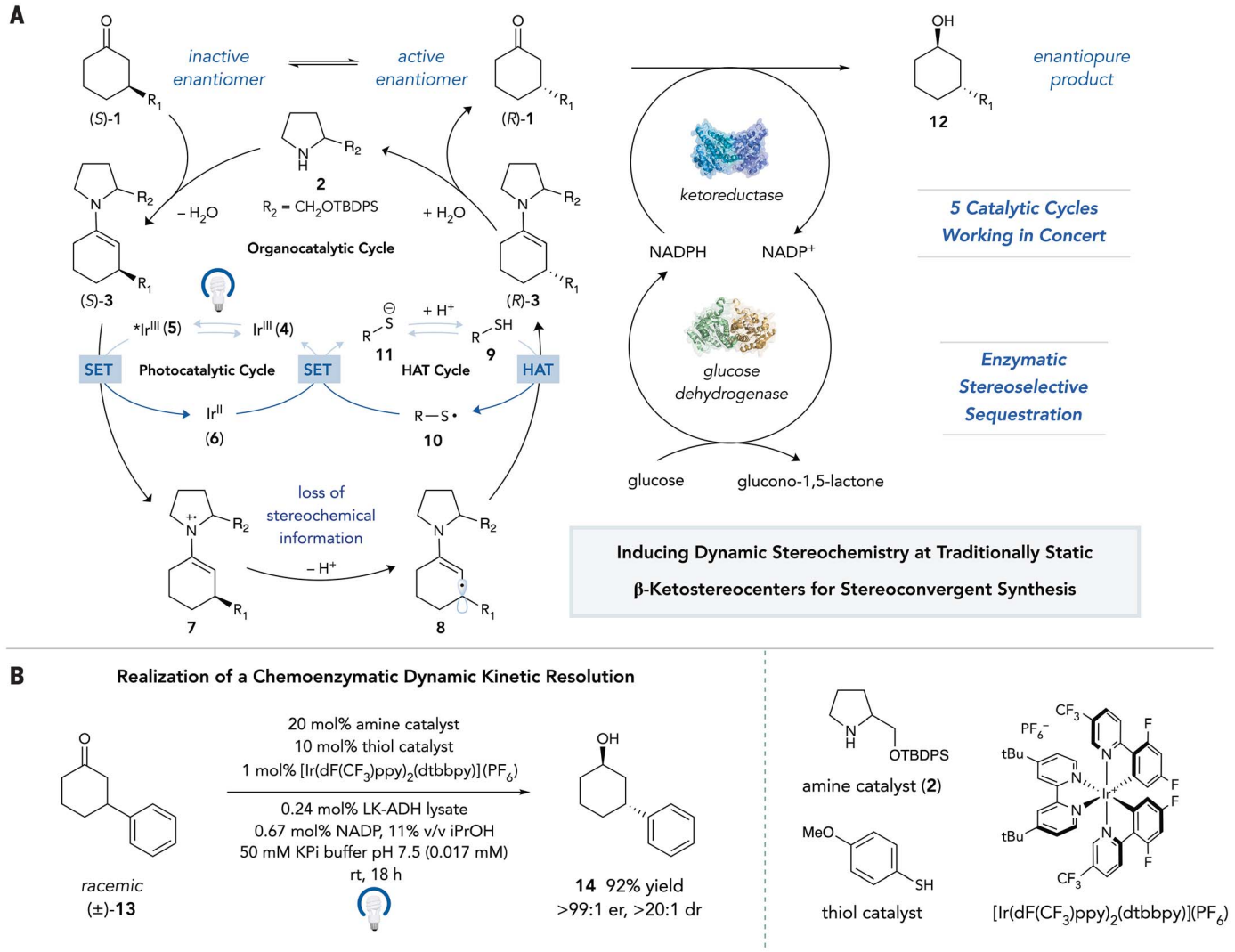


Fig. 2. Proposed mechanism and preliminary studies. (**A**) Proposed mechanism for the photoredox- and enzyme-catalyzed dynamic kinetic resolution of β-substituted ketones. For the purposes of demonstration, the racemization process is depicted as the conversion of (*S*)-**1** to the faster-reacting active enantiomer (*R*)-**1**. (**B**) Successful combination of photoredox and enzyme catalysis to realize a dynamic kinetic resolution. TBDPS, *tert*-butyldiphenylsilyl; tBu, *tert*-butyl; NADPH, reduced nicotinamide adenine dinucleotide phosphate; dF(CF_3)ppy, 2-(2,4-difluorophenyl)-5-(trifluoromethyl)pyridine; dtbbpy, 4,4′-di-*tert*-butyl-2,2′-bipyridine; LK-ADH, *Lactobacillus kefir* alcohol dehydrogenase [Protein Data Bank (PDB) ID 4RF2]; glucose dehydrogenase (PDB ID 1GCO); KPi, potassium phosphate; iPrOH, isopropyl alcohol; mM, millimolar; rt, room temperature; h, hours; er, enantiomeric ratio; dr, diastereomeric ratio.

the ketoreductase employed, either the combination of glucose dehydrogenase (GDH) with glucose (as shown) or excess alcoholic cosolvents, such as isopropanol, can be used. This demonstration of five catalytic cycles working in concert highlights the compatibility of enzymatic stereoselective sequestration with organocatalysis and the radical generation capabilities of photoredox catalysis (*27*–*30*), facilitating the engagement of traditionally static β-ketostereocenters in a stereoconvergent protocol.

We began our investigations into the proposed chemoenzymatic DKR of β-substituted ketone substrates by first examining the racemization of (*R*)-3-phenylcyclohexanone [(*R*)-**13**] (see supplementary materials for more details). After evaluating a variety of photocatalysts, amine catalysts, thiol catalysts, and solvents, we observed that the use of [Ir(dF(CF_3)ppy)$_2$(dtbbpy)](PF_6) as photocatalyst, racemic 2-(((*tert*-butyldiphenylsilyl)-oxy) methyl)pyrrolidine as amine catalyst and 4-methoxythiolphenol as HAT catalyst facilitated efficient racemization after 4 hours of irradiation with visible light [98:2 to 60:40 enantiomeric ratio (er)] in the enzyme-compatible solvent system of isopropanol and phosphate buffer (1:8 ratio) (see supplementary materials). The hydrophobicity of the amine catalyst used was critical for efficient racemization owing to the well-precedented facile hydrolysis of the catalytically generated enamine species under predominantly aqueous conditions (*31*). In addition, preliminary studies also indicated the efficacy of the hydrophobic amine catalyst in mitigating enzymatic deactivation (see supplementary materials). Concurrent with the optimization of the racemization protocol, we also examined the ketoreductase-catalyzed kinetic resolution of racemic 3-phenylcyclohexanone [(±)-**13**] (see supplementary materials for more details). We found that wild-type *Lactobacillus kefir* alcohol dehydrogenase (LK-ADH) was able to successfully kinetically resolve 3-phenylcyclohexanone, giving alcohol **14** with excellent enantio- and diastereoselectivity [51% yield, 99:1 er, and 10:1 diastereomeric ratio (dr)]. With these promising results in hand, we proceeded to combine the racemization and the kinetic resolution

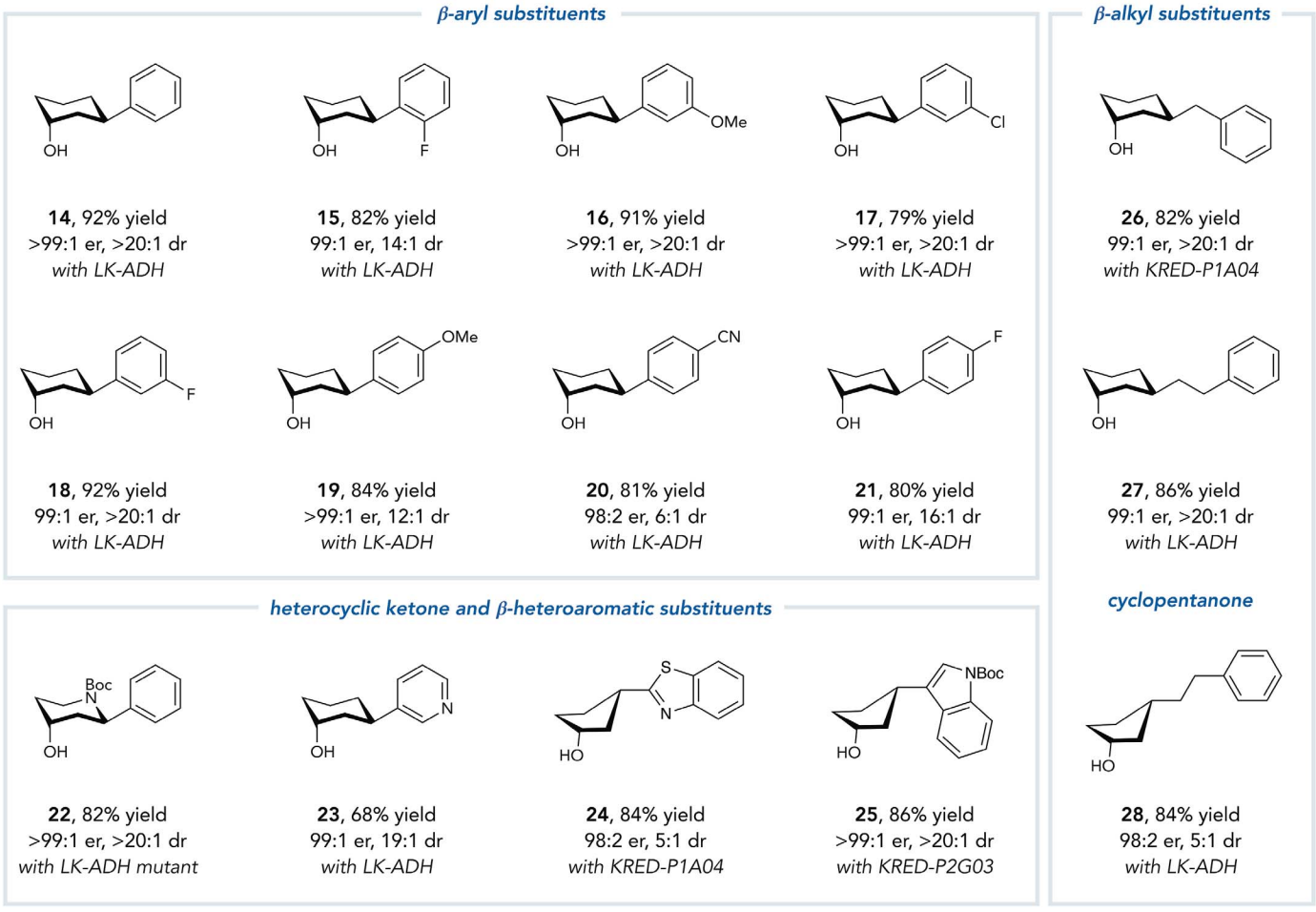


Fig. 3. Scope of dynamic kinetic resolution of β-substituted ketones. See supplementary materials for specific reaction details. Boc, *tert*-butyloxycarbonyl; KRED, ketoreductase.

to realize our desired DKR of β-substituted ketones (Fig. 2B). Upon subjecting (±)-**13** to our photoredox- and enzyme-catalyzed protocol, alcohol **14** was obtained in 92% yield, >99:1 er, and >20:1 dr.

Next, we set out to explore the scope of our DKR protocol (Fig. 3). Given the privileged status of ketoreductases in biocatalysis, we developed a library of 32 nicotinamide-dependent ketoreductases from both commercial sources and from in-house heterologous expression in *Escherichia coli*. Using this library, we were able to quickly identify the appropriate enzymes required for effective kinetic resolutions of a range of substrates. These hits were subsequently combined with our photoredox-catalyzed racemization protocol and translated into DKRs. We found that our racemization protocol was compatible with the different enzymes identified from the kinetic resolution screens, which speaks to the protocol's general biocompatibility.

A range of differentially substituted β-phenylcyclohexanones were found to be excellent substrates and were successfully converted into their corresponding γ-substituted alcohols in good yields and selectivities. Substitution at the ortho position with fluorine was well tolerated (**15**; 82% yield, 99:1 er, 14:1 dr), as were electronically diverse meta- (**16** to **18**; 79 to 92% yield, >99:1 er, >20:1 dr) and para-substituents (**19** to **21**; 80 to 84% yield, 98:2 to >99:1 er, 6:1 to 16:1 dr). Heterocyclic ketones such as piperidones also performed well in this reaction (**22**; 82% yield, >99:1 er, >20:1 dr), while β-heteroaromatic substituents such as pyridines, benzothiazoles, and indoles were good substrates as well (**23** to **25**; 68 to 86% yield, 98:2 to >99:1 er, 5:1 to >20:1 dr). Additionally, β-alkyl substituents such as benzyl and phenethyl also worked well in the DKR (**26** to **28**; 82 to 86% yield, 98:2 to 99:1 er, 5:1 to >20:1 dr). Finally, across these examples, cyclopentanones were demonstrated to be excellent substrates for this protocol (**24**, **25**, and **28**; 84 to 86% yield, 98:2 to >99:1 er, 5:1 to >20:1 dr).

To demonstrate the utility of our DKR in a pharmaceutical drug development setting, we applied it to the synthesis of LNP023, a complement factor B inhibitor recently reported by Novartis (*32*) (Fig. 4A). By subjecting racemic piperidone **29** to our DKR protocol, we were able to obtain alcohol **30** in 85% yield, >99:1 er, and >20:1 dr, shortening the synthesis of LNP023 by two steps, in addition to avoiding the need for two chiral resolutions, drastically simplifying the synthetic sequence.

Beyond streamlining the synthesis of a single enantiomer of a molecule, inducibly dynamic stereocenters also enable stereodivergent synthesis. By allowing molecules to be synthesized racemically and delaying the incorporation of stereochemical information, different stereoisomers of a desired product can be accessed simply by changing the selectivity sense of the chiral catalyst, avoiding the need for independent synthetic approaches (*33*). To showcase the viability of such an approach, we were able to identify three enzymes from our 32-enzyme library that enabled us to access the three remaining stereoisomers of **14** in good

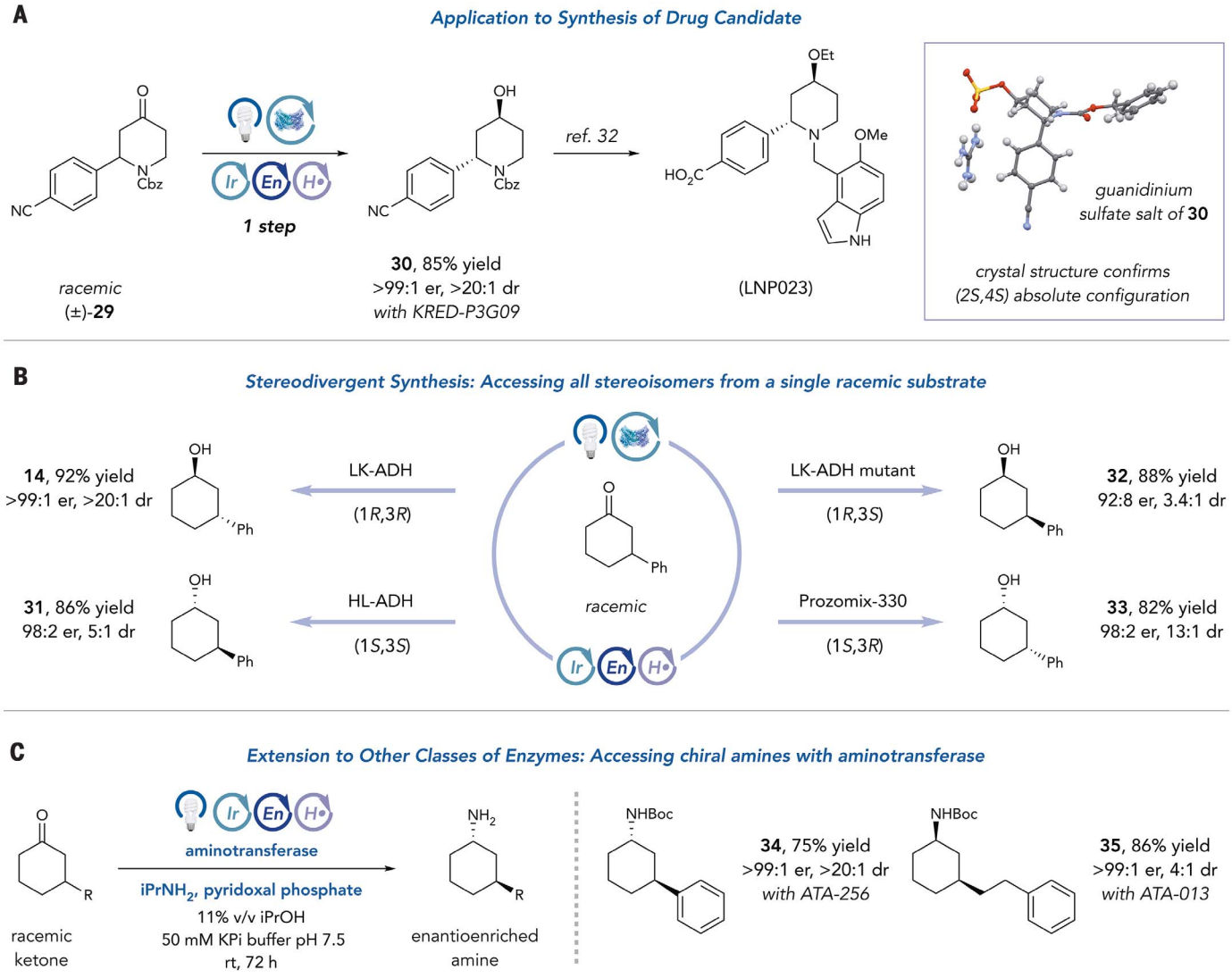


Fig. 4. Further applications of racemization protocol. (**A**) Simplification of synthetic routes to drug candidates. (**B**) Stereodivergent synthesis. (**C**) Extension to other classes of enzymes. Amine products were isolated after a Boc-protection workup. Cbz, benzyloxycarbonyl; HL-ADH, horse liver alcohol dehydrogenase; ATA, amine transaminase.

enantioselectivity and moderate diastereoselectivity (**31** to **33**; 82 to 88% yield, 92:8 to 98:2 er, 3.4:1 to 13:1 dr) (Fig. 4B).

As a final test of the generality and biocompatibility of our racemization protocol, we explored the feasibility of extending our scope to other classes of enzymes beyond ketoreductases and products beyond enantioenriched γ-substituted alcohols. In principle, any enzyme class that alters the carbonyl functional group to prevent enamine formation with the product should be compatible with our racemization strategy. To this end, we identified aminotransferases ATA-256 and ATA-013 to be capable of kinetically resolving 3-phenylcyclohexanone and 3-phenethylcyclohexanone, respectively. By combining these kinetic resolutions with our racemization protocol, we were able to successfully demonstrate the synthesis of enantioenriched amines **34** (75% yield, >99:1 er, >20:1 dr) and **35** (86% yield, >99:1 er, 4:1 dr) via a chemoenzymatic DKR protocol. We anticipate that this powerful platform for racemization will continue to challenge conventional definitions of static and dynamic stereocenters, enabling the development of additional stereoconvergent processes.

REFERENCES AND NOTES

1. D. A. Evans, *Science* **240**, 420–426 (1988).
2. E. N. Jacobsen, A. Pfaltz, H. Yamamoto, *Comprehensive Asymmetric Catalysis I–III* (Springer-Verlag Berlin Heidelberg, 1999).
3. M. S. Taylor, E. N. Jacobsen, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5368–5373 (2004).
4. S. Masamune, W. Choy, J. S. Petersen, L. R. Sita, *Angew. Chem. Int. Ed.* **24**, 1–30 (1985).
5. A. H. Hoveyda, D. A. Evans, G. C. Fu, *Chem. Rev.* **93**, 1307–1370 (1993).
6. V. Andrushko, N. Andrushko, Eds., *Stereoselective Synthesis of Drugs and Natural Products* (Wiley, 2013).
7. F. F. Huerta, A. B. E. Minidis, J. E. Bäckvall, *Chem. Soc. Rev.* **30**, 321–331 (2001).
8. J. I. Day, K. Singh, W. Trinh, J. D. Weaver 3rd, *J. Am. Chem. Soc.* **140**, 9934–9941 (2018).
9. M. Alexeeva, A. Enright, M. J. Dawson, M. Mahmoudian, N. J. Turner, *Angew. Chem. Int. Ed.* **41**, 3177–3180 (2002).
10. N. Y. Shin, J. M. Ryss, X. Zhang, S. J. Miller, R. R. Knowles, *Science* **366**, 364–369 (2019).
11. O. Pàmies, J. E. Bäckvall, *Trends Biotechnol.* **22**, 130–135 (2004).
12. O. Verho, J.-E. Bäckvall, *J. Am. Chem. Soc.* **137**, 3996–4009 (2015).
13. V. Bhat, E. R. Welin, X. Guo, B. M. Stoltz, *Chem. Rev.* **117**, 4528–4561 (2017).
14. R. L. O. R. Cunha, E. A. Ferreira, C. S. Oliveira, Á. T. Omori, *Biotechnol. Adv.* **33**, 614–623 (2015).
15. M. H. Shaw, J. Twilton, D. W. C. MacMillan, *J. Org. Chem.* **81**, 6898–6926 (2016).

16. M. D. Kärkäs, J. A. Porco Jr., C. R. J. Stephenson, *Chem. Rev.* **116**, 9683–9747 (2016).
17. J. Twilton *et al*., *Nat. Rev. Chem.* **1**, 0052 (2017).
18. M. T. Pirnot, D. A. Rankic, D. B. C. Martin, D. W. C. MacMillan, *Science* **339**, 1593–1596 (2013).
19. P. J. Sarver *et al*., *Nat. Chem.* **12**, 459–467 (2020).
20. I. B. Perry *et al*., *Nature* **560**, 70–75 (2018).
21. M. S. Lowry *et al*., *Chem. Mater.* **17**, 5712–5719 (2005).
22. W. W. Schoeller, J. Niemann, P. Rademacher, *J. Chem. Soc. Perkin 2* **1988**, 369–373 (1988).
23. Various radical cations have been shown to exhibit enhanced acidity relative to their ground-state carbon acids (*34*–*36*).
24. S. L. Khursan, D. A. Mikhailov, V. M. Yanborisov, D. I. Borisov, *React. Kinet. Catal. Lett.* **61**, 91–95 (1997).
25. F. G. Bordwell, X.-M. Zhang, A. V. Satish, J.-P. Cheng, *J. Am. Chem. Soc.* **116**, 6605–6610 (1994).
26. A. Godsk Larsen, A. Hjarbaek Holm, M. Roberson, K. Daasbjerg, *J. Am. Chem. Soc.* **123**, 1723–1729 (2001).
27. F. Hollmann, A. Taglieber, F. Schulz, M. T. Reetz, *Angew. Chem. Int. Ed.* **46**, 2903–2906 (2007).
28. Z. C. Litman, Y. Wang, H. Zhao, J. F. Hartwig, *Nature* **560**, 355–359 (2018).
29. X. Guo, Y. Okamoto, M. R. Schreier, T. R. Ward, O. S. Wenger, *Chem. Sci.* **9**, 5052–5056 (2018).
30. K. Lauder *et al*., *Angew. Chem. Int. Ed.* **57**, 5803–5807 (2018).
31. N. Mase *et al*., *J. Am. Chem. Soc.* **128**, 734–735 (2006).
32. A. Schubart *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7926–7931 (2019).
33. I. P. Beletskaya, C. Nájera, M. Yus, *Chem. Rev.* **118**, 5080–5200 (2018).
34. F. G. Bordwell, J.-P. Cheng, M. J. Bausch, J. E. Bares, *J. Phys. Org. Chem.* **1**, 209–223 (1988).
35. F. G. Bordwell, M. J. Bausch, J. C. Branca, J. A. Harrelson, *J. Phys. Org. Chem.* **1**, 225–239 (1988).
36. F. G. Bordwell, A. V. Satish, *J. Am. Chem. Soc.* **114**, 10173–10176 (1992).

ACKNOWLEDGMENTS

We thank Codexis and Prozomix for providing transaminases and ketoreductases, respectively. We also thank V. Bacauanu, B. X. Li, and D. K. Kim for their assistance with the preparation of this manuscript, and P. Jeffery for assistance with x-ray structure determination. **Funding:** Research reported in this publication was supported by BioLEC, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, under award DE-SC0019370, and the Princeton Catalysis Initiative. The authors also acknowledge funding from Merck, Bristol-Myers Squibb (BMS), Eli Lilly, Genentech, Pfizer, and Johnson & Johnson. T.K.H. is supported by a Searle Scholars Award (SSP-2017-1741), a Sloan Research Fellowship, and the ACS-PRF. J.S.D. and Y.Y.L. are supported by BMS graduate fellowships. Y.Y.L. is also supported by an A*STAR graduate fellowship. J.A.K. is supported by a NSF GRFP (DGE 1656466) predoctoral fellowship. K.N. is supported by a JSPS overseas postdoctoral fellowship. M.Y. received funding from Daiichi Sankyo Co., Ltd. **Author contributions:** T.K.H. and D.W.C.M. conceived of the work. All authors designed the experiments. J.S.D., Y.Y.L., J.A.K., K.N., A.J.M., and M.Y. performed and analyzed the experiments. T.K.H., D.W.C.M., J.S.D., Y.Y.L., and J.A.K. prepared the manuscript. **Competing interests:** The authors declare no conflicts of interest. **Data and materials availability:** Commercial enzymes used are available from Codexis [single enzyme batches as well as screening kits such as the Codex Ketoreductase (KRED) Screening Kit and Codex Amine Transaminase (ATA) Screening Kit] and Prozomix [PRO-kREDy-to-go(001-384) Screening Kit]. Crystallographic parameters for the guanidinium sulfate salt of compound **30** are available free of charge from the Cambridge Crystallographic Data Centre under CCDC 2005332. Other data are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1113/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 and S2
HPLC Traces
X-ray Crystallography Data
NMR Spectra
References (*37*–*57*)

25 May 2020; accepted 7 July 2020
10.1126/science.abc9909

CORONAVIRUS

Structural basis of a shared antibody response to SARS-CoV-2

Meng Yuan[1]*, Hejun Liu[1]*, Nicholas C. Wu[1]*, Chang-Chun D. Lee[1], Xueyong Zhu[1], Fangzhu Zhao[2,3,4], Deli Huang[2], Wenli Yu[1], Yuanzi Hua[1], Henry Tien[1], Thomas F. Rogers[2,5], Elise Landais[2,3,6], Devin Sok[3,4,6], Joseph G. Jardine[3,6], Dennis R. Burton[2,3,4,7], Ian A. Wilson[1,3,4,8]†

Molecular understanding of neutralizing antibody responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) could accelerate vaccine design and drug discovery. We analyzed 294 anti–SARS-CoV-2 antibodies and found that immunoglobulin G heavy-chain variable region 3-53 (IGHV3-53) is the most frequently used IGHV gene for targeting the receptor-binding domain (RBD) of the spike protein. Co-crystal structures of two IGHV3-53–neutralizing antibodies with RBD, with or without Fab CR3022, at 2.33- to 3.20-angstrom resolution revealed that the germline-encoded residues dominate recognition of the angiotensin I converting enzyme 2 (ACE2)–binding site. This binding mode limits the IGHV3-53 antibodies to short complementarity-determining region H3 loops but accommodates light-chain diversity. These IGHV3-53 antibodies show minimal affinity maturation and high potency, which is promising for vaccine design. Knowledge of these structural motifs and binding mode should facilitate the design of antigens that elicit this type of neutralizing response.

The ongoing coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in enormous global health and socioeconomic damage and requires urgent development of an effective vaccine (*1*). Although multiple vaccine candidates have entered clinical trials (*2*), the molecular features that contribute to an effective antibody response are not clear. Shared antibody responses to specific microbial pathogens have been found in which the same genetic elements and modes of recognition are observed in multiple individuals against a given antigen. Such responses to microbial pathogens have been observed against influenza (*3*), dengue (*4*), malaria (*5*), and HIV (*6*). Characterization of the molecular interactions between pathogens and cognate antigen can provide insight into how the immune repertoire is able to quickly respond to novel microbial pathogens and will facilitate the rational design of vaccines against them (*7*, *8*).

The spike (S) protein is the major surface antigen of SARS-CoV-2. The S protein uses its receptor-binding domain (RBD) to engage the host receptor angiotensin I converting enzyme 2 (ACE2) for viral entry (*9*–*12*). RBD-targeting antibodies could then neutralize SARS-CoV-2 by blocking ACE2 binding. A number of antibodies that target the RBD of SARS-CoV-2 have now been discovered (*13*–*28*). We compiled a list of 294 SARS-CoV-2 RBD-targeting antibodies for which information on immunoglobulin G heavy-chain variable (IGHV) gene usage is available (*17*–*28*) (table S1) and found that IGHV3-53 is the most frequently used IGHV gene among these antibodies (Fig. 1A), with 10% encoded by IGHV3-53, compared with 0.5 to 2.6% (mean 1.8%) in the repertoire of naïve healthy individuals (*29*, *30*). IGHV3-53 antibodies were found in seven of 12 studies and in 17 of 32 COVID-19 patient samples (*17*–*28*, *31*). These IGHV3-53 antibodies not only had lower somatic mutation rates but also were more potent compared with other germ lines in the cohort investigated here (*27*) (fig. S1). The prevalence of IGHV3-53 in the antibody response in SARS-CoV-2 patients has also been recognized in other antibody studies (*20*, *22*, *27*).

To understand the molecular features that endow IGHV3-53 with favorable properties for RBD recognition, we determined the crystal structures of two IGHV3-53–neutralizing antibodies, CC12.1 and CC12.3, in complex with the SARS-CoV-2 RBD and with the cross-reactive Fab CR3022 to SARS-like CoVs (*17*). CC12.1 and CC12.3 were previously isolated from a SARS-CoV-2–infected patient and were shown to be specific for the RBD (*27*). CC12.1 and CC12.3 [median inhibitory concentration (IC_{50}), ~20 ng/ml] were among the top four highly potent neutralizing antibodies in the panel of

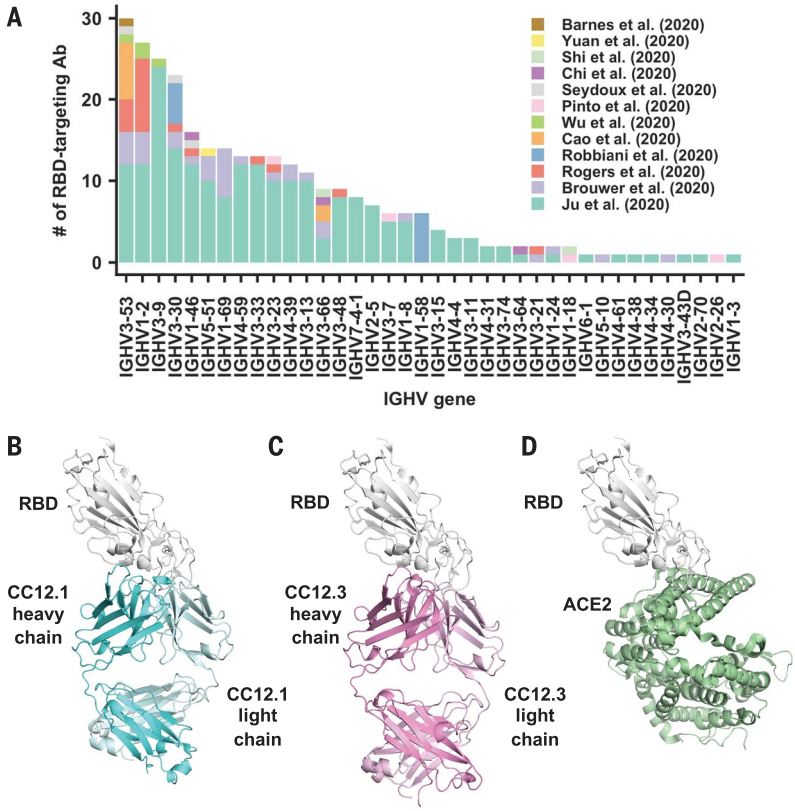


Fig. 1. Structures of two IGHV3-53 antibodies. (**A**) The distribution of IGHV gene usage is shown for a total of 294 RBD-targeting antibodies (*17*–*28*). (**B** to **D**) Crystal structures of (B) CC12.1 in complex with SARS-CoV-2 RBD, (C) CC12.3 with SARS-CoV-2 RBD, and (D) human ACE2 with SARS-CoV-2 RBD (PDB 6M0J) (*12*).

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA. [3]IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, CA 92037, USA. [4]Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, CA 92037, USA. [5]Division of Infectious Diseases, Department of Medicine, University of California, San Diego, La Jolla, CA 92037, USA. [6]IAVI, New York, NY 10004, USA. [7]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Cambridge, MA 02139, USA. [8]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, 92037, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: wilson@scripps.edu

antibodies assayed against live replicating SARS-CoV-2 virus and pseudovirus (*27*). Although CC12.1 and CC12.3 are both encoded by IGHV3-53, CC12.1 uses IGHJ6, IGKV1-9, and IGKJ3, whereas CC12.3 uses IGHJ4, IGKV3-20, and IGKJ1. This variation in IGHJ, IGKV, and IGKJ usage indicates that CC12.1 and CC12.3 belong to different clonotypes but are encoded by a common IGHV3-53 germline gene (fig. S2). IgBlast analysis (*32*) showed that IGHV and IGKV of CC12.1 have acquired only four amino acid changes (somatic mutations) during affinity maturation from the original germline antibody sequence (fig. S2, A and B). Similarly, CC12.3 is also minimally somatically mutated, with three amino acid changes in IGHV and a single amino acid deletion in IGKV (fig. S2, A and C). The binding affinities (K_d) of the Fabs CC12.1 and CC12.3 to SARS-CoV-2 RBD are 17 and 14 nM, respectively (fig. S3). Moreover, competition experiments suggest that CC12.1 and CC12.3 bind to a similar epitope, which overlaps with the ACE2-binding site but not the CR3022 epitope (fig. S4).

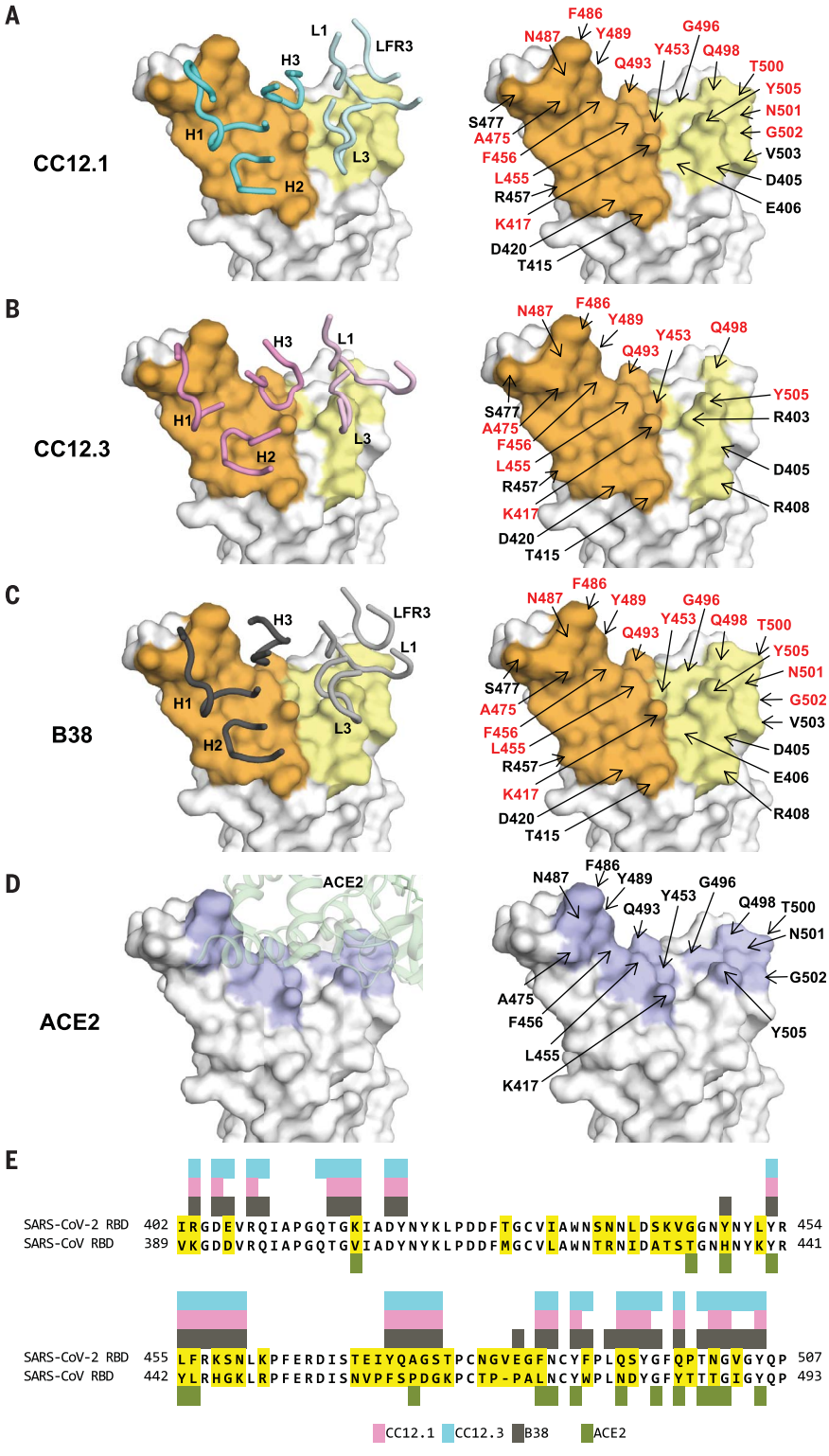


Fig. 2. Epitopes of IGHV3-53 antibodies. (**A** to **C**) Epitopes of (A) CC12.1, (B) CC12.3, and (C) B38 (PDB 7BZ5) (*23*). Epitope residues contacting the heavy chain are shown in orange and those contacting the light chain are shown in yellow. CDR loops are labeled in the left panels; epitope residues are labeled in the right panels. For clarity, only representative epitope residues are labeled. Epitope residues that are also involved in ACE2 binding are shown in red. (**D**) ACE2-binding residues are shown in blue. ACE2 is shown in green in the left panel in a semitransparent cartoon representation. ACE2-binding residues are labeled in the right panel. A total of 17 residues were used for ACE2 binding (*12*), but only 15 are labeled here because the other two are at the back of the structure in this view and do not interact with the antibodies here. (**E**) Epitope residues for CC12.1, CC12.3, and B38 were identified by PISA (*41*) and annotated on the SARS-CoV-2 RBD sequence, which is aligned to the SARS-CoV RBD sequence with nonconserved residues highlighted. The 17 ACE2-binding residues were identified from a SARS-CoV-2 RBD–ACE2 complex structure as described previously (*12*).

We determined four complex crystal structures, CC12.1/RBD, CC12.3/RBD, CC12.1/RBD/CR3022, and CC12.3/RBD/CR3022, at resolutions of 3.20, 2.33, 2.70, and 2.90 Å, respectively (table S2). CC12.1 and CC12.3 bind to the ACE2-binding site on SARS-CoV-2 RBD with an identical angle of approach (Fig. 1, B to D, and fig. S5). Another IGHV3-53 antibody, B38, the structure of which was determined recently (*23*), binds to the ACE2-binding site on SARS-CoV-2 RBD in a similar manner but with a K_d of 70.1 nM (fig. S6). Similar to the ACE2-binding site (*11*), the epitopes of these antibodies can only be accessed when the RBD is in the "up" conformation (fig. S7). Among 17 ACE2-binding residues on RBD, 15 and 16 are within the epitopes of CC12.1 and B38, respectively, and 11 are in the epitope of CC12.3 (Fig. 2, A to D). Many of the epitope residues are not conserved between SARS-CoV-2 and SARS-CoV (Fig. 2E), explaining their lack of cross-reactivity (*27*). The buried surface area (BSA) from the heavy-chain interaction is quite similar in CC12.1 (723 Å^2), CC12.3 (698 Å^2), and B38 (713 Å^2). By contrast, the light-chain interaction is much smaller for CC12.3 (176 Å^2) compared with CC12.1 (566 Å^2) and B38 (495 Å^2), consistent with different light-chain gene usage. Although both CC12.1 and B38 use IGKV1-9, CC12.3 uses IGKV3-20, which suggests that IGHV3-53 can pair with different light chains to target the ACE2-binding site of the SARS-CoV-2 RBD. CC12.1 (56% BSA from the heavy chain) binds the RBD with similar affinity to CC12.3 (80% BSA from the heavy chain) but with a slightly slower dissociation rate (fig. S3), which might be influenced by the different light chain and its greater contribution in CC12.1. Nevertheless, the light-chain identity seems not to be as critical as the heavy chain. In fact, among the RBD-targeting IGHV3-53 antibodies, nine different light chains are observed, although IGKV1-9 and IGKV3-20 are the most frequently found to date (fig. S8).

To understand why IGHV3-53 is elicited as a shared antibody response, the molecular interactions between the RBD and the heavy chains of CC12.1, CC12.3, and B38 were analyzed. The complementarity-determining regions (CDRs) H1 and H2 of these antibodies interact extensively with the RBD mainly through specific hydrogen bonds (Fig. 3, A and B). All residues on CDR H1 and H2 that hydrogen bond with the RBD are encoded by the germ line IGHV3-53 (fig. S2 and table S3). These interactions are almost identical among CC12.1, CC12.3, and B38, with the only difference at the variable region of immunoglobulin heavy chain (V_H) residue 58. A somatic mutation V_H Y58F in CC12.1 and CC12.3, but not in B38 (Fig. 3, A to C, boxed residues, and fig. S9), results in similar van der Waals interactions, with only a loss of a single hydrogen bond from the hydroxyl of the germ line Tyr in B38 to the RBD (Fig. 3C).

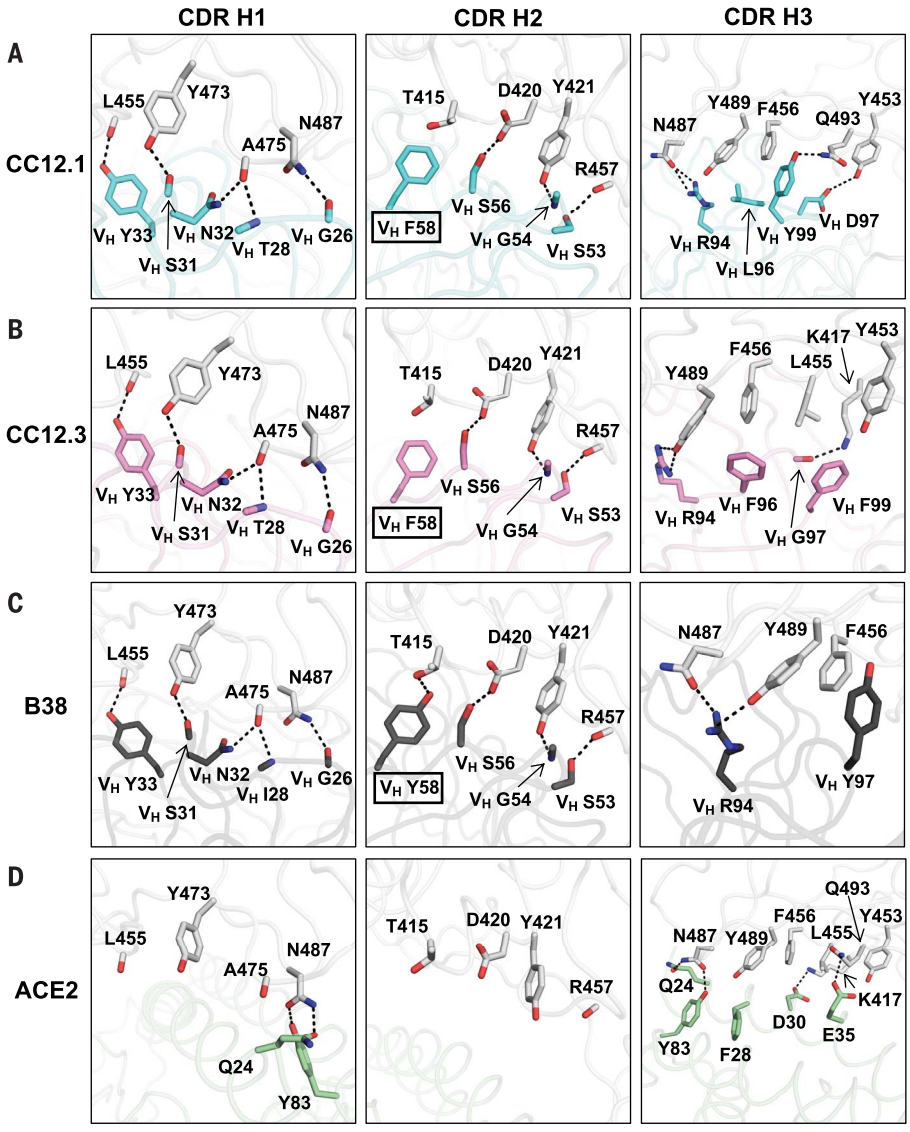


Fig. 3. Interactions between the RBD and heavy-chain CDR loops. (**A** to **C**) Highly similar interaction modes between SARS-CoV-2 RBD and the antibody CDR H1 and H2 loops, but not the H3 loop, are observed for (A) CC12.1, (B) CC12.3, and (C) B38 (PDB 7BZ5) (*23*). The RBD is shown in white and antibody residues in cyan, pink, and dark gray, respectively. Oxygen atoms are shown in red and nitrogen atoms in blue. Hydrogen bonds are represented by dashed lines. (**D**) Interaction between ACE2 (green) and residues of the RBD (PDB 6M0J) (*12*) shown in (A) to (C).

None of these antibody interactions mimics ACE2 binding (Fig. 3D).

Our structural analysis reveals two key motifs in the IGHV3-53 germline sequence that are important for RBD binding: an NY motif at V_H residues 32 and 33 in the CDR H1 and an SGGS motif at V_H residues 53 to 56 in the CDR H2 (Fig. 3 and fig. S10). The side chain of V_H N32 in the NY motif hydrogen bonds with the backbone carbonyl of A475 on the RBD, and this interaction is stabilized by an extensive network of hydrogen bonds with other antibody residues as well as a bound water molecule (Fig. 4A). V_H N32 also hydrogen bonds with V_H R94, which in turn hydrogen bonds with N487 and Y489 on the RBD (Fig. 4A). These polar contacts not only enhance the RBD-Fab interaction but also stabilize the CDR conformations with the surrounding residues (framework). V_H Y33 in the NY motif inserts into a hydrophobic cage formed by RBD residues Y421, F456, and L455 and the aliphatic component of K417 (Fig. 4B). A hydrogen bond between V_H Y33 and the carbonyl oxygen of L455 on the RBD further strengthens the interaction. The second key motif, SGGS, in CDR H2 forms an extensive hydrogen bond network with the RBD (Fig. 4C), including four hydrogen bonds that involve the hydroxyl side chains of V_H S53 and V_H S56 and four water-mediated hydrogen bonds to the backbone carbonyl of V_H G54, the backbone amide of V_H S56, and the

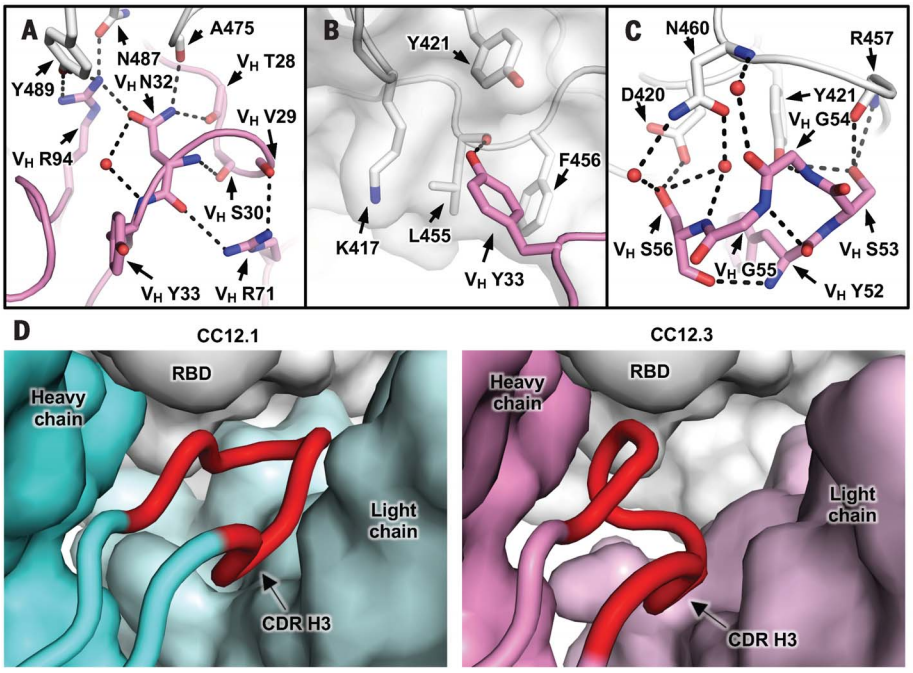


Fig. 4. Two IGHV3-53 germline-encoded motifs with a short CDR H3. (**A**) Illustration of the extensive hydrogen bond network that involves V_H N32 of the NY motif in CDR H1. (**B**) Hydrophobic cage interaction between the RBD and V_H Y33 of the NY motif in CDR H1. (**C**) Hydrogen bond network that involves the SGGS motif in CDR H2. CC12.3 is shown here because its structure is at higher resolution than CC12.1. (**D**) CDR H3 length is constrained to fit in a relatively small pocket on the RBD surface. The heavy and light chains of CC12.1 (cyan) and the RBD (white) are shown in surface representation, with CDR H3 (red) highlighted in the diagram in the left panel. CC12.3 (pink) is shown in the right panel in the same representation.

side chain of V_H S56. Along with V_H Y52, the SGGS motif takes part in a type I beta turn, with a positive Φ angle for V_H G55 at the end of the turn. In addition, the Cα of V_H G54 is only 4 Å away from the RBD, indicating that side chains of other amino acids would clash with the RBD if they were present at this position.

The NY and SGGS motifs, important for RBD binding, are both encoded in the IGHV3-53 germline gene. In addition to IGHV3-53, only the closely related IGHV3-66 contains an NY motif in CDR H1 and an SGGS motif in CDR H2 (*33*). IGHV3-66 is commonly observed (*19*–*22*, *24*, *26*) in antibodies in SARS-CoV-2 patients (Fig. 1A) and is also well represented in the repertoire of healthy individuals (0.3 to 1.7% of total antibodies) (*29*). Overall, our structural analysis has identified two germline-encoded binding motifs that enable IGHV3-53 to target the SARS-CoV-2 RBD, with mutations apparently not required from affinity maturation.

Although the binding mode of CDR H1 and H2 to RBD is highly similar among CC12.1, CC12.3, and B38, the interaction of CDR H3 with the RBD varies (Fig. 3, A to C) because of differences in the CDR H3 sequences and conformations (fig. S1 and Fig. 4D). For example, whereas CDR H3 of CC12.1 interacts with RBD Y453 through a hydrogen bond, CDR H3 of CC12.3 and B38 do not form such a bond (Fig. 3, A to C). Similarly, because of the difference in light-chain gene usage, light-chain interactions with the RBD can vary substantially in IGHV3-53 antibodies (fig. S11). Overall, our structural analysis demonstrates that IGHV3-53 provides a versatile framework with which to target the ACE2-binding site in SARS-CoV-2 RBD.

An interesting feature of CC12.1 and CC12.3 is their relatively short CDR H3. Although the CDR H3 sequences of CC12.1 and CC12.3 differ, both are nine amino acids in length (Kabat numbering) (*34*). Antibody B38 has an even shorter CDR H3 of seven residues (*23*). The average CDR H3 length for human antibodies is ~13 (*35*), although very long CDR H3s (up to 30 residues or more) are found in many broadly neutralizing antibodies to HIV-1 (*36*). Longer CDR H3s cannot be accommodated in these IGHV3-53 antibodies because their epitopes are relatively flat, with only a small pocket to insert their CDR H3 loop (Fig. 4D). A similar conclusion was reached in another recent study, which also reported that SARS-CoV-2 RBD-targeting antibodies encoded by IGHV3-53 or IGHV3-66 tend to have a short CDR H3 (*28*, *37*). Among the RBD-targeting antibodies reported recently (*17*–*28*), IGHV3-53 antibodies consistently have a significantly shorter CDR H3 compared with those encoded by other IGHV genes (P = 6e-8, Mann-Whitney U test) (fig. S12) and are also shorter than IGHV3-53–encoded antibodies in the naïve human antibody repertoire (*30*). Thus, a short CDR H3 length is one of the molecular features of the IGHV3-53–encoded antibody response to SARS-CoV-2 RBD, reminiscent of a below-average five-residue CDR L3 in IGHV1-2 antibodies to the receptor-binding site in HIV-1 Env gp120 (*38*). Nevertheless, a small subset of these IGHV3-53 antibodies have longer CDR H3s that warrant further investigation of their binding mode (fig. S12).

In addition to IGHV3-53, several other IGHV genes, such as IGHV1-2, IGHV3-9, and IGHV3-30, are also more frequently observed than other germ lines in SARS-CoV-2 RBD-targeting antibodies (Fig. 1A). Future work will investigate the molecular mechanisms of these IGHV responses to SARS-CoV-2, as well as whether other germline gene segments, including IGHD and the light chain, contribute in recurring motifs to the SARS-CoV-2 antibody response. The characterization of these IGHV3-53 antibodies to SARS-CoV-2 is a promising starting point for rational vaccine design (*39*), given that limited to no affinity maturation is required to achieve a highly potent neutralizing antibody response to the RBD. Because IGHV3-53 is found at a reasonable frequency in healthy individuals (*29*, *30*), this particular antibody response could be commonly elicited during vaccination (*40*).

REFERENCES AND NOTES

1. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* **368**, 860–868 (2020).
2. N. Lurie, M. Saville, R. Hatchett, J. Halton, *N. Engl. J. Med.* **382**, 1969–1973 (2020).
3. S. F. Andrews, A. B. McDermott, *Curr. Opin. Immunol.* **53**, 96–101 (2018).
4. P. Parameswaran *et al.*, *Cell Host Microbe* **13**, 691–700 (2013).
5. K. Pieper *et al.*, *Nature* **548**, 597–601 (2017).
6. I. Setliff *et al.*, *Cell Host Microbe* **23**, 845–854.e6 (2018).
7. N. C. Wu, I. A. Wilson, *Nat. Struct. Mol. Biol.* **25**, 115–121 (2018).
8. A. Lanzavecchia, A. Frühwirth, L. Perez, D. Corti, *Curr. Opin. Immunol.* **41**, 62–67 (2016).
9. P. Zhou *et al.*, *Nature* **579**, 270–273 (2020).
10. M. Letko, A. Marzi, V. Munster, *Nat. Microbiol.* **5**, 562–569 (2020).
11. R. Yan *et al.*, *Science* **367**, 1444–1448 (2020).
12. J. Lan *et al.*, *Nature* **581**, 215–220 (2020).
13. W. Li *et al.*, bioRxiv 2020.05.13.093088 [Preprint] (2 June 2020); https://doi.org/10.1101/2020.05.13.093088.
14. E. Andreano *et al.*, bioRxiv 2020.05.05.078154 [Preprint] (9 May 2020); https://doi.org/10.1101/2020.05.05.078154.
15. S. J. Zost *et al.*, *Nat. Med.* 10.1038/s41591-020-0998-x (2020).
16. A. Z. Wec *et al.*, *Science* eabc7424 (2020).
17. M. Yuan *et al.*, *Science* **368**, 630–633 (2020).
18. D. Pinto *et al.*, *Nature* **583**, 290–295 (2020).
19. B. Ju *et al.*, bioRxiv 2020.03.21.990770 [Preprint] (26 March 2020); https://doi.org/10.1101/2020.03.21.990770.
20. Y. Cao *et al.*, *Cell* 10.1016/j.cell.2020.05.025 (2020).
21. P. J. M. Brouwer *et al.*, *Science* eabc5902 (2020).
22. D. F. Robbiani *et al.*, *Nature* 10.1038/s41586-020-2456-9 (2020).
23. Y. Wu *et al.*, *Science* **368**, 1274–1278 (2020).
24. X. Chi *et al.*, *Science* **369**, 650–655 (2020).
25. E. Seydoux *et al.*, *Immunity* 10.1016/j.immuni.2020.06.001 (2020).
26. R. Shi *et al.*, *Nature* 10.1038/s41586-020-2381-y (2020).
27. T. F. Rogers *et al.*, *Science* eabc7520 (2020).
28. C. O. Barnes *et al.*, *Cell* 10.1101/2020.05.28.121533 (2020).
29. S. D. Boyd *et al.*, *J. Immunol.* **184**, 6986–6992 (2010).

30. B. Briney, A. Inderbitzin, C. Joyce, D. R. Burton, *Nature* **566**, 393–397 (2019).
31. The highly related IGHV3-66–encoded RBD-targeting antibodies are also frequently found and reported in five out of 12 studies and in 11 of 32 COVID-19 patients (*17*–*28*).
32. J. Ye, N. Ma, T. L. Madden, J. M. Ostell, *Nucleic Acids Res.* **41**, W34-40 (2013).
33. The amino-acid sequences of IGHV3-53 and IGHV3-66 germline genes only differ by one amino acid at residue 12. Specifically, IGHV3-53 contains Ile12, whereas IGHV3-66 contains Val12.
34. CDRH3 lengths with Kabat numbering are two amino acids shorter than those with IMGT numbering.
35. M. Zemlin *et al*., *J. Mol. Biol.* **334**, 733–749 (2003).
36. L. Yu, Y. Guan, *Front. Immunol.* **5**, 250 (2014).
37. IGHV3-53 and IGHV3-66 antibodies in general have slightly shorter than average CDR H3s (by around one residue) and also appear to have a few much shorter CDR H3s (<10 amino acids) than average in the baseline antibody repertoire (*30*).
38. J. G. Jardine *et al*., *Science* **349**, 156–161 (2015).
39. D. R. Burton, L. M. Walker, *Cell Host Microbe* **27**, 695–698 (2020).
40. R. K. Abbott, S. Crotty, *Immunol. Rev.* 10.1111/imr.12861 (2020).
41. E. Krissinel, K. Henrick, *J. Mol. Biol.* **372**, 774–797 (2007).

ACKNOWLEDGMENTS

We thank R. Stanfield for assistance in data collection, B. Briney for naïve antibody germline analysis, and the staff of Stanford Synchrotron Radiation Laboratory (SSRL) Beamline 12-1 for assistance. **Funding:** This work was supported by the NIH (grant no. K99 AI139445 to N.C.W.), the Bill and Melinda Gates Foundation (grant no. OPP1170236 to I.A.W. and D.R.B.), NIH NIAID CHAVD (grant no. UM1 AI44462 to I.A.W., D.S., and D.R.B.), and the IAVI Neutralizing Antibody Center. Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02–76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). **Author contributions:** M.Y., H.L., N.C.W., F.Z., D.H., T.F.R., E.L., D.S, J.G.J., D.R.B., and I.A.W. conceived and designed the study. F.Z., D.H., T.F.R., E.L., D.S., J.G.J., and D.R.B. isolated the antibodies and provided clones and sequences. M.Y., H.L., N.C.W., C.-C.D.L., W.Y., and Y.H. expressed and purified the proteins. M.Y. and C.-C.D.L. performed biolayer interferometry binding assays. M.Y., H.L., N.C.W., X.Z., and H.T. performed the crystallization and x-ray data collection. M.Y. and X.Z. determined and refined the x-ray structures. M.Y., H.L., N.C.W., C.-C.D.L., and X.Z. analyzed the data. M.Y., H.L., N.C.W., and I.A.W. wrote the manuscript, and all authors reviewed and/or edited the manuscript. **Competing interests:** D.R.B., D.H., J.G.J., E.L., T.F.R., D.S., and F.Z. are listed as inventors on pending patent applications describing the SARS-CoV-2 antibodies originally published in Rogers *et al*. (*27*). **Data and materials availability:** X-ray coordinates and structure factors are deposited in the RCSB Protein Data Bank under identification nos. 6XC2, 6XC3, 6XC4, and 6XC7. Antibody plasmids are available from D.R.B. under a material transfer agreement from The Scripps Research Institute. Other materials are available from I.A.W. on request. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1119/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S3
References (*42*–*50*)
MDAR Reproducibility Checklist

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

9 June 2020; accepted 7 July 2020
Published online 13 July 2020
10.1126/science.abd2321

FLAVIVIRUSES

Zika virus infection enhances future risk of severe dengue disease

Leah C. Katzelnick[1], César Narvaez[2], Sonia Arguello[2], Brenda Lopez Mercado[2], Damaris Collado[2], Oscarlett Ampie[2], Douglas Elizondo[2], Tatiana Miranda[2], Fausto Bustos Carillo[1], Juan Carlos Mercado[3,2], Krista Latta[4], Amy Schiller[4], Bruno Segovia-Chumbez[5], Sergio Ojeda[2], Nery Sanchez[2], Miguel Plazaola[2], Josefina Coloma[1], M. Elizabeth Halloran[6,7], Lakshmanane Premkumar[5], Aubree Gordon[4], Federico Narvaez[2], Aravinda M. de Silva[5], Guillermina Kuan[8,2], Angel Balmaseda[3,2], Eva Harris[1]*

The Zika pandemic sparked intense interest in whether immune interactions among dengue virus serotypes 1 to 4 (DENV1 to -4) extend to the closely related Zika virus (ZIKV). We investigated prospective pediatric cohorts in Nicaragua that experienced sequential DENV1 to -3 (2004 to 2015), Zika (2016 to 2017), and DENV2 (2018 to 2020) epidemics. Risk of symptomatic DENV2 infection and severe disease was elevated by one prior ZIKV infection, one prior DENV infection, or one prior DENV infection followed by one ZIKV infection, compared with being flavivirus-naïve. By contrast, multiple prior DENV infections reduced dengue risk. Further, although high preexisting anti-DENV antibody titers protected against DENV1, DENV3, and ZIKV disease, intermediate titers induced by previous ZIKV or DENV infection enhanced future risk of DENV2 disease and severity, as well as DENV3 severity. The observation that prior ZIKV infection can modulate dengue disease severity like a DENV serotype poses challenges to development of dengue and Zika vaccines.

Dengue virus serotypes 1 to 4 (DENV1 to -4) and Zika virus (ZIKV) are closely related mosquito-borne flaviviruses with high global burdens (*1*, *2*). Dengue epidemics often overwhelm health care systems as medical staff respond to life-threatening manifestations of severe dengue disease, including vascular leak syndrome and shock (*3*). ZIKV spread across the Pacific and Americas in 2013 to 2017 and caused rare but devastating clinical outcomes, including congenital microcephaly and Guillain-Barré syndrome in adults (*2*, *4*). Vaccines against both dengue and Zika are undergoing clinical evaluation (*4*, *5*). However, the only licensed dengue vaccine, Dengvaxia, increases risk of severe dengue in previously DENV-naïve individuals (*6*). Other dengue vaccine candidates are being evaluated for possible differences in safety and efficacy on the basis of DENV infection history. There remains concern that ZIKV infection or Zika vaccines could also enhance subsequent dengue disease.

A prior DENV infection is an established risk factor for future symptomatic and severe dengue during infection with a different serotype (*7*, *8*). A first DENV or ZIKV infection induces antibodies that limit disease upon reinfection with the same virus but also generates nonprotective antibodies that bind other viruses (*9*). DENV cross-reactive antibodies can facilitate heterologous DENV infection of myeloid cells by means of antibody-dependent enhancement (ADE) and can increase dengue disease severity in humans (*10*–*14*). Many cross-reactive antibodies target epitopes conserved across flaviviruses, including the envelope protein fusion loop (*15*, *16*). $CD4^+$ and $CD8^+$ T cell responses and nonstructural protein 1 (NS1) also modulate DENV and ZIKV protection and pathogenesis (*17*, *18*).

Emerging evidence suggests that prior DENV infection may not enhance noncongenital Zika disease, but whether prior ZIKV infection increases future dengue disease in humans is unknown. In vitro and mouse challenge studies have shown that antibodies raised against DENV can enhance ZIKV infection (*19*, *20*). However, prior DENV infection was not associated with ZIKV viremia or cytokine expression in experimentally challenged macaques (*21*–*23*) or in humans (*24*–*26*), nor with fetal demise or congenital Zika syndrome in pregnant women (*27*). Further, prior DENV infection was protective against uncomplicated Zika in prospective cohort studies (*28*, *29*). Prior DENV infection was also associated with stronger cytotoxic CD8 T cell responses in ZIKV-infected humans and with protection against ZIKV in mice (*30*). In contrast, anti-ZIKV antibodies increased DENV2 infection, viral output, and migration of myeloid cells in skin explants to the same degree as anti-DENV3 antibodies (*31*). In murine models, transfer of anti-ZIKV antibodies caused greater clinical severity, mortality, proinflammatory cytokine levels, and viral load after DENV2 challenge compared with untreated mice (*32*, *33*). In macaques, prior ZIKV infection produced binding but non-neutralizing antibodies to DENV2, and challenge with DENV2 resulted in elevated viral load and hematological changes associated with severe dengue, although not in all studies (*23*, *34*, *35*).

The Zika epidemic was followed by several years of low DENV transmission, but in 2019, countries across Latin America reported a major resurgence of dengue cases. Since 2004, we have followed an active cohort of ~3800 children 2 to 16 years old living in Managua, Nicaragua, for DENV infection and disease (*36*, *37*). As chikungunya virus (CHIKV) and ZIKV were introduced into Nicaragua in 2014 and 2016, respectively, the cohort was extended to capture cases and infections with these emerging arboviruses (*28*). In 2019 to 2020, Nicaragua experienced the largest dengue epidemic in recorded history (Fig. 1A). An unprecedented number of cohort participants (n = 375) experienced a symptomatic DENV infection. All virologically confirmed dengue cases (n = 293) were caused by DENV2.

The epidemiology and longevity of the Nicaraguan cohort allowed us to test whether ZIKV infection modified subsequent risk of dengue disease. All children presenting with suspected dengue, undifferentiated febrile illness and, after 2016, afebrile rash were tested for DENV, ZIKV, and CHIKV infection by real-time reverse transcription–polymerase chain reaction (RT-PCR) of acute-phase samples and by serological assays run on paired acute and convalescent samples. Since 2004, cohort participants have provided healthy annual serum and/or plasma samples that have been tested for DENV-specific antibodies (DENV-Abs) using the DENV inhibition enzyme-linked immunosorbent assay (iELISA) (n = 8399 children, 57,963 samples tested) (Fig. 1B and table S1) (*12*, *37*). Paired annual samples are tested side by side for seroconversion or a ≥4-fold rise in the DENV iELISA titer (*12*). Since 2016, samples are screened for ZIKV infection and ZIKV-Abs using the ZIKV iELISA (Fig. 1C) and ZIKV NS1 blockade-of-binding assays (n = 14,159 and 14,247 measurements, respectively) (*37*, *38*). In each assay, a titer or percent inhibition is derived by measuring the serum antibody concentration required to compete for binding to the antigen of interest with an antigen-specific, labeled antibody. This detailed information on infections and cases is used to identify complete DENV and ZIKV infection histories for cohort participants (*37*) (Fig. 1D).

We estimated the probability of symptomatic DENV2 infection in the cohort during the 2019 to 2020 epidemic on the basis of ZIKV and DENV infection histories by using log-binomial

[1]Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, Berkeley, CA, USA. [2]Sustainable Sciences Institute, Managua, Nicaragua. [3]Laboratorio Nacional de Virología, Centro Nacional de Diagnóstico y Referencia, Ministry of Health, Managua, Nicaragua. [4]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA. [5]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC, USA. [6]Department of Biostatistics, University of Washington, Seattle, WA, USA. [7]Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [8]Centro de Salud Sócrates Flores Vivas, Ministry of Health, Managua, Nicaragua.
*Corresponding author. Email: eharris@berkeley.edu

generalized linear models (GLMs) adjusted for age and sex. A total of 8.8% of all cohort participants experienced symptomatic DENV infections in 2019 to 2020 (n = 302 cases meeting dengue case definition, of n = 3434 cohort participants with full infection histories). Unexpectedly, children with one prior ZIKV infection had a 12.1% probability of having a symptomatic DENV2 infection in 2019 to 2020 [confidence interval (CI) 9.9 to 14.5%], compared with only 3.5% (CI 2.4 to 4.6%) of flavivirus-naïve children (Fig. 2A and table S2), and similar to the 9.2% (CI 4.6 to 14.5%) of children with one prior DENV infection. The increased risk of dengue disease with one prior ZIKV infection remained when CIs were estimated using alternative modeling approaches; after adjustment for years since previous infection, for neighborhood-level risk of flavivirus infection, and for prior symptomatic ZIKV infection; and when afebrile dengue cases were included in analyses (figs. S1 to S5 and table S3).

After sequential DENV infections, individuals are thought to be at reduced risk of future dengue disease (*9*). Consistent with previous findings, children with two prior DENV infections had only a 2.5% (CI 0.0 to 9.0%) probability of symptomatic DENV2 infection. In contrast, children with one DENV infection followed by ZIKV infection remained at significantly elevated risk, with a 9.5% (CI 6.7 to 13.0%) probability of symptomatic DENV2 infection (Fig. 2A). For children with at least two prior DENV infections, the probability of dengue disease was low for those with (2.9%, CI 0.7 to 6.2%) or without (0.0%, CI 0.0 to 0.0%) a subsequent ZIKV infection.

A history of ZIKV infection was also a significant risk factor for severe dengue disease. The probability of experiencing dengue with warning signs or severe dengue (DwWS/SD) [2009 World Health Organization (WHO) criteria; n = 144] was significantly elevated for individuals with one DENV infection (5.4%, CI 2.0 to 9.6%), one ZIKV infection (5.9%, CI 4.3 to 7.7%), or one DENV and one ZIKV infection (4.8%, CI 2.9 to 7.0%) compared with flavivirus-naïve children (0.7%, CI 0.3 to 1.2%) (Fig. 2B and table S4). The risk of dengue hemorrhagic fever or dengue shock syndrome (DHF/DSS) (1997 WHO criteria; n = 15) was also unusually high (*39*) and was significantly greater for children with one prior ZIKV infection (probability, 1.1%; CI 0.3 to 1.8%) and one DENV and one ZIKV infection (0.9%, CI 0 to 3.3%) compared with flavivirus-naïve children (0%) (Fig. 2C and table S5). These relationships held for individual manifestations of severe dengue disease (fig. S6). In contrast, multiple prior DENV infections did not enhance dengue disease severity.

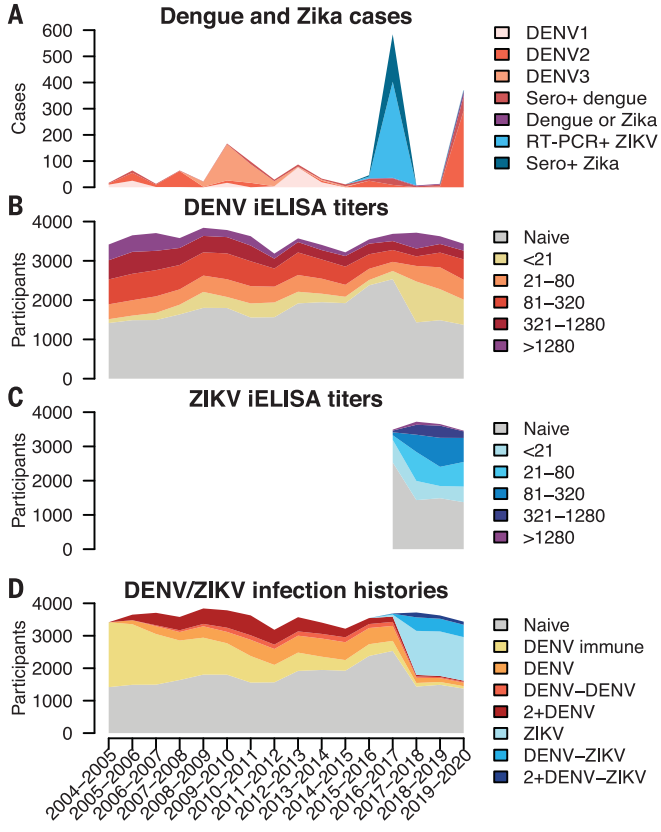


Fig. 1. Dengue and Zika cases, DENV-Ab and ZIKV-Ab titers, and infection histories in the PDCS (2004 to 2020). (**A**) Confirmed dengue and Zika cases by epidemic season and infecting virus. (**B** to **D**) DENV iELISA titers (B), ZIKV iELISA titers (C), and DENV and ZIKV infection histories (D) for cohort participants, measured at the beginning of each epidemic season. Ab titers were grouped by reciprocal serum dilution, ranging from <21 to >1280. Infection histories were grouped as follows: flavivirus-naïve (naive), entered cohort DENV-immune without subsequent infections (DENV immune), entered flavivirus-naïve with one DENV (DENV) or ZIKV (ZIKV) infection, one prior DENV infection followed by a DENV (DENV-DENV) or ZIKV (DENV-ZIKV) infection, ≥2 prior DENV infections without (2+DENV) or with (2+DENV-ZIKV) a subsequent ZIKV infection.

We further tested whether history of ZIKV infection increased risk of dengue disease in a separate study based at the Nicaraguan National Pediatric Reference Hospital. Since 2005, the study has followed 5832 children 6 months to 14 years of age during and after presentation to the hospital for suspected dengue, Zika, or chikungunya. In 2018 and 2019, 388 children were enrolled in the hospital study, and all virologically confirmed cases were caused by DENV2 (n = 277). For a random subset of the hospital study's DENV2 cases (n = 88), we analyzed acute- and convalescent-phase samples (days 1 to 5 and 12 to 26 post–symptom onset, respectively) by the ZIKV and DENV E-domain III ELISA and DENV iELISA (*40*). In combination, these assays had 96% sensitivity and 96% specificity for detecting prior ZIKV infection in a separate set of children from the Pediatric Dengue Cohort Study (PDCS) with known infection histories (n = 53) (*37*). Compared with the full PDCS cohort (53%), significantly more cohort (76%) and hospital study (86%) dengue cases had a prior ZIKV infection, with or without a prior DENV infection (Kruskal-Wallis rank sum test, $P < 2.2 \times 10^{-16}$). By contrast, for those with prior DENV infection only, there were no differences between groups (P = 0.25), suggesting that ZIKV infection history helped explain the higher rate of disease (Fig. 2D). Further, among DENV2 cases in the cohort only or cohort and hospital studies (Fig. 2, E and F; and table S6; logistic regression adjusted for age and sex), the probabilities of experiencing DwWS/SD or DHF/DSS were significantly greater for children with histories of ZIKV with or without DENV (DHF/DSS: 11.6%, CI 6.2 to 17.2%), compared with naïve children (0.0%) and similar to the rate for those with only prior DENV infection (14.0%, CI 2.5 to 31.4%).

We previously showed that the level of preexisting DENV-Abs correlates with subsequent symptomatic and severe dengue disease (*12*, *14*). Here, we found that children with the same number of prior DENV or DENV and ZIKV infections had similar levels of preexisting DENV-Ab titers (Fig. 2G), whereas preexisting ZIKV-Ab titers were higher for children with prior ZIKV infection, with or without prior DENV infection (Fig. 2H), consistent with previous studies (*41*–*43*). We used log-binomial GLMs to test whether preexisting DENV-Abs, including cross-reactive antibodies induced by prior ZIKV infection, and preexisting ZIKV-Abs were associated with dengue disease risk in 2019 to 2020. Children with any level of preexisting cross-reactive DENV-Abs were at significantly greater risk of symptomatic DENV2 than flavivirus-naïve children (Fig. 2I). Those with a range of intermediate preexisting DENV-Abs were at the greatest risk of symptomatic DENV2 disease (titers of 21 to 80: 13.7%, CI 10.3 to 17.4; naïve: 3.5%, CI 2.4 to 4.6) (Fig. 2I), DwWS/SD (titers of 21 to

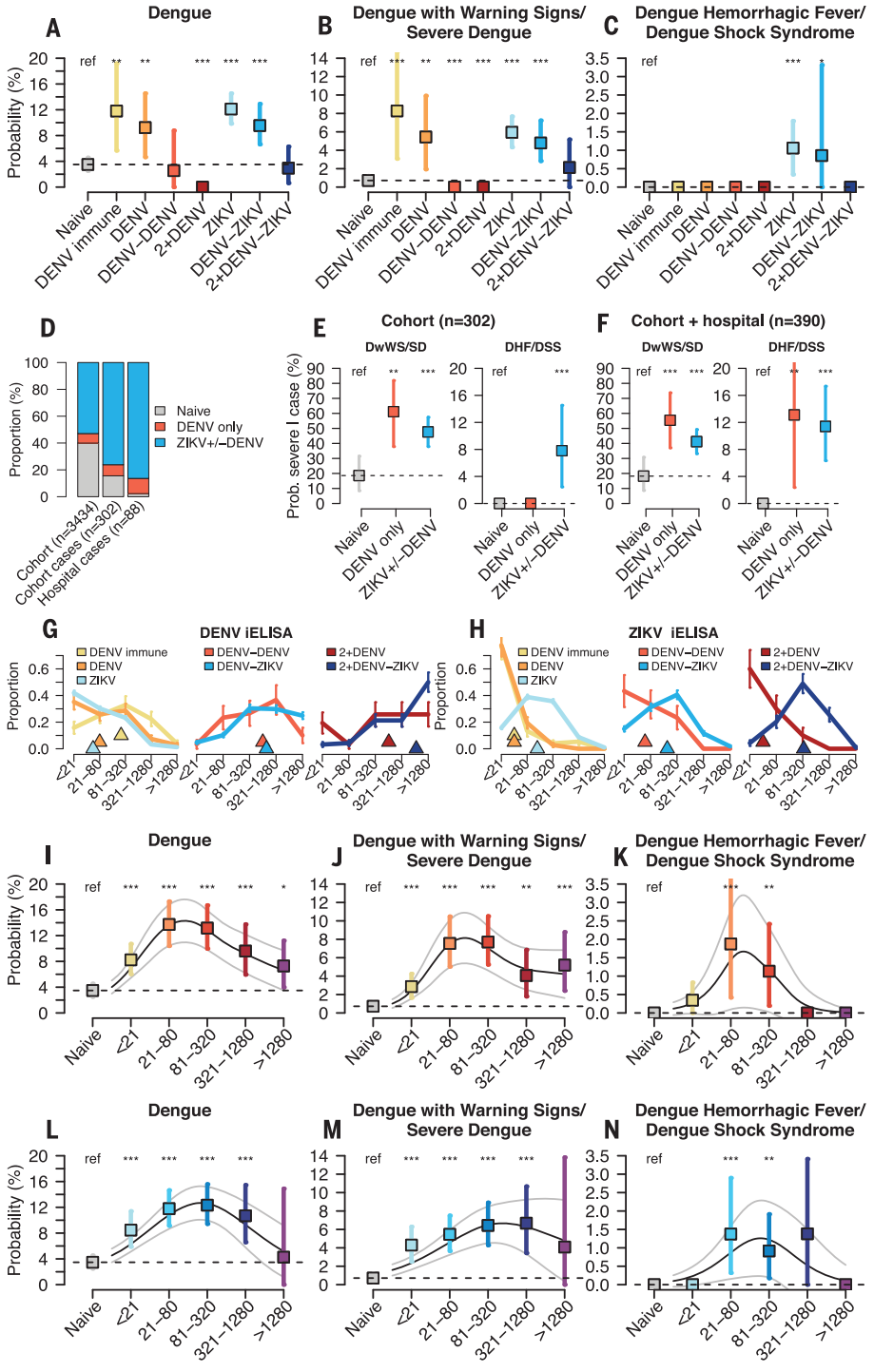


Fig. 2. Probability of symptomatic and severe DENV2 infection by prior DENV and ZIKV infection history and preexisting antibody titer, 2019 to 2020. (**A** to **C**) Log-binomial GLMs were used to estimate the probability of dengue disease in 2019 to 2020 in the cohort study on the basis of DENV and ZIKV infection history with bootstrap resampling (n = 10,000) to construct 95% CIs and calculate P values. (**D**) Infection histories for all children in the cohort, for cohort dengue cases, and for hospital study dengue cases. Differences in the proportion with histories of prior ZIKV and/or DENV or DENV only were tested with a Kruskal-Wallis rank sum test. (**E** and **F**) Probability of severe dengue disease among confirmed dengue cases in cohort (E) or cohort and hospital studies (F) by infection history, estimated using logistic regression. Bootstrap resampling (n = 10,000) was used to construct 95% CIs and calculate P values. (**G** and **H**) DENV iELISA (G) and ZIKV iELISA (H) titer distributions (grouped by reciprocal serum dilution, from <21 to >1280) for cohort participants in 2019 to 2020 by DENV and ZIKV infection history (triangles show median values, vertical bars show ±1 standard deviation). (**I** to **K**) Log-binomial GLMs were used to estimate the probability of dengue disease in 2019 to 2020 in the cohort study on the basis of preexisting DENV iELISA titers. (**L** to **N**) Log-binomial GLMs were used to estimate the probability of dengue disease in 2019 to 2020 in the cohort study on the basis of preexisting ZIKV iELISA titers. Continuous relationships between titers and disease were modeled with log-binomial generalized additive models (GAMs; black lines show probabilities, gray lines show 95% CIs). All models were adjusted for age and sex, and probabilities are shown for an average study participant (male, age 8). *P < 0.05; **P < 0.01; ***P < 0.001. P values indicate significantly different probability estimates from the naïve group.

80: 7.5%, CI 5.0 to 10.4%; naïve: 0.7%, CI 0.3 to 1.2%) (Fig. 2J), and DHF/DSS (titers of 21 to 80: 1.9%, CI 0.4 to 3.9%; naïve: 0%) (Fig. 2K). Intermediate preexisting ZIKV-Ab titers were also associated with enhancement of symptomatic and severe dengue (Fig. 2, L to N).

Epidemiological evidence suggests that prior DENV immunity differentially affects disease caused by each DENV serotype. For instance, DENV2 and DENV4 more commonly manifest as symptomatic or severe disease in secondary infections than do DENV1 and, in some studies, DENV3 (*7*, *8*, *44*, *45*). Immune correlate analyses in natural infection and vaccine studies have shown that high preexisting neutralizing antibody titers protect against DENV1 and DENV3 but not necessarily DENV2 (*46*–*48*). Consistent with these observations, we found that in the pre-Zika era (2004 to 2015), children with intermediate preexisting DENV-Ab titers also had increased probability of symptomatic (0.8%, CI 0.5 to 1.1; vs. naïve, 0.3%, CI 0.2 to 0.4) or severe dengue disease caused by DENV2 (DwWS/SD: 0.3%, CI 0.1 to 0.6% vs. naïve: 0.02%, CI 0.01 to 0.06; DHF/DSS: 0.1%, CI 0.03 to 0.3%, vs. naive, 0%) (Fig. 3, A to C; and table S7). The magnitude of the enhancing effect was greater in the post-Zika era, due in part to the higher incidence of dengue cases in 2019 to 2020 (Fig. 2, I to K; vs. Fig. 3, A to C). In the pre-Zika era, low and

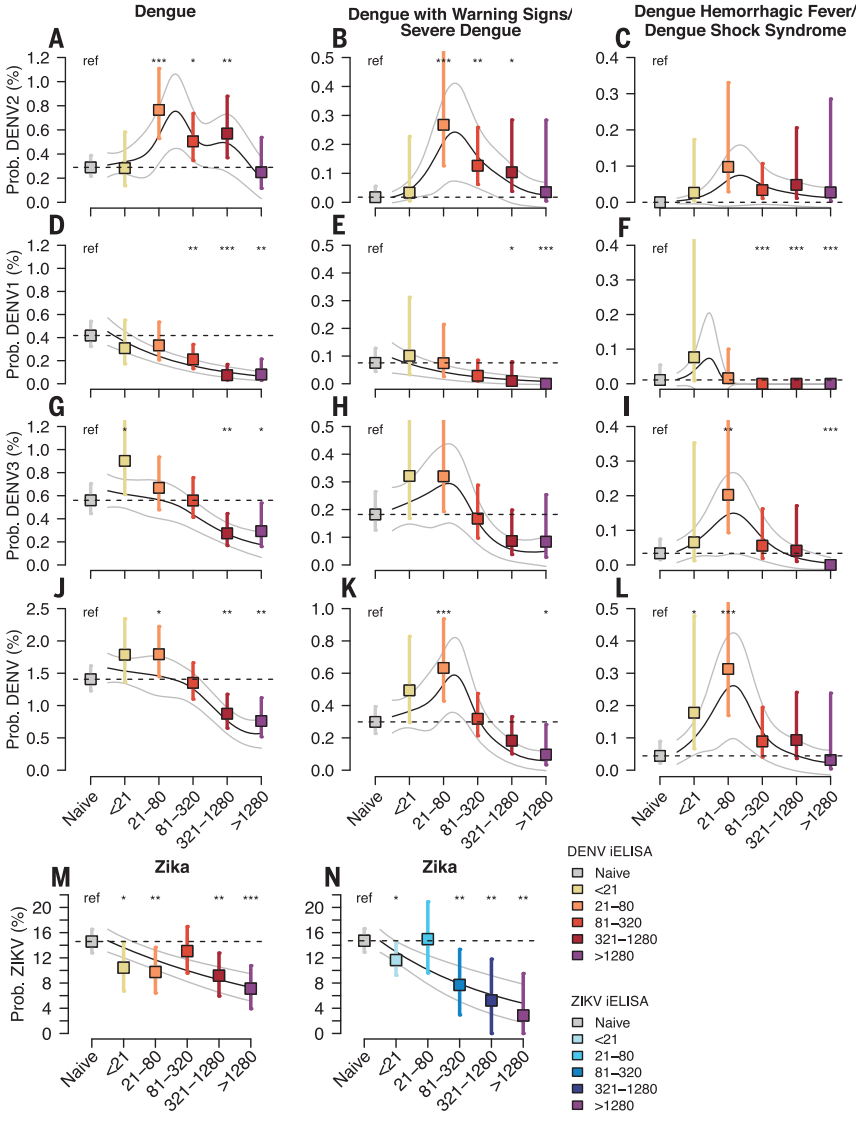


Fig. 3. Probability of disease caused by DENV2, DENV1, DENV3, all DENV, and ZIKV infection by preexisting DENV and ZIKV iELISA titers, 2004 to 2017. Probability of each disease outcome for each infecting virus was modeled as a function of preexisting antibody titer on both discrete (colored bars) and continuous (black lines) scales, shown with 95% CIs. All continuous relationships were modeled using log-binomial GAMs. (**A** to **L**) Probabilities of dengue [(A), (D), (G), and (J)], DwWS/SD [(B), (E), (H), and (K)], and DHF/DSS [(C), (F), (I), and (L)] by discrete DENV iELISA titer bins were modeled separately for [(A) to (C)] DENV2, [(D) to (F)] DENV1, [(G) to (I)] DENV3, and [(J) to (L)] all serotypes in the pre-Zika era (2004 to 2015) (J to L) using generalized estimating equation log-binomial models. Point estimates and confidence intervals correspond to probabilities, and *P* values correspond to iELISA titer bins with significantly different relative risk compared with the naive group. (**M** and **N**) Probabilities of Zika (2016) by preinfection DENV iELISA (M) and ZIKV iELISA (N) titer bins were modeled using log-binomial GLMs, using bootstrap resampling (n = 10,000 to calculate 95% CIs and *P* values.). All models were adjusted for age and sex, and model estimates are shown for an average study participant (male, age 8). $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

intermediate DENV-Ab titers also had an enhancing effect on symptomatic and severe DENV3 infection (0.2%, CI 0.09 to 0.4%; vs. naïve: 0.03%, CI 0.01 to 0.08) (Fig. 3, G to I) but not on symptomatic or severe DENV1 infection (Fig. 3, D to F). By contrast, high preexisting DENV-Abs had a protective effect against symptomatic DENV1 and DENV3 infections (Fig. 3, D and G; and table S7). When all DENV cases were analyzed simultaneously, an overall protective effect was observed against symptomatic disease, whereas the enhancing effect increased as the definition of severity narrowed to DHF/DSS (Fig. 3, J to L), as we observed previously (*12*).

Similar to observations for DENV1, high preexisting DENV-Ab and ZIKV-Ab titers were associated with reduced probability of uncomplicated Zika during the 2016 Zika epidemic (Fig. 3, M and N; and table S8) (we did not observe severe Zika in the cohort). Consistent with our previous epidemiological findings (*28*), one or multiple prior DENV infections were protective against Zika (table S8). However, intermediate DENV-Ab and ZIKV-Ab titers were not significantly protective, possibly indicating more complex relationships (alternative titer bins are provided in fig. S8).

We find that prior ZIKV infection modulates future dengue disease risk to a similar degree as prior infection with a DENV serotype. A single prior ZIKV infection, like one prior DENV infection, increases the probability of symptomatic and severe dengue disease caused by DENV2. Further, one DENV followed by one ZIKV infection also increased future risk of dengue disease, unlike sequential DENV infections, which reduced future risk, suggesting an important difference between secondary flavivirus infection with ZIKV versus a DENV serotype. Our findings also show that the relationship between preexisting anti-flavivirus antibodies and disease depends on the secondary infecting virus. Intermediate preexisting cross-reactive DENV-Ab or ZIKV-Ab titers enhance risk of DENV2 and DENV3 disease severity, but not that of DENV1 or ZIKV. High titers protect against symptomatic DENV1, DENV3, and ZIKV, but not DENV2 infection. Thus, we find that asymmetry exists among DENV serotypes and between DENV and ZIKV infections.

On the basis of our findings and previous literature on DENV1 to -4 (*7*, *8*, *44*–*48*), we posit that prior ZIKV infection, like prior DENV infection, is particularly capable of enhancing DENV2 disease but that enhancement of other serotypes is possible. Mechanistic studies in animal models and human skin explants have shown that prior primary ZIKV infection induces anti-DENV2 antibodies that facilitate classical ADE of infection and increase disease severity during DENV2 challenge (*23*, *31*–*34*). In humans, primary ZIKV infection induces lower heterologous neutralizing antibody titers to DENV1 to -4 than primary DENV1 to -3 infection (*41*, *42*), suggesting the potential for enhancement of multiple serotypes (fig. S9). Further, although secondary ZIKV infection induces highly specific ZIKV-Abs and boosts heterologous DENV binding antibodies, ZIKV infection does not induce the broadly cross-neutralizing antibodies observed after secondary DENV infection (fig. S9) (*41*, *42*, *49*). However, memory B cells and monoclonal antibodies with high DENV1-, DENV2-, and DENV3-ZIKV cross-neutralization activities have been isolated, suggesting cross-protection is possible (*42*, *43*, *50*). The tighter structure of the ZIKV virion may modify the types of antibodies ZIKV induces

and limit their neutralizing potency against DENV, as well as reduce the ability of cross-reactive DENV antibodies to enhance ZIKV infection (*51*, *52*). In addition, anti-ZIKV cytotoxic T cells mostly target structural protein epitopes, which may not protect against dengue, as anti-DENV $CD8^+$ T cells mostly target conserved regions of nonstructural proteins (*53*, *54*). Finally, whether anti-ZIKV immunity enhances other DENV serotypes may depend on characteristics specific to the secondary infection, including virion structure and maturation state, cell infection mechanism, and fucosylation of immunoglobulin G (*51*, *52*, *55*, *56*).

Our findings suggest that protective and pathogenic interactions between DENV1 to -4 and ZIKV could affect vaccine efficacy and safety. If monovalent Zika vaccines induce cross-reactive DENV antibodies such as those observed after natural ZIKV infection, Zika vaccines could increase risk of subsequent symptomatic and severe dengue disease. In contrast, we found that DENV-Abs induced by natural DENV infection modestly protect against uncomplicated Zika, but the effect on vertically transmitted ZIKV in humans and risk of congenital Zika syndrome requires additional study. Further, differences in efficacy in the phase 3 trials of multiple dengue vaccines may be explained not only by lack of protection but also by enhancement of certain serotypes (*6*). Our findings demonstrate why it is important to analyze separately correlates of protection and risk for dengue caused by specific serotypes. We have illustrated critical immunological interactions among ZIKV and the DENV serotypes and revealed differences in the relationship between preexisting cross-reactive antibodies and serotype-specific risk of symptomatic and severe disease. Elucidating how immunity to DENV1 to -4, ZIKV, and possibly other flaviviruses modulates future disease risk is of utmost concern for developing and deploying safe, effective flavivirus vaccines and preventing future epidemics.

REFERENCES AND NOTES

1. L. Cattarino, I. Rodriguez-Barraquer, N. Imai, D. A. T. Cummings, N. M. Ferguson, *Sci. Transl. Med.* **12**, eaax4144 (2020).
2. N. R. Faria *et al.*, *Nature* **546**, 406–410 (2017).
3. WHO/TDR, *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control* (WHO, 2009).
4. M. S. Diamond, J. E. Ledgerwood, T. C. Pierson, *Annu. Rev. Med.* **70**, 121–135 (2019).
5. M. Redoni *et al.*, *Rev. Med. Virol.* **30**, e2101 (2020).
6. S. Sridhar *et al.*, *N. Engl. J. Med.* **379**, 327–340 (2018).
7. A. Nisalak *et al.*, *Am. J. Trop. Med. Hyg.* **94**, 1342–1347 (2016).
8. N. Sangkawibha *et al.*, *Am. J. Epidemiol.* **120**, 653–669 (1984).
9. E. N. Gallichotte, R. S. Baric, A. M. de Silva, *Adv. Exp. Med. Biol.* **1062**, 63–76 (2018).
10. S. B. Halstead, *J. Infect. Dis.* **140**, 527–533 (1979).
11. S. C. Kliks, A. Nisalak, W. E. Brandt, L. Wahl, D. S. Burke, *Am. J. Trop. Med. Hyg.* **40**, 444–451 (1989).
12. L. C. Katzelnick *et al.*, *Science* **358**, 929–932 (2017).
13. H. Salje *et al.*, *Nature* **557**, 719–723 (2018).
14. J. J. Waggoner *et al.*, *J. Infect. Dis.* **221**, 1846–1854 (2020).
15. W. Dejnirattisai *et al.*, *Science* **328**, 745–748 (2010).
16. R. de Alwis *et al.*, *PLOS Pathog.* **10**, e1004386 (2014).
17. A. E. Ngono, S. Shresta, *Annu. Rev. Immunol.* **36**, 279–308 (2018).
18. P. R. Beatty *et al.*, *Sci. Transl. Med.* **7**, 304ra141 (2015).
19. S. V. Bardina *et al.*, *Science* **356**, 175–180 (2017).
20. M. G. Zimmerman *et al.*, *Cell Host Microbe* **24**, 731–742.e6 (2018).
21. P. Pantoja *et al.*, *Nat. Commun.* **8**, 15674 (2017).
22. M. K. McCracken *et al.*, *PLOS Pathog.* **13**, e1006487 (2017).
23. M. E. Breitbach *et al.*, *PLOS Pathog.* **15**, e1007766 (2019).
24. A. C. B. Terzian *et al.*, *Clin. Infect. Dis.* **65**, 1260–1265 (2017).
25. G. A. Santiago *et al.*, *Open Forum Infect. Dis.* **6**, ofz320 (2019).
26. D. Michlmayr *et al.*, *Cell Rep.* **31**, 107569 (2020).
27. U.-A. Halai *et al.*, *Clin. Infect. Dis.* **65**, 877–883 (2017).
28. A. Gordon *et al.*, *PLOS Med.* **16**, e1002726 (2019).
29. I. Rodriguez-Barraquer *et al.*, *Science* **363**, 607–610 (2019).
30. J. Wen *et al.*, *Nat. Commun.* **8**, 1459 (2017).
31. P. M. S. Castanha *et al.*, *JCI Insight* **5**, e133653 (2020).
32. A. M. Fowler *et al.*, *Cell Host Microbe* **24**, 743–750.e5 (2018).
33. S. Watanabe, N. W. W. Tan, K. W. K. Chan, S. G. Vasudevan, *J. Infect. Dis.* **219**, 223–233 (2019).
34. J. George *et al.*, *Sci. Rep.* **7**, 10498 (2017).
35. E. X. Pérez-Guzmán *et al.*, *Nat. Commun.* **10**, 4316 (2019).
36. G. Kuan *et al.*, *Am. J. Epidemiol.* **170**, 120–129 (2009).
37. Materials and methods are available as supplementary materials.
38. A. Balmaseda *et al.*, *J. Clin. Microbiol.* **56**, e01785-17 (2018).
39. M. L'Azou *et al.*, *N. Engl. J. Med.* **374**, 1155–1166 (2016).
40. L. Premkumar *et al.*, *J. Clin. Microbiol.* **56**, e01504-17 (2018).
41. M. Montoya *et al.*, *J. Infect. Dis.* **218**, 536–545 (2018).
42. P. Andrade *et al.*, *Nat. Commun.* **10**, 938 (2019).
43. D. F. Robbiani *et al.*, *Cell* **169**, 597–609.e11 (2017).
44. M. G. Guzmán *et al.*, *Am. J. Epidemiol.* **152**, 793–799 (2000).
45. R. Aguas, I. Dorigatti, L. Coudeville, C. Luxemburger, N. M. Ferguson, *Sci. Rep.* **9**, 9395 (2019).
46. T. P. Endy *et al.*, *J. Infect. Dis.* **189**, 990–1000 (2004).
47. D. Buddhari *et al.*, *PLOS Negl. Trop. Dis.* **8**, e3230 (2014).
48. Z. Moodie *et al.*, *J. Infect. Dis.* **217**, 742–753 (2018).
49. B. Patel *et al.*, *PLOS Negl. Trop. Dis.* **11**, e0005554 (2017).
50. V. Dussupt *et al.*, *Nat. Med.* **26**, 228–235 (2020).
51. L. Goo *et al.*, *Virology* **515**, 191–202 (2018).
52. E. N. Gallichotte *et al.*, *J. Infect. Dis.* **216**, 1196–1204 (2017).
53. A. Grifoni *et al.*, *J. Virol.* **91**, e01469-17 (2017).
54. D. Weiskopf *et al.*, *J. Infect. Dis.* **212**, 1743–1751 (2015).
55. R. Raut *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 227–232 (2018).
56. T. T. Wang *et al.*, *Science* **355**, 395–398 (2017).
57. L. C. Katzelnick *et al.*, Data from Katzelnick *et al.*, "Zika virus infection enhances future risk of severe dengue disease." Zenodo 2020; doi:10.5281/zenodo.3925798

ACKNOWLEDGMENTS

We thank the Pediatric Dengue Cohort Study and Pediatric Dengue Hospital-based Study participants and their families. We are grateful to past and present members of the study team based at the Centro de Salud Sócrates Flores Vivas, Hospital Infantil Manuel de Jesús Rivera, the National Virology Laboratory in the Centro Nacional de Diagnóstico y Referencia, and the Sustainable Sciences Institute in Nicaragua for their dedication and high-quality work. We thank M. Montoya, P. Andrade, and P. Narvekar for providing important data from previous publications for understanding post-ZIKV infection neutralizing antibody titers. **Funding:** This research was funded by National Institutes of Health (NIH) grant P01AI106695 (to E.H.). The Pediatric Dengue Cohort Study was supported by NIH grants P01AI106695 (to E.H.), U19AI118610 (to E.H.), and R01AI099631 (to A.B.), as well as the Pediatric Dengue Vaccine Initiative grant VE-1 (to E.H.) and FIRST grant (to E.H. and J.C.), both from the Bill and Melinda Gates Foundation. L.C.K. was supported in part by the Global Health Equity Scholar FIC/NIH training grant D43TW010540. A.M.d.S. was supported by CDC contract 00HVCLJB-2017-04191 (to A.M.d.S.) and NIH/NIAID R01AI107731 (to A.M.d.S.). **Competing interests:** The E.H. laboratory received research funds from Takeda Vaccines, Inc. to analyze samples from vaccine recipients. E.H. served on one-time advisory boards for Merck and Takeda. **Author contributions:** L.C.K., A.B., and E.H. conceived the project; G.K., S.O., N.S., M.P., F.N., E.H., J.C., and A.G. led and managed the clinical studies; A.B., E.H., J.C.M., L.P., and A.M.d.S designed and oversaw all laboratory analyses; D.C., O.A., T.M., and B.S.-C. collected the serological data; C.N., S.A., B.L.M., D.E., K.L., A.S., and L.C.K. managed the clinical and laboratory databases and organized the data for analysis; L.C.K. E.H., A.B., A.M.d.S., L.P., C.N., A.G., F.B.C., and M.E.H. analyzed and discussed the data; L.C.K. and E.H. wrote the first draft of the manuscript and created the figures; all authors reviewed the manuscript. **Ethical statement:** The Pediatric Dengue Cohort Study and Pediatric Dengue Hospital-based Study were approved by the Institutional Review Boards (IRBs) of the Nicaraguan Ministry of Health, the University of California, Berkeley, and the University of Michigan. Parents or legal guardians of all subjects provided written informed consent, and subjects ≥6 years old provided assent. **Data and materials availability:** Data to understand and assess the conclusions of this research are available in the main text and supplementary materials. Sequence data are available on GenBank (accession numbers KM204119, KM204118, KU050695, KR011349, and KX421193), and code is available on Zenodo (*57*). Individual-level data for reproducing figures may be shared with outside investigators following UC Berkeley IRB approval. Please contact the UC Berkeley Center for the Protection of Human Subjects (ophs@berkeley.edu) and Eva Harris (eharris@berkeley.edu) to arrange for data access. Encoded databases, which do not contain identifiable information, were created and used for all analyses. Collaborating research groups and institutions will be sent coded data with all personal identifiers unlinked. More details are available in the supplementary material. The assay reagents are available from E.H. under a materials transfer agreement with UC Berkeley.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1123/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S8
References (*58*–*81*)
MDAR Reproducibility Checklist

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

8 April 2020; accepted 16 July 2020
10.1126/science.abb6143

IMMUNOGENOMICS

Mosquito cellular immunity at single-cell resolution

Gianmarco Raddi[1,2]*†, Ana Beatriz F. Barletta[2]*, Mirjana Efremova[1], Jose Luis Ramirez[2]‡, Rafael Cantera[3,4], Sarah A. Teichmann[1,5], Carolina Barillas-Mury[2]§, Oliver Billker[1,6]§

Hemocytes limit the capacity of mosquitoes to transmit human pathogens. Here we profile the transcriptomes of 8506 hemocytes of *Anopheles gambiae* and *Aedes aegypti* mosquito vectors. Our data reveal the functional diversity of hemocytes, with different subtypes of granulocytes expressing distinct and evolutionarily conserved subsets of effector genes. A previously unidentified cell type in *An. gambiae*, which we term "megacyte," is defined by a specific transmembrane protein marker (TM7318) and high expression of lipopolysaccharide-induced tumor necrosis factor–α transcription factor 3 (LL3). Knockdown experiments indicate that LL3 mediates hemocyte differentiation during immune priming. We identify and validate two main hemocyte lineages and find evidence of proliferating granulocyte populations. This atlas of medically relevant invertebrate immune cells at single-cell resolution identifies cellular events that underpin mosquito immunity to malaria infection.

Anopheline mosquitoes transmit *Plasmodium* parasites to humans and are responsible for an estimated 219 million cases of malaria, leading to more than 400,000 deaths annually (*1*). Parasites taken up by female mosquitoes from the blood of an infected human transform into motile ookinetes, which traverse the mosquito midgut and establish an infection. The mosquito's immune system limits *Plasmodium* infection in several ways (*2*–*4*), and hemocytes (insect white blood cells) are key players in these defense responses (*5*, *6*). Ookinete invasion triggers a strong nitration response in invaded midgut epithelial cells and their basal lamina (*7, 8*). Hemocytes that come in contact with a nitrated midgut basal lamina release microvesicles into the epithelial basal labyrinth and promote local complement activation, inducing parasite lysis (*6*). An infection with *Plasmodium* primes mosquitoes to mount a stronger immune response to subsequent infections (*9*). Primed mosquitoes release hemocyte differentiation factor (HDF) into the hemolymph (*9*); this factor consists of a complex of lipoxin A4 bound to evokin, a lipocalin carrier (*10*). HDF increases the proportion of circulating hemocytes of the granulocyte type (*2*), promotes microvesicle release, and enhances complement-mediated parasite lysis (*6*). Enhanced immunity is lost if HDF synthesis is blocked (*10*).

Three hemocyte types have been described in *Anopheles gambiae* morphologically (*11*). Granulocytes are highly phagocytic cells of about 10 to 20 μm in diameter. Oenocytoids are 8- to 12-μm round cells that produce melanin, an insoluble pigment involved in wound healing and pathogen containment by encapsulation. Prohemocytes are round cells (4 to 6 μm) with a high nuclear-to-cytoplasmic ratio and are thought to be precursors of the other two cell types. Hemocytes alternate between circulatory and tissue-resident (sessile) states (*12*, *13*). However, the full functional diversity of mosquito hemocytes and their developmental trajectories has not been established, and it is not clear to what extent morphologically similar hemocytes are functionally equivalent.

In this study, we used single-cell RNA sequencing (scRNA-seq) to analyze the transcriptional profile of individual mosquito hemocytes in response to blood feeding or infection with *Plasmodium*. Circulating hemocytes were collected from adult *An. gambiae* M-form (*An. coluzzii*) females that were either kept on a sugar meal or allowed to feed on a healthy or a *P. berghei*–infected mouse (Fig. 1A). Transcriptomes from 5383 cells (collected 1, 3, and 7 days after feeding) revealed nine major cell clusters (Fig. 1B and fig. S1). Two clusters originate from adipose tissue [fat-body clusters (FBCs) 1 and 2] and one from muscle tissue (MusC) (Fig. 1, B and C). FBC1 cells express several immune-modulatory genes such as CLIPs (CLIPA1, -7, -8, -9, and -14), LRIMs (LRIM1, -4A, -8A, -8B, -9, and -17), lectins (CTL4 and MA2), and SRPN2 (Fig. 1, B and C, fig. S2, and table S1), whereas FBC2 cells express high levels of vitellogenin, a canonical fat-body marker (*14*), after blood feeding. On the basis of their distinctive transcriptional profiles, we identified six hemocyte clusters (HCs) (Fig. 1, B and C). Hemocyte cluster 1 (HC1) has high mRNAs levels of prophenoloxidases, including PPO4 (Fig. 1, B and C) and PPO9, characteristic of oenocytoids. This expression profile is consistent with reported scRNA-seq data for 25 hemocytes (*15*). To select markers for major hemocyte lineages, we used bulk RNA-seq data from different tissues to identify hemocyte-specific genes (fig. S3 and tables S2 and S3). HC1 contains low levels of leucine-repeat protein 8 (LRR8) mRNA, whereas HC2 to HC4 have an inverse pattern (i.e., low or absent PPO4 and high LLR8 levels) (Figs. 1, B and C, and 2A, and fig. S2).

In situ hybridization using PPO4 and LLR8 as markers revealed that the morphology of circulating HC1 ($PPO4^{high}/LLR8^{low}$) cells is typical of oenocytoids, with round cells that have few pseudopodia and granules (Fig. 2B), whereas the morphology of HC2 to HC4 ($PPO4^{low}/LLR8^{high}$) cells is typical of prohemocytes and granulocytes (Fig. 2B). HC2 and HC3 shared markers such as SPARC, cathepsin-L, and LRR8 (Fig. 1C). However, HC2 had 73% fewer unique molecular identifiers (UMIs) (mean UMI of 413) than HC3 (mean UMI of 1516), which suggests that HC2 cells are less differentiated and probably constitute prohemocytes. Our observation that cells of prohemocyte morphology did not express markers of the more differentiated HC1 or HC5 cells is consistent with this notion (Fig. 2B). HC3 cells have a typical granulocyte morphology, with prominent pseudopodia and abundant granules (Fig. 2B).

HC4 shares markers with HC3, and a correlation analysis shows these cells to be granulocytes (fig. S6C). However, HC3 cells are also characterized by a specific subset of markers (Fig. 1C). HC4 cells express cyclin B, aurora kinase, and other mitotic markers, which suggests that they are proliferating hemocytes (fig. S6, C to E). Cells in HC5 and HC6 both are negative for PPO4 (Figs. 1C and 2A). HC5 cells express high levels of an uncharacterized transmembrane protein AGAP007318 (TM7318) and lipopolysaccharide-induced tumor necrosis factor–α transcription factor 3 (LL3) (Figs. 1C and 2A), whereas HC6 is negative for those markers and expresses antimicrobial peptides such as defensin 1, cecropins 1, and C-type lysozyme (Fig. 1C). Cells in HC5 and HC6 have both low levels of LLR8 and no PPO4 but have two distinct morphologies (Fig. 2, A and B). Cells negative for TM7318 (HC6) are small granulocytes that express antimicrobial genes (AM Gran) (16.4% of granulocytes), whereas TM7318-positive cells (HC5) are in low abundance (0.5% of granulocytes) and represent a separate giant cell type (25 to 40 μm) that we named "megacytes" (Fig. 2B).

To investigate the differentiation dynamics of *An. gambiae* hemocytes, we reclustered the cellular transcriptomes at higher resolution and performed lineage-tree reconstruction

[1]Wellcome Sanger Institute, Hinxton, Cambridge CB10 2AZ, UK. [2]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA. [3]Zoology Department, Stockholm University, Stockholm S-10691, Sweden. [4]Departamento de Biología del Neurodesarrollo, Instituto de Investigaciones Biológicas Clemente Estable, Montevideo 11600, Uruguay. [5]Institute and Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK. [6]Molecular Infection Medicine Sweden, Molecular Biology Department, Umeå University, Umeå S-90187, Sweden.

*These authors contributed equally to this work. †Present address: School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SP, UK. ‡Present address: Crop Bioprotection Research Unit, U. S. Department of Agriculture, National Center for Agricultural Utilization Research, Agricultural Research Service, Peoria, IL 61604, USA.

§Corresponding author. Email: cbarillas@niaid.nih.gov (C.B.-M.); oliver.billker@umu.se (O.B.)

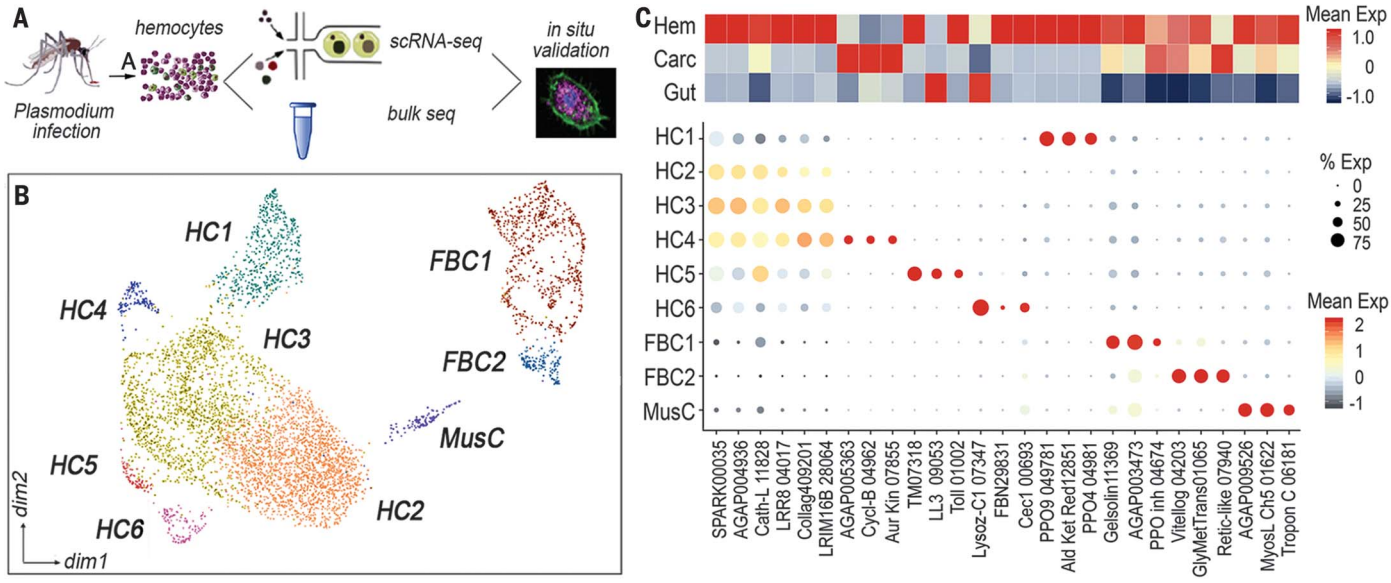


Fig. 1. Molecular identification of hemocyte cell types of *Anopheles* mosquitoes. (**A**) Workflow and experimental scheme. (**B**) UMAP (uniform manifold approximation and projection) of 5383 hemolymph *An. gambiae* (M form) cells colored by cluster (cell type) identity with Seurat clustering. HC, hemocyte cluster; Mus, muscle; FBC, fat body cluster. (**C**) (Top) Heatmap showing mean expression (Exp) of three gene markers per scRNA-seq cluster in bulk RNA-seq of mosquito tissues. Hem, hemocytes; Carc, carcasses; Gut, gut samples. (Bottom) Dot plot showing corresponding expression of the cluster marker genes, where color indicates mean expression and dot size encodes percentage of cells within the cluster expressing the marker. The last five digits of each marker gene are the Vectorbase transcript accession numbers (after AGAP0).

with partition-based graph abstraction (PAGA) (*16*). The major granulocyte population subclustered into three cell populations representing a basal, homeostatic state (Gran1) and two states activated by blood feeding and *Plasmodium* infection (Gran2 and Gran3) (Fig. 2C and fig. S6, A and B). Prohemocytes were subclustered into two populations (PHem1 and PHem2), of which PHem2 appeared to be an intermediate stage between PHem1 and Gran1. We traced a connection between dividing granulocytes (Div Gran) and Gran3 granulocytes, which in turn link to Gran1. Gran1 also links to Gran2, which in turn links to megacytes (Megac), whereas Gran1 links directly to antimicrobial granulocytes (AM Gran) (Fig. 2D). These findings were confirmed with diffusion maps and slingshot cell-trajectory analysis packages (figs. S4 and S5). Together, these analyses suggest the existence of a proliferative, oligopotent cell population that can replenish the pool of granulocytes and differentiate into more specialized hemocytes, such as megacytes and antimicrobial granulocytes.

Granulocytes (HC4 subtype) specifically express mitotic markers (fig. S6, C to E) largely in response to blood feeding (fig. S6, A and B), in agreement with a recent report that blood feeding induced DNA synthesis in hemocytes (*17*). Our data suggests that both granulocyte proliferation and prohemocyte differentiation contribute to the observed increase in granulocyte numbers after blood feeding (*11*, *18*, *19*). However, the placement of prohemocytes in the granulocyte lineage tree should be treated with caution owing to the paucity of distinctive markers.

Prohemocytes are proposed to be stem cell precursors of both granulocytes and oenocytoids (*11*). However, oenocytoids are transcriptionally disconnected from other hemocyte subtypes, and we did not observe transcriptional markers of cell proliferation in oenocytoids, which suggests that they may represent a separate lineage originating from either larval stages or other adult tissues. Alternatively, oenocytoids could be derived from granulocytes (Fig. 1B), but the differentiation rate may be very low when melanization responses are not elicited, which could result in too few cells at intermediate stages of differentiation to be captured in our transcriptomic analysis.

To assess which of the newly discovered putative cell types are shared between anopheline and culicine mosquitoes, we analyzed the transcriptomes of 3123 hemolymph cells from *Aedes aegypti*, a vector for several viral diseases including yellow fever, dengue, chikungunya, and Zika. As with *Anopheles*, a dimensional reduction plot shows both canonical hemocytes and other cell types with mostly fat-body signatures (Fig. 3A and fig. S7). A cross-species correlation analysis reveals two clusters (AaHC1 and AaHC2) with conserved transcriptome signatures for oenocytoids (99 and 77% correlation, respectively, with AgHC1) (Fig. 3, A and B) and different granulocyte types, including antimicrobial peptide–expressing cells (94% with AgHC6) and proliferating granulocytes (87% with AgHC4) (Fig. 3, A and B, and table S4). Granulocytes and prohemocytes are again positioned along a continuum of transcriptomic similarity, with four different cell states, including a proliferating S-phase granulocyte cluster (AaHC6) without a clear *Anopheles* equivalent (Fig. 3, A and B). Granulocytes express laminins, leucine-rich repeat proteins, scavenger receptors, Toll-like receptor 5, and the transcription factor Rel2 (fig. S8 and table S5). However, megacytes (AgHC5) lack an obvious counterpart in *Aedes*, and their characteristic marker (TM7318) is present only in anophelines of the *Cellia* subgenus (malaria vectors in Africa and Asia; fig. S9).

The transcription factor LL3 can be detected in granulocytes from *Plasmodium*-infected *An. gambiae*, and silencing LL3 expression disrupts priming (*20*). However, it is not clear whether LL3 is essential for HDF synthesis or for hemocytes to respond to HDF. We found that LL3 is highly expressed in megacytes (HC5) (Figs. 1C and 4A) and explored whether silencing LL3 affects the HDF response. Transfer of hemolymph from primed *An. gambiae* donors had HDF activity and elicited a strong priming response in control recipients injected with *lacZ* double-stranded (ds) RNA, resulting in a prominent increase in circulating granulocytes, a modest increase in oenocytoids, and a decrease in prohemocytes (Fig. 4B and tables S6 to S9). This response was abolished when LL3 expression was silenced in the recipients

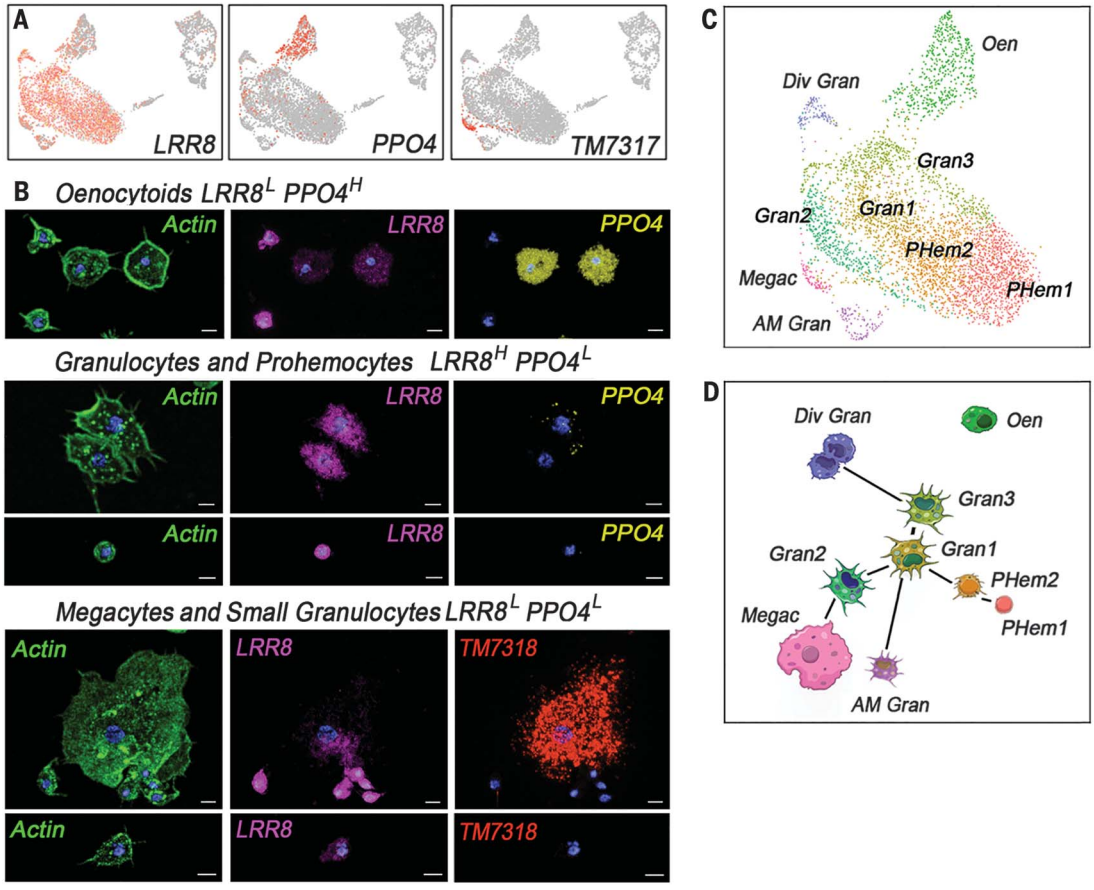


Fig. 2. Cell type validation and characterization of *An. gambiae* hemocyte lineages. (**A**) UMAP visualization of marker gene expression. Cells with more than 1 UMI are shown in red. (**B**) RNA in situ hybridization combined with morphology analysis of circulating hemocytes. Actin, green; LRR8, magenta; PPO4, yellow; TM7318, red; nuclei, blue. L, low; H, high. Scale bars: 5 μm. (**C**) UMAP of 5383 hemolymph cells colored by cluster (cell state) identity with Seurat clustering. (**D**) Unsupervised PAGA network analysis connecting stylized cell states based on UMAP clustering. Oen, oenocytoids; Div Gran, dividing granulocytes; Gran, granulocytes; Megac, megacytes; AM Gran, antimicrobial granulocytes; PHem, prohemocytes.

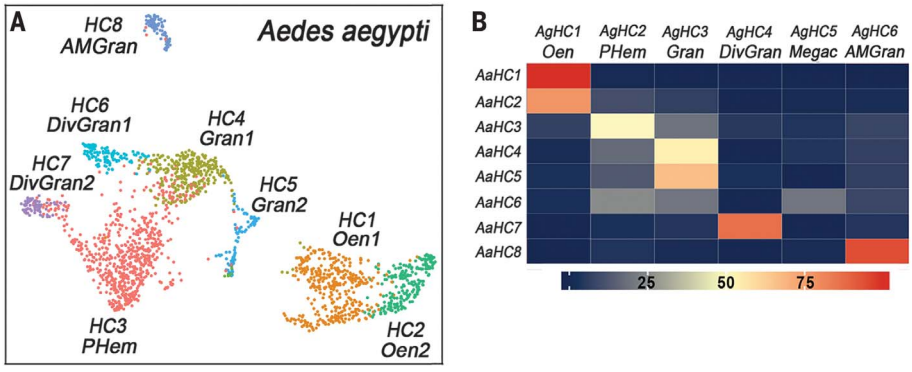


Fig. 3. Characterization of *Ae. aegypti* hemocytes and correlation with *Anopheles* cell types. (**A**) UMAP of 3123 *Ae. aegypti* hemocytes colored by cluster identity with Seurat. (**B**) Heatmap showing probability of each *Ae. aegypti* hemocyte in a cluster belonging to each of the *An. gambiae* (M form) cell types after logistic regression and using a multinomial learning approach. *Ag*, *An. gambiae* (M form); *Aa*, *Ae. aegypti*.

by injection of dsLL3 RNA (Fig. 4B and tables S6 to S9). We cannot rule out the possibility that hemocytes other than megacytes express LL3 at levels below the detection limit of scRNA-seq, which could have led to the notion that LL3 is up-regulated more broadly in granulocytes after *Plasmodium* infection (*20*), in which case LL3 silencing might directly control hemocyte differentiation in response to HDF. However, the discovery of megacytes expressing high levels of LL3 as a defining feature now raises the possibility that this cell type plays a key role in orchestrating hemocyte responses to HDF.

The proportion of circulating granulocytes is low (1 to 3%) under normal conditions but increases in response to *Plasmodium* infection (*2*). Whether the increase is due to granulocyte proliferation, differentiation from prohemocytes, or mobilization of sessile hemocytes was unknown. Transmission electron microscopy of sugar-fed mosquitoes showed individual sessile hemocytes bathed by hemolymph and attached to the basal lamina of the tissues through pseudopods (Fig. 4C), indicative of a dynamic and potentially transient association.

Using whole-tissue–mount in situ hybridization, we found that most sessile hemocytes are $PPO4^{low}/LLR8^{high}$ granulocytes (89.3 ± 6.2% SEM), whereas $PPO4^{high}/LLR8^{low}$ oenocytoids are less abundant (4.2 ± 3.1% SEM) and TM7318-positive megacytes are even rarer (2.7 ± 2.3% SEM) (Fig. 4, D and E, and tables S10 and S11). Furthermore, we found a reduction of sessile $PPO4^{low}/LLR8^{high}$ granulocytes in response to *Plasmodium* infection ($P < 0.0001$, Welch

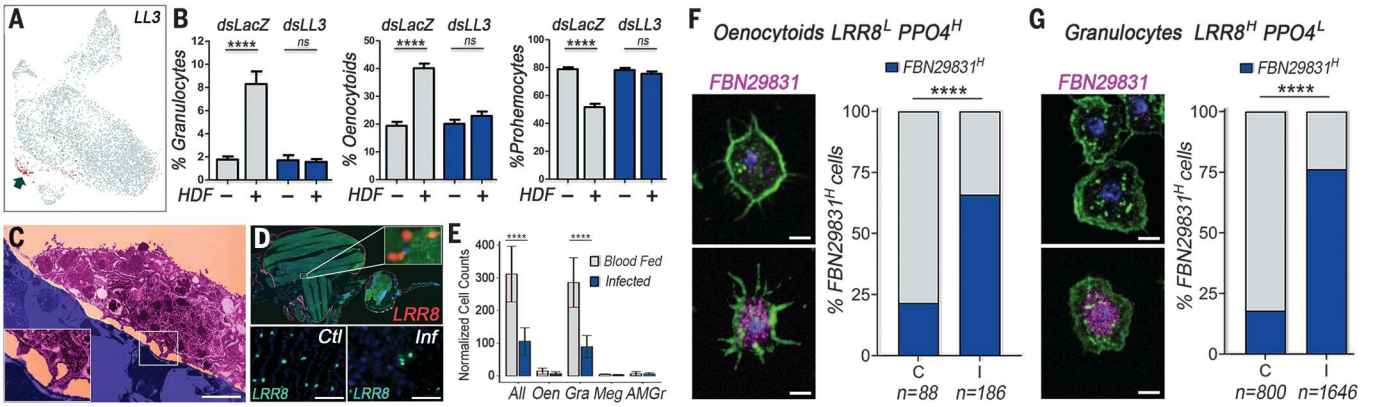


Fig. 4. *An. gambiae* cellular immune responses to *Plasmodium* infection. (**A**) UMAP visualization of all hemocytes by *LL3* expression. Cells with more than 1 UMI are shown in red. The arrow denotes a megacyte cluster. (**B**) Percentage of circulating granulocytes, oenocytoids, and prohemocytes of *LL3*-silenced mosquitoes injected (+) or not (−) with HDF versus double-stranded *lacZ* RNA–injected mosquitoes used as negative controls (mean ± SEM, ****$P < 0.0001$, unpaired *t* test; two independent experiments). ns, not significant. (**C**) False-color transmission electron microscopy images depicting an *An. gambiae* (M form) granulocyte (purple) attached via pseudopodia (inset) to the abdominal fat body (blue). Scale bar: 1.5 μm. (**D**) RNA in situ hybridization of a longitudinal section (top) or carcass whole-mounts (bottom) of control blood-fed (Ctl) and *P. berghei*–infected (Inf) *An. gambiae* mosquitoes. (Top) Glyceraldehyde phosphate dehydrogenase, green; LRR8 in red. (Bottom) LRR8, green. Scale bars: 20 μm. (**E**) Quantification of *Anopheles* hemocytes attached to the mosquito fat body with blood feeding or *P. berghei* infection, normalized by surface area of fat bodies. All, all hemocytes; Oen, oenocytoids; Gra, granulocytes; Meg, megacytes; AM Gr, antimicrobial granulocytes (mean ± SEM, ****$P < 0.0001$, Welch *t* test; three independent experiments). (**F**) Percentage of circulating oenocytoids ($LRR^{L}PPO4^{H}$) and (**G**) granulocytes ($LRR^{H}PPO4^{L}$) positive for *FBN29831* in control (C) and *P. berghei*–infected (I) mosquitoes 48 hours after feeding (right panels). n, number of cells examined. Representative RNA in situ hybridization images of oenocytoids (F) and granulocytes (G) with low and high expression of FBN29831 (left panels). Actin, green; FBN29831, magenta; nuclei, blue. Scale bars: 5 μm (****$P < 0.0001$, χ^2 test; two independent experiments).

t test; Fig. 4E and tables S12 and S13), with no significant difference in the numbers of sessile $PPO4^{high}/LLR8^{low}$ oenocytoids, $TM7318^{+}$ megacytes, or $TM7318^{-}$ AM granulocytes (Fig. 4E and tables S12 and S13). In circulating hemocytes, *P. berghei* infection induced high expression of the fibrinogen-like protein AGAP029831 (FBN29831) (fig. S10). RNA in situ hybridization showed that the proportion of FBN29831-positive cells increased in both $PPO^{low}/LLR8^{high}$ granulocytes ($P < 0.0001$, X^2; Fig. 4F and tables S14 and S15) and $PPO4^{high}/LLR8^{low}$ oenocytoids ($P < 0.0001$, X^2; Fig. 4G and tables S16 and S17), indicating that this is a general marker of hemocyte immune activation. Infecting *An. gambiae* with the human parasite *P. falciparum* produced a similar increase in FBN29831-positive cells (fig. S11). Combined, our results suggest that hemocyte recruitment from the body wall, granulocyte activation and proliferation, and prohemocyte differentiation can all contribute to boost circulating granulocyte numbers in response to immune challenge.

The hemocyte atlas presented here confirms the existence of two canonical hemocyte types in mosquitoes (oenocytoids and granulocytes), and, with the help of transcript markers, we relate cellular morphology to their transcriptome. We show that prohemocytes and granulocytes are closely related cells, and we identify transcriptional profiles and molecular markers that define previously unknown hemocyte subtypes (megacytes, proliferating granulocytes, and antimicrobial granulocytes), as well as a subpopulation of fat-body cells (AgFBC1) with potential immune-modulatory functions.

We define two main hemocyte lineages in *An. gambiae*—the oenocytoid lineage and the prohemocyte-granulocyte lineage—along with three sublineages that lead to differentiated megacytes, antimicrobial granulocytes, or proliferating granulocytes. The cell type markers and fluorescence in situ hybridization probes we identify and validate will allow investigators to probe the immune functions of megacytes and other specialized hemocyte types in detail, which will provide a basis for a better understanding of how cellular immunity limits the ability of malaria parasites to infect mosquitoes. Our analysis cannot resolve the developmental origin of oenocytoids, but we identify two potential origins for the observed expansion of circulating granulocytes in response to blood feeding. One is the mobilization of sessile $PPO4^{low}/LLR8^{high}$ cells from the body wall; the other is a pool of proliferating, oligopotent granulocytes. Whether prohemocytes, which appear transcriptionally related to but less responsive than granulocytes, can transform into granulocytes and whether they need to enter the cell cycle remains to be determined. These questions may now be addressed with the type of lineage tracing experiments commonly performed in mouse immunology to resolve the developmental origins and functions of diverse immune cell populations. The cell type–specific marker genes, reference transcriptomes, and companion website (https://hemocytes.cellgeni.sanger.ac.uk/) we present here will serve as resources for such studies.

The conservation of diverse and molecularly well-defined hemocyte types between distantly related mosquito genera and the apparent absence of megacytes in our *Ae. aegypti* mosquito dataset raise questions as to how the immune systems of these mosquito species have evolved to limit their capacity to transmit parasites and arboviruses to humans. This knowledge will ultimately underpin immunological strategies aimed at interrupting disease transmission by rendering mosquitoes resistant to such pathogens.

REFERENCES AND NOTES

1. World Health Organization, "World malaria report 2019" (2019); www.who.int/publications-detail/world-malaria-report-2019.
2. J. Rodrigues, F. A. Brayner, L. C. Alves, R. Dixit, C. Barillas-Mury, *Science* **329**, 1353–1355 (2010).
3. S. Luckhart, Y. Vodovotz, L. Cui, R. Rosenberg, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5700–5705 (1998).
4. S. Blandin *et al.*, *Cell* **116**, 661–670 (2004).
5. J. L. Ramirez *et al.*, *J. Innate Immun.* **6**, 119–128 (2014).
6. J. C. Castillo, A. B. B. Ferreira, N. Trisnadi, C. Barillas-Mury, *Sci. Immunol.* **2**, eaal1505 (2017).
7. S. Kumar, L. Gupta, Y. S. Han, C. Barillas-Mury, *J. Biol. Chem.* **279**, 53475–53482 (2004).
8. G. de A. Oliveira, J. Lieberman, C. Barillas-Mury, *Science* **335**, 856–859 (2012).
9. B. C. Ho, E. H. Yap, M. Singh, *Parasitology* **85**, 567–575 (1982).
10. J. L. Ramirez *et al.*, *Nat. Commun.* **6**, 7403 (2015).
11. J. C. Castillo, A. E. Robertson, M. R. Strand, *Insect Biochem. Mol. Biol.* **36**, 891–903 (2006).
12. J. G. King, J. F. Hillyer, *PLOS Pathog.* **8**, e1003058 (2012).
13. J. G. King, J. F. Hillyer, *BMC Biol.* **11**, 55 (2013).

14. G. M. Attardo, I. A. Hansen, A. S. Raikhel, . *Insect Biochem. Mol. Biol.* **35**, 661–675 (2005).
15. M. S. Severo *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E7568–E7577 (2018).
16. F. A. Wolf *et al*., *Genome Biol.* **20**, 59 (2019).
17. W. B. Bryant, K. Michel, *J. Exp. Biol.* **217**, 1238–1245 (2014).
18. L. A. Baton, A. Robertson, E. Warr, M. R. Strand, G. Dimopoulos, *BMC Genomics* **10**, 257 (2009).
19. J. Castillo, M. R. Brown, M. R. Strand, *PLOS Pathog.* **7**, e1002274 (2011).
20. R. C. Smith, C. Barillas-Mury, M. Jacobs-Lorena, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E3412–E3420 (2015).
21. G. Raddi, M. Efremova, Mosquito-Atlas, Version 1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.3882128.

ACKNOWLEDGMENTS

The authors are grateful to T. Metcalf for technical support, A. Goldsborough for making available the vivoPHIX reagent and advising on its use, J. Eliasova for help with illustrations, and the Wellcome Sanger Institute CellGen team for creating and maintaining the data visualization website. **Funding:** This work was funded by the Intramural Research Program of the Division of Intramural Research Z01AI000947, National Institute of Allergy and Infectious Diseases (NIAID), NIH to C.B.-M.; an NIH Oxford-Cambridge fellowship and a UCLA-Caltech Medical Scientist Training Program grant (NIGMS T32 GM008042) to G.R.; and a Wellcome core grant 206194/Z/17/Z to the Sanger Institute and funding from the Knut and Alice Wallenberg Foundation and the European Research Council (grant 788516) to O.B. **Author contributions:** Conceptualization: G.R., A.B.F.B., M.E., S.A.T., C.B.-M., and O.B.; Methodology: G.R.; Investigation: G.R., A.B.F.B., J.L.R., and R.C.; Formal analysis: G.R., A.B.F.B., and M.E.; Writing – original draft: G.R., C.B.-M., and O.B.; Supervision: S.A.T., C.B.-M., and O.B.; Funding acquisition: S.A.T., C.B.-M., and O.B. All authors reviewed the manuscript. **Competing interests:** S.A.T. reports consulting and scientific advisory board activities over the past 3 years for Genentech, Roche, Biogen, GSK, and Foresite Labs. The other authors declare no competing interests. **Data and materials availability:** scRNA-seq data for *An. gambiae* and *Ae. aegypti* hemocytes are available in ArrayExpress under accession number E-MTAB-9240. Bulk RNA-seq data are available under accession number E-MTAB-9241. The code to reproduce the analysis of the paper is available on Zenodo (*21*). All other data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1128/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S18
References (*22*–*36*)
MDAR Reproducibility Checklist

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

2 April 2020; accepted 30 June 2020
10.1126/science.abc0322

PROTEIN DESIGN

Expanding the space of protein geometries by computational design of de novo fold families

Xingjie Pan[1,2]*, Michael C. Thompson[1]†, Yang Zhang[1], Lin Liu[1], James S. Fraser[1,3], Mark J. S. Kelly[4], Tanja Kortemme[1,2,3,5]*

Naturally occurring proteins vary the precise geometries of structural elements to create distinct shapes optimal for function. We present a computational design method, loop-helix-loop unit combinatorial sampling (LUCS), that mimics nature's ability to create families of proteins with the same overall fold but precisely tunable geometries. Through near-exhaustive sampling of loop-helix-loop elements, LUCS generates highly diverse geometries encompassing those found in nature but also surpassing known structure space. Biophysical characterization showed that 17 (38%) of 45 tested LUCS designs encompassing two different structural topologies were well folded, including 16 with designed non-native geometries. Four experimentally solved structures closely matched the designs. LUCS greatly expands the designable structure space and offers a new paradigm for designing proteins with tunable geometries that may be customizable for novel functions.

The design of proteins with new and useful architectures and functions requires precise control over molecular geometries (*1*, *2*). In nature, proteins adopt a limited set of protein fold topologies (*3*–*5*) that are reused and adapted for different functions. Here we define "topology" as the identity and connectivity of secondary structure elements (Fig. 1A). Within a given topology, geometric features including length and orientations of secondary structure elements are often highly variable (*3*, *4*). These considerable geometric differences between proteins with the same topology are necessary, as they define the exquisite shape and physicochemical complementarity characteristic of protein functional sites. Creating proteins with new functions de novo therefore requires the ability to design proteins with not only different topologies, but also distinct custom-shaped geometries within these topologies optimal for each function (Fig. 1A).

Computational design has been successful in mimicking the ability of evolution to generate diverse protein structures spanning α-helical (*6*–*10*), α-β (*11*–*13*), and β-sheet (*14*, *15*) fold topologies, including novel folds (*16*). However, most design methods do not include explicit mechanisms to vary geometric features within a topology. For instance, successful design methods assemble protein structures from peptide fragments using a definition of the desired fold and topological rules derived from naturally occurring structures (*12*). Subsequent iterative cycles of fixed-backbone sequence optimization and fixed-sequence structure minimization (*16*) refine atomic packing interactions but do not create substantial changes in geometry. An exception are methods that use parametric equations to sample backbone variation (*17*) or take advantage of modular protein elements, but these methods are restricted to helical bundles (*6*, *8*, *10*) or repeat protein (*18*) architectures, respectively.

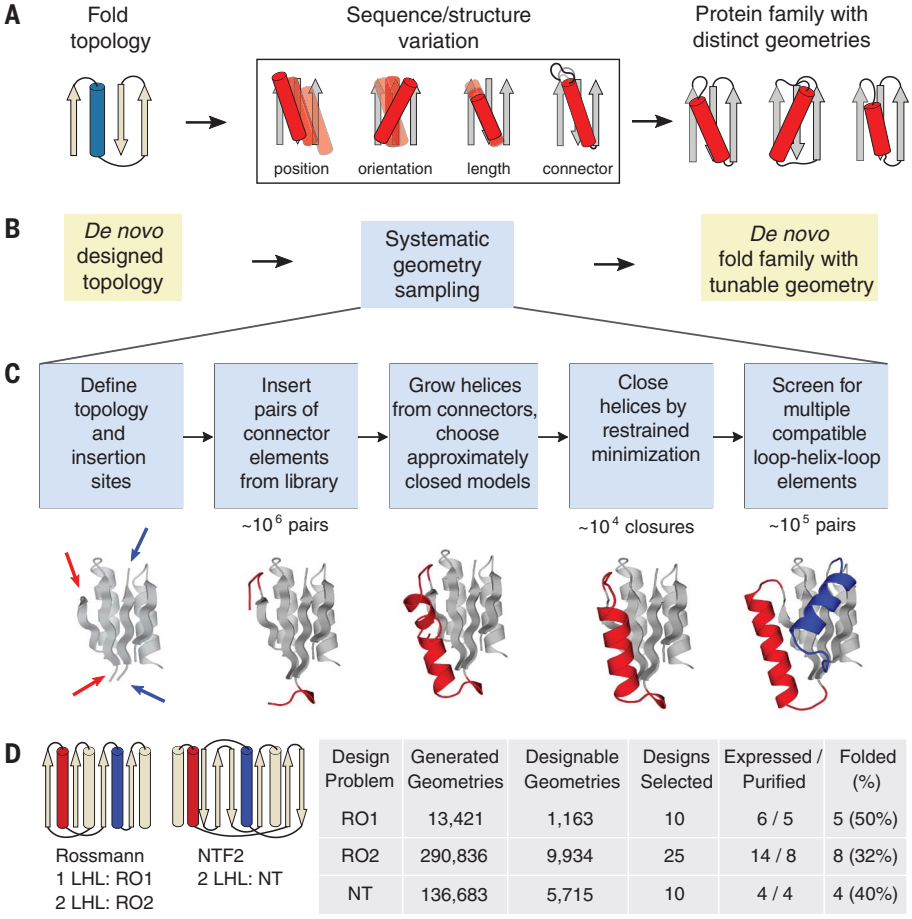
Here, we sought to develop a generalizable computational design approach that mimics the ability of evolution to create geometric variation within a given fold topology (Fig. 1). When analyzing geometric variation in protein fold families, we found that 84% of naturally occurring fold families contain variations in loop-helix-loop (LHL) elements (fig. S1). We hence reasoned that a method that systematically samples geometric variation in these units would be able to recapitulate a large fraction of geometric diversity in naturally occurring structures while also creating fold families of de novo designed proteins with tunable geometries (Fig. 1B).

To develop a generalizable method that systematically samples geometries of LHL elements, we first examined the individual connecting loop elements in native LHL units. For all LHL elements from all CATH superfamilies (*3*) of nonredundant structures, 72.8% of the loops contained five or fewer residues (fig. S2A). We extracted 313,072 loops of length 2 to 5 connecting to helices from the Rosetta nonredundant fragment database (*19*) and sorted loops into 12 libraries according to loop length and type of adjacent secondary structure (table S1). For each library, only nonredundant loops were retained (*20*); this procedure yielded between 224 and 5826 loops per library. The loop libraries had degeneracies (total number of loops divided by number of nonredundant loops in each library) ranging from 4.4 to 202 (fig. S2B), indicating that evolution frequently uses similar loop structures in different proteins. This observation suggests that the identified loop element libraries could also be used to computationally sample novel protein structures that have not been explored by nature.

We developed a protocol called loop-helix-loop unit combinatorial sampling (LUCS; Fig. 1C and fig. S3). LUCS starts with an input protein fold, which can be naturally occurring or (as in our case) de novo designed (*20*), and a definition of gaps in which to insert LHL units. The first step systematically samples all individual loop elements from our libraries (table S1). For each gap, loops are inserted at each end of the gap and any loops that clash with the input structure are removed. In a second step, all pairs of remaining loops are tested for supporting LHL units by growing helices from each loop. If helices grown from the two ends meet in the middle, excess residues are removed in the third step and the gap is closed by energy minimization with a chain-break penalty and hydrogen bond restraints. Closed LHL units with distorted hydrogen bond geometries, steric clashes, or suboptimal interactions between designed backbones and the environment are discarded (*20*). In a fourth step, combinations of LHL units at different positions can be screened to yield final structures that have multiple compatible LHL units with systematically sampled lengths and orientations.

To validate the ability of LUCS to generate distinct geometries within given fold topologies, we applied the method to three design problems (Fig. 1D). In the first two design problems, we varied one (RO1) or two (RO2) LHL units of a de novo designed protein (*12*) (PDB ID 2LV8) with a Rossmann fold topology. In the third problem, we varied two LHL units of a de novo designed protein (*21*) (PDB ID 5TPJ) with a nuclear transport factor 2 (NTF2) fold topology (NT). In principle, LUCS can sample topologies with an arbitrary number of LHL units. For the systems we tested, systematic geometry sampling generated ~10^4 LHL elements for each gap. To limit the required computing power, we screened 10^6 random combinations of LHL units and generated 10^4 to 10^5 final backbone structures for each design problem (table S2). We then applied the Rosetta FastDesign protocol (*20*) to optimize sequences for all residue positions within 10 Å from the new LHL elements. The number of designed residues for each backbone was between 33 and 87. We note that Rosetta FastDesign also introduces structural changes outside the reshaped LHL elements of the designed fold through gradient-based torsion minimization, although these changes are small [backbone heavy-atom root mean square deviation (RMSD) < 1 Å]. After sequence design, we filtered the design models computationally using a set of quality criteria

[1]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, USA. [2]UC Berkeley–UCSF Graduate Program in Bioengineering, University of California, San Francisco, CA, USA. [3]Quantitative Biosciences Institute, University of California, San Francisco, CA, USA. [4]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA. [5]Chan Zuckerberg Biohub, San Francisco, CA, USA.
*Corresponding author. Email: xingjiepan@gmail.com (X.P.); tanjakortemme@gmail.com (T.K.)
†Present address: Department of Chemistry and Chemical Biology, University of California, Merced, CA, USA.



Design Problem	Generated Geometries	Designable Geometries	Designs Selected	Expressed / Purified	Folded (%)
RO1	13,421	1,163	10	6 / 5	5 (50%)
RO2	290,836	9,934	25	14 / 8	8 (32%)
NT	136,683	5,715	10	4 / 4	4 (40%)

Fig. 1. LUCS sampling strategy to create de novo designed protein fold families with tunable geometries. (**A**) In nature, protein fold topologies (left) are diversified to create families of proteins with distinct geometries (right) optimized for function; α helices are shown as cylinders and β strands as arrows. The box shows schematic representations of common types of geometric variation. (**B**) The LUCS computational design protocol seeks to mimic the ability of evolution to diversify protein geometries to generate de novo designed fold families. (**C**) Schematic of the LUCS protocol for sampling LHL geometries. The reshaped LHL units are colored in red and blue. Typical numbers of models generated at major stages of the protocol are indicated. (**D**) Designed fold families. Schematic shows fold topologies and design problems: Rossmann fold with one or two reshaped LHL units (RO1 and RO2) and NTF2 fold with two reshaped LHL units (NT). Also shown are numbers for geometries generated by LUCS, designed models that passed quality filters, and experimentally characterized designs for three design problems. The rightmost column indicates the fraction of experimentally tested designs that adopted folded structures.

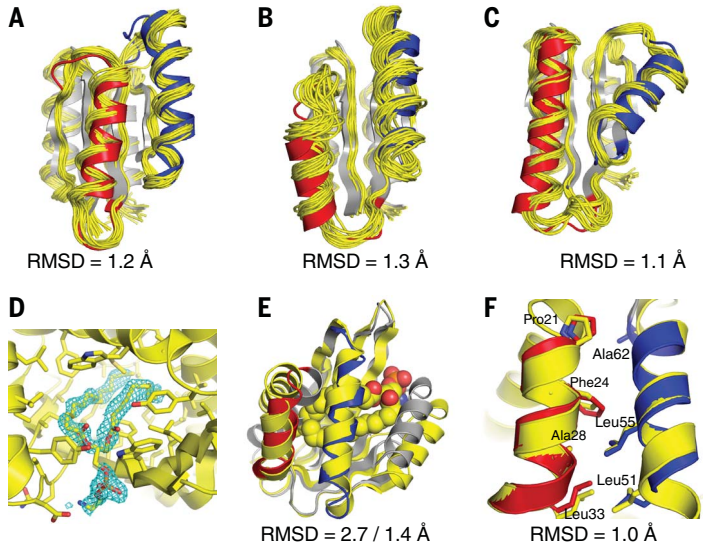


Fig. 2. Close agreement between models and experimentally determined structures of designed proteins. Experimentally determined structures are shown in yellow and design models in gray, with the reshaped LHL elements highlighted in red and blue. (**A** to **C**) Designs for the Rossmann fold topology. Computational models and NMR structures are compared for designs RO2_1 (A), RO2_20 (B), and RO2_25 (C). Also shown are the backbone heavy-atom RMSDs calculated using the lowest-energy structure from the NMR ensemble. (**D** to **F**) Design for the NTF2 fold topology. (D) The binding pocket of a phosphatidylethanolamine ligand. The $2F_{obs} - F_{calc}$ electron density map (cyan) for the ligand molecule is shown at 1.0σ level. (E) Comparison between computational model and x-ray crystal structure for design NT_9. The phosphatidylethanolamine ligand is shown in space-filling representation (carbon atoms in yellow, oxygen atoms in red, phosphorus atoms in orange, and nitrogen atoms in blue). Also shown are the backbone heavy-atom RMSDs calculated including or excluding the terminal helices, respectively. (F) Alignment between the designed helices in the computational model and the experimentally solved structure for design NT-9. The hydrophobic residues at the packing interface are shown in stick representation. The RMSD shown includes the helix backbone heavy atoms and side-chain heavy atoms displayed as sticks.

that included a minimal number of buried unsatisfied hydrogen bond donors and acceptors, tight atomic packing interactions in the protein core, and compatibility between sequences and local structures (*20*).

For each of the three design problems, we selected 50 low Rosetta energy (*22*) designs from models that passed the quality filters and had diverse conformations for further computational characterization. The Rosetta FastDesign simulations optimized low-energy sequences given a desired structure. To determine the converse—whether the desired structure is also a low-energy conformation, given the sequence—we conducted ab initio protein structure prediction simulations in Rosetta (*23*). For the Rossmann fold designs, we required the lowest-energy predicted structure to be within 1 Å Cα RMSD of the design model. For the NTF2 fold designs, we used a less strict criterion requiring a number of low-energy models to be close to the design model; this criterion enabled us to account for the more difficult problem of sampling native-like structures

for proteins larger than 100 amino acids. For each of the three design problems, 10, 25, and 10 designs that passed these tests, respectively, were chosen for experimental characterization (Fig. 1D and data S1 and S2). The designed proteins were recombinantly expressed in *Escherichia coli* and purified using His-tag affinity and size exclusion chromatography. We found that 5/10, 8/25, and 4/10 designs were monomeric and well folded for each of the three design problems, respectively, as determined by far-ultraviolet circular dichroism and one-dimensional ^{1}H and two-dimensional ^{15}N heteronuclear single quantum coherence (HSQC) nuclear magnetic resonance (NMR) spectroscopy (Fig. 1D, fig. S4, and table S3).

To assess whether the designed structures adopted their intended geometries, we solved structures for three designs (RO2-1, RO2-20, and RO2-25) that sampled two LHL units in the Rossmann fold topology by NMR spectroscopy, and one structure for the NTF2 fold topology designs (NT-9) by x-ray crystallography (*20*) (fig. S5 and tables S4 and S5). The experimentally solved RO2 design structures closely matched the computational models (Fig. 2, A to C), with backbone heavy-atom RMSDs between models and solved structures within 1.3 Å, core hydrophobic side chains in good agreements with the designed models (fig. S6), and five of the loops in designed LHL units well converged (fig. S7). In the crystallographic electron density map obtained at 1.5 Å resolution for the NTF2 fold design (NT-9), strong signal was clearly identifiable inside a surface pocket (Fig. 2D), which was interpreted as a bound phospholipid [1,2-diacyl-*sn*-glycero-3-phosphoethanolamine (*20*)]. The two N- and C-terminal helices (residues 1 to 20 and 113 to 128), which had not been reshaped by LUCS, were pushed apart to accommodate the ligand, leading to an overall backbone heavy-atom RMSD between design and model of 2.7 Å. However, when excluding the N- and C-terminal helices and aligning the remainder of the design, the backbone heavy-atom RMSD between the model and the solved structure was 1.4 Å (Fig. 2E). Moreover, the designed side-chain packing interactions between the reshaped helices were in excellent agreement with the design (Fig. 2F). Taken together, our structural analysis confirmed the designed geometry in the reshaped regions for all four designs. The presence of a ligand in the NT-9 design is consistent with the known ability of the NTF2 fold to bind to diverse hydrophobic small molecules. Such a result implies the possibility of introducing new functions such as ligand binding by reshaping protein geometries.

We next analyzed the magnitude of the geometric differences between our designs. We first compared the backbone heavy-atom RMSDs between the reshaped helices of all well-folded designs (Fig. 1D) after aligning the non-reshaped regions using both the design models and experimentally solved structures (Fig. 3A and fig. S8). For the designs with one LHL unit reshaped, 18 of 20 off-diagonal differences were more than 3 Å (Fig. 3A, left). For the designs with two LHL units reshaped, 55 of 68 off-diagonal differences were more than 4 Å (Fig. 3A, center and right). This scale of variation exceeds the backbone changes generated by existing flexible backbone design methods (*24*, *25*), which are typically smaller than 2 Å RMSD. For each well-folded design, we also identified the closest structures in the Protein Data Bank (PDB) using TM-align (*26*). Of the designed LHL regions in the 17 well-folded LUCS designs, 15 were significantly different (RMSD > 3 Å for designs with one LHL unit reshaped; RMSD > 4 Å for designs with two LHL units reshaped) from their closest match in the PDB (Fig. 3A and fig. S9), indicating that the design protocol not

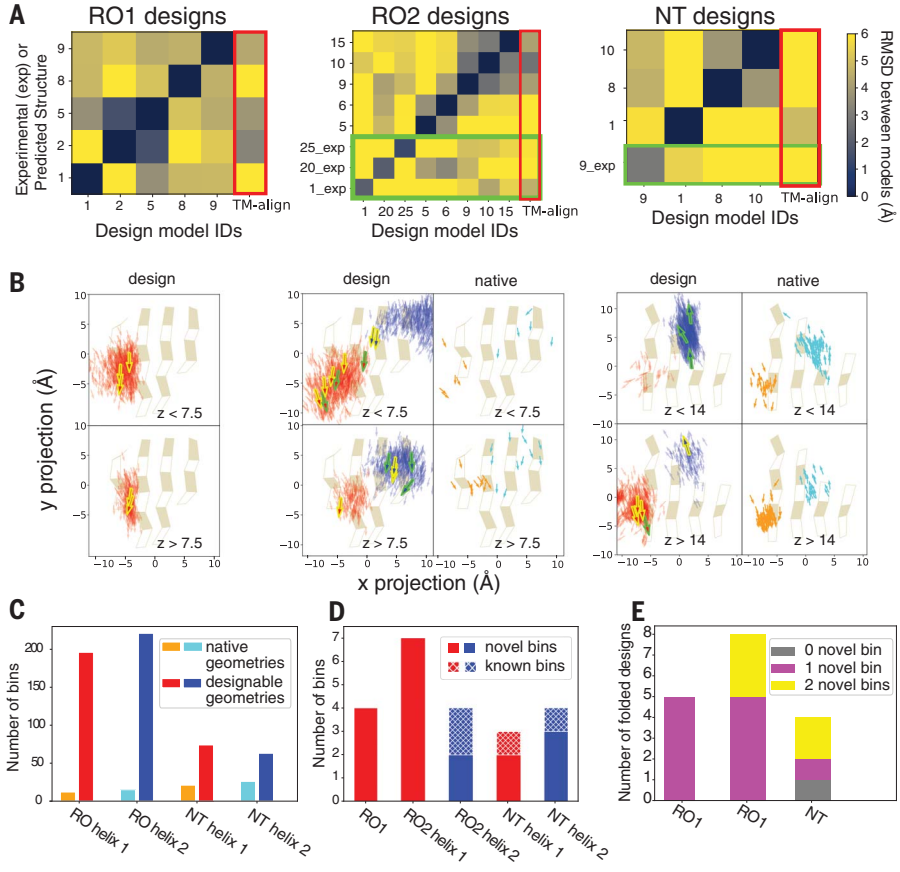


Fig. 3. Geometry space sampled by de novo designed fold families. In (A) and (B), the columns show the three design problems: Left, Rossmann fold with one designed LHL unit (RO1); center, Rossmann fold with two designed LHL units (RO2); right, NTF2 fold with two designed LHL units (NT). (**A**) Heat maps showing backbone RMSDs between the reshaped LHL regions of well-folded designs, comparing design models (*x* axis) with experimentally determined structures (_exp) or lowest-scoring models from Rosetta structure prediction (*y* axis). Green boxes show RMSDs calculated using experimentally solved structures. Red boxes (rightmost column in each grid) show the RMSDs between designs and the closest known structures found by TM-align. (**B**) Projection of centers and directions of designed helices (arrows) onto the underlying β sheets. For the RO2 (center) and NT (right) panels, distributions in designable models (Fig. 1D) are shown on the left (helices colored red and blue), and in known naturally occurring structures on the right (corresponding helices in orange and cyan). The two rows show helices on two *z*-level planes based on their distances from the β-sheet projection plane. For planes that have more than 1000 sampled structures, only 1000 randomly selected helices are shown. For the designs, experimentally confirmed folded designs are represented as bold arrows with yellow boundaries; designs with experimentally solved structures as bold arrows with green boundaries. For the natural proteins, the Rossmann fold structures are from the CATH superfamily 3.40.50.1980 and the NTF2 fold structures are from the CATH superfamily 3.10.450.50. (**C**) Number of structure bins occupied by known structures (orange, cyan) and sampled by designable models generated by LUCS (red, blue). (**D**) Structure bins occupied by well-folded designs. (**E**) Classification of the well-folded structures by the number of novel structure bins they occupy.

only generates stable structures with considerable conformational divergence, but also geometries not observed in known structures.

We further analyzed the distribution of sampled geometries and their coverage of designable backbone structure space, where a structure is defined as designable if at least one sequence folds into that structure. As a computational approximation, we defined the models that passed the quality filters after the first iteration of sequence design (*20*) as designable because they had good core packing, hydrogen bond satisfaction, and local sequence structure compatibility with the designed sequence. We projected the center and directions of the helices onto the underlying β sheets (Fig. 3B). The sampled helices from designable models at each position encompassed the distributions derived from native protein structures in the PDB (Fig. 3B, right panels). For the NTF2 fold, the distributions sampled in the designs were slightly shifted to the upper left when compared to the distributions in known structures (fig. S8). This difference could be a result of the presence of a C-terminal helix in our designs occupying the region shown at the right of the space projection, whereas C-terminal helices were often missing in the ensemble of known structures. Overall, because the number of known protein structures for a given topology is limited, the structure space covered by the known structures is much sparser than the space covered by the sampled structures. We quantified the size of structure space by dividing the six-dimensional space of helix centers and orientations into bins (*20*) (fig. S10). For the geometries sampled in this work, the known structures covered 12 to 26 bins, whereas LUCS-generated structures covered 63 to 221 bins (Fig. 3C); the smaller number of geometries in the NT designs (relative to the RO designs) could be a consequence of the additional C-terminal helix present on our NT designs restricting the accessible space of the two sampled helices. The 17 well-folded designs (Fig. 1D) sampled three to seven bins for each helix, respectively, and the majority (18/22) of these bins were not covered by known structures (Fig. 3D). All but one of the well-folded designs had at least one helix in a novel bin. Five well-folded designs had both helices in novel bins (Fig. 3E). Taken together, these results show that LUCS generates highly diverse geometries encompassing those found in nature but also exceeding known structure space, indicating that a large proportion of designable protein structure space remains unexplored.

We next sought to understand in more detail how the backbone geometries of the designed proteins were defined by the precise details of their noncovalent intramolecular interactions. The three experimentally solved Rossmann fold topology structures had distinct sequence patterns (Fig. 4A), resulting in distinct packing arrangements (Fig. 4, B and C) in their hydrophobic cores. The β sheets favored β-branched residues, as expected, whereas the side-chain sizes varied across different designs and resulted in differential hydrophobic packing. In particular, we observed previously described knob-socket–type packing motifs (*27*) (Fig. 4C and fig. S11) where nonpolar side chains fit into pockets formed by three residues on helices. These arrangements result in matched geometries between the side chains from sheets and helices that likely contribute to specifying the three-dimensional arrangement of the helices (*20*) (fig. S12). We also applied tertiary motif analysis using MASTER (*28*). For all well-folded designs, we were able to match tertiary motifs to both the designed loops and interacting secondary structure elements (fig. S13). Moreover, we identified side chains mediating helix-helix, helix-sheet, and helix-loop interactions that are similar in our designs and the corresponding matched tertiary motifs (Fig. 4D). Despite the close match between the local structures in the design and the tertiary motifs, the source proteins of the motifs had overall structures very different from the designs (fig. S13). Because tertiary motif information was not used directly in LHL backbone sampling or side-chain design, we conclude that recurrent tertiary motifs can be recapitulated solely by our LUCS sampling protocol and the Rosetta energy function (*22*).

Previous key achievements in de novo design (*11*–*15*, *21*) focused on designing one or a

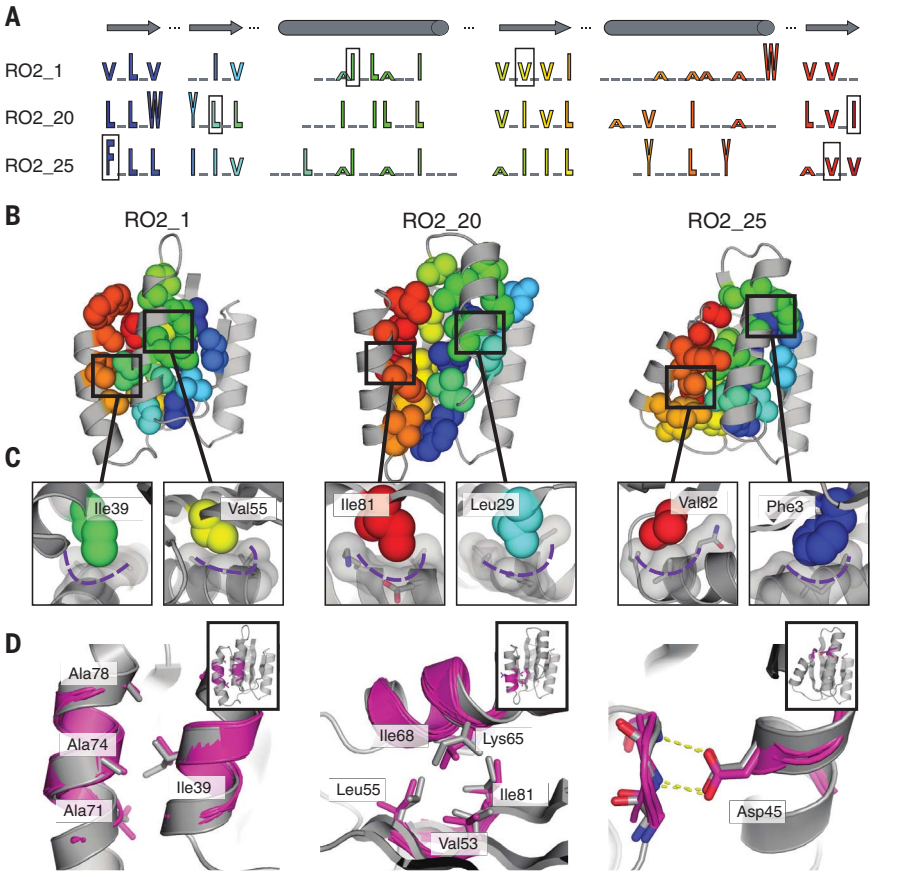


Fig. 4. Structural features encoding distinct protein geometries. (**A**) Sequence patterns of the hydrophobic cores in three designed models for the Rossmann fold, aligned by corresponding secondary structure elements (top). Hydrophobic residues are shown as letters in rainbow colors (A, Ala; F, Phe; I, Ile; L, Leu; V, Val; W, Trp; Y, Tyr) ordered by position in the primary protein sequence and scaled by side-chain size. Gray line segments indicate positions of surface-exposed polar residues. The residues in the boxes are the knob residues shown in (C). (**B**) Atomic packing of hydrophobic cores in the three experimentally determined structures for the Rossmann fold (Fig. 2). The hydrophobic side chains in the designed cores are shown as spheres. (**C**) Knob-socket packing motifs found in the designs. Three residues on a helix (gray sticks and surfaces) form a socket accommodating a knob residue, shown as colored spheres. (**D**) Examples of tertiary motifs matching the designed LHL structures. Designed structures are shown in gray; matched motifs are shown in magenta. Side chains of the best-matched tertiary motifs and design models are shown as sticks. Insets indicate location of the tertiary motif in the structure in the same orientation as in (B).

few structures for diverse non–helical bundle topologies by deriving design rules for specific topologies to identify the most favorable "idealized" geometries. This topology-centric strategy typically finds deep energy minima and thereby succeeds in overcoming errors in energy functions to create highly stable de novo folds. In contrast, natural and LUCS-generated structure families adopt non-ideal geometric features such as diverse helix positions, orientations, lengths, and conformations of connector elements, and exploring these non-ideal regions presents extra challenges (*29*). Nonetheless, we show here that LUCS achieves accurate atom-level control over diverse geometries, and our designs are not notably less stable than their de novo designed starting points (fig. S4). This success could at least partially be explained by the ability of LUCS to recover three-dimensional packing arrangements that are recurrent in nature (Fig. 4D and fig. S13), but without using this information as input.

We envision many applications for LUCS to precisely tune protein geometries for new protein functions that require atom-level control. The generalizable strategy underlying LUCS (Fig. 1C) does not require prior definition of structural variation based on design rules identified in native structures (*21*, *30*). New protocols could exploit this ability to flexibly tune protein geometries during design simulations while simultaneously building new functional sites for ligand binding or protein-protein recognition. The systematic sampling of protein geometries should also enable the design of dynamic proteins (*31*) that can switch between multiple distinct de novo designed conformations. Methods such as LUCS bring control over designable protein geometry space for arbitrary functions within reach.

REFERENCES AND NOTES

1. D. Baker, *Protein Sci.* **19**, 1817–1819 (2010).
2. K. Kundert, T. Kortemme, *Biol. Chem.* **400**, 275–288 (2019).
3. N. L. Dawson *et al.*, *Nucleic Acids Res.* **45**, D289–D295 (2017).
4. N. K. Fox, S. E. Brenner, J. M. Chandonia, *Nucleic Acids Res.* **42**, D304–D309 (2014).
5. J. Hou, S. R. Jun, C. Zhang, S. H. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3651–3656 (2005).
6. P. S. Huang *et al.*, *Science* **346**, 481–485 (2014).
7. T. M. Jacobs *et al.*, *Science* **352**, 687–690 (2016).
8. A. R. Thomson *et al.*, *Science* **346**, 485–488 (2014).
9. R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **33**, 745–754 (2000).
10. P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **282**, 1462–1467 (1998).
11. G. J. Rocklin *et al.*, *Science* **357**, 168–175 (2017).
12. N. Koga *et al.*, *Nature* **491**, 222–227 (2012).
13. P. S. Huang *et al.*, *Nat. Chem. Biol.* **12**, 29–34 (2016).
14. J. Dou *et al.*, *Nature* **561**, 485–491 (2018).
15. E. Marcos *et al.*, *Nat. Struct. Mol. Biol.* **25**, 1028–1034 (2018).
16. B. Kuhlman *et al.*, *Science* **302**, 1364–1368 (2003).
17. F. Crick, *Acta Crystallogr.* **6**, 689–697 (1953).
18. T. J. Brunette *et al.*, *Nature* **528**, 580–584 (2015).
19. D. Gront, D. W. Kulp, R. M. Vernon, C. E. Strauss, D. Baker, *PLOS ONE* **6**, e23294 (2011).
20. See supplementary materials.
21. E. Marcos *et al.*, *Science* **355**, 201–206 (2017).
22. H. Park *et al.*, *J. Chem. Theory Comput.* **12**, 6201–6212 (2016).
23. P. Bradley, K. M. Misura, D. Baker, *Science* **309**, 1868–1871 (2005).
24. J. A. Davey, R. A. Chica, *Methods Mol. Biol.* **1529**, 161–179 (2017).
25. N. Ollikainen, C. A. Smith, J. S. Fraser, T. Kortemme, *Methods Enzymol.* **523**, 61–85 (2013).
26. Y. Zhang, J. Skolnick, *Nucleic Acids Res.* **33**, 2302–2309 (2005).
27. H. Joo, A. G. Chavan, J. Phan, R. Day, J. Tsai, *J. Mol. Biol.* **419**, 234–254 (2012).
28. J. Zhou, G. Grigoryan, *Protein Sci.* **24**, 508–524 (2015).
29. D. Baker, *Protein Sci.* **28**, 678–683 (2019).
30. Y. R. Lin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5478–E5485 (2015).
31. J. A. Davey, A. M. Damry, N. K. Goto, R. A. Chica, *Nat. Chem. Biol.* **13**, 1280–1285 (2017).

ACKNOWLEDGMENTS

We thank M. Wu, N. Hoppe, and members of the Kortemme lab for discussion. **Funding:** Supported by NIH grant R01-GM110089 and NSF grant DBI-1564692 (T.K.); the UCSF Program for Breakthrough Biomedical Research, funded in part by the Sandler Foundation; a UCSF Discovery Fellowship (X.P.); and a NIH F32 Postdoctoral Fellowship (M.C.T.). T.K. is a Chan Zuckerberg Biohub Investigator. **Author contributions:** X.P. conceived the idea for the project; X.P. and T.K. conceived the computational and experimental approach; X.P. developed and performed the computational design; X.P. and Y.Z. performed the majority of the experimental characterization; X.P. and M.J.S.K. determined the NMR structures; X.P., M.C.T., L.L., and J.S.F. determined the crystal structure; J.S.F., M.J.S.K., and T.K. provided guidance, mentorship, and resources; and X.P. and T.K. wrote the manuscript with contributions from the other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Coordinates and structure files have been deposited to the Protein Data Bank (PDB) with accession codes 6VG7, 6VGA, 6VGB, and 6W90. NMR data have been deposited to the Biological Magnetic Resonance Data Bank with accession codes 30706, 30707, and 30708. All other relevant data are available in the main text or the supplementary materials. Rosetta source code is available from rosettacommons.org.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6507/1132/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S5
Data S1 and S2
References (*32*–*61*)

6 April 2020; accepted 14 July 2020
10.1126/science.abc0881

Caltech Division of Chemistry & Chemical Engineering

Assistant Professor of Chemistry or Biochemistry
Division of Chemistry and Chemical Engineering
California Institute of Technology
https://applications.caltech.edu/job/chemistry

The Division of Chemistry and Chemical Engineering at the California Institute of Technology invites applications for a **tenure-track faculty position at the assistant professor level in all areas of chemistry and biochemistry**. Candidates with strong commitments to research and teaching excellence are encouraged to apply. The term of the initial untenured appointment is four years, and the appointment is contingent upon completion of all requirements for a Ph.D. in chemistry, biochemistry, or a related field. Exceptionally well-qualified applicants at the tenured level may also be considered.

Interested candidates should apply at **https://applications.caltech.edu/job/chemistry**. Applications should include a curriculum vita, publication list, teaching statement, a description of proposed research, and three letters of recommendation. Applicants should submit a diversity and inclusion statement that discusses past and/or anticipated contributions to improving diversity, equity, and inclusion in the areas of research, teaching, and/or outreach. Applications should be received by **October 1, 2020**.

We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to age, race, color, religion, sex, sexual orientation, gender identity, or national origin, disability status, protected veteran status, or any other characteristic protected by law.

Caltech Division of Chemistry & Chemical Engineering

Assistant Professor of Chemical Engineering
Division of Chemistry and Chemical Engineering
California Institute of Technology
https://applications.caltech.edu/job/chemeng

The Division of Chemistry and Chemical Engineering at the California Institute of Technology invites applications for a **tenure-track faculty position at the assistant professor level in the area of chemical engineering**. Candidates with strong commitments to research and teaching excellence are encouraged to apply. The term of the initial appointment is four years and is contingent upon completion of all requirements for a Ph.D. in chemical engineering or in a related field.

Interested candidates should apply at **https://applications.caltech.edu/job/chemeng**. Applications should include a curriculum vita, publication list, teaching statement, a description of proposed research, and three letters of recommendation. Applicants should submit a diversity and inclusion statement that discusses past and/or anticipated contributions to improving diversity, equity, and inclusion in the areas of research, teaching, and/or outreach. Applications should be received by **November 1, 2020**.

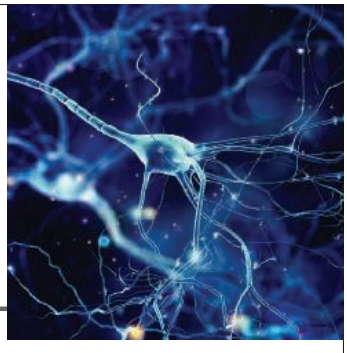
We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to age, race, color, religion, sex, sexual orientation, gender identity, or national origin, disability status, protected veteran status, or any other characteristic protected by law.

SPECIAL JOB FOCUS:

Neuroscience

Issue date: Oct. 2
Reserve ad space by Sept. 17

Ads accepted until Sept. 25 if space allows



Produced by the *Science*/AAAS Custom Publishing Office.

To book your ad, contact: advertise@sciencecareers.org

The Americas
+1 201 748 6702
Europe & ROW
+44 (0) 1273 810850
Japan
+81 3 6459 4174
Greater China, South Korea, Singapore, Thailand
+86 131 4114 0012

Why choose this job focus for your advertisement?

- Relevant ads lead off this career section with a special neuroscience banner.

ScienceCareers AAAS

SCIENCECAREERS.ORG



FOR RECRUITMENT IN SCIENCE, THERE'S ONLY ONE *SCIENCE*.

POST-DOCTORAL FELLOWSHIP IN LNCRNA BIOLOGY

A postdoctoral position is available immediately in the Vascular Biology Center at the Medical College of Georgia at Augusta University (**https://www.augusta.edu/centers/vbc** for the study of long noncoding RNAs in cultured cells as well as genetically tractable mouse models (see PMID IDs **30822570, 30584103, 29976770, 28292896, 28062498, 27444199, 27102963**). The lab is particularly focused on lncRNA discovery in, but not limited to, the cardiovascular system. The newly developed state-of-the-art lab has demonstrable expertise in molecular biology, various screening tools, RNA-FISH, CRISPR/Prime editing in mice, next generation sequencing, and bioinformatics analyses. In addition, the lab harbors its own LSM900 Zeiss confocal microscope, histopathology room for frozen/paraffin sections, and physiology room for contractile studies. The successful candidate will work in a highly interactive and collegial group that fosters rigorous learning and preparation for faculty-hood, including grant writing and submission. The lab is extraordinarily well-funded with internal startup as well as multiple NIH grants thus allowing for ample opportunities to develop an independent research program for future job-hunting here in the USA or abroad. The successful candidate will not only enjoy unusually rich resources, but a higher than average salary with excellent benefits.

The candidate must have demonstrable expertise in molecular biology and biochemistry (eg, RNA pulldown, CLIP-seq) and a yearning for scientific excellence and independence.

Strongly motivated candidates should send an updated CV, a brief statement of research interests, and 3 references to **jmiano@augusta.edu**.

Augusta University/ AU Health is an equal opportunity employer including individuals with disabilities and veterans.

By **Sieun Ruth Lee**

One, two, three, thrive

It was a sunny Friday afternoon, and my smiling labmates were heading home for the weekend. I, however, was glumly sitting in Tristan's office, staring at the spreadsheet on which we had outlined our goals for the week. I shaded my cells red; I had not been able to complete any of them. I was disappointed in myself. Tristan gave me a reassuring smile, reminding me not to beat myself up too much. "It's OK to be ambitious, but make sure your three weekly goals are realistic and measurable," he advised. It was guidance I was still learning to follow.

Halfway through my second year of grad school, I had been feeling overwhelmed by a sea of experimental failures and lost without the structure I was accustomed to as an undergrad. When my mentor Tristan—a postdoc—saw that I was struggling, he suggested an approach from an unlikely source: his experience in the world of elite sport. Introduced to him by a coach during his time as a member of the Australian men's national field hockey team, it's called Three to Thrive (T2T), referring to three goals set each week to help guide progress.

We set up a shared spreadsheet where we would plan our goals, which can be personal as well as academic. At the end of the week, we would meet and discuss our progress in the "Friday Finish." That's how I found myself, a few weeks in, sitting down with Tristan that Friday afternoon.

As we talked through my weekly goals and why I hadn't been able to complete them, we realized the goals themselves were the problem. It was easy for me to see where I wanted to be in 1 year, or five, but I had a hard time figuring out what weekly steps would get me there. As a result, I was setting goals too high to achieve in a week. Since then, Tristan has helped me learn to break up my big-picture goals into smaller tasks that are challenging but realistic, and I've started to make some good progress on my research.

With T2T, I have been able to regain some control over my life. On Sunday afternoons, I sit down at a coffee shop—or, these days, in my cozy plant room sipping coffee—and set my goals, mentally preparing myself for the week ahead. Then, during the week, I can focus on working hard to achieve those goals, not wanting to disappoint myself—or Tristan. T2T has also helped Tristan and me develop a much more honest and authentic mentor-mentee relationship, connecting on everything from cryogenic electron microscopy techniques to tips for not eating too much ice cream.



"'Make sure your three weekly goals are realistic and measurable,' he advised."

Here are five keys for a successful T2T experience.

SET APPROPRIATE GOALS

Goal setting is a skill. Your weekly T2T goals need to be measurable and achievable but challenging at the same time.

BE HONEST

If you don't accomplish a goal, that's OK. But don't make excuses to yourself. Be upfront about why you fell short so you can figure out how to do better in the future. Similarly, if your accountability partner isn't fulfilling their role, you have to be comfortable holding them to account.

MAKE IT ROUTINE

T2T is not a groundbreaking concept; staying accountable and making it habitual are key to its power. Set time aside in your weekly schedule for goal setting and reflection.

RECORD YOUR PROGRESS

Logging your weekly goals and progress allows you to look back and see what you've achieved—or what prevented you from reaching your goals. The exercise can be encouraging when you've had a tough week, reminding you of the progress you've made. Or it can serve as a wake-up call if things aren't going the right direction.

START

With all good habits, sometimes the hardest bit is just getting started. You won't have it perfect from the beginning. But with some practice, you'll figure out what works for you. ■

Sieun Ruth Lee is a Ph.D. student at Northwestern University in Evanston, Illinois. She thanks Tristan Clemons for his mentorship. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS