NASA kicks off first round trip to Mars p. 1416

Reducing SARS-CoV-2 airborne transmission with masks p. 1422

A nearby system of super-Earth exoplanets pp. 1432 & 1477

Science

$15
26 JUNE 2020
sciencemag.org

AAAS

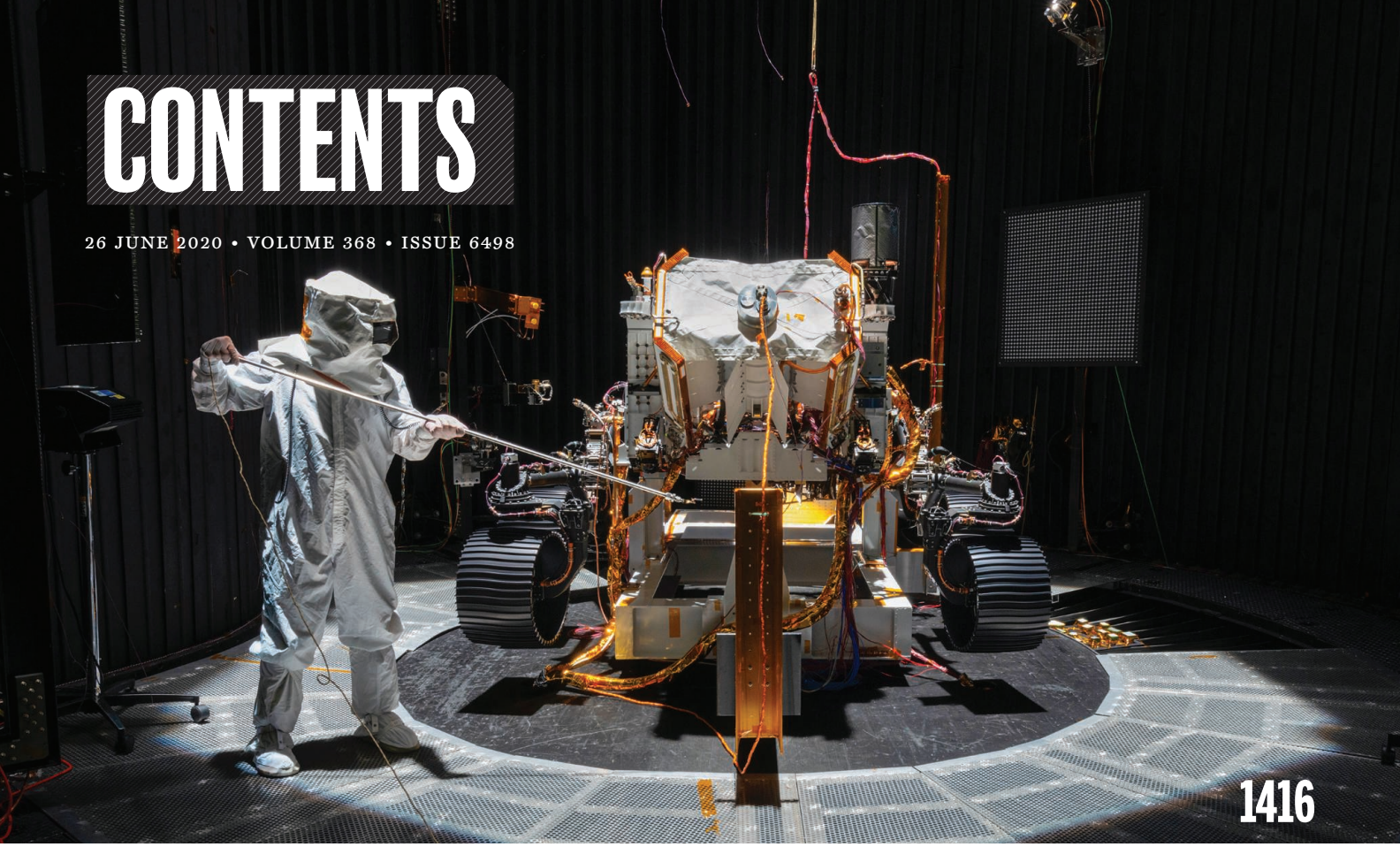


ARCTIC DOGS

Ancient DNA separates wolves from dogs p. 1495

CONTENTS

26 JUNE 2020 • VOLUME 368 • ISSUE 6498



1416

PHOTO: NASA/JPL-CALTECH

NEWS

IN BRIEF

1406 News at a glance

IN DEPTH

1408 Hope grows for targeting the brain with ultrasound
Noninvasive method of tweaking neural activity enters new tests in human *By K. Servick*

1409 Lawsuit alleges scientific misconduct at weapons lab
Dispute underscores the challenges of using simulations to "test" warheads *By A. Cho*

1411 Dawn of the chicken revealed in Southeast Asia
Chicken domestication—debated since Charles Darwin—tracked by genomic survey *By A. Lawler*

1412 Charities that fund research face deep revenue declines
Cancellation of major fundraisers forces groups to lay off staff and delay, trim, or cancel grants to scientists *By E. Cahan*

1413 Could a blood 'observatory' stop pandemics?
Proposal calls for screening blood from around the world for thousands of antibodies *By R. Bazell*

1414 Online GRE test heightens equity concerns
Pandemic spurs graduate programs to drop standardized test requirements *By J. C. Hu*

1415 Senate bill to curb foreign threats raises alarms
Sweeping changes to protect research could hamper academic collaboration, critics say *By J. Mervis*

FEATURES

1416 Martian chronicler
NASA's Perseverance rover aims to find out whether ancient Mars was warm and wet or cold and dry *By P. Voosen*

1420 Mars mission would put China among space leaders
By D. Normile

INSIGHTS

PERSPECTIVES

1422 Reducing transmission of SARS-CoV-2
Masks and testing are necessary to combat asymptomatic spread in aerosols and droplets *By K. A. Prather* et al.
PODCAST

1424 Evolution after genome duplication
Genetic interactions in yeast reveal factors governing duplicate gene retention and divergence
By I. M. Ehrenreich
RESEARCH ARTICLE p. 1446

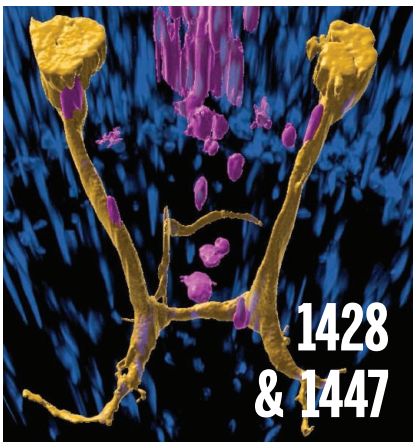
1426 Action spectra of chiral secondary structure
Mass spectra of DNA complexes photoreduced with polarized light reveal their handedness *By P. Barran*
REPORT p. 1465

1427 The simplicity of robust light harvesting
A universal design principle underlies photosynthetic antenna systems
By C. D. P. Duffy
REPORT p. 1490

1428 Guide cells help navigate axon regeneration
Rebuilding the flatworm visual system after injury requires guidepost-like muscle cells *By R. Roberts-Galbraith*
RESEARCH ARTICLE p. 1447

1430 Microplastic in terrestrial ecosystems
Research shifts from ecotoxicology to ecosystem effects and Earth system feedbacks
By M. C. Rillig and A. Lehmann

Published by AAAS



Muscle guidepost-like cells (magenta) affiliate with the axons of the planarian visual system in yellow (nuclei, blue).

1432 Multiple, quiet, and close by
A planetary system around a nearby star offers the prospect for atmospheric study
By M. B. Davies
REPORT p. 1477

POLICY FORUM

1433 "Explaining" machine learning reveals policy challenges
The need to make objectives explicit may expose policy trade-offs that had previously been implicit and obscured
By D. Coyle and A. Weller

BOOKS *ET AL.*

1435 Confronting bias
Prejudice and discrimination are deeply personal, but their effects are profound
By C. K. Lai

1436 Like other species, we have always been on the move
Human migration and cultural exchange represent progress, not peril, argues a journalist *By K. Koser*

LETTERS

1438 COVID-19's unsustainable waste management
By S. You et al.

1438 Misguided forest action in EU Biodiversity Strategy
By N. Selva et al.

1439 Afforestation falls short as a biodiversity strategy
By S. Gómez-González et al.

RESEARCH

IN BRIEF

1443 From *Science* and other journals

RESEARCH ARTICLES

1446 Genomic duplications
Exploring whole-genome duplicate gene retention with complex genetic interaction analysis *E. Kuzmin* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAZ5667
PERSPECTIVE p. 1424

1447 Developmental biology
Muscle and neuronal guidepost-like cells facilitate planarian visual system regeneration
M. L. Scimone et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABA3203
PERSPECTIVE p. 1428

1448 Neuroscience
Flexible recruitment of memory-based choice representations by the human medial frontal cortex *J. Minxha* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABA3313

1449 Chromatin sequencing
Single-molecule regulatory architectures captured by chromatin fiber sequencing
A. B. Stergachis et al.

1454 Structural biology
Cryo-EM structure of the human cohesin-NIPBL-DNA complex *Z. Shi* et al.

1460 Stuctural biology
Mechanisms of OCT4-SOX2 motif readout on nucleosomes *A. K. Michael* et al.

REPORTS

1465 Spectroscopy
Mass-resolved electronic circular dichroism ion spectroscopy *S. Daly* et al.
PERSPECTIVE p. 1426

1469 Solar physics
Meridional flow in the Sun's convection zone is a single cell in each hemisphere
L. Gizon et al.

1472 Plasmonic materials
Micelle-directed chiral seeded growth on anisotropic gold nanocrystals
G. González-Rubio et al.

1477 Exoplanets
A multiple planet system of super-Earths orbiting the brightest red dwarf star GJ 887
S. V. Jeffers et al.
PERSPECTIVE P. 1432

1481 Coronavirus
Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China *J. Zhang* et al.

1487 Quantum optics
Metalens-array–based high-dimensional and multiphoton quantum source *L. Li* et al.

1490 Photosynthesis
Quieting a noisy antenna reproduces photosynthetic light-harvesting spectra *T. B. Arp* et al.
PERSPECTIVE p. 1427

1495 Dog genomics
Arctic-adapted dogs emerged at the Pleistocene–Holocene transition
M.-H. S. Sinding et al.
PODCAST

1499 Coronavirus
Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir *W. Yin* et al.

DEPARTMENTS

1405 Editorial
Persuasive words are not enough
By H. Holden Thorp

1506 Working Life
The pressure to assimilate *By Montrai Spikes*

ON THE COVER



Greenland sled dogs do the same job their ancestors have done for countless generations. Sequencing of a 9500-year-old dog genome reveals that modern dogs—particularly sled dogs—were genetically distinct from wolves by the early Holocene. Of all modern sled dogs, the Greenland sled dog shares the most genomic ancestry with the first sled dogs, emphasizing the conservation value of this breed. See page 1495. *Photo: Carsten Egevang*

Science Staff 1402
AAAS News & Notes 1440
Science Careers 1505

PHOTO: M. LUCILA SCIMONE, DENIZ ATABAY, PETER REDDIEN, MIT, WHITEHEAD INSTITUTE

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

News Editor Tim Appenzeller

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

Creative Director Beth Rakouskas

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

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

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

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

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

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

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

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

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

---

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

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

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

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

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

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

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

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

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

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

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

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

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

Published by AAAS

EDITORIAL

Persuasive words are not enough



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

Communicating the findings of science plays a vital role in shaping our lives and the planet. From my vantage point at the *Science* family of journals, I witness daily how editors, journalists, and scientists work together to deliver scientific information to the world. Given the recent rapid pace of discovery, relaying remarkable findings has never been busier. So why aren't these efforts having a bigger positive effect on the public acceptance of science?

It's baffling that as the world struggles to tame coronavirus disease 2019 (COVID-19), a large portion of the population ignores the fact that wearing masks and practicing social distancing dampen the spread of the pandemic. Even when a vaccine for COVID-19 becomes available, the benefits of achieving herd immunity will be endangered if growing antivaccine sentiment leads some folks to refuse to get vaccinated. American science denialism, in particular, persists, even at the highest level of leadership, with a president who denies climate change and a vice president—a devout creationist—who believes that Earth is only 6000 years old.

I hear a lot from our readers and stakeholders about how to solve this problem of science denial. Most of them suggest that *Science* and science "need to do a better job at telling its stories." I don't buy it. For more than a century, this journal has been delivering insightful and reliable scientific information; today, our articles have the highest readership ever. Sure, we can do a better job of simplifying messages and making them accessible to more people. There is always room for improvement. But is this really the crux of this dangerous problem?

The scientific community is up against a sophisticated, data-driven machine that is devoted to making sure that science doesn't fully succeed, and the history of this is quite clear. A recent *Science* editorial pointed out that U.S. Republican politicians embraced Earth Day when it started 50 years ago. But in the 1980s, digital analysis of political polling data based on location fostered the formation of the anti-science movement in the United States—politicians and their supporters who did not like the results criticized the findings and the process. It became fashionable and politically expedient to run against science. Any carve-outs for the environment in the conservative government went away in favor of the dismantling of regulations grounded in evidence. Over time, digital technologies have become more sophisticated, and now there is a massive, churning, finely tuned digital misinformation machine that has seized social media to ensure that a portion of the population doesn't accept science. And this battle between science fact and fiction isn't just being waged in the United States—the United Kingdom, Russia, India, and Brazil all face a similar predicament.

> **"The scientific community is losing the battle against... misinformation."**

The current implications of this battle in the United States are everywhere. The administration has promulgated the idea that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, the cause of COVID-19) was engineered in China's Wuhan Institute of Virology, in part based on a non–peer-reviewed preprint that was later retracted. The misinformation about masks and social distancing is spurring dangerous bar gatherings and choir practices. Unsubstantiated claims in a "plandemic" video are convincing citizens that Dr. Anthony Fauci, the longtime leader of the U.S. National Institute of Allergy and Infectious Diseases, is hiding a secret business deal from which he stands to profit from COVID-19. The anti-science movement started with the environment, which could hurt our long-term survival, but in the era of COVID-19, it threatens our immediate survival.

The scientific community is losing the battle against this digital leviathan of misinformation. A well-reasoned and highly placed op-ed on this topic is not going to move the needle, no matter how well it is crafted to adhere to the best practices in science communication. Neither is a perfect trade book, television appearance, or speaking tour by a scientific leader. The only way to win this fight is to harness the same sophisticated tools in the name of science that are being used to tear science down. With social media companies afraid to challenge the misinformation machine, even when their own platforms are being misused, the task is daunting. But we can at least move on from the idea that if we could just find those perfect, persuasive words, everyone would suddenly realize that facts are facts with no alternatives.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

10.1126/science.abd4085

Published by AAAS

NEWS



People exercise on 15 June behind plastic sheets at a gym in Redondo Beach, California, that reopened under state guidelines that call for continuing social distancing. Public health experts worry the reopenings are increasing COVID-19 infections in many U.S. states.

PHOTO: FREDERIC J. BROWN/AFP VIA GETTY IMAGES

IN BRIEF

Edited by **Jeffrey Brainard**

DISPATCHES FROM THE PANDEMIC

Mask study draws fire

PUBLIC HEALTH | Dozens of epidemiologists are calling for the retraction of a paper claiming that face masks are "the most effective means" to prevent transmission of the virus that causes COVID-19. In the study, published on 11 June in the *Proceedings of the National Academy of Sciences*, Nobel laureate and atmospheric chemist Mario Molina of the University of California, San Diego, and colleagues analyzed COVID-19 case counts and mask policies in areas with large outbreaks, concluding that social distancing and other measures to control coronavirus transmission were "insufficient by themselves." Last week, more than 40 scientists signed a letter to the journal acknowledging the importance of masks but calling the study flawed. For example, they argue it didn't properly account for an "enormous set of changes across society" that could have affected COVID-19 incidence, independent of mask use. The scientists also ask that the journal reassess its long-standing policy of allowing U.S. National Academy members such as Molina to select their own reviewers for some submissions. (Neither of the paper's listed reviewers is an epidemiologist.) Molina has defended the paper, saying the critics failed to understand the role of aerosol particles in viral transmission and simply distrust work by scientists outside their field.

Antimalarial drug trials end

CLINICAL RESEARCH | The World Health Organization (WHO) and U.S. National Institutes of Health (NIH) decided separately last week to stop large clinical trials evaluating whether hydroxychloroquine can safely and effectively treat people hospitalized with COVID-19, an idea touted by U.S. President Donald Trump early in the pandemic. In an interim analysis announced on 20 June, NIH's trial data and safety monitoring board said the antimalarial drug was safe but "very unlikely to be beneficial." WHO cited similar grounds on 17 June in permanently halting the hydroxychloroquine arm of its Solidarity trial. Both decisions came after investigators running the U.K. Recovery trial announced on 5 June they had stopped giving hydroxychloroquine to patients because the data showed no benefit. That study, the largest investigating drugs to treat COVID-19, continues.

Steroid shortage feared

DRUG THERAPY | The World Health Organization (WHO) has called for increased production of dexamethasone, a commonly used corticosteroid recently shown to reduce death rates of COVID-19. "Hoarding and speculative procurement appear to have already started," says Emer Cooke, WHO's head of regulation of medicines and other health technologies. But she says it's "probably too early to say if there will be a global shortage." The greatest concern is for the injectable version of the drug, which some physicians say is the preferred formulation, especially for critically ill patients. But it is more complicated to produce than oral dexamethasone, and a major Indian manufacturer has faced regulatory scrutiny for production problems.

SCIENCEMAG.ORG/TAGS/CORONAVIRUS

Read additional *Science* coverage of the pandemic.

Published by AAAS

PARTICLE PHYSICS

Another step toward a new megacollider

CERN, the European particle physics laboratory near Geneva, announced last week it will study the technical and financial feasibility of an atom smasher 80 to 100 kilometers in circumference. To be built in the 2040s, the megacollider would succeed the lab's 27-kilometer Large Hadron Collider (LHC), smashing electrons into positrons to study in detail the Higgs boson, the famous particle the LHC discovered in 2012. But the 23 nations that fund CERN have made no commitment to build the mammoth new collider, for which researchers produced a rough design 18 months ago, says CERN Director-General Fabiola Gianotti. The feasibility study should be done by 2026 or 2027. The electron-positron collider could open the way to later building an even more powerful proton smasher like the LHC in the same tunnel.

NSF director starts work

LEADERSHIP | Sethuraman "Panch" Panchanathan began his 6-year term this week as head of the National Science Foundation (NSF). The 59-year-old computer scientist, who had been a senior administrator at Arizona State University, Tempe, was confirmed by the U.S. Senate on 18 June, 6 months after President Donald Trump nominated him to be the agency's 15th director. Kelvin Droegemeier, Trump's science adviser, had been double hatting as acting NSF director since France Córdova completed her term on 31 March.

Critics ask to strip Watson name

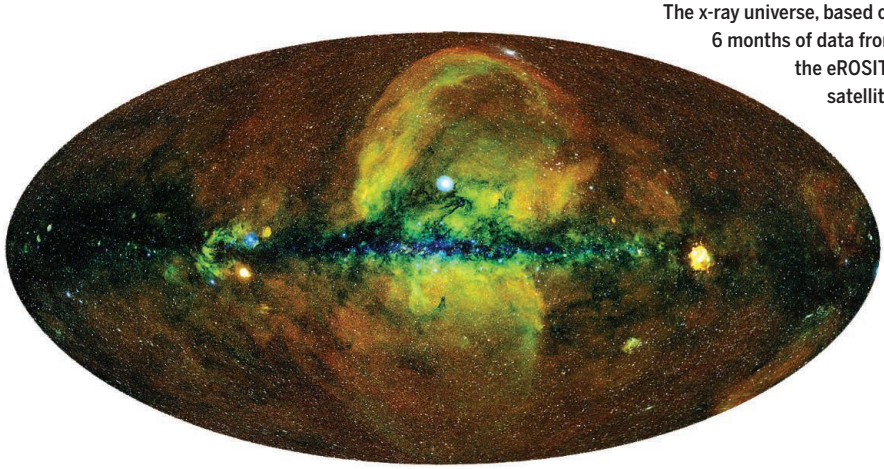
RACIAL JUSTICE | Spurred by the recent groundswell of the Black Lives Matter movement, scores of students and alumni of a prestigious doctoral program are asking its trustees to remove Nobel laureate James Watson's name from the program, called the Watson School of Biological Sciences. The request from 133 students and alumni of the Cold Spring Harbor Laboratory (CSHL) program, in a letter sent 21 June, says Watson has made racist comments since at least 2007, when a newspaper quoted him saying Black people were of inferior intelligence, leading CSHL to remove him as chancellor. In 2019, after he told PBS his views had not changed, the lab stripped his remaining honors except for the school's name. Watson's name, the letter states, "is now inextricably linked with racism." It adds: "This name change should be the first step in a larger process to make [CSHL] an inclusive and supportive environment, especially for people of color." A spokesperson for Marilyn Simons, chair of CSHL's Board of Trustees, wrote in an email that changing the school's name requires 75% of the board to vote to amend CSHL's charter; a meeting is scheduled for early July. Watson, 92, who co-discovered DNA's structure and was CSHL's director and president for decades, could not be reached for comment.

Trump suspends high-tech visas

IMMIGRATION | President Donald Trump's new executive order limiting the entry of foreign workers through the end of the year spares two visa categories important to U.S. universities but halts new visas in a third category. The 22 June executive order, which Trump says will protect U.S. jobs, does not affect the Optimal Practical Training program that allows graduating international students to stay in the country for up to 3 years. It also exempts visiting researchers and scholars coming under the J-1 visa program. But the order does block the issuance of any new so-called H-1B visas, a category that both universities and high-tech companies use to attract skilled foreign workers. None of the new restrictions applies to visa holders already in the country. Google CEO Sundar Pichai tweeted that he's "disappointed" by the order and that the company "will continue to stand with immigrants."

X-rays reveal black hole bonanza

ASTRONOMY | An orbiting x-ray telescope launched last year has released its first all-sky image, recording more than 1 million objects, twice as many sources as previously detected in 60 years of x-ray astronomy. Most of the dots in the image recorded by eROSITA (extended Roentgen Survey with an Imaging Telescope Array) are supermassive black holes at the centers of galaxies, where they gorge on gas heated so hot that it can be seen across the universe. The project's 4-year goal is to map millions of such galaxies to understand how gravity gloms them together in clusters and how the mysterious dark energy pushes against gravity to accelerate expansion of the universe.



The x-ray universe, based on 6 months of data from the eROSITA satellite.

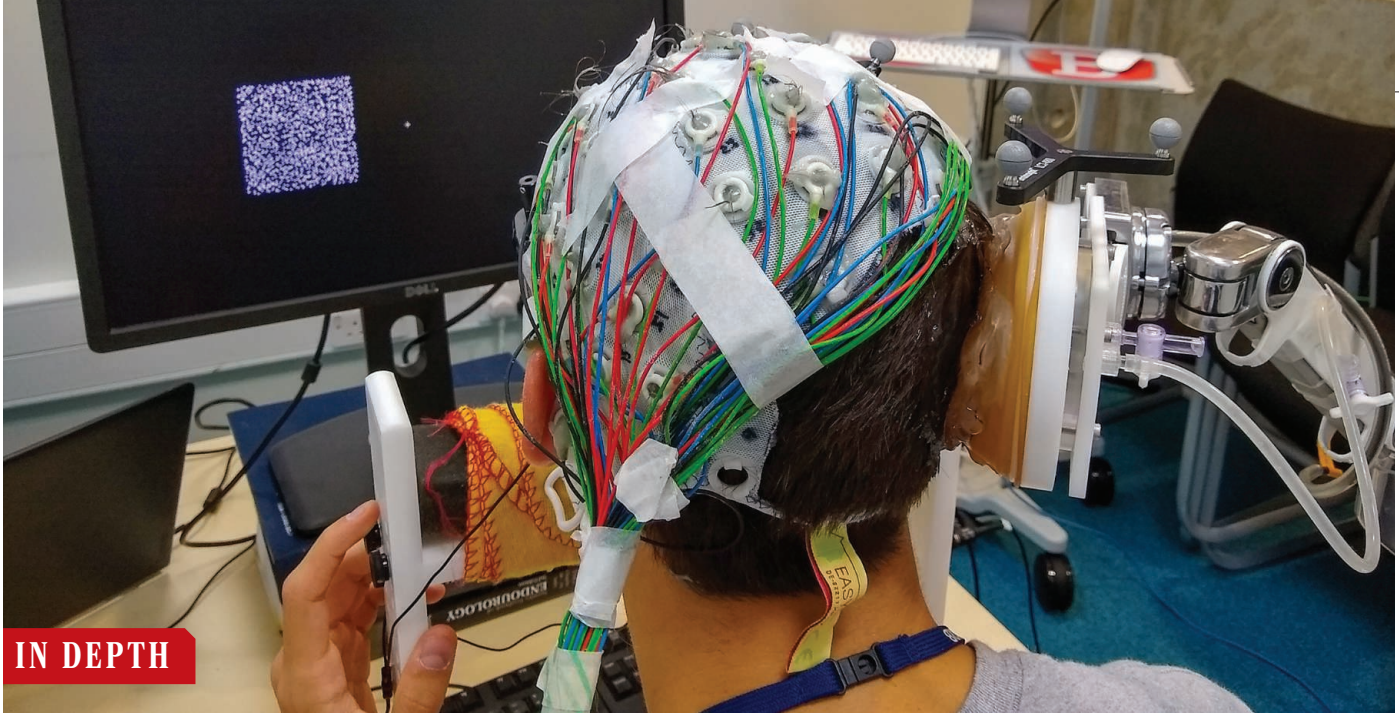
IMAGE: MPE/IKI

BY THE NUMBERS

38°C

Temperature on 20 June in the town of Verkhoyansk, Russia, the highest ever recorded north of the Arctic Circle. The average June high there is just 20°C. Unusually warm temperatures in Siberia this year have melted permafrost and fed wildfires.

Published by AAAS



IN DEPTH

Pulses of ultrasound aimed at a motion-sensitive brain area (from brown device, right) improved people's ability to judge dots' direction of motion.

PHOTO: CHRIS BUTLER

NEUROSCIENCE

Hope grows for targeting the brain with ultrasound

Noninvasive method of tweaking neural activity enters new tests in humans

By **Kelly Servick**

As a way to see inside the body, revealing a tumor or a fetus, ultrasound is tried and true. But neuroscientists have a newer ambition for the technology: tinkering with the brain. At frequencies lower than those of a sonogram but still beyond the range of human hearing, ultrasound can penetrate the skull and boost or suppress brain activity. If researchers can prove that ultrasound safely and predictably changes human brain function, it could become a powerful, noninvasive research tool and a new means of treating brain disorders.

How ultrasound works on the brain remains mysterious. But recent experiments have offered reassurance about safety, and small studies hint at meaningful effects in humans—dampening pain, for example, or subtly enhancing perception. "I've seen a lot of tantalizing data," says Mark Cohen, a neuroscientist at the University of California, Los Angeles (UCLA). "While the challenges are very large, the potential of this thing is so much larger that we really have to pursue it."

Scientists can already modulate the brain noninvasively by delivering electric current or magnetic pulses across the skull. The U.S. Food and Drug Administration (FDA) has approved transcranial magnetic stimulation (TMS) to treat depression, migraine pain, and obsessive-compulsive disorder (OCD).

But unlike magnetic or electric fields, sound waves can be focused—like light through a magnifying glass—on a point deep in the brain without affecting shallower tissue. For now, that combination of depth and focus is possible only with a surgically implanted wire. But ultrasound could temporarily disrupt a deep human brain region—the almond-shaped amygdala, a driver of emotional responses, for example, or the thalamus, a relay station for pain and regulator of alertness—to test its function or treat disease.

Results in animals are encouraging. Experiments in the 1950s first showed ultrasound waves could suppress neural activity in a visual region of the cat brain. In rodents, aiming ultrasound at motor regions has triggered movements such as a twitch of a paw or whisker. And focusing it on a frontal region of monkey brains can change how the animals perform at eye movement tasks.

But it's technically tricky to aim ultrasound through thick, dense skull bone and to show its energy has landed at the intended point. And ultrasound's effects on the brain can be hard to predict. How much it boosts or suppresses neural activity depends on many parameters, including the timing and intensity of ultrasound pulses, and even characteristics of the targeted neurons themselves. "I have tremendous excitement about the potential," says Sarah Hollingsworth Lisanby, a psychiatrist at the National Institute of Mental Health who studies noninvasive neuromodulation. "We also need to acknowledge that there's a lot we have to learn," she says.

For one thing, researchers are largely in the dark about how sound waves and brain cells interact. "That's the million-dollar question in this field," says Mikhail Shapiro, a biochemical engineer at the California Institute of Technology. At high intensities, ultrasound can heat up and kill brain cells—a feature neurosurgeons have exploited to burn away sections of brain responsible for tremors.

Even at intensities that don't significantly increase temperature, ultrasound exerts a mechanical force on cells. Some studies suggest this force alters ion channels on neurons, changing the cells' likelihood of firing a signal to neighbors. If ultrasound works primarily via ion channels, "That's great news," Shapiro says, "because that means we can look at where those channels are expressed and make some predictions about what cell types will be excited." In a preprint on bioRxiv last month, Shapiro's team reported that exposing mouse neurons in a dish to ultrasound opens a particular set of calcium ion channels to render certain cells more excitable.

But these channels alone won't explain ultrasound's effects, says Seung-Schik Yoo, a neuroscientist at Harvard University. He notes that ultrasound also appears to affect

Published by AAAS

receptors on nonneuronal brain cells called glia. "It's very hard to [develop] any unifying theory about the exact mechanism" of ultrasound, he says.

Regardless of mechanism, ultrasound is starting to show clear, if subtle, effects in humans. In 2014, a team at Virginia Polytechnic Institute and State University showed focused ultrasound could increase electrical activity in a sensory processing region of the human brain and improve participants' ability to discern the number of points being touched on their fingers. Neurologist Christopher Butler at the University of Oxford and colleagues have tested ultrasound during a more complex sensory task: judging the motion of drifting, jiggling dots on a screen. Last month at the Cognitive Neuroscience Society's annual meeting online, he reported that stimulating a motion-processing visual region called MT improved subjects' ability to judge which way the majority of the dots drifted.

Ultrasound's effects have so far been subtler than those of TMS, says Mark George, a psychiatrist at the Medical University of South Carolina, who helped develop and refine that technology. With TMS, "you put it on your head and turn it on and your thumb moves," he says. But the ultrasound experiments that prompted paw twitches in mice used intensities "so, so, so much higher than what we're being allowed to use in humans."

> **"While the challenges are very large, the potential of this thing is so much larger that we really have to pursue it."**
> **Mark Cohen**, UCLA

Regulators have limited human studies in part because ultrasound has the potential to cook the brain or cause damage through cavitation—the creation of tiny bubbles in tissue. In 2015, Yoo and colleagues found microbleeds, a sign of blood vessel damage, in sheep brains repeatedly exposed to ultrasound. "This was a huge speed bump," says Kim Butts Pauly, a biophysicist at Stanford University. But in February in *Brain Stimulation*, her group reported microbleeds in control animals as well, suggesting this damage might result from dissection of the brains. Butts Pauly and Yoo now say they're confident the technology can be used safely.

Cohen and collaborators recently tested safety in people by aiming ultrasound at regions slated for surgical removal to treat epilepsy. With FDA's OK, they used intensities up to eight times as high as the limit for diagnostic ultrasound. As they reported in a preprint on medRxiv in April, they found no significant damage to brain tissue or blood vessels. However, to find the limit of safety, researchers will likely need to go all the way to levels that damage tissue, Cohen says.

Several teams are cautiously moving into tests of ultrasound as treatment. In 2016, UCLA neuroscientist Martin Monti and colleagues reported that a man in a minimally conscious state regained consciousness following ultrasound stimulation of his thalamus. Monti is preparing a publication on a follow-up study of three people with chronically impaired states of consciousness. After ultrasound, they showed increased responsiveness over a period of days—much faster than expected, Monti says, although the study included no control group.

That research and the tests in epilepsy patients used an ultrasound device developed by BrainSonix Corporation. Its founder, UCLA neuropsychiatrist Alexander Bystritsky, hopes ultrasound can disrupt neural circuits that drive symptoms of OCD. A team at Massachusetts General Hospital and Baylor College of Medicine is planning a study in humans using the BrainSonix device, he says.

Columbia University biomedical engineer Elisa Konofagou hopes to use ultrasound to treat Alzheimer's disease. Before COVID-19 interrupted participant recruitment, she and colleagues were preparing a pilot study to inject tiny gas-filled bubbles into the bloodstream of six people with Alzheimer's and use pulses of ultrasound to oscillate the microbubbles in blood vessels lining the brain. The mechanical force of those vibrations can temporarily pull apart the cells lining these vessels. The researchers hope opening this blood-brain barrier will help the brain clear toxic proteins. (Konofagou's team and others are also exploring this ultrasound-microbubble combination to deliver drugs to the brain.)

In his first test of ultrasound after years of studying TMS, George looked to reduce pain. His team applied increasing heat to the arms of 19 participants, who tended to become more sensitive over repeated tests, reporting pain at lower temperatures by the last test. But if, between the first and last test, they had pulses of ultrasound aimed at the thalamus, their pain threshold dipped half as much. "This is definitely a double green light" to keep pursuing the technology, George says.

George regularly treats depressed patients with TMS and has seen the technology save lives. "But everybody wonders if we could go deep with a different technology—that would be a game changer," he says. "Ultrasound holds that promise, but the question is can it really deliver?" ■

NATIONAL LABORATORIES

Lawsuit alleges scientific misconduct at weapons lab

Dispute underscores the challenges of using simulations to "test" warheads

By **Adrian Cho**

An unusual lawsuit alleges scientific misconduct at Lawrence Livermore National Laboratory in California, one of the United States's three nuclear weapons labs. Peter Williams, a 50-year-old physicist, worked at Livermore from January 2016 until May 2017, when he says he was fired in retaliation for complaining that his superiors were mishandling a computer program that simulates the detonation of high explosives, undermining their ability to predict how a particular nuclear weapon would perform if used. Williams, who now works at a private research lab, has sued Livermore and seven individuals for reinstatement and $600,000 in damages.

Researchers familiar with the labs say Williams's allegations should be taken seriously. "If there's been a cover-up, that's something that ought to be looked into," says Raymond Jeanloz, a geophysicist at the University of California, Berkeley, who has been involved with the weapons labs. But he also says the labs implement internal reviews and other measures to ensure the integrity of their work and head off the kind of problem Williams alleges. "This is exactly the kind of thing the people at the lab worry about," Jeanloz says. Livermore declined to comment on the suit, but in a statement said: "Rigorous debate is a part of the scientific process—the Laboratory does not retaliate against individuals for holding differing opinions."

The suit, which Williams filed on 22 May, seems quixotic. He is representing himself; to make his case, he needs documents that only the lab can provide; and his complaint centers on a differential equation. Williams spent only a short time at Livermore before

Published by AAAS



Peter Williams claims tweaks to a program for modeling a bomb's explosive trigger make its predictions unreliable.

he was fired. (In a 12-month performance review Williams included in his suit, his superiors state he wasn't keeping up with assignments.) Before joining the lab he did two postdocs, taught at City College of San Francisco and Sonoma State University, and worked for 8 years at Agilent Technologies. But Williams is a talented scientist, says Craig Wheeler, an astrophysicist at the University of Texas, Austin, who was his graduate adviser and has subsequently published with him. "He's a deep, independent thinker," Wheeler says. "He's definitely not a crackpot."

At Livermore, Williams was given the unclassified task of modeling the behavior of a high explosive. Called PBX 9502, the polymer-bonded explosive is intended to be used in refurbishing a 40-year-old thermonuclear warhead called the W80, which is about the size of a garbage can and fits on a cruise missile. The explosive, when detonated, would compress the warhead's plutonium pit to set off a nuclear fission explosion, which in turn would trigger an even more powerful fusion explosion. Modeling a high explosive is difficult, researchers say, because its behavior depends on its density, which can vary from batch to batch and within a single piece. But such modeling is crucial to be sure that stored weapons will work as intended because the United States gave up nuclear testing in 1992 under the still unratified Comprehensive Nuclear-Test-Ban Treaty.

PBX 9502 is especially tricky because it's an insensitive explosive, meaning it won't go up if smacked or set on fire. But that also means that as the detonation front moves through the material, the chemical reactions behind it progress more slowly than in older explosives, Williams says. So instead of assuming the reactions are instantaneous, researchers must model their evolution, too.

To make sure they can do so correctly, researchers detonate samples of the stuff in different geometries, observe the explosions with ultra–high-speed cameras and other tools, and compare the results with simulations. For example, in one test, the sample resembles a mallet, with a thin cylindrical handle that leads to a fatter cylindrical head. Because the detonation front cannot instantly turn the corner where skinny cylinder meets fat, the explosion doesn't consume the entire sample, but leaves a ring of material whose thickness reveals the speed of the detonation front and chemical reactions.

> **"The motive was to keep money flowing to help develop and improve CHEETAH."**
>
> **Peter Williams**, formerly of Lawrence Livermore National Laboratory

Livermore used such test explosions to refine a modeling program called ARES-CHEETAH. But Williams says he became aware that one of his supervisors was essentially changing parameters in the program after the fact to make sure the simulation fit the data from each experimental setup. That's "pseudoscience," Williams charges, as it guarantees the model will look accurate, even if it isn't. He says he repeatedly asked the researcher to explain the rationale for the tweaking. "It became clear that he didn't have a rationale, he had a motive," Williams says. "And the motive was to keep money flowing to help develop and improve CHEETAH."

Such "curve fitting" renders the program incapable of making meaningful predictions, he says. If CHEETAH can't properly model the bench tests of PBX 9502, larger, far more expensive integrated test explosions of the weapon's nonnuclear components could fail, he says. And modelers may not be able to correctly calculate the yield of the refurbished weapon, he says.

That would be a big problem, says Robert Rosner, a theoretical astrophysicist at the University of Chicago and former director of Argonne National Laboratory who is on Livermore's advisory board. The uncertainty in the yields of various weapons directly affects calculations of how big the U.S. nuclear arsenal should be and could determine the course of future arms control negotiations, Rosner says. "The extent to which we can decrease [the stockpile] is completely related to the ability to be certain about the yield."

But Rosner says he's neither surprised nor alarmed to hear researchers had to adjust the high-explosives models to fit different experiments. What Williams sees as unacceptable tweaking is an inevitable part of the empirical modeling of any complex system, such as the climate or a star, Rosner says. "The idea that the model has to be tuned, I find that unsurprising," he says. "I would be stunned if it didn't."

Both Rosner and Jeanloz say they suspect Williams may not know the whole story behind ARES-CHEETAH, as the lab dismissed him before he obtained a clearance to work on classified material. The weapons program also takes explicit measures to ensure that researchers adhere to scientific standards, Jeanloz says, such as running its own classified peer-reviewed journals and encouraging researchers to keep a hand in unclassified research. The rivalry between Livermore and Los Alamos National Laboratory in New Mexico provides a particularly valuable check, scientists say.

Yet Jeanloz and Rosner agree that models like CHEETAH need careful scrutiny. That's because some policymakers may view uncertainties in modeling as a reason to resume nuclear testing. In fact, Jeanloz says, some arms control experts argue that to discourage calls for new tests, the United States should make no changes to its existing nuclear arms—even ones that make them safer, such as switching to insensitive explosives.

Williams does not yet have a court date. He's suing, he says, not out of anger, but out of a sense of scientific duty: "I couldn't look myself in the mirror if I didn't do it." ■

PHOTO: CLAUDIA WILLIAMS

Published by AAAS

Corrected 25 June 2020. See full text.

ANIMAL DOMESTICATION

Dawn of the chicken revealed in Southeast Asia

Chicken domestication—debated since Charles Darwin—tracked by genomic survey

By **Andrew Lawler**

It is the world's most common farm animal as well as humanity's largest single source of animal protein. Some 24 billion strong, it outnumbers all other birds by an order of magnitude. Yet for 2 centuries, biologists have struggled to explain how the chicken became the chicken.

Now, the first extensive study of the bird's full genome concludes that people in northern Southeast Asia or southern China domesticated a colorful pheasant sometime after about 7500 B.C.E. Migrants and traders then carried the bird across Asia and on to every continent except Antarctica.

"Our results contradict previous claims that chickens were domesticated in northern China and the Indus Valley," researchers led by Ming-Shan Wang from the Chinese Academy of Sciences's Kunming Institute of Zoology write in a paper published today in *Cell Research*. They also found that the modern chicken's chief ancestor is a subspecies of red jungle fowl named *Gallus gallus spadiceus*.

"This is obviously a landmark study," says Dorian Fuller, an archaeologist at University College London who was not involved in the effort. He adds that the results could shed light on the emergence of agriculture and early trade networks, and what features of the bird made it so attractive to people.

Charles Darwin argued the chicken descended from the red jungle fowl because the birds resemble each other and can make fertile offspring; he speculated that domestication happened in India. But five varieties of the pheasant inhabit a broad arc extending from the jungles of Indonesia to the Himalayan foothills of Pakistan. Which variety led to the chicken, and where, was uncertain. Based on presumed chicken bones, archaeologists claimed, variously, that people domesticated the bird 9000 years ago in northern China and 4000 years ago in Pakistan.

DNA studies promised to resolve the issue, but researchers had few samples from the bird's wild relatives. So Jianlin Han, a geneticist at the Joint Laboratory on Livestock and Forage Genetic Resources, embarked on a 20-year project to sample indigenous village chickens and wild jungle fowl near more than 120 villages across Asia and Africa (*Science*, 23 November 2012, pp. 1020 and 1022).

Wang's team sequenced the full genomes of 863 birds and compared them. The results suggest modern chickens descend primarily from domesticated and wild varieties in what is now Myanmar, Laos, Thailand, and southern China (see map, below). "This region is a center of domestication," says co-author and geneticist Olivier Hanotte of the University of Nottingham. The results confirm a hypothesis put forward in 1994 by Japan's Crown Prince Akishino, an ornithologist, on the basis of mitochondrial DNA data.

Wang's team did find some evidence for a South Asian contribution: A jungle fowl native to the Indian subcontinent may have interbred with the chicken after its initial domestication in Southeast Asia, the team says.

The new DNA data link domesticated chickens most closely to the Southeast Asian subspecies *G. g. spadiceus*, however. They suggest the lineage that became the modern chicken branched off from the jungle fowl between 12,800 and 6200 years ago, with domestication occurring sometime after the lineages split.

Fuller doubts the bird was fully domesticated before the arrival of rice and millet farming in northern Southeast Asia about 4500 years ago. Hanotte acknowledges that "we need the help of archaeologists" to understand the human events that triggered domestication.

But Jonathan Kenoyer, an archaeologist and Indus expert at the University of Wisconsin, Madison, remains skeptical that the chicken arose in Southeast Asia. "They need to get ancient DNA" to back up their claims, he says, because genomes of modern birds may provide limited clues to early events in chicken evolution.

Nor does the DNA show what first enticed people to tame the bird. Early varieties were far scrawnier and produced fewer eggs than today's industrial varieties, and their predators were legion. Some researchers suggest the bird was initially prized for its exotic plumage or for cockfighting. Selling prize fighting cocks remains a lucrative business in Southeast Asia, and the birds' high value may have spurred traders to carry them long distances.

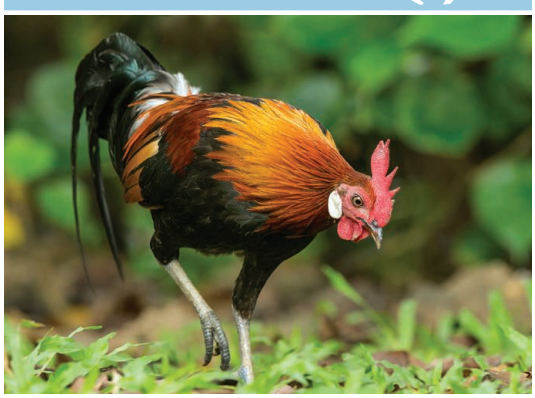
Smithsonian Institution archaeozoologist Melinda Zeder calls the new paper "fascinating" and says it shows "the domestication and dispersal story is more complicated than we thought." She urges combining genetic and archaeological data to flesh out the tale. Archaeologists are now gathering chicken bones that suggest farmers in southern China and Southeast Asia first domesticated the bird some 3500 years ago—findings that bolster the genetic work.

Han's group, meanwhile, is creating a massive data set based on more than 1500 modern chicken genomes from Asia, Europe, and Africa. The researchers plan to analyze chicken dispersal into Europe and Africa, as well as the genetic variations behind traits such as the ability to withstand disease or produce more eggs. "This study opens a whole new page in chicken genomics," Han says. ■

Early bird

One subspecies of the red jungle fowl (*Gallus gallus spadiceus*, bottom), found in northern Southeast Asia, likely led to the first domesticated chickens, genomes of wild and tame birds show.





CREDITS: (MAP) X. LIU/*SCIENCE*; (PHOTO) DAVID IRVING

Published by AAAS



A fundraiser in which firefighters collect donations for muscular dystrophy research was canceled this year.

COVID-19

Charities that fund research face deep revenue declines

Cancellation of major fundraisers forces groups to lay off staff and delay, trim, or cancel grants to scientists

By **Eli Cahan**

Early this year, University of Colorado, Denver, cancer researcher Patricia Ernst was thrilled when her postdoc Therese Vu won a grant from the Leukemia & Lymphoma Society, a nonprofit that has pumped more than $1.2 billion into blood cancer research since its founding in 1949. The funding would allow the scientists to launch studies using a technique to generate malignant leukemia from immature blood cells—an approach that Ernst had been eager to try for more than a decade. Then, last month, the pair got bad news: The philanthropic group canceled the grant, citing "unprecedented" revenue losses caused by the COVID-19 pandemic.

"I did anticipate there would be cutbacks," Ernst says. "But I didn't think it would be that serious, and I didn't think it would happen to us."

Many researchers are having similar experiences. Foundations that fund biomedical research in the United States, the United Kingdom, and elsewhere are reporting record revenue drops because of the pandemic. One major factor: It has forced them to cancel key fundraising events, including galas, walks, Broadway partnerships, and even an event that sends thousands of U.S. firefighters into the streets, asking passersby to support medical research by dropping donations into a rubber boot. Many groups are trying to stem the losses by cutting staff and delaying, trimming, or outright canceling grants to researchers.

The chaos imperils a small, but pivotal, part of the scientific ecosystem. Although nonprofits provide just 5% of overall U.S. research funding, they often support small, high-risk pilot studies that later enable researchers to attract larger grants from government funders. And many of the grants go to young researchers, helping them launch their careers. "If you're in a room with researchers of vascular disease, almost all of them will say their first grant came from [us]," says Mariell Jessup, chief science and medical officer at the American Heart Association (AHA).

Science's COVID-19 coverage is supported by the Pulitzer Center.

So far, Jessup says, AHA has been fortunate: Although donations have dropped, the $890 million organization hasn't had to lay people off or rescind grants—but it has postponed awarding a new round of grants.

The red ink is drowning other U.S. groups. At the National Multiple Sclerosis Society, which last year spent about $40 million of its $190 million budget on research, officials forecast a $60 million shortfall in 2020; they've given 78 of their 198 grantees a 15% "haircut." Susan G. Komen, the largest nonprofit funder of breast cancer research, has laid off about 20% of its 211 employees, closed 30 of its 61 local affiliates, and tabled future grant cycles. The Muscular Dystrophy Association, which counts on its annual "Fill the Boot" fundraiser with firefighters for 25% of its annual revenue, has furloughed many of its development staff and canceled plans to award new grants. At the $724 million American Cancer Society (ACS), a $200 million drop in revenue has prompted it to lay off 1000 of its some 4300 employees. "If current trends continue," chief medical officer William Cance says, ACS could temporarily cut research funding by half.

In the United Kingdom, the Association of Medical Research Charities (AMRC), whose members last year sent £1.9 billion to biomedical researchers (compared with £1.8 billion in funding provided by the U.K. government), is reporting a 38% drop in fundraising revenue. That has forced groups to withdraw or defer grants, says AMRC CEO Aisling Burnand. Cancer Research UK, which funds half of the country's noncommercial cancer research, has cut its funding by about 10%, or £44 million, says CEO Michelle Mitchell. The cuts will get deeper if charities do not receive more support from the government, she adds. And the crisis could have long-lasting ripple effects on the next generation of research. "We're in danger of destroying a decade's worth of work, infrastructure, and future talent," Mitchell says.

Vu, for example, had hoped to have gathered enough data from her pilot study by October to apply for a grant from the National Institutes of Health (NIH). Now, even if she can find replacement funding, she thinks it will be an additional 12 to 18 months before she can apply to NIH. And because the leukemia researcher is from Australia, a funding cutoff could imperil her U.S. work visa. "I don't want to be all 'woe is me,' but … the junior people have gotten hammered" by the disruption, Vu says.

The loss of nonprofit grants could also hurt researchers seeking funding for high-risk ideas that can't get support from government funders, says Maryrose Franko, CEO of the Health Research Alliance, which represents 85 nonprofit research funders. "We derisk research for the government, and we embrace failure," Franko says. "If we're not funding it, who will?" ■

With reporting by Cathleen O'Grady.

PHOTO: ZUMA PRESS INC./ALAMY STOCK PHOTO

Corrected 24 June 2020. See full text.

Published by AAAS

COVID-19

Could a blood 'observatory' stop pandemics?

Proposal calls for screening blood from around the world for thousands of antibodies

By **Robert Bazell**

Michael Mina is out for blood—millions of samples, which a nascent effort dubbed the Global Immunological Observatory (GIO) would monitor for signs of pathogens spreading through the population. Instead of a telescope, it will rely on technology that can measure hundreds of thousands of distinct antibodies in a microliter of blood. If the GIO can overcome technical and logistical hurdles and find sustained funding, he says, it could provide a powerful tool for monitoring and responding to disease outbreaks.

For now, the idea is just a pilot project to track the spread of COVID-19. The stealthy spread of that disease through the population underscored the need for such a monitoring system, says Mina, an immunologist and epidemiologist at Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health, who with colleagues outlines the GIO concept this month in *eLife*. (The co-authors include Jeremy Farrar, an infectious disease specialist and director of the Wellcome Trust, as well as vaccine and immunology specialists Adrian McDermott and Daniel Douek of the National Institutes of Health.)



Antibodies in donated blood or other samples could reveal where previously identified pathogens are returning or new ones emerge.

PHOTO: MOUSSA81/ISTOCK.COM

Disease surveillance in the United States now relies on a patchwork of hospitals, clinics, and doctors to report unusual events to state health departments, which pass the information on to the Centers for Disease Control and Prevention. The need for faster, more comprehensive surveillance, Mina says, "was starkly clear with the inability to identify and model local circulation of COVID-19 in a timely fashion."

Mina wants to watch for outbreaks by looking for antibodies to infectious agents in regularly collected, anonymized blood samples from every possible source—blood banks, plasma collection centers, even the heel needle sticks of newborns, which are taken in most states from every baby in order to identify genetic diseases. The samples would be identified only by geographical area. Chip-based platforms that can identify hundreds of thousands of antibodies are already produced commercially by companies including VirScan and Luminex. Mina says these could easily be scaled up to look at huge numbers of samples, either individually or in batches.

"This is an extraordinary and exciting concept," says infectious disease specialist William Schaffner of the Vanderbilt University Medical Center. "It is an example of the kind of fresh new thinking we need in public health." But, Schaffner adds, "The logistical challenges for such an endeavor could be daunting."

Mina and his co-authors envision initially testing about 10,000 samples per day and later, if they secure funding to build up the project, some 100,000 per day for the United States alone. Even the smaller number could detect—far faster than the current reporting system—an outbreak of Zika virus in rural Louisiana, for example, or an eruption of West Nile virus in Colorado. The GIO could also accelerate the monitoring of seasonal influenza, allowing hospitals to prepare for patient surges and for public health officials to efficiently distribute vaccines.

When a new infectious disease such as COVID-19 appears, the GIO could track its spread. The antibody-detecting chips wouldn't necessarily have to be updated to spot a new pathogen, such as SARS-CoV-2, the virus that causes COVID-19. Researchers might see a rise in antibodies that nonspecifically target known pathogens—and might flag their unknown relatives. For example, a burst of antibodies that cross-react with known coronaviruses would likely have been seen in Chinese cities, such as Wuhan, as SARS-CoV-2 began to spread there.

The idea of regularly monitoring entire populations for antibodies arose in the lab of evolutionary biologist Bryan Grenfell at Princeton University, where Mina worked as a postdoctoral fellow. Now, Mina has joined Grenfell and Jessica Metcalf, also an evolutionary biologist at Princeton, in expanding the concept. Antibodies can signal not just people who are currently infected, but also those who had the disease and recovered. The GIO would also be able to distinguish different strains of a bacterium or virus infecting people because each produces a unique antibody signature.

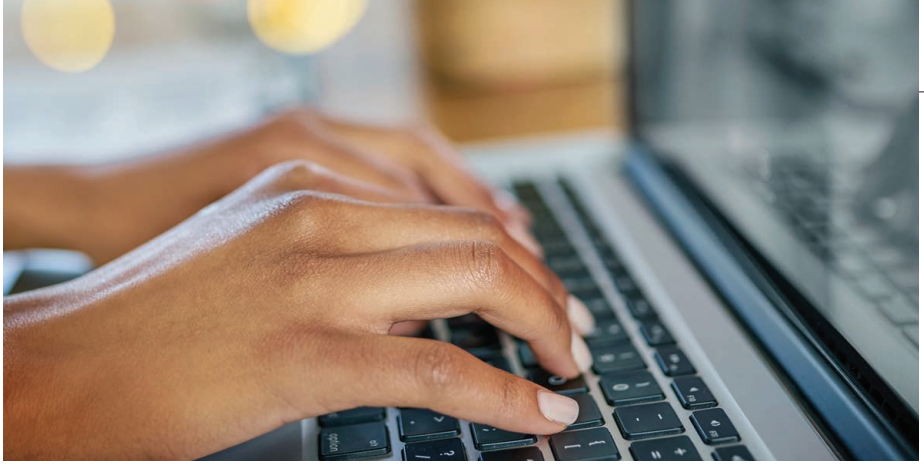
The GIO team is already building a pilot laboratory in Massachusetts, while it looks to secure financial support. "Given the importance we believe this could have, we are beginning to look for funding from some of the major philanthropic donors of public health work," Mina says. "We are currently exploring and open to all options."

Meanwhile, the team's pilot project, supported by the Open Philanthropy foundation, is gathering many thousands of anonymous blood samples from the plasma-collecting company Octapharma. By screening them for antibodies to SARS-CoV-2, Mina and his colleagues hope to learn how useful widespread antibody testing can be in tracing the spread of the new coronavirus and possibly predicting future "hot spots" or localized outbreaks.

People often do not develop antibodies until well after infections; for SARS-CoV-2 it takes 1 or 2 weeks. But Mina says the antibody testing still provides valuable information. "A week into an outbreak isn't huge," he said. "For example, if we were doing this with [blood from] just a small fraction of New York, we would have detected that [SARS-CoV-2] was there in February and could have given [Governor Andrew] Cuomo plenty of ammunition to close down the city March 1 instead of March 19." ■

Robert Bazell, an adjunct professor at Yale University, is a journalist based in New Haven, Connecticut.

Published by AAAS



At-home GRE test takers must have a computer, private room, and stable internet connection.

PHOTO: PEOPLEIMAGES/ISTOCK.COM

SCIENTIFIC COMMUNITY

Online GRE test heightens equity concerns

Pandemic spurs graduate programs to drop standardized test requirements

By **Jane C. Hu**

As COVID-19 swept across the United States, standardized testing centers closed and the GRE General Test—an exam that's required for admission to many U.S. graduate programs—went online. The Educational Testing Service (ETS), which offers the GRE, "completely revamped its delivery model so [aspiring graduate students] can test from the safety of home," it declared in May. Since then, though, scores of academics have raised concerns about the equity of the online version of the test, arguing it disadvantages prospective students from rural and low-income backgrounds. "If I were … a student trying to take this exam, satisfying [the online testing] criteria would be extremely difficult for me," says Emily Levesque, an assistant professor of astronomy at the University of Washington, Seattle.

> **"It was simply a question of access."**
>
> **Emily Levesque**, University of Washington, Seattle

Levesque wrote about ETS's testing requirements in a Twitter thread this month, detailing what she sees as "a shopping list of hurdles." Test takers must have access to a computer with a webcam—"tablets and smartphones won't cut it," she wrote—as well as a private room in a home with a stable internet connection. Libraries and other public spaces are out. "We already know from virtual teaching this spring that not all students/prospective grads have access to [computers] in their homes," she wrote.

On top of that, test takers must have a whiteboard if they want to take notes, sit in a standard—not "overstuffed"—chair, and ensure that no one enters their room for the duration of the 4-hour test. In a statement, Alberto Acereda, executive director of higher education at ETS, wrote that the rules are "necessary to ensure the testing experience is similar to that in a test center, as well as to maintain the security and integrity of the test."

Yet to some academic departments already questioning the value of the GRE, the burdensome requirements of the at-home test are a tipping point. Levesque's department decided to temporarily suspend requiring GRE scores. "It was simply a question of access," she says. "If we require the exam this year, that puts an excessive burden on folks we want to encourage to apply."

Other departments have decided to forgo the GRE for good. "We've been thinking about [eliminating it] for a long time," says Chrissy Wiederwohl, assistant department head for engagement and graduate affairs for Texas A&M University, College Station's oceanography department, which voted to stop requiring GRE scores earlier this month. "COVID is what helped frontburner it."

Levesque's and Wiederwohl's departments join a growing list of U.S. graduate programs that have moved away from the GRE in recent years. In 2018 alone, 44% of the country's top molecular biology programs dropped the GRE as an application requirement, according to an investigation by *Science* (31 May 2019, p. 816). Dubbed "GRExit," the movement has been fueled by concerns that the GRE doesn't predict student success in graduate school, and that its use in admissions decisions disadvantages applicants from underrepresented groups.

Delia Shelton, a Black psychologist at the University of Tennessee, Knoxville, says when she was applying to graduate schools, taking the GRE was a hardship; it was expensive and she had to drive 2 hours to her testing location. Her scores prevented her from applying to some programs that specified a cutoff for applicants. Yet she's done well for herself in academia, winning a prestigious fellowship from the U.S. National Institutes of Health and an assistant professorship at her university. She and her colleagues voted earlier this month to eliminate the GRE from admissions requirements. The scores "don't speak to how well you can do in graduate school," she says.

"The socioeconomic constraints [of] standardized testing … [are] well documented, and I think this at-home test exacerbates some of those," agrees Joshua Hall, director of admissions for the biological and biomedical science program at the University of North Carolina, Chapel Hill, who maintains a list of roughly 300 life science programs that have joined GRExit. Hall estimates that 15 programs have contacted him since the start of the pandemic asking to be added to his list.

Acereda emphasized that GRE scores can be valuable as part of a holistic review process, one that looks at reference letters, essays, and other application materials. He acknowledged that some prospective students might be unable to take the test online, but he added that more than 1000 test centers have already reopened. "As the world continues to reopen after COVID, test takers will have greater choice regarding where they would like to test."

This year, however, some simply gave up. Natasha Hodges says she ran into problems when she couldn't install the proctoring software on her Apple laptop. "No matter how many people I chatted with, or how many times I've called or emailed them, no one can explain to me or even address [my problem]," she says. But after failing to resolve her technical issues, she was pleased to discover that many of the microbiology Ph.D. programs she wants to apply to have waived the GRE as an application requirement. "It ends up working out that I don't end up having to take it anyway," she says. ■

Jane C. Hu is a science journalist in Seattle.

Published by AAAS

SCIENCE AND SECURITY

Senate bill to curb foreign threats raises alarms

Sweeping changes to protect research could hamper academic collaboration, critics say

By **Jeffrey Mervis**

A bipartisan group of U.S. senators last week proposed sweeping—and controversial—changes in how the federal government manages academic research in the face of foreign threats.

The authors of the legislation, more than 1 year in the making, claim it will stop China and other countries from stealing the fruits of federally funded research without weakening a system that has made the United States a global leader in innovation. But research advocates worry the bill, if enacted, would restrict the exchange of talent and ideas.

Drafted by Senators Rob Portman (R–OH) and Tom Carper (D–DE) and with eight Republican and five Democratic co-sponsors, the Safeguarding American Innovation Act is the latest and most substantive attempt in Congress to reconcile scientific security and openness. One contentious provision would give the State Department grounds to reject a visa application from anyone with ties to a foreign government seen as hostile to the United States. Critics worry such power could be used to keep out the tens of thousands of Chinese graduate students and postdocs who seek to study in the United States each year.



Senator Rob Portman (R–OH) targets China in new legislation.

Other provisions would expose scientists who fail to disclose ties to foreign governments to criminal penalties including jail time, require international partners to embrace U.S. scientific norms, lower the size of foreign gifts that universities must report, and give the White House budget office new powers to oversee research security.

"For nearly 2 decades, the federal government has been asleep at the wheel while foreign governments have exploited the lack of transparency in our education system and bought access and influence," says Portman, who leads the Permanent Subcommittee on Investigations, which has issued several reports sharply critical of current federal efforts. "This bill will help us stop foreign governments from stealing our research and innovation."

The sponsors don't hide their intended target. "America's research enterprise is the best in the world and the Chinese Communist Party knows it," says Senator Josh Hawley (R–MO). "That's why they've spent the last 20 years stealing American taxpayer-funded intellectual property."

Carper, the top Democrat on the investigative panel and a co-sponsor, uses more judicious language. The legislation, he says, is a "common sense [approach] to protect American intellectual property and better leverage our international research partnerships."

A Portman press release claims "widespread support" from academia. But all of the supportive statements come from institutions in his home state of Ohio, and many also hint at the need to tweak some of its provisions. "We endorse [its] goals to modernize the safety and security of our nation and we look forward to continuing to collaborate with Sen. Portman as the legislation moves forward," says Barbara Snyder, who is stepping down this fall as president of Case Western Reserve University in Cleveland to lead the Association of American Universities (AAU), a coalition of 65 major research institutions.

In private, research advocates express grave reservations. "It violates the culture of openness that is fundamental to academic research," one says. "I don't think the higher education community is going to like any of this," says another, who, like many advocates, requested anonymity because they were not authorized to speak for their organization.

Although the bill's language is subtle, it contains "key provisions … [that] are overly broad and will only serve to harm American science without improving national security," says AAU's Tobin Smith. One such provision, Smith and others say, would give the State Department the authority to reject a visa application from anyone based on their "cooperation with … military organizations adversarial to the United States, foreign institutions involved in the theft of United States research, [or] a government that seeks to undermine the integrity and security of the United States research community."

But staffers on the subcommittee that Portman leads say the language would apply to fewer researchers than the critics fear. "The focus of the bill is on bad actors," one staffer noted. "The vast majority of foreign researchers [asking to come to the United States] are benign, and we need their talents."

The bill would also empower the State Department to reject or restrict the activities of a visa applicant if officials decide that giving the applicant access to "goods, technologies, or sensitive information" would harm the United States. Extensive rules already limit the sharing and export of research products deemed sensitive. But lobbyists say the new provision could require universities to impose additional restrictions on visa holders, such as blocking them from attending open lectures or visiting laboratories doing unclassified research.

Again, the staffers accuse the research community of overreacting. "We're not locking down campuses," one staffer says. But universities and other federally funded institutions "don't need to give everyone access to everything."

There is no companion bill in the House of Representatives, and the House's Democratic leadership is more skeptical that the problem of foreign influence warrants wholesale legislative changes. Little time remains for the Senate to act on the legislation before the November elections. But research advocates expect the debate in Congress to continue regardless of the outcome of the vote. ■

PHOTO: SENATE TELEVISION/AP IMAGES

Published by AAAS



MARTIAN CHRONICLER

NASA's Perseverance rover aims to find out whether ancient Mars was warm and wet or cold and dry

By **Paul Voosen**

NASA's newest Mars rover, Perseverance, is going back in time to the bottom of a vanished lake. If all goes well, in February 2021 it will land in Jezero crater and pop the dust covers off its camera lenses. Towering in front of it, in all likelihood, will be a 60-meter cliff of mudstone: the edge of a fossilized river delta. These lithified martian sediments could hold answers to urgent questions about the earliest days of Earth's chilly, parched neighbor: How did this pint-size planet, so distant from a faint young Sun, support liquid water on its surface? How much water was there, and how long did it persist? And did Mars ever spawn life?

The 45-kilometer-wide crater is an intriguing target. Billions of years ago, when life was just beginning on Earth, water broke through its western rim and spilled into its interior, carrying sediments that settled and piled up in thick, meandering braids that today can be seen from space, as plain as day. "It's kind of like the Mississippi delta, but smaller," says Raymond Arvidson, a planetary geologist at Washington University in St. Louis. The water filled the crater like a bathtub until, 250 meters deep, it breached the eastern rim. And then, just as mysteriously as it arrived, the water disappeared.

Scientists have traced the tracks of ancient water across Mars ever since the 1970s, when orbiters revealed branching valley networks that matched the dendritic shape of water-eroded valleys on Earth. In the 1990s, the Mars Global Surveyor zoomed in on deeply incised gullies that could only have been

PHOTO: NASA/JPL-CALTECH

Published by AAAS



Perseverance is thermally tested in artificial sunlight. It will explore more than 500 million years of martian climate.

carved by powerful flows of water—and may even have glimpsed shorelines from an ancient ocean. Later orbiters found evidence of abundant clay-bearing minerals that need water to form. More recently, the Curiosity rover, Perseverance's predecessor, has charted the existence of a long-lived lake at the bottom of its adopted home, Gale crater.

Some scientists believe the water shows ancient Mars was warm for millions of years, a favorable climate for life to emerge. Others say the climate was cold and dry, punctuated by sporadic bursts of water that only lasted for hundreds or thousands of years—a much more difficult environment for life to take root. Along with the question of past life, says Ken Farley, the mission's project scientist and a geologist at the California Institute of Technology (Caltech), Mars's ancient climate "is the biggest unanswered question."

Perseverance will tackle both questions, although the search for life will take longer. The rover, developed by NASA's Jet Propulsion Laboratory (JPL) and set for launch next month from Cape Canaveral Air Force Station in Florida, is also the start of an audacious campaign that will ferry to Earth about 30 samples of martian rock and grit. Perseverance will gather the samples, and NASA and the European Space Agency (ESA) are designing two follow-up missions to retrieve them, aiming for launches in 2026 (*Science*, 22 November 2019, p. 932).

The complex mechanisms needed to drill and store these cores limited the room on board for tools to chemically analyze samples and look for organic molecules. Until the samples reach labs on Earth, the question of whether life once existed in Jezero will probably go unanswered. "We'll have to be patient," says Tanja Bosak, a geobiologist at the Massachusetts Institute of Technology and member of the rover's science team.

The story of the martian climate, on the other hand, will be etched across Jezero's surface, visible to an array of rover instruments. Scientists can only make a rough guess at the lake's age, but they think it formed 3.8 billion years ago, about the same time as the valley networks, over hundreds or thousands of years. Unlike Curiosity's target, Gale crater, which offers a snapshot of a moment some 3.5 billion years ago when Mars was likely drying out, Jezero and its surroundings will grant access to more than 500 million years of martian history, including some of the planet's oldest terrain, says Bethany Ehlmann, a Caltech planetary scientist and member of the science team. "We have the potential for a really rich history of climate."

BY DESIGN, PERSEVERANCE borrows much from Curiosity: a six-wheeled chassis the size of a small SUV, an imaging turret, a radioisotope power source. "From the outside it looks the same," says Allen Chen, one of the rover's lead engineers at JPL. "But it's got it where it counts." That includes advanced new imaging instruments, landing capabilities, and a complex drilling system—innovations that led its budget, originally pitched as a bargain at $1.5 billion, to balloon and end up matching Curiosity's $2.7 billion price tag, which includes operations.

To analyze samples in its onboard lab, Curiosity's drill only needed to pulverize rock. Perseverance, in contrast, must drill intact cores, each about the size of a thick piece of school chalk, and store them within titanium tubes. The system also has to keep the cores safe and clean, to prevent Earth-borne microbes and molecules from being mistaken for martian ones when the cores finally arrive back on Earth. In the end, engineers dreamed up a system involving two robotic arms, nine drill bits, 43 sample tubes, and a rotating carousel. "When you look at it, you won't think of it being simple," says Adam Steltzner, the rover's chief engineer at JPL, "but it was the simplest we could imagine."

Building and testing that system nearly delayed a mission straining to meet tight deadlines. In October 2019, engineers discovered the tubes seized up inside the drill bit when tested in martian conditions. "For me it was a moment of despair," Farley says. "How were we ever going to fix this?"

The problem, it turned out, was that the rover was too clean. The tubes had been baked at 350°C for 1 hour, which not only sterilized them, but also vaporized a hydrocarbon film. The team hadn't realized that the film, a patina that forms on nearly any metal exposed to Earth's atmosphere, was needed as a grease. After several stressful months, they developed a cleaning routine that limited the baking to 150°C and included a series of chemical washes. That left a small amount of the film on the outside of the tubes but no trace inside, where it might contaminate samples. "We leave nothing behind, like a good hiker," Steltzner says.

After the issue was resolved, another mote of organic material began to threaten the mission's launch. By February, the coronavirus pandemic had postponed the launch of ESA's Rosalind Franklin rover, which already had parachute problems, until 2022, the next Mars launch window. Determined to hit its window, NASA shuttled a skeleton crew to and from Florida for the rover's final inspections, while most JPL engineers did what they could from their California homes. "It's fascinating how much of it you can do from your living room," says Jennifer Trosper, the rover's deputy project manager for surface operations. "We're used to remote operation. We just had to move it back a little earlier."

In May, with the rover already stacked on the spacecraft that would ferry it to Mars, a C-130 transport plane delivered the cleaned sample tubes, quarantined in nitrogen-filled cases, to Cape Canaveral. Engineers loaded the tubes just before a heat shield sealed the rover within its landing capsule. A last-minute arrival of the tubes was always the plan to limit contamination risks. Also, Trosper adds, "We just finished them."

Published by AAAS

On 20 July, a 3-week launch window opens up. Seven months after an Atlas V rocket puts it on a path to Mars, the rover will plunge through the barely there martian atmosphere. Just as for Curiosity, a "sky crane" hovering on retrorockets will unspool Perseverance on a tether and lower it to the ground. But there's an important improvement: A camera on the rover's belly will assess the landscape as it descends and compare it to a stored map of safe landing spots. The sky crane will fire its thrusters to divert to one of these zones, enabling the rover to land far closer to its target than Curiosity did in 2012, in a nearly circular, 8-kilometer-wide landing ellipse at the delta's edge.

FROM THAT MOMENT it will be a player in what *Nature Geoscience* dubbed a "war" over Mars's ancient climate. What Curiosity saw at Gale crater convinced some geologists that ancient Mars remained warm for millions of years. Sediments probably built up more slowly on Mars than on Earth, so the thick sediments at Gale suggested "this lake almost certainly existed for tens of millions of years, maybe longer," says John Grotzinger, the Caltech geologist who led Curiosity's science for its first few years.

If so, the lake would have endured climate variations driven by chaotic wobbles in the planet's tilt, which varies from 10° to 60°. Something must have kept the planet warm while the lake shifted between tropical and arctic latitudes. "Did we land in one weirdo place on Mars? Probably not," Grotzinger says. But what warmed the climate is a mystery, he admits. "Something is missing, and we don't know what that is yet."

To the opposing camp, that's grounds for skepticism about a warm early Mars. In 1991, James Kasting, a planetary scientist at Pennsylvania State University, University Park, reported that an atmosphere of carbon dioxide (CO_2) and water vapor, both greenhouse gases, was not capable of keeping the ancient planet wet and warm for millions of years. The atmosphere would have been too thin, and the early Sun too weak. Mars "must have had a phenomenal greenhouse effect," Arvidson says, double what exists now on Earth. To this day, even with more sophisticated models, "The climatologists haven't figured out how to do it," he adds.

That has led these scientists to argue that martian water flowed in bursts lasting just thousands of years—brief exclamations in an eternal deep freeze. That is a Mars that climate models can simulate, says Robin Wordsworth, a planetary scientist at Harvard University. Its ancient volcanoes could have belched a lot of hydrogen, a strong but short-lived greenhouse gas. Periodic bursts of water could have rusted iron-bearing minerals, releasing more hydrogen to the air. Or asteroid strikes, more common in that era, could have released hydrogen if they hit regions rich in ice or subsurface water. "For all of them you can make episodic warming work," Wordsworth says. "But not warm and wet."

The rocks in Gale crater can also support this view, Ehlmann says. They lack certain minerals that should be present if they were exposed to water for 1 million years or more. Jim Bell, a planetary scientist at Arizona State University, Tempe, has concluded that ancient Mars was probably like Antarctica, icy and dry, with spurts of melt. "More Earth-like does not mean like most of the Earth."



Sample tubes, baked and washed clean of microbes, were the last rover parts installed. NASA aims to return about 30 to Earth.

PERSEVERANCE WILL NEED the head start pro-vided by a precise landing to try to settle the issue. During its 2-year primary mission, it will take advantage of upgraded wheels and autonomous navigation capabilities to briskly traverse more than 15 kilometers—a distance Curiosity took more than 4 years to cover. The rover will collect its first 20 samples for an eventual return mission from the geologically diverse terrain it will cross.

The first samples are likely to be rocks thought to come from an eruption that covered parts of the crater after the lake dried up. Volcanic rocks contain trace radioactive elements that decay at a certain rate, a clock that lab scientists on Earth can use to date the eruption, putting a lower limit on the age of the lake. Mission scientists also hope to find outcrops of older volcanic rocks that sit below the delta mudstones, marking an eruption that occurred before the water arrived. Those would provide an upper age limit, making it possible to roughly bracket the lake's existence. "When were these habitable environments in absolute time, and how quickly did they come and go?" Ehlmann asks.

As the rover rolls along the lake bottom, a ground-penetrating radar mounted on its belly will fire, recording echoes that reveal the textures of sediment up to 10 meters below the surface. "We'll be creating a giant ribbon of data," says David Paige, a planetary scientist at the University of California, Los Angeles, and the instrument's deputy principal investigator. The reflections could help determine whether the lake was open water or covered in ice. Fine mud would suggest open water; anomalously large stones would suggest ice, which could have carried them to the middle of the lake before dropping them.

From there the rover will visit the fine-grained clay-bearing mudstones of the lower parts of the delta. Here the hunt for past life will take the lead. On Earth, such clays blanket living things and preserve them as fossils. In similar clays at Gale crater, Curiosity scientists detected traces of complex organic compounds that resembled kerogen, the feedstock of oil. But they could not determine whether the compounds, detected at levels of a few dozen parts per million, were produced by ancient life, or deposited on the martian surface by meteorites, which often contain complex organic molecules.

Two instruments mounted at the end of Perseverance's main robotic arm may help tell the difference. One will fire an ultraviolet laser at the rocks; the other will bombard them with x-rays. The radiation re-emitted by atoms in the rocks could reveal organic chemistry. Mapping any organics in a rock could also say something about their origin. A uniform signal would favor meteoritic fallout, whereas a lumpier distribution, and the presence of minerals that hint at microbe-fueling reactions, could be a sign of life—and a green light to drill a sample.

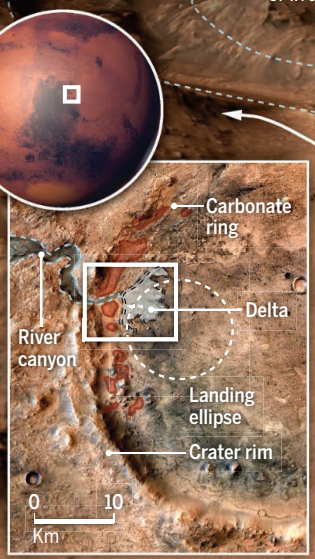
As the rover forges a path up the delta, the fine mudstone will give way to rough sandstone. The team will keep an eye out for exposures of opal-like rocks that have recently been spotted from space. Opal forms from a solution of silica and water, and on Earth the deposits are classic fossil-hunting

CREDITS: (PHOTO) NASA/JPL-CALTECH; (GRAPHIC, OPPOSITE PAGE) C. BICKEL/*SCIENCE*; (MAP) NASA/JPL/MSSS/ESA/DLR/FU-BERLIN/J. COWART, CC-BY-SA 3.0 IGO; (DATA) FERGASON *ET AL.*/PLANETARY DATA SYSTEM EXPERIMENTAL DATA RECORD

Published by AAAS

A trip through time

Next month, NASA's Perseverance rover will head to Mars to explore the remains of a muddy river delta more than 3 billion years old. Scientists don't know whether the water came during a brief thaw on a cold, dry planet or in a lasting period of warmth. The rover's path, crossing more than 500 million years of geologic history, could help resolve the debate. *Graphic by* **Chris Bickel**



River canyon

Delta

Rover's path

0 1 Km

Jezero crater

After landing near—or maybe on—the delta, the rover will begin a 15-kilometer climb to the crater rim, where a lake may have left a ring of carbonate rocks.

1 Volcanic rocks
Dating lava that may have covered parts of the crater could bound the end of the wet episode.

2 Delta mudstone
Mud can smother and fossilize microbes. Samples drilled here will be prized on return to Earth.

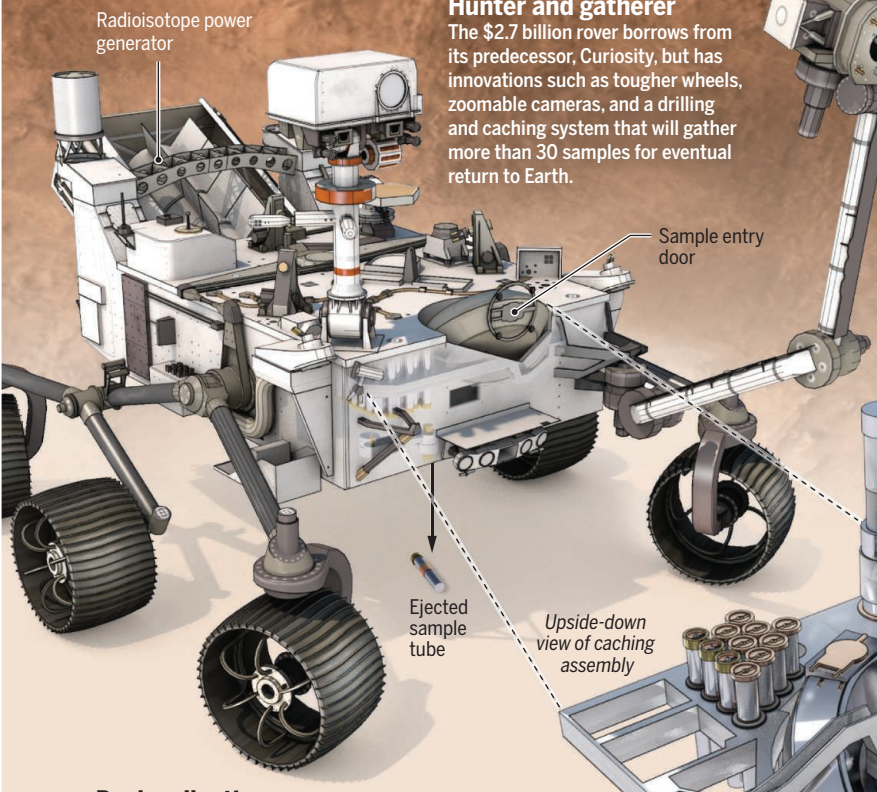
3 Delta sandstone
Higher up the delta, sand grains may hold imprints of a lost magnetic field that may have protected a thick, warming atmosphere.

4 Carbonates
Rocks formed in the lake's shallows could contain biosignatures and clues to an ancient greenhouse effect.

5 Canyon mouth
River deposits may reveal the strength of the ancient water flows, or how often they froze.

Hunter and gatherer

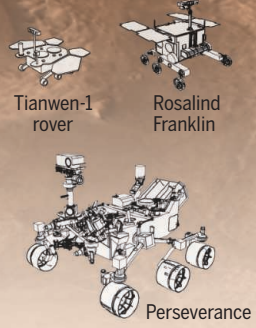
The $2.7 billion rover borrows from its predecessor, Curiosity, but has innovations such as tougher wheels, zoomable cameras, and a drilling and caching system that will gather more than 30 samples for eventual return to Earth.

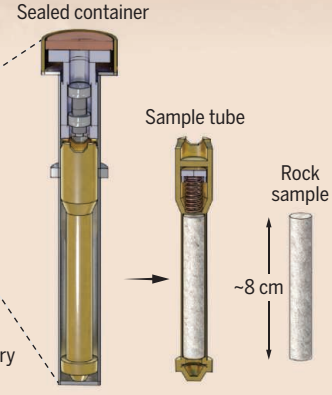


Helicopter
A drone the size of a softball will take photos in short test flights, its rotors spinning 10 times faster than a helicopter on Earth.

Three's company

In July, China will launch Tianwen-1, an orbiter, lander, and rover. Europe has delayed the launch of its Rosalind Franklin rover to 2022.





Rock collection

A drill will feed samples to a carousel and a second arm that helps seal and store them. Some tubes will be dropped and others kept on board; a later rover will retrieve them for a return flight to Earth.

Published by AAAS



Microbes in Turkey's Lake Salda create layered mounds of carbonate called stromatolites. Jezero crater may have hosted a similar environment billions of years ago.

spots. That's because the mineral creeps into organic layers and preserves fossil structures, Bosak says. "That's where we find the most beautifully preserved microbial mats." The rover's cameras will search for such structures, but Bosak doubts they will be seen—even on Earth, they are not often apparent until polished in the lab.

The sand grains, washed in by the long-lost river, could also say something about what caused Mars's early warmth—whether steady or intermittent—to dissipate after the delta formed. Some of the sand grains, eroded from volcanic rocks, will contain radioactive isotopes that make it possible to date them. Scientists on Earth will also examine certain minerals to look for the frozen imprint of a magnetic field. Mars is believed to have had a magnetic field early in its history, generated by a molten dynamo in the planet's interior. The field would have failed as the dynamo cooled and shut down, and some believe that explains Mars's radical change in climate. A weakening field could have allowed charged particles from the Sun to erode the planet's once-thick atmosphere. Water would have escaped to space, making the planet colder and drier. Magnetic signatures teased out of the sand grains could show whether the decline of the field preceded—and perhaps caused—the climate change.

AFTER NEARLY 2 YEARS of frantic drilling, the rover will climb one of the delta's fingers to reach the shores of Jezero's paleolake, fast against the crater's edge. Orbiters have spotted a bathtub ring of carbonate rocks running around the crater rim in a narrow band, likely where the lake's warm shallows were. On Earth, such deposits are known to preserve fossilized stromatolites, bumpy cauliflowerlike mounds formed by the growth of bacteria. "They are an ideal place to look for past life," says Briony Horgan, a planetary scientist at Purdue University. That is, she says, if the deposits were formed by the lake, and not by hot water created by the crater-forming impacts.

If the lake was responsible for the carbonate deposits, they will offer a window on the ancient martian atmosphere, which supplied the CO_2 that formed them. By comparing carbon isotopes from carbonates in the bathtub ring and in older rocks outside the crater, scientists could learn how levels of atmospheric CO_2—and the greenhouse effect it drives—changed over this time.

The outcrops around these shallows, near the entry point of the river, could also betray something about the climate at the time the delta was laid down, says Timothy Goudge, a planetary scientist at the University of Texas, Austin. Layering in the outcrops will reveal how much water was needed to form the delta, how long it flowed, and whether it came in the brief floods or steady flows. Cracks in river bottom rocks could be wedges opened up by continuous freeze-thaw cycles—a sign of persistent frigid conditions.

The shallows will likely mark the end of the primary mission. But the rover ought to have many more years left on its odometer. Engineers want the rover, while still healthy, to drop some samples on flat, accessible terrain where a later mission can retrieve them. But it may drill some sites twice, and it will continue to collect—in case the rover itself is healthy enough to deliver samples to the Earth return mission.

After the primary mission, many on the team will be eager to escape the delta for the ancient, mysterious terrain to the west. Deposits of clay and carbonate seen there also needed water to form. If bands of water-weathered rock capped by water-deposited sediments are visible on its mesas, a temperate climate may have prevailed. Alternatively, as Ehlmann and others believe, this landscape could be what's left of an icy subsurface that was heated by nearby giant impacts 4 billion years ago and turned into an underground hydrothermal system capable of fostering life. "That would point to an ancient Mars that was habitable, but not so warm," she says.

Whatever answers the rover finds, it will mark the end of an era on Mars. For decades, NASA has dominated exploration of the planet's surface, culminating in the increasingly ambitious rovers of the past 2 decades. "We've had the privilege and responsibility to do a systematic investigation of a planet," says Jim Watzin, director of NASA's Mars Exploration Program. Other nations will soon add their rovers, starting with China this year (see sidebar, p. 1420).

But Perseverance also kicks off a new era. The sample return effort it anchors "will be the first round trip to another planet by humanity," Watzin says. Bringing Mars to Earth will enable scientists to probe the secrets of the Red Planet more deeply. If humans follow in the rover's tracks in the coming decades, as the United States and China have vowed, the terrain they encounter may seem strange. But it will be familiar ground. ■

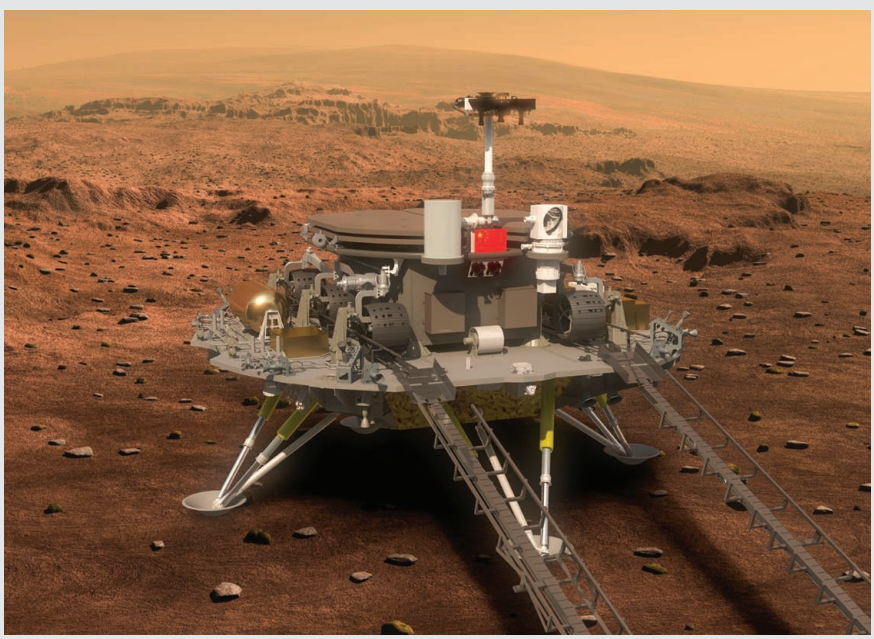
PHOTO: BILAL ALTIOK/ANADOLU AGENCY/GETTY IMAGES

Published by AAAS

Mars mission would put China among space leaders

By **Dennis Normile**

NASA's Perseverance rover may have company on the Red Planet. China aims to leap to the front ranks in planetary exploration with an ambitious Mars mission, its first independent bid to reach the planet. Tianwen-1—"quest for heavenly truth"—consists of not only an orbiter, but also a lander and a rover, a trifecta no other nation has accomplished on its first Mars bid. "A successful landing would put China among elite company," says Mason Peck, an aerospace engineer at Cornell University.



The as-yet-unnamed Tianwen-1 rover is the size of a small golf cart and one-quarter the weight of NASA's Perseverance. The solar-powered rover is expected to operate for at least 90 days.

Due to launch in July, the mission, if successful, would mark dramatic progress for China's space program. In recent years it has fielded several lunar landers but made only one attempt on Mars, an orbiter that piggybacked on a failed 2011 Russian mission to the martian moon Phobos.

A Mars landing is among the most challenging feats in spaceflight. Unlike the Moon, Mars has an atmosphere, which means landers need protection from the heat generated during descent. But its air is too thin for a parachute alone to slow a lander; retrorockets are needed as well. And the entire sequence must be executed autonomously. Of 18 lander or rover missions to Mars, only 10 have been successful. Nine of those 10 were NASA missions. A Russian probe landed successfully, but almost immediately lost communications.

Scientists involved in Tianwen-1 said they did not have permission from the China National Space Administration (CNSA) to speak to the press, and the agency did not respond to questions. Although state media have run stories about the mission, there is nothing like the fanfare that accompanies a NASA Mars landing. Several sources within China's space community believe the agency is muting publicity to temper expectations for a risky mission.

China has not yet announced which of two candidate landing sites it prefers. Both are flat, smooth plains not far from where NASA's Viking 1 and Viking 2 landers touched down in 1976. The low-lying sites give the lander's parachute more time to work. Although scientists might have preferred a more rugged site at higher elevations with more interesting geology, "I speculate [CNSA engineers] are looking to particularly demonstrate a safe landing," says Jim Bell, a planetary scientist at Arizona State University, Tempe, and veteran of several Mars rover missions.

Landing is not the only objective, however. "Our goal is to explore and gather as much scientific data as possible," CNSA chief mission architect Zhang Rongqiao said during a July 2019 lecture on the mission. The orbiter aims to study the martian magnetic field and atmosphere. With a high-resolution camera, it will map the surface and characterize its geology.

The as-yet-unnamed, 240-kilogram rover, the size of a small golf cart and one-quarter the weight of Perseverance, carries six scientific instruments. Among them is a ground-penetrating radar (GPR) that, along with one on Perseverance, will be the first such devices on Mars, able to map subsurface features that orbiting radars see only dimly. "You can really investigate layering, structures, and the presence of permafrost or ice," says Elena Pettinelli, a geophysicist at Roma Tre University, who has helped analyze GPR data from China's Chang'e 3 and 4 missions to the Moon.

Tianwen-1 will take 7 months to reach Mars, and it will be several more months before the orbiter releases the lander, according to a 2017 paper outlining the mission in *Science China Technological Sciences*. After trundling off a ramp on the lander, the solar-powered rover is expected to operate for at least 90 martian days, using the orbiter as a communications relay. The orbiter will keep going for about one martian year, or roughly 23 months.

Dean Cheng, a China policy expert at the Heritage Foundation, a conservative U.S. think tank, says beyond demonstrating technological prowess, China wants to contribute "to the global pool of knowledge." It believes "great powers are also scientific powers," he says.

Tianwen-1 is not the only upcoming demonstration of those ambitions. Later this year, China plans to launch its Chang'e 5 mission, which would return the first Moon rocks since the last Soviet Union Luna mission in 1976; it will likely attempt a far-side sample return mission after that.

CNSA officials have suggested that if Tianwen-1 and Chang'e 5 go well, China could attempt to return samples from Mars beginning around 2030. That timeline puts it on the heels of the NASA–European Space Agency sample return mission—but not by much.

With reporting by Bian Huihui.

PHOTO: XINHUA VIA GETTY IMAGES

Published by AAAS



Correctly fitted masks are an important tool to reduce airborne transmission of SARS-CoV-2, particularly in enclosed spaces, such as on this Moscow Metro train in Russia.

PHOTO: SERGEI FADEICHEV/TASS VIA GETTY IMAGES

VIEWPOINT: COVID-19

Reducing transmission of SARS-CoV-2

Masks and testing are necessary to combat asymptomatic spread in aerosols and droplets

By **Kimberly A. Prather[1], Chia C. Wang[2,3], Robert T. Schooley[4]**

Respiratory infections occur through the transmission of virus-containing droplets (>5 to 10 μm) and aerosols (≤5 μm) exhaled from infected individuals during breathing, speaking, coughing, and sneezing. Traditional respiratory disease control measures are designed to reduce transmission by droplets produced in the sneezes and coughs of infected individuals. However, a large proportion of the spread of coronavirus disease 2019 (COVID-19) appears to be occurring through airborne transmission of aerosols produced by asymptomatic individuals during breathing and speaking (*1*–*3*). Aerosols can accumulate, remain infectious in indoor air for hours, and be easily inhaled deep into the lungs. For society to resume, measures designed to reduce aerosol transmission must be implemented, including universal masking and regular, widespread testing to identify and isolate infected asymptomatic individuals.

Humans produce respiratory droplets ranging from 0.1 to 1000 μm. A competition between droplet size, inertia, gravity, and evaporation determines how far emitted droplets and aerosols will travel in air (*4*, *5*). Larger respiratory droplets will undergo gravitational settling faster than they evaporate, contaminating surfaces and leading to contact transmission. Smaller droplets and aerosols will evaporate faster than they can settle, are buoyant, and thus can be affected by air currents, which can transport them over longer distances. Thus, there are two

Published by AAAS

major respiratory virus transmission pathways: contact (direct or indirect between people and with contaminated surfaces) and airborne inhalation.

In addition to contributing to the extent of dispersal and mode of transmission, respiratory droplet size has been shown to affect the severity of disease. For example, influenza virus is more commonly contained in aerosols with sizes below 1 μm (submicron), which lead to more severe infection (*4*). In the case of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), it is possible that submicron virus-containing aerosols are being transferred deep into the alveolar region of the lungs, where immune responses seem to be temporarily bypassed. SARS-CoV-2 has been shown to replicate three times faster than SARS-CoV-1 and thus can rapidly spread to the pharynx, from which it can be shed before the innate immune response becomes activated and produces symptoms (*6*). By the time symptoms occur, the patient has transmitted the virus without knowing.

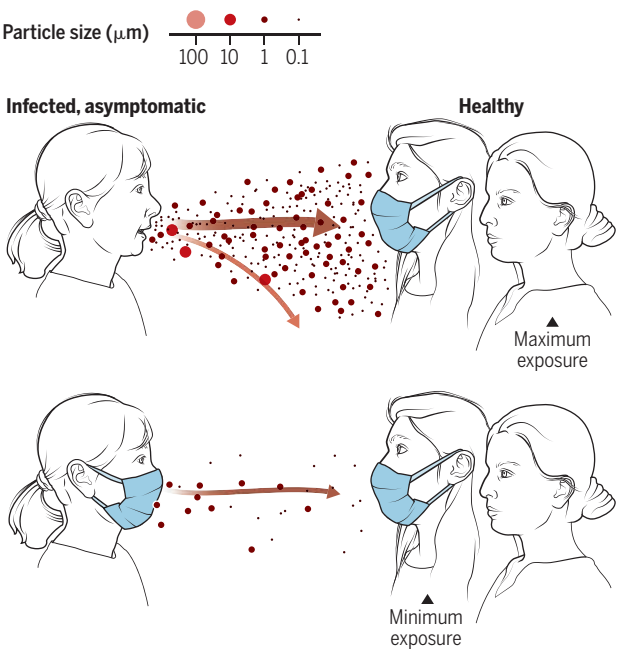
Identifying infected individuals to curb SARS-CoV-2 transmission is more challenging compared to SARS and other respiratory viruses because infected individuals can be highly contagious for several days, peaking on or before symptoms occur (*2, 7*). These "silent shedders" could be critical drivers of the enhanced spread of SARS-CoV-2. In Wuhan, China, it has been estimated that undiagnosed cases of COVID-19 infection, who were presumably asymptomatic, were responsible for up to 79% of viral infections (*3*). Therefore, regular, widespread testing is essential to identify and isolate infected asymptomatic individuals.

Airborne transmission was determined to play a role during the SARS outbreak in 2003 (*1, 4*). However, many countries have not yet acknowledged airborne transmission as a possible pathway for SARS-CoV-2 (*1*). Recent studies have shown that in addition to droplets, SARS-CoV-2 may also be transmitted through aerosols. A study in hospitals in Wuhan, China, found SARS-CoV-2 in aerosols further than 6 feet from patients, with higher concentrations detected in more crowded areas (*8*). Estimates using an average sputum viral load for SARS-CoV-2 indicate that 1 min of loud speaking could generate >1000 virion-containing aerosols (*9*). Assuming viral titers for infected super-emitters (with 100-fold higher viral load than average) yields an increase to more than 100,000 virions in emitted droplets per minute of speaking.

The U.S. Centers for Disease Control and Prevention (CDC) recommendations for social distancing of 6 feet and hand washing to reduce the spread of SARS-CoV-2 are based on studies of respiratory droplets carried out in the 1930s. These studies showed that large, ~100 μm droplets produced in coughs and sneezes quickly underwent gravitational settling (*1*). However, when these studies were conducted, the technology did not exist for detecting submicron aerosols. As a comparison, calculations predict that in still air, a 100-μm droplet will settle to the ground from 8 feet in 4.6 s, whereas a 1-μm aerosol particle will take 12.4 hours (*4*). Measurements now show that intense coughs and sneezes that propel larger droplets more than 20 feet can also create thousands of aerosols that can travel even further (*1*). Increasing evidence for SARS-CoV-2 suggests the 6 feet CDC recommendation is likely not enough under many indoor conditions, where aerosols can remain airborne for hours, accumulate over time, and follow airflows over distances further than 6 feet (*5, 10*).

Masks reduce airborne transmission

Infectious aerosol particles can be released during breathing and speaking by asymptomatic infected individuals. No masking maximizes exposure, whereas universal masking results in the least exposure.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

In outdoor environments, numerous factors will determine the concentrations and distance traveled, and whether respiratory viruses remain infectious in aerosols. Breezes and winds often occur and can transport infectious droplets and aerosols long distances. Asymptomatic individuals who are speaking while exercising can release infectious aerosols that can be picked up by airstreams (*10*). Viral concentrations will be more rapidly diluted outdoors, but few studies have been carried out on outdoor transmission of SARS-CoV-2. Additionally, SARS-CoV-2 can be inactivated by ultraviolet radiation in sunlight, and it is likely sensitive to ambient temperature and relative humidity, as well as the presence of atmospheric aerosols that occur in highly polluted areas. Viruses can attach to other particles such as dust and pollution, which can modify the aerodynamic characteristics and increase dispersion. Moreover, people living in areas with higher concentrations of air pollution have been shown to have higher severity of COVID-19 (*11*). Because respiratory viruses can remain airborne for prolonged periods before being inhaled by a potential host, studies are needed to characterize the factors leading to loss of infectivity over time in a variety of outdoor environments over a range of conditions

Given how little is known about the production and airborne behavior of infectious respiratory droplets, it is difficult to define a safe distance for social distancing. Assuming SARS-CoV-2 virions are contained in submicron aerosols, as is the case for influenza virus, a good comparison is exhaled cigarette smoke, which also contains submicron particles and will likely follow comparable flows and dilution patterns. The distance from a smoker at which one smells cigarette smoke indicates the distance in those surroundings at which one could inhale infectious aerosols. In an enclosed room with asymptomatic individuals, infectious aerosol concentrations can increase over time. Overall, the probability of becoming infected indoors will depend on the total amount of SARS-CoV-2 inhaled. Ultimately, the amount of ventilation, number of people, how long one visits an indoor facility, and activities that affect airflow will all modulate viral transmission pathways and exposure (*10*). For these reasons, it is important to wear properly fitted masks indoors even when 6 feet apart. Airborne transmission could account, in part, for the high secondary transmission rates to medical staff, as

[1]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92037, USA. [2]Department of Chemistry, National Sun Yat-sen University, Kaohsiung, Taiwan 804, Republic of China. [3]Aerosol Science Research Center, National Sun Yat-Sen University, Kaohsiung, Taiwan 804, Republic of China. [4]Department of Medicine, Division of Infectious Diseases and Global Public Health, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA. Email: kprather@ucsd.edu

Published by AAAS

well as major outbreaks in nursing facilities. The minimum dose of SARS-CoV-2 that leads to infection is unknown, but airborne transmission through aerosols has been documented for other respiratory viruses, including measles, SARS, and chickenpox (*4*).

Airborne spread from undiagnosed infections will continuously undermine the effectiveness of even the most vigorous testing, tracing, and social distancing programs. After evidence revealed that airborne transmission by asymptomatic individuals might be a key driver in the global spread of COVID-19, the CDC recommended the use of cloth face coverings. Masks provide a critical barrier, reducing the number of infectious viruses in exhaled breath, especially of asymptomatic people and those with mild symptoms (*12*) (see the figure). Surgical mask material reduces the likelihood and severity of COVID-19 by substantially reducing airborne viral concentrations (*13*). Masks can also protect uninfected individuals from SARS-CoV-2 aerosols and droplets (*13, 14*). Thus, it is particularly important to wear masks in locations with conditions that can accumulate high concentrations of viruses, such as health care settings, airplanes, restaurants, and other crowded places with reduced ventilation. The aerosol filtering efficiency of different materials, thicknesses, and layers used in properly fitted homemade masks was recently found to be similar to that of the medical masks that were tested (*14*). Thus, the option of universal masking is no longer held back by shortages.

From epidemiological data, places that have been most effective in reducing the spread of COVID-19 have implemented universal masking, including Taiwan, Japan, Hong Kong, Singapore, and South Korea. In the battle against COVID-19, Taiwan (population 24 million, first COVID-19 case 21 January 2020) did not implement a lockdown during the pandemic, yet maintained a low incidence of 441 cases and 7 deaths (as of 21 May 2020). By contrast, the state of New York (population ~20 million, first COVID case 1 March 2020), had a higher number of cases (353,000) and deaths (24,000). By quickly activating its epidemic response plan that was established after the SARS outbreak, the Taiwanese government enacted a set of proactive measures that successfully prevented the spread of SARS-CoV-2, including setting up a central epidemic command center in January, using technologies to detect and track infected patients and their close contacts, and perhaps most importantly, requesting people to wear masks in public places. The government also ensured the availability of medical masks by banning mask manufacturers from exporting them, implementing a system to ensure that every citizen could acquire masks at reasonable prices, and increasing the production of masks. In other countries, there have been widespread shortages of masks, resulting in most residents not having access to any form of medical mask (*15*). This striking difference in the availability and widespread adoption of wearing masks likely influenced the low number of COVID-19 cases.

Aerosol transmission of viruses must be acknowledged as a key factor leading to the spread of infectious respiratory diseases. Evidence suggests that SARS-CoV-2 is silently spreading in aerosols exhaled by highly contagious infected individuals with no symptoms. Owing to their smaller size, aerosols may lead to higher severity of COVID-19 because virus-containing aerosols penetrate more deeply into the lungs (*10*). It is essential that control measures be introduced to reduce aerosol transmission. A multidisciplinary approach is needed to address a wide range of factors that lead to the production and airborne transmission of respiratory viruses, including the minimum virus titer required to cause COVID-19; viral load emitted as a function of droplet size before, during, and after infection; viability of the virus indoors and outdoors; mechanisms of transmission; airborne concentrations; and spatial patterns. More studies of the filtering efficiency of different types of masks are also needed. COVID-19 has inspired research that is already leading to a better understanding of the importance of airborne transmission of respiratory disease. ■

REFERENCES AND NOTES

1. L. Morawska, *J. Cao, Environ. Int.* **139**, 105730 (2020).
2. E. L. Anderson *et al.*, *Risk Anal.* **40**, 902 (2020).
3. S. Asadi *et al.*, *Aerosol Sci. Technol.* **54**, 635 (2020).
4. R. Tellier *et al.*, *BMC Infect. Dis.* **19**, 101 (2019).
5. R. Mittal, R. Ni, J.-H. Seo, *J. Fluid Mech.* 10.1017/jfm.2020.330 (2020).
6. H. Chu *et al.*, *Clin. Infect. Dis.* 10.1093/cid/ciaa410 (2020).
7. X. He *et al.*, *Nat. Med.* **26**, 672 (2020).
8. Y. Liu *et al.*, *Nature* 10.1038/s41586-020-2271-3 (2020).
9. V. Stadnytskyi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 202006874 (2020).
10. G. Buonanno, L. Stabile, L. Morawska, *Environ. Int.* **141**, 105794 (2020).
11. E. Conticini, B. Frediani, D. Caro, *Environ. Pollut.* **261**, 114465 (2020).
12. N. H. L. Leung *et al.*, *Nat. Med.* **26**, 676 (2020).
13. J. F.-W. Chan *et al.*, *Clin. Infect. Dis.* 10.1093/cid/ciaa644 (2020).
14. A. Konda *et al.*, *ACS Nano* 10.1021/acsnano.0c03252 (2020).
15. C. C. Leung, T. H. Lam, K. K. Cheng, *Lancet* **395**, 945 (2020).

ACKNOWLEDGMENTS

The authors thank S. Strathdee, D. Petras, and L. Marr for helpful discussions. K.A.P. is supported by the NSF Center for Aerosol Impacts on Chemistry of the Environment (CHE1801971). R.T.S. is supported by the National Institute of Allergy and Infectious Diseases (R01 AI131424). C.C.W. is supported by the Ministry of Science and Technology (MOST 108-2113-M-110-003) and the Higher Education Sprout Project of the Ministry of Education, Taiwan, ROC.

Published online 27 May 2020
10.1126/science.abc6197

GENETICS

Evolution after genome duplication

Genetic interactions in yeast reveal factors governing duplicate gene retention and divergence

By **Ian M. Ehrenreich**

Genome duplication generates an extra copy of nearly all genes carried by an organism, providing a potential substrate for evolution. Although many duplicate genes will be eliminated after a genome duplication, those that are retained may evolve distinct functions over time. This process can be studied by characterizing the shared and divergent functions of duplicate genes in present-day organisms whose ancestors experienced genome duplication in the past. However, such work requires examining the functional relationships between each copy of a duplicated gene and the other genes in the genome. This is inherently difficult for duplicate genes because of their redundancy. However, on page 1445 of this issue, Kuzmin *et al.* (*1*) show that systematic analysis of di- and trigenic genetic interactions in budding yeast can overcome this challenge. With this approach, they discover general constraints that influence the retention and divergence of duplicate genes.

> **"...duplicate gene retention and divergence heavily depended on whether the different functions of these genes were entangled (i.e., unable to evolve independently)."**

In 1970, it was proposed that gene duplication provides the critical fuel for evolution (*2*). Owing to their functional redundancy, it was hypothesized that duplicate genes are able to evolve in ways that single-copy genes cannot. Over time, different

Published by AAAS

copies of a duplicate gene might diverge to have distinct or even new functions. These ideas preceded genome sequencing by decades and were difficult to study empirically when first proposed. However, duplicate genes were subsequently found to be prevalent in the genomes of many organisms, supporting their potential importance to evolution (*3*, *4*).

Gene duplication arises through a variety of mechanisms and can occur on scales ranging from individual genes to entire genomes (*2*). Evidence of genome duplication exists in a number of eukaryotic lineages (*5*). These genome duplications likely enhanced organisms' robustness to genetic and environmental perturbations, although they might have also facilitated evolutionary innovations (*5*). The budding yeast *Saccharomyces cerevisiae*—the same organism used to make beer, bread, and wine—is an excellent model system for studying how duplicate genes evolve after genome duplication. The ancestor of this organism experienced a genome duplication, and hundreds of duplicate genes produced by this event have been retained (*6*). These duplicate genes provide a foundation for research on duplicate gene retention and divergence.

Numerous approaches have been used to study these duplicate genes in budding yeast, including comparative genomics (*7*), transcriptomics (*8*), and single gene deletions (*9*). This work suggests that retained duplicate genes are enriched for essential cellular components (*7*). They also show much greater transcriptional divergence than protein sequence divergence (*8*). Potentially consistent with this last point, individual deletion of many duplicate genes has little phenotypic effect, which implies that both copies maintain overlapping function (*9*).

This past work was unable to directly probe the functional relationships between duplicate genes and the other genes in the genome. Such information is the most useful for understanding how duplicate genes have functionally diverged, but is also more difficult to obtain (*10*). Analysis of genetic interactions is one potential high-throughput strategy for mapping these functional relationships (*11*). A genetic interaction occurs when a combination of mutations, such as deletions, exhibits unexpected phenotypic consequences (*12*). Although most studies focus on digenic interactions, analysis of genetic interactions among three or more genes can also provide important biological insights (*13*, *14*).

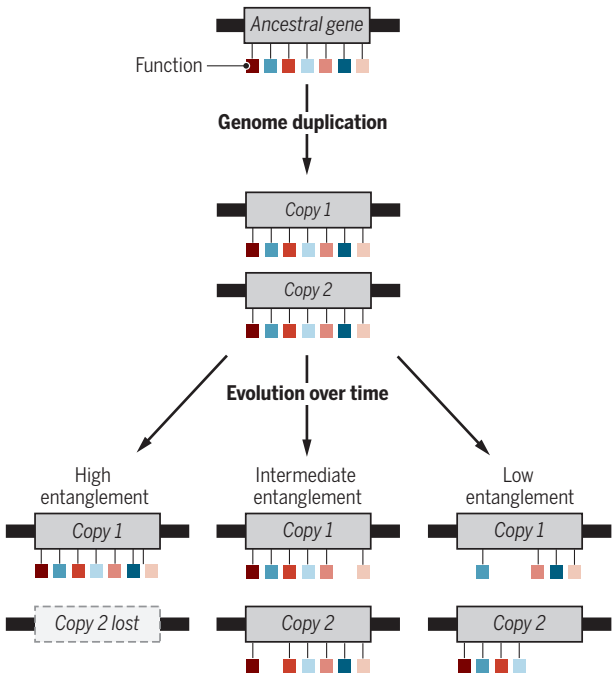
Kuzmin *et al*. used trigenic genetic interaction analysis to measure the shared and distinct functional relationships of duplicate genes. They examined 240 pairs of nonessential duplicate genes in budding yeast. To do this, they combined single and double gene deletions for each duplicate gene pair with a library of ~1200 single gene deletions spanning all major cellular processes. This produced 537,911 double-mutant and 256,861 triple-mutant yeast strains, resulting in the detection (through phenotypic changes) of 7197 digenic and 4557 trigenic genetic interactions. On the basis of these genetic interactions, Kuzmin *et al*. distinguished two main classes of duplicate genes: those that are more functionally divergent, and those that have retained a high degree of functional overlap.

In silico modeling informed by the genetic interaction data revealed that duplicate gene retention and divergence heavily depended on whether the different functions of these genes were entangled (i.e., unable to evolve independently). Such functional entanglements influenced how the copies of a duplicate gene accumulated degenerative mutations over time. High entanglement tended to result in one duplicate gene copy amassing degenerative mutations and being lost (see the figure). When entanglement was lower, duplicate genes were more likely to be retained, with their functional divergence inversely related to their entanglement. This suggests that the two classes of duplicate genes that were found in budding yeast possess different levels of entanglement.

Increasingly, evidence shows that evolution after genome duplication is not entirely random (*1*, *10*, *15*). For many duplicate genes, the potential for retention and divergence is constrained by functional entanglements (*1*). This finding provides further insight about how genome duplication contributes to evolution. For example, it bolsters the idea that many duplicate genes, as a result of their limited ability to evolve, may contribute more to robustness than to innovation (*3*–*5*). Yet it also indicates that particular duplicate genes have greater potential to facilitate innovation because they may be less affected by entanglements. These points illustrate how determining the shared and divergent functions of duplicate genes in organisms such as budding yeast can improve the understanding of the evolutionary process. ■

Duplicate gene evolution

After genome duplication, how duplicate genes evolve depends on functional entanglements. When entanglement is high, one copy is likely to be lost. By contrast, when entanglement is low, the chance that both copies will be retained and able to diverge is higher. Duplicate genes with intermediate levels of entanglement may be retained but are less able to diverge.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

Molecular and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA. Email: ian.ehrenreich@usc.edu

REFERENCES AND NOTES

1. E. Kuzmin *et al.*, *Science* **368**, aaz5667 (2020).
2. S. Ohno, *Evolution by Gene Duplication* (Springer, 1970).
3. A. Force *et al.*, *Genetics* **151**, 1531 (1999).
4. M. Lynch, J. S. Conery, *Science* **290**, 1151 (2000).
5. K. D. Crow, G. P. Wagner, SMBE Tri-National Young Investigators, *Mol. Biol. Evol.* **23**, 887 (2006).
6. K. H. Wolfe, D. C. Shields, *Nature* **387**, 708 (1997).
7. I. Wapinski, A. Pfeffer, N. Friedman, A. Regev, *Nature* **449**, 54 (2007).
8. X. Gu, Z. Zhang, W. Huang, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 707 (2005).
9. Z. Gu *et al.*, *Nature* **421**, 63 (2003).
10. G. Diss *et al.*, *Science* **355**, 630 (2017).
11. B. VanderSluis *et al.*, *Mol. Syst. Biol.* **6**, 429 (2010).
12. M. Costanzo *et al.*, *Science* **353**, aaf1420 (2016).
13. E. Kuzmin *et al.*, *Science* **360**, eaao1729 (2018).
14. M. B. Taylor, I. M. Ehrenreich, *Trends Genet.* **31**, 34 (2015).
15. A. Marchant *et al.*, *eLife* **8**, e46754 (2019).

ACKNOWLEDGMENTS

Thanks to the Ehrenreich lab members for feedback. The author is supported by NIH grant R35GM130381.

10.1126/science.abc1796

Published by AAAS

SPECTROSCOPY

Action spectra of chiral secondary structure

Mass spectra of DNA complexes photoreduced with polarized light reveal their handedness

By **Perdita Barran**

Louis Pasteur has particular resonance in the midst of the coronavirus disease 2019 (COVID-19) pandemic, given his work on the germ theory of disease that led to the development of vaccines for rabies and anthrax. Some years before these better-known discoveries, Pasteur had already made a great contribution to science with his observation of molecular chirality in crystals (*1*). Chirality is a fundamental property of molecular asymmetry, which in turn dictates the secondary structure of proteins and gives rise to handedness in DNA helices. Some 170 years later, on page 1465 of this issue, Daly *et al.* (*2*) present a new method, mass-resolved electronic circular dichroism (CD) ion spectroscopy, that allows the chirality of secondary structures—in this case, isolated guanine-rich DNA quadruplexes—to be unambiguously defined.

In the 1830s, the acceptance of the concept of chemical structure was some decades away, but developments in optics were revealing regularity in the ordering of crystalline materials at a submicrometer level. Light could be filtered through polarizing optics, such as Nicol prisms. Although not understood at the time, this process causes the oscillating electric field of the light to move in only one direction. Using a polarized light source, Jean-Baptiste Biot (*3*) observed that a saturated solution of tartaric acid would rotate the polarized light clockwise or anticlockwise, but he presented no explanation for this phenomenon.

Some 20 years later, in 1848, Pasteur crystallized paratartaric acid, produced from overboiled wine, and noticed that it produced two types of crystals, one that looked like tartaric acid crystals and one that did not. He repeated Biot's experiment with a 50-50 mixture of these crystals and found that this solution would not rotate the polarized light. From these observations, Pasteur concluded that each crystal had an intrinsic property, and that the material in each crystal was the mirror image of the other, like a left and a right hand. Pasteur named this phenomenon chirality (from *kheir*, the Greek word for hand). It soon became evident that this is a property of fundamental importance to chemistry, where different substituents on tetravalent carbon create a chiral center.

Naturally occurring biopolymers such as proteins and DNA are made of monomeric units with such chiral centers. When these repeat at regular intervals, they favor repeating noncovalent interactions that give rise to distinctive secondary structural motifs with characteristic optical signatures. Circular dichroism, the modern version of the experiments of Biot and Pasteur, measures the absorbance of polarized light and can verify, for example, that secondary structures of proteins are correctly folded. However, CD measures the aggregate effect of a mixture and can be insensitive to the presence of an enantiomer at lower concentrations, or to the particular form of the molecule that is causing the optical rotation.

To address this limitation, Daly *et al.* built an experimental apparatus that can make precise measurements of chirality in biomolecules (see the figure). Unlike an optical absorption experiment in solution, the optical source (laser excitation) releases characteristic photoelectrons that lead to charge-reduced ions, so this method is an action spectroscopy. Electrospray ionization introduces guanine-rich DNA quadruplex molecules of different chirality into the vacuum chamber of a mass spectrometer. This transfer is performed carefully under conditions that aim to leave the solvated structure unperturbed. A few thousand molecular ions are trapped and irradiated with polarized laser light.

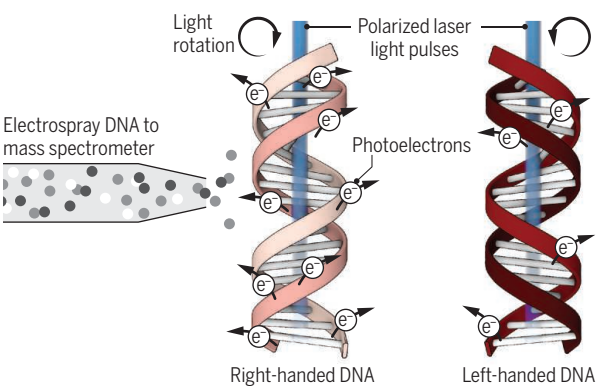
The masses of the left- and right-handed forms of these DNA complexes are identical, so mass spectrometry alone cannot determine chirality. Further, unlike solution experiments where the concentration of molecules is high enough

An exciting way to measure handedness

Molecular handedness (chirality) is usually measured with circular dichroism, in which the direction of polarized light is rotated by absorption of light by molecules in solution. Daly *et al.* determined such spectra for individual molecules in the gas phase.

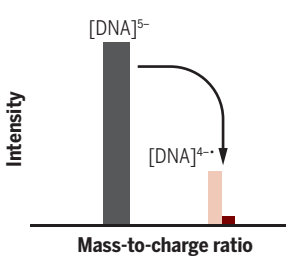
Gas-phase photoexcitation

Charged DNA molecules are introduced to the mass spectrometer with an electrospray source and irradiated with circularly polarized laser light, which excites photoelectrons with an efficiency that is dependent on the rotation.



Taking action

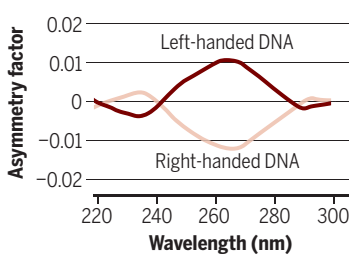
Photoexcitation reduces the negative charge state of the parent ion. The product ion has the same mass and thus a higher mass-to-charge ratio, and reduced intensity as not all parent ions photoexcite.



● Parent ion
◐ Charge-reduced product

Circular dichroism spectra

A circular dichroism spectrum is recorded by monitoring the relative intensity of charge-reduced species (for right- versus left-handed excitation) at different laser wavelengths.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

Department of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester, UK.
Email: perdita.barran@manchester.ac.uk

Published by AAAS

that the absorbance of light can be readily measured, the gas-phase DNA ions are dilute. Instead, Daly *et al.* use the mass spectrometer to measure the action of the polarized light on the DNA molecule. Facile release of a photoelectron upon irradiation gives a signal caused by the formation of the charge-reduced anion. The laser light pulses are polarized, so the efficiency of photoelectron removal depends on the helix handedness. By monitoring the intensity of the distinctive charge-reduced species as a function of the polarization, they are able to determine the chirality of the molecule. This type of experiment can isolate any given molecular complex because each analyte is unequivocally defined by its mass-to-charge ratio.

Prior work by this group revealed coexisting structures formed by oligonucleotides in the presence of cations or organic molecules (*4*); this work was aimed at understanding the balance of forces that guide folding and self-assembly. Stoichiometry and quantitative measurements of nucleic acid complexes were readily made by combining mass and ion-mobility spectrometries, but this approach could not determine chirality or details of secondary structure. Mass-resolved electronic CD ion spectroscopy complements other structural mass spectrometry methods because it can provide secondary structure information in addition to molecular identity.

Daly *et al.* also studied complexes of DNA with ammonium and potassium counterions and obtained a tantalizing glimpse of the effects of individual molecule solvation. The wavelength dependence of the action of polarized light followed the same trend as data obtained in solution, albeit with differences in magnitude, indicating that structures are preserved and that these gas-phase results are relevant to molecules in solution. They extended their measurement on human telomeric DNA sequences to determine enantiomeric ratios of mixtures of G-quadruplex topologies. Intriguingly, the difference in electron-detachment efficiency for the left-handed molecule and the right-handed molecule under a given circularly polarized light is equivalent to the slight enantiomeric excess in the products—up to 1% at some wavelengths—thereby demonstrating the high sensitivity as well as the potential of this new method. ■

REFERENCES AND NOTES

1. L. Pasteur, *C. R. Séances Acad. Sci.* **26**, 535 (1848).
2. S. Daly, F. Rosu, V. Gabelica, *Science* **368**, 1465 (2020).
3. J.-B. Biot, *Mem. Acad. Sci. Inst. Fr.* **15**, 93 (1836).
4. M. Porrini *et al.*, *ACS Cent. Sci.* **3**, 454 (2017).

10.1126/science.abc1294

PHOTOSYNTHESIS

The simplicity of robust light harvesting

A universal design principle underlies photosynthetic antenna systems

By **Christopher D. P. Duffy**

Photosynthetic light harvesting can achieve a quantum efficiency that approaches 100% (that is, the conversion of 100 photons of light into 100 chemically available electrons), and yet it displays notable robustness in the face of ever-changing external light conditions. Although light harvesting varies in structure and composition across the range of photosynthetic life, there is an ongoing effort to uncover a set of common "design" principles for these systems. On page 1490 of this issue, Arp *et al.* (*1*) have revealed the first hints of a simple, seemingly universal set of rules that define the robustness of natural light harvesters. These rules should inform the design of future solar technology.

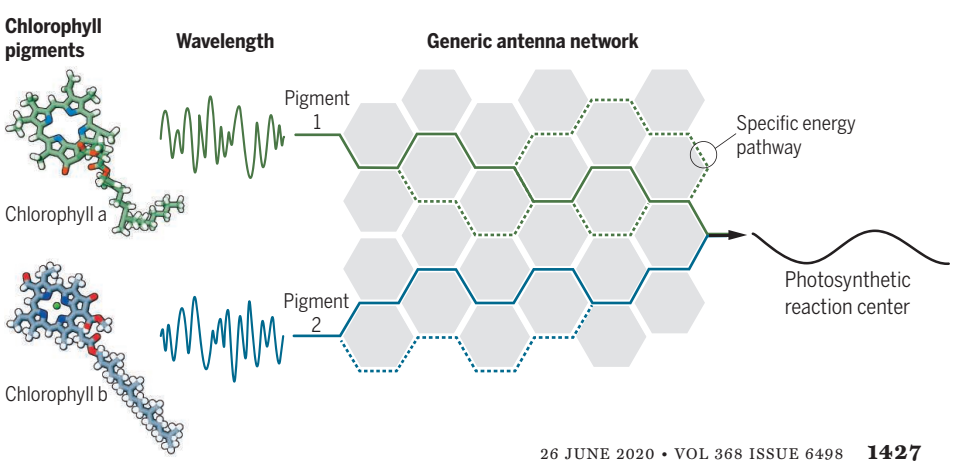
The main challenge facing early photosynthetic life was the paucity of light that was characteristic of marine environments. This led to the evolution of the light-harvesting "antenna"—a large modular assembly of various pigment-binding proteins that absorbs and delivers energy to the photosynthetic reaction centers (*2*). Despite major structural and compositional differences, it has become apparent that antenna systems share several common features (*3, 4*). Each antenna harbors more than one type of pigment. For example, light-harvesting complex II (LHCII), the major antenna protein of plants, binds chlorophyll a and b and several carotenoids. This allows broader coverage of the solar spectrum, which is further enhanced by the fine-tuning of each pigment by its local protein environment. A mixture of pigments are bound at a high density but with specific relative distances and orientations, resulting in energy transfer that is fast and directional. The distribution of different pigments and their specific interactions with the protein ensures irreversibility, with energy flowing downhill to the photosynthetic reaction center rather than lingering in the antenna. It has been argued that these systems owe their efficiency to nontrivial quantum effects (*5*), although this has been recently challenged (*6*). Regardless, photosynthetic light harvesting is considered to be a biological mechanism that is finely tuned for efficiency.

Efficiency, however, can be detrimental under strong light conditions. A mismatch between an antenna's ability to deliver energy and the finite maximum working rate of reaction centers can result in oxidative damage called photoinhibition (*7*). Although effective repair mechanisms exist, they are slow and metabolically costly. Therefore, a suite of reg-

School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK. Email: c.duffy@qmul.ac.uk

Photosynthetic antenna that handles the noise

An antenna binding at least two different types of pigments, whose absolute and relative absorption wavelengths are finely tuned, can convert a fluctuating (i.e., noisy) input into a consistent (i.e., quiet) output.



GRAPHIC: N. CARY/*SCIENCE*

Published by AAAS

ulatory and protective processes is in place to minimize damage. The fastest regulators activate within minutes and relax on a similar time scale when light returns to normal (*8*), but these arguably still lag behind the fastest fluctuations in light intensity (*9*).

As well, light harvesting is not finely tuned in a molecular sense. Instead, the picture is one of robustness in the face of external conditions (*10*) and even genetic manipulation. An extreme example is found in the model plant *Arabidopsis thaliana*, from which LHCII is eliminated through genetic manipulation. In response, an ersatz antenna is formed from assemblies of the remaining minor light-harvesting proteins (*11*). Although these mutant plants may not thrive like their wild-type counterparts, they are certainly photosynthetically competent. If there is some universal organizational principle behind light harvesting, it does not appear to lie in the molecular detail.

"…the evolutionary driving force behind the development of photosynthetic antennae is not maximization of efficiency but the cancellation of noise."

Arp *et al.* have taken a new approach to investigating the question of universal organization. By applying network theory, they have determined the most basic organizational requirements necessary for a light-harvesting system to function optimally. The authors used a very generalized model of the antenna as a network of interconnected sites that can represent individual pigment states within delocalized multipigment states across the antenna. Energy enters this network from one or more input channels that reflect light absorption through different types of pigments. Two input channels were chosen as a minimal model (see the figure), which is reasonably representative of several real antenna systems, such as chlorophyll a and b in LHCII or bacteriochlorophyll c and e in the antenna of green sulfur bacteria (*12*). Absorbed energy can take many different paths through the antenna. However, all paths terminate at a single output channel representing photoinduced charge separation. Such a system is subject to a high degree of input noise. External noise comes from rapid fluctuations in the incident radiation, whereas internal noise originates from the structural dynamics of the antenna. This results in a noisy output that fluctuates between underpowered, optimal, and overpowered states. The underpowered state is metabolically insufficient, whereas the overpowered state risks photoinhibition. The question then is whether the inputs are arranged in such a way that the output spends most of its time at the optimal state, thereby minimizing output noise.

The input channels represent different groups of light-absorbing pigments. As such, they are each defined by the wavelength and rate at which they absorb. If their intrinsic rates (determined by chemical properties and stoichiometry) are assumed to be identical, then the overall rate of absorption is determined solely by the intensity of available light at that wavelength. Consider, for example, two limiting regimes. If the two channels are identical (i.e., a single channel), then antenna homogeneity minimizes internal noise. However, having a single absorbing species makes the system very sensitive to external noise. Conversely, if the two channels differ strongly in both wavelength and absorption rate, then internal noise dominates. Optimum output is generated in an intermediate regime where the two inputs have similar wavelengths but different absorption rates. This is achieved by locating the pair in a region of the spectrum of available light that has the steepest gradient. With only this consideration in mind, Arp *et al.* predicted, with a high degree of accuracy, the absorption profile of green plants, purple bacteria, and green sulfur bacteria.

The finding of Arp *et al.* is important because it suggests that the evolutionary driving force behind the development of photosynthetic antennae is not maximization of efficiency but the cancellation of noise. Moreover, the finding indicates that to build such a system, one must start from the simple requirements of two similar absorbing species that are tuned to the steepest region (not the strongest) of the spectrum of available light. Fine structural details are important, but they come as refinements to this simple underlying principle. ■

REFERENCES AND NOTES

1. T. B. Arp *et al.*, *Science* **368**, 1490 (2020).
2. G. D. Scholes *et al.*, *Nat. Chem.* **3**, 763 (2011).
3. R. Croce, H. van Amerongen, *Nat. Chem. Biol.* **10**, 492 (2014).
4. T. P. J. Krüger, R. van Grondelle, *Physica B* **480**, 7 (2016).
5. G. R. Engel *et al.*, *Nature* **446**, 782 (2007).
6. D. M. Wilkins, N. S. Dattani, *J. Chem. Theory Comput.* **11**, 3411 (2015).
7. B. Kok, *Biochim. Biophys. Acta* **21**, 234 (1956).
8. A. V. Ruban, M. P. Johnson, C. D. P. Duffy, *Biochim. Biophys. Acta Bioenerg.* **1817**, 167 (2012).
9. J. Kromdijk *et al.*, *Science* **354**, 857 (2016).
10. P. Malý, A. T. Gardiner, R. J. Cogdell, R. van Grondelle, T. Mančal, *Phys. Chem. Chem. Phys.* **20**, 4360 (2018).
11. A. V. Ruban *et al.*, *J. Biol. Chem.* **281**, 14981 (2006).
12. C. M. Borrego *et al.*, *Photosynth. Res.* **60**, 257 (1999).

10.1126/science.abc8063

NEUROBIOLOGY

Guide cells help navigate axon regeneration

Rebuilding the flatworm visual system after injury requires guidepost-like muscle cells

By **Rachel Roberts-Galbraith**

During embryonic development, neurons project nascent axons that navigate through space to find their targets. Axonal-growth paths forged by these pioneer axons depend on chemical and physical cues from cells that can include guidepost cells, which directly interact with nascent axons and induce them to grow, stop, or turn (*1*). A series of guidepost cells can serve as "stepping stones" to organize the complex growth of an axon in space (*2*, *3*). Although guidepost cells direct axon formation during embryonic development in diverse model organisms, their existence often is transient. Thus, it was unclear whether guidepost-like cells promote axon regeneration and whether regenerative guidepost cells are constitutive or induced by injury. On page 1447 of this issue, Scimone *et al.* (*4*) pinpoint regenerative guidepost-like cells in the visual system of freshwater flatworms called planarians.

Planarians are well known for their ability to regenerate diverse cell types in predictable patterns even in the face of severe injuries or amputations (*5*). One of the most recognizable features of the planarian body is a pair of crossed eyespots, which consist of pigment cells that form optic cups and photoreceptor neurons that nestle their photosensitive elements within the pigmented cups (*6*). From the eyespots, planarian photoreceptor neurons project axons through space to connect to the cephalic ganglia (a bilobed, horseshoe-shaped brain). The eyespots are positioned dorsally and the brain ventrally in the planarian head. Thus, photoreceptor axons must project ventrally and also turn to make either contralateral or ipsilateral

Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA. Email: robertsgalbraith@uga.edu

Published by AAAS

contacts (*7*). This highly predictable structure made planarian photoreceptor neurons ideal for the identification of cellular guideposts that direct axon organization during regeneration.

Scimone *et al.* thus began their search for regenerative guidepost-like cells that shape the planarian visual system (see the figure). By investigating *notum* gene–expressing cells (*notum*$^+$) that are associated with photoreceptor neurons, the authors discovered three distinct cell types specifically positioned at axonal decision points (*4*). The authors first identified clusters of muscle cells near the eyespots, called NMEs (*notum*$^+$ muscle cells near the eye), and proposed that these cells promote the bundling of axons as they exit the posterior of the eyespot. A second group of muscle cells called NMCs (*notum*$^+$ muscle cells at the choice point) were found near decision points for axons as they diverged to project toward ipsilateral or contralateral contact sites. The third cluster of guidepost-like cells were neurons called NBCs (*notum*$^+$ brain cells), which localized medially and worked either alone or with other medial neurons to promote midline crossing of photoreceptor axons at the optic chiasm.

A combination of physical and molecular manipulations revealed that NMEs, NMCs, and NBCs serve as guidepost-like cells during regeneration to assist photoreceptor axons toward their final arrangement. This demonstrates that highly regenerative organisms like planarians can use constitutive guidepost-like cells to promote a lifelong capacity for neuronal replacement and axon guidance. The new research also addresses long-standing questions about the ubiquity of guidepost cells across the animal kingdom.

One surprise in the new study is the identity of the guidepost-like cells. Whereas guideposts are often neurons, glia, or epithelial cells (*1*), Scimone *et al.* unexpectedly determined that two of the three guidepost-like populations in the planarian visual system express markers of muscle identity. Although previous studies revealed a global role for muscles in body axis patterning (*8*), the current work illustrates that planarian muscle cells also act locally to shape individual organ systems and even to direct the physical arrangement of single cells.

The third population of planarian guidepost-like cells, NBCs, were neurons. Although less surprising, this finding fits into an emerging theme of planarian neurons influencing brain regeneration in a multitude of ways, from patterning to fate choice. *notum*$^+$ neurons at the anterior-most end of the planarian brain influence brain scaling (*9*), and a population of medial neurons produces Hedgehog ligand to promote neurogenesis (*10*). Scimone *et al.* also built on the prior identification of *arrowhead* as a key transcription factor–encoding gene that is expressed in medial neurons (including NBCs) and influences medial patterning of photoreceptors and the anterior commissure, the largest connection between halves of the planarian brain (*11*).

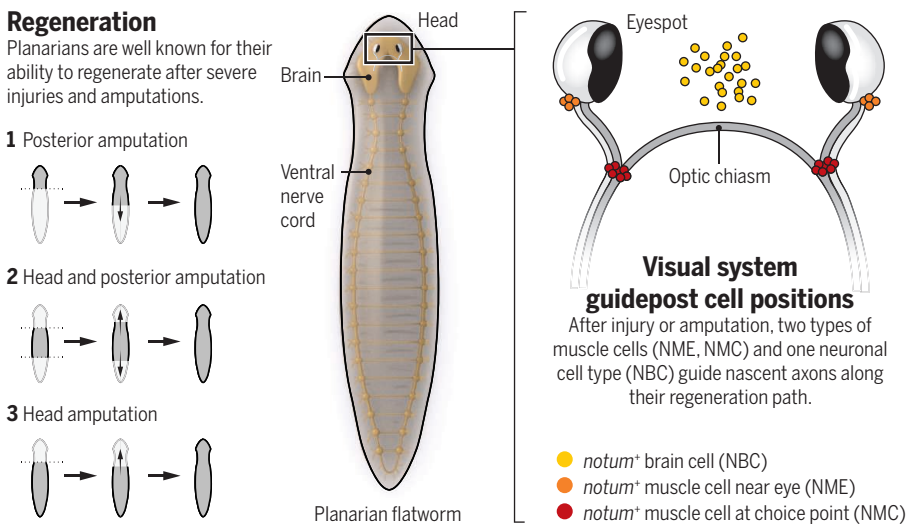
It remains unknown whether guidepost-like cells organize the positioning of other planarian neuronal structures during regeneration (for example, axons that run along the ventral nerve cords, axons that project into the planarian pharynx, or neural processes that bundle together to form sensory structures) (*6*). Planarian guidepost-like cells might also regulate other aspects of neuronal cell behavior during regeneration, including cell migration or synapse formation.

After amputation, planarian guidepost-like cells must themselves be regenerated and positioned properly. Scimone *et al.* determined that these cells regenerate independently of photoreceptor neurons; they discovered factors (intrinsic and extrinsic) that regulate regeneration of guidepost-like cells and ensure their proper arrangement in space. However, the precise mechanisms that direct regeneration of each cell type remain unknown.

The new work raises the intriguing possibility that human neural regeneration might be improved by mimicking the guideposts that direct vertebrate axon formation during development. Human neurons in the brain and spinal cord usually fail to regenerate axons after injury. However, introducing cellular or molecular "stepping stones" along a desired axonal path might coax axons toward better regrowth. Current and future research investigating this theme—in fields ranging from bioengineering to basic developmental biology—might reveal the extent to which cues or matrices meant to mimic guidepost-like cells can improve axon guidance for various populations of adult human neurons (*12*). ■

Guidepost-like cells govern planarian regeneration

Three specialized cell types direct the growth and positioning of nascent axons during regeneration of the visual system.



GRAPHIC: C. BICKEL/*SCIENCE*

REFERENCES AND NOTES

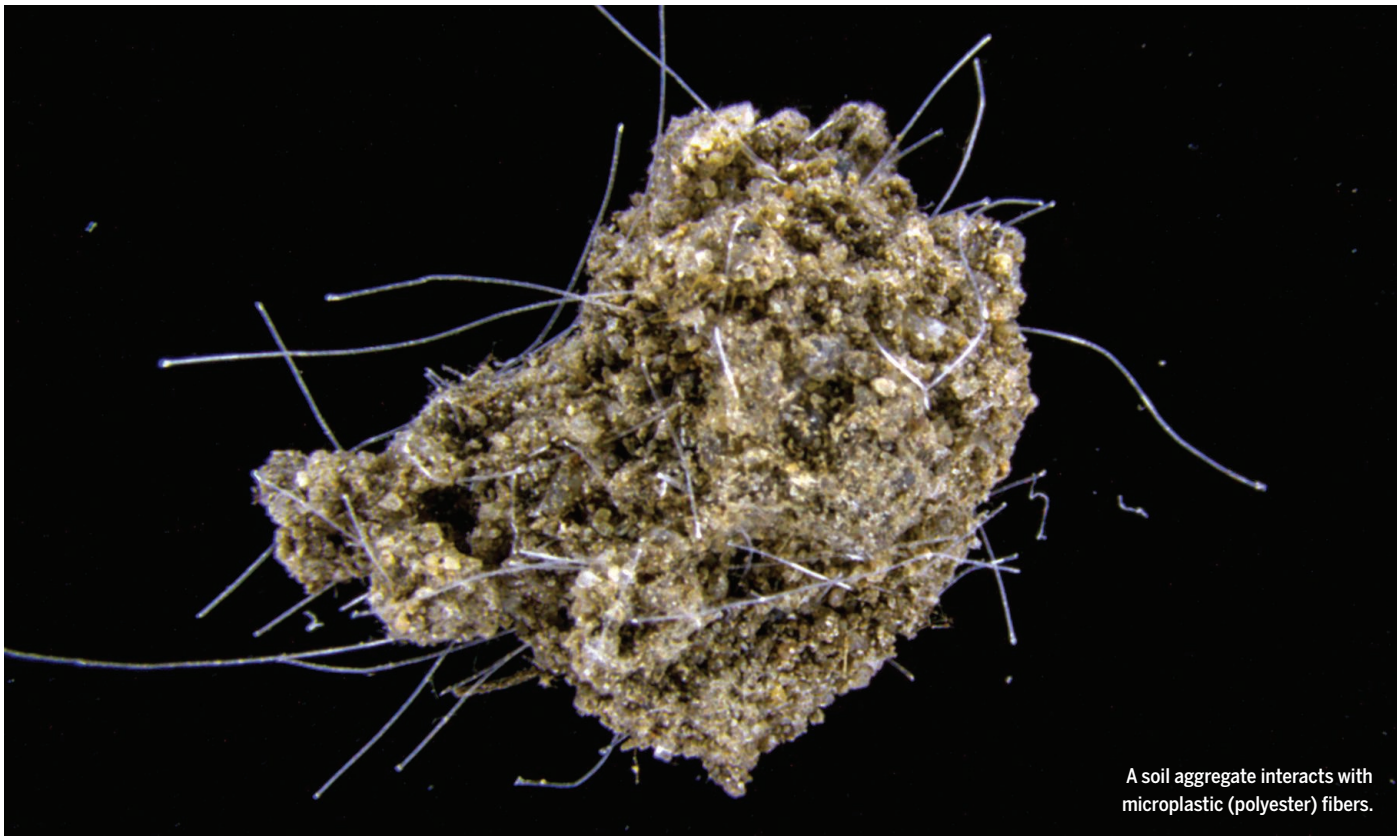
1. D. L. Chao, L. Ma, K. Shen, *Nat. Rev. Neurosci.* **10**, 262 (2009).
2. C. M. Bate, *Nature* **260**, 54 (1976).
3. J. Palka, K. E. Whitlock, M. A. Murray, *Curr. Opin. Neurobiol.* **2**, 48 (1992).
4. M. L. Scimone *et al.*, *Science* **368**, eaba3203 (2020).
5. M. Ivankovic *et al.*, *Development* **146**, dev167684 (2019).
6. K. G. Ross, K. W. Currie, B. J. Pearson, R. M. Zayas, *Wiley Interdiscip. Rev. Dev. Biol.* **6**, e266 (2017).
7. K. Okamoto, K. Takeuchi, K. Agata, *Zool. Sci.* **22**, 535 (2005).
8. J. N. Witchley, M. Mayer, D. E. Wagner, J. H. Owen, P. W. Reddien, *Cell Rep.* **4**, 633 (2013).
9. E. M. Hill, C. P. Petersen, *Development* **142**, 4217 (2015).
10. K. W. Currie, A. M. Molinaro, B. J. Pearson, *eLife* **5**, e19735 (2016).
11. R. H. Roberts-Galbraith, J. L. Brubacher, P. A. Newmark, *eLife* **5**, e17002 (2016).
12. B. J. Hilton, F. Bradke, *Development* **144**, 3417 (2017).

ACKNOWLEDGMENTS

R.R.-G. is supported by funding from the Alfred P. Sloan and McKnight Foundations.

10.1126/science.abc8066

Published by AAAS



A soil aggregate interacts with microplastic (polyester) fibers.

ECOLOGY

Microplastic in terrestrial ecosystems

Research shifts from ecotoxicology to ecosystem effects and Earth system feedbacks

By **Matthias C. Rillig** *and* **Anika Lehmann**

Concern about microplastics (plastic particles <5 mm) polluting different environmental compartments is mounting. Research has recently begun to embrace terrestrial systems, having initially focused at least a decade earlier on marine and aquatic ecosystems (*1*–*3*). The early research agenda on microplastics in both aquatic and terrestrial systems was mainly ecotoxicological. It included laboratory tests on individual organisms, often well-established test species (*4*), and also targeted selected soil properties and processes. Such research is necessary to establish baseline mechanisms, which is important because microplastics differ from other pollutants. Many of their effects appear to be mediated by physical parameters, such as particle shape and size, rather than overt chemically mediated toxicity. Moreover, their effects are mostly sublethal or even nominally positive. Although the study of other global change factors has tended to focus at the level of the ecosystem, research on microplastic is only now on the verge of this wider view.

The first step in this direction has been the conceptualization of the "plastic cycle" (*5*, *6*), acknowledging that microplastics can move between different large-scale compartments, including the air, terrestrial habitats, rivers and other freshwater bodies, and the ocean, including its sediments (*7*) (see the figure). The cycle framework places these movements in the context of the pools and fluxes that are inherent to ecosystem ecology. The current challenge is to understand how microplastic flows affect such pools and fluxes in terrestrial ecosystems.

Microplastics are mostly composed of carbon, among other elements. Microplastic addition to ecosystems thus represents a source of carbon independent of photosynthesis and net primary production. This polymer carbon likely has a slow turnover, because the material is mostly inert; however, the behavior and residence time of microplastics in soil are currently unknown. We also do not know the input rate of microplastic-carbon into ecosystems itself, because research hitherto has largely focused on quantifying particle numbers and types, rather than on the microplastic-derived carbon itself. Originally, most of this carbon is of fossil origin, rather than having recently been fixed from the atmosphere. Because of the resistance of microplastic to decomposition, it would be expected to accumulate in soils, where it needs to be accounted for in assessments of soil carbon storage (*8*), a major ecosystem function.

Other effects of microplastics will be indirect and likely depend on particle shape and size. For microplastic fibers, effects on soil aggregation, a key process governing soil structure, are quite well established (*9*). Soil aggregates are the crumbs contributing to soil structure and have a central role in shaping the habitat of soil orga-

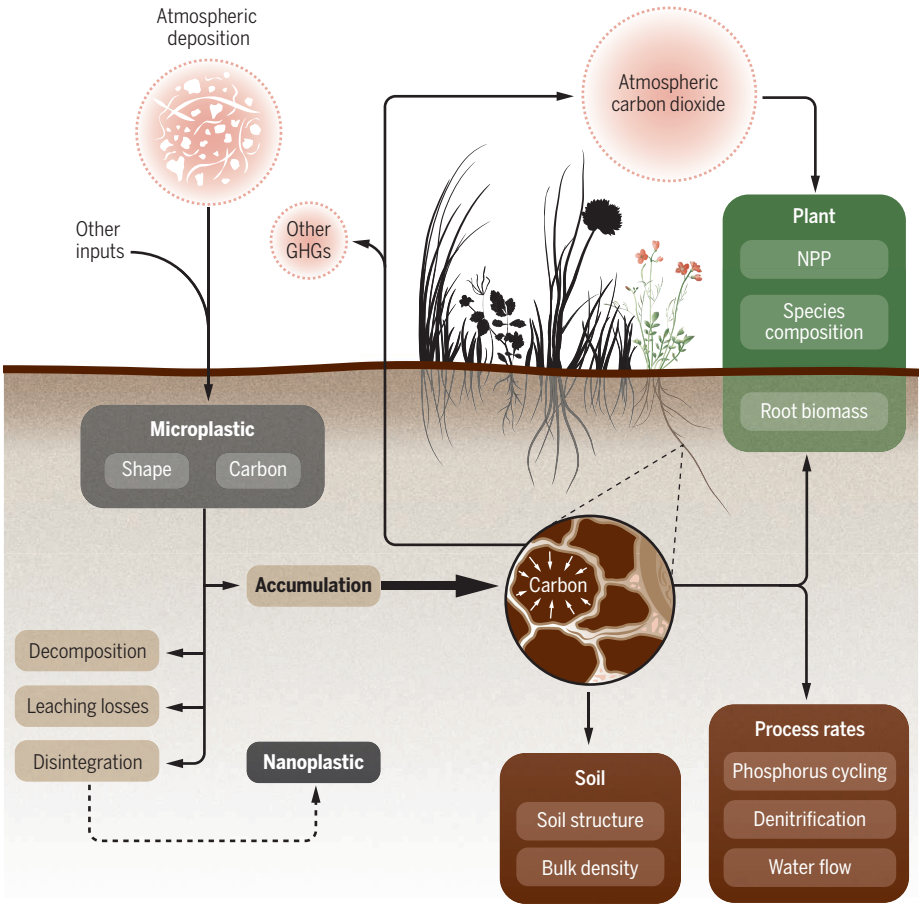
Institute of Biology, Freie Universität Berlin, Berlin, Germany. Email: rillig@zedat.fu-berlin.de; lehmann.anika@googlemail.com

PHOTO: ANIKA LEHMANN

Published by AAAS

Microplastic fluxes and associated ecosystem feedbacks

Deposition and accumulation of microplastics can affect soil properties, with consequences for process rates and net primary production (NPP), causing feedbacks to the atmosphere, including greenhouse gases (GHGs). So far, nanoplastic has unknown consequences for this system.



GRAPHIC: A. KITTERMAN/*SCIENCE* FROM M. C. RILLIG AND A. LEHMANN

nisms. Additionally, carbon compounds are stored within aggregates, where they are physically protected from being rapidly decomposed. Soil aggregates also determine the pore space in the soil overall, in turn influencing movement of gases and water, and the activity of associated microbial communities. A completely different indirect effect occurs because of lower soil bulk density in the presence of fibers. This can lead to enhanced plant growth, probably because roots experience less resistance when growing (*10*). However, negative effects on plants, likely related to plastic additives, are also possible (*11*). Which types of microplastics could promote or inhibit plant biomass production is an important area for future research.

Consequences for other element cycles are more uncertain. Direct effects will likely be minimal, because microplastics contain mostly negligible amounts of nitrogen and phosphorus (even though there are exceptions, such as polyamide). However, alterations to soil structure would be expected to change the rates of many microbial processes, including those in the nitrogen cycle. An example is denitrification, a process that occurs anaerobically, within the center of soil aggregates, which reduces nitrate and nitrite to gaseous forms of nitrogen, including nitrous oxide and nitrogen. Effects on emissions of nitrous oxide and other important greenhouse gases are only now being examined (*12*). One study (*10*) found an increase in arbuscular mycorrhizal fungi, a key symbiont group that associates with plant roots. If generally true, this could affect phosphorus cycling, because these symbionts transport nutrients, including phosphorus, to their plant hosts.

Plastic films, and likely fibers, may alter the flow of water in soils, including evaporation (*13*). Thus, effects on ecosystem water dynamics and energy balance, mediated by direct effects in soils or indirectly through plants, are also likely. Other possible ecosystem-level effects include altered rates of erosion owing to changes in soil aggregate stability.

There are some critical unknowns that need to be addressed before the impacts of microplastic pollution on terrestrial ecosystems and the subsequent feedbacks can be understood. Accurate, sensitive, low-cost, and harmonized detection methods and high-throughput sample processing are needed (*14*), to better understand the effects on turnover and transformation processes in the soil. Beyond this, research needs to cover more ecosystem types. Most research has so far focused on agricultural systems, which are expected to contain the largest amount of microplastics (*15*) because of input pathways (including sewage sludge, compost, and plastic mulching). We know much less about microplastics in other ecosystems, such as drylands or forests, where the microplastic dynamics might be quite different because of the different ecosystem structures.

Feedbacks to the Earth system can be expected. Microplastic itself represents fossil carbon, which might indirectly affect rates of net primary production and carbon storage in soils and alter the fluxes of greenhouse gases. The direction, magnitude, and balance of these effects should be a focus of future research. Empirical work will need to adopt tools such as mesocosm studies (outdoor experimental systems that examine the natural environment under controlled conditions) and carefully designed field experiments to address these problems. Microplastic pollution is an international problem, and international cooperation in research will be key. ■

REFERENCES AND NOTES

1. M. C. Rillig, *Environ. Sci. Technol.* **46**, 6453 (2012).
2. R. C. Thompson *et al.*, *Science* **304**, 838 (2004).
3. W. Wang, J. Ge, X. Yu, H. Li, *Sci. Total Environ.* **708**, 134841 (2020).
4. E. Huerta Lwanga *et al.*, *Environ. Sci. Technol.* **50**, 2685 (2016).
5. M. S. Bank, S. V. Hansson, *Environ. Sci. Technol.* **53**, 7177 (2019).
6. J. Brahney *et al.*, *Science* **368**, 1257 (2020).
7. D. Mohrig, *Science* **368**, 1055 (2020).
8. M. C. Rillig, *Environ. Sci. Technol.* **52**, 6079 (2018).
9. A. A. de Souza Machado *et al.*, *Environ. Sci. Technol.* **52**, 9656 (2018).
10. A. A. de Souza Machado *et al.*, *Environ. Sci. Technol.* **53**, 6044 (2019).
11. M. van Kleunen, A. Brumer, L. Gutbrod, Z. Zhang, *Plants, People Planet* **2**, 157 (2020).
12. X. Ren, J. Tang, X. Liu, Q. Liu, *Environ. Pollut.* **256**, 113347 (2020).
13. Y. Wan, C. Wu, Q. Xue, X. Hui, *Sci. Total Environ.* **654**, 576 (2019).
14. J. Li, Y. Song, Y. Cai, *Environ. Pollut.* **257**, 113570 (2020).
15. N. Weithmann *et al.*, *Sci. Adv.* **4**, eaap8060 (2018).

ACKNOWLEDGMENTS

M.C.R. acknowledges funding from an ERC Advanced Grant (694368), from the Federal Ministry of Education and Research (BMBF; projects BIBS and uPlastic), and from the Deutsche Forschungsgemeinschaft.

10.1126/science.abb5979

Published by AAAS

ASTRONOMY

Multiple, quiet, and close by

A planetary system around a nearby star offers the prospect for atmospheric study

By **Melvyn B. Davies**

Planetary systems around stars are common, with many or most stars possessing them. On page 1477 of this issue, Jeffers *et al.* (*1*) report the discovery of a multiplanet system around the low-mass star GJ 887. GJ 887 is an unusual host star, as it is relatively inactive, and the planetary system happens to be very close to our Sun. This makes the system a good candidate for study with the to-be-launched James Webb Space Telescope (JWST).

Stars form with a range of masses, the vast majority being less massive than the Sun. GJ 887 is a so-called red dwarf and has a mass of about one-half that of the Sun. Many red dwarfs are found to host planetary systems, typically containing multiple, low-mass planets (*2*).

Two methods have been used to discover most planets. In the transit method, planets are found when they periodically block out a fraction of the starlight as they pass in front of their host star. This allows many stars to be searched for orbiting planets simultaneously. For example, some one hundred thousand stars were observed at the same time in the Kepler mission (*3*). However, for planets to be found in this way, their planetary orbits must be close to edge-on as seen from Earth in order for them to be seen to pass in front of their host star. This means that only a small fraction of planets will be found using this method.

The planets around GJ 887 were discovered through a second approach: using the Doppler effect to measure the radial velocity of the host star over time. A planet and star orbit around a common center of mass, which means that the presence of a planet orbiting a star can be deduced by observing the periodic variation in the radial speed of the star. The radial velocity is measured from the Doppler shift of lines observed in the stellar spectra. Jeffers *et al.* observed GJ 887 nightly over a period of 3 months and combined their data with other observations. They found that GJ 887 has at least two planets, with orbital periods of 9.3 and 21.8 days, and a possible third planet farther out, with an orbital period of ~50 days.

GJ 887 is not unusual in possessing planets so close to it. Planets are generally believed to form farther away from their host star and migrate inward owing to a drag force from the protoplanetary disk from which they formed. As planets move inward, interactions between them can act to tune the ratio of their orbital periods, for example, leading to a 2:1 ratio where one planet goes around once in the time it takes the other to go around twice (*4*).

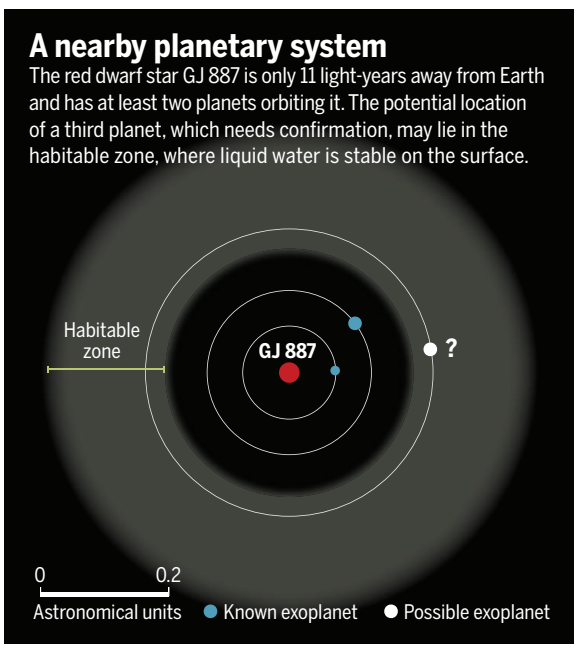
Planetary systems have been seen with neighboring planets in or close to these so-called orbital resonances. The TRAPPIST-1 system is a classic example of a set of planets in a resonance chain where all neighboring planets have periods with simple integer ratios (*5*). Planetary systems also appear to be full, with no room left for additional planets (*6*). The tugging and pulling of neighboring planets can act to change planetary orbits. This effect is damped out by the protoplanetary gaseous disk. Once the disk is evaporated by the host star, the planetary system might become unstable, leading to planet scattering and collisions that leave a system with the planets no longer in resonance chains (*7*). This scenario may well have happened for GJ 887.

The potential for life on any of the planets can be evaluated by looking at the likely surface conditions. The inner planets are probably too close to their host star, receiving about 2.5 and 8 times more energy from their host star than we receive from the Sun on Earth. However, red dwarfs are a good deal fainter than our own Sun, and the tentative third planet with a period of ~50 days could be in the so-called habitable zone, where the surface temperature is suitable to allow liquid water (*8*).

Solar flares and coronal mass ejections pose a serious threat to our life on Earth. For example, they could disrupt the electrical power supply grid. But serious events are rare, with the largest recent event striking Earth in 1859. Because red dwarfs have a higher magnetic activity, they have a much higher frequency of flares (*9*). Their flares can be lethal to life on planets in close orbit. But GJ 887 is a relatively inactive red dwarf, which makes the system particularly interesting. If someone had to live around a red dwarf, they would want to choose a quieter star like GJ 887.

GJ 887 is also attractive to observe because, at a distance of only ~11 light-years, it is one of the closest planetary systems to Earth. This opens the prospect of atmospheric study of its planets by looking for reflected light from them using the JWST.

These types of observations could tell us about the atmospheric makeup of these planets (*10*–*12*). If further observations confirm the presence of the third planet in the habitable zone, then GJ 887 could become one of the most studied planetary systems in the Solar neighborhood. ■



GRAPHIC: N. CARY/*SCIENCE*

Lund Observatory, Department of Astronomy and Theoretical Physics, Lund University, Lund, Sweden. Email: mbd@astro.lu.se

REFERENCES AND NOTES

1. S. V. Jeffers *et al.*, *Science* **368**, 1477 (2020).
2. G. D. Mulders, I. Pascucci, D. Apai, *Astrophys. J.* **814**, 130 (2015).
3. J. J. Lissauer *et al.*, *Astrophys. J. Suppl. Ser.* **197**, 8 (2011).
4. G. A. L. Coleman, R. P. Nelson, *Mon. Not. R. Astron. Soc.* **457**, 2480 (2016).
5. M. Gillon *et al.*, *Nature* **542**, 456 (2017).
6. L. M. Weiss *et al.*, *Astron. J.* **155**, 48 (2018).
7. B. Pu, Y. Wu, *Astrophys. J.* **807**, 44 (2015).
8. R. K. Kopparapu *et al.*, *Astrophys. J.* **765**, 131 (2013).
9. K. France *et al.*, *Astrophys. J.* **820**, 89 (2016).
10. L. Kreidberg, A. Loeb, *Astrophys. J.* **832**, L12 (2016).
11. I. A. G. Snellen *et al.*, *Astron. J.* **154**, 77 (2017).
12. E. M. R. Kempton *et al.*, *Publ. Astron. Soc. Pac.* **130**, 114401 (2018).

ACKNOWLEDGMENTS

M.B.D. is supported by the project grant 2014.0017 "IMPACT" from the Knut and Alice Wallenberg Foundation.

10.1126/science.abb0217

Published by AAAS

POLICY FORUM

MACHINE LEARNING

"Explaining" machine learning reveals policy challenges

The need to make objectives explicit may expose policy trade-offs that had previously been implicit and obscured

By **Diane Coyle**[1] *and* **Adrian Weller**[2,3]

There is a growing demand to be able to "explain" machine learning (ML) systems' decisions and actions to human users, particularly when used in contexts where decisions have substantial implications for those affected and where there is a requirement for political accountability or legal compliance (*1*). Explainability is often discussed as a technical challenge in designing ML systems and decision procedures, to improve understanding of what is typically a "black box" phenomenon. But some of the most difficult challenges are nontechnical and raise questions about the broader accountability of organizations using ML in their decision-making. One reason for this is that many decisions by ML systems may exhibit bias, as systemic biases in society lead to biases in data used by the systems (*2*). But there is another reason, less widely appreciated. Because the quantities that ML systems seek to optimize have to be specified by their users, explainable ML will force policy-makers to be more explicit about their objectives, and thus about their values and political choices, exposing policy trade-offs that may have previously only been implicit and obscured. As the use of ML in policy spreads, there may have to be public debate that makes explicit the value judgments or weights to be used. Merely technical approaches to "explaining" ML will often only be effective if the systems are deployed by trustworthy and accountable organizations.

The promise of ML is that it could lead to better decisions, yet concerns have been raised about its use in policy contexts such as criminal justice and policing. A fundamental element of the demand for explainability is for explanation of what the system is "trying to achieve." Most policy decision-making makes extensive use of constructive ambiguity to pursue shared objectives with sufficient political consensus. There is thus a tension between political or policy decisions, which trade off multiple (often incommensurable) aims and interests, and ML, typically a utilitarian maximizer of what is ultimately a single quantity and which typically entails explicit weighting of decision criteria.



Will the demand for explainable ML systems for decisions in the justice system require more explanations from policy-makers, too?

We focus on public policy decision-making using ML algorithms that learn the relationships between data inputs and decision outputs. As a first step, policy-makers need to decide among a number of possible meanings of explainability. These range from causal accounts and post hoc interpretations of decisions (*3*) to assurance that outcomes are reliable or fair in terms of the specified objectives for the system (*4*). For example, the explainability requirements for ML systems used by local authorities to determine benefit payments will differ greatly from those required for the enforcement of competition policy with respect to pricing by online merchants. Each of the specific meanings of explainability has different technical requirements, which will imply choices about where efficiency and cost might need to be sacrificed to deliver both explainability and the desired outcomes. Choosing which meaning is relevant is far from a technical question (though what can be provided depends on what is technically feasible). Thus, those seeking explainability will need to specify, in terms translatable to how ML systems operate, what exactly they mean, and what kind of evidence would satisfy their demand (*5*). It must also be possible to monitor whatever explanations are provided, and there must be practical methods to enforce compliance.

Furthermore, policy institutions starting to deploy algorithmic or ML-based decision systems, such as the police, courts, and government agencies, are operating in the context of declining trust in some aspects of public life. This context is important for understanding demands for explainability, as these may in part reflect broader legitimacy demands of the policy-making process. If an organization is not trusted, its automated decision procedures will likely also be distrusted. This implies a broader need for trustworthy processes and institutions, for "intelligent accountability" as the result of informed and independent scrutiny, communicated clearly to the public (*6*). Satisfying the demand for explainability implies testing the trustworthiness of the organizations using ML systems to make decisions affecting individuals. Evaluation requires comparing outcomes against a benchmark, which can be the baseline situation, or a specified desired outcome.

Taking the demand for explainability as a demand for accountability, the promise of ML is that it could lead to more legitimate and better decisions than humans can make, on some measure. Potential benefits are clearly demonstrable in some forms of medical diagnosis (*7*) or monitoring attempted financial fraud (*8*). In these domains, there is general agreement on a straightforward quantity to optimize, and the incentives of principals (citizens or customers) and agents (public or corporate decision-makers) are aligned. Public concern about the use of ML focuses on other domains, such as marketing or policing, where there may be less agreement about (or trust in) the aim of either the ML system or the organization using it.

[1]Bennett Institute for Public Policy, University of Cambridge, Cambridge, UK. [2]University of Cambridge, UK. [3]The Alan Turing Institute, London, UK.
Email: dc700@cam.ac.uk; aw665@cam.ac.uk

PHOTO: THE WOLF LAW LIBRARY/FLICKR/CC BY-NC-ND

Published by AAAS

These concerns highlight a key challenge posed by the use of ML in policy decisions, which is that ML processes are almost always set up to optimize an objective function; this optimization goal can be described in anthropomorphic terms as the "intention" of the system. Yet there is often little or no explicit discussion by policy-makers when considering using ML systems about what conflicting goals, benefits, and risks may trade off against each other as a result. One reason for this is that it is inherently challenging to specify a concrete objective function in sociopolitical domains (*9*). For example, like current ML systems, economists' decisions are informed by estimates of statistical relationships between directly observable and unobservable variables, derived from data generated by a complex environment. Yet economic policies such as tax changes often fail to take into account all relevant factors in the decision environment, or likely behavior changes, in specifying the objective function (*10*). The use of ML systems in other policy contexts will expand the scope of such unintended consequences.

Given that the dominant paradigm of machine learning is based on optimization, the use of ML in policy decisions thus speaks to a fundamental debate about social welfare. From the perspective of ethical theories, ML is largely consequentialist: A machine system is configured on the basis of its ability to achieve a desired outcome. Conventional policy analysis is similarly typically based on consequentialist economic social welfare criteria. The well-known impossibility theorems in social choice theory (*11*) establish that when the goal is to aggregate individual choices under a set of reasonable social decision rules, it is impossible to satisfy a set of desirable criteria simultaneously, and thus impossible to achieve a set of desired outcomes by optimizing a single quantity. Critics of consequentialist economic policy analysis argue that people have multidimensional, probably incommensurable, and possibly contradictory objectives, so that imposing utilitarian decision-making procedures will conflict both with reality and with ethical intuitions (*12*).

Nevertheless, policy choices are made, so there has always been an unavoidable, albeit often implicit, trade-off or weighting of different objectives (*12*). For example, cost-benefit analysis can incorporate environmental and cultural, as well as financial, considerations, but converts all of these into monetary values. Any choice made when there are multiple interests or trade-offs will imply weights on the different components. As these trade-offs are codified into ML objective functions, the weights given to competing objectives comprise a first-line characterization of how conflicts will be resolved. Using ML systems in political contexts is extending the use of optimization; progress in making these ML systems more understandable to policy-makers will make the de facto choices between competing objectives more explicit than they have been previously (*13*). Greater explainability is therefore likely to have to lead to a more explicit political, not wholly technical, debate.

Distilling concrete, unambiguous objectives in this way may turn out to be extremely challenging, for ambiguity about objectives is often useful in policy-making precisely because it blurs uncomfortable conflicts of interest. In many domains, policies generally emerge as a pragmatic compromise between fundamentally conflicting aims. For example, people who disagree about whether the justice system should be retributive or rehabilitative may well be able to agree on specific sentencing policies. Such incompletely theorized agreements "Play an important function in any well-functioning democracy consisting of a heterogeneous population" (*14*, p. 1738). The omission of discussion of ultimate aims can make it easier to achieve consensus on difficult issues. As there is some (limited) scope to interpret means to achieve the objective with flexibility, the "weighting" of different fundamental aims remains implicit, and diverse political communities can make progress.

An optimistic conclusion would be that being forced by the use of ML systems to be more explicit about policy objectives could promote useful debate leading in the long run to more considered outcomes. ML systems can be used to explore choices and outcomes on different counterfactual high-level objectives, such as retribution or rehabilitation in justice, enabling considered human judgments. However, it may in practice be impossible to specify what we collectively truly want in rigid code. For example, many local governments do not seem to be engaging in public consultation when they adopt predictive ML systems, such as to flag "troubled" families that are likely to need interventions. Although steps such as explicitly adding uncertainty to the ML objective might address this challenge of imperfectly specified objectives in future, ML systems are unable at present to offer wisely moderated solutions to ambiguous objectives (*15*).

Human decision-makers can make use of common sense or tacit knowledge, and often override decisions indicated by an economic model or other formal policy analysis, and they will be able to do the same when assisted by ML. Yet, demanding that ML systems be explainable is likely to make the trade-offs between different objectives far more explicit than has been the norm previously. Ultimately, the use of explainable ML systems in the public sector will make a broader debate about social objectives and social justice newly salient. Providing explanations requires being transparent about the systems' objectives — forcing clarity about choices and trade-offs previously often made implicitly — and how their predictions or decisions draw on patterns revealed by a fundamentally biased social and institutional system. Moreover, whereas democratic political systems often look to resolve conflicts through constructive ambiguity—or in other words, the failure to explain—ML systems may require ambiguous objectives to be resolved unequivocally. So, although the need for explainability certainly poses technical challenges, it poses political challenges too, which have not to date been widely acknowledged. Yet, the increasing scope of ML, and progress in delivering explainability, in politically salient areas of policy could shine a helpful spotlight on the conflicting aims and the implicit trade-offs in policy decisions, just as it already has on the biases in existing social and economic systems. ■

REFERENCES AND NOTES

1. B. Dattner, T. Chamorro-Premuzic, R. Buchband, L. Schittler, *The legal and ethical implications of using AI in hiring*, *Harv. Bus. Rev.* April, 25 (2019); https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in-hiring.
2. R. Richardson, J. Schultz, K. Crawford, *New York Univ. Law Rev.* **192**, 204 (2019).
3. Z. Lipton, The mythos of model interpretability (2017); https://arxiv.org/pdf/1606.03490.pdf .
4. T. Miller, Explanation in artificial intelligence: Insights from the social sciences (2018); https://arxiv.org/pdf/1706.07269.pdf.
5. P. Madumal, T. Miller, L. Sonenberg, F. Vetere, A grounded interaction protocol for explainable artificial intelligence (2019); https://arxiv.org/pdf/1903.02409.pdf.
6. O. O'Neill, *Int. J. Philos. Stud.* **26**, 293 (2018).
7. J. De Fauw *et al.*, *Nat. Med.* **24**, 1342 (2018).
8. S. Aziz, M. Dowling, in *Disrupting Finance: FinTech and Strategy in the 21st Century*, T. Lynn, G. Mooney, P. Rosati, M. Cummins, Eds. (Palgrave, 2019), pp. 33–50.
9. P. Samuelson, *Foundations of Economic Analysis* (Harvard University Press, 1979), chap. 8, pp. 203–252.
10. J. Le Grand, *Br. J. Polit. Sci.* **21**, 423 (1991).
11. A. Sen, *Am. Econ. Rev.* **89**, 349 (1999).
12. E. Anderson, *Value in Ethics and Economics* (Harvard Univ. Press, 1993).
13. S. Grover, C. Pulice, G. I. Simari, V. S. Subrahmanian, *IEEE Trans. Comput. Soc. Syst.* **6**, 350 (2019).
14. C. R. Sunstein, *Harv. Law Rev.* **108**, 1733 (1995).
15. M. Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar, 2016).

ACKNOWLEDGMENTS

We are grateful to M. Kenny and N. Rabinowitz for helpful comments. A.W. acknowledges support from the David MacKay Newton research fellowship at Darwin College, The Alan Turing Institute under EPSRC grants EP/N510129/1 and TU/B/000074, the Leverhulme Trust via CFI, and the Centre for Data Ethics and Innovation.

10.1126/science.aba9647

Published by AAAS



BOOKS *et al.*

A woman speaks to security personnel in Washington, D.C., on 3 June 2020. Biases can affect how individual law enforcement officers interact with citizens.

PHOTO: OLIVIER DOULIERY/AFP/GETTY IMAGES

PSYCHOLOGY

Confronting bias

Prejudice and discrimination are deeply personal, but their effects are profound

By **Calvin K. Lai**

On an idyllic summer day in 2009, Pragya Agarwal and her 9-year-old daughter went shopping for a new school uniform. As they were walking back to their car, an armed police officer stopped them. The officer told them that a customer had reported them as people who "looked like shoplifters" and were "suspicious." Agarwal and her daughter were eventually allowed to leave, but the consequences of that incident were enduring. Experiences like this underscore the prevalence of group-based biases that are the focus of Agarwal's new book, *Sway*.

She begins by describing the origins of biases. In everyday life, we often have either too little or too much information to make optimal decisions. Our biases serve as mental shortcuts that help us make "good enough" decisions in a complex world. From an evolutionary perspective, this ability to quickly separate good from bad and friend from foe was essential for survival.

These biases, however, are prone to systematic errors that can have grave consequences in modern society. Take our preference for similarity: We seek out information that confirms views we already hold, prefer to live with people who are like us, and discriminate against those who do not share our views. We also prefer the status quo, meaning that we are biased toward current conditions, even when they reinforce the oppression of a minority group.

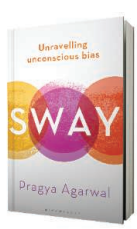
Agarwal documents biases across many social distinctions. She notes the double standards to which women in leadership are held, health care workers' beliefs that Black individuals experience less severe pain than non-Black individuals, and the insidious stereotype of Asian people as "model minorities." She describes how anti-fat stigma is internalized, how elderly people are perceived as burdens, and how accents can confer a host of hidden stereotypes. Although her approach is exhaustive, it falters in its synthesis.

Research tells us that prejudice, stereotyping, and discrimination are pervasive. Less attention is paid to how to connect these findings together. What common psychological principles underlie the prevalence of unconsciously biased discrimination?

I suspect that Agarwal's lack of emphasis on common principles is due to concept creep: an expansion of the idea of unconscious bias that goes beyond its original meaning (*1*). She assumes that many social disparities are necessarily caused by unconscious biases. However, many of the disparities she describes are better explained by structural causes or conscious prejudices. She writes, for example, that "implicit aversive attitudes" are an explanation for racial housing segregation. Racial housing segregation can often be more parsimoniously explained by institutional forces, such as a history of redlining, exclusionary zoning regulations, and economic inequalities. When psychological biases are considered, there is clearer evidence of real estate agents consciously steering homebuyers to certain neighborhoods on the basis of the homebuyers' race rather than as a result of unconscious bias.

Another form of concept creep is in the use of the term "unconscious" to encompass biases that are subtle but conscious. Some may just be hidden (for example, an avowed racist who lies about the reasons for treating a racial minority poorly), or they may be fast to arise (for example, a visceral reaction to a person who speaks with a foreign accent).

In the book's final chapter, Agarwal describes approaches for debiasing, including raising awareness of bias, confronting others tactfully, removing biasing information from job applications, perspective-taking, and role models. Many of these approaches have firm scientific backing and a track record of success in reducing prejudice and discrimination. But most of what she offers is psychologically and individually oriented. If the disparities Agarwal discusses have structural causes, then these individualistic approaches may fall short.



Sway: Unravelling Unconscious Bias
Pragya Agarwal
Bloomsbury Sigma, 2020. 448 pp.

Promoting equality takes time, effort, and systemic change. The enduring influence of structural factors, such as unequal education and wealth, creates social environments that cannot be escaped. At the same time, structural inequalities depend on the countless individual choices that people make for others in everyday life. An integrated perspective that deeply considers the relationship between individuals and society would have helped illustrate a key lesson of research on inequality: Bias is deeply personal, but it is also universal. ■

The reviewer is at the Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO 63130, USA. Email: calvinlai@wustl.edu

REFERENCES AND NOTES

1. N. Haslam, *Psych. Inquiry* **27**, 1 (2016).

10.1126/science.abb9316

Published by AAAS

DEMOGRAPHY

Like other species, we have always been on the move

Human migration and cultural exchange represent progress, not peril, argues a journalist

By **Khalid Koser**

It is ironic to be reviewing Sonia Shah's latest book, a thoughtful and thought-provoking defense of migration, under coronavirus disease 2019 (COVID-19) lockdown. One irony is that this is the perfect time to be reading her excellent earlier book, *Pandemic: Tracking Contagions, from Cholera to Ebola and Beyond*. Another is that migration—at least by people and across borders—has virtually come to a standstill. But this is no reason to shelve *The Next Great Migration: The Beauty and Terror of Life on the Move*. International migration will rebound.

Shah would probably contend that this is because people have never been fixed in place and that their migration is natural. One of the more original aspects of her book is that it depicts human migration in the context of migration by plants, animals, and even viruses. My analysis is more prosaic. I believe that the forces of globalization that drive migration—inequality; disparities in development, democracy, and demography; the global jobs crisis; segmented labor markets—will be exacerbated by COVID-19 and government responses to it.

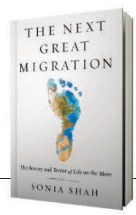
I also disagree with the book's premise that, the current global standstill notwithstanding, climate change will trigger a new great migration. Most sensible scholars agree, of course, that the effects of climate change will accelerate migration. But there remains little consensus concerning how many people will be forced to move, where they will move from and to, or how soon this migration will happen. But whatever the numbers, Shah is right to assert that we need to start respecting the rights of migrants and seizing the opportunities of migration.

Some of the book's most powerful sections reflect on the author's own migrant heritage: Fifty years after her parents' arrival in the United States, she recounts, their migration remains the central fact of their lives. She also shares the stories of migrants whom she has met around the world in the course of her research. She speaks, for example, with individuals who have fled Afghanistan and Eritrea and taken the precarious journey across mountain ranges, deserts, and the sea to Europe, as well as with Latin American migrants who have made it to the United States without authorization. Shah does not overdramatize or sentimentalize their stories. Instead, the reader is left with an overwhelming sense that people who migrate out of desperation are resigned to experiencing great hardship as a means to improve their lives.

Shah seeks to normalize human migration by situating it within the broader context of other species' migratory behavior. Humans, she contends, do not belong in a particular place any more than do butterflies, fish, or trees. She demonstrates how migration has been integral to the evolution of humankind, by injecting change and innovation into cultural practices. The mixing of people from different places represents progress, not peril.

Shah methodically dismantles the racial "science" that still underlies certain attitudes toward those who migrate and rejects arguments for controlling migration on the grounds that it could potentially lead to overpopulation, an idea that originated in the work of Thomas Malthus. It is not particularly groundbreaking to take such attitudes and assumptions to task, but Shah does so in an engaging manner, weaving history with geography and storytelling with science.

The Next Great Migration: The Beauty and Terror of Life on the Move
Sonia Shah
Bloomsbury, 2020. 400 pp.



Still, I think Shah overestimates the importance of such ideas and beliefs in explaining the xenophobia, discrimination, and antimigrant sentiment and policies that abound today, which I believe are more driven by perceptions about the economic impact and security implications of immigration. I do, however, agree with her pessimistic prognosis that unless we shift attitudes, our default response to more migration will be to build more walls and enact more stringent border controls. Her closing chapter is a grim litany of migrant deaths and the detention of immigrant children.

Countless scholars, analysts, and researchers have produced evidence that migration overwhelmingly benefits economies and societies and that there is a higher rate of criminality and violent extremism among nationals than among migrants, but seemingly to little effect. Perhaps Shah's more fundamental plea—that migration is normal, that we are all migrants, and that, like nature, migration can be both beautiful and terrifying—will have more traction. ■

10.1126/science.abb9025

The reviewer is at the Maastricht Graduate School of Governance, Maastricht University, 6211 AX Maastricht, Netherlands. Email: k.koser@gcerf.org



A woman delivers groceries to a senior citizen.

PHOTO: ISTOCK.COM/ALEX POTEMKIN

PODCAST

Humankind: A Hopeful History
Rutger Bregman
Little, Brown, 2020. 480 pp.

Difficult as it might be to believe, humans are hardwired for kindness, cooperation, and trust, argues historian Rutger Bregman in his latest book. This week on the *Science* podcast, Bregman discusses the evolutionary origins of humanity's goodness and its implications for our future.

10.1126/science.abd0367

Published by AAAS



The COVID-19 pandemic has revealed vulnerabilities in waste management chains, which could hinder disease containment and increase environmental pollution.

Edited by **Jennifer Sills**

COVID-19's unsustainable waste management

The coronavirus disease 2019 (COVID-19) pandemic has led to an abrupt collapse of waste management chains. Safely managing medical and domestic waste is crucial to successfully containing the disease (*1*). Mismanagement can also lead to increased environmental pollution. All countries facing excess waste should evaluate their management systems to incorporate disaster preparedness and resilience.

Wuhan, the COVID-19 epicenter of China, experienced a massive increase of medical waste from between 40 and 50 tons/day before the outbreak to about 247 tons on 1 March (*2*). Cities such as Manila, Kuala Lumpur, Hanoi, and Bangkok experienced similar increases, producing 154 to 280 tons more medical waste per day than before the pandemic (*3*). Meanwhile, the widespread lockdown has caused a substantial increase in domestic waste in the United Kingdom (*4*). These large amounts of waste require collection and recycling, both of which are compromised as a result of manpower shortages and efforts to enforce infection control measures (*5*, *6*).

Disrupted services have led to waste mismanagement increases of 300% in some rural UK communities (*7*). With fewer options available, traditional waste management practices such as landfills and incineration are replacing more sustainable measures such as recycling, with adverse effects on the environment (*8*). The U.K. Environment Agency further threatens the environment by allowing temporary storage of waste and incineration ash at sites that have not been granted a permit, as is usually required (*9*, *10*).

To address the overflow of medical waste, the United Kingdom and other affected countries should install mobile treatment systems near hospitals and health care centers (*2*). The design and analysis of sustainable waste management chains, including logistics, recycling, and treatment technologies and policies, should be prioritized (*11*). To reduce the socio-economic and environmental impacts of waste management, the whole system must be considered, including waste generation, collection, transport, recycling and treatment, recovered resource use, and disposal of remains. Protecting waste management chains will help achieve sustainable cities and communities as outlined in the UN Sustainable Development Goals (*12*).

Siming You[1], Christian Sonne[2], Yong Sik Ok[3,4]*
[1]University of Glasgow, Glasgow, UK. [2]Aarhus University, Roskilde, Denmark. [3]Korea University, Seoul, Korea. [4]APRU Sustainable Waste Management Program, Korea University, Seoul, Korea.
*Corresponding author:
Email: yongsikok@korea.ac.kr

REFERENCES AND NOTES

1. World Health Organization, "Water, sanitation, hygiene, and waste management for the COVID-19 virus: Interim guidance" (2020).
2. Z. H. Si, Y. Li, "Medical waste treatment in Wuhan from emergency to stability," *Xin Hua Net* (2020); www.xinhuanet.com/local/2020-04/01/c_1125796126.htm [in Chinese].
3. Asian Development Bank, "Managing infectious medical waste during the COVID-19 pandemic" (2020).
4. Association of Directors of Environment, Economy, Planning, and Transport, "COVID 19—waste survey results w/c 27 April" (2020).
5. European Commission, "Waste management in the context of the coronavirus crisis" (2020).
6. Association of Cities and Regions for Sustainable Resource Management, "Municipal waste management and COVID-19" (2020).
7. K. P. Roberts *et al.*, "Rubbish is piling up and recycling has stalled—waste systems must adapt," *The Conversation* (2020).
8. J. J. Klemeš *et al.*, *Renew. Sustain. Energ. Rev.* **127**, 109883 (2020).
9. UK Environment Agency, "COVID-19 and temporary storage of incinerator bottom ash aggregate: RPS C16" (2020).
10. UK Environment Agency, "COVID-19 and storing waste at unpermitted sites due to exceeding your storage limits: RPS C17" (2020).
11. R. Djalante, R. Shaw, A. DeWit, *Prog. Disaster Sci.* **6**, 100080 (2020).
12. United Nations, "Sustainable development goal 11: Make cities and human settlements inclusive, safe, resilient, and sustainable" (2020).

10.1126/science.abc7778

Misguided forest action in EU Biodiversity Strategy

After failing to achieve the 2020 Aichi biodiversity targets, governing bodies are preparing targets for the future. In doing so, they must acknowledge that effective policies address not only the quantity of protected areas but also the quality of protection and the management of the surrounding matrix. The European Union's Biodiversity Strategy for 2030 (*1*), launched in May, is an opportunity to enact successful conservation. The Biodiversity Strategy commits to protecting the European Union's

PHOTO: ISTOCK.COM/AVIGATORPHOTOGRAPHER

Published by AAAS



remaining forests and restoring forest ecosystem resilience. However, the plans for implementing these goals are misguided.

The EU Strategy commits to "strictly protecting…primary and old-growth forests." Known primary forests represent 0.7% of Europe's forest area, of which only 46% are strictly protected (*2*). Strictly protecting the remaining forests is the right strategy, given that forest management seriously threatens remaining primeval and old forests in Europe, including the Białowieża Forest, Carpathian forests, and Scandinavian Mountains Green Belt (*3*, *4*). However, the EU Strategy overlooks the need for spatial coherence. Tiny protected patches of forest will be of little relevance if intensive forestry prevails in the surrounding matrix. An ecosystem-based approach to forest management must complement protection efforts.

The EU Strategy also commits to restoring forest ecosystems but only offers planting more than 3 billion trees as a concrete action. Planting trees in forest habitats may have detrimental effects (*5*, *6*). The EU Strategy doesn't specify what kinds of trees should be planted, a necessary detail given that the forestry sector in most countries increasingly undermines ecosystem functionality by introducing exotic species believed to be better adapted to future climates (*7*).

Instead of planting trees, conservationists should focus on reducing the rate of forest degradation (namely, tree harvesting) and supporting natural renewal processes. The EU Strategy does not consider natural restoration of spontaneously regenerating forests, the surface of which has decreased 7% worldwide over the past 30 years (*8*). Simply allowing the forests to naturally regrow often results in more trees at much lower costs than planting (*6*). The EU Strategy should advocate a hands-off strategy for safeguarding ecosystemic responses to degradation and environmental change. Policies should support the natural recruitment and selection for trees with greater resistance to insect attacks or extreme events (*9*) and biomass-rich forests with closed canopies (*10*), and they should prevent road construction through valuable forest patches (*11*). Actions aimed at promoting the adaptation of forests to global change and increasing their resilience should be based on all available ecological science and require a more complex view than a simplistic planting strategy. The EU Forest Strategy planned for 2021 should focus on developing a holistic approach with a clearly defined timetable of actions.

Nuria Selva[1]*, Przemysław Chylarecki[2], Bengt-Gunnar Jonsson[3,4], Pierre L. Ibisch[5]
[1]Institute of Nature Conservation Polish Academy of Sciences, 31-120 Kraków, Poland. [2]Museum and Institute of Zoology, Polish Academy of Sciences, 00-679 Warszawa, Poland. [3]Department of Natural Sciences, Mid Sweden University, SE-851 70 Sundsvall, Sweden. [4]Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden. [5]Centre for Econics and Ecosystem Management, Eberswalde University for Sustainable Development, 16-225 Eberswalde, Germany.
*Corresponding author.
Email: nuriselva@gmail.com

REFERENCES AND NOTES

1. European Commission, "EU Biodiversity Strategy for 2030," (2020); https://ec.europa.eu/environment/nature/biodiversity/strategy/index_en.htm.
2. F. M. Sabatini *et al.*, *Divers. Distrib.* **24**, 1426 (2018).
3. E. Stokstad, *Science* **358**, 1240 (2017).
4. B.-G. Jonsson *et al. Forests* **10**, 564 (2019).
5. M. Żmihorski *et al.*, *Science* **361**, 238 (2018).
6. K. D. Holl, P. H. S. Brancalion, *Science* **368**, 580 (2020).
7. F. Krumm, L. Vítková, Eds., "Introduced tree species in European forests: opportunities and challenges" (European Forest Institute, 2016).
8. Food and Agriculture Organization of the United Nations and UN Environment Programme, *The State of the World's Forests: Forests, Biodiversity and People* (FAO and UNEP, 2020).
9. D. L. Six, C. Vergobbi, M. Cutter, *Front. Plant Sci.* **9**, 993 (2018).
10. F. Zellweger *et al.*, *Science* **368**, 772 (2020).
11. P. L. Ibisch *et al.*, *Science* **354**, 1423 (2016).

10.1126/science.abc9892

Afforestation falls short as a biodiversity strategy

The recent EU Biodiversity Strategy for 2030 (*1*) recognizes the importance of biodiversity for increasing our resilience to natural disasters and pandemics and, thus, for human well-being. Although it proposes ambitious measures such as reversing pollinator decline and controlling invasive species, it also introduces the ill-advised idea of planting 3 billion trees.

Massive tree plantation programs (*2*, *3*) have been strongly criticized by the scientific community for their negative ecological and economic impacts and their limited role in climate change and CO_2 mitigation (*4*–*8*). The specific number of trees proposed in the EU Strategy suggests a lack of a serious, science-based ecological assessment of actual restoration needs. Meeting such a target could threaten biodiverse treeless ecosystems (*4*, *6*, *7*, *9*) and would waste an opportunity to implement ecologically sound management practices to restore fully functionally integrated mosaics of natural, seminatural, and sustainable agricultural ecosystems.

Massive tree planting could also substantially change the fire regime, especially given the increasing frequency of heat waves and droughts in an area with high population density (*10*). The probability of large intense fires that threaten biodiversity and human assets is largely influenced by the type, amount, and continuity of biomass. Therefore, determining how many trees should be planted is less important than figuring out the most safe and effective conservation strategy.

We need to move toward optimizing our landscapes. A diverse mosaic of nature-based production systems should be interspersed with protected natural areas to maximize biodiversity, resilience, and ecosystem services. Trees are not synonymous with biodiversity. Policy-makers and society need to internalize this message to make proper decisions in the context of environmental and health crises.

Susana Gómez-González[1,2]*, Raúl Ochoa-Hueso[1], Juli G. Pausas[3]
[1]Departamento de Biología-IVAGRO, Universidad de Cádiz, Puerto Real, Spain. [2]Center for Climate and Resilience Research (CR)2, Santiago, Chile. [3]Centro de Investigaciones sobre Desertificación (CIDE-CSIC), Valencia, Spain.
*Corresponding author.
Email: susana.gomez@uca.es

REFERENCES AND NOTES

1. European Commission, "EU Biodiversity Strategy for 2030" (2020); https://ec.europa.eu/environment/nature/biodiversity/strategy/index_en.htm.
2. International Union for Conservation of Nature, The Bonn Challenge (2020); www.bonnchallenge.org.
3. J.-F. Bastin *et al.*, *Science* **365**, 76 (2019).
4. W. J. Bond *et al.*, *Trends Ecol. Evol.* **34**, 963 (2019).
5. D. Baldocchi, J. Penuelas, *Glob. Change Biol.* **25**, 1191 (2019).
6. J. W. Veldman *et al.*, *Science* **366**, eaay7976 (2019).
7. K. D. Holl, P. H. Brancalion, *Science* **368**, 580 (2020).
8. M. Jiang *et al.*, *Nature* **580**, 227 (2020).
9. J. Scurlock, D. Hall, *Glob. Change Biol.* **4**, 229 (1998).
10. D. de Rigo *et al.*, "Forest fire danger extremes in Europe under climate change: Variability and uncertainty" (JRC Tech. Rep. 28926, Publications Office of the European Union, Luxembourg, 2017).

10.1126/science.abd3064

Published by AAAS



AAAS NEWS & NOTES

A view from my basement

AAAS CEO reflects on his first 6 months on the job

By **Sudip S. Parikh**

On 6 January, when I first walked through the doors of our beautiful living monument to science on New York Avenue in downtown Washington as the American Association for the Advancement of Science's new chief executive officer and executive publisher of the *Science* family of journals, I had plenty of ideas for how my first 90 days, 6 months, and year on the job would go.

I articulated this vision on the editorial page of the 31 January issue of *Science* in the context of "Envisioning Tomorrow's Earth," the theme of our 2020 Annual Meeting. I noted the extraordinary advances of our era, including treating and curing disease and deepening our understanding of the Universe. I also identified some of the critical topics of our time in prioritizing the need to safeguard the scientific ecosystem for diverse voices and inclusivity and leverage the power of all great minds.

At the time that editorial went to press, the United States had fewer than 10 confirmed cases of COVID-19 and George Floyd, Ahmaud Arbery, and Breonna Taylor were alive.

When we gathered in Seattle in February to envision tomorrow's Earth, most of us had no idea how the tomorrows we were about to experience would change our Earth forever. COVID-19 is now a global pandemic that has infected more than 7 million people, killed more than 400,000, and fundamentally changed the way the world operates. The tragic and needless deaths of George Floyd, Ahmaud Arbery, and Breonna Taylor have brought to the fore the need for decisive and lasting action to declare the importance of Black lives, voices, and contributions to all aspects of our society—and, central to our work at AAAS, to the STEM ecosystem.

Parikh calls for decisive and lasting action to declare the importance of Black lives, voices, and contributions to all aspects of our society.

Almost every generation has arrived at similar moments of opportunity born of tragedy. AAAS is uniquely positioned as an organization to be at the forefront of these generation-defining issues, and we are taking action to make sure these moments do not slip away.

Science, engineering, and medicine are not immune to the discrimination, subjugation, and silencing of minority people and voices. We are too often unwitting perpetuators of the status quo, and the reasons are deeply ingrained in the systems that govern the conduct of these fields. When we hold up a mirror to the scientific enterprise, we see that it is not only politicians and law enforcement that need to be reminded that Black lives matter.

On 10 June, AAAS participated in a grassroots movement called #ShutDownSTEM with the goal to get our community to stop business as usual and consider the facts—and, with those facts in hand, to act. Despite so many in our community being learned and well-read, outside of the social science circles that have studied this issue for years, few of us know very much about the systemic racism that has kept generations of Black Americans from realizing success in the STEM workforce. #ShutDownSTEM encouraged us to take responsibility in our own lives and circles of influence to be actively anti-racist and to recognize the time and space required for our Black colleagues to heal.

#ShutDownSTEM is just one step toward following through on fulfilling the hope for positive change in our community and across society. Other steps are already under way within AAAS, and many more have yet to be conceived.

For decades, under the leadership of Shirley Malcom, AAAS has been at the forefront of the discussion about making the STEM enterprise more diverse, equitable, and inclusive. As an African American growing up in Birmingham, Alabama, in the 1950s and 1960s, Malcom went through segregated schools that were underresourced because of the assumption that no scientists or engineers would ever be produced from within their walls. She defied those odds, eventually completing her doctoral work in ecology at Penn State, dedicating her life's work to achieving equality in STEM education and the workforce, and going on to serve as a member of the National Science Board, a trustee of two of our nation's most prestigious research universities, and a world-renowned leader in the conversation that we've been having for quite some time—but to which too few have paid sufficient attention.

Progress to increase the participation and advancement of underrepresented groups in STEM has been incremental. For the past 10 years, AAAS, with the support of the U.S. National Science Foundation (NSF), has convened the Emerging Researchers National (ERN) Conference in STEM. Each year, more than a thousand undergraduate and graduate students from underrepresented communities come together to share their research and develop their careers—with the goal of broadening participation in STEM fields. In describing the kind of experience that ERN attendees have, Malcom has said, "For some students it's their first time on a plane,

PHOTO: ROBB COHEN PHOTOGRAPHY & VIDEO

Published by AAAS

never mind their first time presenting at a scientific meeting."

For the past 3 years AAAS has convened, also with NSF support, the HBCU Making and Innovation Showcase. The event brings together more than 80 students and faculty from historically Black colleges and universities (HBCUs) for 2 days of workshops and training on invention and entrepreneurship. The students are divided into teams that create an innovative solution to a community problem that relates to one of the United Nations Sustainable Development Goals. This year's winning team, comprised of students from Clark Atlanta University and Morehouse College, conceived a network of communication devices for use during natural disasters and other emergency situations.

Convenings like ERN and the HBCU Making and Innovation Showcase have surely helped. Bit by bit, person by person, they have encouraged many students and professionals who might have otherwise abandoned STEM careers to stay the course, knowing that their voices and contributions are valued and essential to the long-term success of the STEM enterprise. As Malcom posited in her 9 May 2019 testimony before the U.S. House Committee on Science, Space, and Technology: "How do we ensure a steady flow of talent for STEM while also responding to the larger need for a workforce and citizenry with knowledge and skills to address emerging challenges and opportunities? We can only do this by expanding that pool of talent, tapping into the vast well of women, minorities, and persons with disabilities currently underrepresented in STEM."

We must acknowledge that our efforts thus far have fallen short of what is truly necessary: systemic change that transforms institutions—not just individuals. We must tackle the issue where it is most oppressive: deeply ingrained institutional systems. Through AAAS's SEA (STEM Equity Achievement) Change program, institutions of higher education commit to a self-reflection process with the aim of disentangling themselves from practices of the past that made inequities possible—indeed, almost inevitable. The program incentivizes institutions' alignment with SEA Change principles by publicly recognizing them for their commitment to and creation of sustainable systemic change through self-assessment. "It's a transformative national vision," said Paula Rayman, a sociologist at the University of Massachusetts, Lowell, who chairs the SEA Change advisory board. Now more than ever, we must embrace transformative national visions over piecemeal, individual-focused interventions.

The rapid response of AAAS and *Science* to the COVID-19 pandemic also has been notable.

Science and its family of journals provide credible, evidence-based information, share the latest research, and disseminate up-to-the-minute, science-informed news coverage. Our editorial team continues to deliver seminal papers showing how the structure of the coronavirus informs vaccine development and how the virus bonds to human cells, exploring the beginning of new therapies, and examining how the public health system and social distancing can mitigate the spread of COVID-19. Journalists at *Science* are covering the science and responses to the pandemic around the world, often highlighting aspects that are picked up by mainstream news outlets.

Among AAAS programs focusing on COVID-19 is SciLine, which connects journalists with vetted scientific experts. SciLine began 2020, its second full year of operation, with plans that included two "boot camps"—one to help journalists understand the science behind key electoral issues and another for reporters covering adolescent health. In the space of about 2 weeks, the SciLine team quickly pivoted to organizing a series of online media briefings for journalists on COVID-19 and developed a resources webpage to provide ready-to-use quotes from scientists.

SciLine also enlisted Margaret Hamburg, past president of AAAS, former commissioner of the U.S. Food and Drug Administration, and former health commissioner of New York City. On one day in April, from her home office, Hamburg conducted interviews with 20 local TV stations and 13 local radio stations across the country. In doing so, she reached nearly 1.6 million people in 17 states with timely, accurate, and authoritative scientific and public health information.

AAAS's Science and Technology Policy Fellowships program has adopted virtual platforms to connect with current policy fellows and to open doors for the incoming cohort of scientists and engineers. The program, which places some 280 scientists and engineers in federal agencies and congressional offices each year, places a high value on in-person professional development and networking, switching to virtual convenings at a time when the need for fellows to provide scientific and technical advice is greater than ever. The team recently conducted thousands of interviews virtually for the incoming class and is preparing to host a virtual orientation in the fall.

"#ShutDownSTEM is just one step toward following through on fulfilling the hope for positive change in our community and across society."

Sudip S. Parikh,
AAAS CEO

For the team that produces the AAAS Annual Meeting, moving from in-person to virtual convenings has been top of mind. After announcing that the February 2021 meeting would be a virtual event, the team has not looked back, rolling out an entirely new format for scientific symposia that serves as a model for other organizations.

The Science for Seminaries program, part of our Dialogue on Science, Ethics, and Religion, which assists theologically diverse seminaries to integrate science into their required courses of study, also adjusted the focus of one of its recent discussions, shifting it to an examination of the science behind the spread of COVID-19 when audiences sing or speak in loud voices, such as at worship services.

These initiatives display the amazing number of influential audiences we reach as a scientific society. Researchers, policy-makers, journalists, science communicators, students, seminarians—all of these groups play critical roles in our mission to advance science and serve society.

I have not stepped foot in AAAS headquarters in downtown Washington since 13 March—which means I have now spent more time leading AAAS from my basement or my kids' playroom than I have from my office. Even so, I am more confident than ever in the vitality of our mission and in those working daily to execute that mission. We will be a force for science, a force for good, and a support for one another. Our programs, publications, and advocacy are critical to a better and more just world, and what we do during this time will define a generation.

My 31 January *Science* editorial called on us all to rise to the challenges of our time to ensure that the next generation has the opportunity to rise to theirs. On so many levels, those words ring truer today.

Screeners needed for journalism awards

Scientists from the United States and abroad are needed to review the scientific accuracy of entries in the prestigious AAAS Kavli Science Journalism Awards competition. The screening sessions in late August and September will be online this year, opening them to participation by scientists beyond the Washington, D.C., area. We need additional screeners with expertise in virology, epidemiology, and public health. If you can volunteer, please contact Emily Hughes at ehughes@aaas.org.

Published by AAAS

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



ECOLOGY

Easier polar plunge for penguins

In polar regions affected by a warming climate, it is increasingly important to understand what effect yearly variations in sea ice have on animal foraging and reproductive success. Watanabe *et al.* instrumented 175 Adélie penguins—a sentinel species for Southern Ocean ecosystems—with activity monitors and video cameras during four Antarctic field seasons with variable amounts of ice. During ice-free years, penguins traveled more by swimming rather than walking, lowering their energetic cost per unit distance traveled. Forage area increased in ice-free years, and more krill were captured per unit dive time, increasing foraging efficiency, growth rates, and breeding success. Adélie penguin populations in the continental Antarctic region are thus likely to grow in the coming decades as sea ice declines. —BB *Sci. Adv.* 10.1126/sciadv.aba4828 (2020).

Sea-ice loss in the Antarctic has led to greater success in foraging and breeding for Adélie penguins.

CREDITS (FROM TOP): NATURE PICTURE LIBRARY/ALAMY STOCK PHOTO; GONZÁLEZ-RUBIO *ET AL.*

CHROMATIN SEQUENCING

Primary architecture of chromatin fibers

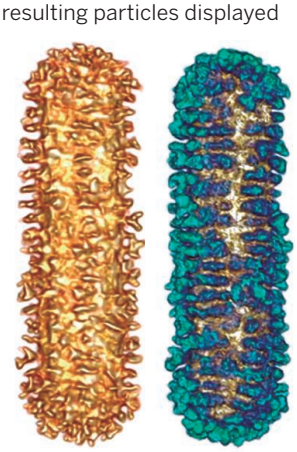
The organization of chromosomal DNA, including the positioning of nucleosomes and nucleosome-free regions harboring regulatory proteins along single chromatin fibers, is fundamental to genome function. However, most sequencing methods cannot elucidate this organization at the nucleotide level. Stergachis *et al*. present an approach, Fiber-seq, that maps chromatin fibers onto the underlying DNA template using methyltransferases to create a kind of stencil in fly and human cells. This method identifies chromatin structure at nearly a single-molecule level and can monitor the position of nucleosomes. Using Fiber-seq, the authors identify how regulatory DNA activation is related to nucleosome positioning and DNA variation. —LMZ

Science, this issue p. 1449

PLASMONIC MATERIALS

Groovy chiral gold particles

Although plasmonic optical activity can arise from chiral assemblies of gold and silver nanoparticles, there are few examples of gold nanoparticles with intrinsic chirality and high optical activity. González-Rubio *et al*. show that morphological chirality can be induced during the seeded growth of gold nanoparticles, particularly for highly anisotropic nanorods. Chiral additives as cosurfactants formed helical micelles that directed the seeded growth to create grooves that maintained a chiral morphology. The resulting particles displayed high-intensity circular dichroism with anisotropy factors near 0.2 at near-infrared wavelengths. —PDS



Electron tomography reconstructions of a gold nanorod grown with a chiral surfactant

Science, this issue p. 1472

SOLAR PHYSICS

The flow of plasma inside the Sun

The activity of the Sun, including the occurrence of sunspots, is driven by magnetic fields that originate from the motion of charged plasma beneath the surface. Helioseismology uses acoustic oscillations to probe the Sun's interior, analogous to seismology's use of earthquakes to investigate Earth's interior. Gizon *et al*. analyzed helioseismology data from 1996 to 2019, covering two 11-year solar cycles.

Published by AAAS

They measured the latitudinal and radial flow of plasma as a function of depth within the Sun and how it varies with time. The results support magnetic flux-transport dynamo models, which can explain the distribution of sunspots over each solar cycle. —KTS

Science, this issue p. 1469

QUANTUM OPTICS

Metalens-array–based quantum source

Spontaneous down-conversion is an exotic optical process in a nonlinear crystal in which a high-energy photon splits into two lower-energy photons that are quantum mechanically entangled. These entangled pairs are valuable commodities for quantum information processing and quantum communications. Because the experimental setup is usually performed with bulk optical components, there is a need to decrease the size scale for application. Li *et al*. combined an array of specialized metalenses with a nonlinear crystal and show that the scale of the process can be shrunk substantially. The approach should prove useful for developing miniaturized integrated quantum optical technologies. —ISO

Science, this issue p. 1487

CORONAVIRUS

Who and what next?

The coronavirus 2019 (COVID-19) pandemic has brought tighter restrictions on the daily lives of millions of people, but we do not yet understand what measures are the most effective. Zhang *et al*. modeled virus transmission in Wuhan, China, in February 2020, investigating the effects of interventions ranging from patient management to social isolation. Age-mixing patterns were estimated by contact surveys conducted in Wuhan and Shanghai at the beginning of February 2020. Once people reduced their average daily contacts from 14 to 20 down to 2, transmission rapidly fell below the epidemic threshold. The model also showed that preemptive school closures helped to reduce transmission, although alone they would not prevent a COVID-19 outbreak. Limiting human mixing to within households appeared to be the most effective measure. —CA

Science, this issue p. 1481

CORONAVIRUS

A wrench in the works of COVID-19

Understanding the inner workings of the virus that causes coronavirus disease 2019 (COVID-19) may help us to disrupt it. Yin *et al*. focused on the viral polymerase essential for replicating viral RNA. They determined a structure of the polymerase bound to RNA and to the drug remdesivir. Remdesivir mimics an RNA nucleotide building block and is covalently linked to the replicating RNA, which blocks further synthesis of RNA. The structure provides a template for designing improved therapeutics against the viral polymerase. —VV

Science, this issue p. 1499

MACROPHAGES

Sex, age, and the macrophage

Peritoneal macrophages are known to contribute to pathology of peritonitis, endometriosis, and cancer, but the effects of age and sex on the biology of these cells are not well understood. Bain *et al.* show that replenishment of peritoneal macrophages from the bone marrow is much higher in male mice than in female mice. This disparity results in marked sexual dimorphisms in the phenotypic identity of peritoneal macrophages, including differential expression of a receptor involved in response to bacterial infection. These findings provide insight into the effects of sex and age on peritoneal inflammation and infection. —CNF

Sci. Immunol. **5**, eabc4466 (2020).

IN OTHER JOURNALS

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



Climate change is expected to alter the ecology of high-altitude plants in the Andes, like this Andean blueberry (*Vaccinium floribundum*).

PHOTO: FRANZ X AVER CC-BY-SA

MOUNTAIN ECOLOGY

Seeds of change

Dispersal of seeds is a vital process for static plants. High-altitude plant communities are especially vulnerable to the effects of changing climate, and it is important to understand how climate affects their dispersal. Tovar *et al.* investigated the seed dispersal strategies of plants living close to the summits of the tropical Andes. Wind-dispersed species were common throughout; plants with unspecialized, gravity-dispersed seeds were more frequent at sites with low minimum temperatures; and plants with animal-dispersed seeds were more frequent at sites with milder climates. As climate warms in the future, the spectrum of dispersal strategies and the composition of these plant communities may change accordingly. —AMS *J. Ecol.* 10.1111/1365-2745.13416 (2020).

GENE REGULATION

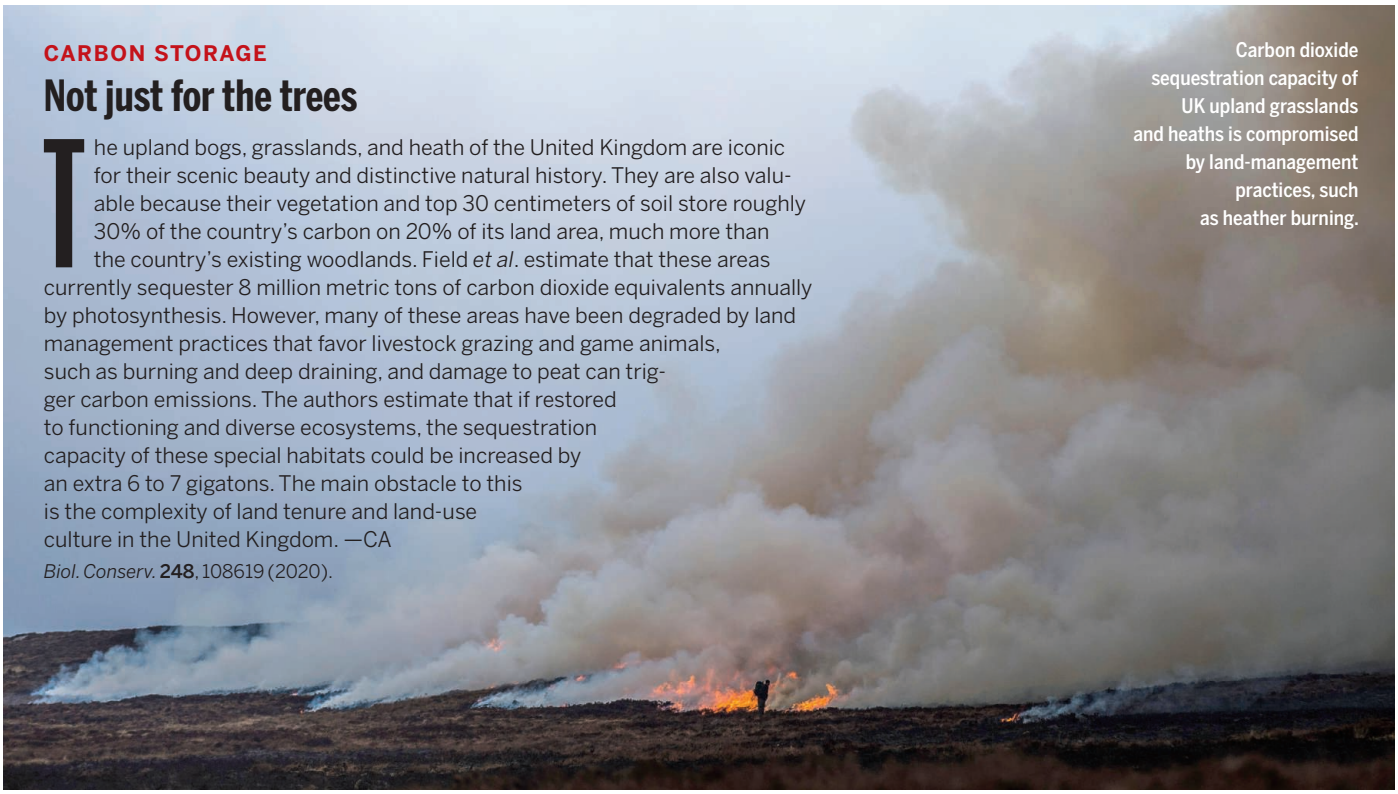
Pol II kicks out histone variant

When we think of histone proteins, we generally think of H2A, H2B, H3, and H4 because they are the histones that package DNA within our cells. However, other histone variants perform important regulatory and structural roles. Ranjan *et al.* used single-particle tracking of fluorescently tagged proteins to examine histone variants in yeast cells in vivo and discovered that RNA polymerase II (Pol II) itself evicts H2A.Z from chromatin. In addition, the kinase Kin28/Cdk7, which phosphorylates serine-5 of heptapeptide repeats in the

Published by AAAS

CARBON STORAGE

Not just for the trees



Carbon dioxide sequestration capacity of UK upland grasslands and heaths is compromised by land-management practices, such as heather burning.

The upland bogs, grasslands, and heath of the United Kingdom are iconic for their scenic beauty and distinctive natural history. They are also valuable because their vegetation and top 30 centimeters of soil store roughly 30% of the country's carbon on 20% of its land area, much more than the country's existing woodlands. Field *et al*. estimate that these areas currently sequester 8 million metric tons of carbon dioxide equivalents annually by photosynthesis. However, many of these areas have been degraded by land management practices that favor livestock grazing and game animals, such as burning and deep draining, and damage to peat can trigger carbon emissions. The authors estimate that if restored to functioning and diverse ecosystems, the sequestration capacity of these special habitats could be increased by an extra 6 to 7 gigatons. The main obstacle to this is the complexity of land tenure and land-use culture in the United Kingdom. —CA

Biol. Conserv. **248**, 108619 (2020).

carboxy-terminal domain of the Pol II subunit Rpb1, is required for this eviction. These findings indicate a general mechanism coupling eukaryotic transcription to erasure of the H2A.Z epigenetic signal. —BAP

eLife **9**, e55667 (2020).

ULTRAFAST METHODS

Avoiding multiphoton artifacts

Ultrafast pump-probe crystallography experiments with x-ray free-electron lasers have now resolved structural changes in several light-sensitive proteins. A major experimental question has been how to set up the pump laser to increase the occupancy of excited states in the crystal but not cause multiphoton excitation. Grünbein *et al*. studied how incident laser pulses would be refracted and reflected by the crystal transport medium and the crystals themselves. They found that processes that would attenuate laser power are minimal, whereas refraction increases power density in a portion of the jet. Power-density titrations and assessment of the spectroscopic properties of each experimental setup are therefore essential to avoid multiphoton artifacts at high laser power. —MAF

Nat. Methods 10.1038/s41592-020-0847-3 (2020).

ICE SHELVES

Loss of shelf

An unusually warm and prolonged inflow of warm ocean water intruded beneath Antarctica's Filchner Ronne Ice Shelf in 2017. Ryan *et al*. describe a 4-year-long time series of ocean temperature measurements that documents the surge and its unusual persistence. They discuss the hydrographic properties of the event and possible forcing mechanisms, suggesting that it was caused by anomalous summer sea ice melting. Much of the thinning of ice shelves like the Filchner Ronne is caused by warm water intrusions. Similar intrusions beneath other ice shelves will affect the future stability of the Antarctic Ice Sheet, a pressing concern in our warming climate. —HJS

Geophys. Res. Lett. 10.1029/2020GL088119 (2020).

STRUCTURAL BIOLOGY

The double punch of perchlorate

To make thyroid hormones, which are critical in development, iodine ions (I^-) must be transported into thyroid cells. This is achieved by the sodium ion (Na^+)/I^- symporter (NIS), which couples the energetically unfavorable import of one I^- with the energetically favorable import of two Na^+. NIS also transports the environmental pollutant perchlorate ($ClO4^-$), but in this case couples transport of one $ClO4^-$ with that of one Na^+. Based on theory and experiments, Llorente-Esteban *et al*. show that $ClO4^-$ binds not only to the transport site but also to a nontransport site. Binding at the second site precludes binding of one of the Na^+. Therefore, $ClO4^-$ not only inhibits I^- transport by direct competition, it also reduces the driving force for I^- transport by binding at the nontransport site. —VV

Nat. Struct. Mol. Biol. **27**, 533 (2020).

PHYSICS

Charmed excitations

Unlike the general purpose detectors used by the ATLAS and CMS experiments at the Large Hadron Collider, the LHCb experiment has equipment designed specifically to study particles that contain the so-called beauty and charm quarks. By tracking processes that result in the formation of a charmed lambda baryon and a negatively charged kaon, Aaij *et al*. (the LHCb Collaboration) have now uncovered evidence for at least two previously unobserved baryons. These baryons have been identified as the excited states of a Xi baryon consisting of a charm quark and two other quarks. The findings may help to constrain theories on how quarks bind together to form baryons. —JS

Phys. Rev. Lett. **124**, 222001 (2020).

PHOTO: REUBEN TABNER / ALAMY STOCK PHOTO

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CORONAVIRUS

Reducing airborne transmission

Respiratory infectious diseases are thought to be transmitted mostly through contact with surfaces that are contaminated with virus-laden droplets produced from coughs and sneezes of infected individuals, but this is not the only transmission route. When people speak and breathe, they produce tiny droplets called aerosols, as well as larger droplets, which evaporate and become buoyant in airflows. In a Perspective, Prather *et al*. discuss the accumulating evidence that virus-containing aerosols could lead to airborne transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19). Given the lengthy presymptomatic phase and asymptomatic infections, the authors argue that wearing well-fitted masks, especially in enclosed indoor spaces, is important to help prevent SARS-CoV-2 transmission. —GKA

Science, this issue p. 1422

GENOMIC DUPLICATIONS

The fate of genes after duplication

Gene duplication within an organism is a relatively common event during evolution. However, we cannot predict the fate of the duplicated genes: Will they be lost, evolve, or overlap in function within an organismal lineage or species? Kuzmin *et al*. explored the fate of duplicated gene function within the yeast *Saccharomyces cerevisiae* (see the Perspective by Ehrenreich). They examined how experimental deletions of one or two duplicated genes (paralogs) affected yeast fitness and were able to determine which genes have likely evolved new essential functions and which retained functional overlap, a condition the authors refer to as entanglement. On the basis of these results, they propose how entanglement affects the evolutionary trajectory of gene duplications. —LMZ

Science, this issue p. 1446; see also p. 1424

DEVELOPMENTAL BIOLOGY

Guiding regeneration

Many adult organisms can regenerate neural circuits after injury. However, it is not clear which guidance mechanisms operate to promote axon path finding in the adult. Scimone *et al*. addressed this question by investigating regeneration of the planarian visual system (see the Perspective by Roberts-Galbraith). Distinct muscle cell populations were found in close association with photoreceptor axons that, together with a neuron class, facilitated visual system assembly after diverse injuries or eye transplantations. These cells exhibited features similar to embryonic guidepost cells and were specified independently of eyes in precise locations by the action of adult positional information cues. Absence of these guidepost-like cells was associated with defective neuronal wiring in regeneration. —BAP

Science, this issue p. 1447; see also p. 1428

NEUROSCIENCE

The adaptive human frontal cortex

Flexibly switching between different tasks is a fundamental human cognitive ability that allows us to make selective use of only the information needed for a given decision. Minxha *et al*. used single-neuron recordings from patients to understand how the human brain retrieves memories on demand when needed for making a decision and how retrieved memories are dynamically routed in the brain from the temporal to the frontal lobe. When memory was not needed, only medial frontal cortex neural activity was correlated with the task. However, when outcome choices required memory retrieval, frontal cortex neurons were phase-locked to field potentials recorded in the medial temporal lobe. Therefore, depending on demands of the task, neurons in different regions can flexibly engage and disengage their activity patterns. —PRS

Science, this issue p. 1448

STRUCTURAL BIOLOGY

Engaging the nucleosome

Cell identity is defined by gene expression patterns that are established through the binding of specific transcription factors. However, nucleosomal units limit access of transcription factors to specific DNA motifs within the mammalian genome. To study how transcription factors bind such chromatinized, nucleosome-embedded motifs, Michael *et al*. focused on the pluripotency factors OCT4 and SOX2. They systematically quantified the relative affinities of these factors at different motif positions throughout the nucleosome, enabling structure determination of OCT4-SOX2–bound nucleosomes by cryo–electron microscopy. OCT4 and SOX2 bound cooperatively to strengthen DNA-binding affinity and resulted in DNA distortions that destabilized the nucleosome. This analysis reveals position-dependent binding modes that were validated in vivo, providing insights on how transcription factors read out chromatinized motifs. —BAP

Science, this issue p. 1460

STRUCTURAL BIOLOGY

A blueprint to understand cohesin

Cohesin is a multiprotein complex that entraps sister chromatids for chromosome segregation and regulates transcription by extruding DNA loops to shape DNA organization. Shi *et al*. determined the structure of human cohesin bound to the protein NIBPL, which helps load cohesin onto DNA, and DNA at medium resolution by cryo–electron microscopy. Two adenosine triphosphatase domains play a key role in cohesin function. The structure explains how NIBPL and DNA synergistically activate these domains and gives insight into how DNA is trapped by cohesin. —VV.

Science, this issue p. 1454

SPECTROSCOPY

DNA circular dichroism in gas phase

Circular dichroism spectroscopy is widely used to distinguish between nonidentical mirror-image molecules. The technique relies on differential absorption of left versus right circularly polarized light and therefore tends to require solution-phase samples for adequate sensitivity. Daly *et al*. now report gas-phase circular dichroism spectra of DNA oligonucleotides based on detection of photodetached electrons rather than transmitted light (see the Perspective by Barran). The salient spectral features matched those in solution. Pairing the technique with mass spectrometry enables prior mass selection of particular molecules for analysis. —JSY

Science, this issue p. 1465; see also p. 1426

Published by AAAS

EXOPLANETS

A nearby multiplanet system

Exoplanets can interact gravitationally with other objects orbiting the same star, affecting their evolution and stability. Studying these effects requires locating systems with multiple planets. Monitoring the nearby red dwarf star GJ 887, Jeffers *et al*. detected periodic radial velocity signals, indicating the presence of two planets on orbits with periods of about 9 and 22 days and a further candidate planet (see the Perspective by Davies). The inclinations of the orbits are unknown, so only minimum masses could be determined, but those were consistent with both planets being super-Earths—more massive than Earth but less than Neptune. This system is only 3.3 parsecs from the Sun, which should facilitate follow-up with other techniques. —KTS

Science, this issue p. 1477; see also p. 1432

PHOTOSYNTHESIS

Pairs of peaks stabilize output power

A counterintuitive feature of photosynthesis is that the primary pigments involved in absorbing light—for example, chlorophyll a and b in plants—do not all absorb right at the peak of the spectrum but instead are offset from the peak and each other. Arp *et al*. formulated a network model that explains how using pigments with this absorption-peak pattern can mitigate internal and external fluctuations in energy transfer, minimizing noise in output power (see the Perspective by Duffy). The model accurately reproduces absorption peaks for three diverse photosynthetic systems from different spectral environments. Such a mechanism may provide an underlying robustness to biological photosynthetic processes that can be further tuned and tweaked to adapt to longer-scale fluctuations in light intensity. —MAF

Science, this issue p. 1490; see also p. 1427

DOG GENOMICS

Sled dog arctic adaptations go far back

Dogs have been used for sledding in the Arctic as far back as ~9500 years ago. However, the relationships among the earliest sled dogs, other dog populations, and wolves are unknown. Sinding *et al*. sequenced an ancient sled dog, 10 modern sled dogs, and an ancient wolf and analyzed their genetic relationships with other modern dogs. This analysis indicates that sled dogs represent an ancient lineage going back at least 9500 years and that wolves bred with the ancestors of sled dogs and precontact American dogs. However, gene flow between sled dogs and wolves likely stopped before ~9500 years ago. —LMZ

Science, this issue p. 1495

CALCIUM SIGNALING

Lighting a spark from the lysosome

Calcium ion (Ca^{2+}) sparks in vascular smooth muscle cells are critical for blood vessel dilation and for regulating blood pressure. Thakore *et al*. showed that these Ca^{2+} sparks were initiated by Ca^{2+} released from lysosomes through the cation channel TRPML1 (see the Focus by Nieves-Cintron *et al*.). Lysosomes in contractile vascular smooth muscle cells were localized in close proximity to this ion channel. Vascular smooth muscle cells from mice deficient in TRPML1 did not generate Ca^{2+} sparks, and these mice were spontaneously hypertensive. —WW

Sci. Signal. **13**, eaba1015, eabc0993 (2020).

CANCER

Intuitive $CD8^+$ T cells sense DNA

Stimulator of interferon genes (STING) agonism is an area of active exploration for cancer immunotherapy. Li *et al*. examined a DNA-sensing cascade called cGAS-STING in antitumor $CD8^+$ T cells. They observed dampened STING activity in $CD8^+$ T cells from patients with cancer or mice implanted with tumors. By contrast, STING signaling in transferred T cells supported a stem-like memory phenotype, which is known to be beneficial for responses to immunotherapy. Cytosolic DNA was enriched in activated T cells, and STING agonism improved the efficacy of adoptive cell therapy in multiple mouse models. These results highlight that $CD8^+$ T cell DNA sensing could be exploited for therapeutic benefit in immunotherapy. —LP

Sci. Transl. Med. **12**, eaay9013 (2020).

Published by AAAS

RESEARCH ARTICLE SUMMARY

GENOMIC DUPLICATIONS

Exploring whole-genome duplicate gene retention with complex genetic interaction analysis

Elena Kuzmin*, Benjamin VanderSluis*, Alex N. Nguyen Ba, Wen Wang, Elizabeth N. Koch, Matej Usaj, Anton Khmelinskii, Mojca Mattiazzi Usaj, Jolanda van Leeuwen, Oren Kraus, Amy Tresenrider, Michael Pryszlak, Ming-Che Hu, Brenda Varriano, Michael Costanzo, Michael Knop, Alan Moses, Chad L. Myers†, Brenda J. Andrews†, Charles Boone†

INTRODUCTION: Whole-genome duplication (WGD) events are pervasive in eukaryotes, shaping the genomes of simple single-celled organisms, such as yeast, as well as those of more complex metazoans, including humans. Most duplicated genes are eliminated after WGD because one copy accumulates deleterious mutations, leading to its loss. However, a significant proportion of duplicates persists, and factors that result in duplicate gene retention are poorly understood but critical for understanding the evolutionary forces that shape genomes.

RATIONALE: Quantifying the functional divergence of paralog pairs is of particular interest because of the strong selection against functional redundancy. Negative genetic interactions identify functional relationships between genes and provide a means to directly capture the functional relationship between duplicated genes. Genetic interactions occur when the phenotype associated with a combination of mutations in two or more different genes deviates from the expected combined effect of the individual mutations. A negative genetic interaction refers to a combination of mutations that generates a stronger fitness defect than expected, such as synthetic lethality. Here, we used systematic analysis of digenic and trigenic interaction profiles to assess the functional relationship of retained duplicated genes.

ON OUR WEBSITE

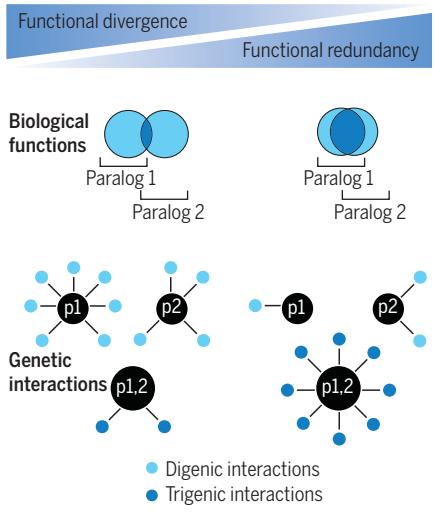
Read the full article at https://dx.doi.org/10.1126/science.aaz5667

RESULTS: To map both digenic and trigenic interactions of duplicated genes, we profiled query strains carrying single-deletion mutations and the corresponding double-deletion mutations for 240 different dispensable paralog pairs originating from the yeast WGD event. In total, we tested ~550,000 double and ~260,000 triple mutants for genetic interactions, and identified ~4700 negative digenic interactions and ~2500 negative trigenic interactions. We quantified the trigenic interaction fraction, defined as the ratio of negative trigenic interactions to the total number of interactions associated with the paralog pair. The distribution of the resulting trigenic interaction fractions was distinctly bimodal, with two-thirds of paralogs exhibiting a low trigenic interaction fraction (diverged paralogs) and one-third showing a high trigenic interaction fraction (functionally redundant paralogs). Paralogs with a high trigenic interaction fraction showed a relatively low asymmetry in their number of digenic interactions, low rates of protein sequence divergence, and a negative digenic interaction within the gene pair.

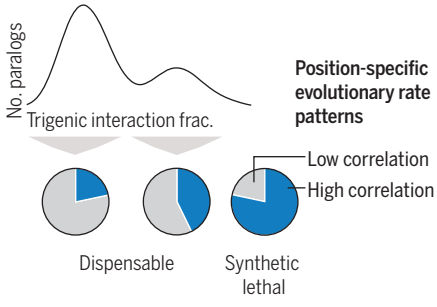
We correlated position-specific evolutionary rate patterns between paralogs to assess constraints acting on their evolutionary trajectories. Paralogs with a high trigenic interaction fraction showed more correlated evolutionary rate patterns and thus were more evolutionarily constrained than paralogs with a low trigenic interaction fraction. Computational simulations that modeled duplicate gene evolution revealed that as the extent of the initial entanglement (overlap of functions) of paralogs increased, so did the range of functional redundancy at steady state. Thus, the bimodal distribution of the trigenic interaction fraction may reflect that some paralogs diverged, primarily evolving distinct functions without redundancy, while others converged to an evolutionary steady state with substantial redundancy due to their structural and functional entanglement.

CONCLUSION: We propose that the evolutionary fate of a duplicated gene is dictated by an interplay of structural and functional entanglement. Paralog pairs with high levels of entanglement are more likely to revert to a singleton state. In contrast, unconstrained paralogs will tend to partition their functions and adopt divergent roles. Intermediately entangled paralog pairs may partition or expand nonoverlapping functions while also retaining some common, overlapping functions, such that they can both adopt paralog-specific roles and maintain functional redundancy at an evolutionary steady state. ■

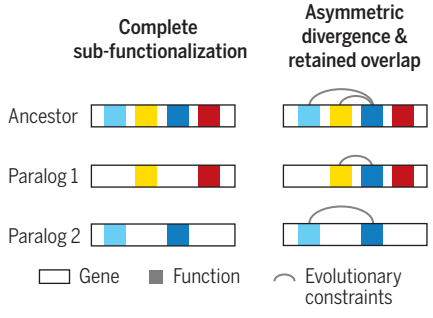
Digenic and trigenic interactions reveal functional relationships of duplicated genes



Evolutionary constraints acting on duplicated genes



Structural and functional entanglement model of paralog divergence



Complex genetic interaction analysis of duplicated genes. The trigenic interaction fraction, which incorporates digenic and trigenic interactions, captures the functional relationship of duplicated genes and follows a bimodal distribution. Paralogs with a high trigenic interaction fraction are under evolutionary constraints reflecting their structural and functional entanglement.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: charlie.boone@utoronto.ca (C.B.); brenda.andrews@utoronto.ca (B.J.A.); chadm@umn.edu (C.L.M.)
Cite this article as E. Kuzmin *et al.*, *Science* **368**, eaaz5667 (2020). DOI: 10.1126/science.aaz5667

RESEARCH ARTICLE

GENOMIC DUPLICATIONS

Exploring whole-genome duplicate gene retention with complex genetic interaction analysis

Elena Kuzmin[1,2]*†, Benjamin VanderSluis[3]*, Alex N. Nguyen Ba[4,5]‡, Wen Wang[3], Elizabeth N. Koch[3], Matej Usaj[1], Anton Khmelinskii[6]§, Mojca Mattiazzi Usaj[1], Jolanda van Leeuwen[1]¶, Oren Kraus[1,2], Amy Tresenrider[7], Michael Pryszlak[1,2], Ming-Che Hu[1], Brenda Varriano[1], Michael Costanzo[1], Michael Knop[6,8], Alan Moses[4,5,9], Chad L. Myers[3]#, Brenda J. Andrews[1,2]#, Charles Boone[1,2]#

Whole-genome duplication has played a central role in the genome evolution of many organisms, including the human genome. Most duplicated genes are eliminated, and factors that influence the retention of persisting duplicates remain poorly understood. We describe a systematic complex genetic interaction analysis with yeast paralogs derived from the whole-genome duplication event. Mapping of digenic interactions for a deletion mutant of each paralog, and of trigenic interactions for the double mutant, provides insight into their roles and a quantitative measure of their functional redundancy. Trigenic interaction analysis distinguishes two classes of paralogs: a more functionally divergent subset and another that retained more functional overlap. Gene feature analysis and modeling suggest that evolutionary trajectories of duplicated genes are dictated by combined functional and structural entanglement factors.

Most eukaryotic genomes, including the human genome, contain a substantial fraction of duplicated genes (*1*–*7*). Gene duplication is generated by two main mechanisms: segmental duplication (small-scale duplication) due to error-prone DNA replication, and simultaneous duplication of all genomic segments (whole-genome duplication) due to a variety of polyploidy events (*3*, *8*). Gene duplication provides a source of new genes (*9*), and duplicate retention may lead to the development of specialized functional modules involving paralogous proteins through "subfunctionalization," which promotes biological complexity (*10*). Nonetheless, after duplication most paralogs are eliminated from the genome because the functional redundancy of duplicated genes is evolutionarily unstable, as one gene copy may accumulate intrinsically deleterious mutations and be removed from the genome by selection (*11*). However, a significant fraction of duplicates has been retained over the course of evolution. Thus, understanding the molecular mechanisms that underlie duplicate gene retention may provide insights into the evolutionary forces that shape genomes.

About 100 million years ago, the budding yeast *Saccharomyces cerevisiae* arose from a whole-genome duplication (WGD) event, and, after a massive gene loss, retained 551 duplicate gene pairs (*6*, *7*, *12*). Quantifying the functional divergence of each paralog pair is of particular interest because of the strong selection against functional redundancy. Paralog functional divergence has been estimated by the rates of evolutionary divergence of coding and regulatory regions (*6*, *12*–*14*), Gene Ontology (GO) semantic distance (*15*–*17*), metabolic flux analysis (*18*, *19*), similarity of gene expression profiles (*3*, *10*, *20*–*23*), changes in the encoded protein abundance of one sister upon perturbation of another (*24*), and analysis of similarity of partners within the protein-protein interaction network (*25*).

Genetic interaction analysis provides a powerful means to directly capture the functional relationship between duplicated genes. Genetic interactions occur when a combination of mutations in different genes results in an unexpected phenotype, deviating from a model based on the integration of the individual mutant phenotypes. A negative genetic interaction occurs when a combination of mutations leads to a more severe fitness defect than expected (*26*). An extreme example of a negative digenic interaction is synthetic lethality, which occurs when two mutations, neither of which is lethal on its own, combine to generate an inviable double-mutant phenotype (*27*, *28*). Negative genetic interactions often occur between genes that impinge on a common essential function. A positive genetic interaction occurs when a combination of mutations results in a phenotype that is less severe than expected from the phenotypes associated with the single mutants. For example, digenic suppression is observed when a double mutant exhibits a greater fitness than the sickest single mutant (*29*). A previous global digenic interaction network in yeast identified ~550,000 negative and ~350,000 positive genetic interactions (*30*). The profile of genetic interactions for a specific query gene provides a quantitative measure of function (*31*, *32*), and a network based on genetic interaction profile similarity reveals a hierarchy of functional modules, including pathways and complexes, bioprocesses, and cellular compartments (*30*).

A systematic analysis of the digenic interactions between duplicated gene pairs in yeast revealed functional redundancy, whereby ~30% of pairs interacted (relative to ~3.6% for random gene pairs) (*33*–*37*). However, this observation may indicate that the overall contribution of duplicate retention to the ability of an organism to tolerate mutations, known as mutational robustness, is relatively low because total functional compensation is only observed for a minor fraction of duplicates (*38*). Mechanisms that drive duplicate retention may be influenced by gene dosage effects or by functional divergence through asymmetric evolution (*33*). Retention of duplicates may result from subfunctionalization, such that duplicates degenerate in some of their function differentially and result in a pair of genes that function fully as the single ancestral copy; this outcome is postulated by the duplication-degeneration-complementation model (*39*). Evidence of functional partitioning of ancestral functions includes but is not limited to biochemical function (*40*), gene expression regulatory elements (*22*), and subcellular localization patterns (*41*).

Here, we expanded upon the use of genetic interaction profiles to capture the functional relationship of duplicated genes. We compared trigenic interactions for double-mutant query strains deleted for both members of nonessential paralog pairs to the corresponding digenic interaction profiles for each single-mutant sister gene, and we quantified a spectrum of functional redundancy among paralogs. A correlative analysis of the gene features suggests that the evolutionary trajectories of retained duplicated genes can be driven by genes encoding functionally and structurally

[1]Donnelly Centre, University of Toronto, Toronto, Ontario M5S 3E1, Canada. [2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 3E1, Canada. [3]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA. [4]Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada. [5]Center for Analysis of Evolution and Function, University of Toronto, Toronto, Ontario, Canada. [6]Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance, 69120 Heidelberg, Germany. [7]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. [8]Cell Morphogenesis and Signal Transduction, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany. [9]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada.
*These authors contributed equally to this work.
†Present address: Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, Quebec H3A 1A3, Canada.
‡Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA.
§Present address: Institute of Molecular Biology, Mainz, Germany.
¶Present address: Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland.
#Corresponding author. Email: charlie.boone@utoronto.ca (C.B.); brenda.andrews@utoronto.ca (B.J.A.); chadm@umn.edu (C.L.M.)

constrained proteins, which we refer to as "entangled."

Mapping trigenic interactions for duplicated genes

We constructed 240 double-mutant query yeast strains, each deleted for a pair of nonessential WGD paralog genes (tables S1 to S4). These are dispensable paralog pairs and represent 44% (240/551) of unique WGD paralog pairs (*12*), a number of which were not included because the pair was either essential or refractory to double-mutant query strain construction (*42*). Using colony size as a proxy for cell fitness, we measured the growth phenotypes of the set of 240 double-mutant query strains and the corresponding 480 single mutants (table S5) (*42*), which correlated well with previous large-scale measurements of single-mutant fitness (Pearson correlation coefficient $r = 0.51$, $P = 3 \times 10^{-30}$) and double-mutant fitness ($r = 0.72$, $P = 2 \times 10^{-23}$) (*30*) (fig. S1, A to D).

We used the set of single- and double-mutant query strains to score digenic and trigenic interactions, respectively, using trigenic synthetic genetic array analysis (τ-SGA) (Fig. 1). We crossed the queries to a diagnostic array of nonessential gene deletion mutants and essential gene mutants, carrying temperature-sensitive alleles, which span all major cellular processes and cover ~1200 genes representing ~20% of the yeast genome (*37*). In total, we examined 537,911 double and 256,861 triple mutants for genetic interactions. Query strains deleted for an individual paralog gene were screened for digenic interactions, and double-mutant query strains deleted for both paralogs were screened for trigenic interactions in two replicate screens with four colonies per screen (fig. S2, A to C) (*37*). Negative and positive interactions were quantified for digenic and trigenic interactions (*30*, *32*), which were determined from validation of trigenic interactions of the *CLN1-CLN2* double-mutant query, as previously described (*37*). This resulted in an estimated recall (sensitivity) of ~60% and a precision of ~75% (*37*). Additionally, we used replicate screens to independently estimate the false discovery rate (FDR) as a function of recall, which resulted in a consistent estimate of >75% precision (<25% FDR) (fig. S2D).

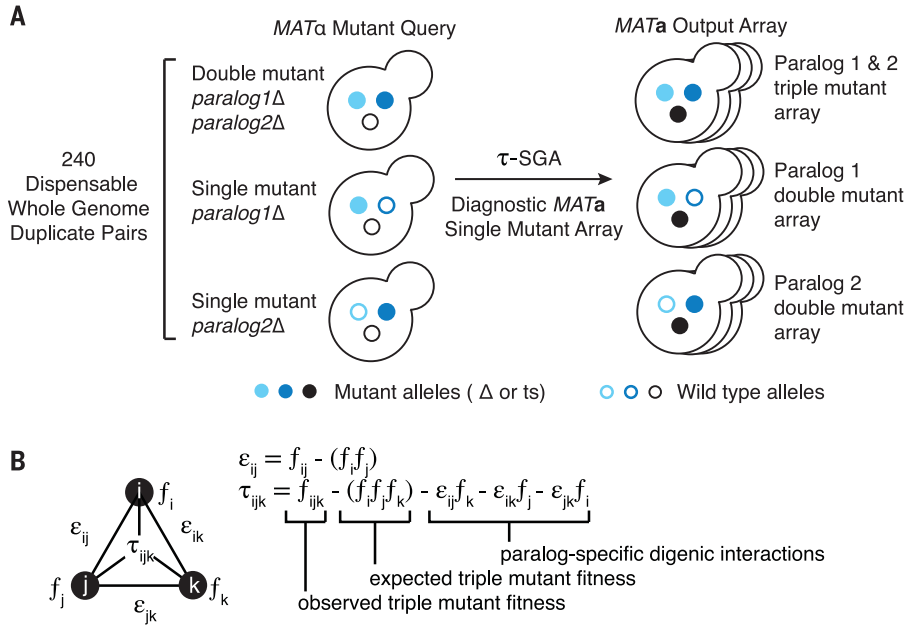


Fig. 1. Trigenic synthetic genetic array (τ-SGA) analysis for paralogs. (**A**) An illustration of the τ-SGA experimental approach in which a query set of 240 dispensable paralog pairs originating from whole-genome duplication (WGD) in yeast was screened for trigenic interactions. Three types of screens were carried out in parallel, whereby triple-mutant fitness was estimated by crossing a double-mutant query strain deleted for both paralogs (light and dark blue solid circles) into a diagnostic array of single mutants (black solid circles) (*37*). After induction of meiosis in heterozygous triple mutants, sequential replica-pinning steps are used to select haploid triple-mutant progeny. Single-mutant control query strains are screened in parallel to estimate paralog-specific double-mutant fitness. (**B**) We used the τ-SGA scoring method to identify trigenic interactions quantitatively by combining double- and triple-mutant fitness estimates derived from colony size measurements (*37*). The digenic interaction score is expressed as ε_{ij}, where f_{ij} is the observed double-mutant fitness, and the expected double-mutant fitness is expressed as the product of single-mutant fitness estimates $f_i f_j$. The trigenic interaction score is expressed as τ_{ijk}, where f_{ijk} is the observed triple-mutant fitness and $f_i f_j f_k$ is the triple-mutant fitness expectation expressed as the product of three single-mutant fitness estimates.

This analysis identified 4650 negative and 2547 positive digenic interactions, as well as 2466 negative and 2091 positive trigenic interactions (tables S1, S2, and S4). About one-third of negative and one-fourth of positive trigenic interactions were of the "novel" class, identifying connections that were not observed in their corresponding digenic interaction network (Fig. 2). Indeed, for 129 paralog pairs, the corresponding single genes displayed only sparse digenic interaction profiles on the global genetic interaction network (*30*) (table S6). However, they exhibited more novel trigenic interactions (~40% negative and ~31% positive trigenic interactions) than average, indicating that paralog pair trigenic interactions expanded the known global genetic interaction network. The remaining two-thirds of negative and three-fourths of positive trigenic interactions overlapped a previously identified digenic interaction and thus represent a "modified" class of trigenic interactions (Fig. 2). Paralogs with more negative trigenic interactions than digenic interactions showed predominantly novel and modified negative trigenic interactions, which overlapped exclusively with negative digenic interactions indicating functional redundancy (fig. S3A).

Genetic interaction profiles highlight functional divergence of duplicated genes

The functional relationship between paralogs should be captured by their negative genetic interactions. Consistent with previous observations (*32*, *33*), we observed that duplicates showed fewer negative digenic interactions than singletons (Wilcoxon rank-sum test, $P = 6 \times 10^{-8}$; fig. S3B), which is suggestive of a general trend in which paralogs retain functional redundancy. We predicted that highly divergent gene pairs should exhibit a relatively high number of paralog-specific negative digenic interactions (Fig. 3). In contrast, functionally overlapping paralogs should be biased toward trigenic interactions and should display few paralog-specific interactions (Fig. 3). To assess this possibility, we computed the trigenic interaction fraction, defined as the ratio of negative trigenic interactions to the total number of all negative interactions (digenic and trigenic) associated with the paralog pair (Fig. 3).

From this analysis, we observed that double mutants involving duplicates showed a range of trigenic and digenic interaction degrees. This indicated that complex genetic interactions may reveal their functional redundancy (fig. S3B). For example, *SKI7-HBS1* is a paralog pair that showed a relatively low trigenic interaction fraction (Fig. 3). Although the *SKI7* and *HBS1* gene products both recognize stalled ribosomes and initiate mRNA degradation, they do so in a different manner (*43*). Ribosomes that stall upon encountering an in-frame stop

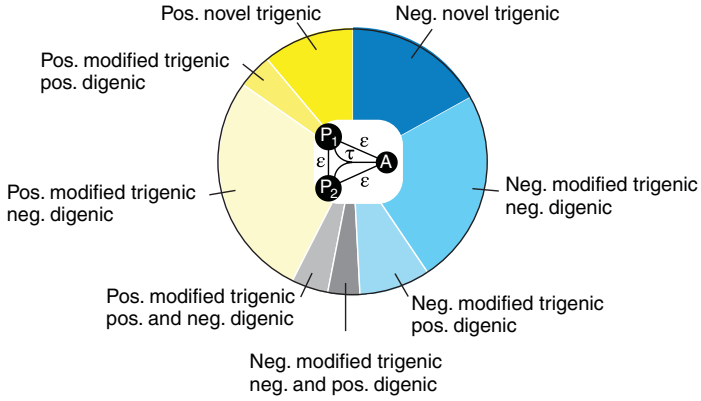


Fig. 2. Distribution of different types of trigenic interactions for paralogs. The different types of trigenic interactions for all paralogs are compared in a pie chart. Negative [(τ or ε) < −0.08, P < 0.05] and positive [(τ or ε) > 0.08, P < 0.05] genetic interactions are shown in blue and yellow, respectively. A trigenic interaction between a double-mutant query and the array strain is called "novel" (dark blue/dark yellow) if there is no significant digenic interaction between either single-mutant control query and the array strain or between the query gene pair. Trigenic interactions that overlap with one or more negative or positive digenic interactions are called "modified" and are further classified by the type of digenic interaction. All trigenic interactions of double-mutant query strains that show a negative or a positive digenic interaction between the members of a query gene pair (P_1-P_2) ($|\varepsilon|$ > 0.08, P < 0.05) are considered "modified." Interactions may be further classified by digenic interactions (if any) between a single-mutant query control strain and the array strain (P_1 and/or P_2-A negative, P_1 and/or P_2-A positive). Digenic interactions of the same sign are in medium blue/yellow, digenic interactions of the opposite sign are in light blue/yellow, and mixed positive and negative digenic interactions are depicted in different shades of gray according to whether their modified trigenic interactions are positive or negative.

codon at the 3′ end of a transcript are recognized by Ski7, which in turn recruits the RNA exosome (*44*). In contrast, ribosomes stalled within the coding region of a transcript, possibly because of an unusual structural conformation or damage in the mRNA, are recognized by Hbs1, which initiates mRNA cleavage in an RNA exosome–independent manner (*45*). Our model (Fig. 3) predicts that the digenic interaction profiles of the paralogs should reflect their independent function. Indeed, *SKI7* showed digenic interactions with genes involved in mRNA 3′ end protection and 5′-3′ mRNA decay, such as *PAT1* and *LSM1*, whereas *HBS1* interacted with numerous genes involved in ribosome biogenesis and recycling. Thus, the low trigenic interaction fraction of the *SKI7-HBS1* gene pair (Fig. 3) appears to reflect the functional divergence of these paralogs (*39*, *40*, *43*).

Conversely, the *MRS3-MRS4* duplicate pair showed a high trigenic interaction fraction (~0.85) (Fig. 3). These paralogs are members of the eukaryotic-specific mitochondrial carrier family, which transports compounds, including nucleotides, amino acids, carboxylates, small inorganic ions, and vitamins, across the inner mitochondrial membrane linking the cytosolic and mitochondrial biochemical pathways (*46*). *MRS3* and *MRS4* encode highly similar mitochondrial carrier proteins with high affinity for Fe^{2+}, which they transport across the inner mitochondrial membrane (*47*). The corresponding vertebrate homolog, mitoferrin, is involved in erythropoiesis by maintaining mitochondrial iron homeostasis (*48*). The *MRS3-MRS4* trigenic interactions involved genes related to cell redox homeostasis, such as *GRX3*, *TSA1*, and *TRX3*. The processes that regulate Fe^{2+} homeostasis are important components of the cellular defense mechanism against oxidative damage. Indeed, the *MRS3-MRS4* trigenic interactions were also enriched for genes involved in DNA replication and repair, including genes encoding members of the Rad51-Rad57 complex (*RAD51*, *RAD54*, *RAD55*, *RAD57*), the Rad5-Rad6-Rad18 complex (*RAD5*), the DNA replication factor C complex (*CTF8*, *CTF18*), the MRX complex (*MRE11*, *XRS2*), the Holliday junction resolvase complex (*MUS81*, *MMS4*), and the nucleotide-excision repair factor 3 complex (*TFB1*, *SSL1*) (*49*). Together these examples illustrate how the trigenic interaction fraction may reflect the degree of functional overlap of paralogs.

Distribution of trigenic interaction fraction among retained paralog pairs

In total, we measured the trigenic interaction fraction for 161 paralog pairs that showed at least six total trigenic or digenic interactions (table S7). These paralog pairs displayed a range of trigenic interaction fraction values (Fig. 4A) that showed a distinctly bimodal distribution, with 114 paralog pairs exhibiting a relatively low trigenic interaction fraction (below 0.4) while a smaller subset of 47 paralogs displayed a higher trigenic interaction fraction (above 0.4). This finding suggests that comparison of digenic and trigenic interaction profiles is an effective way to differentiate paralog pairs and can provide insight into the extent of functional overlap between members of a given duplicated gene pair (Fig. 4A and table S7). This distribution of trigenic interaction did not differ for subsets of duplicated genes that originated by distinct mechanisms, such as ohnologs, which originated from WGD, or homeologs, which originated from hybridization between species (*42*, *50*). We confirmed that the genetic interaction profiles we generated are robust to array size, in that we observed a significant correlation for the trigenic interactions obtained from the diagnostic array and the genome-wide array, which were derived from screening 11 double mutants with their single-mutant control query strains in two replicates each (fig. S4). The strength of correlation between replicates was not affected appreciably with decreasing stringency for either digenic or trigenic interactions (fig. S2, A to C), indicating that our conclusions are not dependent on an interaction score threshold. The remaining 79 paralog pairs, representing about one-third of all screened pairs, were characterized by sparse genetic interaction profiles. These paralogs span a diverse set of biological processes and tend to belong to larger gene families (one-sided t test, P = 0.04; table S9), which may confer higher-order redundancy and reduce second- and third-order genetic interactions.

Trigenic interaction fraction, digenic interaction profiles, and paralog properties

The trigenic interaction fraction of paralog pairs was also associated with several fundamental physiological and evolutionary properties (Fig. 4B and tables S5, S7, and S10). Paralogs with a high trigenic interaction fraction tend to exhibit low asymmetry with respect to their number of digenic interactions. The asymmetry score measures whether the digenic profile of one paralog is composed of substantially more interactions than its corresponding duplicate, and it correlates with divergent evolution of paralog gene sequences (*33*). Consistent with this observation, a high trigenic interaction fraction also correlated with relatively low rates of protein sequence divergence (Fig. 4B). We also observed a high trigenic interaction fraction in paralogs whose double mutant showed a negative digenic interaction, which is often associated with functionally related genes (Fig. 4B) (*30*, *37*).

Conversely, paralogs with a low trigenic interaction fraction often showed a high asymmetric score for their digenic degrees (Fig. 4B). Consistent with our hypothesis that a low trigenic fraction is indicative of functional divergence (Figs. 3 and 4A), a high asymmetry score may reflect that one paralog has evolved a specialized role. For example, a paralog displaying relatively few digenic interactions under standard conditions may only be expressed and functional under a different environmental condition or during a specialized developmental program. Indeed, in the case of four

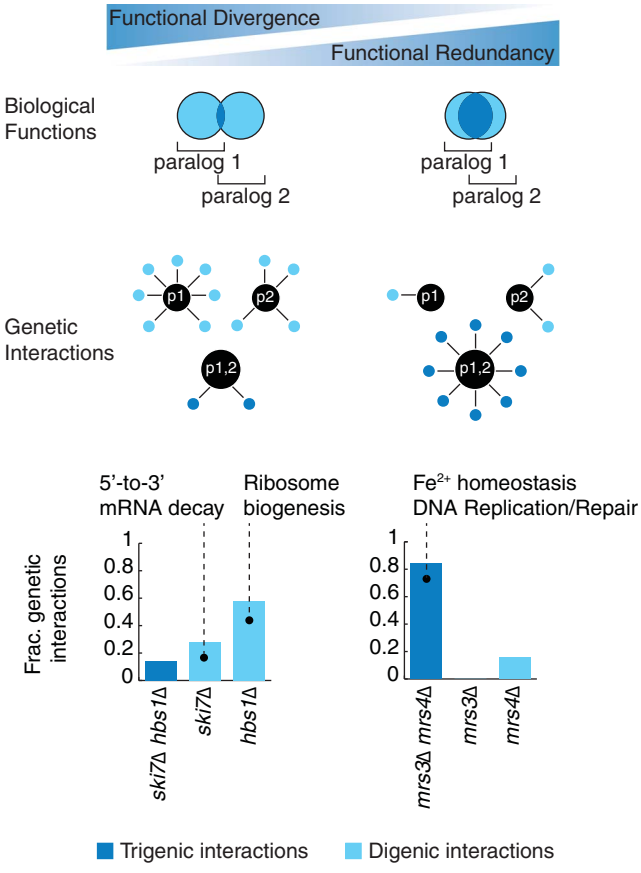


Fig. 3. Mapping the functional relationship of paralogs through their digenic and trigenic interactions. This schematic depicts highly divergent paralogs with little functional overlap and functionally redundant paralogs with extensive functional overlap, as represented by the Venn diagrams. Diverged paralogs are predicted to exhibit many digenic interactions, indicative of their paralog-specific functions and few trigenic interactions, whereas functionally redundant paralogs are expected to show sparse digenic interactions and numerous trigenic interactions, indicative of their functional overlap. Divergent paralogs such as *SKI7-HBS1* behave in a manner consistent with the expectation and display fewer trigenic than digenic interactions. However, functionally redundant paralogs such as *MRS3-MRS4* display a higher fraction of trigenic interactions with a corresponding drop in the fraction of paralog-specific digenic interactions. The fraction of different types of genetic interactions is illustrated using bar graphs. The fraction of total genetic interactions attributed to the trigenic interactions associated with a *par1*Δ *par2*Δ double-mutant query, deleted for both paralogs, is shown as a dark blue bar; the fraction of digenic interactions associated with each paralog single-deletion mutant, *par1*Δ or *par2*Δ, is shown as a light blue bar.

asymmetric paralog pairs, expression of the low-degree sister was induced during sporulation, a meiotic developmental program that cells enter in response to a low-carbon and low-nitrogen environment. Specifically, *GIS1* in paralog pair *GIS1-RPH1*, *HES1* in *HES1-KES1*, *ECI1* in *ECI1-DCI1*, and *DON1* in *DON1-CUE5* showed higher meiotic gene expression (table S11) (*51*). For another seven asymmetric paralog pairs, the low-degree paralog was required for filamentous growth (*52*), and for another 15 asymmetric pairs, the gene expression of the low-degree sister was induced under glucose starvation (table S11) (*53*). In total, 22 of 63 asymmetric pairs (~35%) have a paralog that may have been retained for function in a different condition or during a specialized program (table S11).

The retention of gene duplicates may also be related to gene dosage (*54*). For example, some duplicates appear to be maintained as a result of selection for high levels of expression, as in the case of metabolic genes that provide high enzymatic flux (*18*). WGD events also enable maintenance of stoichiometric balance of protein complex members, consistent with the high rate of duplication among components of the ribosome (*35*). Dosage duplicates are associated with a severe fitness defect when either one of the paralogs is deleted, as well as a greater digenic interaction profile similarity than other duplicates, indicating that they have retained functional redundancy (*33*, *38*). Because dosage duplicates should have substantial functional redundancy, we reasoned that they might also tend to have a higher trigenic interaction fraction. Although the paralogs with a greater digenic interaction profile similarity also tend to have a higher trigenic interaction fraction (Fig. 4C), potential dosage duplicates (*33*) were found within both the high and low trigenic interaction fraction distributions (table S10). Thus, the functional redundancy associated with the subset of paralogs with a higher trigenic interaction fraction is probably not driven solely by dosage mechanisms.

The evolution of regions important for protein localization may lead to differential subcellular localization of some paralogous proteins and may play a role in their retention by enabling specific paralog functions in distinct parts of the cell (*41*, *55*). However, because on average duplicated genes do not appear to evolve a relocalization more frequently than singletons, this may not represent a major mechanism driving paralog retention (*56*). We observed that paralogs with different subcellular localization patterns tended to show a modestly higher trigenic interaction fraction than those with the same subcellular localization patterns (Wilcoxon rank-sum test, $P < 0.05$), which suggests that differentially localized paralogs may retain some functional overlap.

Defining paralog function in terms of trigenic interactions

To characterize the roles of paralog pairs with overlapping functions, we mapped their trigenic interactions onto the global digenic interaction profile similarity network (*30*) (Fig. 5A). Using this approach, trigenic interactions associated with paralog pairs that have a relatively high trigenic interaction fraction can be examined for enrichment within defined bioprocesses. For example, the *sbe2*Δ *sbe22*Δ double mutant showed few paralog-specific interactions and numerous trigenic interactions, with a trigenic interaction fraction of ~0.69 (table S7). The proteins encoded by the *SBE2-SBE22* pair share ~51% amino acid sequence identity and have relatively well-characterized functions in the transport of cell wall components from the Golgi to the cell surface (*57*). The *SBE2-SBE22* negative trigenic interactions involved genes enriched for GO annotations related to vesicle-mediated transport and cell wall organization (Fig. 5B) and include interactions with an ARF-like small guanosine triphosphatase (GTPase) secretion pathway (*ARL1*, *ARL3*, *SYS1*, *YPT6*), the exocyst (*SEC4*), and the chitin biosynthesis pathway (*CHS5*, *CHS6*).

Previously, we showed that trigenic interactions, like digenic interactions, are functionally informative and are enriched among genes annotated to the same biological process (*37*). Mapping paralog trigenic interaction profiles onto the global digenic interaction profile similarity network enables the functional annotation of previously uncharacterized paralogs. For example, *ECM13-YJR115W* showed a trigenic interaction fraction of 0.77 with 12 negative digenic and 40 negative trigenic interactions, suggesting high functional redundancy (fig. S5A and table S7). These paralogs are fungal-specific and share 39% identity and 70% similarity at the amino acid sequence level. The genes that constitute the *ECM13-YJR115W* trigenic interaction profile were positioned mainly within the mitosis and DNA replication and repair clusters on the global digenic profile network (Fig. 5C). The *ECM13-YJR115* trigenic interaction profile is correlated to the digenic interaction profiles

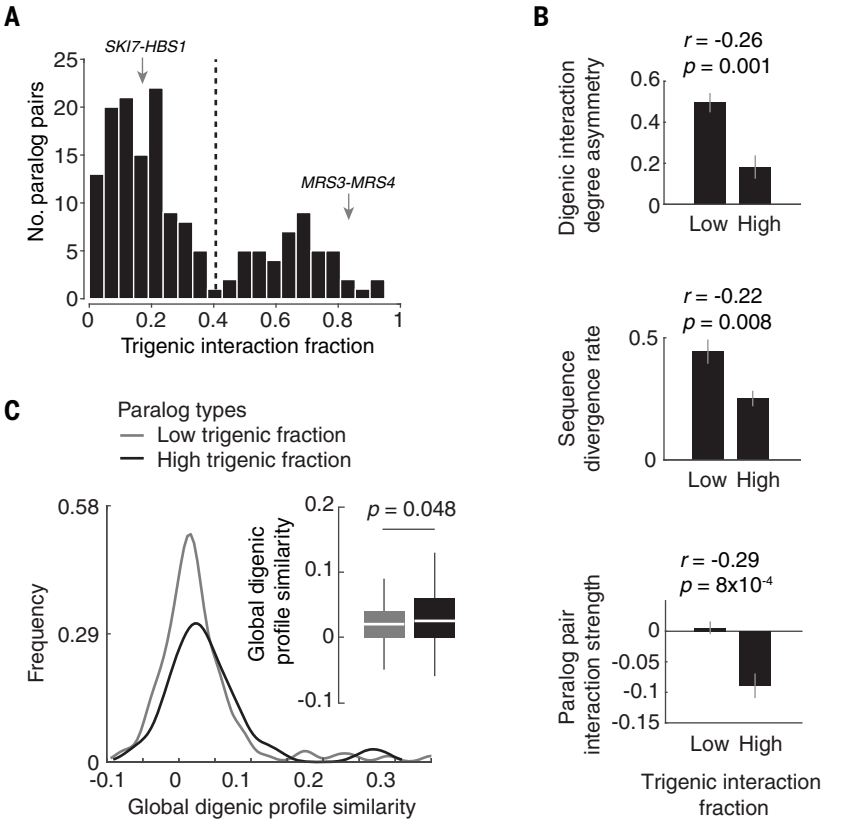


Fig. 4. The trigenic interaction fraction correlates with fundamental physiological and evolutionary properties. (**A**) Negative trigenic interaction fraction distribution of screened paralogs, (τ or ε) < −0.08, P < 0.05; paralogs with at least six trigenic or digenic interactions in one of the screens are considered. Representative examples of paralogs with a low (*SKI7-HBS1*) and high (*MRS3-MRS4*) trigenic interaction fraction are marked with an arrow. (**B**) Physiological and evolutionary properties for paralogs characterized by varying fraction of trigenic interactions. Spearman correlation coefficients (r) and associated P values are measures of the strength of the correlation between the trigenic interaction fraction and the three features being examined: digenic interaction degree asymmetry, sequence divergence rate, and paralog pair interaction strength. The correlation was measured on the entire dataset. The bar plots serve to visualize the trend, in which a trigenic interaction fraction cutoff of 0.4 was used to identify paralogs with low and high trigenic interaction fraction, based on negative interactions (τ or ε) < −0.08, P < 0.05. Means of specified features are shown; error bars denote SEM. (**C**) The distribution of global digenic profile correlation similarity (*30*) was compared for paralogs with high and low trigenic interaction fraction. A trigenic interaction fraction cutoff of 0.4 was used as in (B). Box plots serve to summarize the data of the distributions. Analyses are restricted to paralogs with at least six total trigenic or digenic interactions in one of the screens. Significance was assessed using one-tailed Wilcoxon rank-sum test.

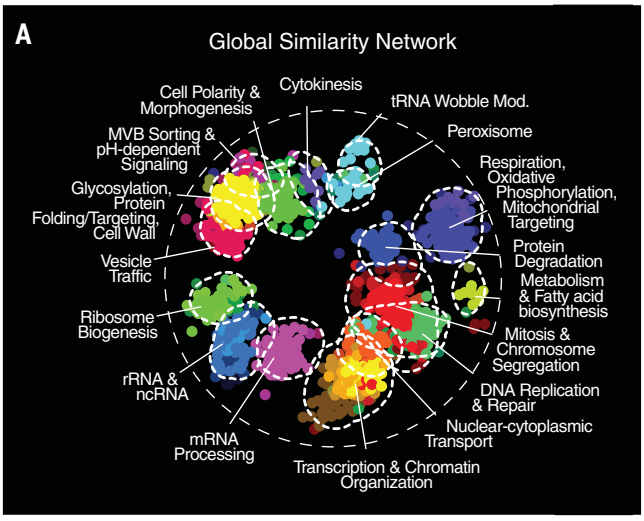
of *CCT4* and *CCT5*, which encode members of the cytosolic chaperonin Cct ring complex that participates in the assembly of tubulin, and *RBL2*, which is involved in microtubule morphogenesis, suggesting a possible microtubule-related role for this paralog pair (*30*) (fig. S5B). The *ecm13*Δ *yjr115w*Δ double mutant was specifically sensitive to benomyl, a microtubule-disrupting agent (fig. S5, C and D), and showed a delay in spindle nucleation and polymerization (fig. S5E). Moreover, the Ecm13 protein interacts with Dad2 and Dad4, which are components of the Dam1 complex that links kinetochores to microtubules to facilitate chromosome segregation (*58*). Consistent with a role in spindle function and chromosome segregation, we found that both GFP-Ecm13 and GFP-Yjr115w showed a distinct nuclear localization (fig. S5F). Another poorly characterized pair is *STB6-STB2*, which encodes proteins that bind the *SIN3* transcriptional repressor and may impinge on the MTC (maintenance of telomere capping) pathway, suggesting a role in aromatic amino acid permease secretion (*30*) (fig. S5, G to J).

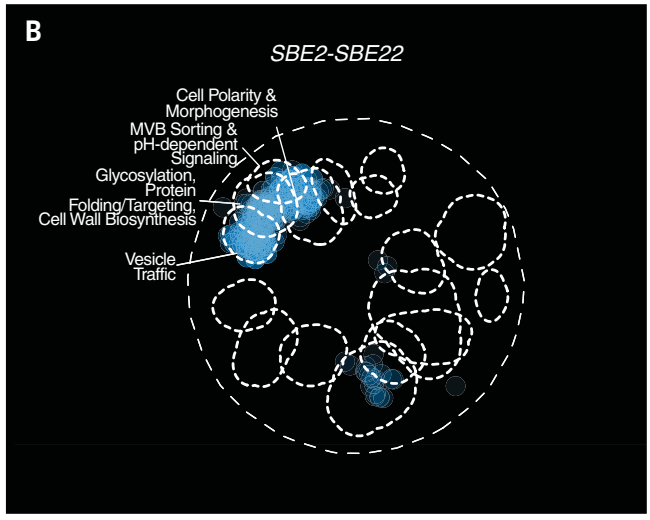
Correlation analysis of position-specific evolutionary rate patterns between paralogs

To further explore the factors that drive paralog retention and produce a spectrum of functional redundancy, we considered the relationship among divergent evolution, structure, and function. A negative correlation of trigenic interaction fraction with measures of divergent evolution may be indicative of the constraints that govern this process (Fig. 4B). Amino acid conservation in specific positions indicates evolutionary constraints on amino acid residues that are important for protein function, including those involved in oligomerization, protein-protein interactions, and protein-substrate interactions (*59*). We therefore used the correlation of position-specific evolutionary rate patterns between paralogs as a measure of evolutionary divergence of the sequence constraints (*42*). We reasoned that if both paralogs share constraints on specific residues of protein domains during evolution and evolve similarly after the WGD, then they would be expected to be more similar to each other than to the pre-WGD species. On the other hand, if paralogs do not share constraints on specific residues of protein domains during evolution and evolve differently after the WGD, then they would be less similar to each other than to the pre-WGD species. We therefore designated a set of pairs with a high correlation of position-specific evolutionary rate patterns as evolutionarily constrained pairs for which the correlation of rates between extant sisters was greater than or equal to both the correlations between each sister and the pre-WGD homolog. Paralog pairs with little correlation in their position-specific evolutionary rate patterns have different sequence constraints and are therefore likely to have different structure and function (Fig. 6A and table S12).

For example, the proteins encoded by the *MRS3-MRS4* gene pair show a high correlation in their position-specific evolutionary rate patterns and also have high protein sequence identity (~76%) (Fig. 6B and fig. S6A). Although both *MRS3* and *MRS4* encode high-affinity mitochondrial Fe^{2+} transporters, the *MRS3* protein displays an additional Cu^{2+} transport function (*60*). The general mitochondrial carrier function is structurally constrained because all mitochondrial carrier proteins have a tripartite structure with three similar segments, each of which is ~100 amino acids in length and forms two membrane-spanning α helices (fig. S7A) (*61*). Interestingly, these proteins appear to have properties that are favorable for retention after duplication because multiple mitochondrial carrier protein genes were retained after the yeast WGD. Indeed, of the 35 different mitochondrial carrier protein genes in the yeast genome, 10 are encoded by five WGD paralog pairs (*62*).

Our study interrogated four mitochondrial carrier protein WGD pairs: *MRS3-MRS4*, *YIA6-YEA6*, *YMC1-YMC2*, and *ODC1-ODC2*. Like the *MRS3-MRS4* paralog pair, *YIA6* and *YEA6* are connected by a negative digenic interaction and show a high trigenic interaction fraction, which is consistent with a retained functional





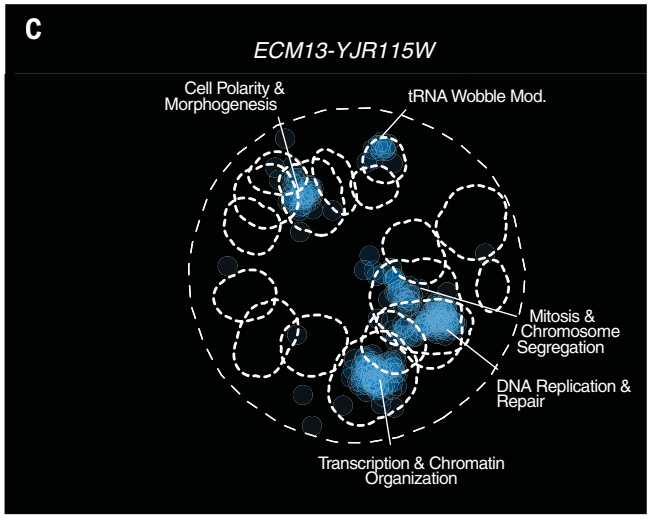


Fig. 5. The trigenic interaction fraction reveals the functional divergence of duplicated genes and illuminates gene function. (**A** to **C**) SAFE (*70*) analysis (A) was used to visualize regions of the global digenic interaction profile similarity network (*30*) that were enriched for genes in the trigenic interaction profiles of the paralog pairs *SBE2-SBE22* (B) and *ECM13-YJR115W* (C). Blue indicates the enrichment related to negative trigenic interactions ($\tau < -0.08$, $P < 0.05$).

redundancy. In contrast, both the *YMC1-YMC2* and *ODC1-ODC2* paralog pairs display relatively few trigenic interactions. Indeed, the *YMC1-YMC2* pair displays a low correlation of position-specific evolutionary rate patterns and thus a low level of functional redundancy (table S12). On the other hand, *ODC1-ODC2* displays a relatively high correlation of position-specific evolutionary rate patterns (table S12); however, their functional overlap may be masked by the presence of other mitochondrial carrier proteins because they belong to a large gene family with multiple paralog members, expanded also by small-scale duplications (table S9). Moreover, we combined our genetic analysis with literature curated data to map a genetic network underlying numerous mitochondrial carrier protein genes, and the *YMC1-YMC2* and *ODC1-ODC2* paralogs appear to form a highly connected subnetwork (fig. S8), which suggests that these two paralog pairs display a more complex functional redundancy.

In contrast to *MRS3-MRS4*, the *SKI7-HBS1* proteins show a relatively low correlation in their position-specific evolutionary rate patterns and thus show a low sequence identity (~26%) with an asymmetric rate of sequence divergence, whereby Ski7 appears to be diverging faster than Hbs1 (table S10). Detailed inspection of these proteins revealed that the Hbs1 protein resembles the pre-WGD homolog, whereas Ski7 has adopted a more divergent fate, which suggests that its evolved role was not constrained structurally by the pre-WGD ancestor (Fig. 6B and fig. S6B). Despite retaining the EF-Tu GTP-binding domain (PF00009), it is present in a highly divergent form. Ski7 also lost critical sequences encoding an Hbs1-like N-terminal motif (PF8938) and sequences encoding the EF-Tu C-terminal domain (PF03143), highlighting the evolutionary divergence of Ski7 from Hsb1 and the pre-WGD homolog (Fig. 6B and fig. S6B).

By calculating the level of correlation of position-specific evolutionary rate patterns between members of the duplicate pair in relation to the pre-WGD homolog (Fig. 6A), we assessed the evolutionary constraints acting on paralogs (Fig. 6C). We found that paralogs with a high trigenic interaction fraction were composed of a significantly higher number of paralogs with correlated evolutionary rate patterns and thus were more evolutionarily constrained than those characterized by a low trigenic interaction fraction (Fisher exact test, $P = 0.01$) (Fig. 6C). Moreover, paralogs that show a synthetic lethal genetic interaction are considered highly functionally redundant. This subset of essential paralogs shows a higher correlation in their position-specific evolutionary rate patterns, which suggests that these genes are more evolutionarily constrained than those within the subset of highly redundant nonessential paralogs displaying a high trigenic interaction fraction (Fisher exact test, $P = 0.02$) (Fig. 6C and table S13). Hence, it appears that some paralogs are highly structurally constrained or "entangled," which limits their divergence leading to the maintenance of functional overlap, but presumably within a context that also enables the evolution of novel functions.

Modeling simulates the divergent evolution of paralogs with retained functional redundancy

We also explored functional redundancy and paralog retention using in silico modeling in an attempt to test two hypotheses. Under the first hypothesis, the retained nonessential paralog pairs with a high trigenic interaction fraction—and thus a high functional overlap—are inherently unstable over evolutionary time and would eventually diverge completely, losing any common functionality. Under the alternative hypothesis, retained paralogs may converge to an evolutionary steady state, in which paralogs with retained functional overlap cannot segregate certain functional regions without a fitness cost. We computationally generated "genes" of fixed length, in which regions of random length were assigned responsibility for a function, and a random number of such functions was generated for each gene, such that these functional regions were allowed to overlap. Then we duplicated each gene and began introducing random "degenerative" mutations, which would render the affected paralog unable to perform any function associated with the mutated region (Fig. 7A). We discarded any lineage as unfit when any one of the original functions could not be carried out by at least one sister, and continued simulating mutations until the pair reached steady state and could tolerate no additional mutations. The extent of overlap of functions in each randomly generated ancestral gene at the start of paralog evolution provides a measure of their initial structural and functional "entanglement," generating a baseline from which we assessed their evolutionary trajectories (*42*).

These simulations revealed that for a large fraction of paralog pairs, the mutation process resulted in a singleton state with only one of the sisters being retained. A sizable fraction of

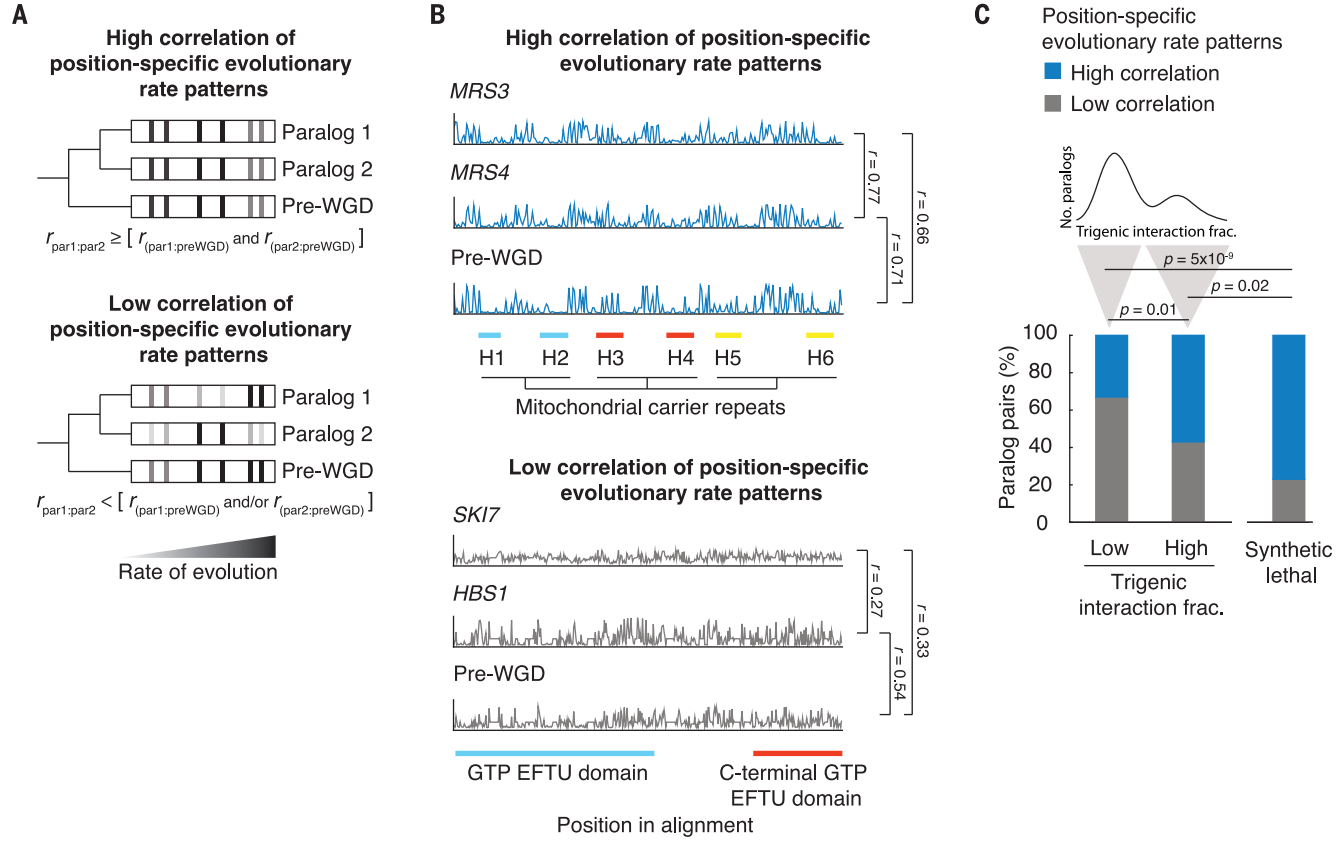


Fig. 6. Evolution of retained overlap due to evolutionary constraints acting on duplicated gene sequences. (**A**) Schematic depiction of the analysis of correlated evolutionary sequence changes across paralog sequences reflecting evolutionary constraints on paralogs. Correlated rates of evolution for specific columns in multiple sequence alignments for the pre-WGD homolog and each paralog are denoted with a gray-to-black gradient, from low to high, respectively. High correlation of position-specific evolutionary rate patterns identifies residues with similar evolutionary constraints. Paralogs with correlated rates ($r_{par1:par2}$) that are greater than or equal to that of each paralog and with the corresponding pre-WGD ($r_{par1:preWGD}$ and $r_{par2:preWGD}$) were designated as having a high correlation of position-specific evolutionary rate pattern, and paralogs with correlated rates ($r_{par1:par2}$) that were less than that of either paralog or both paralogs with the pre-WGD ($r_{par1:preWGD}$ and/or $r_{par2:preWGD}$) were designated as having a low correlation of position-specific evolutionary rate pattern; *r* refers to the Pearson correlation coefficient between the respective sequences. (**B**) Examples of evolutionary rates for positions in the alignments for representative paralogs, showing a high correlation of position-specific evolutionary rate patterns (*MRS3-MRS4*) or a low correlation of position-specific evolutionary rate patterns (*SKI7-HBS1*). The position in the alignment is plotted on the *x* axis; the rate of evolution at a particular position divided by the average rate of evolution for all residues in the given sister paralog is plotted on the *y* axis. The scale of the *y* axis has been fixed for each paralog pair. Pfam domains are annotated. The *MRS3-MRS4* alignment shows three mitochondrial carrier repeats, each composed of two α helices (blue, H1 and H2; red, H3 and H4; yellow, H5 and H6) followed by a characteristic motif Pro-X-[Asp/Glu]-X-X-[Lys/Arg]-X-[Lys/Arg]-(20 to 30 residues)-[Asp/Glu]-Gly-X-X-X-X-[Trp/Tyr/Phe]-[Lys/Arg]-Gly connecting each pair of membrane-spanning domains by a loop. The *SKI7-HBS1* alignment shows GTP EF-Tu (blue) and C-terminal GTP EF-Tu (red) domains. The Hbs1-like N-terminal motif lies outside of the alignment window. (**C**) Fraction of nonessential and essential paralogs that show a high or low correlation of position-specific evolutionary rate patterns. The paralogs with low and high trigenic interaction fraction belong to the part of the distribution shown above; a trigenic interaction fraction cutoff of 0.4 was used, based on negative interactions (τ or ε) < −0.08, $P < 0.05$, and contains the set of paralogs that were used for the correlated evolution analysis. Significance was assessed by Fisher exact test.

simulations, however, ended with paralogs in a stable steady state in which no more mutations could be tolerated in either paralog, while still maintaining viability. Analysis of the simulation results revealed that the particular trajectory followed by a given paralog pair was correlated with the level of functional entanglement. Specifically, paralog pairs that started with the highest levels of entanglement immediately upon duplication were more likely to revert to a singleton state. This suggests that duplicated genes generally cannot tolerate genetic perturbations when they lack functionally independent regions (Fig. 7B and fig. S9A). Among paralog pairs that were retained at steady state, increased entanglement at the point of duplication also led to a broader bias in the functional asymmetry (ratio of functional responsibilities) at steady state. Thus, paralogs diverge asymmetrically when they begin their evolutionary trajectory with a protein sequence containing extensive entanglement (Fig. 7C and fig. S9, B and C). Consistent with this observation from our simulations, paralog evolution can show asymmetric bias with respect to functional redundancy (Fig. 4B).

Our modeling further revealed that as the extent of the initial entanglement of paralogs increased, so did the range of steady-state functional overlap, which is represented by constrained domains at steady state (Fig. 7D and fig. S9D). This suggests that the bimodal distribution of the trigenic interaction fraction (Fig. 4A) may indicate that one subset of WGD paralogs diverged substantially so that each of the sister paralogs has a distinct function, and another subset of retained WGD paralogs reached an evolutionary steady state despite retained functional overlap, perhaps as a result of their structural and functional entanglement (Fig. 7E). For example, in the case of *SKI7-HBS1*, Ski7 diverged from Hbs1 by losing the Hbs1-like N-terminal motif and the EF-Tu C-terminal domain while retaining a highly diverged form of the EF-Tu GTP-binding domain, reflecting a modular, structural, and functional organization of the protein (Fig. 6B). On the other hand, *MRS3* and *MRS4* encode mitochondrial carrier proteins dedicated to the transport of small inorganic ions, and thus their divergence would be predicted to occur in specific residues that modulate ion specificity (Fig. 6B and fig. S7).

We propose that the evolutionary fate of a duplicated gene can be governed by an interplay of structural and functional entanglement (Fig. 7E). If a duplicated gene contains several easily partitioned functions, then it will most likely subfunctionalize; on the other hand, an entangled pair, which is highly restricted structurally and functionally, would have a tendency to revert to a singleton state because one of the genes is predicted to quickly become nonfunctional. However, given multiple functions and an

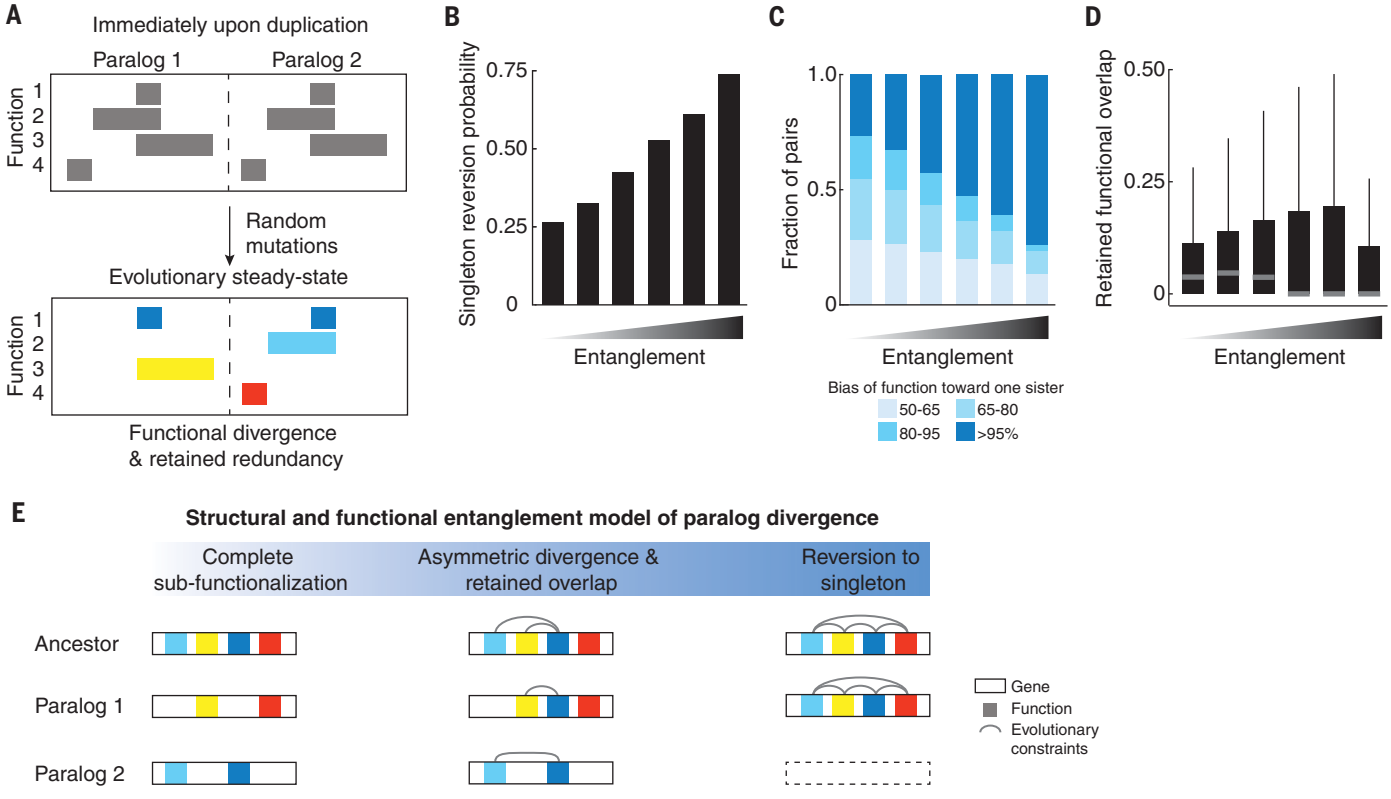


Fig. 7. In silico evolutionary model. (**A**) Schematic depiction of the in silico evolutionary model. The pair evolves through random mutations until it reaches an evolutionarily stable state that can sustain no further mutations without a loss of function. Top, a pair at the start of the evolutionary trajectory; bottom, a pair that achieves a division of labor with a retention of a common function (dark blue blocks), the loss of which is prevented because it would compromise the unique functions of each paralog (yellow, light blue, red). (**B** to **D**) Evolutionary fates of paralogs with functional and structural entanglement. Paralogs were generated to represent a range of overlapping functional domains at the onset of their evolutionary trajectory, and the propensity to assume specific paralog properties was quantified. In each case, the *x* axis represents bins of initial functional overlap as a fraction of "gene" length at the start of the simulations (< 10%, 30%, 50%, 70%, 90%, 100%, respectively); the *y* axis depicts the propensities of paralogs to revert to a singleton state (B), evolve functional asymmetry (C), or retain functional overlap at the evolutionary steady state (D). (**E**) The structural and functional entanglement model of paralog divergence. A pair will evolve by subfunctionalization if it is modular and is composed of partitionable functions (left). A paralog pair that is very structurally and functionally entangled will have a high probability of reversion to a singleton state because one of the sisters will quickly degenerate (right). Paralogs with an intermediate level of entanglement at the time of duplication will tend to partition some and retain some overlapping functions, allowing for specialization of a common activity (center).

intermediate level of entanglement, a gene pair has a chance of partitioning or expanding some nonoverlapping functions while retaining others in common, and remaining evolutionarily stable.

Discussion

The complex genetic interaction network of dispensable WGD paralogs provides insight into the long-standing question of why paralogs with overlapping functions are retained on an evolutionary time scale. By combining both paralog-specific digenic interactions and the paralog pair trigenic interactions in a single metric, the trigenic interaction fraction, we captured the spectrum of the retained functional redundancy of dispensable paralogs.

Definitions of functionally redundant or divergent paralogs using genetic interactions appear to be consistent with classification from protein-protein interaction studies (*25*). For example, in the case where a paralog is deleted and the sister responds by gaining specific protein-protein interactions, then the paralogs should compensate for each other's loss and thus should exhibit a high trigenic interaction fraction. Indeed, four such paralogs were examined in our study and they showed a propensity to exhibit high trigenic interaction fractions. In particular, *NUP53-ASM4* and *OSH6-OSH7* showed high trigenic interaction fractions of 0.49 and 0.80, respectively. On the other hand, some paralogs share protein-protein interactions that are lost for both paralogs when only one sister is deleted, which suggests that although these paralogs may cooperate, they do not fully compensate for each other (*25*). We examined two such known paralogs, *PEX25-PEX27* and *GSY1-GSY2*, and they exhibited low trigenic interaction fractions of 0.19 and 0.25, respectively. Beyond these examples, genetic interaction profiling provides a functional readout and allows assessment of pairs of genes that do not have extensive protein-protein interaction profiles, and therefore it provides a complementary view of functional redundancy.

Our framework enabled us to interrogate how WGD paralog evolution relates to the evolutionary stability of retained common functions and asymmetric functional divergence. By computing the extent of correlated evolution in sister paralogs (Fig. 6), we identified paralogs that show highly correlated position-specific evolutionary rate patterns and thus are under strong evolutionary constraints to retain some of their ancestral function, reflecting their structural and functional entanglement. This was further explored by our in silico model (Fig. 7), which predicts that low levels of entanglement are sufficient to drive asymmetric subfunctionalization, whereas more complex sequence-function relationships with higher structural entanglement can result in fixation of functional redundancy. Indeed, our modeling shows that given some level of moderate structural entanglement and the potential for multifunctionality, a substantial fraction of duplicate pairs converge to a steady state in which they retain functional overlap. This result offers a possible explanation as to the persistence of the functional overlap in paralogs, which is not simply due to paralogs diverging slowly from one another. We propose that the results of our in silico modeling may explain why the trigenic interaction fraction tends to follow a bimodal distribution (Figs. 4A and 7). The upper mode of the distribution represents the set of duplicate pairs that will likely remain fixed in a partially functionally redundant state, whereas the lower mode represents duplicates that already have or are diverging in function. Because ohnologs and homeologs (*42*, *50*) show the same distribution of trigenic interaction fraction, this model of paralog divergence and retention of functional redundancy also likely applies to gene duplicates of various ages and origins beyond WGD, which may include small-scale duplicates.

In the simulation analysis, we modeled functions as being supported by contiguous sequence domains. However, because our model treats every position along a "gene" as statistically independent, positions contributing to a common function would not need to be contiguous, and the conclusions would remain the same for functional domains encoded by discontinuous sequences. Therefore, the model has the flexibility to capture a wide variety of different physiological scenarios that might display structural or functional entanglement, such as independent modular domains; linearly distant contact sites within a secondary, tertiary, or quaternary structure; or even regulatory regions beyond coding boundaries or elsewhere within the genome. It is important to note that this definition of structural and functional entanglement is distinct from the simple physical entanglement of proteins restricted to the basic organization of polypeptide chains (*63*).

The question of why subfunctionalization does not proceed to completion—leading to fixation of duplicated genes with some specific functions, yet exhibiting a certain level of functional redundancy—remains an outstanding problem in evolutionary biology. Constraints that prevent complete divergence may allow paralogs to retain the ability to function in parallel biochemical pathways or macromolecular complexes, thus resulting in a retention of redundancy (*18*, *43*, *64*). More specifically, despite functional divergence of independent domains, incomplete subfunctionalization of paralogs could be driven by a structurally and functionally overlapping ancestral domain (*35*). There are few existing models of duplicate evolution that specifically address the existence of redundancy in a steady state. There are models that address different potential modes of functional divergence, such as neo- or subfunctionalization (*39*, *65*). However, reasons for the persistence of functional redundancy have remained elusive. It is noteworthy that previous simple computer simulations, which incorporated mutation rates of genes and the varying contribution of their functions to overall fitness, have also identified situations in which redundancy can be maintained indefinitely (*66*–*68*). It has been shown that paralogs that are selected to function as distinct homomers also retain the ability to heterodimerize, which may prevent functional divergence between paralogs (*69*). In general, our findings support fixation of overlapping functional redundancy for a substantial proportion of yeast paralogs.

Methods summary

To study the functional divergence of duplicated genes, 240 double mutants and 480 corresponding single mutant control "query" strains, involving dispensable WGD pairs in the budding yeast *Saccharomyces cerevisiae* S288C, were constructed using PCR-mediated gene deletion followed by tetrad analysis. Paralog 1 deletion was marked with *natMX4*, while paralog 2 was deleted and replaced with *K. lactis URA3*. Single mutant control strains deleted for each one of the paralogs were also marked with the relevant control marker, which was inserted at the benign *HO* locus. Query strain fitness and query gene pair genetic interactions were measured using high-density synthetic genetic array (SGA) analysis. To obtain trigenic interactions, double mutant query strains along with their respective single mutant control query strains were subjected to trigenic-SGA analysis (τ-SGA), which involves a number of automated replica pinning steps. Each query strain was mated to a diagnostic array of 1200 strains, consisting of deletion mutants of nonessential genes and temperature-sensitive alleles of essential genes, providing a representative view of the global digenic interaction network. Briefly, the query and array strains were mated on rich media, *MAT***a**/α diploids were selected on media containing G418 and clonNAT, sporulation was induced by transferring to media with low levels of nitrogen and carbon sources, and *MAT***a** meiotic haploid progeny was selected on haploid selection media. Triple mutants were then selected by first pinning onto haploid selection media containing G418, lacking uracil, and then onto haploid selection media containing both G418 and clonNAT, lacking uracil. Every query strain was screened in two independent replicates.

Colony size was measured as a proxy for fitness, and digenic and trigenic interactions were scored using a quantitative model. Trigenic

interactions were classified into novel versus modified by overlapping with digenic interactions. Functional information embedded within digenic and trigenic interactions was assessed by their enrichment with external functional standards, such as protein-protein interactions, subcellular localization, coexpression, and co-annotation. Trigenic interaction fraction was calculated as the ratio of the negative trigenic interaction degree relative to the total negative digenic and trigenic interaction degree. Correlation of trigenic interaction fraction with physiological and evolutionary features included quantification of genetic interactions within a paralog pair, asymmetry of digenic interactions of members of each paralog pair, and sequence divergence rate, which was calculated as the raw difference between the fold-changes in substitutions per site in post-WGD clades. The potential for paralog induction during developmental programs was assessed in (i) meiosis using published meiotic mRNA-seq and ribosome profiling datasets, (ii) filamentous growth using a published measure of invasion, as well as (iii) glucose starvation conditions using published gene expression dataset. Dosage selection was estimated using global digenic interaction profile correlation similarity.

SAFE (Functional annotations based on the Spatial Analysis of Functional Enrichment) of the global genetic interaction profile similarity network was used to annotate gene function. Enrichment was calculated using the overlap of trigenic interactions with a neighborhood on the global digenic interaction similarity network. Novel paralog function for *ECM13-YJR115W* was interrogated using a drug sensitivity spot assay on media containing benomyl, and a liquid growth curve analysis on media containing latrunculin B. Spindle morphology was monitored by expressing Tub1-GFP, as well as sfGFP fusion proteins of Ecm13 and Yjr115w and imaging the resulting strains using a spinning-disc confocal microscope. Novel paralog function for *STB2-STB6* was monitored by Bap2-GFP localization in *stb2*Δ *stb6*Δ double mutant deletion strains and quantified using CellProfiler.

To measure evolutionary constraints on paralogs, evolutionary rates for specific amino acid columns in multiple sequence alignments were computed using the discrete gamma model of protein evolution, as implemented in PAML for the pre-WGD sequences and for each paralog separately. Pearson correlation coefficients were computed between the rates of the pre-WGD clade to each paralog (pre-WGD & Paralog 1 and pre-WGD and Paralog 2), and between the two paralogs (Paralog 1 and Paralog 2) to classify paralogs into those with low and high correlation of position specific evolutionary rate patterns. BioGRID was used to curate genetic interactions for the mitochondrial carrier protein family. Paralog divergence was simulated using a computational framework in which a gene of fixed length was generated, annotated with hypothetical functions and subjected to random degenerative mutations at a constant rate. Evolution to a steady state was achieved when no more divergence mutations could be tolerated while maintaining viability. The resulting paralogs were binned according to each pair's initial level of structural entanglement, which is the level of mutable positions within a gene that carry out two or more functions to quantify the number of paralogs that reverted to singleton state, completely diverged or retained functional overlap. For a more detailed description of the experimental and computational analyses, refer to the supplementary materials.

REFERENCES AND NOTES

1. J. E. Bowers, B. A. Chapman, J. Rong, A. H. Paterson, Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438 (2003). doi: 10.1038/nature01521; pmid: 12660784
2. P. Dehal, J. L. Boore, Two rounds of whole genome duplication in the ancestral vertebrate. *PLOS Biol.* **3**, e314 (2005). doi: 10.1371/journal.pbio.0030314; pmid: 16128622
3. Y. Guan, M. J. Dunham, O. G. Troyanskaya, Functional analysis of gene duplications in Saccharomyces cerevisiae. *Genetics* **175**, 933–943 (2007). doi: 10.1534/genetics.106.064329; pmid: 17151249
4. S. Maere *et al*., Modeling gene and genome duplications in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5454–5459 (2005). doi: 10.1073/pnas.0501102102; pmid: 15800040
5. E. E. Eichler, Recent duplication, domain accretion and the dynamic mutation of the human genome. *Trends Genet.* **17**, 661–669 (2001). doi: 10.1016/S0168-9525(01)02492-1; pmid: 11672867
6. M. Kellis, B. W. Birren, E. S. Lander, Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature* **428**, 617–624 (2004). doi: 10.1038/nature02424; pmid: 15004568
7. K. H. Wolfe, D. C. Shields, Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997). doi: 10.1038/42711; pmid: 9192896
8. J. Thornton, in *Evolutionary Genetics: Concepts and Case Studies*, C. W. Fox, J. B. Wolf, Eds. (Oxford Univ. Press, New York, 2006), pp. 160–161.
9. S. Ohno, in *Evolution by Gene Duplication* (Springer, 1970), chap. 10, p. 59.
10. I. Wapinski, A. Pfeffer, N. Friedman, A. Regev, Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–61 (2007). doi: 10.1038/nature06107; pmid: 17805289
11. C. Seoighe, K. H. Wolfe, Yeast genome evolution in the post-genome era. *Curr. Opin. Microbiol.* **2**, 548–554 (1999). doi: 10.1016/S1369-5274(99)00015-6; pmid: 10508730
12. K. P. Byrne, K. H. Wolfe, The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* **15**, 1456–1461 (2005). doi: 10.1101/gr.3672305; pmid: 16169922
13. Z. Gu *et al*., Role of duplicate genes in genetic robustness against null mutations. *Nature* **421**, 63–66 (2003). doi: 10.1038/nature01198; pmid: 12511954
14. L. Grassi *et al*., Identity and divergence of protein domain architectures after the yeast whole-genome duplication event. *Mol. Biosyst.* **6**, 2305–2315 (2010). doi: 10.1039/c003507f; pmid: 20820472
15. A. Baudot, B. Jacq, C. Brun, A scale of functional divergence for yeast duplicated genes revealed from analysis of the protein-protein interaction network. *Genome Biol.* **5**, R76 (2004). doi: 10.1186/gb-2004-5-10-r76; pmid: 15461795
16. L. Hakes, J. W. Pinney, S. C. Lovell, S. G. Oliver, D. L. Robertson, All duplicates are not equal: The difference between small-scale and genome duplication. *Genome Biol.* **8**, R209 (2007). doi: 10.1186/gb-2007-8-10-r209; pmid: 17916239
17. J. Li, Z. Yuan, Z. Zhang, The cellular robustness by genetic redundancy in budding yeast. *PLOS Genet.* **6**, e1001187 (2010). doi: 10.1371/journal.pgen.1001187; pmid: 21079672
18. B. Papp, C. Pál, L. D. Hurst, Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661–664 (2004). doi: 10.1038/nature02636; pmid: 15190353
19. D. Vitkup, P. Kharchenko, A. Wagner, Influence of metabolic network structure and function on enzyme evolution. *Genome Biol.* **7**, R39 (2006). doi: 10.1186/gb-2006-7-5-r39; pmid: 16684370
20. R. Kafri, A. Bar-Even, Y. Pilpel, Transcription control reprogramming in genetic backup circuits. *Nat. Genet.* **37**, 295–299 (2005). doi: 10.1038/ng1523; pmid: 15723064
21. X. Gu, Z. Zhang, W. Huang, Rapid evolution of expression and regulatory divergences after yeast gene duplication. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 707–712 (2005). doi: 10.1073/pnas.0409186102; pmid: 15647348
22. G. C. Conant, K. H. Wolfe, Functional partitioning of yeast co-expression networks after genome duplication. *PLOS Biol.* **4**, e109 (2006). doi: 10.1371/journal.pbio.0040109; pmid: 16555924
23. I. Tirosh, N. Barkai, Comparative analysis indicates regulatory neofunctionalization of yeast duplicates. *Genome Biol.* **8**, R50 (2007). doi: 10.1186/gb-2007-8-4-r50; pmid: 17411427
24. A. DeLuna, M. Springer, M. W. Kirschner, R. Kishony, Need-based up-regulation of protein levels in response to deletion of their duplicate genes. *PLOS Biol.* **8**, e1000347 (2010). doi: 10.1371/journal.pbio.1000347; pmid: 20361019
25. G. Diss *et al*., Gene duplication can impart fragility, not robustness, in the yeast protein interaction network. *Science* **355**, 630–634 (2017). doi: 10.1126/science.aai7685; pmid: 28183979
26. R. Mani, R. P. St. Onge, J. L. Hartman 4th, G. Giaever, F. P. Roth, Defining genetic interaction. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3461–3466 (2008). doi: 10.1073/pnas.0712255105; pmid: 18305163
27. P. Novick, D. Botstein, Phenotypic analysis of temperature-sensitive yeast actin mutants. *Cell* **40**, 405–416 (1985). doi: 10.1016/0092-8674(85)90154-0; pmid: 3967297
28. A. Bender, J. R. Pringle, Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **11**, 1295–1305 (1991). doi: 10.1128/MCB.11.3.1295; pmid: 1996092
29. J. van Leeuwen *et al*., Exploring genetic suppression interactions on a global scale. *Science* **354**, aag0839 (2016). doi: 10.1126/science.aag0839; pmid: 27811238
30. M. Costanzo *et al*., A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, aaf1420 (2016). doi: 10.1126/science.aaf1420; pmid: 27708008
31. A. H. Tong *et al*., Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004). doi: 10.1126/science.1091317; pmid: 14764870
32. M. Costanzo *et al*., The genetic landscape of a cell. *Science* **327**, 425–431 (2010). doi: 10.1126/science.1180823; pmid: 20093466
33. B. VanderSluis *et al*., Genetic interactions reveal the evolutionary trajectories of duplicate genes. *Mol. Syst. Biol.* **6**, 429 (2010). doi: 10.1038/msb.2010.82; pmid: 21081923
34. G. Musso *et al*., The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast. *Genome Res.* **18**, 1092–1099 (2008). doi: 10.1101/gr.076174.108; pmid: 18463300
35. E. J. Dean, J. C. Davis, R. W. Davis, D. A. Petrov, Pervasive and persistent redundancy among duplicated genes in yeast. *PLOS Genet.* **4**, e1000113 (2008). doi: 10.1371/journal.pgen.1000113; pmid: 18604285
36. A. DeLuna *et al*., Exposing the fitness contribution of duplicated genes. *Nat. Genet.* **40**, 676–681 (2008). doi: 10.1038/ng.123; pmid: 18408719
37. E. Kuzmin *et al*., Systematic analysis of complex genetic interactions. *Science* **360**, eaao1729 (2018). doi: 10.1126/science.aao1729; pmid: 29674565
38. J. Ihmels, S. R. Collins, M. Schuldiner, N. J. Krogan, J. S. Weissman, Backup without redundancy: Genetic interactions reveal the cost of duplicate gene loss. *Mol. Syst. Biol.* **3**, 86 (2007). doi: 10.1038/msb4100127; pmid: 17389874
39. A. Force *et al*., Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999). pmid: 10101175
40. A. van Hoof, Conserved functions of yeast genes support the duplication, degeneration and complementation model for gene duplication. *Genetics* **171**, 1455–1461 (2005). doi: 10.1534/genetics.105.044057; pmid: 15965245
41. A. C. Marques, N. Vinckenbosch, D. Brawand, H. Kaessmann, Functional diversification of duplicate genes through

subcellular adaptation of encoded proteins. *Genome Biol.* **9**, R54 (2008). doi: 10.1186/gb-2008-9-3-r54; pmid: 18336717
42. See supplementary materials.
43. A. N. Marshall, M. C. Montealegre, C. Jiménez-López, M. C. Lorenz, A. van Hoof, Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. *PLOS Genet.* **9**, e1003376 (2013). doi: 10.1371/journal.pgen.1003376; pmid: 23516382
44. A. van Hoof, P. A. Frischmeyer, H. C. Dietz, R. Parker, Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262–2264 (2002). doi: 10.1126/science.1067272; pmid: 11910110
45. M. K. Doma, R. Parker, Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006). doi: 10.1038/nature04530; pmid: 16554824
46. E. R. Kunji, The role and structure of mitochondrial carriers. *FEBS Lett.* **564**, 239–244 (2004). doi: 10.1016/S0014-5793(04)00242-X; pmid: 15111103
47. E. M. Froschauer, R. J. Schweyen, G. Wiesenberger, The yeast mitochondrial carrier proteins Mrs3p/Mrs4p mediate iron transport across the inner mitochondrial membrane. *Biochim. Biophys. Acta* **1788**, 1044–1050 (2009). doi: 10.1016/j.bbamem.2009.03.004; pmid: 19285482
48. G. C. Shaw *et al.*, Mitoferrin is essential for erythroid iron assimilation. *Nature* **440**, 96–100 (2006). doi: 10.1038/nature04512; pmid: 16511496
49. G. G. Perrone, S. X. Tan, I. W. Dawes, Reactive oxygen species and yeast apoptosis. *Biochim. Biophys. Acta* **1783**, 1354–1368 (2008). doi: 10.1016/j.bbamcr.2008.01.023; pmid: 18298957
50. M. Marcet-Houben, T. Gabaldón, Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage. *PLOS Biol.* **13**, e1002220 (2015). doi: 10.1371/journal.pbio.1002220; pmid: 26252497
51. Z. Cheng *et al.*, Pervasive, Coordinated Protein-Level Changes Driven by Transcript Isoform Switching during Meiosis. *Cell* **172**, 910–923. e16 (2018). doi: 10.1016/j.cell.2018.01.035; pmid: 29474919
52. O. Ryan *et al.*, Global gene deletion analysis exploring yeast filamentous growth. *Science* **337**, 1353–1356 (2012). doi: 10.1126/science.1224339; pmid: 22984072
53. P. H. Bradley, M. J. Brauer, J. D. Rabinowitz, O. G. Troyanskaya, Coordinated concentration changes of transcripts and metabolites in Saccharomyces cerevisiae. *PLOS Comput. Biol.* **5**, e1000270 (2009). doi: 10.1371/journal.pcbi.1000270; pmid: 19180179
54. F. A. Kondrashov, A. S. Kondrashov, Role of selection in fixation of gene duplications. *J. Theor. Biol.* **239**, 141–151 (2006). doi: 10.1016/j.jtbi.2005.08.033; pmid: 16242725
55. A. N. Nguyen Ba *et al.*, Detecting functional divergence after gene duplication through evolutionary changes in posttranslational regulatory sequences. *PLOS Comput. Biol.* **10**, e1003977 (2014). doi: 10.1371/journal.pcbi.1003977; pmid: 25474245
56. W. Qian, J. Zhang, Protein subcellular relocalization in the evolution of yeast singleton and duplicate genes. *Genome Biol. Evol.* **1**, 198–204 (2009). doi: 10.1093/gbe/evp021; pmid: 20333190
57. B. Santos, M. Snyder, Sbe2p and sbe22p, two homologous Golgi proteins involved in yeast cell wall formation. *Mol. Biol. Cell* **11**, 435–452 (2000). doi: 10.1091/mbc.11.2.435; pmid: 10679005
58. S. Westermann *et al.*, Formation of a dynamic kinetochore-microtubule interface through assembly of the Dam1 ring complex. *Mol. Cell* **17**, 277–290 (2005). doi: 10.1016/j.molcel.2004.12.019; pmid: 15664196
59. A. N. Nguyen Ba *et al.*, Proteome-wide discovery of evolutionary conserved sequences in disordered regions. *Sci. Signal.* **5**, rs1 (2012). doi: 10.1126/scisignal.2002515; pmid: 22416277
60. K. E. Vest *et al.*, Overlap of copper and iron uptake systems in mitochondria in Saccharomyces cerevisiae. *Open Biol.* **6**, 150223 (2016). doi: 10.1098/rsob.150223; pmid: 26763345
61. F. Palmieri, C. L. Pierri, A. De Grassi, A. Nunes-Nesi, A. R. Fernie, Evolution, structure and function of mitochondrial carriers: A review with new insights. *Plant J.* **66**, 161–181 (2011). doi: 10.1111/j.1365-313X.2011.04516.x; pmid: 21443630
62. R. Belenkiy, A. Haefele, M. B. Eisen, H. Wohlrab, The yeast mitochondrial transport proteins: New sequences and consensus residues, lack of direct relation between consensus residues and transmembrane helices, expression patterns of the transport protein genes, and protein-protein interactions with other proteins. *Biochim. Biophys. Acta* **1467**, 207–218 (2000). doi: 10.1016/S0005-2736(00)00222-4; pmid: 10930523
63. Y. Zhao, M. Cieplak, Stability of structurally entangled protein dimers. *Proteins* **86**, 945–955 (2018). doi: 10.1002/prot.25526; pmid: 29790597
64. B. Papp, C. Pál, L. D. Hurst, Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194–197 (2003). doi: 10.1038/nature01771; pmid: 12853957
65. S. Ohno, in *Evolution by Gene Duplication* (Springer, 1970), chap. 13, pp. 71–72.
66. M. A. Nowak, M. C. Boerlijst, J. Cooke, J. M. Smith, Evolution of genetic redundancy. *Nature* **388**, 167–171 (1997). doi: 10.1038/40618; pmid: 9217155
67. A. Wagner, The role of population size, pleiotropy and fitness effects of mutations in the evolution of overlapping gene functions. *Genetics* **154**, 1389–1401 (2000). pmid: 10757778
68. T. Vavouri, J. I. Semple, B. Lehner, Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends Genet.* **24**, 485–488 (2008). doi: 10.1016/j.tig.2008.08.005; pmid: 18786741
69. A. Marchant *et al.*, The role of structural pleiotropy and regulatory evolution in the retention of heteromers of paralogs. *eLife* **8**, e46754 (2019). doi: 10.7554/eLife.46754; pmid: 31454312
70. A. Baryshnikova, Systematic Functional Annotation and Visualization of Biological Networks. *Cell Syst.* **2**, 412–421 (2016). doi: 10.1016/j.cels.2016.04.014; pmid: 27237738

ACKNOWLEDGMENTS

We thank H. Friesen, E. Ünal, A. Caudy, and J. Hanchard for discussions and experimental input, and A. Baryshnikova for discussions and critical comments on the manuscript. **Funding:** Supported by NIH grant R01HG005853 (C.B., B.J.A., and C.L.M.), Canadian Institutes of Health Research grants FDN-143264 and FDN-143265 (C.B. and B.J.A.), NIH grants R01HG005084 and R01GM104975 (C.L.M.), and NSF grant DBI\0953881 (C.L.M.). Computing resources and data storage services were partially provided by the Minnesota Supercomputing Institute and the UMN Office of Information Technology, respectively. Additional support was provided by Natural Science and Engineering Research Council of Canada Postgraduate Scholarship–Doctoral PGS D2 (E.K. and A.N.N.B.), a University of Toronto Open Fellowship (E.K.), a University of Minnesota Doctoral Dissertation Fellowship (B.V.), and Deutsche Forschungsgemeinschaft grant CRC1036/TP10 (A.K. and M.K.). C.B. is a fellow of the Canadian Institute for Advanced Research (CIFAR). **Author contributions:** Conceptualization: E.K., B.V., C.L.M., B.J.A., and C.B.; methodology and investigation: E.K., B.V., A.N.N.B., W.W., E.N.K., M.U., A.K., M.M.U., J.v.L., O.K., A.T., M.P., M.-C.H., B.V., M.C., M.K., and A.M.; formal analysis: E.K., B.V., A.N.N.B., W.W., E.N.K., M.U., J.v.L., O.K., A.T., and A.M.; data Curation: M.U.; writing–original draft: E.K., B.V., C.L.M., B.J.A., and C.B.; writing–review and editing: E.K., B.V., A.N.N.B., E.N.K., A.K., M.M.U., J.v.L., A.T., M.C., M.K., A.M., C.L.M., B.J.A., and C.B.; supervision: C.L.M., B.J.A., and C.B.; funding acquisition: C.L.M., B.J.A., and C.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data associated with this study are available in the supplementary materials. The genetic interaction data are available in a searchable format at http://boonelab.ccbr.utoronto.ca/paralogs/. Tables S1 to S13 were also deposited in the DRYAD Digital Repository (https://doi.org/10.5061/dryad.g79cnp5m9). MATLAB routines that produce SGA digenic and trigenic interaction scores are available at https://doi.org/10.5281/zenodo.3665423.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/eaaz5667/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S13
References (71–96)
MDAR Reproducibility Checklist

19 September 2019; accepted 6 May 2020
10.1126/science.aaz5667

RESEARCH ARTICLE SUMMARY

DEVELOPMENTAL BIOLOGY

Muscle and neuronal guidepost-like cells facilitate planarian visual system regeneration

M. Lucila Scimone*, Kutay D. Atabay*, Christopher T. Fincher, Ashley R. Bonneau, Dayan J. Li, Peter W. Reddien†

INTRODUCTION: Multiple strategies exist to promote precise wiring of developing neuronal circuits. One strategy involves guidepost cells, which exist transiently in embryos. Guidepost cells can act as intermediate guidance targets for axons or by providing a scaffold that facilitates axonal targeting. Most guidance mechanisms become dispensable once the circuit is assembled. Loss of guidance mechanisms creates a potential limitation on regeneration of neuronal patterns—yet some animals are capable of functional regeneration of their nervous system.

RATIONALE: Assuming some adult animals have the ability to regenerate functional neuronal circuits, they must possess mechanism(s) for de novo repair of neuronal patterns. In this study, we aimed to characterize such mechanisms by studying regeneration of the planarian visual system after diverse injuries.

RESULTS: We identified a rare subset of muscle cells (*notum*$^+$*; frizzled 5/8-4*$^+$) concentrated at two precise anatomical locations and in tight association with photoreceptor axons. The first group of these cells was found near the eye, where visual axons project and fasciculate to form a bundle. The second group of these cells was found near choice points, where sorting of contralateral and ipsilateral axons occurs. Both groups of muscle cells were formed during regeneration of the visual system and were always tightly associated with axonal projections, consistent with a possible role in attraction to facilitate visual system assembly. In addition, we found that a *notum*$^+$ set of neurons, located at the adult anterior brain commissure, regenerated before axonal midline crossing and was associated with optic chiasm regeneration.

We reasoned that if the photoreceptor axon–associated *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells have a guidepost-like function, their formation should be independent of eye cells. Eyes transplanted to ectopic anatomical locations did not result in the formation of *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells. Furthermore, animals that were unable to generate eyes [*ovo* RNA interference (*ovo* RNAi) animals] were still capable of specifying these muscle cells at the right locations. In addition, we predicted that if these muscle cells were indeed guidepost-like cells, visual axon trajectories should be associated with them after eye transplantation into eyeless heads. In all instances, axons from transplanted eyes projected toward *notum*$^+$*; frizzled 5/8- 4*$^+$ muscle cells and often adjusted their trajectories after encountering them.

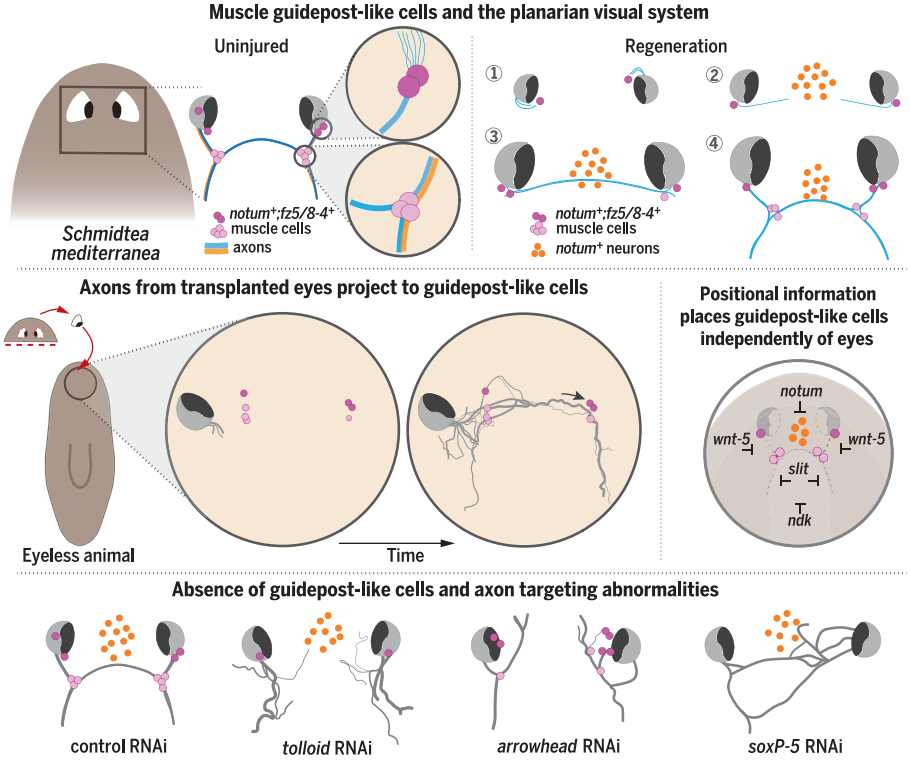
ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aba3203

We found that an array of signaling cues, which provide positional information essential for planarian patterning, was required for dictating the precise location of these guidepost-like cells. This provides a visual system–extrinsic mechanism for placing guidepost-like cells in the adult.

Finally, with single-cell RNA sequencing and fluorescent in situ hybridization screening, we identified molecules and transcription factors expressed in these cells. RNAi studies reduced or eliminated muscle or neuronal guidepost-like cell subsets and resulted in aberrant patterns of visual axonal trajectories.

CONCLUSION: Adult molecular and cellular strategies for regenerating neuronal pattern in the absence of embryo-specific contexts must exist to overcome damage or loss after injury. Understanding these mechanisms might provide important insights for regenerative medicine. Here, we found adult guidepost-like cell populations, extrinsic to the visual system and placed by adult positional information, that promote normal visual system regeneration in planarians. ■



Adult guidepost-like cells facilitate visual system regeneration in planarians. Muscle and neuron guidepost-like cells are present at key locations near the planarian visual system and are formed independently of photoreceptor axons. Regenerating and transplanted eyes target projections to guidepost-like cells. Positional information cues provide an eye-extrinsic mechanism to place guidepost-like cells in the adult. Loss of guidepost-like cells is associated with visual system disruption. *wnt-5*, *slit*, and *ndk* are involved in positional control of guidepost-like cell placement. *fz5/8-4*, *frizzled 5/8-4*.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: reddien@wi.mit.edu
Cite this article as M. L. Scimone *et al.*, *Science* **368**, eaba3203 (2020). DOI: 10.1126/science.aba3203

RESEARCH ARTICLE

DEVELOPMENTAL BIOLOGY

Muscle and neuronal guidepost-like cells facilitate planarian visual system regeneration

M. Lucila Scimone[1,2,3]*, Kutay D. Atabay[1,2,3]*, Christopher T. Fincher[1,2,3], Ashley R. Bonneau[1,2,3], Dayan J. Li[1,2,3], Peter W. Reddien[1,2,3]†

Neuronal circuits damaged or lost after injury can be regenerated in some adult organisms, but the mechanisms enabling this process are largely unknown. We used the planarian *Schmidtea mediterranea* to study visual system regeneration after injury. We identify a rare population of muscle cells tightly associated with photoreceptor axons at stereotyped positions in both uninjured and regenerating animals. Together with a neuronal population, these cells promote de novo assembly of the visual system in diverse injury and eye transplantation contexts. These muscle guidepost-like cells are specified independently of eyes, and their position is defined by an extrinsic array of positional information cues. These findings provide a mechanism, involving adult formation of guidepost-like cells typically observed in embryos, for axon pattern restoration in regeneration.

Formation of neural circuits during development requires the orchestration of multiple events, such as the specification of neurons, the precise navigation of axons through the extracellular environment toward targets, and the generation of specific synapses. The first axons to extend in developing nervous systems, known as pioneer axons, grow along stereotyped routes and fasciculate with each other to form tracts that are followed by subsequent axons. Ablation of pioneer neurons can perturb the guidance of subsequent axons, causing delay or misrouting, although in most cases these axons still locate their final targets (*1*–*4*).

Growing axons navigate with the assistance of cell-extrinsic guidance cues, such as Netrins and their DCC/UNC-40 or UNC-5 receptors; Slits and their Robo receptors; Semaphorins and their Plexin receptors; and Ephrins and their receptors (*5*–*8*). In addition, transient cell-cell interactions have important roles in neural circuit assembly. Such interactions involve cells referred to as guideposts. Guidepost cells, which were originally identified in the grasshopper limb (*9*, *10*), are discrete, early-born, and transient specialized cell populations located at decision or choice points along axonal trajectories (*6*, *11*, *12*). Most identified guidepost cells are either glia or neurons (*11*, *13*–*16*). On encountering guidepost cells, axon growth cones can change their responsiveness to extracellular guidance cues and modify their trajectory (*6*). In most systems, positioning of guidepost cells at choice points along the axonal path is instructed by canonical axon guidance cues themselves (*5*–*7*), and the proper location of guidepost cells is fundamental for precise axonal tract development (*17*–*20*). Since the discovery of guidepost cells in the grasshopper limb, guidepost cells in several other organisms have been identified, including midline glia cells in *Drosophila*, floor plate cells in the vertebrate spinal cord, and CD44$^+$ neurons and retinal glial cells found at the developing optic chiasm in mammals (*21*–*24*).

The transient developmental programs that guide pioneer axons to their ultimate targets become dispensable once the neuronal circuit is assembled. In regenerative species, however, neuronal circuits that are lost to injury must reestablish themselves in the adult animal. This presents a puzzle: How can an adult neuronal circuit be formed de novo by regeneration in the absence of embryonic and potentially transient guidance mechanisms, such as guidepost cells? In some instances, damage to the nervous system can involve repair by new neuron production but is limited by the inability of new neurons to wire correctly (*25*). However, in some other animals, regeneration of the nervous system seemingly restores normal function, although study of the regeneration of wiring patterns in highly regenerative models is sparse. Animals that have the ability to regenerate their neuronal architecture must possess adult mechanism(s) for repairing circuit patterns in the context of diverse injuries. To address how this can occur, we studied these mechanisms in the planarian *Schmidtea mediterranea*.

Planarians are freshwater flatworms that belong to the Spiralia superphylum (*26*). They are capable of whole-body regeneration and undergo constant tissue turnover. These processes involve neoblasts, a proliferating cell population containing pluripotent stem cells. In addition to dividing cells, proper regeneration requires positional information. Genes proposed to encode positional information in planarians are predominantly expressed in muscle cells in a regionally restricted manner across body axes (*27*). Whereas the cells and molecules guiding the fate choices of stem cells into new differentiated cells in planarian regeneration have received substantial attention, the mechanisms by which the architecture of the different planarian tissues is restored remains little studied.

We identified a set of muscle cells and neurons tightly associated with the planarian photoreceptor axons with guidepost-like attributes. These cells were constitutively present and capable of being regenerated after injury. The positioning of these muscle cells was guided by a constitutive system of positional information enabling their replacement at suitable positions after injury to promote precision in the regeneration of visual system architecture. Our results describe a strategy for de novo formation of precise axonal projection pattern after injury in the adult.

notum$^+$*; frizzled 5/8-4*$^+$ muscle cells are tightly associated with the visual system

Many planarian species have a pair of true cerebral eyes that are symmetrically located dorsal to the cephalic ganglia (*28*). Planarian eyes contain rhabdomeric photoreceptor neurons that extend their axons ipsi- and contralaterally in axon bundles and pigmented cells that build an optic cup (Fig. 1A). The visual axons follow a stereotypic path that can be visualized using an anti-Arrestin monoclonal antibody (*28*). Contralateral axons defasciculate from ipsilateral axons at choice points and project toward the midline, forming the optic chiasm. Both contralateral and ipsilateral axons project ventrally and posteriorly to the visual targets in the brain (*29*, *30*).

Using fluorescence in situ hybridization (FISH) in combination with immunostaining with an anti-Arrestin antibody, we noticed a previously unknown and small population of cells that were in close association with the visual system, marked by expression of the genes *notum* and *frizzled 5/8-4* (Fig. 1, B and C). We found that these cells were mostly concentrated in two regions: (i) near the eye, where the photoreceptor axons form a bundle projecting out from the eye, and (ii) near the axon choice points, located more ventrally (Fig. 1, B to D). The number of these *notum*$^+$*; frizzled 5/8-4*$^+$ cells found in each animal was variable but scaled with animal size (Fig. 1D). In a transverse cross section at the level of the eye, these cells were clearly distinguishable from other *notum*$^+$ cells known to exist in the

[1]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Whitehead Institute, 455 Main Street, Cambridge, MA 02142, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: reddien@wi.mit.edu

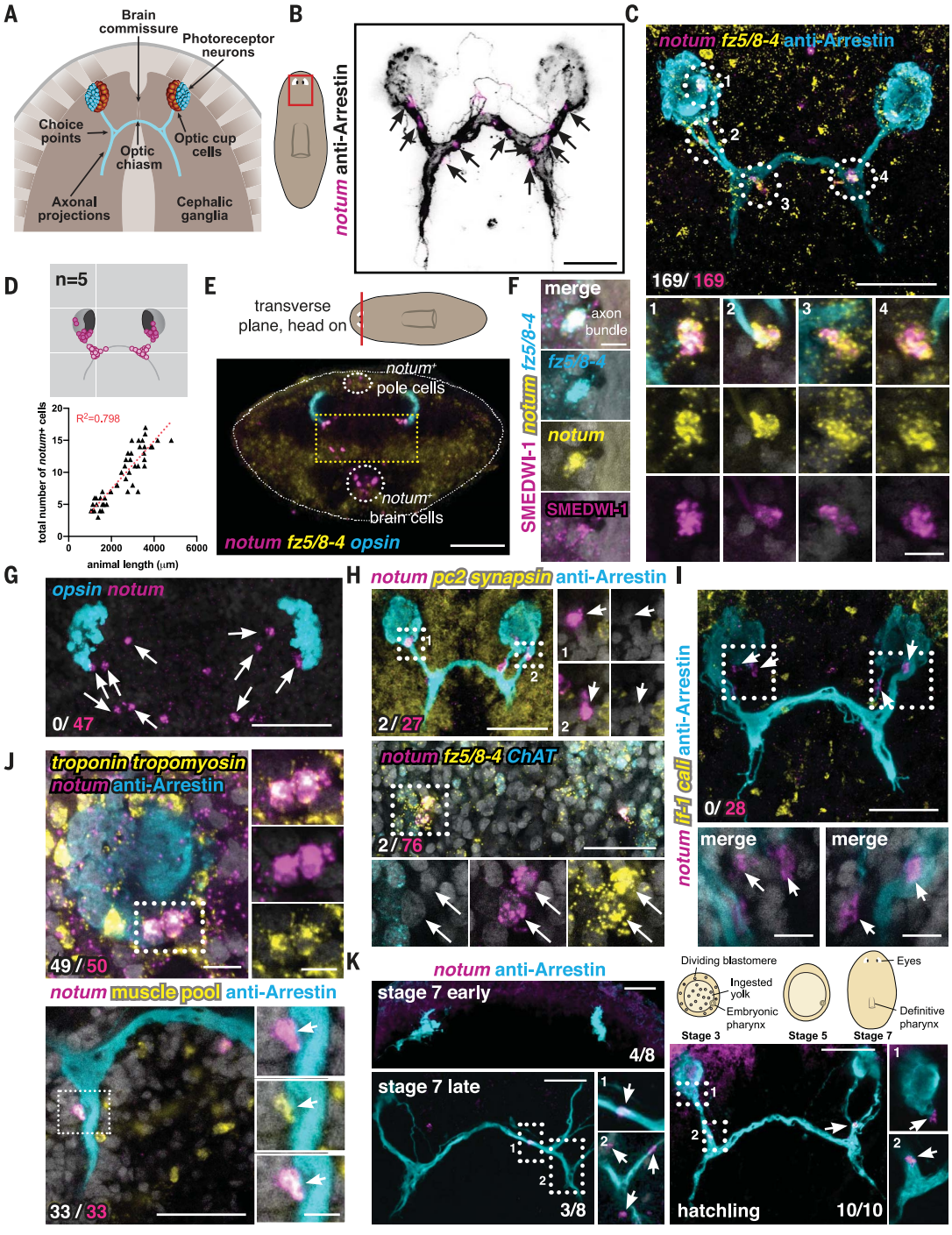


Fig. 1. *notum*$^+$; *frizzled 5/8-4*$^+$ muscle cells are tightly associated with the planarian visual system. (**A**) Schematic diagram of planarian visual system. (**B**) *notum*$^+$ cells associated with visual axons. Red box, location of image taken. (**C** and **D**) Coexpression (C) and density map [(D), top] of *notum* and *frizzled 5/8-4* (*fz5/8-4*) in cells associated with visual axons. Dark pink, cells near the eye; light pink, cells at choice points. The graph [(D), bottom] shows a positive correlation between the number of cells associated with the visual system and animal size (length). Red line, best fit for linear regression. R^2, coefficient of determination. (**E**) Transverse cut (red line) shows nuclei position of three *notum*$^+$ cell subsets. Yellow dotted line, *notum*$^+$; *frizzled 5/8-4*$^+$ cells. Dorsal, top; ventral, bottom. (**F**) SMEDWI-1 presence in a cell associated with visual system indicates recent specification. (**G** to **J**) *notum*$^+$ cells associated with visual axons do not express eye-specific markers (G), neuron-specific markers (H), or glia-specific markers (I) but do express muscle-specific markers [(J); pool: *troponin*, *tropomyosin*, *colF-2*, *colF-10*]. (**K**) Muscle *notum*$^+$ cells are associated with visual axons in *S. polychroa* embryos. The cartoon shows embryonic stages. In (G) to (K), white arrows point to *notum*$^+$ cells associated with the visual system. Ratios represent the fraction of double-positive cells (white) of total (magenta). When both numbers are in the same color, the ratio indicates the fraction of animals with that phenotype. Scale bars are 50 μm in (B), (C), (E), and (G) to (K) and 10 μm for all magnified views in (C), (F), and (H) to (J).

planarian head, such as anterior pole cells (*31*) and *notum*$^+$ neurons (*32*) located at the anterior commissure of the brain (Fig. 1E and fig. S1A). The nuclei of the *notum*$^+$; *frizzled 5/8-4*$^+$ cells found near the eye were located dorsal to the nuclei of the *notum*$^+$; *frizzled 5/8-4*$^+$ cells found at the choice points, which were observed immediately dorsal to the cephalic ganglia (movie S1). To make this dorsal-ventral–location distinction clear, we have colored these two populations differently in illustrations throughout this article.

The clear association of these cells at two discrete and important locations near axons of the visual system raised the possibility of a role for these previously unidentified cells in visual system wiring.

Planarian eye cells, like all other cell types assessed in this organism, are replaced during normal tissue turnover. This process requires neoblasts, the only proliferating somatic cells in the animal (*33*). Neoblasts express the *piwi* homolog *smedwi-1*. *smedwi-1* transcription ceases as cells leave the neoblast state and begin to differentiate, but the SMEDWI-1 protein transiently perdures as cells differentiate. An antibody that recognizes SMEDWI-1 therefore allows visualization of newly generated cells (*34*, *35*). We observed a small number of *notum*$^+$; *frizzled 5/8-4*$^+$ cells positive for SMEDWI-1 protein, indicating that these cells are constantly specified and replaced in uninjured adult animals (Fig. 1F).

To determine the identity of *notum*$^+$; *frizzled 5/8-4*$^+$ cells, we examined them for coexpression of different cell-specific markers. These

cells did not express the photoreceptor neuron marker *opsin* or the optic cup marker *tyrosinase* (Fig. 1G and fig. S1, B and C); the neuronal markers *pc2*, *synapsin*, or *ChAT* (Fig. 1H and fig. S1D); or the planarian glia markers *if-1* and *cali* (Fig. 1I). Some *fz5/8-4*$^+$ *ChAT*$^+$ cells were found near the eye but did not express *notum* (fig. S1E). Instead, *notum*$^+$*; frizzled 5/8-4*$^+$ cells expressed muscle markers (Fig. 1J) and could occasionally be detected in cross sections colabeled with an antibody for planarian muscle (6G10; fig. S2A). Muscle cells are broadly distributed within the animal (fig. S2, A and B); however, *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells were a rare subpopulation present in association with visual axons (movie S2). *notum*$^+$*; frizzled 5/8-4*$^+$ cells expressed low levels of muscle markers (*troponin*, *tropomyosin*, *colF-2*, and *mp-1*; fig. S2, C to E) compared with other muscle cell subsets such as body wall, intestinal, and other dorsal-ventral muscle cells (*36*, *37*), suggesting that these cells could be specialized muscle cells serving a different role. Muscle cells in general are known to secrete signaling factors in planarians (*27*). Despite lower levels of expression of muscle markers per cell, signal from pooled probes for canonical muscle markers, such as *troponin* and *tropomyosin*, was detected in essentially all of the *notum*$^+$*; frizzled 5/8-4*$^+$ cells (Fig. 1J and fig. S2F). Whether these cells retain contractile function is unknown. *notum*$^+$ muscle cells associated with the visual axons were also found during planarian embryonic development. Such *notum*$^+$ cells were first observed in prehatchling stage 7 *S. polychroa* embryos at the time of optic chiasm formation, but not before (Fig. 1K). After hatching, all free-swimming juveniles displayed a similar array of *notum*$^+$ muscle cells and visual axons, as did their adult counterparts (Fig. 1K). The tight association of *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells with visual axons raised the possibility of a role for these cells in the wiring of the planarian photoreceptor axons.

Regenerating visual axons project toward *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells

We defined *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells near the eye as NMEs (*notum*$^+$ muscle cells near the eye) and those at axon choice points as NMCs (*notum*$^+$ muscle cells at the choice point) (Fig. 2A). If NMEs and/or NMCs facilitate wiring of photoreceptor axons, they should be present before or concurrently with key axon-pathfinding processes and should also be tightly spatially associated with such events. To examine this possibility, we studied the regeneration of these cells and the dynamics of photoreceptor axons after different regenerative challenges.

First, we examined these cells in the context of a unilateral eye resection, which does not perturb the preexisting optic chiasm [Fig. 2B; (*38*)]. This injury removed NMEs but did not remove NMCs (Fig. 2B). Within 2 to 4 days after eye resection, new photoreceptor neurons nucleated dorsally (regenerating the resected eye) and projected their axons ventrally and posteriorly. NMC numbers stayed constant during this time (Fig. 2C). Despite variance in the relative positions of NMCs with respect to the new photoreceptor neurons and in the initial paths of projecting axons, the trajectories of photoreceptor axons essentially always coincided with the NMCs (Fig. 2, B and D, and fig. S3, A and B). This tight association between NMC position and axon path is consistent with the possibility that NMCs have an attractive influence on photoreceptor axons. NMEs regenerated 4 days after eye resection, suggesting that NMEs are not necessary for early projections during eye regeneration after resection, a scenario where NMCs remained present throughout the repair process.

We next studied the behavior of photoreceptor axons, NMEs, and NMCs after head amputation. In this context, the animals need to regenerate the entire visual system, including NMEs and NMCs. Between days 2 and 3 after decapitation, we observed nucleation of new eyes and the appearance of *notum*$^+$*; frizzled 5/8-4*$^+$ cells in close proximity to the regenerating eyes (Fig. 2E and fig. S3C). At this early time point, the regenerating blastema is small, making distinction between *notum*$^+$*; frizzled 5/8-4*$^+$ NMEs and NMCs not possible. These cells displayed detectable *frizzled 5/8-4* transcripts first, followed by *notum* expression (Fig. 2E). At this time, photoreceptor axons from the regenerating eyes projected toward where *notum*$^+$*; frizzled 5/8-4*$^+$ cells were observed (Fig. 2, E and F, and fig. S3, A and C). The pattern of axonal projections and *notum*$^+$*; frizzled 5/8-4*$^+$ cell location was variable, with some axon tracts projecting anterior-medially and others posteriorly, suggesting that the initial steps in axonal pattern formation are noisy, but the spatial association of axons and *notum*$^+$*; frizzled 5/8-4*$^+$ cells was apparent nonetheless. This is consistent with the possibility that in the context of de novo visual system formation, these cells have an attractive influence on early axonal projections. It is also possible that photoreceptor axons and *notum*$^+$*; frizzled 5/8-4*$^+$ cells might stabilize each other on interaction.

We noticed that the population of *notum*$^+$ neurons in the medial brain, normally located ventral to the optic chiasm in a fully formed brain, was in close association with photoreceptor axons during early head regeneration. We refer to this third population of cells associated with photoreceptor axons during regeneration as *notum*$^+$ brain cells (NBCs) (Fig. 2A). NBCs were apparent within 3 to 4 days after amputation at the regenerating anterior brain commissure [Fig. 2E; (*32*)]. Visual axons only projected to the midline, forming the optic chiasm, after NBCs appeared, and despite variation in their paths, they were typically associated with NBCs (Fig. 2, E and G, and fig. S3, A and C). This process is reminiscent of optic chiasm formation in mouse embryonic development, in which the site of the future optic chiasm is first populated by a subset of neurons expressing L1 and CD44. Subsequently, retinal ganglion axons grow and project toward the midline using these neurons as a landmark or a scaffold (*24*). Other cell types, such as radial glial cells at the optic chiasm, or midline cells in insects have a similar scaffolding role for axons crossing the midline during development (*39*).

In planarian head regeneration, early photoreceptor axonal projections cross the midline [Fig. 2, E and G; (*40*)]. At the time of optic chiasm formation, the pattern of axonal projections with choice points and *notum*$^+$*; frizzled 5/8-4*$^+$ cells near them became apparent (Fig. 2E). We also observed axonal projections to the brain targets at this time. By day 4 of regeneration, the pattern of the photoreceptor axons was similar across animals, indicating that initial patterning during de novo regeneration resolves with time into a stereotypical pattern. The events that occur can be approximated into the following stages: (i) eye nucleation, (ii) NME and/or NMC regeneration, and axon bundle formation and projection toward NMEs and/or NMCs, (iii) NBC regeneration, (iv) axonal projections toward NBCs and optic chiasm formation, and (v) axonal projections toward the visual areas in the brain. The close association between photoreceptor axons and NMEs, NMCs, and NBCs in this process is consistent with the possibility that these cells have a function in facilitating visual system wiring.

NMEs and NMCs are specified independently of eyes

If NMEs and NMCs have a guidepost-like function, their formation should at least in part be independent of the photoreceptor neurons themselves. To examine this possibility, we transplanted wild-type eyes (*41*) at different positions in prepharyngeal and postpharyngeal animal regions. We did not observe any cells coexpressing *notum* and *frizzled 5/8-4* near the transplanted eyes 10 days after transplantation (Fig. 3A and fig. S4A). Interestingly, we observed that axon bundles from transplanted eyes were commonly defasciculated (Fig. 3A and fig. S4A). These findings suggest that eyes in ectopic locations are insufficient to induce *notum*$^+$*; frizzled 5/8-4*$^+$ muscle cells.

Planarian eye cells (both photoreceptor neurons and optic cup cells) are specified by the transcription factor *ovo* (*42*). *ovo* RNA interference (RNAi) animals cannot replace eye

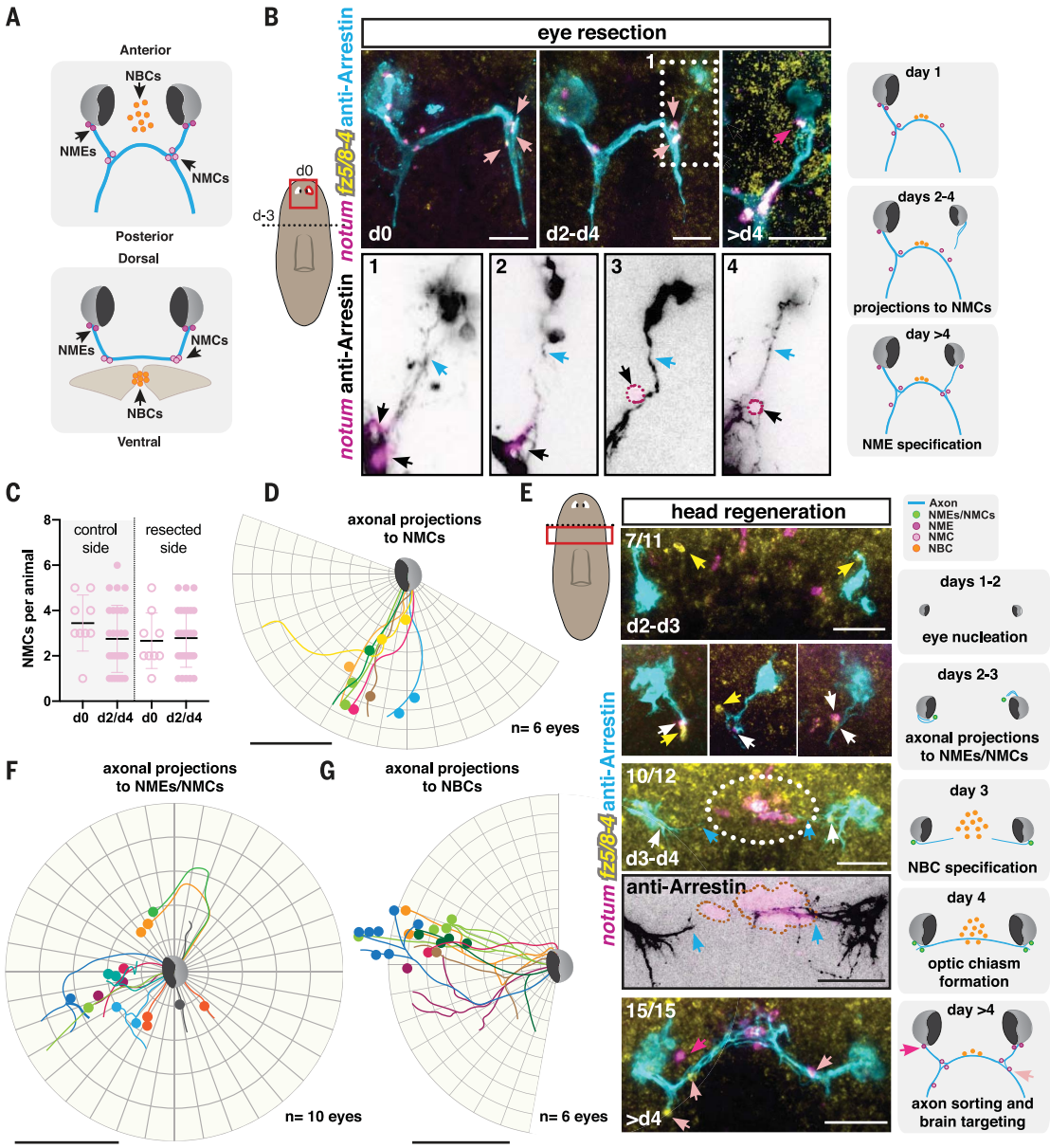


Fig. 2. *notum*$^{+}$; *frizzled 5/8-4*$^{+}$ muscle cells are associated with regenerating visual axons. (**A**) Illustrations show NME, NMC, and NBC positions across different axes. (**B**) Regeneration of visual system at different time points after unilateral eye resection. The bottom images are magnified views of visual axonal projection examples (1 to 4) 2 to 4 days after eye resection. Dotted outline, NMEs and NMCs. The left cartoon shows the surgical procedure: head decapitation (dotted line; day −3) and unilateral eye resection 3 days later [day 0 (d0); red line]. Red box, location of image taken. The cartoons on the right summarize events observed after eye resection. (**C**) There is no change in NMC numbers between resected and control sides. (**D**, **F**, and **G**) Circular plots show tracing of photoreceptor axonal trajectories (lines) from independent right eyes during regeneration (d2 to d4) of a resected eye (D) or decapitation [(F) and (G)]. Colored dots represent NMCs (D), NMEs and NMCs (F), or NBCs (G). (**E**) Regenerating visual system after decapitation. The dotted line in the left cartoon indicates the amputation line, and the red box shows the location of the image taken. Illustrations (right) summarize events observed after decapitation. Blue arrows, axons; white arrows, NMEs and/or NMCs; dark pink arrows or dots, NMEs; light pink arrows or dots and black arrows, NMCs. Orange dots or dotted outline, NBCs; yellow arrows, NMEs and NMCs expressing only *frizzled 5/8-4*. Scale bars are 50 μm in (B) and (D) to (G).

cells during normal tissue turnover or regenerate eyes after eye resection or decapitation. Nonetheless, NMEs and NMCs were both present in uninjured *ovo* RNAi animals, including in animals completely lacking photoreceptor neurons (Fig. 3B). However, their numbers were significantly reduced compared with control animals (Fig. 3B), suggesting a potential role for photoreceptor axons in the maintenance of normal NME and NMC numbers. Next, to determine whether NMEs and NMCs could regenerate independently of eye cells, we performed different injuries in *ovo* RNAi animals. We found that NMEs were able to regenerate after unilateral eye resection in *ovo* RNAi animals, indicating that eye cells were not required for NME specification in this context (Fig. 3C). Furthermore, NMEs were still observed 30 days after double-eye resection in *ovo* RNAi animals, suggesting that these cells were still specified even when no axons were left (fig. S4B). Moreover, NMEs and NMCs also regenerated after decapitation of *ovo* RNAi animals (Fig. 3D and fig. S4, C and D), although the total number of NMEs and NMCs in this context was significantly lower compared with controls (Fig. 3D). In conclusion, both NMEs and NMCs were observed in uninjured or regenerating animals that completely lacked photoreceptor neurons and associated axons. Together, these data suggest that NMEs and NMCs can be formed in the absence of eyes, but their homeostatic numbers are positively influenced by photoreceptor axons.

Axonal projections from transplanted eyes associate with NMEs and NMCs

If visual circuit–associated muscle cells (NMEs and NMCs) have guidepost-like activity, we predicted that visual axonal projections should be correlated to their location not just during regeneration but also after eye transplantation in the head. We performed unilateral eye transplantations at approximately the original

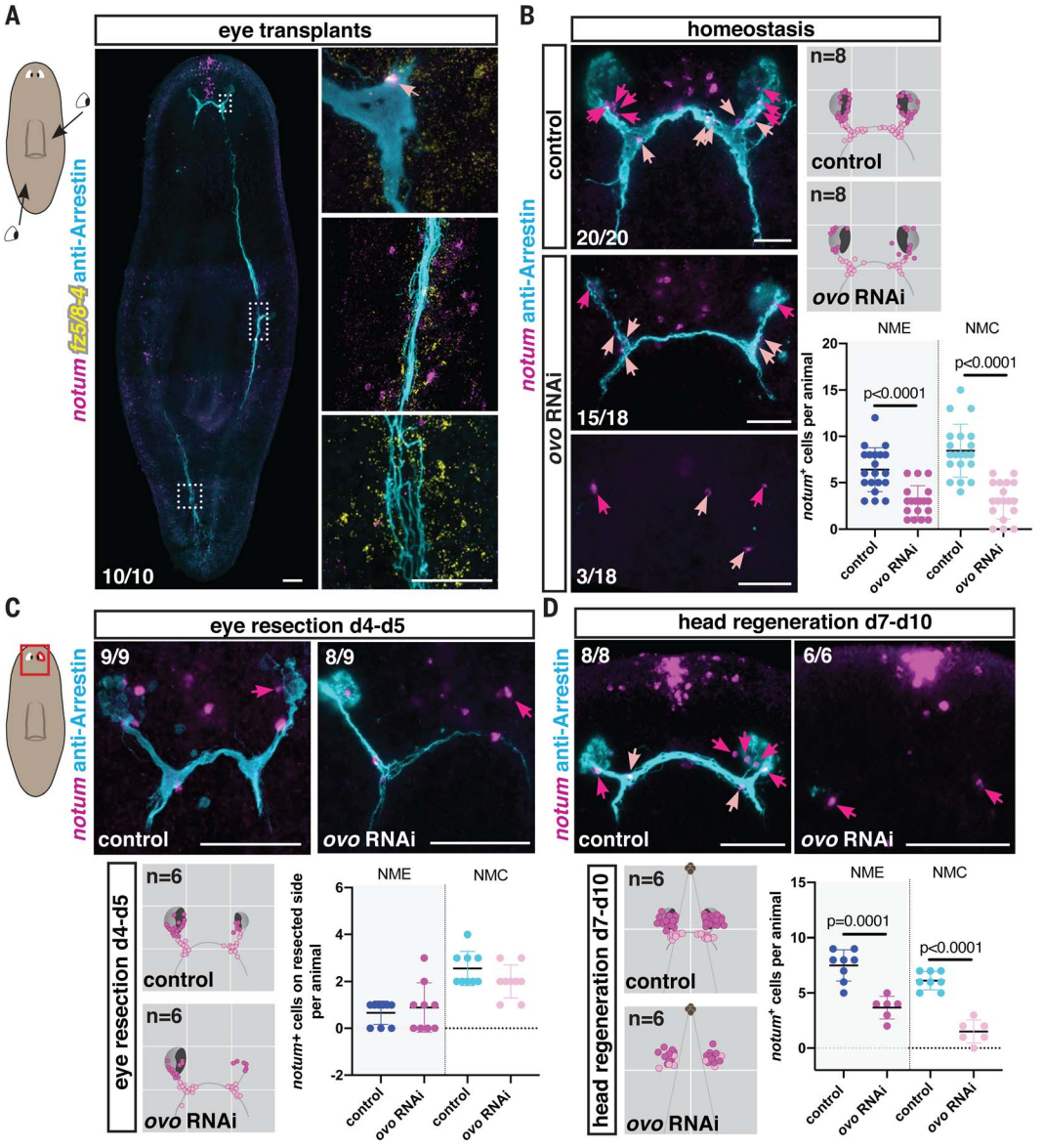


Fig. 3. Visual axons are not required for NME and NMC specification but are required for maintenance. (**A**) NMCs are observed associated with original visual axons but not with axons from transplanted eyes. The cartoon on the left shows the location of transplanted eyes. (**B**) NMEs and NMCs are present in uninjured animals with few or no visual axons (*ovo* RNAi). The density map at the top right shows NME and NMC distributions in an idealized visual system cartoon. n indicates the number of animals mapped. The graph at the bottom right shows NME and NMC numbers in uninjured RNAi animals. (**C** and **D**) Regeneration of NMEs (C) and NMEs and NMCs (D) in *ovo* RNAi animals after eye resection (C) or head amputation (D). Mapping at the bottom left shows NME and NMC distributions in an idealized visual system cartoon. n indicates the number of animals mapped. The graph at the bottom right shows NME and NMC numbers. The red box shows the location of the image taken. Dark pink arrows or dots, NMEs; light pink arrows or dots, NMCs. Scale bars are 100 μm in (A) and 50 μm in (B) to (D).

eye position in double-eye-resected *ovo* RNAi recipients (fig. S5A). As shown above (Fig. 3B), some NMEs and NMCs perdured in intact *ovo* RNAi animals, allowing assessment of correlation between the location of eye projections and NMEs and NMCs. One to 3 days after double-eye resection was not sufficient time to eliminate all preexisting photoreceptor axons in *ovo* RNAi recipients (fig. S5, B and C). Axonal projections from all transplanted eyes at this time point were able to navigate and follow the stereotypical path (fig. S5, B and C). To eliminate a potential role for remaining axons in serving as a scaffold for transplanted axons, we transplanted eyes 10 to 12 days after double-eye resection of *ovo* RNAi recipients. Within this time window, 87.5% of the *ovo* RNAi recipients had no remnant axons left, although some NMEs and NMCs were present at the time of transplantation (d0 in Fig. 4A). Axons from all transplanted eyes were prone to defasciculation and errors but projected ipsilaterally and crossed the midline in a stereotypical manner (Fig. 4A and fig. S5D). We observed that some NMEs were cotransplanted with the wild-type eye, explaining, at least in part, the increased number of observed NMEs at day 7 after transplantation (Fig. 4B). Importantly, in all instances, even when transplanted eyes were slightly mispositioned, the visual axon tracts extended toward the NMCs present in recipient animals and bifurcated at these choice points (magnified views in Fig. 4A and fig. S5D). Moreover, in some cases, we observed that axons locally deviated from the stereotypical route of the main axon bundle toward NMEs or NMCs that were present in the recipient animals (magnified views in Fig. 4A and fig. S5D). Trace mappings of axonal trajectories showed that nearly all axons passed close by NMEs and NMCs (within two cell diameters), and many directly interacted with NMEs and NMCs (Fig. 4C and fig. S6, A and B). The locations of eyes, axonal projections, and NMEs and NMCs were variable in these experiments, but tight association of projections with NMEs and NMCs was observed nonetheless. These results are consistent with the possibility that NMEs and NMCs might attract and/or stabilize visual axons. Moreover, our observations show that formation of an optic chiasm in adult planarians does not depend on interactions between the axons coming from each eye, in contrast to vertebrates (*23*, *39*).

Extrinsic axial patterning cues control NME and NMC positioning

If NMEs and NMCs facilitate a photoreceptor axonal projection pattern in regeneration, then

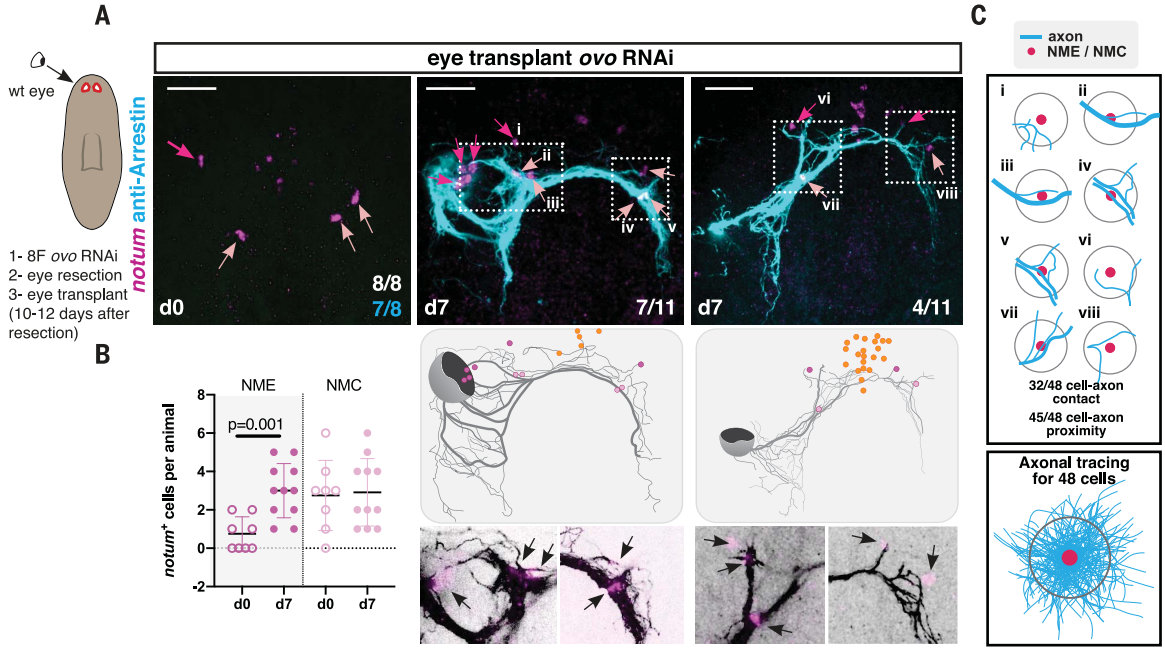


Fig. 4. NMEs and NMCs facilitate visual system patterning after eye transplantation. (**A**) Shown at the top, NMEs and NMCs are still present 10 to 12 days after double-eye resection in *ovo* RNAi animals before a wild-type (wt) eye is transplanted (left). Complete (middle) or incomplete (right) recapitulation of stereotypical visual axonal projections after eye transplantation in double-eye-resected *ovo* RNAi animals. The cartoon on the left shows the summarized surgical procedure: Animals were RNAi fed eight times (8F), double-eye resection was performed, and wild-type eyes were transplanted 10 to 12 days after eye resection. Animals were fixed just before (d0) or 7 days after (d7) transplantation. In the middle, cartoons show axonal tracings of the image shown above and the positions of NMEs, NMCs, and NBCs. The images at the bottom are magnified views (white dotted box) showing axonal projections from transplanted eyes in close association with NMEs and NMCs (black arrows). (**B**) NME and NMC numbers before and after transplantation. (**C**) Cartoons show examples from (A) (i to viii) of NME and NMC associations with visual axons. Numbers of cells that contact an axonal tract or found within a two-cell-diameter distance to the axonal tract are shown. Axonal tracing observed near 48 NMEs and NMCs after transplantation is shown at the bottom. Dark pink arrows or dots, NMEs; light pink arrows or dots, NMCs. Scale bars are 50 μm in (A).

the positioning of these muscle cells should be regulated, at least in part, independently from the visual system itself. That NMEs and NMCs can be regenerated independently of photoreceptor neurons is consistent with this possibility and suggests that an extrinsic mechanism should exist to position these cells. Correct positioning of guidepost cells is required for the precise wiring of neural circuits in other organisms. In most examined cases, canonical axon guidance cues (Netrins and DCC or UNC-5; Slits and Robo; Semaphorins and Plexins; and Ephrins and Ephrin receptors) also affect guidepost cell positioning (*12*). To examine the role of conserved guidance cues in NME and NMC positioning, we inhibited previously reported and newly identified planarian genes encoding guidance cue homologs with RNAi (fig. S7, A and B). Inhibition of these genes perturbed photoreceptor axonal projections [fig. S8, A and B, and (*43*–*46*)] but did not alter overall numbers or positions of NMEs and NMCs (fig. S8, A to C), except when planarian axial patterning was disrupted (see below).

Positional information is essential during both planarian regeneration and normal tissue turnover to inform neoblasts and/or their progeny about the location of tissues that need to be replaced. In planarians, a variety of signaling molecules known as position control genes (PCGs) have important roles in axial patterning and constitute adult positional information. PCGs are predominantly expressed in muscle cells along different planarian axes (*27*, *47*), and inhibition of some PCGs results in notable patterning defects, including hypercephalized animals, animals with duplication of eyes or pharynges (*33*, *47*–*49*). Inhibition of PCGs affected the pattern of NMEs and NMCs during normal tissue turnover (Fig. 5A) and regeneration (fig. S8A).

The planarian medial-lateral (ML) axis is regulated by medial *slit* and lateral *wnt5*, with *wnt5* restricting *slit* expression to the midline (*50*). Inhibition of *slit* or *wnt5* affected the ML positioning of NMEs and NMCs in both uninjured and regenerating animals (Fig. 5A and fig. S8A). *slit* RNAi animals showed a shift of these cells toward the midline, where an ectopic or cyclopic eye was observed (Fig. 5A and fig. S8A). *roboC* RNAi resulted in a similar phenotype to that of *slit* RNAi, suggesting that *roboC* encodes a major Slit receptor (fig. S8, A and D). *wnt5* RNAi animals showed increased numbers of NMEs and NMCs laterally, in conjunction with the appearance of ectopic lateral eyes (Fig. 5A). Regenerating *wnt5* RNAi animals did not form an optic chiasm, and NMEs and NMCs were displaced more laterally (fig. S8A). *ror-1* RNAi resulted in a phenotype similar to that of *wnt5* RNAi animals during both regeneration and homeostasis (fig. S8, A and D), suggesting that *ror1* encodes the Wnt5 receptor in this process.

RNAi of genes involved in the patterning of the anterior-posterior (AP) axis affected the AP positioning of NMEs and NMCs (Fig. 5A). *notum* RNAi causes anteriorization of brain and ectopic anterior eye appearance (*32*), and these animals showed ectopic anterior NMEs and NMCs. Because of the axial patterning role of *notum*, it was not possible to readily determine the function of *notum* within NME and NMCs.

nou darake (*ndk*) RNAi animals develop ectopic posterior eyes (*47*, *48*) and shift anterior PCG expression domains posteriorly (*47*). *ndk* RNAi caused ectopic posterior appearance of NMEs and NMCs. Major surgery leading to extensive PCG shifts in head fragments (*41*) also resulted in the appearance of ectopic NMEs and NMCs at positions consistent with the locations of ectopic de novo eyes (Fig. 5A). In PCG RNAi experiments, NMEs and NMCs were sometimes observed in ectopic locations before the nucleation of a new ectopic eye or the projection of a new axon bundle (fig. S9A). This is consistent with the possibility that PCG alteration itself leads to ectopic positioning of NMEs and NMCs as opposed to these cells being induced in ectopic locations only after ectopic eyes or axonal projections appeared.

To further test this hypothesis, we generated posteriorized *ndk* RNAi animals. After allowing RNAi to wear off, these animals maintained all ectopic posterior eyes and posteriorly expanded

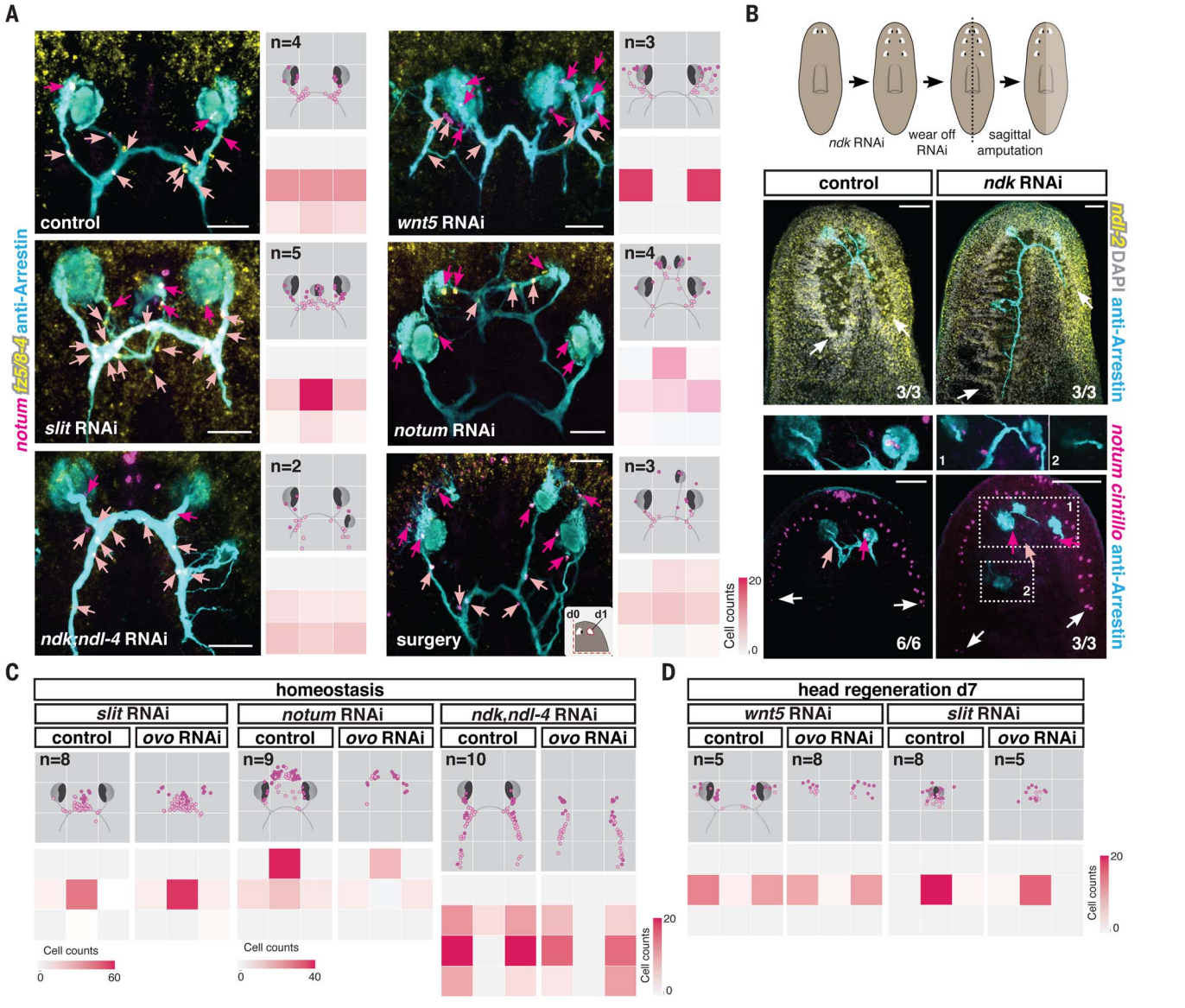


Fig. 5. Axial patterning genes are required for positioning of NMEs and NMCs. (**A**) Visual system and NMEs and NMCs in different uninjured RNAi conditions or following surgery are shown on the left. The density map (top right) shows NME and NMC distributions in an idealized visual system illustration. n indicates the number of animals mapped. The heatmap (bottom right) shows numbers of NMEs and NMCs in each quadrant. (**B**) Illustration (top) summarizing the experimental procedure. Rescaling of *ndl-2* PCG expression, maintenance of posteriorized brain lobe in old tissue, and regeneration of normal size brain lobe in blastema are shown in the middle. Shown at the bottom, NMEs and NMCs are present near eyes located at the correct position after PCG rescaling (1) but not near ectopic posterior eyes (2) in an *ndk* RNAi animal. White arrows point to brain lobes or neurons (*cintillo*$^{+}$). (**C**) Mapping shows NME and NMC distributions in an idealized visual system cartoon of uninjured PCG RNAi animals in the presence (control) or absence (*ovo* RNAi) of eyes. n indicates the number of animals mapped. The heatmap shows the total number of NMEs and NMCs located in each quadrant. *slit* RNAi leads to medialization, *notum* RNAi leads to anteriorization, and *ndk; ndl-4* RNAi leads to posteriorization of NME and NMCs—in each case independently of eyes. (**D**) *wnt5* RNAi animals show a lateralized NME and NMC distribution, whereas *slit* RNAi animals show medialized cell distribution independently of visual axons during regeneration. Dotted lines in cartoons show amputation planes. Dark pink arrows or dots, NMEs; light pink arrows or dots, NMCs. Scale bars are 50 μm in (A) and 100 μm in (B).

brain lobes, even though *ndk* function returns and the PCG map scales back to its normal proportions (*41*). Sagittal amputations at this time result in animals regenerating a side with a new eye and brain lobe at the correct scale and anterior position, generating an asymmetrical body plan [Fig. 5B and (*41*)]. Under these conditions, NMEs and NMCs only regenerated at the correct AP position with respect to the rescaled PCG domains, near the regenerating eye within the blastema. Within the uninjured old tissue side, which maintained a posteriorly expanded cephalic ganglion and ectopic eyes, NMEs and NMCs were also only observed near the eye that was found at the correct position based on the rescaled PCG map. Posterior ectopic eyes did not maintain NMEs and NMCs (Fig. 5B and fig. S9B), unlike animals that were kept under *ndk* RNAi conditions (Fig. 5A). These results further indicate a role for PCGs, rather than ectopic eyes themselves, in instructing the positioning of newly specified NMEs and NMCs.

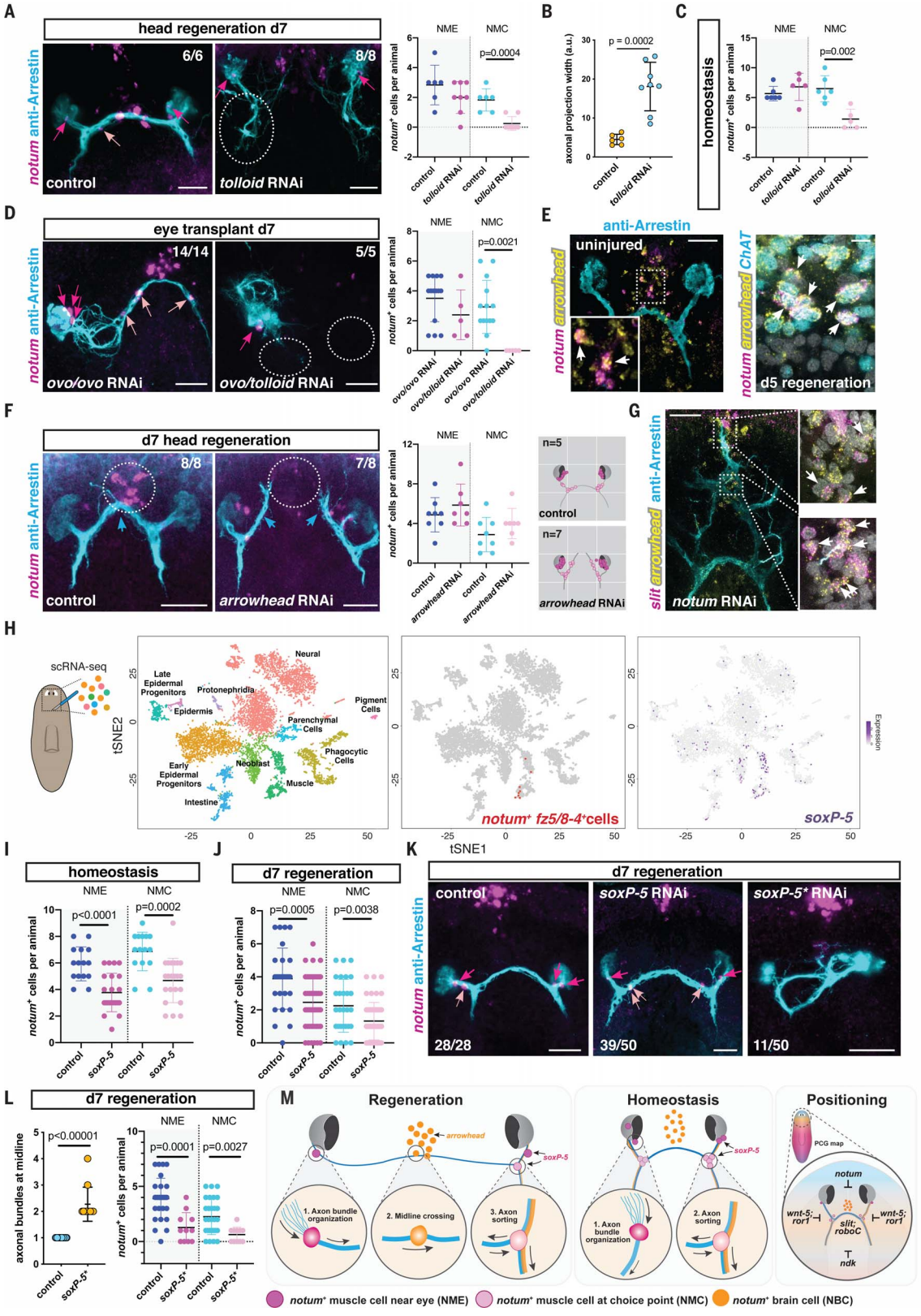


Fig. 6. *tolloid*, *arrowhead*, and *soxP-5* are required for specification of guide-post-like cells and precise wiring of the visual system. (**A**) Defasciculation of visual axons (white dotted circle) and absence of NMCs in a regenerating *tolloid* RNAi animal. The graph on the right shows NME and NMC numbers after head regeneration. (**B**) Axonal projection width in RNAi animals. (**C**) NME and NMC numbers in uninjured animals. (**D**) Lack of NMCs and inability of transplanted eyes to cross the midline in *tolloid* RNAi animals. The graph on the right shows NME and NMC numbers after transplantation. Dotted ovals indicate lack of NMCs. (**E**) NBCs expressed the transcription factor *arrowhead* in intact (left) or regenerating (right) animals. White arrows show coexpression of *arrowhead* and *notum* in NBCs. (**F**) Absence of NBCs and lack of optic chiasm (dotted circles) in a regenerating *arrowhead* RNAi animal are shown on the left. In the middle, the graph shows normal NME and NMC numbers in regenerating *arrowhead* RNAi animals. On the right, mappings show the distribution of NMEs and NMCs in an idealized visual system. n indicates the total number of animals mapped. (**G**) *slit*$^{+}$*; arrowhead*$^{+}$ cells (white arrows) in ectopic anterior brain commissures in a *notum* RNAi animal coinciding with ectopic optic chiasm. (**H**) scRNA-sequencing analysis of cells from region depicted in the cartoon. t-distributed stochastic neighbor embedding (t-SNE) representation of clustered cells (dots) colored based on planarian cell types is shown on the left. t-SNE representation of clustered muscle cells expressing *notum; frizzled 5/8-4* (red) is shown in the middle. t-SNE plot colored by expression of the transcription factor *soxP-5* is shown on the right. (**I** and **J**) Total NME and NMC numbers in uninjured (I) or regenerating (J) RNAi animals. (**K**) Visual system in presence of normal (middle) or severely reduced (right, *soxP-5**) numbers of NMEs and NMCs. (**L**) The number of axon bundles at the midline (left) and the number of NMEs and NMCs (right) in *soxP-5** RNAi animals. (**M**) Model summarizing the findings (see text for details). Dark pink arrows or dots, NMEs; light pink arrows or dots, NMCs. Scale bars are 50 μm in (A), (D) to (G), and (K).

Because NMEs and NMCs can be formed in the absence of eyes (Fig. 3D), we inhibited both *ovo* and PCGs at the same time to determine whether the ectopic positioning of NMEs and NMCs is still observed in PCG RNAi conditions in the absence of eye formation (Fig. 5, C and D, and fig. S9, C and D). In both uninjured and regenerating RNAi animals, NMEs and NMCs were mispositioned regardless of the presence (in control) or absence (in *ovo* RNAi) of new photoreceptor neurons. Taken together, our results indicate that PCGs constitute a positional information system that instructs the positioning of NMEs and NMCs independently of eye cells, providing a mechanism to place cells with candidate guidepost-like function in specific locations.

NMCs are absent in *tolloid* RNAi animals

Bmp signaling ligands have been shown to guide commissural axons of the spinal cord in vertebrates ventrally, through repulsion (*51*, *52*). Inhibition of the planarian *bmp4* homolog did not grossly affect the pathfinding of visual axons during regeneration (*53*). However, inhibition of *tolloid*, a metalloproteinase that can cleave the Bmp inhibitor Chordin in vertebrates and *Drosophila* (*54*, *55*), resulted in severe defects in the trajectory of photoreceptor axons (*53*). In regenerating *tolloid* RNAi animals, axons were more defasciculated and were not able to cross the midline, resulting in disorganized posterior projections (Fig. 6, A and B, and fig. S10A). The midline cues *slit* and *netrin-2* were normally expressed in these animals (fig. S10B). NMEs and NBCs were still present, but NMCs were completely absent in regenerating *tolloid* RNA animals (Fig. 6A). NMCs were also severely reduced in *tolloid* RNAi animals during eye regeneration after double-eye resection and during homeostasis (Fig. 6C and fig. S10C). Under all of these conditions, photoreceptor axons showed higher levels of defasciculation.

We next transplanted eyes into *tolloid; ovo* RNAi recipients, which lack NMCs. Independently of the presence or absence of remaining axons (transplantation 3 or 10 days after double-eye resection), wild-type photoreceptor axons were not able to robustly project in a fasciculated bundle or cross the midline in *tolloid; ovo* RNAi recipients (Fig. 6D and fig. S10D). These results are consistent with the possibility that NMCs can attract and/or stabilize axons to the choice points and facilitate axon bundle formation. Similar results were observed when transplanting a wild-type eye into a *netrin-1; netrin-2; ovo* RNAi recipient or when transplanting an eye from a *DCC* RNAi animal into an *ovo* RNAi recipient (fig. S10E). Metalloproteinases have been shown to cleave the DCC receptor (*56*), raising the possibility of direct or indirect interaction between Tolloid and DCC in planarians. The midline defect in *tolloid* RNAi animals indicates that additional *tolloid*-dependent processes likely exist and affect midline axonal crossing.

arrowhead is required for NBC specification and optic chiasm formation

Specific ablation of CD44-expressing chiasm neurons in mice demonstrated a role for these cells in the formation of the optic chiasm during embryonic development (*57*). Cell-specific ablation strategies are not currently available in planarians. However, in some instances, RNAi of transcription factor–encoding genes involved in cell-fate specification can be used to deplete animals of specific cell populations (*34*, *42*, *58*, *59*). Inhibition of *arrowhead*, which is expressed medially to the cephalic ganglia and encodes a Lim homeodomain protein, resulted in failed regeneration of the optic chiasm and defective axonal crossing at the midline (*60*). *arrowhead* RNAi did not affect the expression of normal midline guidance cues, such as *slit* (*60*) and *netrin-2* (fig. S11A). Visual axonal projections did not overtly follow other axon bundles at the anterior commissure in regenerating *arrowhead* RNAi animals (fig. S11B). We hypothesized that *arrowhead* might be required for the regeneration of NBCs at the anterior commissure and thereby affect midline crossing of visual axons. Indeed, NBCs expressed *arrowhead* both in intact and regenerating animals (Fig. 6E and fig. S11C), and *arrowhead* RNAi animals failed to regenerate NBCs (Fig. 6F). However, we cannot rule out the possibility that other cells specified by *arrowhead* also contribute to optic chiasm formation. Numbers of NMEs and NMCs were not affected in *arrowhead* RNAi animals during head regeneration (Fig. 6F) or after double-eye resection (fig. S11D). Intact *arrowhead* RNAi animals that lost all NBCs during normal tissue turnover were still able to properly rewire the visual system after double-eye resection (fig. S11D), suggesting that NBCs might be dispensable for maintenance of the optic chiasm. Importantly, visual axons after double-eye resection perdure for around 3 days, which gives sufficient time for new photoreceptors to regenerate and project their axons to the brain targets following the preexisting axonal tract. Alternatively, the photoreceptor axons could also be influenced by the existing brain commissure. In addition, RNAi animals with duplicated optic chiasms [*notum* RNAi animals; (*32*)] also had ectopic anterior brain commissures with a duplication of NBCs (Fig. 6G).

soxP-5 is required for NME and NMC specification

To examine the transcriptomic signature of NMEs and NMCs, we performed single-cell RNA sequencing (scRNA-seq) using Drop-seq (*61*) on cells isolated from the region containing the visual system in intact animals (Fig. 6H). Cells were clustered using Seurat into 28 clusters, each of which could be assigned to a certain tissue class on the basis of tissue-specific gene expression [(*62*); Fig. 6H]. We recovered eight *notum*$^{+}$*; frizzled5/8-4*$^{+}$ cells that were negative for brain markers (*pc2* and *ChAT*) and anterior pole markers (*foxD*) and that clustered together with muscle cells (Fig. 6H, fig. S12A, and table S1). We performed single-cell differential expression (SCDE) analysis with these cells, comparing their transcriptomes with the transcriptomes of other muscle cells, and identified enriched transcripts (table S2). A similar analysis was also performed for NBCs (table S3). One of the few transcription factors expressed in NMEs and NMCs in the scRNA-seq data was *soxP-5*. (*soxP-5* expression was difficult to detect by FISH.) *soxP-5* was also broadly expressed in muscle cells, including muscle progenitors, but not in photoreceptor neurons or other cell types (Fig. 6H, fig. S12A, and table S2).

RNAi of *soxP-5* led to a significant decrease in the numbers of both NMEs and NMCs in uninjured (Fig. 6I) or double-eye-resected animals (fig. S12B), indicating that it is required for production of normal numbers of these specialized muscle cells. After head amputation, *soxP-5* RNAi animals regenerated head blastemas that were normal in macroscopic size, appearance, and brain architecture (fig. S12C). However, these blastemas displayed lower numbers of NMEs and NMCs—in most *soxP-5* RNAi animals, multiple NMEs and NMCs did regenerate, but a minority of animals had very few or no detectable NMEs and NMCs (Fig. 6J). In eye-resected *soxP-5* RNAi animals, the axonal projection pattern remained mostly unaffected, as expected, because new axons could follow existing axonal tracts in their development (fig. S12B). In regeneration after amputation, the visual system must form de novo, allowing assessment of pattern in animals with reduced NMEs and NMCs. Whereas most *soxP-5* RNAi animals regenerated the stereotypical axonal projection pattern, around 20% of the *soxP-5* RNAi animals exhibited defects in this pattern (labeled as *soxP-5** in Fig. 6, K and L, and fig. S12, D to F). *soxP-5** RNAi animals showed ectopic axon bundles at the optic chiasm, abnormal fasciculation, and misrouting of the axon bundles at the choice points (Fig. 6, K and L, and fig. S12, D and E). We found that the *soxP-5** animals were those that had a severe reduction in NMEs and NMCs: *soxP-5** RNAi animals had significantly fewer NMEs ($p = 0.0029$) and NMCs ($p = 0.02$) than did *soxP-5* RNAi animals with a relatively normal axonal projection pattern (fig. S12F). Normal numbers of muscle cells were observed in intact (fig. S12B) and regenerating *soxP-5** RNAi animals (fig. S13, A and B), indicating that *soxP-5* is not essential for general muscle

regeneration or maintenance per se. Furthermore, NBCs, as well as other neuronal subsets, regenerated normally in *soxP-5** RNAi animals, indicating a specific effect on the muscle cells associated with the photoreceptor axons (fig. S13, C and D). The correspondence between low numbers or absence of NMEs and NMCs and the quality of the regenerated axonal projections provides further evidence for a guidepost-like role for these muscle cells. Our results suggest that NME, NMC, and NBC guidepost-like cells, together with constitutive guidance cues expressed broadly in the animal, contribute to the precise overall assembly and pattern of the visual system.

Discussion

Orchestration of precisely wired neural circuits relies on multiple developmental processes. Axon pathfinding is a highly precise process, and conserved mechanisms of guidance are used by different organisms. Guidance mechanisms can act in concert and include long-range or local chemoattractants and chemorepellents (guidance cues), transient cell-cell interactions (guidepost cells), and fasciculation of axons with pioneer or preceding axon tracts that form an initial scaffold. Once the circuit is established, some of these mechanisms become dispensable for the maintenance of pattern. The transient nature of these developmental mechanisms creates a problem for repair after injury: If an injury removes a patterned circuit, the information for bringing the circuit back would be lost and might result in irreparable damage to the nervous system, even if the mechanisms for new neuron production existed. However, functional regeneration of the nervous system in some animals suggests that mechanisms for de novo restoration of the precise pattern of neuronal circuits in adults must exist.

We identified a population of muscle cells expressing *notum* and *frizzled 5/8-4* located in close proximity to axon tracts at key locations in the planarian visual system. The tight association and precise positioning of these muscle cells suggest that they have a role in the assembly of the planarian visual system. This finding led us to investigate the possible existence of guidepost-like cells in nervous system regeneration. Guidepost cells in other organisms are often transient neurons or glia. These planarian guidepost-like cells were not only present in regeneration but also constitutively maintained in the adult. It is atypical for a muscle cell to exert a guidepost-like role. However, planarian muscle cells are not only contractile, they also secrete extracellular matrix proteins (*63*) and harbor positional information required for regeneration and tissue turnover (*27*, *47*). In addition, specialized subsets of muscle cells serve as anterior (*64*) and posterior organizers (*65*, *66*). Our findings suggest that non-neuronal connective tissue–like cells can promote neuronal circuit formation in regeneration.

Our results suggest that *notum*$^{+}$*; frizzled 5/8-4*$^{+}$ muscle cells can locally attract and/or stabilize photoreceptor axons, facilitating formation of visual axon bundles and possibly helping with axon sorting at choice points (Fig. 6M). Axon-axon interactions are critical for organizing axons into bundles and are required for proper topographic pathfinding. The mechanisms involved in axon bundle formation described in other organisms include either interactions of adhesion molecules expressed on the axons or repulsive interactions between axons and surrounding tissues [reviewed in (*67*)]. The molecules required in planarians to facilitate this process remain unknown, although members of the immunoglobulin or cadherin superfamilies are appealing candidates to play roles in this process. Guidepost cells in other systems exert their function in guiding pioneer axon tracts by secreting attractive or repulsive guidance cues. Inhibition of several genes encoding homologs of canonical guidance cues did not result in defective attraction of the visual axons toward the planarian guidepost-like cells, suggesting that different molecules are at play or that there exists some redundancy within the system. Our findings suggest a model in which guidepost-like cells act in concert with constitutively active global axon guidance cues to facilitate precision in the regeneration of the visual system.

Our study also identified two different transcription factors required for the specification of distinct subsets of cells with guidepost-like attributes. *arrowhead*, a Lim domain homeobox gene, is involved in the specification of NBCs, whereas *soxP-5* is required for the specification of both NMEs and NMCs. NMEs and NMCs show differences not only regarding their position but also, based on our observations from *tolloid* RNAi animals, in the expression of different programs to exert their function. NMCs were sensitive to *tolloid* inhibition, whereas NMEs remained unaffected. Some of these genes likely regulate systems beyond the NMEs and NMCs. Previously identified targets of the Tolloid protease include the BMP inhibitor Chordin, a homolog for which has not been found in planarians. However, increasing evidence suggests that members of the Astacin-like family of metalloproteases can also process extracellular matrix components and certain TGFβ-family ligands, and they are likely candidates for modulating the activity of guidance cues or their receptors (*56*).

How intermediate target or guidepost cells in different organisms are specified at precise locations has not been fully elucidated. Our observations suggest a model in which muscle guidepost-like cells are coarsely specified in a permissive region proximal to the eyes, which is, at least in part, defined by an axial PCG expression map (Fig. 6M). Density maps for the muscle guidepost-like cells suggest that these cells might be maintained only at the positions where photoreceptor axons are present, as indicated by their tight association with photoreceptor axons. This feature may allow dynamic maintenance of muscle guidepost-like cells in precise locations and allow regeneration of these cells in the face of an unlimited array of regenerative challenges.

Strategies for neuronal patterning in the absence of embryo-specific contexts likely will also be crucial to overcome challenges in regenerative medicine. By studying a naturally occurring context of de novo formation of a neural system in regeneration, we found that small populations of cells, extrinsic to the system itself, provide information to help precisely pattern and assemble axonal projections in regeneration. Our results expand the diversity of strategies used by different organisms to organize and properly build the nervous system, especially in the context of injury repair.

Materials and methods

Animal husbandry

S. mediterranea clonal asexual strain CIW4 animals, starved for 7 to 14 days before experimentation, were used for all experiments except for in situ hybridizations in embryos (Fig. 1). All animals used were healthy, not previously used in other procedures, and were of wild-type genotype. Animals were cultured in plastic containers or petri dishes for experiments, in 1x Montjuic water (1.6 mmol/liter NaCl, 1.0 mmol/liter $CaCl_2$, 1.0 mmol/liter $MgSO_4$, 0.1 mmol/liter $MgCl_2$, 0.1 mmol/liter KCl, and 1.2 mmol/liter $NaHCO_3$ prepared in Milli-Q water) at 20°C in the dark. Animals were fed blended calf liver. *S. polychroa* lines were generated by amputation of a starter animal, and the line was initially propagated through successive rounds of amputation. After generating a large colony, sexual reproduction ensued. Animals were fed homogenized calf liver once a week and cleaned twice weekly. Animals were kept in the dark at 20°C and maintained in 1x Montjuic water. Egg capsules were collected and staged daily (*68*).

Replication, size estimation, and randomization

At least two independent FISH and immunostaining experiments with a minimum of three animals per experiment were performed for the characterization of muscle guidepost cells in intact and regenerating adults and in embryos. For RNAi phenotype characterization, numbers of animals used in each staining are indicated in each panel. No sample size estimation was performed. Animals for all experiments were randomly selected from a large

collection of clonal animals. All animals have been included in statistical analyses, and no exclusions have been done. Images were randomized before quantification.

Gene nomenclature

Genes that encode proteins with a clear domain structure have been assigned a name accordingly but have also been identified using a transcriptome contig id number to facilitate identification (figs. S7 and S8).

Drop-seq clustering and SCDE analysis

To obtain transcriptomes for NMEs and NMCs, data from targeted single-cell sequencing of the planarian brain [(62), GEO accession number GSE111764, BrainClusteringDigitalExpressionMatrix.dge.txt.gz] was used to identify cells positive for expression of *notum* (dd_Smed_v4_24180_0_1) and *fz5/8-4* (dd_Smed_v4_11823_0_1) and negative for expression of the anterior pole marker *foxD* (dd_Smed_v4_23249_0_1) and the neuronal markers *ChAT* (dd_Smed_v4_6208_0_1) and *pc2* (dd_Smed_v4_1566_0_1) (log-scale expression of 0.5 for all genes). To identify these cells within the single-cell sequencing data, the Seurat function WhichCells [subset.name=contig ID, accept.low=0.5] was used with Seurat package, v2.2 (69). Fifteen cells satisfied all gene expression thresholds (Cells_Head1_AAGTCTCACGCC, Cells_Head1_CAGACCTTCCCC, Cells_Head1_CGTGACTAAGAA, Cells_Head1_GGCGTGGTGACN, Cells_Head1_TAAATTCGATAG, Cells_Head1_TTTACTTTCGAT, Cells_Head2_AACGCCATTTCC, Cells_Head2_AGTATGAATATG, Cells_Head2_AGTCACTAACAA, Cells_Head2_CACCGTGTACTA, Cells_Head2_CCAGATAACGCA, Cells_Head2_CGTAACTATCGT, Cells_Head2_GGTTACAGCTTT, Cells_Head2_GTTCCATGAAGN, Cells_Head2_TGCTGTGCATCT). Of these 15 cells, eight cells were assigned a muscle identity in (62) (Cells_Head2_AACGCCATTTCC, Cells_Head2_AGTATGAATATG, Cells_Head2_AGTCACTAACAA, Cells_Head2_CACCGTGTACTA, Cells_Head2_CCAGATAACGCA, Cells_Head2_CGTAACTATCGT, Cells_Head2_GGTTACAGCTTT, Cells_Head2_TGCTGTGCATCT). Additional data on identified cells can be found in table S1 of (62). An expression matrix for all cells identified as muscle in the targeted brain sequencing from (62) was generated, tagging the eight muscle cells above as positive and the remaining 522 muscle cells as negative. This expression matrix was used as input for the R package SCDE (70). SCDE analysis revealed a list of genes enriched in these putative guidepost cells. Genes were ranked by their "conservative expression" value. To obtain transcriptomes for *notum*$^+$ brain neurons, cells identified as neural from the targeted brain sequencing were combined with all cells identified as neural in the principal single-cell sequencing from (62). Cells positive for expression of both *ChAT* (log-scale expression of 2.5) and *notum* (log-scale expression of 2) were identified, as above. Thirteen cells satisfied both gene expression thresholds (Cells_Head_TCATCCACGCTT, Cells_Head_ACAATGTTTGGT, Cells_Head_GTCAGACTCAGN, Cells_Head_CCTGTCGGCTCN, Cells_Head_AGGGCGTAGTAA, Cells_Head2_GTCGTCGTTGCG, Cells_Head2_GGCCCGAGGATG, Cells_Whole_CGAACCAATAGT, Cells_Pharynx_AGTACAATGTGN, Cells_Head2_AAACCAAGCCAG, Cells_Head2_AGTTAGGCACAN, Cells_Pharynx_GGTAGGTTATCG, Cells_Pharynx_AAGTAGGCATCG). Of these 13 cells, three had been isolated from the planarian pharynx (Cells_Pharynx_AGTACAATGTGN, Cells_Pharynx_GGTAGGTTATCG, Cells_Pharynx_AAGTAGGCATCG) and were thus discarded. Additional data on identified cells can be found in table S1 of (62). An expression matrix for all *ChAT*$^+$ neurons (2103 cells) was generated, tagging the name of each neuron as positive or negative for expression of *notum*. This expression matrix was used as input for analysis by SCDE, as above, revealing a list of genes enriched in *notum*$^+$ neurons. Genes were ranked by their "conservative expression" value.

For all cells analyzed in table S1, unique molecular identifier (UMI) counts were log normalized by dividing by the total number of UMIs per cell, then multiplying by 10,000. All calculations were performed in log space [i.e., ln(UMIs-per-10,000+1)]. The 50 transcripts with the highest average normalized expression in the eight putative guidepost-like cells were then identified, and their expression was compared with 10 randomly chosen non-guidepost muscle cells (chosen from cluster 9 of the Drop-seq data), 10 randomly chosen ciliated epidermis cells (chosen from cluster 19 of the Drop-seq data), and 10 randomly chosen neural cells (chosen from clusters 0, 1, 5, 17, 20, 23, 24, 25, 26, 27, and 28 of the Drop-seq data). Averages indicate the average normalized UMI counts for each group of cells. Muscle enrichment and broad expression determinations were made using the online resource digiworm.wi.mit.edu from (62). Namely, genes enriched in clusters 7, 13, 14, or 16 of the main clustering data from (62), but not broadly expressed across most clusters, were designated muscle enriched. All of the top 50 transcripts expressed in the putative guidepost cells were either muscle enriched or broadly expressed, including in muscle.

Gene cloning

Homologs of guidance cues and transcription factors in planarians were cloned using the following primers:

roboC (dd_Smed_v4_7921_0_1), fwd: 5′ atggtgccattgtcccgg; rv: 5′ aaccgagagttgccggtg;

roboD (dd_Smed_v4_14150_0_1), fwd: 5′ tgctcaatcgtcagataccg; rv: 5′ accgggaattcgaaaagact

unc-5A (dd_Smed_v4_10380_0_1), fwd: 5′ ttgctcctagcggtcttcat; rv: 5′ tgtacgcggaattgctactg

unc-5B (dd_Smed_v4_10585_0_1), fwd: 5′ tcttgagccacaacccttt; rv: 5′ ccagttcgatatccgaagga

unc-5C (dd_Smed_v4_10730_0_1), fwd: 5′ccaactcgggaaattgaaga; rv: 5′ ccgaaacaaaaggtggagaa

unc-5D (dd_Smed_v4_16435_0_1), fwd: 5′ ccctcaaggaacaaaatgga; rv: 5′ aaatttcccaatcgggtttc

ephR-1 (dd_Smed_v4_16483_0_1), fwd: 5′ ccgatcacttttcagccaat; rv: 5′ gtgaggttggctgattccat

ephR-2 (dd_Smed_v4_16928_0_1), fwd: 5′ gttcctctgatgtgcccagt; rv: 5′ agatccggcatgaatctgac

ephrin (dd_Smed_v4_16552_0_1), fwd: 5′ tccagcaagatatgccgata; rv: 5′ tgctgaaaaactgataattgaaaca.

ephrin (dd_Smed_v4_10687_0_1), fwd: 5′agtcattcacggtccaggc; rv: 5′ cccatgaaaaacaggttcaaaag

arrowhead (dd_Smed_v4_47123_0_1), fwd: 5′ gttgcaaagctagctcaatttca; rv: 5′ acgggatatggattaacttgaca

plexin (dd_Smed_v4_11934_0_1), fwd: 5′ gcgagtgtggtgggaaaaat; rv: 5′ ccaagacgcaccaagaacaa

semaphorin-1 (dd_Smed_v4_12018_0_1), fwd: 5′acctgaaatccctttcactcgt; rv: 5′ tccagctgtgtaagagagga

soxP-5 (dd_Smed_v4_9050_0_1) accession number: GenBank JX010525.1.

All constructs were cloned from cDNA into the pGEM vector (Promega). These constructs were used to synthesize RNA probes and double-stranded RNA (dsRNA) for RNAi experiments.

RNAi

For RNAi experiments, dsRNA was synthesized by in vitro transcription reactions (Promega) using polymerase chain reaction (PCR)–generated templates with flanking T7 promoters, followed by ethanol precipitation, and annealed after resuspension in water. The concentration of dsRNA varied in each prep between 4 and 7 μg/ml. dsRNA was then mixed with planarian food (liver) (36), and 2 μl of this mixture per animal (liver containing dsRNA) was used for feedings. For the guidance cue screen (fig. S8), animals were fed six times in 3 weeks. For the PCG screen (Fig. 5 and fig. S9), animals were fed between 6 and 10 times in 3 to 5 weeks until phenotype was observed. *ovo* RNAi animals for homeostasis experiments (Figs. 3 and 5 and fig. S9) were fed 8 to 12 times during a 4- to 6-week period. *ovo* RNAi regeneration experiments (eye resection and decapitation; Fig. 3) animals were fed eight times in 4 weeks. *ovo* RNAi animals for transplantation experiments were fed eight times in 4 weeks (Fig. 4 and figs. S5 and S6). *ndk; ndl4* RNAi animals were fed until posterior eyes appeared (8 to 12 feedings). For RNAi "wear off" experiments, animals were not fed for 2 months before a midline sagittal cut was performed (Fig. 5). *arrowhead* RNAi animals (Fig. 6 and fig. S11) were fed six times over a period of 3 weeks. *tolloid* RNAi animals were fed only once for regeneration

experiments and twice for homeostasis experiments (Fig. 6 and fig. S10). *soxP-5* RNAi animals were fed eight times in a 4-week period of time (Fig. 6 and figs. S12 and S13). All feedings were performed every 3 days. In all cases, animals were fixed 7 days after the last feeding. For regeneration experiments, animals were amputated into three pieces (head, trunk, and tail pieces) 1 week after the last RNAi feeding. Seven days after amputation, trunk pieces were scored and fixed for further analysis.

FISHs and immunostainings

RNA probes were synthesized, and whole-mount FISH was performed (*36*). Briefly, animals were killed in 5% *N*-acetyl cysteine (NAC) and treated with proteinase K (2 μg/ml). After overnight hybridizations, samples were washed twice in each of prehybridization buffer, 1:1 prehybridization-2X SSC, 2X SSC, 0.2X SSC, and phosphate-buffered saline (PBS) with Triton-X (PBST). Subsequently, blocking was performed in 10% Western Blocking Reagent (Roche, 11921673001) PBST solution for digoxigenin (DIG) probes, or in 5% horse serum and 5% casein for dinitrophenol (DNP) and fluorescein (FITC) probes. Antibody washes were then performed for 1 hour followed by tyramide development. Peroxidase inactivation with 1% sodium azide was done for 90 min at room temperature. Brightfield images were taken with a Zeiss discovery microscope. Fluorescent images were taken with a Leica SP8 confocal microscope. Colocalization analyses of FISH signals were performed using Fiji/ImageJ. Three-dimensional (3D) reconstruction for the movies was performed using Imaris 3/4D Image Visualization and Analysis software. For each channel, histograms of fluorescence intensity were used to determine the cut-off between signal and background. All FISH images shown are representative of all images taken in each condition and are maximal intensity projections. All images, unless otherwise indicated, are anterior up. For immunostainings with anti-Arrestin antibody (VC-1) or anti-muscle antibody (6G10), animals were fixed as for in situ hybridizations, blocked in 3% bovine serum albumin (BSA)–PBST or in 10% Western Blocking Reagent (Roche, 11921673001) PBST solution, respectively, for 1 hour and then stained with the antibody of interest. The anti-muscle mouse monoclonal antibody 6G10 (RRID: AB_2619613) (*71*) was used in a 1:1000 dilution, the anti-Arrestin mouse monoclonal antibody VC-1 was used in a 1:5000 dilution, the anti-Arrestin rabbit polyclonal antibody and the anti–α tubulin antibody (Lab Vision catalog no. MS-581, RRID: AB_144075) were used in a 1:500 dilution, and anti-mouse Alexa conjugated antibodies (Life Tech) were used in a 1:500 dilution. The pool of probes used for labeling muscle cells included *troponin*, *tropomyosin*, *colF-2*, and *colF-10* (Fig. 1 and fig. S2).

S. polychroa (*Spol*) FISH

Whole-mount FISH was performed as described above for *S. mediterranea* with the following modifications: (i) Stage 2 to 5 embryos were dissected out of egg capsules and fixed in 4% formaldehyde for 2 hours. Embryos were washed in PBST-0.5% for 10 min and dehydrated in 25, 50, 75, and 100% PBST: methanol for 10 min. Fixed embryos were not bleached and were stored at −20°C. Proteinase K treatment was extended to 20 min. (ii) Stage 6 and 7 embryos were dissected out of egg capsules and fixed in 4% formaldehyde for 1 hour. Embryos were washed in PBST-0.5% for 10 min and dehydrated in 25, 50, 75, and 100% PBST:methanol for 10 min each. Embryos were washed in PBST-0.5% for 10 min and dehydrated in 25, 50, 75, and 100% PBST: methanol for 10 min. Fixed embryos were not bleached and were stored at −20°C. Proteinase K treatment was extended to 15 min. (iii) Hatchlings were treated with 5% NAC PBS with gentle rotation for 5 min and then fixed in 4% formaldehyde for 1 hour. Hatchlings were then washed in PBST-0.5% for 10 min and dehydrated in 25, 50, 75, and 100% PBST: methanol for 10 min. Hatchlings were stored at −20°C and bleached in formamide bleaching solution for 1 hour at room temperature. Proteinase K treatment was extended to 12 min. All embryos were mounted in Vectashield between two coverslips and both sides were imaged on a Leica SP8 confocal microscope.

Eye resections and eye transplantations

To selectively resect eyes, animals were placed on moist filter paper on a cold block to limit movement, and the tip of a microsurgery blade was used to remove eyes. For eye transplants, after anesthesia using 0.2% chloretone in planarian H_2O for 2 min, a thin slit cut was made to desired locations of the recipient animals and a small hole was generated by gently moving the surgical blade up and down within the slit cut. Recipient animals were washed in Holtfreter's solution for 2 min and rested in 0.2% chloretone for 2 min, briefly washed in Holtfreter's solution, and transferred on the cold block to introduce the excised eyes. Eye resections were performed using a dissecting microscope by trimming the pigmented tissue around the eyes with a surgical blade, leaving only the white area and the visible optic cup of the eyes. The pigmented ventral side of this tissue was also trimmed away before transplantation. The eyes were gently pushed inside the previously generated holes in the recipient animals. Transplanted animals were immobilized using Type IV, 5% ultra-low melting agarose (Sigma) on top of WhatmanTM (GE Healthcare, Life Sciences) filter paper. Solidified gel was covered using Rasta Royale ultrathin rolling paper soaked in Holtfreter's solution. Transplanted animals were kept in 10°C overnight and recovered by cutting the gel around and also on top of the animal. Animals were placed in planarian H_2O and kept at 22°C for recovery.

Quantifications and statistical analysis

Total numbers of NMEs and NMCs were counted based on expression of the markers *notum* and *fz5/8-4*, the location of their nuclei on the dorsal-ventral axis, and by using, as a reference, the brain architecture with a 4′,6-diamidino-2-phenylindole (DAPI) nuclei staining. Total numbers of muscle guidepost cells were also counted and graphed relative to the total length of the animal expressed in micrometers.

NME and NMC cell position and density maps were generated by building an idealized planarian visual circuit and positioning it in relation to the brain commissure and the anterior pole determined by DAPI staining and *notum* expression, respectively. This trace was placed on four (for *ovo* RNAi only) or nine (for all other conditions) equal zones in the Adobe Illustrator software and the approximate positions of NMEs and NMCs in relation to the photoreceptor axons, brain commissures, and anterior poles was mapped manually. Heatmaps indicating relative positions of NMEs and NMCs were generated by counting the NMEs and NMCs that fall into each of the nine zones. The range was determined by the highest and the lowest total number of cells that were counted in each zone.

Axonal trajectories were manually traced in the Adobe Illustrator software after transferring a maximum intensity projection of the image containing the planarian visual circuitry into a new document. NMEs and NMCs were mapped onto the traces based on their position in relation to the visual axons using FIJI's channels tool, determining the dorsal-ventral position for each cell. To determine the significance of axonal interactions with NMEs, NMCs, and NBCs (Fig. 2), axons from 6 to 10 separate eyes, for each group, were traced manually in the same manner described above. The trace maps were placed on a circular area that is divided into 36 equal sections (10° each). NME, NMC, and NBC cell distributions in reference to the center point of the circle (eye) were mapped, and angular median values and 95% confidence intervals for cell positions were calculated for each group. NME, NMC, and NBC numbers and the number of axonal intersections in each pie shaped section shown in Fig. 2 were counted, and bar graphs were generated using GraphPad Prism software after linearizing the circular arenas. Coincidence of NMEs, NMCs, and NBCs and axons was

represented by a black box around the area they all occupy (fig. S3). Local axonal trajectories were traced in relation to NMEs and NMCs in "d7 after eye transplant" animals, and axons that extend within two NME or NMC cell diameters were mapped in circular fields. An overlay map of these circular fields was generated to build a density map of axons that project to the vicinity of NMEs and NMCs (Fig. 4 and fig. S6).

One-way analysis of variance (ANOVA) test followed by Dunnett's multiple comparison test was used when analyzing more than two conditions. Unpaired Student's *t* test was used when comparing two conditions. Mean ± SD is shown in all graphs.

REFERENCES AND NOTES

1. J. S. Edwards, S. W. Chen, M. W. Berns, Cercal sensory development following laser microlesions of embryonic apical cells in *Acheta domesticus*. *J. Neurosci.* **1**, 250–258 (1981). doi: 10.1523/JNEUROSCI.01-03-00250.1981; pmid: 7264719
2. H. Hutter, Extracellular cues and pioneers act together to guide axons in the ventral cord of *C. elegans*. *Development* **130**, 5307–5318 (2003). doi: 10.1242/dev.00727; pmid: 13129845
3. M. Klose, D. Bentley, Transient pioneer neurons are essential for formation of an embryonic peripheral nerve. *Science* **245**, 982–984 (1989). doi: 10.1126/science.2772651; pmid: 2772651
4. S. H. Pike, E. F. Melancon, J. S. Eisen, Pathfinding by zebrafish motoneurons in the absence of normal pioneer axons. *Development* **114**, 825–831 (1992). pmid: 1618146
5. A. L. Kolodkin, M. Tessier-Lavigne, Mechanisms and molecules of neuronal wiring: A primer. *Cold Spring Harb. Perspect. Biol.* **3**, a001727 (2011). doi: 10.1101/cshperspect.a001727; pmid: 21123392
6. E. T. Stoeckli, Understanding axon guidance: Are we nearly there yet? *Development* **145**, dev151415 (2018). doi: 10.1242/dev.151415; pmid: 29759980
7. M. Tessier-Lavigne, C. S. Goodman, The molecular biology of axon guidance. *Science* **274**, 1123–1133 (1996). doi: 10.1126/science.274.5290.1123; pmid: 8895455
8. T. W. Yu, C. I. Bargmann, Dynamic regulation of axon guidance. *Nat. Neurosci.* **4** (Suppl), 1169–1176 (2001). doi: 10.1038/nn748; pmid: 11687826
9. D. Bentley, M. Caudy, Pioneer axons lose directed growth after selective killing of guidepost cells. *Nature* **304**, 62–65 (1983). doi: 10.1038/304062a0; pmid: 6866090
10. D. Bentley, H. Keshishian, Pathfinding by peripheral pioneer neurons in grasshoppers. *Science* **218**, 1082–1088 (1982). doi: 10.1126/science.218.4577.1082; pmid: 17752851
11. D. L. Chao, L. Ma, K. Shen, Transient cell-cell interactions in neural circuit formation. *Nat. Rev. Neurosci.* **10**, 262–271 (2009). doi: 10.1038/nrn2594; pmid: 19300445
12. P. Squarzoni, M. S. Thion, S. Garel, Neuronal and microglial regulators of cortical wiring: Usual and novel guideposts. *Front. Neurosci.* **9**, 248 (2015). doi: 10.3389/fnins.2015.00248; pmid: 26236185
13. T. Hirata *et al*., Guidepost neurons for the lateral olfactory tract: Expression of metabotropic glutamate receptor 1 and innervation by glutamatergic olfactory bulb axons. *Dev. Neurobiol.* **72**, 1559–1576 (2012). doi: 10.1002/dneu.22030; pmid: 22539416
14. G. López-Bendito *et al*., Tangential neuronal migration controls axon guidance: A role for neuregulin-1 in thalamocortical axon navigation. *Cell* **125**, 127–142 (2006). doi: 10.1016/j.cell.2006.01.042; pmid: 16615895
15. M. Niquille *et al*., Transient neuronal populations are required to guide callosal axons: A role for semaphorin 3C. *PLOS Biol.* **7**, e1000230 (2009). doi: 10.1371/journal.pbio.1000230; pmid: 19859539
16. Y. Sato, T. Hirata, M. Ogawa, H. Fujisawa, Requirement for early-generated neurons recognized by monoclonal antibody lot1 in the formation of lateral olfactory tract. *J. Neurosci.* **18**, 7800–7810 (1998). doi: 10.1523/JNEUROSCI.18-19-07800.1998; pmid: 9742149
17. F. Bielle *et al*., Slit2 activity in the migration of guidepost neurons shapes thalamic projections during development and evolution. *Neuron* **69**, 1085–1098 (2011). doi: 10.1016/j.neuron.2011.02.026; pmid: 21435555
18. K. Ito *et al*., Semaphorin 3F confines ventral tangential migration of lateral olfactory tract neurons onto the telencephalon surface. *J. Neurosci.* **28**, 4414–4422 (2008). doi: 10.1523/JNEUROSCI.0372-08.2008; pmid: 18434520
19. T. Kawasaki, K. Ito, T. Hirata, Netrin 1 regulates ventral tangential migration of guidepost neurons in the lateral olfactory tract. *Development* **133**, 845–853 (2006). doi: 10.1242/dev.02257; pmid: 16439477
20. T. Nomura, J. Holmberg, J. Frisen, N. Osumi, Pax6-dependent boundary defines alignment of migrating olfactory cortex neurons via the repulsive activity of ephrin A5. *Development* **133**, 1335–1345 (2006). doi: 10.1242/dev.02290; pmid: 16510508
21. R. R. Bernhardt, C. K. Patel, S. W. Wilson, J. Y. Kuwada, Axonal trajectories and distribution of GABAergic spinal neurons in wildtype and mutant zebrafish lacking floor plate cells. *J. Comp. Neurol.* **326**, 263–272 (1992). doi: 10.1002/cne.903260208; pmid: 1479075
22. C. Klämbt, J. R. Jacobs, C. S. Goodman, The midline of the Drosophila central nervous system: A model for the genetic analysis of cell fate, cell migration, and growth cone guidance. *Cell* **64**, 801–815 (1991). doi: 10.1016/0092-8674(91)90509-W; pmid: 1997208
23. R. C. Marcus, R. Blazeski, P. Godement, C. A. Mason, Retinal axon divergence in the optic chiasm: Uncrossed axons diverge from crossed axons within a midline glial specialization. *J. Neurosci.* **15**, 3716–3729 (1995). doi: 10.1523/JNEUROSCI.15-05-03716.1995; pmid: 7751940
24. D. W. Sretavan, L. Feng, E. Puré, L. F. Reichardt, Embryonic neurons of the developing optic chiasm express L1 and CD44, cell surface molecules with opposing effects on retinal axon growth. *Neuron* **12**, 957–975 (1994). doi: 10.1016/0896-6273(94)90307-7; pmid: 7514428
25. R. Amamoto *et al*., Adult axolotls can regenerate original neuronal diversity in response to brain injury. *eLife* **5**, e13998 (2016). doi: 10.7554/eLife.13998; pmid: 27156560
26. C. E. Laumer *et al*., Spiralian phylogeny informs the evolution of microscopic lineages. *Curr. Biol.* **25**, 2000–2006 (2015). doi: 10.1016/j.cub.2015.06.068; pmid: 26212884
27. J. N. Witchley, M. Mayer, D. E. Wagner, J. H. Owen, P. W. Reddien, Muscle cells provide instructions for planarian regeneration. *Cell Rep.* **4**, 633–641 (2013). doi: 10.1016/j.celrep.2013.07.022; pmid: 23954785
28. K. Agata *et al*., Structure of the planarian central nervous system (CNS) revealed by neuronal cell markers. *Zool. Sci.* **15**, 433–440 (1998). doi: 10.2108/zsj.15.433; pmid: 18466009
29. K. Okamoto, K. Takeuchi, K. Agata, Neural projections in planarian brain revealed by fluorescent dye tracing. *Zool. Sci.* **22**, 535–546 (2005). doi: 10.2108/zsj.22.535; pmid: 15930826
30. F. Sakai, K. Agata, H. Orii, K. Watanabe, Organization and regeneration ability of spontaneous supernumerary eyes in planarians—eye regeneration field and pathway selection by optic nerves—. *Zool. Sci.* **17**, 375–381 (2000). pmid: 18494593
31. C. P. Petersen, P. W. Reddien, Polarized *notum* activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science* **332**, 852–855 (2011). doi: 10.1126/science.1202143; pmid: 21566195
32. E. M. Hill, C. P. Petersen, Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. *Development* **142**, 4217–4229 (2015). doi: 10.1242/dev.123612; pmid: 26525673
33. P. W. Reddien, The cellular and molecular basis for planarian regeneration. *Cell* **175**, 327–345 (2018). doi: 10.1016/j.cell.2018.09.021; pmid: 30290140
34. S. W. Lapan, P. W. Reddien, *dlx* and *sp6-9* control optic cup regeneration in a prototypic eye. *PLOS Genet.* **7**, e1002226 (2011). doi: 10.1371/journal.pgen.1002226; pmid: 21852957
35. M. L. Scimone, M. Srivastava, G. W. Bell, P. W. Reddien, A regulatory program for excretory system regeneration in planarians. *Development* **138**, 4387–4398 (2011). doi: 10.1242/dev.068098; pmid: 21937596
36. M. L. Scimone, L. E. Cote, P. W. Reddien, Orthogonal muscle fibres have different instructive roles in planarian regeneration. *Nature* **551**, 623–628 (2017). doi: 10.1038/nature24660; pmid: 29168507
37. M. L. Scimone *et al*., *foxF-1* controls specification of non-body wall muscle and phagocytic cells in planarians. *Curr. Biol.* **28**, 3787–3801.e6 (2018). doi: 10.1016/j.cub.2018.10.030; pmid: 30471994
38. S. A. LoCascio, S. W. Lapan, P. W. Reddien, Eye absence does not regulate planarian stem cells during eye regeneration. *Dev. Cell* **40**, 381–391.e3 (2017). doi: 10.1016/j.devcel.2017.02.002; pmid: 28245923
39. T. J. Petros, A. Rebsam, C. A. Mason, Retinal axon growth at the optic chiasm: To cross or not to cross. *Annu. Rev. Neurosci.* **31**, 295–315 (2008). doi: 10.1146/annurev.neuro.31.060407.125609; pmid: 18558857
40. T. Inoue *et al*., Morphological and functional recovery of the planarian photosensing system during head regeneration. *Zool. Sci.* **21**, 275–283 (2004). doi: 10.2108/zsj.21.275; pmid: 15056922
41. K. D. Atabay, S. A. LoCascio, T. de Hoog, P. W. Reddien, Self-organization and progenitor targeting generate stable patterns in planarian regeneration. *Science* **360**, 404–409 (2018). doi: 10.1126/science.aap8179; pmid: 29545509
42. S. W. Lapan, P. W. Reddien, Transcriptome analysis of the planarian eye identifies *ovo* as a specific regulator of eye regeneration. *Cell Rep.* **2**, 294–307 (2012). doi: 10.1016/j.celrep.2012.06.018; pmid: 22884275
43. F. Cebrià, P. A. Newmark, Planarian homologs of *netrin* and *netrin receptor* are required for proper regeneration of the central nervous system and the maintenance of nervous system architecture. *Development* **132**, 3691–3703 (2005). doi: 10.1242/dev.01941; pmid: 16033796
44. P. A. Newmark, P. W. Reddien, F. Cebrià, A. Sánchez Alvarado, Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. *Proc. Natl. Acad. Sci. U.S.A.* **100** (suppl. 1.), 11861–11865 (2003). doi: 10.1073/pnas.1834205100; pmid: 12917490
45. F. Cebrià, T. Guo, J. Jopek, P. A. Newmark, Regeneration and maintenance of the planarian midline is regulated by a *slit* orthologue. *Dev. Biol.* **307**, 394–406 (2007). doi: 10.1016/j.ydbio.2007.05.006; pmid: 17553481
46. F. Cebrià, P. A. Newmark, Morphogenesis defects are associated with abnormal nervous system regeneration following *roboA* RNAi in planarians. *Development* **134**, 833–837 (2007). doi: 10.1242/dev.02794; pmid: 17251262
47. M. L. Scimone, L. E. Cote, T. Rogers, P. W. Reddien, Two FGFRL-Wnt circuits organize the planarian anteroposterior axis. *eLife* **5**, e12845 (2016). doi: 10.7554/eLife.12845; pmid: 27063937
48. F. Cebrià *et al*., FGFR-related gene *nou-darake* restricts brain tissues to the head region of planarians. *Nature* **419**, 620–624 (2002). doi: 10.1038/nature01042; pmid: 12374980
49. R. Lander, C. P. Petersen, Wnt, Ptk7, and FGFRL expression gradients control trunk positional identity in planarian regeneration. *eLife* **5**, e12850 (2016). doi: 10.7554/eLife.12850; pmid: 27074666
50. K. A. Gurley *et al*., Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. *Dev. Biol.* **347**, 24–39 (2010). doi: 10.1016/j.ydbio.2010.08.007; pmid: 20707997
51. A. Augsburger, A. Schuchardt, S. Hoskins, J. Dodd, S. Butler, BMPs as mediators of roof plate repulsion of commissural neurons. *Neuron* **24**, 127–141 (1999). doi: 10.1016/S0896-6273(00)80827-2; pmid: 10677032
52. S. J. Butler, J. Dodd, A role for BMP heterodimers in roof plate-mediated repulsion of commissural axons. *Neuron* **38**, 389–401 (2003). doi: 10.1016/S0896-6273(03)00254-X; pmid: 12741987
53. P. W. Reddien, A. L. Bermange, A. M. Kicza, A. Sánchez Alvarado, BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. *Development* **134**, 4043–4051 (2007). doi: 10.1242/dev.007138; pmid: 17942485
54. P. Blader, S. Rastegar, N. Fischer, U. Strähle, Cleavage of the BMP-4 antagonist chordin by zebrafish Tolloid. *Science* **278**, 1937–1940 (1997). doi: 10.1126/science.278.5345.1937; pmid: 9395394
55. S. Piccolo *et al*., Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. *Cell* **91**, 407–416 (1997). doi: 10.1016/S0092-8674(00)80424-9; pmid: 9363949
56. M. J. Galko, M. Tessier-Lavigne, Function of an axonal chemoattractant modulated by metalloprotease activity. *Science* **289**, 1365–1367 (2000). doi: 10.1126/science.289.5483.1365; pmid: 10958786
57. D. W. Sretavan, E. Puré, M. W. Siegel, L. F. Reichardt, Disruption of retinal axon ingrowth by ablation of embryonic mouse optic chiasm neurons. *Science* **269**, 98–101 (1995). doi: 10.1126/science.7541558; pmid: 7541558
58. M. W. Cowles *et al*., Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult

neurogenesis and neuronal regeneration. *Development* **140**, 4691–4702 (2013). doi: 10.1242/dev.098616; pmid: 24173799
59. M. L. Scimone, K. M. Kravarik, S. W. Lapan, P. W. Reddien, Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. *Stem Cell Reports* **3**, 339–352 (2014). doi: 10.1016/j.stemcr.2014.06.001; pmid: 25254346
60. R. H. Roberts-Galbraith, J. L. Brubacher, P. A. Newmark, A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *eLife* **5**, e17002 (2016). doi: 10.7554/eLife.17002; pmid: 27612384
61. E. Z. Macosko *et al*., Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015). doi: 10.1016/j.cell.2015.05.002; pmid: 26000488
62. C. T. Fincher, O. Wurtzel, T. de Hoog, K. M. Kravarik, P. W. Reddien, Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. *Science* **360**, eaaq1736 (2018). doi: 10.1126/science.aaq1736; pmid: 29674431
63. L. E. Cote, E. Simental, P. W. Reddien, Muscle functions as a connective tissue and source of extracellular matrix in planarians. *Nat. Commun.* **10**, 1592 (2019). doi: 10.1038/s41467-019-09539-6; pmid: 30962434
64. I. M. Oderberg, D. J. Li, M. L. Scimone, M. A. Gaviño, P. W. Reddien, Landmarks in existing tissue at wounds are utilized to generate pattern in regenerating tissue. *Curr. Biol.* **27**, 733–742 (2017). doi: 10.1016/j.cub.2017.01.024; pmid: 28216315
65. T. Hayashi *et al*., A LIM-homeobox gene is required for differentiation of Wnt-expressing cells at the posterior end of the planarian body. *Development* **138**, 3679–3688 (2011). doi: 10.1242/dev.060194; pmid: 21828095
66. C. P. Petersen, P. W. Reddien, Wnt signaling and the polarity of the primary body axis. *Cell* **139**, 1056–1068 (2009). doi: 10.1016/j.cell.2009.11.035; pmid: 20005801
67. A. Jaworski, M. Tessier-Lavigne, Autocrine/juxtaparacrine regulation of axon fasciculation by Slit-Robo signaling. *Nat. Neurosci.* **15**, 367–369 (2012). doi: 10.1038/nn.3037; pmid: 22306607
68. J. M. Martín-Durán, E. Amaya, R. Romero, Germ layer specification and axial patterning in the embryonic development of the freshwater planarian *Schmidtea polychroa*. *Dev. Biol.* **340**, 145–158 (2010). doi: 10.1016/j.ydbio.2010.01.018; pmid: 20100474
69. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015). doi: 10.1038/nbt.3192; pmid: 25867923
70. P. V. Kharchenko, L. Silberstein, D. T. Scadden, Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014). doi: 10.1038/nmeth.2967; pmid: 24836921
71. K. G. Ross *et al*., Novel monoclonal antibodies to study tissue regeneration in planarians. *BMC Dev. Biol.* **15**, 2 (2015). doi: 10.1186/s12861-014-0050-9; pmid: 25604901

ACKNOWLEDGMENTS

We thank C. McQuestion for help in scRNA-seq surgeries and members of the Reddien lab for discussions and comments on the manuscript. **Funding:** P.W.R. is an investigator of HHMI and an associate member of the Broad Institute. We thank the Eleanor Schwartz Charitable Foundation for support. **Author contributions:** M.L.S., K.D.A., and P.W.R. designed the study; M.L.S., K.D.A., and A.R.B. carried out experiments; D.J.L. originally found the *ror-1* RNAi phenotype; M.L.S., K.D.A., and C.T.F. analyzed data; and M.L.S., K.D.A., and P.W.R. wrote the manuscript. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/eaba3203/suppl/DC1
Figs. S1 to S13
Tables S1 to S3
References
MDAR Reproducibility Checklist
Movies S1 and S2

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

22 November 2019; accepted 6 May 2020
10.1126/science.aba3203

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Flexible recruitment of memory-based choice representations by the human medial frontal cortex

Juri Minxha, Ralph Adolphs, Stefano Fusi, Adam N. Mamelak, Ueli Rutishauser*

INTRODUCTION: Decision-making in complex environments relies on flexibly combining stimulus representations with context, goals, and memories. A central component of cognitive flexibility is to selectively retrieve information from memory and utilize the retrieved information to make decisions. The medial frontal cortex (MFC) plays a critical role in this process by representing task sets, context, and outcomes. During decision-making, the MFC is thought to selectively engage memory retrieval by representing memory-based choices and mediating interactions between the frontal lobes and the hippocampus and amygdala (HA) through phase-locking of MFC activity to oscillations in the HA. It remains unknown what features of decisions and context are represented in the human MFC and what functional interactions between the MFC and HA mediate dynamic memory retrieval during a task.

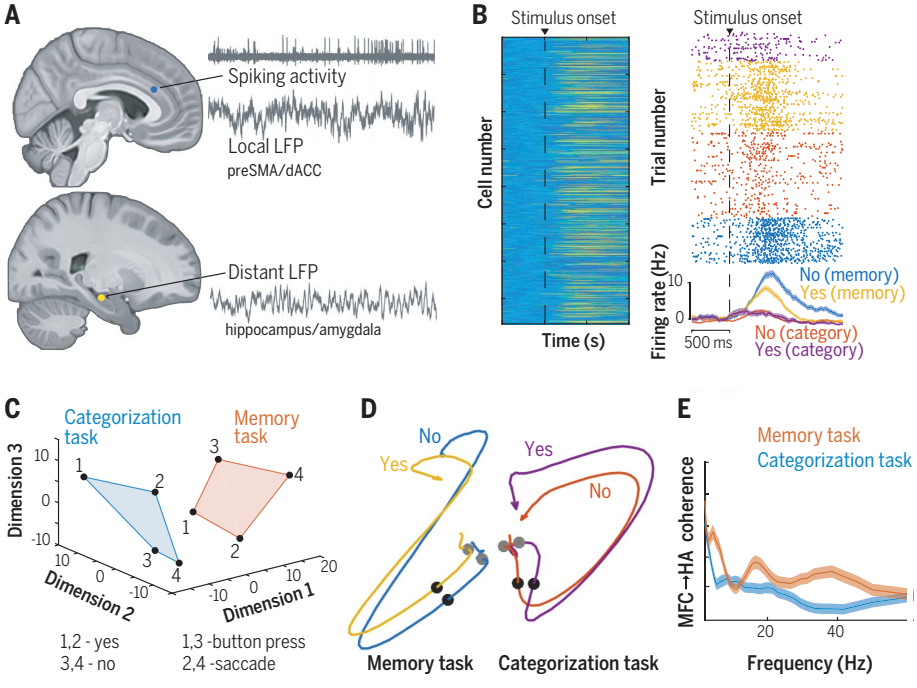
RATIONALE: We recorded single neurons and local field potentials in the human MFC and HA in patients implanted with depth electrodes. Subjects switched between two tasks: recognition memory and categorization. To identify signatures of task demands, we compared the strength of encoding of stimulus familiarity, category, and choices between tasks and tested whether decoders trained in one task generalized to the other task. Such cross-task generalizability would indicate abstract representations of the underlying variables. We hypothesized that this approach would reveal neural signatures of the representations and functional interactions that permit memory-based decisions.

RESULTS: We recorded from 1430 single neurons in the HA and MFC [dorsal anterior cingulate cortex (dACC) and the pre-supplementary motor area (pre-SMA)] across 13 patients. Subjects made "yes" or "no" decisions using button presses or saccades (eye movements) to indicate whether an image was novel or familiar, or whether an image belonged to a given visual category. Instructions were given before each block of trials, explaining the task and response modality to use (i.e., task set). Examining the underlying neural representations at the single-neuron and population levels revealed the following: (i) Cells in the MFC represented task set during baseline periods. These contextual signals emerged rapidly after a task switch and generalized across all response and task-type combinations in the MFC but not the HA. (ii) The strength and geometry of representations of familiarity were task-insensitive in the HA but not in the MFC. The responses of these memory-selective cells were a reflection of memory strength rather than decisions about the memory. (iii) The visual category of stimuli was represented more strongly during the memory task in both the MFC and HA. This encoding of category generalized across tasks fully in the HA but not the MFC. (iv) Choices in both tasks were most strongly represented by cells in the MFC. This choice representation differed in its population-level geometry between the two tasks but was insensitive to response modality (button press or saccade). One subset of MFC cells signaled only memory-based choices, and these cells signaled decisions about the memory. (v) MFC cells phase-locked their activity to theta-frequency band oscillation in the HA preferentially in the memory task, with memory-choice cells also phase-locking in the gamma-frequency band. The strength of this interareal phase-locking in both frequency bands of the MFC cells that signaled memory-based choices was predictive of behavior.

ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aba3313

CONCLUSION: We leveraged the opportunity to record from single neurons in humans to identify representations of choices, task sets, stimulus category, and familiarity in the human MFC and HA. We found that neuronal populations within the MFC formed two separate decision axes: one for memory-based decisions and another for categorization-based decisions. MFC-HA theta-frequency functional connectivity was selectively enhanced during memory retrieval. This work reveals a neuronal mechanism in the human brain whereby oscillation-mediated coordination of activity between distant brain regions and accompanying changes in strength of representation and/or geometry implements task-dependent retrieval of memory. ■



Flexible representations of choices in the human frontal lobe. (**A**) Recording locations. LFP, local field potential. (**B**) Population response of all recorded neurons (left) and example of a cell signaling memory-based choices (right). (**C** and **D**) Representational geometry analysis reveals that different subspaces are used by the two tasks, establishing a memory-specific decision axis. (**E**) Theta- and gamma-band coherence of MFC choice cells with HA LFPs increased during the memory task.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: ueli.rutishauser@csmc.edu
Cite this article as J. Minxha *et al.*, *Science* **368**, eaba3313 (2020). DOI: 10.1126/science.aba3313

RESEARCH ARTICLE

NEUROSCIENCE

Flexible recruitment of memory-based choice representations by the human medial frontal cortex

Juri Minxha[1,2,3], Ralph Adolphs[2,4], Stefano Fusi[3], Adam N. Mamelak[1], Ueli Rutishauser[1,2,5,6]*

Decision-making in complex environments relies on flexibly using prior experience. This process depends on the medial frontal cortex (MFC) and the medial temporal lobe, but it remains unknown how these structures implement selective memory retrieval. We recorded single neurons in the MFC, amygdala, and hippocampus while human subjects switched between making recognition memory–based and categorization-based decisions. The MFC rapidly implemented changing task demands by using different subspaces of neural activity and by representing the currently relevant task goal. Choices requiring memory retrieval selectively engaged phase-locking of MFC neurons to amygdala and hippocampus field potentials, thereby enabling the routing of memories. These findings reveal a mechanism for flexibly and selectively engaging memory retrieval and show that memory-based choices are preferentially represented in the frontal cortex when required.

Behavior in complex environments requires decisions that flexibly combine stimulus representations with context, goals, and memory. Two key aspects of cognitive flexibility are the selective utilization of relevant information depending on task demands and the retrieval of information from memory, when needed (*1*). We are beginning to understand the neural mechanisms that underlie flexible decisions in the case of perceptual decision-making (*2*–*4*), with evidence for both early gating, mediated by top-down attention (*5*), and late selection of relevant features in the prefrontal cortex (*3*). In contrast, little is known about the decision mechanisms that also depend on associated category knowledge and memory. In particular, it is not clear how memory retrieval is selectively engaged when decision-relevant information needs to be actively searched for in memory (*6*–*8*).

The medial frontal cortex (MFC) is critical for complex behavior and registers cognitive conflict, errors, and choice outcomes (*9*–*11*). It supports flexible decision-making in two ways: (i) by representing task sets (*12*–*14*) and context (*15*), and (ii) by selectively engaging memory retrieval through functional interactions with other brain areas (*16*–*18*), specifically the hippocampus (*19*–*21*) and amygdala (*22*, *23*). A mechanism that facilitates such interareal interactions is phase-locking of MFC activity to osciollations in the hippocampus or amygdala. This mechanism has been extensively investigated in rodents during spatial behavior (*24*–*26*) and fear conditioning (*27*, *28*), but its broader function remains poorly understood (*29*), particularly in humans. Similarly, human neuroimaging studies indicate that the MFC is involved in memory search (*8*, *18*, *30*–*34*) and that patterns and level of activity and connectivity assessed by functional magnetic resonance imaging (fMRI) vary as a function of retrieval intentionality (*35*–*38*). It is not yet known what features of decisions and context are represented in the human MFC, whether memory retrieval selectively engages synchrony between the MFC and the hippocampus and/or amygdala, and whether synchrony can be engaged dynamically when required. This lack of knowledge stands in stark contrast to the patent behavioral ability of humans to flexibly recruit memory processes in everyday life (*39*, *40*) and to our detailed knowledge of memory representations in the human hippocampus and amygdala, where cells represent aspects of declarative memories, such as the familiarity and the identity of a stimulus (*41*–*43*). To address these open questions, we used simultaneous recordings of single neurons and local field potentials (LFPs) in the human MFC, hippocampus, and amygdala.

[1]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [2]Computation and Neural Systems Program, Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [3]Center for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University, New York, NY, USA. [4]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, USA. [5]Center for Neural Science and Medicine, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [6]Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA, USA.
*Corresponding author. Email: ueli.rutishauser@csmc.edu

Task and behavior

We recorded from 1430 single neurons across four brain areas (Fig. 1, C and D; see table S1; 33 sessions in 13 subjects): n = 203, 460, 329, and 438 neurons from anterior hippocampus (HF), amygdala (AMY), dorsal anterior cingulate cortex (dACC), and pre-supplementary motor area (pre-SMA), respectively. For brevity, we refer to HF and AMY together as HA (n = 663 neurons) and to dACC and pre-SMA together as MFC (n = 767 neurons).

Human subjects viewed a sequence of 320 images, grouped into eight blocks of 40 images each, in each session (Fig. 1, A and B). At the beginning of each block, subjects were instructed which decision to make and which response modality to use to communicate their decision. Subjects made a "yes" or "no" decision for each trial to indicate whether an image belonged to a given visual category ("categorization task") or whether an image had been seen before in the task or not ("memory task"). No feedback was provided (see Materials and methods section for details on the task). Each image shown belonged to one of four visual categories: human faces, monkey faces, fruits, or cars. In each block, half of the images shown were repeated and half were novel (except in the first block, in which all images were novel).

Subjects indicated choices using either saccades (leftward or rightward eye movement) or button press while maintaining fixation at the center of the screen (Fig. 1, E and F; mean ± SD, 94 ± 15% of all gaze positions fell within the image shown). Reaction times (RTs) were significantly longer in the memory task than in the categorization task [Fig. 1G, mean RT of 1.48 ± 1.1 s versus 1.19 ± 1.2 s, respectively, $P < 1 \times 10^{-20}$, two-sample Kolmogorov-Smirnov (KS) test, mean ± SD across all trials in a given task]. Subjects performed with an average accuracy of 97 ± 6% versus 71 ± 6% in the categorization and memory tasks, respectively (mean ± SD across n = 33 sessions). This difference in accuracy remained after we matched for RT between the two tasks (96 ± 6% versus 72 ± 8% with matched RTs of 1.23 ± 0.60 s versus 1.24 ± 0.60 s for the categorization and memory task, respectively). Even without RT matching, the initial response in terms of arousal was not different between tasks, as assessed by pupillometry (fig. S1, J to L). In the memory task, accuracy increased as a function of how many times an image had been shown (Fig. 1H, $\beta_{appearances}$ = 0.56, $P < 1 \times 10^{-20}$, mixed effects logistic regression; also see fig. S1, C and D, for effect of target versus nontarget on memory performance). Subjects had shorter RTs on "yes" (seen before) decisions than on "no" (novel stimulus) decisions in the memory task (fig. S1A, see legend for statistics), as expected from a medial temporal lobe (MTL)–dependent recognition memory task (*41*). In the categorization task, RT was not significantly different between the two responses (fig. S1A), showing the absence of oddball effects.

Effects of task type and response modality in the MFC

Instructions about the task type and response modality were shown at the beginning of each

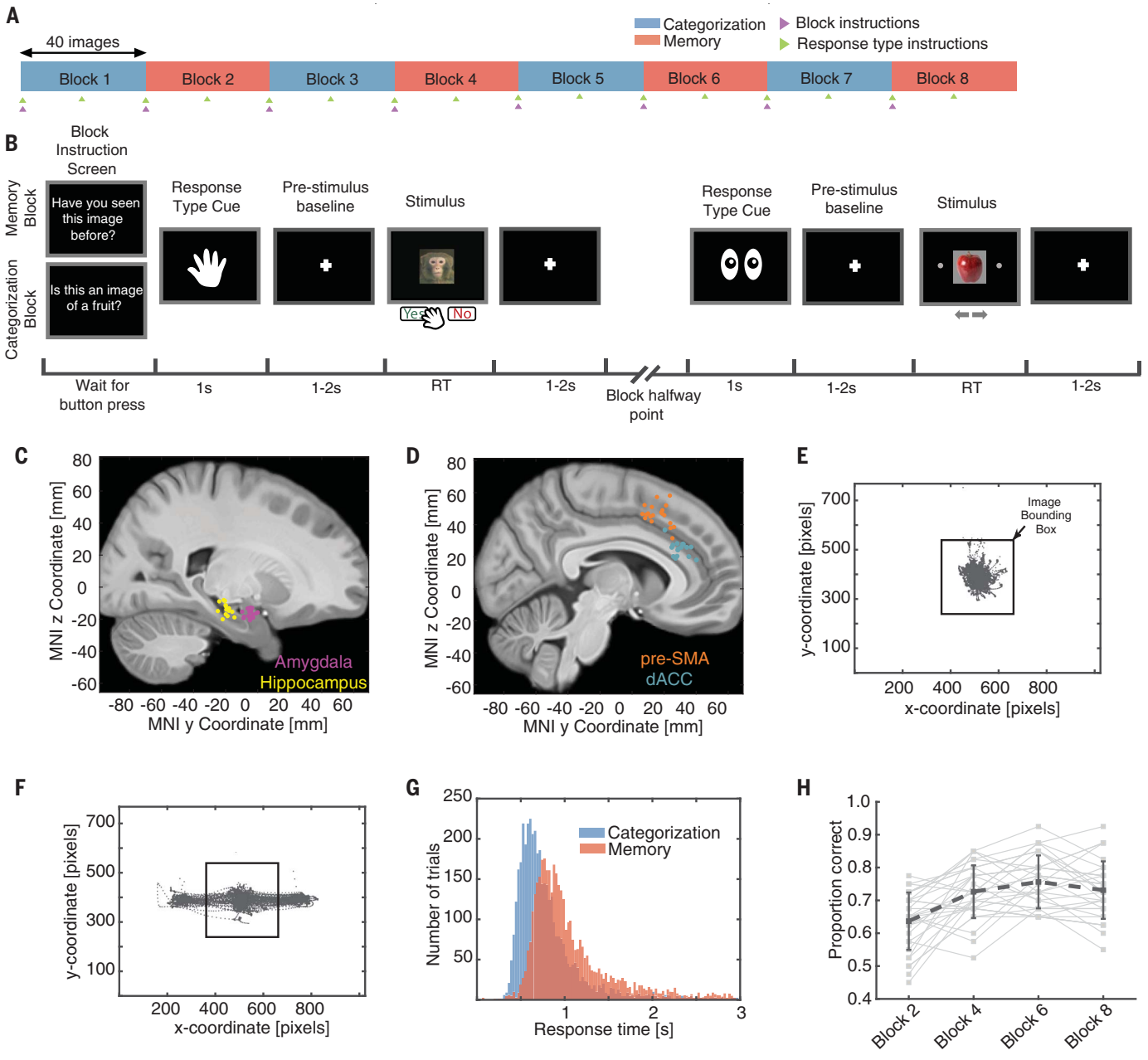


Fig. 1. Task, electrode locations, and behavior. (**A**) Task structure. A session consisted of eight blocks of 40 trials. The task switched with each block (blue = categorization, red = memory), and the response modality switched halfway through each block (saccade or button press, randomly assigned at the beginning of the block). The subject was instructed about the task at the beginning of each block (purple arrows) and how to respond at the beginning and halfway points of each block (green arrows). (**B**) Example of screens shown to subjects for two example trials. (**C** and **D**) Electrode locations. Each dot is the location of a microwire bundle in one subject. Coordinates are in Montreal Neurological Institute (MNI) 152 space. (**E** and **F**) Eye tracking data from one session from the button press (E) and eye movement (F) trials. (**G**) Reaction times as a function of task across all sessions (memory, μ ± SEM, 1.27 ± 0.02 s; categorization, 0.90 ± 0.02 s; $P = 7.6 \times 10^{-228}$, two-sample KS test). (**H**) Memory performance improves over the course of the experiment ($\beta = 0.56$, $P = 8.42 \times 10^{-130}$, logistic mixed effects model). See fig. S1 for an extended summary of the behavior.

block (Fig. 1, A and B). Cells showed significant modulation of their firing rate during the baseline period as a function of task type (Fig. 2, A and B, shows an example in pre-SMA). At the single-neuron level, significantly more cells were modulated by task type in the MFC than in the HA: 25% of MFC cells (165/767, 82 in dACC, 83 in pre-SMA; see fig. S2B) versus 12% of HA cells (79/663, 21 in HF, 58 in AMY), χ^2 test of proportions, $P < 1.5 \times 10^{-6}$. Similarly, at the population level, population decoding accuracy was significantly higher in the MFC than in the HA [Fig. 2C; 90% versus 70%, respectively; $P < 1 \times 10^{-3}$; true difference (Δ_{true}) = 20% versus empirical null distribution; see Materials and methods], a conclusion that held regardless of the number of neurons used (fig. S2H). Cells also modulated their activity as a function of response modality during the baseline period (fig. S2E shows an example). As with task-type encoding, significantly more cells encoded response modality in the MFC, and this signal could be decoded with higher accuracy in the MFC than in the HA (14% versus

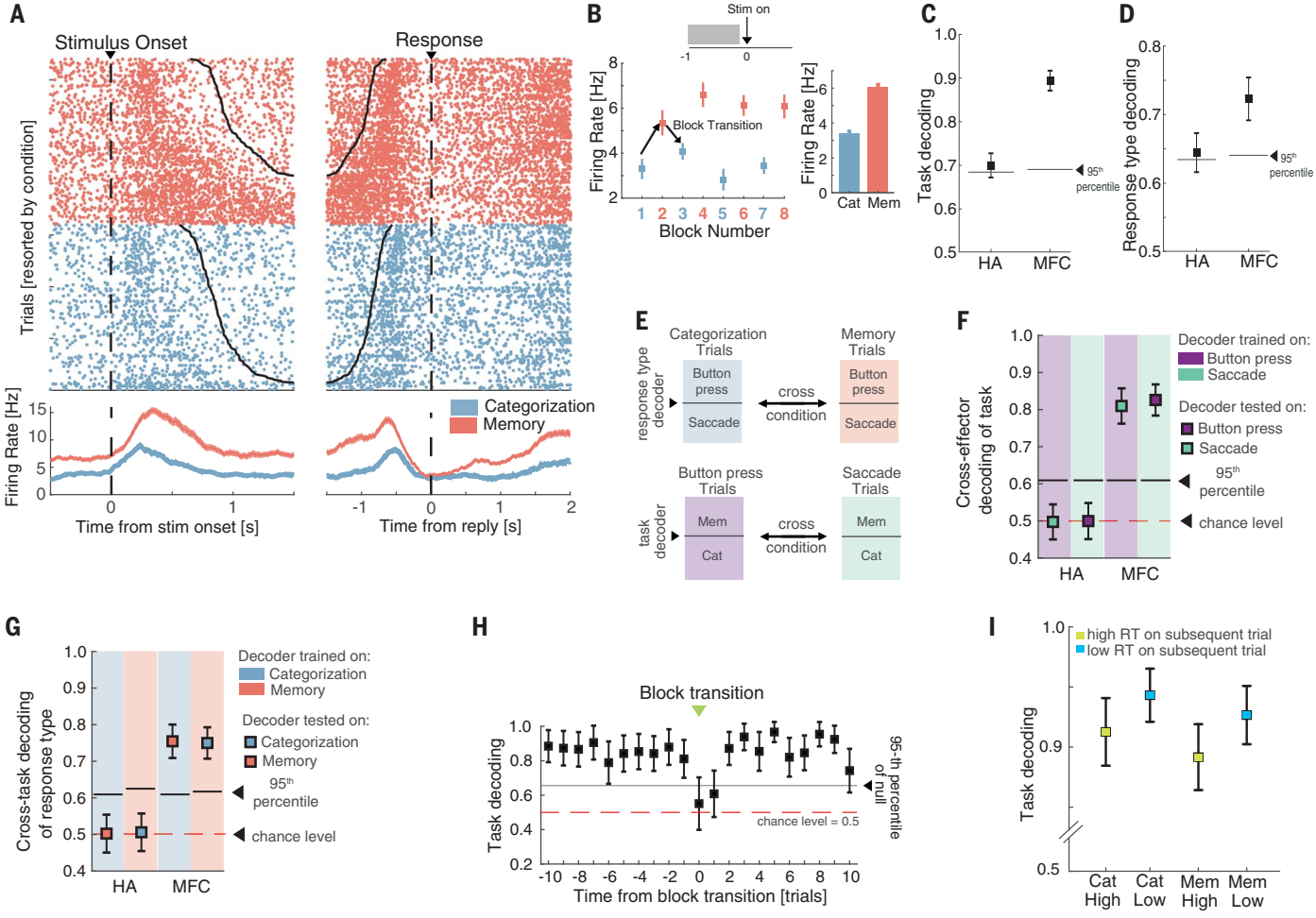


Fig. 2. Representations of task type and response modality. (**A** and **B**) Example pre-SMA neuron. (B) Average firing rate during the baseline period (−1 to 0 s relative to stimulus onset) for each block for the cell shown in (A). The average baseline firing rate across all blocks of the same type is shown. (**C** and **D**) Population decoding of task type (C) and response modality (D). (**E**) Cross-condition decoding approach. The background color denotes the type of trials that were used to train a given decoder. (**F**) Cross-response modality decoding of task type from the baseline firing rate of all recorded cells. (**G**) Cross-task decoding of response modality. (**H**) Decoding performance as a function of trial number relative to a task type switch (green arrows in Fig. 1A; transitions from categorization to memory and vice versa were pooled). Error bars indicate SD in all panels, with the exception of (B), where they indicate SEM. (**I**) Baseline decoding of task type for subsequent trials with short reaction times was more accurate than decoding on long reaction time trials. Performance is shown separately for categorization (Cat) and memory (Mem) trials ($P = 2 \times 10^{-11}$ and 7×10^{-13}, respectively, Wilcoxon rank sum test). Error bars denote standard error in decoding accuracy across trials (80 trials in each of the four groups). See fig. S2 for additional analyses that break down context effects by specific anatomical regions.

10% of cells; 84/593 versus 59/586 in the MFC and HA, respectively; 33 in ACC, 51 in pre-SMA, 27 in HF, 32 in AMY; χ^2 test of proportions, $P = 0.03$; population decoding performance 72% versus 64%, Fig. 2D; $P < 0.05$, $\Delta_{true} = 8\%$ versus empirical null distribution); this conclusion held regardless of number of neurons used (fig. S2I).

After a task switch, contextual signals emerged rapidly within one to two trials in the new context in the MFC (Fig. 2H). This was not a result of ongoing poststimulus processing, because the task could still be decoded even if only considering the subset of task cells in the MFC, which did not differentiate between the tasks around response time (see fig. S2G). Task-switching costs were also reflected in the subjects' longer reaction times shortly after a change in task or effector type (fig. S2A). Task-type representations during the baseline period were stronger on trials where the subject subsequently produced a fast response than on those where the response time was slow (Fig. 2I), indicating behavioral relevance. We also tested whether the two types of contextual signals were sufficiently robust to avoid interference with one another, using a cross-condition generalization decoding analysis (*44*) (Fig. 2E). We first trained a decoder to discriminate task type on trials where the subject was instructed to reply with a button press, and then we tested the performance of this decoder on trials where the subject was instructed to use saccades (and vice versa). The two decoders generalized in the MFC but not in the HA (Fig. 2, F and G). For this reason, we focused on the MFC when conducting the analysis above.

Cross-condition generalization of familiarity and image category

Next, we examined whether the neural representations of image category and familiarity are sensitive to task demands. We assessed two consequences of task demands: generalization across tasks and strength of representations within each task. At the single-unit level, we examined visually selective (VS) cells (*42*), whose responses are thought to reflect input from high-level visual cortex, and memory-selective (MS) cells (*41*), whose response signals stimulus familiarity (Fig. 3, A and B, shows

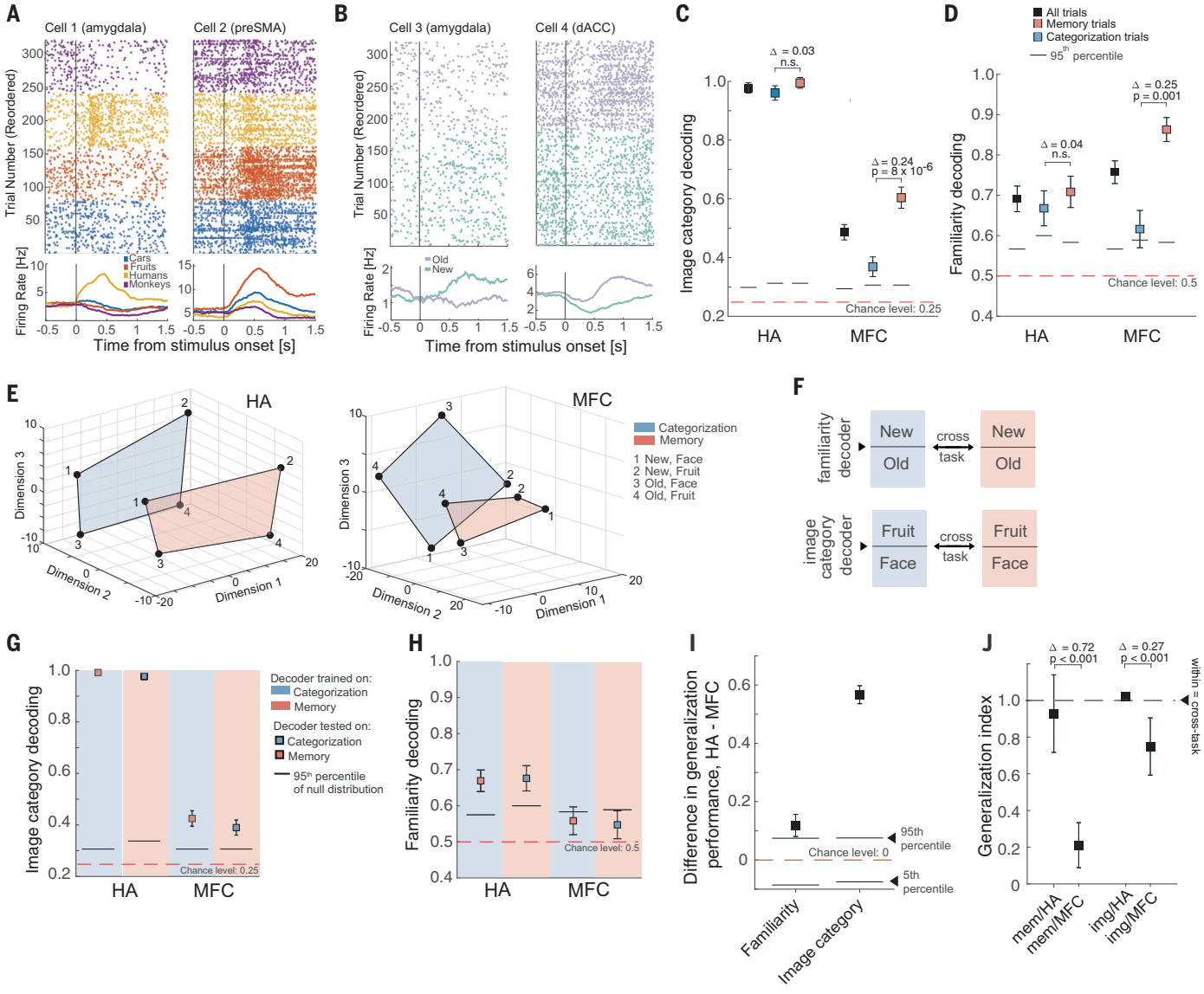


Fig. 3. Representations of image category and familiarity (new versus old). (**A** and **B**) Example cells that (A) represent image category and (B) differentiate between new and old stimuli. (**C**) Decoding accuracy of image category from all recorded cells was significantly higher in the HA relative to the MFC (Δ_{true} = 49%, $P < 0.001$). (**D**) Decoding of new versus old (ground truth) was similarly accurate in the HA and MFC (Δ_{true} = 7%, $P = 0.13$). For new versus old decoding, trials with images of monkeys were excluded, because the recognition performance for these images was at chance (fig. S1B). (**E**) Population activity of all recorded HA (left) and MFC (right) cells, plotted in 3D using MDS. Individual points show the mean activity of the population for that specific condition. The highlighted plane contains all locations of state space occupied by a given task for the case of fruits versus faces as the binary category distinction (for illustration only; all analysis uses all categories). The geometry of the representation allows for a decoder that is trained on one task to generalize to the other task (see fig. S4C for example decoder hyperplanes). (**F**) Approach used for the cross-condition generalization analysis. Color indicates task (blue = categorization, red = memory). (Top) We trained a decoder to discriminate between new and old trials on categorization trials and then tested its performance on new and old stimuli encountered during the memory condition (and vice versa). (Bottom) Similarly, a decoder that is trained to discriminate between image categories (in this example, faces versus fruits; all results include all six possible pairs) on categorization trials was tested on memory trials. (**G**) Cross-condition generalization performance for image category. (**H**) Cross-condition generalization performance for new versus old. (**I**) Difference in cross-task generalization decoding accuracy for familiarity and image category between HA and MFC. Difference is computed between the average cross-task performances in each area (i.e., average of memory→categorization and categorization→memory). The null distribution for the average was estimated from the empirical null estimated by shuffling the labels used to train the decoders. For both variables, decoding from HA had significantly greater cross-task generalization performance than decoding from MFC (the difference in both cases is positive and outside of the 95th percentile of the null distribution). (**J**) Generalization index (see Materials and methods) for memory (two data points on the left) and image category (two data points on the right). For both image category and familiarity, generalization across task was higher in the HA population than in the MFC population (see figure for statistics; Δ, difference).

examples). Of the HA cells, 40% were visually selective (264/663, 35 in HF and 229 in AMY) and 11% were memory-selective (73/663, 10 in HF, 63 in AMY; 24/73 were old>new selective and 49/73 were new>old selective; see Materials and methods for selection model). In the MFC, 13% (103/767) of the cells were visually selective and 11% (84/767) were memory-selective. First, we performed single-neuron analysis of the selected HA cells. Visual and memory selectivity were present in both the memory and categorization blocks (figs. S3, D, E, H, and I). MS cell responses reflected a memory process: they strengthened over blocks as memories became stronger (fig. S3G), and they differed between forgotten old (false negative, FN) and correctly identified new (true negative, TN) stimuli for both new>old (n = 49) and old>new (n = 24) preferring MS cells (fig. S10F). Furthermore, these memory signals were behaviorally relevant: new/old decoding was significantly weaker in incorrect trials than in correct trials (fig. S10G).

To answer the question of whether coding for visual and memory information generalizes across tasks, we took a population-level approach (over all single units, without selection). We used decoding performance on individual trials and single-neuron analysis (fig. S4, D and E) to assess whether the neural encoding of visual category and/or familiarity of a stimulus depended on task demands. In both the HA and MFC, image category could be decoded (Fig. 3C, 98 and 49% in the HA and MFC, respectively; chance level = 25%). Category decoding performance was significantly higher in the HA than the MFC (Δ_{true} = 49% versus empirical null distribution, $P < 1 \times 10^{-3}$). In the HA, the ability to decode category was not significantly different between the two tasks (Fig. 3C, 96% versus 99% in categorization and memory, respectively; Δ_{true} = 3%, P = 0.25) and could be decoded above chance in both the HF and AMY (fig. S4, A and B). At the single-neuron level, HA neurons encoded significantly more information about category in the memory task (fig. S4D), an effect that decoders were not sensitive to because of saturation. In the MFC, decoding accuracy for image category was significantly higher in the memory task (Fig. 3C, 60% versus 36%, Δ_{true} = 24% versus empirical null distribution, $P < 0.001$). Memory was decodable in both the HA and MFC (Fig. 3D, 69% versus 76%, respectively), with no significant difference in decoding accuracy between the two tasks in the HA (Fig. 3D, 67% versus 71% in categorization and memory trials, respectively; Δ_{true} = 4%, P = 0.3) and significantly better decoding ability in the MFC in the memory task (Fig. 3D, 86% versus 61%; Δ_{true} = 25% versus empirical null, P = 0.001). Single-neuron analysis confirmed the impression from decoding that the strength of memory signals in the HA was not modulated by task demands, whereas in the MFC it was (fig. S4E). In either task, at the population level, memory decoding in the HA was only possible in the amygdala (fig. S4B). The population-level decoding of memory in the HA was principally a reflection of the signal carried by the MS cells (fig. S10E) and was not due to repetition suppression of VS cells (fig. S10, B and C), because it was not possible to decode familiarity from VS cells alone (fig. S10D).

To gain insight into the geometry of the population-level representations, we assessed whether the decoders trained to report familiarity and the category of the stimuli in one task would generalize to the other task (Fig. 3F schematizes our approach). Cross-task generalizability would indicate that familiarity and visual category are represented in an abstract format (*44*). First, cross-task generalization performance was greater in the HA than MFC for both image category (Fig. 3, G and I; 98% versus 41%, averaged across the two cross-condition decoding performances; Δ_{true} = 57% versus empirical null, $P < 1 \times 10^{-3}$) and familiarity (Fig. 3, H and I; 67% versus 55%, Δ_{true} = 12% versus empirical null, $P < 0.05$). Second, to help understand the geometry of these neural representations, we projected the average HA and MFC population activity for all possible pairings of familiarity, image category, and task (eight different conditions) into a three dimensional (3D) state space using multidimensional scaling. For illustration purposes, we show this 3D state space for the two image categories (fruits and faces) for which memory performance was the best. In the HA (Fig. 3E, left), the relative positions of a "new face" with respect to an "old face" were preserved across tasks (shown as differently colored planes). The relatively parallel location of the subspace of neural activity occupied by the two tasks permits cross-task generalization for both image category and familiarity. In contrast, in the MFC (Fig. 3E, right), the relative positions of the four conditions were not preserved. This is consistent with the weaker cross-task generalization performance observed in the MFC relative to the HA (Fig. 3, G and H), resulting in reduced generalization indices in the MFC compared with the HA (Fig. 3J; this metric takes into account different levels of within-task decoding accuracy, which is an upper bound for the achievable generalization performance; see the supplementary text section of the supplementary materials for details).

Representation of choice

We next investigated how the subject's choice (yes or no) is represented by single neurons (Fig. 4A shows examples) and the population. Decoding accuracy for choices was highest in the MFC, with an average population decoding performance of 89% compared with 68% in the HA (Fig. 4B; Δ_{true} = 19% versus empirical null, $P < 1 \times 10^{-3}$; 61% in AMY and 57% in HF when trained separately; fig. S7G shows this result as a function of number of neurons used). Further single-neuron (fig. S11, A and B) and population (Fig. 4E and fig. S11C) analysis confirmed that the choice signal was significantly stronger in the MFC regardless of selection threshold and quantification method. We therefore first analyzed choice information in the MFC (see below for results in the HA). Choice decoding in the MFC was strongest shortly after stimulus onset, well before the response was made (fig. S7A). To disassociate representation of choice (yes or no) from the representation of ground truth (old or new) during the memory recognition task, we fit a choice decoder to a subset of trials, half of which were correct and half of which were incorrect. The activity of MFC cells predicted choice but not the ground truth at levels significantly different from chance (Fig. 4C; choice decoding is above the 95th percentile of the null distribution, whereas new/old decoding is not; see fig. S7D for this analysis shown over time). Choice could be decoded from the MFC separately for both the correct and the incorrect trials (fig. S5I). As a control for potential confounds due to RT differences between tasks (see fig. S1A), we acquired data from a separate control task in which we eliminated RT differences behaviorally by adding a waiting period (six sessions in five subjects; n = 180 and 162 neurons in the HA and MFC, respectively; see Materials and methods and fig. S6). As in the original task, MFC cells represented the subject's choice (fig. S6, C to G), thereby confirming that this separation is not due to RT differences.

We used multidimensional scaling to visualize the population activity for the eight combinations of choices, task types, and response modality (Fig. 4D, and see Materials and methods). The resulting geometrical configuration indicates that choice decoding generalizes across response modality (Fig. 4D, top) but not across task types (Fig. 4D, bottom). We therefore computed the cross-task generalization performance of a decoder trained on choices during one task and tested on the other. We performed this analysis across time (Fig. 4E; see fig. S7B for this analysis shown separately for pre-SMA and dACC) and also in a single poststimulus time bin (Fig. 4F). To avoid confounds due to response time differences, we performed the fixed window analysis (Fig. 4F) only for the control task, where the timing between tasks was identical (fig. S6B). While the choice signal did not generalize across task types (Fig. 4F), it did generalize across response modality within the same task type (Fig. 4G). Quantifying this observation with the generalization index confirmed this impression (Fig. 4H). We also examined choice signals in the HA. Although choice signals were comparatively weak in the HA (fig. S7E

and Fig. 4B; see Fig. 4B for statistics), they nevertheless exhibited a pattern of generalization similar to that seen in the MFC (fig. S7, E and F).

To test the possibility that any task might yield a distinctive choice axis that does not generalize to any other task, we considered the four subtasks that make up the categorization trials in a given session (the target category can be any one of the four possible image categories). We tested whether the choice signal generalizes across these subtasks by training and testing across blocks requiring different categorizations. Choice decoding generalized across all subtasks in the categorization task but not the memory task (Fig. 4I). Next, we compared the dynamics of the population activity between the eight conditions arising

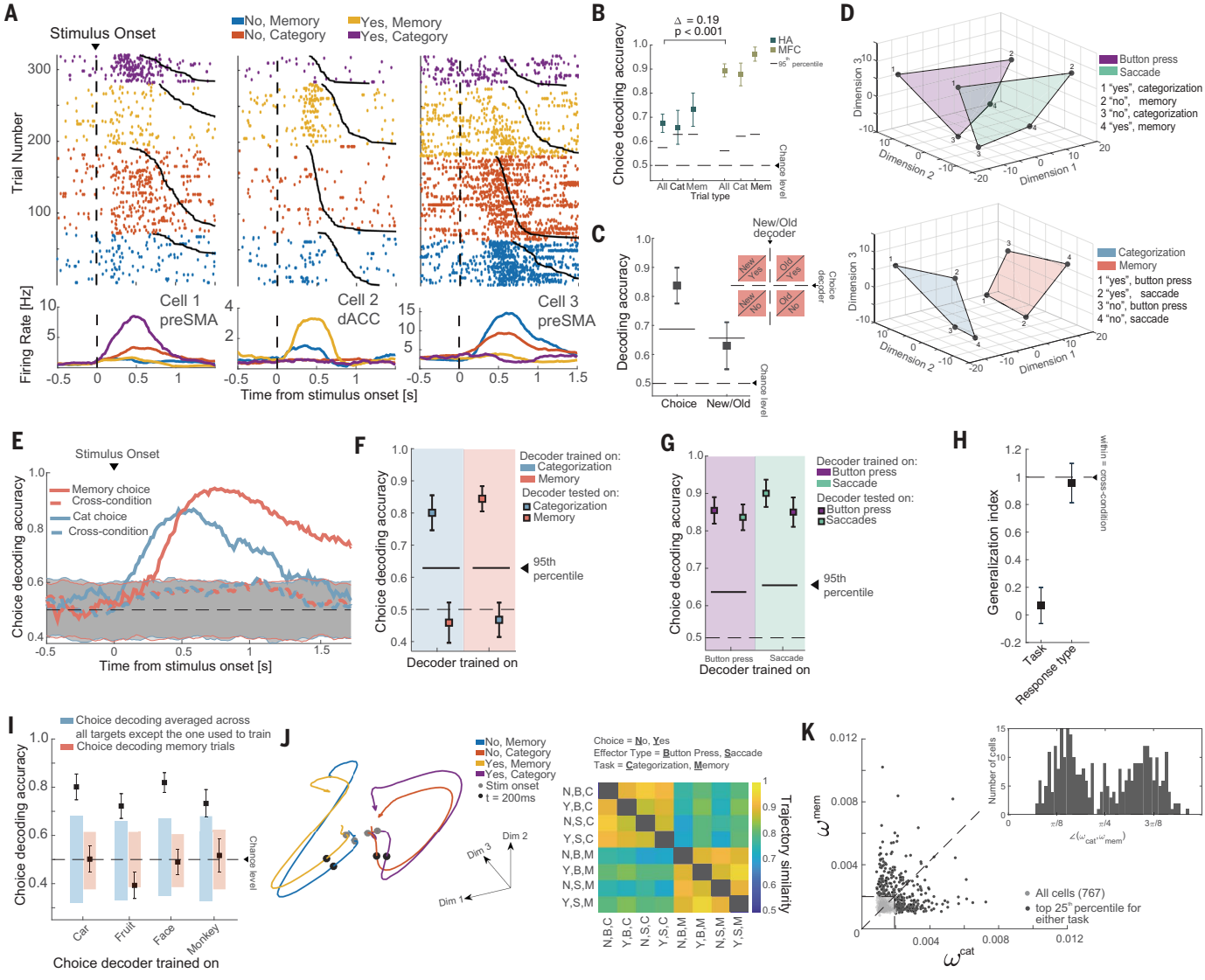


Fig. 4. Task-specific representation of choice. (**A**) Example MFC choice cells, split by choice (yes or no) and task. (**B**) Population choice decoding accuracy was significantly greater in MFC than in HA (across all trials, Δ_{true} = 19% versus empirical null, $P < 1 \times 10^{-3}$). (**C**) MFC cells represent choice and not the ground truth (i.e., new or old; memory trials only). (**D**) Population summary (neural state space) of choice-related activity in MFC, plotted in 3D space derived using MDS. (Top) Variability due to response modality. The highlighted planes connect the points of state space occupied by activity when using button presses (purple) or saccades (green). (Bottom) Variability due to task type. The highlighted planes connect the points of state space occupied by activity in the same task. (**E**) Choice-decoders trained in one task do not generalize to the other task (bin size: 500 ms, step size: 16 ms). (**F**) Same as (E), but for a fixed 1-s time window starting at 0.2 s after stimulus onset. (**G**) Choice decoding generalizes across effectors [see (D)]. (**H**) Generalization index of choice decoding (see Materials and methods) to summarize (F) and (G). The representation of choices generalized across response modality but not task. (**I**) Generalization between different subtasks of the categorization task but not between task types. The colored bars indicate the 5th to 95th percentile of the null distribution. (**J**) (Left) State-space trajectories for the four conditions arising from the combination of response (yes or no) and task (categorization or memory). (Right) Trajectory similarity, computed in an 8D latent space (recovered using GPFA, see Materials and methods) across the eight conditions arising from the combinations of choice, effector type, and task. (**K**) Decoder weight of each cell for decoding choice during the categorization and memory task. The cells in the top 25th percentile are shown in black. The inset shows the angle created by the vector $[\omega_i^{cat}, \omega_i^{mem}]$ with respect to the *x* axis of the cells marked in black.

from the combination of choice, effector type, and task in a state-space model recovered using Gaussian process factor analysis (GPFA) [(45), and see Materials and methods]. Comparing the pairwise similarity between the trajectories in state space (Fig. 4J, left) within the first 500 ms after the stimulus onset revealed that the patterns of dynamics in state space first cluster by task type (Fig. 4J, right; see movie S1).

We next examined whether the population-level analysis relied on different sets of neurons to decode choice in each of the two tasks. We determined how individual cells are recruited by a linear decoder (44, 46). For each cell, we quantified its importance (46) for both the memory and categorization choice decoder (Fig. 4K). We then plotted the degree of specialization for each cell on the basis of its importance in each task (see Materials and methods). Cells that report choice independently of task should lie on the diagonal (i.e., an angle of $\pi/4$). Instead, the distribution of angles was significantly bimodal across all cells (Fig. 4K, inset plot; $P < 1 \times 10^{-5}$, Hartigan dip test), with modes centered away from the diagonal. Despite this bimodality, we could still use the cells that are the most "useful" in one task to train a new decoder that can predict choice well above chance (although significantly weaker compared with using all cells) in the other task (fig. S7C). This is not an example of cross-task generalization, because we are fitting a new decoder.

Task-dependent spike-field coherence between MFC cells and HA LFP

It is thought that the selective routing of decision-related information between the HA and MFC is coordinated by interareal spike-field coherence (24). We therefore asked whether MFC neurons phase-lock to ongoing oscillations in the LFP in the HA and, if so, whether the strength of such interactions is modulated by task demands. We performed this analysis for the 13 subjects and 33 sessions for which we simultaneously recorded from both areas (Fig. 5A). In the following, we only used neural activity from the 1-s baseline period that precedes stimulus onset, to avoid confounds related to stimulus onset–evoked activity. Individual cells in the MFC showed strong task modulation of MFC to HA spike-field coherence (Fig. 5B shows a single-cell example in the dACC). At the population level, MFC cells showed significantly stronger theta-band coherence with HA oscillations during the memory task than during the categorization task (Fig. 5C, 8822 cell electrode pairs; $P < 1.3 \times 10^{-7}$, paired t test, measured at 5.5 Hz; see fig. S9, B and C, for additional controls). This was the case for both MFC-hippocampus and MFC-amygdala interactions (Fig. 5D, $n = 3939$ and 4884 and $P = 8.8 \times 10^{-4}$ and 4.3×10^{-5}, respectively, paired t test). This effect was due to changes in phase preference, as there was no significant difference in HA LFP power between the tasks (Fig. 5E, $P = 0.08$, paired t test of signal power at 5.5 Hz, estimate across all 8822 cell-electrode pairs). Of the 767 MFC cells, a significant subset of ~100 cells were phase-locked to the theta-band HA LFP (fig. S9A), with the largest proportion preferring 3 to 8 Hz.

To determine whether there is a relationship between the tuning of cells in MFC and their interarea coherence with HA, we selected for choice cells independently in the categorization and memory task (see Materials and methods for selection model; note selection controls for RT differences). This approach revealed that 101/767 and 82/767 cells were significantly modulated by choice during the memory and categorization task, respectively ($P < 0.001$ versus chance for both; see Fig. 4A, cells 2 and 1, respectively, as an example). Single-neuron decoding showed that it was not possible to decode the subject's choice in the categorization task from choice cells selected in the memory task and vice versa (fig. S5, A to D). Single-neuron analysis revealed that cells preferring either "no" or "yes" choices were present in approximately equal proportions in both tasks (fig. S5B). The removal of the selected choice cells from a population decoding analysis with access to all recorded neurons significantly diminished decoding performance (fig. S5, F and G). Notably, each of the selected cells had a high importance index, as determined from population decoding (Fig. 5F). Considering only the MFC choice cells revealed that this subset of cells similarly increased their phase-locking during the memory task (Fig. 5G, top), with the strongest effect, again, in the theta range [peak frequency (f_{peak}) = 5 Hz, $P = 1 \times 10^{-6}$, paired t test]. Both categorization and memory choice cells showed this pattern of modulation (Fig. 5G, bottom). The memory choice cells exhibited an increase in gamma-band coherence (Fig. 5G, f_{peak} = 38.5 Hz, $P = 2 \times 10^{-6}$, paired t test). The extent of phase-locking of choice cells following stimulus onset (0.2 to 1.2 s) during the memory task was significantly stronger for correctly retrieved trials than for forgotten old trials, indicating behavioral relevance for memory retrieval (Fig. 5H). Lastly, to exclude the possibility that this interarea effect was due to task-dependent changes within the HA, we examined the phase-locking properties of HA cells to their own locally recorded LFP (LFP and spiking activity is recorded on separate electrodes; see Materials and methods). The spiking activity of 331/663 HA cells was significantly related to the theta-frequency band LFP (Fig. 5I, shown for $f = 5.5$ Hz). The strength of this local spike-field coherence was, however, not significantly different between the two tasks (Fig. 5J, $P = 0.61$, paired t test, $n = 2321$ cell-electrode pairs).

Discussion

We investigated the nature of flexible decision-making in the human brain by probing how the strength and/or geometry (44) of neural representations of stimulus memory, stimulus category, and choice is modified when subjects switch between a memory and a categorization task. We found evidence for both kinds of neural representation changes resulting from changing task demands for a subset of the studied variables. In the MFC, both the strength and geometry of representation of stimulus memory changed as a function of task demands. In contrast, in the HA, both the strength and geometry of the representation of stimulus memory were insensitive to task demands (Fig. 3, D and H). Our finding of memory signals in the amygdala supports the hypothesis (47, 48) that the amygdala contributes to recognition memory by signaling stimulus familiarity. Representation strength of stimulus category in both the HA and MFC was stronger in the memory task, but the geometry of this representation was also modulated by the task in the MFC (Fig. 3G, right). Overall, these results show that the geometry of the representations (as assessed by cross-task generalization) of stimulus familiarity and memory were significantly less sensitive to task demands in the HA compared with the MFC (Fig. 3, G and H).

At the population level in the MFC, choices in both the memory and categorization task were decodable with high reliability, but these decoders did not generalize across the two tasks. Choice decoding generalized across subtasks of the categorization task and across changes of response modality in both tasks, indicating that changes in representations were due to switching between a task that requires memory retrieval and one that does not. While the choice signal was significantly weaker in the HA, this same pattern of generalization also held for the HA, suggesting the possibility that the task demand–dependent choice representation we found in the MFC is widely distributed in the brain. A group of task demand–dependent cells in the MFC were choice cells, which preferentially signal behavioral decisions for either memory or categorization decisions irrespective of response modality and regardless of the ground truth of the decision. Thus, from the point of view of downstream areas, neurons formed two separate decision axes: one for memory-based decisions and one for categorization-based decisions. These two decision axes were instantiated selectively so that they were only present when required by the current task.

These findings contrast with prior work on task switching involving different tasks that required purely perceptual decisions, which found a single decision axis in the monkey prefrontal cortex, with task-irrelevant attributes

also represented (*3*). We found that memory-based choices add a second decision axis, which is present only when decisions engage memory retrieval processes. Although task-sensitive representations of choice have been shown in recordings from rodents and nonhuman primates during perceptual decision-making (*2*, *3*, *49*, *50*) and in human neuroimaging (*51*), our data reveal choice representations that specifically signal recognition memory–based choices at the single-cell level. It has long been appreciated that the frontal lobes are critical for initiating and controlling memory retrieval (*30*, *52*–*54*). Neuroimaging reveals that patterns of activity in some frontal and parietal areas are modulated

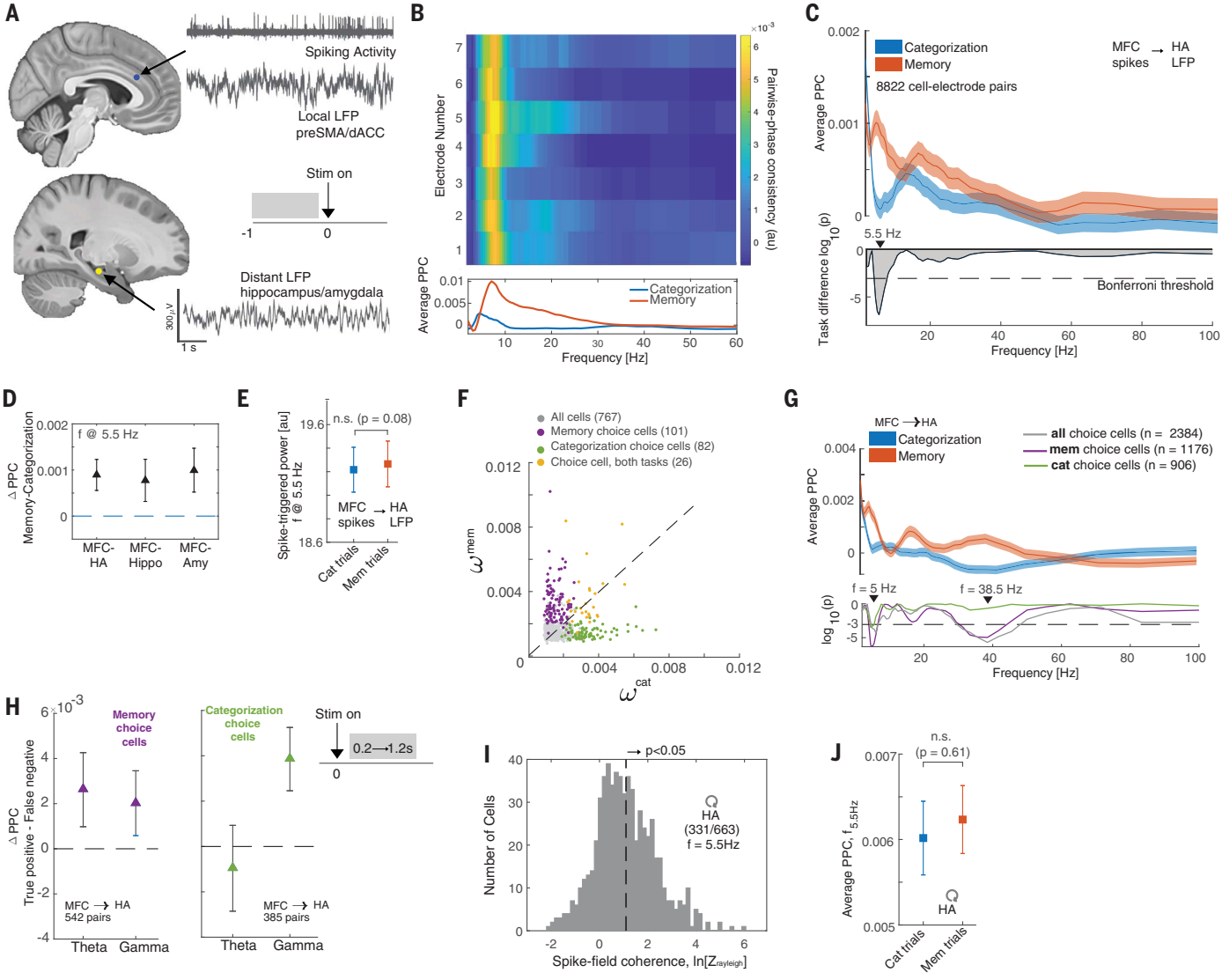


Fig. 5. Modulation of interareal spike-field coherence by task demands. (**A**) Analysis approach. Inset shows that only data from the baseline was used [except in panel (H)]. (**B**) Spike-field coherence for a cell in dACC relative to all channels in the ipsilateral hippocampus. (**C**) Phase-locking of MFC cells to HA LFP. (Top) Average interarea PPC of all cell-electrode pairs for each task. (Bottom) Significance of difference between tasks; peak difference was at f = 5.5 Hz. Dashed line shows the threshold (P = 0.05/56, Bonferroni corrected). (**D**) Difference in average interarea PPC at f = 5.5 Hz between task conditions for all possible cell-electrode pairs (from left to right, n = 8822 electrode pairs, $P = 1.3 \times 10^{-7}$; n = 3938, $P = 8.8 \times 10^{-4}$; n = 4884, $P = 4.3 \times 10^{-5}$; paired t test). (**E**) Average spike-triggered power was not significantly different between the two tasks (paired t test, n = 8822 cell electrode pairs, P = 0.08). (**F**) Single-neuron analysis of choice cells. Importance index assigned by the decoder to each cell for decoding choices in either task. Selected choice cells are indicated in color. (**G**) MFC-HA spike-field coherence for choice cells. (Top) Average PPC for all choice cells in MFC (209 cells, 2384 cell-electrode pairs) for the two tasks. (Bottom) Significance of difference between tasks, shown separately for memory and categorization choice cells (n = 1176 and 906 cell-electrode pairs, respectively). Only memory choice cells show a significant difference in the gamma band ($P = 2 \times 10^{-6}$, t test). (**H**) Difference in spike-field coherence between true-positive (correct retrieval) and false negative (incorrect retrieval) trials measured in the [0.2 s, 1.2 s] window after the stimulus onset, shown separately for memory choice cells (left panel) and categorization choice cells (right panel) in the theta-frequency (4 to 10 Hz) and gamma-frequency (30 to 80 Hz) bands. PPC was significantly stronger in correctly retrieved trials in the theta band for memory choice cells (Δ_{mem} = 0.003, P = 0.002; Δ_{cat} = −0.001, P = 0.3; paired t test) and in the gamma band for both types of choice cells (Δ_{mem} = 0.002, P = 0.005; Δ_{cat} = 0.004, $P = 7.2 \times 10^{-8}$; paired t test). (**I**) Spike times of HA cells are coherent with local theta-band (3 to 8 Hz) LFP. (**J**) Average local PPC in the HA did not differ significantly as a function of the task (f = 5.5 Hz; P = 0.51, paired t test, n = 2321 cell-electrode pairs). Error bars in panels (D) and (H) denote the 95% confidence interval (bootstrap, n = 10,000 iterations). All other error bars are SEM.

by memory retrieval demands (*34*, *35*, *37*, *38*, *55*), whereas memory-related activity patterns in the MTL are comparatively insensitive to retrieval demands (*34*). These findings have led to the proposal that the memory retrieval network consists of specialized processes separate from those used for other kinds of decisions (*6*, *56*, *57*). The memory-choice axis we describe is a potential cellular substrate for this critical aspect of human cognitive flexibility. Future work is needed to investigate whether similar principles also apply to hippocampus-dependent associate or source memory–based (*58*–*60*) decisions, which we did not assess here (we probed recognition memory).

A second group of cells that we characterized in the MFC signal the currently relevant goal (task type and response modality) throughout the task. These cells switched their activity pattern when instructions indicated a change in task demands. Although these switches were rapid, they were not instantaneous, likely reflecting the cost of switching between memory retrieval and categorization modes (*61*–*63*). We hypothesize that these cells facilitate the holding of the active task set in working memory (*64*, *65*) and configure brain networks in preparation for appropriate execution of the instructed task (*12*, *34*, *66*, *67*). Task-switching costs are a much investigated aspect of cognitive flexibility (*39*, *61*–*63*), but how they arise and why some task switches are more difficult than others remain poorly understood. The MFC cells we describe here offer an avenue to directly investigate these questions.

Finally, we uncovered a possible mechanism by which memory-based information can be routed dynamically between the MFC and HA when a task requires memory retrieval. Changing long-range synchronization of neural activity is thought to be a way by which functional connectivity between brain areas can be changed flexibly (*68*–*71*). Here, we reveal a specific instance of this phenomenon at the cellular level in humans in the form of changes in the strength of cortico-hippocampal and cortico-amygdala functional connectivity. Hippocampus–medial prefrontal cortex (mPFC) functional connectivity in rodents supports spatial working memory (*24*) and is prominent during both navigation and rest (*72*–*74*). Similarly, amygdala-mPFC functional connectivity supports flexibly switching between aversive and neutral behaviors depending on learned cues (*75*). But it is not known whether these pathways serve a role in long-term memory retrieval in humans and, if so, whether this retrieval can be engaged selectively. Similarly, in humans, MTL-PFC connectivity changes, as measured by fMRI, have been related to control demands over memory retrieval (*36*, *76*), but it remains unclear how these indirect metrics relate to long-range synchronization as measured in rodents. Here, we show that MFC-HA connectivity is selectively enhanced during the memory task, indicating that patterns of interareal connectivity change in preparation of initiating memory retrieval (*77*, *78*). The extent of synchrony after stimulus onset is stronger when a memory is successfully retrieved compared with when it is forgotten. Memory choice cells in the MFC exhibited enhanced gamma-frequency band coordination of their spiking activity with the HA LFP, and this modulation was behaviorally relevant after stimulus onset, which reveals a specific cellular-level instance of a role for gamma oscillation–mediated coordination of activity between distant brain regions (*24*, *79*) in human memory retrieval.

Materials and methods

Subjects

Subjects were 13 adult patients being evaluated for surgical treatment of drug-resistant epilepsy that provided informed consent and volunteered for this study (see table S1). The institutional review boards of Cedars-Sinai Medical Center and the California Institute of Technology approved all protocols. We excluded potential subjects who did not have at least one depth electrode in medial frontal cortex.

Electrophysiology

We recorded bilaterally from the amygdala, hippocampus, dACC, and pre-SMA using microwires embedded in hybrid depth electrodes (*81*). From each microwire, we recorded the broadband 0.1 to 9000 Hz continuous extracellular signals with a sampling rate of 32 kHz (ATLAS system, Neuralynx Inc.). Subjects from which at least one well-identified single neuron could not be recorded were excluded.

Spike sorting and single-neuron analysis

The raw signal was filtered with a zero-phase lag filter in the 300 to 3000 Hz band, and spikes were detected and sorted using a semiautomated template-matching algorithm (*82*, *83*). All PSTH diagrams were computed using a 500-ms window with a step size of 7.8 ms. No smoothing was applied.

Electrode localization (relevant for Fig. 1)

Electrode localization was performed based on postoperative MRI scans. These scans were registered to preoperative MRI scans using Freesurfer's mri_robust_register (*84*) to allow accurate and subject-specific localization. To summarize electrode positions and to provide across-study comparability, we also aligned the preoperative scan to the MNI152-aligned CIT168 template brain (*85*) using a concatenation of an affine transformation followed by a symmetric image normalization (SyN) diffeomorphic transform (*86*). This procedure provided the MNI coordinates that are reported here for every recording location. The electrode locations shown on the Atlas Brain (Fig. 1, C and D) are for illustration only. Apparent localization outside the target area or in white matter is due to usage of an Atlas Brain alone.

Eye tracking (relevant for fig. S1)

Gaze position was monitored using an infrared-based eye tracker with a 500-Hz sampling rate (EyeLink 1000, SR Research) (*87*). Calibration was performed using the built-in nine-point calibration grid and was only used if validation resulted in a measurement error of <1 degree of visual angle (dva) (average validation error was 0.7 dva). We used the default values for the thresholds in the Eyelink system that determine fixation and saccade (eye movement) onsets.

Task

Each session consisted of eight blocks of 40 trials shown in randomized order. At the beginning of each block, an instruction screen told subjects verbally the task to be performed for the following 40 trials (categorization or recognition memory), the response modality to use (button presses or eye movements), and which visual category is the target (for categorization task only; either human faces, monkey faces, fruits, or cars; order was pseudorandom so that each image type was selected as the target at least once) (see Fig. 1). The task to solve was either "Have you seen this image before, yes or no?" or "Does this image belong to the target category, yes or no." Odd-numbered blocks (1, 3, 5, and 7) were categorization blocks; even numbered blocks were memory blocks (2, 4, 6, and 8). Button presses (yes or no) were recorded using a response box (RB-844, Cedrus Inc.). Eye movements to the left or right of the image served as responses in the eye movement modality (left = yes, right = no). The mapping between button and screen side and yes/no responses was fixed and did not change; "yes" was on the left and "no" was on the right. Subjects were reminded that left = yes, and right = no, at the beginning of each of the eight blocks. In the first block, all images were novel (40 unique images). In all subsequent blocks, 20 new images were shown randomly intermixed with 20 repeated images (the "old set"). The 20 repeated images remained the same throughout a session. We used entirely nonoverlapping image sets for patients that completed multiple sessions. The response modality (button presses or eye movements) was initially selected randomly and then was switched in the middle of each block (an instruction screen in the middle of each block showed the response modality to be used for the remainder of the block). In sessions where eye tracking was not possible because of problems with calibration (five sessions in three patients; see table S1), all trials used the button presses as the response modality. No trial-by-trial feedback was given. In between image presentations, subjects were

instructed to look at the fixation cross in the center of the screen.

Control task (relevant for fig. S6)

In 5 of the 13 subjects in this dataset (6/33 sessions), we ran an additional control task in order to help determine whether neural responses reflected processing of stimuli, of decision variables, or of motor response plans. Unlike the standard task where the subjects could respond at any time after the stimulus onset (thus making it difficult to distinguish decision from choice), in this control task the subjects were instructed to wait until the response cue in order to register their answer, either with a button press or with a saccade. The stimulus was presented for a fixed amount of time (1-s duration) and after a 0.5-to-1.5-s delay period, the subjects were asked to respond to the question relevant for that block.

Mixed-effects modeling of behavior (relevant for Fig. 1 and fig. S1)

For the group analysis of behavior, we used mixed-effects models of the form $y = X\beta + Zb + \varepsilon$, where y is the response, X is the fixed-effects design matrix, β is the fixed-effects coefficients, Z is the random-effects design matrix, b is the random-effects coefficient, and ε is the error vector. In all analyses, we used a random intercept model with a fixed slope. The grouping variable for the random effects was the session ID. The reported P values in the main text correspond to the fixed-intercept for the relevant variable. In the case of measuring the effect of number of expositions (i.e., number of times an image was seen) on the subject's accuracy during the memory trials, we used a mixed-effects logistic regression with the independent variable as an ordinal-valued whole number ranging from 1 to 7. The response was a logical value indicating success or failure on each memory question. Prior to running any analysis of reaction time data, we excluded outliers from the distribution using the following procedure: A sample was considered an outlier if it was outside the 99th percentile of the empirical distribution.

Reaction time matching procedure

As a control, we matched for RTs between the two tasks (categorization and memory) to exclude for potential differences due to difficulty. To achieve this, we first added noise to all reaction times (SD = 1 ms), followed by locating pairs of trials with RTs that were equal to within a tolerance of 0.1 s. Matching pairs were then removed, and this procedure was repeated iteratively until no further matches could be found. Unmatched trials were excluded (resulting in reduced statistical power owing to fewer available trials). We only used the resulting match if the RTs between the two groups were not significantly different. If not, the procedure above was repeated.

Selection of visually (VS) and memory-selective (MS) cells (relevant for fig. S3)

A cell was considered a VS cell if its response covaried significantly with visual category, as assessed using a 1×4 analysis of variance (ANOVA) test at $P < 0.05$. For each selected cell, the preferred image category was set to be the image category for which the mean firing rate of the cell was the greatest. All trials were used for this analysis. MS cells were selected using the following linear model

$$\mathrm{fr}_{\mathrm{cell}} \sim 1 + \beta_{1 \bullet \mathrm{category}} + \beta_{2 \bullet \mathrm{new/old}} + \beta_{3 \bullet \mathrm{rt}}$$

where "category" is a categorical (1×4) variable, "new/old" is a binary variable, and "rt" is a continuously valued variable. A cell was determined to be memory-selective if the t-statistic for $\beta 2$ was significant with $P < 0.05$. We excluded the first block of trials (40 images) from the analysis, in order to keep the number of new and old stimuli the same. Spikes were counted for every trial in a 1-s window starting at 200 ms after stimulus onset.

Selection of choice cells (relevant for fig. S5)

Choice cells were selected using a regression model applied to the firing rate in a 1-s-size window starting 200 ms after stimulus onset. We fit the following regression model

$$\mathrm{fr}_{\mathrm{cell}} \sim 1 + \beta_{1 \bullet \mathrm{category}} + \beta_{2 \bullet \mathrm{response}} + \beta_{3 \bullet \mathrm{rt}}$$

where the response is binary (yes or no), category is a categorical variable with four levels, and RT is the reaction time. We fit this model separately to trials in the memory and categorization condition, assuring independent selection of cells. RT was included as a nuisance regressor to control for reaction time differences between the two possible responses (see fig. S1A). A cell qualified as a choice cell if the t-statistic of the β_2 term was significant at $P < 0.05$ for at least one of the two task conditions. The response preference of significant cells for either yes or no was determined based on the sign of β_2 (positive = yes, negative = no). Notice that the selection process uses separate trials for memory choice cells and categorization choice cells. All trials, regardless of whether the answer was correct or incorrect, were used for selection. To estimate the significance of the number of selected cells, we generated a null distribution by repeating the above selection process 1000 times after randomly reshuffling the response label. We estimated this null distribution separately for choice cells in the memory and categorization condition and used each to estimate the significance of the number of selected cells of each type.

Chance levels for cell selection (relevant for figs. S2, S3, and S5)

To estimate the chance levels for cell selection, we repeated above procedures for selection of visual category, memory selective, and choice cells after randomly scrambling the order of the labels determining the category membership being selected for (yes/no response, visual category, and new/old ground truth, respectively). We repeated this procedure 1000 times.

Single-cell decoding (relevant for fig. S5)

Single-cell decoding was done using a Poisson naïve Bayes decoder. The features used were spike counts in a 1-s window, in the interval [0.2 s, 1.2 s] relative to stimulus onset. The decoder returns the probability of a class label, given the observed spike count. The class label was binary ("yes" or "no"). The model assumes that the spike count is generated by a univariate Poisson distribution, and a separate mean rate parameter (λ) is fit to each feature-class pair. For a new observation, class membership is determined on the basis of the likelihood value. Notice that we used a single spike count as a feature, so the naïve assumption of the decoder is no longer relevant in this case.

Population decoding (relevant for Figs. 2, 3, and 4 and figs. S2 to S5, S7, and S8)

Single-trial population decoding was performed on a pseudo-population assembled across sessions (*88*). We present decoding results for a variety of task variables: (i) image category, (ii) new versus old, (iii) choice during memory trials, (iv) choice during categorization, (v) task, and (vi) response type. In order to estimate the variance of the decoding performance, on each iteration of the decoder (minimum of 250 iterations), we randomly selected 75% of the cell population that was being analyzed. For example, to measure choice decoding in MFC (as shown in Fig. 4), we would randomly select 575/767 cells on each iteration of the decoder. The total number of available cells depended on the variable that was being decoded. For example, for response type decoding, the number of cells in MFC was 593, because 28/33 sessions included both response types. We matched the number of trials per condition contributed by each cell that was selected to participate in the population decoding. For most task variables (image category, new/old, context, effector-type), the number of samples from each cell was equal, because the task structure remained the same across all subjects and sessions. For choice decoding, however, the number of instances varied, because the subjects were free to respond with a "yes" or a "no" for each stimulus. We therefore matched the numbers for the smallest group across all subjects. This matching procedure can further reduce the number of cells we included in the decoding that do not have the minimum

number of trials needed per condition. For the population decoding and cross-condition generalization of familiarity presented in Fig. 3, we used all image categories for which the subjects showed above-chance recognition performance (see fig. S1B). Therefore we used images of cars, fruits, and faces, excluding images of monkeys from the analysis. For all other analysis, we used all available categories.

A series of preprocessing steps were carried out before training the decoder. Firing rates for each cell were first de-trended (to account for any drift in the baseline-firing rate) and then normalized (z-scored) using the mean and standard deviation estimated from the training set. We then performed 10-fold cross validation using a linear support vector machine (SVM) decoder to estimate performance, as implemented by the "fitcecoc" function in MATLAB. We used an SVM with a linear kernel and a scale of 1. Decoding results are reported either as a function of time or in a fixed time window. Time-resolved decoding was done on spike counts measured in a 500-ms moving window, with a 16-ms step size. For fixed-window decoding, we used spike counts in a 1-s window. The location of the window depended on the analysis. In Fig. 2, for example, we used a [−1, 0] window relative to stimulus onset for task type and response type decoding. In Fig. 3, we used spike counts in a [0.2, 1.2] window relative to stimulus onset for decoding image category and new/old.

Null models for testing significance of decoding performance

Throughout the manuscript, we compare the performance of our decoders against the 95th percentile of a null distribution. The way that this null distribution is generated depends on the variable being decoded. For variables such as image category, new versus old, and response (i.e., yes versus no), we used a simple shuffling procedure for the labels. For variables such as task type, which had structure over time (memory blocks were always preceded by categorization block), small drifts in firing rate might lead to inflated decoding accuracy. Therefore, for such variables, the shuffling was done in such a way as to preserve their temporal relationship. Specifically, we offset (i.e., circular shift) the labels by a random integer value (sampled from the range ±10 to 20 trials). In the case of task decoding from the baseline firing rate, this is a very conservative measure of the null decoding performance, because many trials retain their original label, thereby inflating the accuracy. This also means that the mean performance of the null distribution will not be the theoretical chancel level. In the case of task decoding, the theoretical chance level is 50% (binary classification). Using the circular shift method for scrambling labels, the mean of the null distribution was ~60%.

To compare the performance between different decoders, for example choice decoding from the HA versus MFC population, we constructed an empirical null distribution from the pairwise differences in the performance of these two decoders trained using the shuffled labels. For example, if we get N estimates of the null performance (i.e., after shuffling the labels) of the HA decoder and N estimates of the null performance of the MFC decoder, we construct a distribution of the N•N = N^2 pairwise differences. We can then compute the significance of the true difference in decoding performance between MFC and HA, Δ_{true}, relative to this distribution. The variance of the null distribution is sensitive to the number of trials available for decoding because it changes the resolution (step size) by which decoding accuracy can change. For example, for 10 trials, the accuracy can take values from 0 to 1 in increments of 0.1. This results in different values for the 95th percentile of the null distribution and is the reason why in some cases a given difference in decoding accuracy is significant while it is not in others. Unless otherwise specified, all *P* values for comparing decoding performance between conditions or brain areas are calculated using this approach. In the one case where the number of trials in a condition was too low to reliably estimate the null distribution (fig. S5I) and for comparing the generalization index (fig. 3J) we used a bootstrap test for equality of means (*89*) to compare the two conditions to assign a *P* value to the true difference (repeated 1000 times to estimate the null distribution).

Multidimensional scaling (relevant for figs. 3 and 4)

Multidimensional scaling (MDS) was used only for visualization. We computed MDS using Euclidean distances (MATLAB function "mdscale') on z-scored spike count data in the [0.2 s, 1.2 s] window relative to image onset. In Fig. 3E, for example, MDS was computed on the activity across the entire population of HA and MFC cells, averaged across the eight conditions plotted (new/old ⊗ task ⊗ image category, where ⊗ denotes the Cartesian product). Here the image category was restricted to images of human faces and fruits, for visualization purposes. For the cross-condition generalization performance, we use all four image categories. In Fig. 4D we compute MDS on the population of MFC cells, averaged across eight conditions (response ⊗ task ⊗ effector, where ⊗ denotes the Cartesian product). In all cases, we use MDS to map the neural activity to a 3D space.

Normalized weight metric (relevant for Figs. 4 and 5 and figs. S5, S7, S8, and S10)

The normalized weight metric is computed from the weight that a decoder assigns to a particular cell for a given classification. This weight is denoted as w_i^t, where the index i denotes the cell, and the index t denotes the condition (for example, categorization or memory). The weight is converted into a normalized measure called an importance index, defined as

$$\omega_i^t = \frac{\left|w_i^t\right|}{\sum_i^n \left|w_i^t\right|}$$

State-space analysis (relevant for Fig. 4I)

We used Gaussian-process factor analysis (GPFA) (*45*) to analyze the dynamics of the average population activity for the eight conditions arising from the combination of choice (yes, no), response modality (button press, saccade), and task (memory, categorization). The recovered latent space was eight-dimensional (8D), and all similarity measurements between trajectories were performed in this space (not in the 3D projections shown in the figure). The activity was binned using 20-ms windows. All analysis was computed and visualized using the DataHigh (*90*) MATLAB toolbox. Similarity measurements between two conditions were computed and averaged over the first 500ms after stimulus onset as follows

$$sim\left(\boldsymbol{r_1}(t), \boldsymbol{r_2}(t)\right) = \frac{\boldsymbol{r_1'}(t)}{\left\|\boldsymbol{r_1'}(t)\right\|} \cdot \frac{\boldsymbol{r_2'}(t)}{\left\|\boldsymbol{r_2'}(t)\right\|}$$

where $\boldsymbol{r_1}(t)$ and $\boldsymbol{r_2}(t)$ are the 8D state-space trajectories for conditions 1 and 2 [$\boldsymbol{r_1'}(t)$ and $\boldsymbol{r_2'}(t)$ indicate the derivatives over time], respectively.

ANOVA model (relevant for figs. S4 and S11)

We used a single-cell ANOVA model to tease apart the contributions of choice, visual category, memory, and response time on the firing rate of a cell. The model was of the following form

$$\mathrm{fr_{cell}} \sim \beta_1 \bullet \mathrm{category} + \beta_2 \bullet \mathrm{familiarity} + \beta_3 \bullet \mathrm{choice} + \beta_4 \bullet \mathrm{rt}$$

where $\mathrm{fr_{cell}}$ is the mean firing rate in a fixed window (0.2 to 1.2 s following stimulus onset) or a moving window of 500 ms to analyze the time course. The ANOVA model is fit independently at each point of time. We then compute the F-statistic for each of the regressors and report the average F-statistic across the entire population of recorded cells, fit twice to each cell for the memory and categorization task (figs. S4, D and E, fig. S11). To compare the effects of task on the representation of individual variables, we compare the distribution of F-statistics estimated separately on each task, for each cell in the population. We use this approach as a measure of modulation in the strength of representation for a variable induced by task switching. This comparison does not make predictions about generalizability from one task to the next, because the model is fit independently.

Generalization index (relevant for Figs. 3 and 4)

To compare the within-condition decoding to the across condition generalization, we used a generalization index defined as following

$$g = \frac{\text{cross} - \text{chance}}{\text{within} - \text{chance}}$$

where "within" is the decoding performance within condition, "cross" is the decoding across condition, and "chance" is the chance decoding performance for the variable of interest (choice = 0.5, new/old = 0.5, familiarity = 0.5, image category = 0.25).

Spike-field coherence analysis (relevant for Fig. 5 and fig. S9)

LFP preprocessing

The local-field potential recordings were high-pass filtered at 1 Hz. The raw recordings, sampled at 32 kHz, were then downsampled to 500 Hz. The downsampling procedure was done with the "resample" command in MATLAB, which applies the appropriate antialiasing filter prior to reducing the sampling rate. For each session, we screened all MFC and HA electrodes in order to make sure that there were no artifacts that could contaminate the spike-field metrics. We excluded all electrodes with interictal discharges (IEDs) visible in the raw trace (by visual inspection). Specifically, in screening for IEDs, we looked for large stereotyped, recurring transients in the raw recording that did not correspond to cellular spiking activity. The presence of such transients would disqualify an electrode from further consideration.

Spike-field coherence (SFC)

All spike-field coherence analysis was performed on snippets of the LFP extracted around the spike. We extract snippets for every cell-electrode pair. For example, to measure interarea SFC between a single cell in pre-SMA and HA LFPs, we extracted n snippets each (n = number of spikes) from each of the eight ipsilateral electrodes in hippocampus and eight electrodes in the ipsilateral amygdala. For sessions where we used a local reference (i.e., bipolar referencing), we exclude the reference wire. For intra-area coherence (for example, HA spikes to HA field) we also exclude the wire on which the cell was recorded to avoid contamination by spike waveform. For each snippet and for each cell-electrode pair, we compute the spike-triggered spectrum using the FieldTrip "mtmconvol" method, which computes the Fourier spectrum of the LFP around the spikes using convolution of the complete LFP traces. The spectrum was computed with a single "hanning" taper, at 56 logarithmically spaced frequencies ranging from 2 Hz on the low end to 125 Hz on the high end. The length of the snippet window was dynamic as a function of the frequency examined; the snipped length was set to equal to two cycles of the underlying frequency at which the spectrum was estimated (i.e., 2 Hz→2 s snippet). We estimated the phase for each snippet and for each of the 56 frequencies from the complex-valued Fourier coefficients (i.e., phasor). We used the pairwise phase consistency (PPC) metric as the measure of coherence. For the spike-triggered power, we compute the magnitude of the spectral coefficients returned by the Fourier transform (also computed for each cell-electrode pair) for each snippet and averaged the spectra. Unless otherwise stated, all SFC results in the paper are based on spikes recorded during the baseline period between trials (1-s window preceding stimulus onset).

Group comparisons using the SFC metric

When comparing two or more groups using PPC (such as memory versus categorization), we balanced the number of spikes between the two groups. To reduce bias involved in subsampling the larger group, we resampled the spikes from the two groups 200 times and computed the PPC metric on each iteration. The final coherence measure for a given cell-electrode pair was an average across all 200 iterations.

To ensure that the underlying local field potential does not vary in a consistent way across conditions, we compare the distribution of average voltage values for each of the conditions in our spike-field coherence analysis. In the case of the task contrast during baseline (i.e., memory versus categorization), we show the distribution of area under the curve (AUC) values computed separately for each electrode in the amygdala and hippocampus (fig. S9D shows that there was no significant difference). The AUC for each electrode is computed using the average baseline magnitude across memory and categorization trials. In the case of the spike-field coherence results during the stimulus onset (Fig. 5H), to reduce any potential confounds related to event-related potentials, we used only sessions with local referencing (bipolar). The local reference (set to one of the eight microwires in the electrode cluster implanted in each brain area) significantly diminishes the magnitude of any event-related potentials after stimulus onset. To confirm this, we repeated the AUC analysis mentioned above, for the contrast in Fig. 5H [i.e., true positive (TP) versus false negative (FN)]. The results (shown in fig. S9E) show that there is no significant difference between the two conditions of interest.

REFERENCES AND NOTES

1. L. R. Squire, J. T. Wixted, The cognitive neuroscience of human memory since H.M. *Annu. Rev. Neurosci.* **34**, 259–288 (2011). doi: 10.1146/annurev-neuro-061010-113720; pmid: 21456960
2. D. J. Freedman, J. A. Assad, Neuronal mechanisms of visual categorization: An abstract view on decision making. *Annu. Rev. Neurosci.* **39**, 129–147 (2016). doi: 10.1146/annurev-neuro-071714-033919; pmid: 27070552
3. V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013). doi: 10.1038/nature12742; pmid: 24201281
4. J. I. Gold, M. N. Shadlen, The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007). doi: 10.1146/annurev.neuro.29.051605.113038; pmid: 17600525
5. R. Desimone, J. Duncan, Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995). doi: 10.1146/annurev.ne.18.030195.001205; pmid: 7605061
6. M. D. Rugg, K. L. Vilberg, Brain networks underlying episodic memory retrieval. *Curr. Opin. Neurobiol.* **23**, 255–260 (2013). doi: 10.1016/j.conb.2012.11.005; pmid: 23206590
7. A. M. Bornstein, K. A. Norman, Reinstated episodic context guides sampling-based decisions for reward. *Nat. Neurosci.* **20**, 997–1003 (2017). doi: 10.1038/nn.4573; pmid: 28581478
8. M. N. Shadlen, D. Shohamy, Decision making and sequential sampling from memory. *Neuron* **90**, 927–939 (2016). doi: 10.1016/j.neuron.2016.04.036; pmid: 27253447
9. Z. Fu *et al.*, Single-neuron correlates of error monitoring and post-error adjustments in human medial frontal cortex. *Neuron* **101**, 165–177.e5 (2019). doi: 10.1016/j.neuron.2018.11.016; pmid: 30528064
10. A. Shenhav, M. M. Botvinick, J. D. Cohen, The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013). doi: 10.1016/j.neuron.2013.07.007; pmid: 23889930
11. C. B. Holroyd *et al.*, Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.* **7**, 497–498 (2004). doi: 10.1038/nn1238; pmid: 15097995
12. N. U. Dosenbach *et al.*, A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006). doi: 10.1016/j.neuron.2006.04.031; pmid: 16731517
13. M. F. Rushworth, M. E. Walton, S. W. Kennerley, D. M. Bannerman, Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* **8**, 410–417 (2004). doi: 10.1016/j.tics.2004.07.009; pmid: 15350242
14. J. Duncan, The structure of cognition: Attentional episodes in mind and brain. *Neuron* **80**, 35–50 (2013). doi: 10.1016/j.neuron.2013.09.015; pmid: 24094101
15. J. M. Hyman, L. Ma, E. Balaguer-Ballester, D. Durstewitz, J. K. Seamans, Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5086–5091 (2012). doi: 10.1073/pnas.1114415109; pmid: 22421138
16. K. L. Anderson, R. Rajagovindan, G. A. Ghacibeh, K. J. Meador, M. Ding, Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory. *Cereb. Cortex* **20**, 1604–1612 (2010). doi: 10.1093/cercor/bhp223; pmid: 19861635
17. A. J. Watrous, N. Tandon, C. R. Conner, T. Pieters, A. D. Ekstrom, Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nat. Neurosci.* **16**, 349–356 (2013). doi: 10.1038/nn.3315; pmid: 23354333
18. J. S. Simons, H. J. Spiers, Prefrontal and medial temporal lobe interactions in long-term memory. *Nat. Rev. Neurosci.* **4**, 637–648 (2003). doi: 10.1038/nrn1178; pmid: 12894239
19. J. Y. Yu, L. M. Frank, Hippocampal-cortical interaction in decision making. *Neurobiol. Learn. Mem.* **117**, 34–41 (2015). doi: 10.1016/j.nlm.2014.02.002; pmid: 24530374
20. J. A. Gordon, Oscillations and hippocampal-prefrontal synchrony. *Curr. Opin. Neurobiol.* **21**, 486–491 (2011). doi: 10.1016/j.conb.2011.02.012; pmid: 21470846
21. S. L. Brincat, E. K. Miller, Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nat. Neurosci.* **18**, 576–581 (2015). doi: 10.1038/nn.3954; pmid: 25706471
22. E. Likhtik, R. Paz, Amygdala-prefrontal interactions in (mal) adaptive learning. *Trends Neurosci.* **38**, 158–166 (2015). doi: 10.1016/j.tins.2014.12.007; pmid: 25583269
23. A. Burgos-Robles *et al.*, Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nat. Neurosci.* **20**, 824–835 (2017). doi: 10.1038/nn.4553; pmid: 28436980
24. T. Spellman *et al.*, Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature* **522**, 309–314 (2015). doi: 10.1038/nature14445; pmid: 26053122
25. T. Sigurdsson, S. Duvarci, Hippocampal-prefrontal interactions in cognition, behavior and psychiatric disease. *Front. Syst. Neurosci.* **9**, 190 (2016). pmid: 26858612
26. M. Remondes, M. A. Wilson, Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. *Neuron* **80**,

1277–1289 (2013). doi: 10.1016/j.neuron.2013.08.037; pmid: 24239123
27. N. Karalis *et al*., 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. *Nat. Neurosci.* **19**, 605–612 (2016). doi: 10.1038/nn.4251; pmid: 26878674
28. E. Likhtik, J. M. Stujenske, M. A. Topiwala, A. Z. Harris, J. A. Gordon, Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. *Nat. Neurosci.* **17**, 106–113 (2014). doi: 10.1038/nn.3582; pmid: 24241397
29. J. M. Hyman, M. E. Hasselmo, J. K. Seamans, What is the functional relevance of prefrontal cortex entrainment to hippocampal theta rhythms? *Front. Neurosci.* **5**, 24 (2011). doi: 10.3389/fnins.2011.00024; pmid: 21427795
30. M. Lepage, O. Ghaffar, L. Nyberg, E. Tulving, Prefrontal cortex and episodic memory retrieval mode. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 506–511 (2000). doi: 10.1073/pnas.97.1.506; pmid: 10618448
31. I. G. Dobbins, H. Foley, D. L. Schacter, A. D. Wagner, Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron* **35**, 989–996 (2002). doi: 10.1016/S0896-6273(02)00858-9; pmid: 12372291
32. N. U. Dosenbach *et al*., Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11073–11078 (2007). doi: 10.1073/pnas.0704320104; pmid: 17576922
33. Y. Miyashita, Cognitive memory: Cellular and network machineries and their top-down control. *Science* **306**, 435–440 (2004). doi: 10.1126/science.1101864; pmid: 15486288
34. M. D. Rugg, P. C. Fletcher, C. D. Frith, R. S. Frackowiak, R. J. Dolan, Brain regions supporting intentional and incidental memory: A PET study. *Neuroreport* **8**, 1283–1287 (1997). doi: 10.1097/00001756-199703240-00045; pmid: 9175130
35. S. E. Favila, R. Samide, S. C. Sweigart, B. A. Kuhl, Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. *J. Neurosci.* **38**, 7809–7821 (2018). doi: 10.1523/JNEUROSCI.0564-18.2018; pmid: 30054390
36. A. J. Westphal, S. Wang, J. Rissman, Episodic memory retrieval benefits from a less modular brain network organization. *J. Neurosci.* **37**, 3523–3531 (2017). doi: 10.1523/JNEUROSCI.2509-16.2017; pmid: 28242796
37. K. L. Vilberg, M. D. Rugg, The neural correlates of recollection: Transient versus sustained FMRI effects. *J. Neurosci.* **32**, 15679–15687 (2012). doi: 10.1523/JNEUROSCI.3065-12.2012; pmid: 23136408
38. X. Xiao *et al*., Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986–2998 (2017). doi: 10.1523/JNEUROSCI.2324-16.2017; pmid: 28202612
39. D. Badre, A. D. Wagner, Computational and neurobiological mechanisms underlying cognitive flexibility. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7186–7191 (2006). doi: 10.1073/pnas.0509550103; pmid: 16632612
40. S. M. Polyn, M. J. Kahana, Memory search and the neural representation of context. *Trends Cogn. Sci.* **12**, 24–30 (2008). doi: 10.1016/j.tics.2007.10.010; pmid: 18069046
41. U. Rutishauser *et al*., Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nat. Neurosci.* **18**, 1041–1050 (2015). doi: 10.1038/nn.4041; pmid: 26053402
42. G. Kreiman, C. Koch, I. Fried, Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* **3**, 946–953 (2000). doi: 10.1038/78868; pmid: 10966627
43. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, I. Fried, Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005). doi: 10.1038/nature03687; pmid: 15973409
44. S. Bernardi *et al*., The geometry of abstraction in hippocampus and pre-frontal cortex. bioRxiv408633 [Preprint]. 4 October 2019; https://doi.org/10.1101/408633.
45. M. Y. Byron *et al*., Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Adv. Neural Inf. Process. Syst.* **21**, 1881–1888 (2008).
46. F. Stefanini *et al*., A distributed neural code in the dentate gyrus and in CA1. *Neuron* 10.1016/j.neuron.2020.05.022 (2020).
47. A. Farovik, R. J. Place, D. R. Miller, H. Eichenbaum, Amygdala lesions selectively impair familiarity in recognition memory. *Nat. Neurosci.* **14**, 1416–1417 (2011). doi: 10.1038/nn.2919; pmid: 21946327
48. F. A. Wilson, E. T. Rolls, The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Exp. Brain Res.* **93**, 367–382 (1993). doi: 10.1007/BF00229353; pmid: 8519331
49. G. N. Pho, M. J. Goard, J. Woodson, B. Crawford, M. Sur, Task-dependent representations of stimulus and choice in mouse parietal cortex. *Nat. Commun.* **9**, 2596 (2018). doi: 10.1038/s41467-018-05012-y; pmid: 29968709
50. E. Hoshi, K. Shima, J. Tanji, Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *J. Neurophysiol.* **83**, 2355–2373 (2000). doi: 10.1152/jn.2000.83.4.2355; pmid: 10758139
51. C. Sestieri, G. L. Shulman, M. Corbetta, The contribution of the human posterior parietal cortex to episodic memory. *Nat. Rev. Neurosci.* **18**, 183–192 (2017). doi: 10.1038/nrn.2017.6; pmid: 28209980
52. D. Badre, R. A. Poldrack, E. J. Paré-Blagoev, R. Z. Insler, A. D. Wagner, Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* **47**, 907–918 (2005). doi: 10.1016/j.neuron.2005.07.023; pmid: 16157284
53. A. P. Shimamura, "Memory and frontal lobe function" in *The Cognitive Neurosciences*, M. S. Gazzaniga, Ed. (MIT Press, 1995), pp. 803–813.
54. M. Moscovitch, Memory and working-with-memory: A component process model based on modules and central systems. *J. Cogn. Neurosci.* **4**, 257–267 (1992). doi: 10.1162/jocn.1992.4.3.257; pmid: 23964882
55. D. I. Donaldson, S. E. Petersen, J. M. Ollinger, R. L. Buckner, Dissociating state and item components of recognition memory using fMRI. *Neuroimage* **13**, 129–142 (2001). doi: 10.1006/nimg.2000.0664; pmid: 11133316
56. A. D. Wagner, B. J. Shannon, I. Kahn, R. L. Buckner, Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* **9**, 445–453 (2005). doi: 10.1016/j.tics.2005.07.001; pmid: 16054861
57. C. Sestieri *et al*., Memory accumulation mechanisms in human cortex are independent of motor intentions. *J. Neurosci.* **34**, 6993–7006 (2014). doi: 10.1523/JNEUROSCI.3911-13.2014; pmid: 24828652
58. L. Davachi, J. P. Mitchell, A. D. Wagner, Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2157–2162 (2003). doi: 10.1073/pnas.0337195100 pmid: 12578977
59. C. E. Stark, P. J. Bayley, L. R. Squire, Recognition memory for single items and for associations is similarly impaired following damage to the hippocampal region. *Learn. Mem.* **9**, 238–242 (2002). doi: 10.1101/lm.51802; pmid: 12359833
60. H. Eichenbaum, A. P. Yonelinas, C. Ranganath, The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* **30**, 123–152 (2007). doi: 10.1146/annurev.neuro.30.051606.094328; pmid: 17417939
61. S. Monsell, Task switching. *Trends Cogn. Sci.* **7**, 134–140 (2003). doi: 10.1016/S1364-6613(03)00028-7; pmid: 12639695
62. U. Mayr, R. Kliegl, Task-set switching and long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* **26**, 1124–1140 (2000). doi: 10.1037/0278-7393.26.5.1124; pmid: 11009248
63. M. L. Dixon, K. C. Fox, K. Christoff, A framework for understanding the relationship between externally and internally directed cognition. *Neuropsychologia* **62**, 321–330 (2014). doi: 10.1016/j.neuropsychologia.2014.05.024; pmid: 24912071
64. J. Kamiński *et al*., Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* **20**, 590–601 (2017). doi: 10.1038/nn.4509; pmid: 28218914
65. S. Kornblith, R. Quian Quiroga, C. Koch, I. Fried, F. Mormann, Persistent single-neuron activity during working memory in the human medial temporal lobe. *Curr. Biol.* **27**, 1026–1032 (2017). doi: 10.1016/j.cub.2017.02.013; pmid: 28318972
66. E. K. Miller, J. D. Cohen, An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001). doi: 10.1146/annurev.neuro.24.1.167; pmid: 11283309
67. 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
68. G. Hahn, A. Ponce-Alvarez, G. Deco, A. Aertsen, A. Kumar, Portraits of communication in neuronal networks. *Nat. Rev. Neurosci.* **20**, 117–127 (2019). doi: 10.1038/s41583-018-0094-0; pmid: 30552403
69. M. Siegel, T. H. Donner, A. K. Engel, Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* **13**, 121–134 (2012). doi: 10.1038/nrn3137; pmid: 22233726
70. A. Z. Harris, J. A. Gordon, Long-range neural synchrony in behavior. *Annu. Rev. Neurosci.* **38**, 171–194 (2015). doi: 10.1146/annurev-neuro-071714-034111; pmid: 25897876
71. P. Fries, A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474–480 (2005). doi: 10.1016/j.tics.2005.08.011; pmid: 16150631
72. M. W. Jones, M. A. Wilson, Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLOS Biol.* **3**, e402 (2005). doi: 10.1371/journal.pbio.0030402; pmid: 16279838
73. A. G. Siapas, E. V. Lubenov, M. A. Wilson, Prefrontal phase locking to hippocampal theta oscillations. *Neuron* **46**, 141–151 (2005). doi: 10.1016/j.neuron.2005.02.028; pmid: 15820700
74. H. T. Ito, E. I. Moser, M.-B. Moser, Supramammillary nucleus modulates spike-time coordination in the prefrontal-thalamo-hippocampal circuit during navigation. *Neuron* **99**, 576–587.e5 (2018). doi: 10.1016/j.neuron.2018.07.021; pmid: 30092214
75. O. Klavir, R. Genud-Gabai, R. Paz, Functional connectivity between amygdala and cingulate cortex for adaptive aversive learning. *Neuron* **80**, 1290–1300 (2013). doi: 10.1016/j.neuron.2013.09.035; pmid: 24314732
76. R. G. Benoit, M. C. Anderson, Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron* **76**, 450–460 (2012). doi: 10.1016/j.neuron.2012.07.025; pmid: 23083745
77. Y. Ezzyat *et al*., Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. *Nat. Commun.* **9**, 365 (2018). doi: 10.1038/s41467-017-02753-0; pmid: 29410414
78. Y. Ezzyat *et al*., Direct brain stimulation modulates encoding states and memory performance in humans. *Curr. Biol.* **27**, 1251–1258 (2017). doi: 10.1016/j.cub.2017.03.028; pmid: 28434860
79. J. Yamamoto, J. Suh, D. Takeuchi, S. Tonegawa, Successful execution of working memory linked to synchronized high-frequency gamma oscillations. *Cell* **157**, 845–857 (2014). doi: 10.1016/j.cell.2014.04.009; pmid: 24768692
80. J. Minxha, R. Adolphs, S. Fusi, A. Mamelak, U.Rutishauser, Data for: Flexible recruitment of memory-based choice representations by humanmedial-frontal cortex, OSF (2020); https://doi.org/10.17605/OSF.IO/U3KCP.
81. J. Minxha, A. N. Mamelak, U. Rutishauser, "Surgical and electrophysiological techniques for single-neuron recordings in human epilepsy patients" in *Extracellular Recording Approaches*, R. Sillitoe, Ed. (Springer, 2018), pp. 267–293.
82. U. Rutishauser, E. M. Schuman, A. N. Mamelak, Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci. Methods* **154**, 204–224 (2006). doi: 10.1016/j.jneumeth.2005.12.033; pmid: 16488479
83. U. Rutishauser, M. Cerf, G. Kreiman, "Data analysis techniques for human microwire recordings: Spike detection and sorting, decoding, relation between neurons and local field potentials" in *Single Neuron Studies of the Human Brain*, I. Fried, U. Rutishauser, M. Cerf, G. Kreiman, Eds. (MIT Press, 2014), pp. 59–98.
84. M. Reuter, H. D. Rosas, B. Fischl, Highly accurate inverse consistent registration: A robust approach. *Neuroimage* **53**, 1181–1196 (2010). doi: 10.1016/j.neuroimage.2010.07.020; pmid: 20637289
85. J. M. Tyszka, W. M. Pauli, In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template. *Hum. Brain Mapp.* **37**, 3979–3998 (2016). doi: 10.1002/hbm.23289; pmid: 27354150
86. B. Avants *et al*., Multivariate analysis of structural and diffusion imaging in traumatic brain injury. *Acad. Radiol.* **15**, 1360–1375 (2008). doi: 10.1016/j.acra.2008.07.007; pmid: 18995188
87. S. Wang, N. Chandravadia, A. N. Mamelak, U. Rutishauser, Simultaneous eye tracking and single-neuron recordings in human epilepsy patients. *J. Vis. Exp.* **2019**, e59117 (2019). doi: 10.3791/59117; pmid: 31259902
88. E. M. Meyers, D. J. Freedman, G. Kreiman, E. K. Miller, T. Poggio, Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* **100**, 1407–1419 (2008). doi: 10.1152/jn.90248.2008; pmid: 18562555

89. B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*, vol. 57 of *Monographs on Statistics and Applied Probability* (Chapman & Hall, 1993).
90. B. R. Cowley *et al.*, DataHigh: Graphical user interface for visualizing and interacting with high-dimensional neural activity. *J. Neural Eng.* **10**, 066012 (2013). doi: 10.1088/1741-2560/10/6/066012; pmid: 24216250

ACKNOWLEDGMENTS

We thank the members of the Adolphs and Rutishauser labs for discussion, Columbia Theory Center members F. Stefanini and M. Rigotti for sharing their population decoding analysis expertise, and the staff and J. M. Chung and C. M. Reed of the Cedars-Sinai Epilepsy Monitoring unit for their support. We thank all subjects and their families for their participation. **Funding:** This work was supported by NIMH (R01MH110831 to U.R.), the BRAIN initiative through the NIH Office of the Director (U01NS103792 to U.R.), the Caltech NIMH Conte Center (P50MH094258 to R.A. and U.R.), the National Science Foundation (CAREER Award BCS-1554105 to U.R. and NeuroNex Program award DBI-1707398 to S.F.), a Memory and Cognitive Disorders Award from the McKnight Foundation for Neuroscience (to U.R.), and the Simons Foundation Collaboration on the Global Brain (542941 to R.A. and PG007079 to S.F.). **Author contributions:** J.M., U.R., and R.A. designed the study. J.M. performed the experiments. J.M., S.F., and U.R. analyzed the data. J.M., U.R., R.A., and S.F. wrote the paper. A.N.M. performed surgery and supervised clinical work. **Competing interests:** None. **Data and materials availability:** Data needed to reproduce results (firing rates of all neurons versus time, precomputed phase-coherence measures for all neurons) have been deposited at OSF (*80*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/eaba3313/suppl/DC1
Supplementary Text
Figs. S1 to S11
Table S1
Movie S1

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

25 November 2019; accepted 4 May 2020
10.1126/science.aba3313

RESEARCH ARTICLE

CHROMATIN SEQUENCING

Single-molecule regulatory architectures captured by chromatin fiber sequencing

Andrew B. Stergachis[1]*, Brian M. Debo[1,2], Eric Haugen[3], L. Stirling Churchman[2], John A. Stamatoyannopoulos[3,4]*

Gene regulation is chiefly determined at the level of individual linear chromatin molecules, yet our current understanding of cis-regulatory architectures derives from fragmented sampling of large numbers of disparate molecules. We developed an approach for precisely stenciling the structure of individual chromatin fibers onto their composite DNA templates using nonspecific DNA N^6-adenine methyltransferases. Single-molecule long-read sequencing of chromatin stencils enabled nucleotide-resolution readout of the primary architecture of multikilobase chromatin fibers (Fiber-seq). Fiber-seq exposed widespread plasticity in the linear organization of individual chromatin fibers and illuminated principles guiding regulatory DNA actuation, the coordinated actuation of neighboring regulatory elements, single-molecule nucleosome positioning, and single-molecule transcription factor occupancy. Our approach and results open new vistas on the primary architecture of gene regulation.

The primary architecture of chromatin comprises nucleosome arrays punctuated by short regulatory regions containing transcription factors (TFs) and other non-histone proteins. This architecture is foundational for genome function yet remains undefined at the level of individual chromatin fibers, the fundamental units of gene regulation. For example, it is largely unknown how regulatory DNA is actuated on individual chromatin fibers, or the degree to which nearby regulatory regions are coordinately actuated on the same chromatin fiber. It is also unknown how regulatory DNA actuation affects nucleosome positioning on individual chromatin fibers, and how TF occupancy modulates regulatory DNA actuation and function on single templates. Addressing these questions requires nucleotide-resolution analysis of individual multikilobase chromatin fibers, which is not obtainable with current single-cell or bulk profiling approaches.

We sought to develop a method for recording the primary architecture of chromatin onto its underlying DNA template at single-nucleotide resolution, thereby enabling the simultaneous identification of genetic and epigenetic features along multikilobase segments of the genome. Current approaches to mapping chromatin and regulatory architectures sample large populations of chromatin fibers and rely on dissolution of chromatin using nucleases such as deoxyribonuclease (DNase I) (*1*, *2*), micrococcal nuclease (*3*, *4*), restriction enzymes (*5*), transposases (*6*), or mechanical shearing (*7*). CpG and GpC methyltransferases are capable of marking accessible cytosines in a dinucleotide context without digesting DNA (*8*–*10*), and approaches using GpC methyltransferases have been extended to single-pass nanopore sequencing (*11*, *12*). However, the utility of these approaches for gaining insights into the biology of individual chromatin fibers is limited because of (i) the sporadic occurrence and linear clustering of CpG and GpC dinucleotides in animal genomes as a result of mutation and selection (fig. S1); (ii) the confounding influence of endogenous cytosine methylation machineries (*13*); (iii) the marked DNA degradation induced by bisulfite conversion (*14*); and (iv) the intrinsically limited ability of nanopore sequencing to accurately identify modified bases on a per-molecule basis (*11*, *12*, *15*).

Unlike cytosine, adenine bases in DNA are almost completely devoid of endogenous methylation in eukaryotes (*16*) and occur at an average frequency approaching one in every two DNA base pairs in animal genomes without the clustering and extended deserts characteristic of guanine-cytosine dinucleotides (fig. S1). The sequence-specific adenine methyltransferase *Dam* demonstrates some preference for nonchromatinized DNA (*17*–*19*), suggesting that nonspecific members of this class of enzymes could have similar properties. We therefore sought to identify a nonspecific (that is, non–sequence context dependent) N^6-adenine DNA methyltransferase (m6A-MTase) with high efficiency, high stability, and a molecular weight similar to that of nonspecific nucleases such as DNase I (~30 kD) that are able to access protein-DNA interfaces at nucleotide resolution (Fig. 1A). We isolated five distinct nonspecific DNA m6A-MTases (*20*–*23*) and demonstrated that treatment of nonchromatinized (fig. S2) and chromatinized DNA templates (Fig. 1B) with increasing amounts of each enzyme resulted in monotonic increases in adenine methylation.

m6A-MTases are quantitative probes of DNA accessibility

To establish selectivity of nonspecific m6A-MTases for accessible DNA templates in nuclear chromatin, we treated *Drosophila melanogaster* S2 cell nuclei with increasing concentrations of five adenine methyltransferases and compared the genomic distribution of m6A-DNA with the distribution of DNase I cleavages, defined using DNase I–seq [the established standard for marking accessible DNA templates (*1*, *2*)] (Fig. 1C). Immunoprecipitation and sequencing of short [median, 110 base pairs (bp)] DNA fragments harboring m6A (AdMTase-seq) showed the genomic distribution of m6A-DNA to mirror the density of DNase I cleavages quantified by DNase I–seq (Fig. 1D). m6A-MTases showed high selectivity for accessible DNA and quantitative representation of DNase I–hypersensitive sites (DHSs) (Fig. 1D and fig. S3). Further, m6A-MTases demonstrated decreasing selectivity toward DHSs with increasing amounts of enzyme (Fig. 1E and fig. S3, A and B), analogous with the enzymatic action of DNase I (or micrococcal nuclease) on chromatin substrates due to increasing marking of the far more numerous internucleosomal linker regions (*24*–*26*). Quantification of DNA accessibility by AdMTase-seq was highly reproducible at both promoter-proximal and promoter-distal regulatory elements (Fig. 1F and fig. S3C), with the enzyme Hia5 demonstrating the highest efficiency (figs. S2 and S3). These results show that nonspecific m6A-MTases provide quantitative probes of DNA accessibility in chromatin.

m6A stencils reveal single-molecule chromatin architectures

We next aimed to reconstruct the primary architecture of chromatin fibers at nucleotide resolution by sequencing the linear pattern of m6a along multikilobase chromatin stencils; we termed this process Fiber-seq (Fig. 2A). To implement Fiber-seq, we capitalized on the ability of single-molecule DNA sequencers to discriminate methylated from unmethylated adenine residues based on the DNA polymerase kinetics at that base during sequencing (*23*, *27*). To achieve nucleotide resolution, we leveraged single-molecule circular consensus sequencing (CCS) (*27*, *28*) performed on a Pacific Biosciences (PacBio) instrument, which enables resequencing of each chromatin fiber stencil at least 10 times, resulting in highly accurate base calling of both modified and unmodified nucleotides.

To create chromatin stencils, we performed m6A-MTase treatment of S2 cell nuclei, followed by polymerase chain reaction–free library construction on high–molecular-weight DNA extracted from either treated or untreated nuclei.

[1]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [2]Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA, USA. [3]Altius Institute for Biomedical Sciences, Seattle, WA, USA. [4]Departments of Genome Sciences and Medicine, University of Washington, Seattle, WA, USA.
*Corresponding author. Email: absterga@u.washington.edu (A.B.S.); jstam@altius.org (J.A.S.)

The resulting libraries were subjected to single-molecule CCS, providing very high base-calling accuracy (fig. S4, A and B) across chromatin fibers of up to ~30 kb in length (average fiber length, 10.9 kb) (Fig. 2B). Whereas untreated nuclei showed minimal m6A signal, nearly all fibers from m6A-MTase–treated cells demonstrated a high degree of m6A methylation (Fig. 2, C to E, and fig. S4A). The m6A-MTase–treated sample yielded an average coverage of 43 fibers for each DHS, with the average number of m6A-marked bases on each fiber overlapping a DHS mirroring the density of DNase I cleavages quantified by DNase I–seq from bulk nuclei (Fig. 2F). We observed marked clustering of adenine methylations into short contiguous regions spanning tens to hundreds of nucleotides, flanked by extended stretches of unmodified nucleotides (Fig. 2, D and E). Two categories of methylase-accessible DNA sequences (MADs) were evident: (i) sequence elements with an average length of 272 bp that coincided with DHSs and (ii) far more numerous shorter sequence elements with an average length of 67 bp and regularized spacing, paralleling the expected size and distribution of internucleosomal linker regions (*29*) (Fig. 2G and fig. S4D). m6A modifications were largely absent from nucleosome-wrapped DNA between linker regions (Fig. 2, D and E), consistent with the requirement of m6A-MTases for base flipping to modify adenines (*30*), which is likely suppressed by nucleosome wrapping. Notably, extended stretches of unmodified bases were absent from deproteinized chromatin fibers treated with m6A-MTases (fig. S5). Together, these results show that Fiber-seq accurately translates m6A single-molecule chromatin stencils into a linear, nucleotide-precise readout of the primary architecture of chromatin.

All-or-none actuation of regulatory DNA on individual fibers

We next sought to identify how regulatory information encoded in genomic DNA is actuated on individual chromatin fibers. Regulatory DNA actuation (the all-or-none adoption of a nucleosome-free state with resulting hyperaccessibility of the underlying DNA) is fundamental to cell state and fate decisions (*31*) as well as phenotypic traits and disease risk (*32*), yet it is currently unknown whether regulatory elements are actuated in an all-or-none fashion on any given chromatin fiber (in place of a canonical nucleosome) or whether regulatory DNA is actuated to varying degrees on all chromatin fibers as alternative structures with varying DNA accessibility. Evaluation of individual chromatin fibers overlapping each DHS demonstrated that overall, only 81% of these fibers had large MADs consistent with an actuated or open

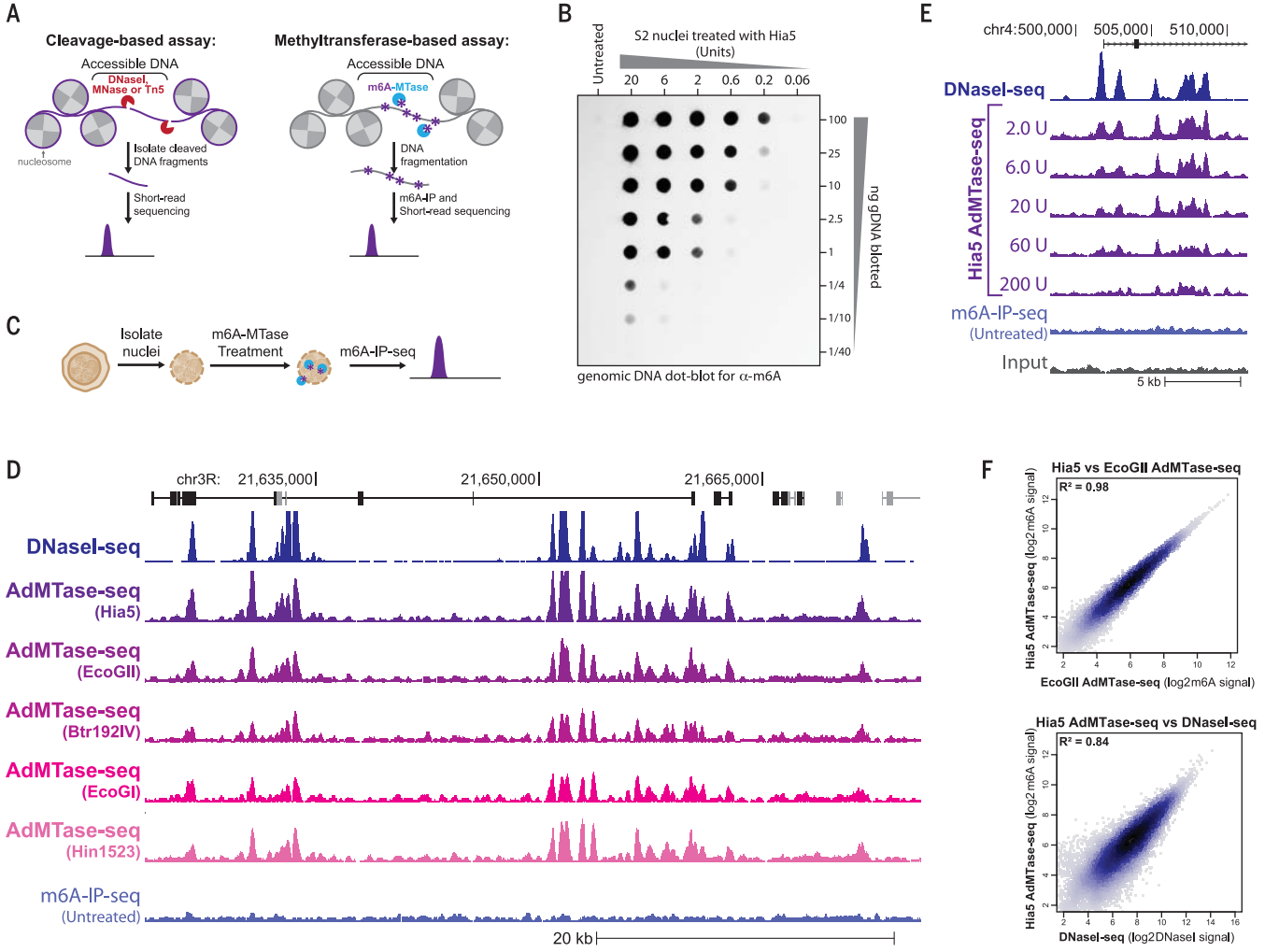


Fig. 1. Nonspecific m6A-MTases selectively mark sites of chromatin accessibility. (**A**) Schematic of cleavage- and m6A-MTase-based methods for marking sites of chromatin accessibility. Tn*5*, transposon 5. (**B**) Dot-blot quantification of m6A-modified genomic DNA (gDNA) from S2 cell nuclei after treatment with the m6A-MTase Hia5. (**C**) Experimental schematic for AdMTase-seq. (**D** and **E**) Genomic loci comparing the relationship between DNase I–seq signal and AdMTase-seq signal after treatment of S2 cell nuclei with (D) five separate m6A-MTases or (E) increasing amount of the m6A-MTase Hia5. The *y* axis is identical for all AdMTase-seq and m6A-immunoprecipitation-sequencing (m6A-IP-seq) experiments. (**F**) Comparison of AdMTase-seq signal for S2 cell DHSs from cells treated with (top) Hia5 versus EcoGII or (bottom) bulk DNase I–seq signal. The *y* axis is identical for all AdMTase-seq and m6A-IP-seq experiments.

chromatin state overlapping the DHS, with the remaining fibers showing nucleosome demarcation indicative of a closed state at the DHS (Fig. 3, A and B, and fig. S6, A to C). Transcription start site (TSS)–distal DHSs were preferentially maintained in a closed state when compared with promoter DHSs (Fig. 3B), and the rate of DNA actuation at individual TSS-distal DHSs mirrored the density of DNase I cleavages quantified by DNase I–seq from bulk nuclei (Fig. 3C). By contrast, wider TSS-distal DHSs, as well as promoter DHSs, were preferentially maintained in an accessible state (Fig. 3B and fig. S6D), with the actuated DNA content of these DHSs demonstrating widespread heterogeneity across individual chromatin fibers (Fig. 2, E and G) because of the variable punctuation of co-occupying nucleosomes (fig. S6E). Together, these findings demonstrate that the regulatory DNA content at TSS-distal regulatory elements is predominately modulated by all-or-none DNA actuation, with the regulatory DNA content at larger promoter elements additionally modulated by the variable punctuation of co-occupying nucleosomes.

Co-actuation of neighboring elements on single chromatin fibers

Linear clustering of gene regulatory elements along complex genomes is well described, canonically in the context of locus control regions (LCRs) or "super-enhancers" (e.g., the beta-globin LCR) or gene-specific control regions (e.g., the *BCL11A* enhancer region) (*33*–*36*). It is currently unknown whether the individual elements in such clusters are actuated coordinately or independently or, more generally, whether actuation of one genomic

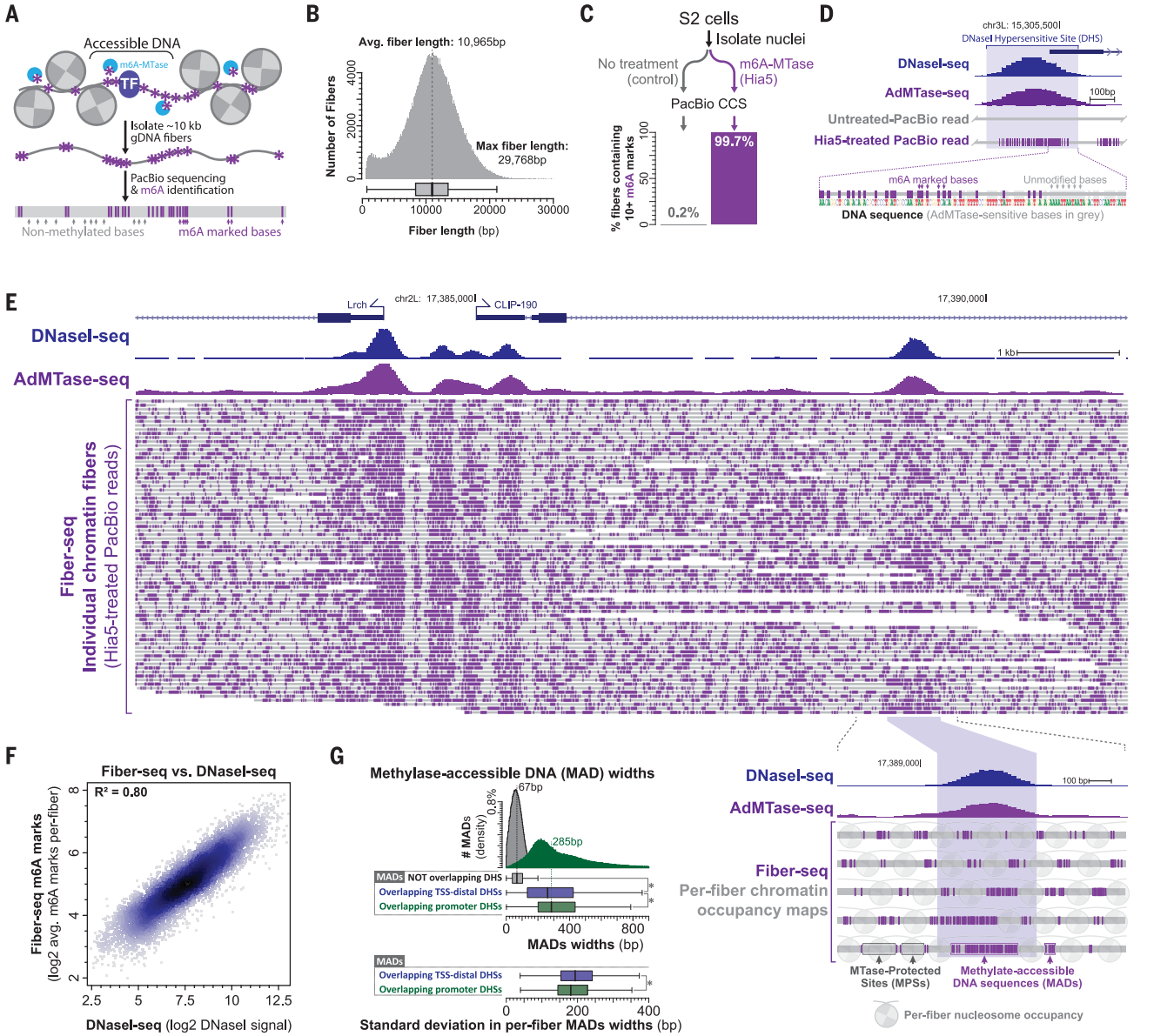


Fig. 2. Fiber-seq exposes base pair–resolution maps of individual chromatin fiber architecture. (**A**) Experimental schematic for Fiber-seq. (**B**) Histogram of read lengths for chromatin fibers sequenced using Fiber-seq. Avg., average. (**C**) Percentage of chromatin fibers with m6A-methylated bases from DNA isolated from untreated and Hia5-treated S2 cell nuclei. (**D** and **E**) Genomic loci comparing the relationship between DNase I–seq, AdMTase-seq, and Fiber-seq. Individual PacBio reads (chromatin fibers) are marked with gray lines, and m6A-modified bases are marked in purple dashes. White regions separate individual reads. Insets show (D) DHS with individual m6A-modified bases and (E) a comparison of multiple fibers overlapping a single DHS. (**F**) Scatterplot comparing DNase I–seq signal and the average Fiber-seq m6A signal at each DHS. (**G**) (Top) Histogram of MAD widths for all MADs identified outside of DHSs (gray) or in promoter DHSs (green). Box-and-whisker plots for the aforementioned in addition to MAD widths for all MADs identified in TSS-distal DHSs (blue). (Bottom) Box-and-whisker plots showing the standard deviation in MAD widths across multiple fibers overlapping a DHS. *P value < 0.001 (Wilcoxon test).

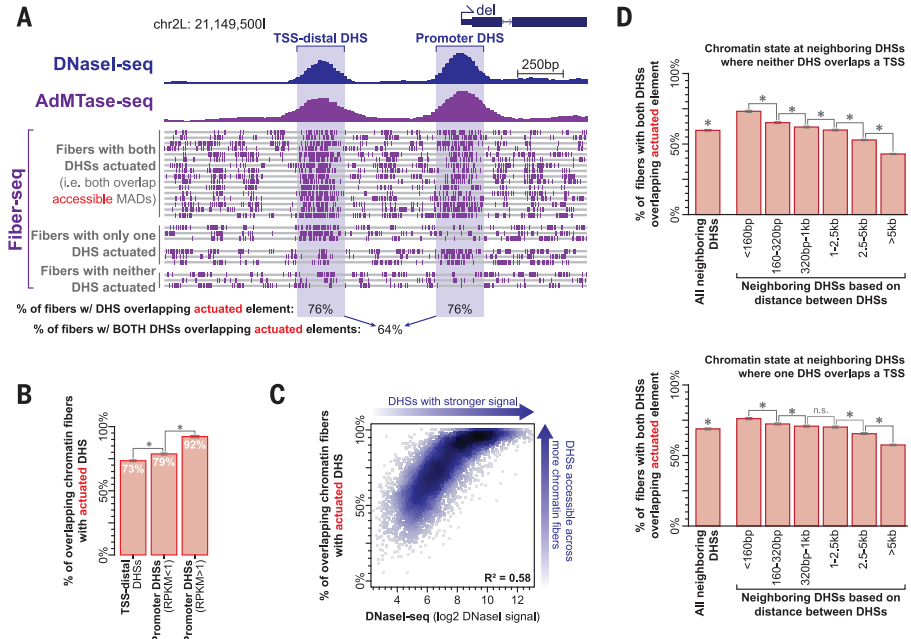


Fig. 3. Preferential co-actuation of neighboring regulatory elements on the same chromatin fiber. (**A**) Representative genomic locus comparing the relationship between DNase I–seq, AdMTase-seq, and Fiber-seq at neighboring DHSs exposes the predominant all-or-none actuation of individual DHSs. (**B**) Proportion of DHSs that overlap actuated versus closed fibers for DHSs divided based on their proximity to TSSs as well as the gene expression level of the TSS. Error bars represent 95% confidence intervals. *P value < 0.001 (z-test). RPKM, reads per kilobase per million reads. (**C**) Scatterplot showing the relationship between DHS actuation and the bulk DNase I–seq signal for TSS-distal DHSs. (**D**) For chromatin fibers encompassing two DHSs, shown is the percentage of fibers containing accessible MADs at both DHSs as a function of the distance between the DHSs. Error bars represent 95% confidence intervals. *P < 0.001 (z-test).

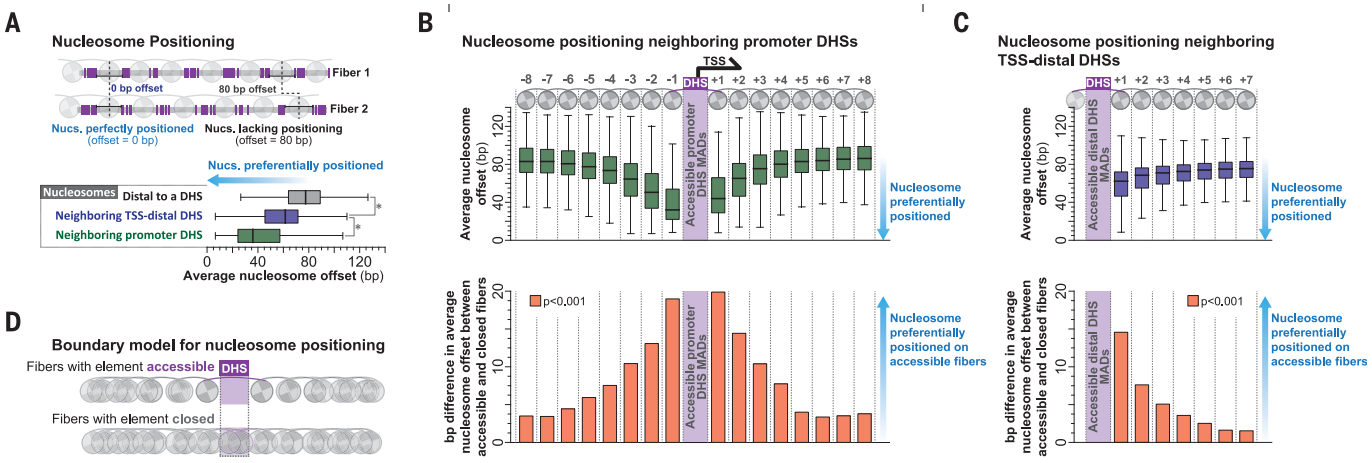


Fig. 4. Impact of regulatory DNA actuation on nucleosome positioning. (**A**) Schematic showing the calculation of nucleosome (Nucs.) positioning using overlapping reads as well as a box-and-whisker plot of nucleosome offsets at different genomic loci. *P value < 0.001 (Wilcoxon test). (**B** and **C**) (Top) Box-and-whisker plots showing nucleosome offsets neighboring (B) promoter DHSs and (C) TSS-distal DHSs. (Bottom) Bar plots demonstrating the difference in nucleosome positioning neighboring (B) promoter DHSs and (C) TSS-distal DHSs between actuated and closed fibers. P < 0.001 (Wilcoxon test). (**D**) Schematic of the boundary model for nucleosome placement surrounding regulatory elements.

element can influence actuation of neighboring elements. As 29% of Fiber-seq reads overlap multiple neighboring DHSs (Figs. 2E and 3A), we next asked whether regulatory DNA actuation at one genomic DNA element could influence the actuation of neighboring elements on the same chromatin fiber. Overall, we found that neighboring DHSs were significantly more likely to be co-actuated on the same chromatin fiber (Fig. 3D). Co-actuation was most strongly enriched for the most tightly clustered elements (Fig. 3D), indicating that the actuation of regulatory DNA accessibility at one distal element appears to potentiate accessibility at neighboring elements in a distance-dependent manner, independent of whether both elements were bound by the same TF (fig. S7). These results provide a possible physicochemical basis for the observed clustering of distal regulatory elements in animal genomes. Moreover, they suggest that genetic variants affecting regulatory DNA accessibility and function may create local knock-on effects in cis, a feature not accounted for in current models of the architecture and evolution of gene regulation.

Nucleosome arrays are bounded by actuated regulatory DNA

Nucleosome positioning is fundamental to gene regulation and is specified by a combination of factors, including DNA sequence, the competitive occupancy of TFs, the action of nucleosome remodelers, and interactions with RNA polymerases (*37*). Although nucleosomes surrounding accessible promoters and TSS-distal DHSs are generally well positioned (*38*–*40*), it remains unclear whether this positioning results from underlying nucleosome-favoring DNA sequences or from a boundary condition imposed by accessibility and TF occupancy at regulatory DNA. We reasoned that the boundary model of nucleosome positioning could be tested directly by comparing nucleosome positions surrounding a regulatory element on overlapping fibers in which the regulatory element is in an actuated or open state versus overlapping fibers in which the regulatory element is in a closed state (as such fibers have the same DNA content and should differ only in the presence or absence of an actuated element). Although nucleosomes

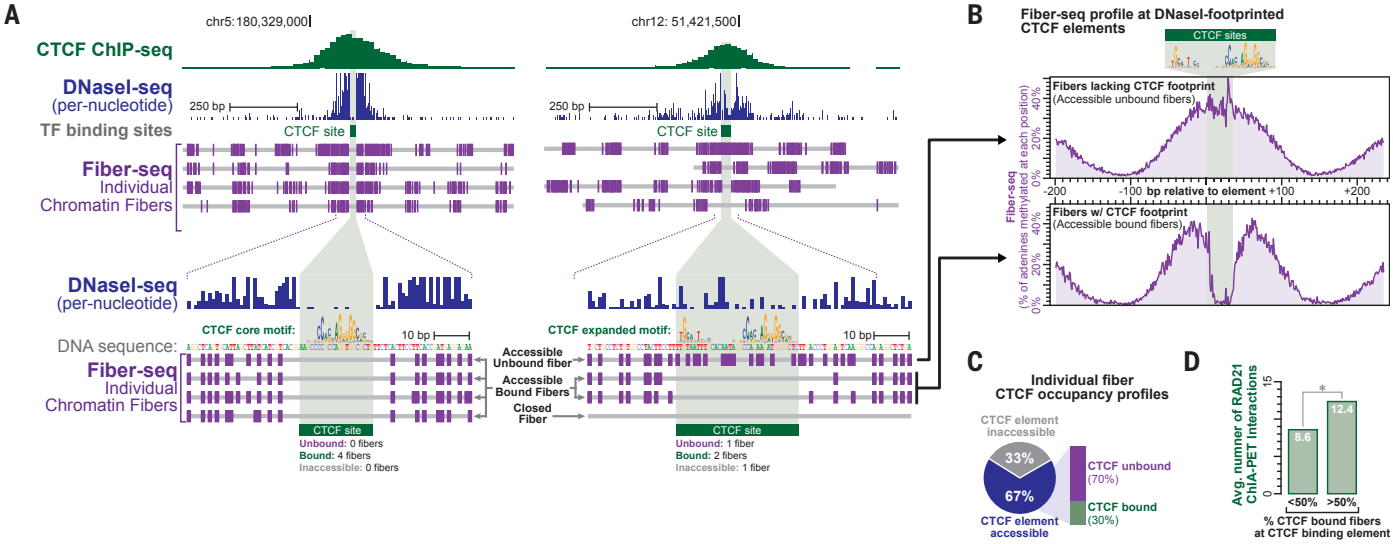


Fig. 5. Heterogeneous single-molecule CTCF occupancy impacts long-range CTCF interactions. (**A**) Genomic loci comparing the relationship between CTCF chromatin immunoprecipitation sequencing (ChIP-seq), DNase I–seq, and Fiber-seq at CTCF binding sites in K562 cells. The inset shows a CTCF binding site and per-fiber m6A methylation. (**B**) Plot of adenine methylation frequency surrounding DNase I–footprinted CTCF binding sites for individual fibers that (top) lack a CTCF footprint or (bottom) have a CTCF footprint. (**C**) Percentage of fibers overlapping CTCF ChIP-seq peaks containing a bulk CTCF DNase I footprint that lack an accessible MADs or have an accessible MADs with or without a CTCF footprint. (**D**) Average number of RAD21 ChIA-PET long-range chromatin interactions at CTCF binding sites based on the percentage of overlapping fibers that contain a single-molecule CTCF footprint. *P value < 0.01 (z-test).

surrounding DHSs were collectively well positioned (Fig. 4A), analysis of single-fiber data showed that well-positioned nucleosomes largely originated from fibers in which the regulatory element is in an actuated state (Fig. 4, B and C), indicating that nucleosome positioning at these locations is largely dependent on the actuation of regulatory DNA, not the DNA sequence itself. As such, nucleosome positioning appears to largely result from a boundary condition imposed by regulatory DNA actuation on individual chromatin fibers (Fig. 4D).

Fiber-seq of human chromatin

To test Fiber-seq on the human genome, we first validated that m6A-MTases can selectively mark cell type–specific accessible DNA in human cell types (fig. S8A). We then performed Fiber-seq on human K562 nuclei, resulting in robust demarcation of both accessible regulatory elements and internucleosomal linker regions (fig. S8, B to E) with an average coverage of 3.7 high-quality reads per DHS. Consistent with the findings from *Drosophila* cells, we found that regulatory DNA accessibility in K562 cells is predominantly actuated as an all-or-none process (fig. S8F), with larger promoter elements additionally demonstrating variably placed co-occupying nucleosomes (fig. S8G). Furthermore, neighboring TSS-distal DHSs were significantly more likely to be co-actuated on the same chromatin fiber (fig. S8H), and nucleosome positioning surrounding regulatory DNA was largely dependent on regulatory DNA accessibility as a boundary condition (fig. S8, I and J), indicating that these are likely universal effects of regulatory DNA actuation.

Single-molecule TF occupancy at nucleotide resolution

Analysis of K562 Fiber-seq patterns in accessible regulatory elements revealed that m6A-MTase methylation was not uniform but rather was punctuated by short gaps, consistent with nucleotide-precise binding of TFs that mirror bulk TF footprints on non–cross-linked chromatin (Fig. 5, A and B, and fig. S9) (*41*). We next explored these single-molecule TF occupancy events to determine the relationship between TF occupancy and regulatory DNA actuation and function, focusing on CTCF because most of its binding sites lack additional co-bound TFs, and CTCF occupancy is considered essential for establishing long-range chromatin interactions (*42*). TFs are known to undergo rapid exchange at regulatory DNA elements (*43*), yet it remains unclear the extent to which regulatory elements can remain actuated during the transient absence of TF occupancy (*44*–*46*). We observed that most of the actuated chromatin fibers overlapping CTCF-bound regulatory elements lacked a CTCF footprint (Fig. 5, B and C), suggesting that regulatory DNA actuation can be maintained in the absence of immediate occupancy by a gating TF. Notably, CTCF elements that preferentially contained fibers with a CTCF footprint were more likely to participate in long-range chromatin interactions (Fig. 5D), providing a mechanistic link between single-molecule CTCF occupancy and function.

Outlook

In conclusion, we have shown that Fiber-seq provides a robust and scalable approach for stenciling the primary architecture of individual chromatin fibers onto their underlying DNA templates with high precision and resolution. As both sequencing read lengths and throughput increase, it should be possible in the near future to transcribe and resolve the primary regulatory architectures of entire genetic haplotypes by simultaneously mapping both the primary genetic sequence and overlying chromatin state. As such, Fiber-seq has the potential to provide a unifying tool for analyzing the gene regulatory impact of both rare and common regulatory DNA variation, and for resolving extended regulatory alleles.

REFERENCES AND NOTES

1. D. S. Gross, W. T. Garrard, *Annu. Rev. Biochem.* **57**, 159–197 (1988).
2. R. E. Thurman *et al.*, *Nature* **489**, 75–82 (2012).
3. M. Noll, R. D. Kornberg, *J. Mol. Biol.* **109**, 393–404 (1977).
4. D. E. Schones *et al.*, *Cell* **132**, 887–898 (2008).
5. E. Lieberman-Aiden *et al.*, *Science* **326**, 289–293 (2009).
6. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, *Nat. Methods* **10**, 1213–1218 (2013).
7. T. Kouzarides, *Cell* **128**, 693–705 (2007).
8. T. K. Kelly *et al.*, *Genome Res.* **22**, 2497–2506 (2012).
9. A. R. Krebs *et al.*, *Mol. Cell* **67**, 411–422.e4 (2017).
10. N. H. Nabilsi *et al.*, *Genome Res.* **24**, 329–339 (2014).
11. Y. Wang *et al.*, *Genome Res.* **29**, 1329–1342 (2019).
12. Z. Shipony *et al.*, *Nat. Methods* **17**, 319–327 (2020).
13. A. P. Bird, *Nature* **321**, 209–213 (1986).
14. K. Tanaka, A. Okamoto, *Bioorg. Med. Chem. Lett.* **17**, 1912–1915 (2007).
15. A. B. R. McIntyre *et al.*, *Nat. Commun.* **10**, 579 (2019).
16. G. Z. Luo, C. He, *Nat. Struct. Mol. Biol.* **24**, 503–506 (2017).
17. J. Singh, A. J. S. Klar, *Genes Dev.* **6**, 186–196 (1992).
18. D. E. Gottschling, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4062–4065 (1992).
19. M. P. Kladde, R. T. Simpson, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1361–1365 (1994).

20. M. Drozdz, A. Piekarowicz, J. M. Bujnicki, M. Radlinska, *Nucleic Acids Res.* **40**, 2119–2130 (2012).
21. B. P. Anton, G. P. Harhay, T. P. L. Smith, J. Blom, R. J. Roberts, *PLOS ONE* **11**, e0161499 (2016).
22. I. A. Murray *et al*., *Nucleic Acids Res.* **46**, 840–848 (2018).
23. G. Fang *et al*., *Nat. Biotechnol.* **30**, 1232–1239 (2012).
24. H. Weintraub, M. Groudine, *Science* **193**, 848–856 (1976).
25. K. S. Bloom, J. N. Anderson, *Cell* **15**, 141–150 (1978).
26. J. Mieczkowski *et al*., *Nat. Commun.* **7**, 11485 (2016).
27. J. Eid *et al*., *Science* **323**, 133–138 (2009).
28. K. J. Travers, C. S. Chin, D. R. Rank, J. S. Eid, S. W. Turner, *Nucleic Acids Res.* **38**, e159 (2010).
29. R. V. Chereji *et al*., *Nucleic Acids Res.* **44**, 1036–1051 (2016).
30. J. R. Horton, K. Liebert, M. Bekes, A. Jeltsch, X. Cheng, *J. Mol. Biol.* **358**, 559–570 (2006).
31. A. B. Stergachis *et al*., *Cell* **154**, 888–903 (2013).
32. M. T. Maurano *et al*., *Science* **337**, 1190–1195 (2012).
33. W. A. Whyte *et al*., *Cell* **153**, 307–319 (2013).
34. P. Diaz, D. Cado, A. Winoto, *Immunity* **1**, 207–217 (1994).
35. L. Madisen, M. Groudine, *Genes Dev.* **8**, 2212–2226 (1994).
36. F. Grosveld, G. B. van Assendelft, D. R. Greaves, G. Kollias, *Cell* **51**, 975–985 (1987).
37. K. Struhl, E. Segal, *Nat. Struct. Mol. Biol.* **20**, 267–273 (2013).
38. S. Baldi *et al*., *Mol. Cell* **72**, 661–672.e4 (2018).
39. G.-C. Yuan *et al*., *Science* **309**, 626–630 (2005).
40. C. Jiang, B. F. Pugh, *Nat. Rev. Genet.* **10**, 161–172 (2009).
41. S. Neph *et al*., *Nature* **489**, 83–90 (2012).
42. J. E. Phillips, V. G. Corces, *Cell* **137**, 1194–1211 (2009).
43. T. C. Voss *et al*., *Cell* **146**, 544–554 (2011).
44. C. C. Adams, J. L. Workman, *Mol. Cell. Biol.* **15**, 1405–1421 (1995).
45. J. A. Miller, J. Widom, *Mol. Cell. Biol.* **23**, 1623–1632 (2003).
46. L. A. Mirny, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22534–22539 (2010).
47. A. B. Stergachis, Code for mapping single-molecule m6A methylations, Zenodo (2020); doi: 10.5281/zenodo.3743228

ACKNOWLEDGMENTS

We thank M. Radlinska for kindly providing the Hia5 and Hin1523 plasmids (*20*). We also thank L. Tallon and L. Sadzewicz for their assistance in PacBio sequencing as well as R. S. Isaac for his help with FPLC purification of the Hia5 enzyme. **Funding:** This work was supported by NIH grants UM1HG009444 to J.A.S. and T32GM007748 supporting A.B.S. **Author contributions:** A.B.S., B.M.D., and J.A.S. designed the experiments. A.B.S. and B.M.D. performed the experiments. A.B.S. and E.H. performed the computational analyses. A.B.S., B.M.D., and J.A.S. wrote the manuscript. L.S.C. provided ongoing support and many helpful critiques. **Competing interests:** A.B.S. and J.A.S. are coinventors on U.S. patent application 63/004,361 that includes discoveries described in this manuscript. **Data and materials availability:** All data are available in the manuscript or the supplementary materials or at GEO accession GSE146942. Code for generating single-fiber chromatin architectures is available at Zenodo.org (*47*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1449/suppl/DC1
Materials and Methods
Figs. S1 to S9
References (*48*–*54*)
MDAR Reproducibility Checklist

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

17 August 2019; resubmitted 12 January 2020
Accepted 24 April 2020
10.1126/science.aaz1646

STRUCTURAL BIOLOGY

Cryo-EM structure of the human cohesin-NIPBL-DNA complex

Zhubing Shi[1], Haishan Gao[1], Xiao-chen Bai[2,3]*, Hongtao Yu[1,4]*

As a ring-shaped adenosine triphosphatase (ATPase) machine, cohesin organizes the eukaryotic genome by extruding DNA loops and mediates sister chromatid cohesion by topologically entrapping DNA. How cohesin executes these fundamental DNA transactions is not understood. Using cryo–electron microscopy (cryo-EM), we determined the structure of human cohesin bound to its loader NIPBL and DNA at medium resolution. Cohesin and NIPBL interact extensively and together form a central tunnel to entrap a 72–base pair DNA. NIPBL and DNA promote the engagement of cohesin's ATPase head domains and ATP binding. The hinge domains of cohesin adopt an "open washer" conformation and dock onto the STAG1 subunit. Our structure explains the synergistic activation of cohesin by NIPBL and DNA and provides insight into DNA entrapment by cohesin.

Cohesin is a multisubunit adenosine triphosphatase (ATPase) machine with key roles in genome organization and chromosome segregation (*1*–*4*). Cohesin mediates the formation of chromatin loops and topologically associating domains by loop extrusion (*5*–*12*). During DNA replication, cohesin topologically entraps both sister chromatids to establish sister chromatid cohesion (*13*, *14*), which is critical for accurate chromosome segregation.

Human cohesin consists of four subunits: SMC1, SMC3, RAD21, and either STAG1 or STAG2 (Fig. 1A). The homologous ATPases SMC1 and SMC3 heterodimerize through their hinge domains. The kleisin subunit RAD21 (Scc1/Mcd1 in yeast) links their ATPase head domains (HDs) to form a ring (*15*–*17*). The huntingtin, elongation factor 3, A subunit, and TOR (HEAT) repeat proteins STAG1 and STAG2 (Scc3 in yeast) bind RAD21 and provide docking sites for other regulators (*18*, *19*). The NIPBL-MAU2 loader complex (Scc2-Scc4 in yeast) loads cohesin onto chromosomes and associates with chromosome-bound cohesin to promote loop extrusion (*6*, *9*, *20*). The HEAT repeat protein NIPBL (Scc2) alone can stimulate the ATPase activity of cohesin in the presence of DNA and is sufficient for cohesin loading onto DNA and loop extrusion (*12*, *21*, *22*). Chromosome-bound cohesin is dynamic and can be released by the releasing factor WAPL with the help of the scaffold protein PDS5 (*23*, *24*). Acetylation of SMC3 by the acetyltransferases ESCO1 and ESCO2 antagonizes WAPL-PDS5–dependent cohesin release and stabilizes cohesin on chromosomes (*25*–*29*).

The crystal structures of each cohesin subunit alone or in subcomplexes have been determined (*15*, *17*, *18*, *30*–*38*). Low-resolution negative-stain electron microscopy (EM) maps of an engineered human cohesin complex have also been reported (*39*). However, the lack of high-resolution structures of large cohesin complexes hinders our understanding of how cohesin interacts with DNA and how NIPBL promotes the loading of cohesin onto DNA.

Here, we report the cryo–electron microscopy (cryo-EM) structure of human cohesin bound to the C-terminal HEAT repeat domain of NIPBL (NIPBL^{C}; residues 1163 to 2804) and DNA. This structure, together with crystal structures of the human SMC1-SMC3 hinge domains, reveals how cohesin-NIPBL embraces DNA and establishes the basis for NIPBL-DNA–dependent activation of cohesin's ATPase activity.

Overall structure of human cohesin-NIPBL-DNA complex

NIPBL^{C} is sufficient to stimulate the ATPase activity of cohesin in the presence of DNA and to promote loop extrusion by cohesin (*12*). We thus reconstituted the cohesin-NIPBL^{C}-DNA complex for structural studies. To capture cohesin in the ATP-bound state, we used the ATPase-deficient SMC1 Glu^{1157}→Gln (E1157Q)/SMC3 E1144Q (cohesin^{EQ}) mutant for complex assembly (fig. S1). We incubated cohesin^{EQ} and NIPBL^{C} with the nonhydrolyzable ATP analog adenylyl-imidodiphosphate (AMP-PNP) and a 72–base pair double-stranded DNA (dsDNA), subjected the complex to mild cross-linking to stabilize the complex, and performed cryo-EM analysis (figs. S2 and S3).

The three-dimensional (3D) reconstruction of the cohesin^{EQ}-NIPBL^{C}-DNA complex was first performed at an overall resolution of 4.0 Å (fig. S3). To further improve the resolution, we performed a focused refinement of the more stable core of the complex, including NIPBL^{C}, RAD21, and the ATPase HDs of SMC1-SMC3. The resulting cryo-EM map at 3.9 Å showed improved density for this core, allowing us to build a nearly complete model, with the help of crystal structures of individual subunits (*15*, *17*, *33*, *37*) (fig. S4, table S1, and movie S1).

The cryo-EM density for STAG1 was less well resolved, presumably because of low binding occupancy and structural flexibility (fig. S3). Subsequent application of further 3D classification with local angular search into six classes identified a single class with good density for STAG1 (fig. S3). Further 3D refinement of the corresponding particles from this class produced a map with an overall resolution of 5.3 Å, which

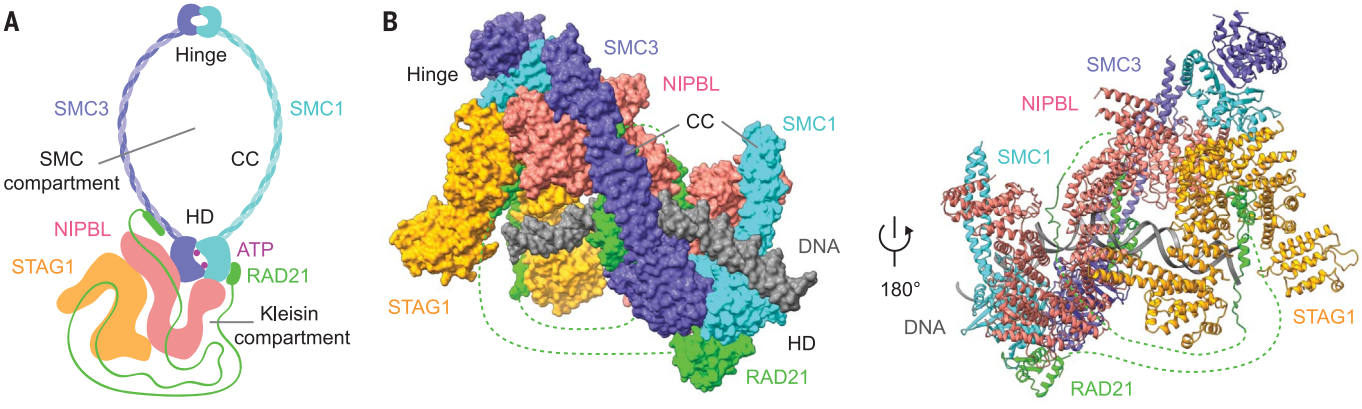


Fig. 1. Overall structure of the human cohesin-NIPBL^{C}-DNA complex. (**A**) Schematic representation of the human cohesin-NIPBL complex. The SMC and kleisin compartments are indicated. (**B**) Surface (left) and cartoon (right) representations of the cryo-EM structure of the cohesin-NIPBL^{C}-DNA complex in two different views. SMC1, SMC3, RAD21, STAG1, NIPBL, and DNA are colored cyan, purple, green, orange, pink, and gray, respectively.

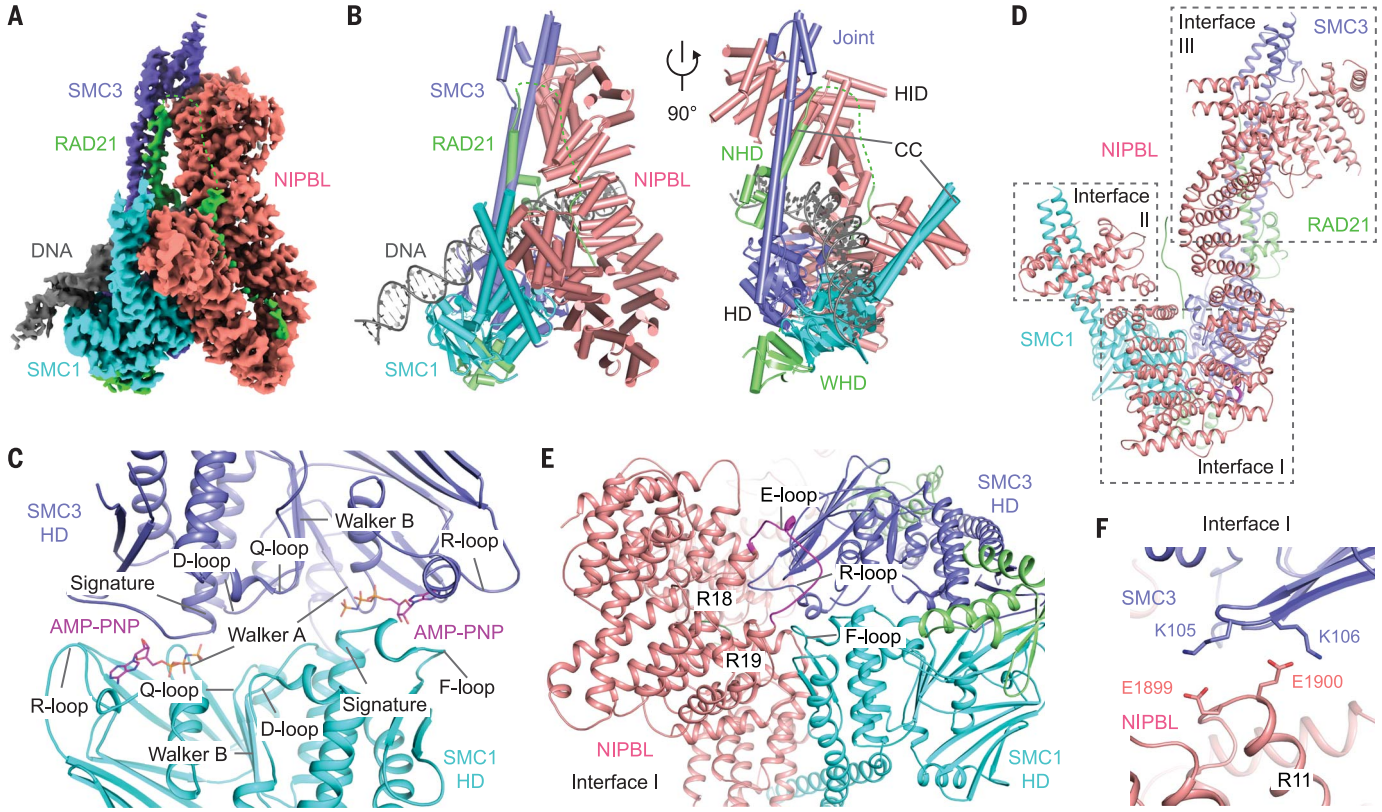


Fig. 2. Interactions between NIPBL^C and the SMC1-SMC3 heterodimer. (**A** and **B**) Cryo-EM map at a resolution of 3.9 Å (A) and cartoon illustration (B) of the cohesin-NIPBL^C-DNA complex. (**C**) ATPase active sites of the SMC1-SMC3 heterodimer. AMP-PNP is shown as sticks. Conserved motifs in SMC1 and SMC3 HDs are indicated. (**D**) Interfaces between NIPBL^C and the SMC1-SMC3 heterodimer. (**E**) Close-up view of interface I. The NIPBL E-loop is colored purple. (**F**) Close-up view of the electrostatic interactions involving the acetylation acceptor lysines (K105 and K106) of SMC3 at interface I.

showed improved density for STAG1 (fig. S5, table S1, and movie S2). The structural model of STAG1, generated by homology modeling using the crystal structure of STAG2 as template (*18*), can be unambiguously docked into the cryo-EM density on the basis of clear secondary structural features (Fig. 1B).

In this cryo-EM map, we identified additional density docked on the outer surface of STAG1 (fig. S5). On the basis of the overall shape and secondary structure features, this density could be attributed to the SMC1-SMC3 hinge domain heterodimer (*30*) (movie S3). Most of the SMC1 and SMC3 coiled coils (CCs), large segments of the flexible region of RAD21, and the N- and C-terminal unstructured parts of STAG1 and NIPBL^C were completely unresolved. As the shape of each DNA base was not clearly defined at 3.9-Å resolution, we used A-T base pairs, instead of the original DNA sequence, for DNA modeling.

The cohesin-NIPBL^C-DNA complex consists of three layers that tightly pack against each other (Fig. 1B and movie S2). The first layer comprises the V-shaped SMC1-SMC3 heterodimer and the N-terminal helical domain (NHD) and C-terminal winged helix domain (WHD) of RAD21. NIPBL^C and a flexible segment of RAD21 make up the middle layer. The third layer contains STAG1, a central segment of RAD21, and the SMC1-SMC3 hinge domains. Being the main component of the middle layer, NIPBL^C packs against the SMC1-SMC3 heterodimer on one side and STAG1 on the other.

The highly extended protein RAD21 tethers all three layers through interactions with SMC1-SMC3, NIPBL^C, and STAG1 (Fig. 1B). RAD21 interactions with SMC1, SMC3, and STAG1 are consistent with those previously reported for homologous subcomplexes (*15*, *17*, *18*). DNA binds at a central tunnel perpendicular to the three layers and bends ~45° at the NIPBL^C-SMC3 interface (Fig. 1B). It contacts all three layers and stabilizes the entire complex.

Structure of the SMC1-SMC3 heterodimer

Both SMC1 and SMC3 ATPase HDs are composed of N- and C-lobes and contain key elements for ATP binding and hydrolysis. In the structure of the cohesin-NIPBL^C-DNA complex, SMC1 and SMC3 HDs form a heterodimer and bind two AMP-PNP molecules at the dimer interface through conserved motifs (Fig. 2, A to C, and fig. S6A). Consistent with previous results (*15*, *17*), RAD21 interacts with SMC1 and SMC3 through its WHD and NHD, respectively (fig. S6, B to D).

The major parts of SMC1 and SMC3 CCs are missing in the EM map; however, a small CC segment connected to the SMC3 HD can be modeled (fig. S6, B and C). The distal end of this CC segment folds into a short four-helix bundle because of breaks in the C-terminal helix of the CC. The positions of these breaks correspond to the "SMC joint," which was first defined in yeast Smc3 and bacterial Smc proteins (*15*, *40*) (fig. S6, C and D). The SMC joint is more likely to undergo bending motions, as only one helix in the CC is contiguous at this position.

Interactions between NIPBL^C and cohesin

NIPBL^C adopts a U-shaped fold, which contains 24 HEAT repeats (R1 to R24) and a helical insert domain (HID) (*33*, *37*) (Fig. 2B and fig. S7A). A continuous and narrow density was observed in the central cleft of the "U" (fig. S7A). We previously showed that a *Chaetomium*

[1]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [2]Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [3]Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [4]School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China.
***Corresponding author. Email: xiaochen.bai@utsouthwestern.edu (X.-c.B.); hongtao.yu@utsouthwestern.edu (H.Y.)**

thermophilum Scc1 fragment (residues 126 to 230) interacts with Scc2 in vitro (*33*). This region of Scc1 corresponds to residues 120 to 194 of human RAD21. On the basis of clearly defined side-chain densities, we could assign the density in the NIPBL cleft to residues 154 to 171 in RAD21 (fig. S7A). Extra unassigned densities extending from this RAD21 segment to the bottom of the NIPBL "U" likely also belong to RAD21 (Fig. 2A and movie S1). RAD21 binds to NIPBL through both electrostatic and hydrophobic interactions (fig. S7, B and C). The relevance of the NIPBL-RAD21 interface is supported by previous mutagenesis data on the *C. thermophilum* proteins (*33*).

NIPBL^{C} forms three interfaces with SMC1-SMC3 (denoted as interfaces I, II, and III) (Fig. 2D). Interface I is mainly formed by both SMC1 and SMC3 HDs and the C-terminal arm of the "U" in NIPBL—i.e., HEAT repeats R18 to R21. NIPBL simultaneously contacts the SMC1 HD C-lobe and SMC3 HD N-lobe and strengthens the heterodimeric engagement of the two HDs through an extensive interaction network (Fig. 2E and fig. S8, A and B). This network involves the SMC1 F-loop, which resembles the W-loop of SMC4 (*41*), the SMC3 R-loop, and the loop connecting R18 and R19 in NIPBL^{C}, which is termed the E-loop. Notably, Lys^{105} and Lys^{106} in the N-lobe of SMC3 HD, which are the acetylation sites of ESCO1 and ESCO2 (*25*–*27*), appear poised to form electrostatic interactions with Glu^{1899} and Glu^{1900} in R11 of the NIPBL^{C} N-terminal arm (Fig. 2F). Acetylation of these two lysines is expected to neutralize their positive charges, disrupt favorable electrostatic interactions with NIPBL, and inhibit cohesin loading and possibly loop extrusion. Indeed, Eco1 limits loop expansion in yeast (*42*). Mutations of the SMC3 deacetylase HDAC8 have been linked to human developmental diseases termed cohesinopathies (*43*). HDAC8-deficient cells exhibit elevated levels of SMC3 acetylation and deficient cohesin loading at certain genomic loci. The functional importance of Glu^{1899} and Glu^{1900} of NIPBL remains to be validated.

Interface II is between the SMC1 CC and the C-terminal region of NIPBL^{C} (Fig. 2D and fig. S8C). It mainly involves hydrogen bonding and hydrophobic interactions. Interface III is between the SMC3 joint and the N-terminal region of NIPBL^{C}—namely, HEAT repeats R2 to R6—and also involves the RAD21 NHD (Fig. 2D and fig. S8D). This interface comprises extensive electrostatic and hydrophobic interactions. Compared with the Scc2 HEAT repeat alone (*33*), the N-terminal arm of NIPBL^{c} undergoes a large conformational change upon binding to cohesin, which may facilitate their interaction. Overall, NIPBL associates with both SMC1 and SMC3 HDs and CCs to stabilize the HD-engaged conformation of SMC1-SMC3 and helps to create a closed SMC compartment for DNA entrapment.

Similar to NIPBL, STAG1 also adopts a U-shaped fold. In the cohesin-NIPBL^{C}-DNA complex, the N-terminal arm of the "U" in NIPBL packs against the N-terminal arm of the "U" in STAG1 in an antiparallel manner (Fig. 3A). A major interface is formed by HEAT repeats R8 to R10 of STAG1 and R1 to R5 and HID of NIPBL (fig. S9A). At a second, smaller interface, two long helices from R1 and R2 in the N-terminal region of STAG1 make contacts with the HEAT repeat R13 in NIPBL (Fig. 3A and fig. S9B). Through performing 3D classification, we observed short-range hinge motions of the N-terminal region of STAG1 relative to NIPBL (fig. S9C), suggesting that this interface is fluid.

The SMC1-SMC3 hinge domains directly contact STAG1

The SMC1-SMC3 hinge heterodimer makes direct contact with the bottom part of the "U" of STAG1 (Fig. 3A). This unexpected interaction

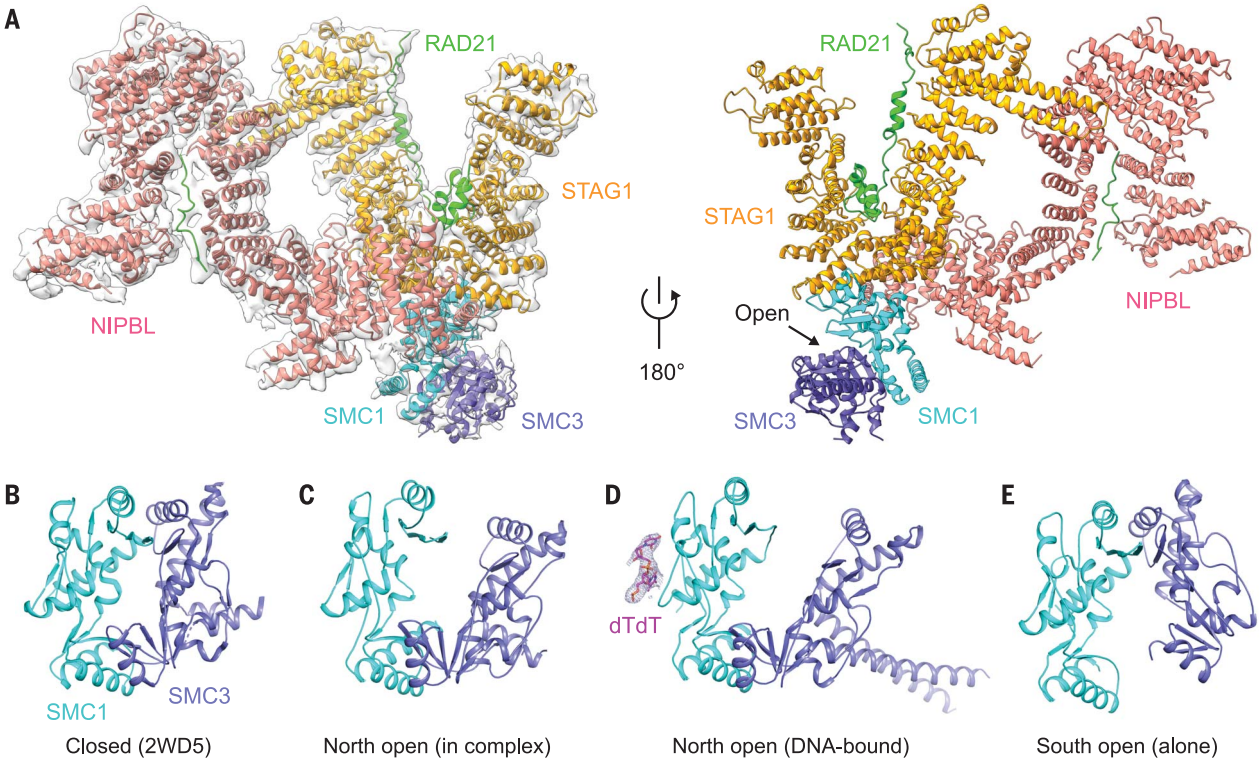


Fig. 3. Interactions among STAG1-RAD21, NIPBL^{C}, and the SMC1-SMC3 hinge heterodimer. (**A**) Cartoon representation of the subcomplex of STAG1-RAD21, NIPBL^{C}, and the SMC1-SMC3 hinge in two different views. The EM map of these components (colored gray) is overlaid on the cartoon in the left panel. (**B** to **E**) Structures of the SMC1-SMC3 hinge heterodimer in the closed conformation (PDB ID 2WD5) (B), the north-open conformation in the cryo-EM structure of the cohesin-NIPBL^{C}-DNA complex (C), the north-open conformation in the crystal structure of the hinge with short CCs bound to ssDNA (D), and the south-open conformation in the crystal structure of the hinge (E). The electron density of the two deoxythymidines (dT) is shown as gray mesh (2mFo-DFc map contoured at 1.2σ).

is mainly mediated by the SMC1 hinge and HEAT repeats R8 and R9 of STAG1 (fig. S10A). Consistent with these structural observations, a previous study showed that the hinge domain of Psm1-Psm3, the SMC1-SMC3 homolog in fission yeast, can bind to Psc3, the fission yeast homolog of STAG1 and STAG2 (*44*). The N-terminal region of NIPBL^{C} is also located in close proximity to the SMC1 hinge (Fig. 3A). An extra unassigned density extends from the N-terminal region of NIPBL^{C} and inserts into a surface groove connecting the SMC1 hinge and STAG1 (fig. S10B). Thus, there may also be interactions between the SMC1-SMC3 hinge heterodimer and NIPBL.

In the reported structure of the SMC1-SMC3 hinge heterodimer (*30*), the two subunits are tightly packed through two distinct interfaces (termed north and south interfaces), forming a donut-shaped structure with a central pore (Fig. 3B). The north interface of the hinge heterodimer is disrupted in the cohesin-NIPBL^{C}-DNA complex, producing a half-open conformation that resembles an open washer (Fig. 3C). The hinge domains of SMC complexes, including cohesin, have been shown to bind single-stranded DNA (ssDNA) or dsDNA (*45*–*47*). We did not, however, identify any density close to the SMC1-SMC3 hinge that could be attributed to DNA.

We next determined the crystal structures of the isolated SMC1-SMC3 hinge heterodimer in the presence of DNA. The structures of the two different versions of the SMC1-SMC3 hinge heterodimer adopt different open conformations: one with the north interface open and the other with the south interface open (Fig. 3, D and E, and table S2). The conformation with the north interface open closely resembles that observed in the cohesin-NIPBL^{C}-DNA complex (Fig. 3, C and D). Thus, both interfaces of the SMC1-SMC3 hinge are intrinsically dynamic and can undergo close-open transitions. Whether these transitions are coupled to the ATPase cycle of cohesin or further regulated by binding to STAG1 or DNA remains unknown.

The crystal structure of the SMC1-SMC3 hinge with the north interface open contained a two-nucleotide ssDNA bound to the outer surface of the SMC1 hinge (Fig. 3D and fig. S10C). The open conformation of the hinge in the cohesin-NIPBL^{C}-DNA complex is likely capable of binding ssDNA. The open hinge may act to capture ssDNA and then allow its entry into the SMC chamber through the central pore (*48*, *49*). Notably, STAG1 and ssDNA contact overlapping surfaces in the SMC1 hinge (fig. S10D), suggesting that they might compete for hinge binding. The potential function of this hinge surface remains to be determined.

DNA binding by cohesin and NIPBL

DNA is bound at the central tunnel of the cohesin-NIPBL^{C} complex, contacting all subunits (Fig. 4A). The major DNA binding pocket is formed by the heterodimer of SMC1-SMC3 HDs, NIPBL, and the RAD21 NHD (which is bound to the CC region of SMC3) (Fig. 4B). DNA is tightly gripped by the V-shaped SMC1-SMC3 HD heterodimer, in a manner similar to DNA binding by the DNA repair protein Rad50 (*50*). The R-loop and the succeeding β sheet in the N-lobe of both SMC1 and SMC3 are the major sites for DNA binding (fig. S11). The acetylation sites Lys^{105} and Lys^{106} in SMC3 that contact NIPBL are also located in close proximity to the DNA and may contribute to DNA binding (Fig. 4C). Acetylation of Lys^{105} and Lys^{106} of SMC3 may weaken DNA binding to the SMC3 HD and block the ability of DNA to stimulate the ATPase activity of cohesin, as previously proposed (*44*).

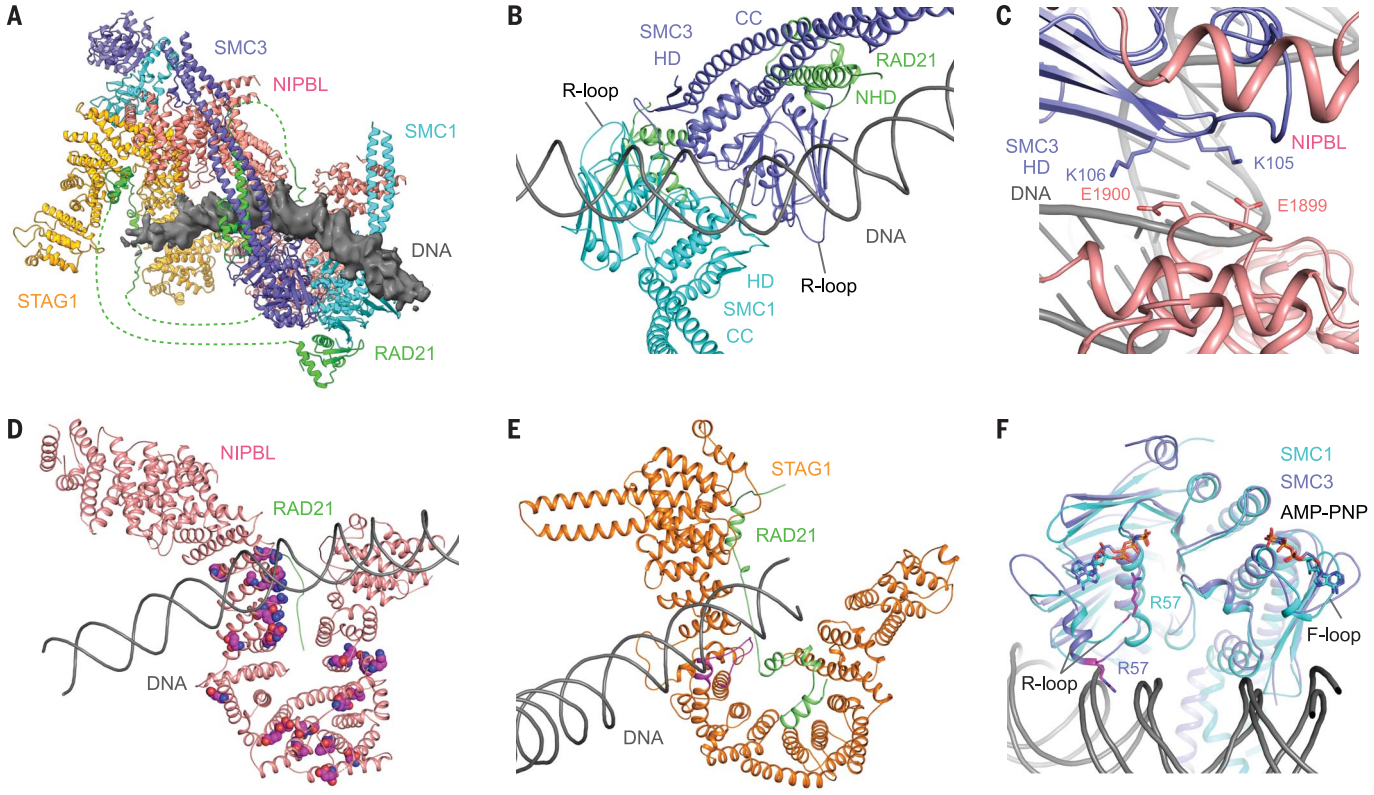


Fig. 4. DNA binding by cohesin and NIPBL^{C}. (**A**) Structure of the cohesin-NIPBL^{C}-DNA complex, with protein subunits shown in cartoon representation and the EM map of DNA shown as gray surface. (**B**) Interface between DNA and the SMC1-SMC3 heterodimer. (**C**) Close-up view of the DNA and NIPBL^{C} interactions involving the acetylation acceptor lysines of SMC3. (**D**) Interactions between DNA and NIPBL^{C}. NIPBL mutations in patients with Cornelia de Lange syndrome are highlighted as spheres. (**E**) Interactions between DNA and STAG1. The DNA binding element of STAG1 is colored purple. (**F**) Structural comparison of SMC1 and SMC3 HDs bound to DNA. AMP-PNP and the conserved arginines in the R-loops of SMC1 and SMC3 are shown as sticks.

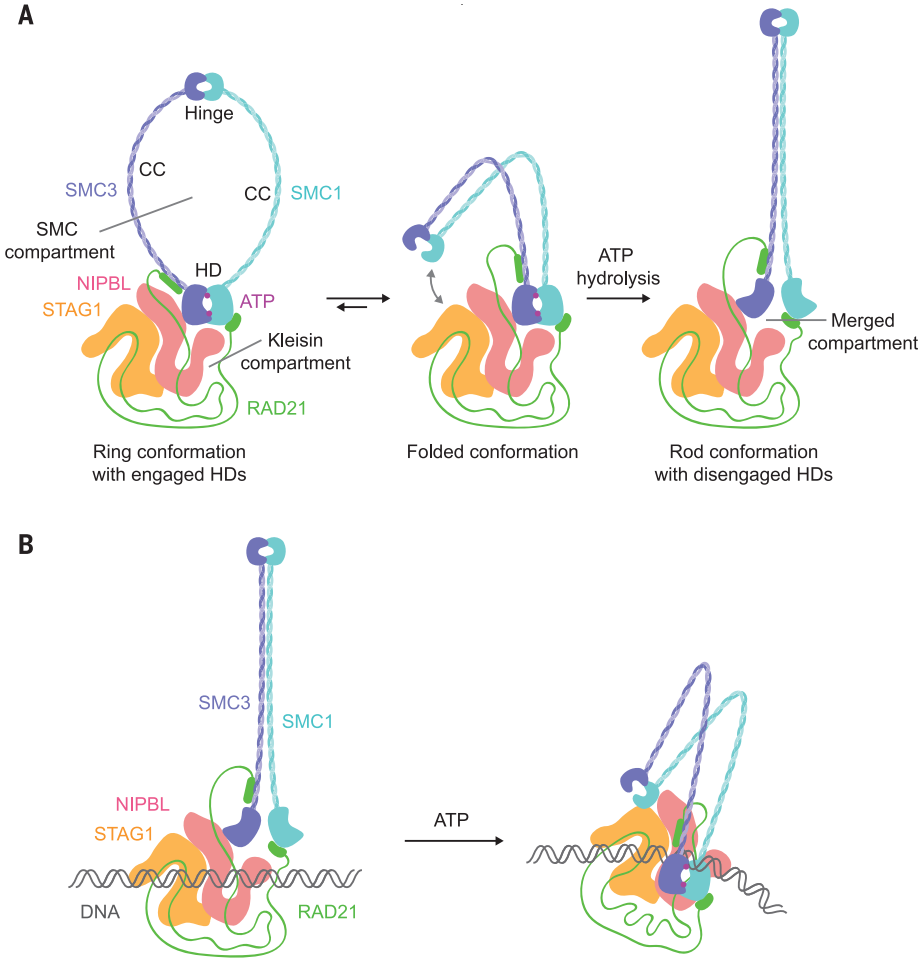


Fig. 5. DNA entrapment by cohesin. (**A**) Schematic representation of the cohesin-NIPBL complex in HD-engaged and HD-disengaged conformations. The HD-engaged ring conformation might exist in an equilibrium with a folded conformation with the hinge domains docked on STAG1. The SMC and kleisin compartments of HD-engaged cohesin are indicated. These compartments merge in the HD-disengaged conformation. Owing to the lack of cryo-EM density for major parts of SMC CCs, it is unclear whether the interaction between STAG1 and SMC hinge domains is intramolecular (depicted) or intermolecular (not depicted). (**B**) Model for cohesin loading onto DNA. DNA is captured by NIPBL and STAG1 in the HD-disengaged cohesin. Upon ATP binding, SMC1 and SMC3 HDs engage, creating a DNA binding site between their CCs and an interface for NIPBL binding. RAD21-tethered NIPBL binds the SMC1-SMC3 heterodimer and, together with STAG1, forms the central tunnel to entrap DNA.

The HEAT repeat subunit Ycg1 in the related condensin complex directly interacts with DNA (*51*). Similarly, the HEAT repeat proteins in cohesin, NIPBL and STAG1, serve as major DNA binding receptors (Fig. 4, D and E). DNA binding to NIPBL buries a large solvent-exposed surface of NIPBL (fig. S12A). Several NIPBL missense mutations found in patients with Cornelia de Lange syndrome (*52*) affect residues in this DNA binding site, as well as those in the RAD21 binding site (Fig. 4D). Mutations of yeast Scc2 residues that correspond to several DNA binding residues in NIPBL are lethal (*22*). These findings confirm the functional importance of this NIPBL-DNA binding interface.

STAG1 binds to DNA through the N-terminal part of a helix in R8 and an extended loop that follows the helix (Fig. 4E and fig. S12B). The DNA binding interface of STAG1 in our structure is different from that in the reported structure of the yeast Scc3-DNA complex (*53*) (fig. S12, C to F). Most DNA binding interfaces of cohesin-NIPBL^{C} are mediated by electrostatic interactions with the phosphate backbone. This feature explains the salt sensitivity and lack of sequence specificity in initial DNA binding by the cohesin-NIPBL complex (*11*, *12*, *21*).

Cohesin activation by NIPBL and DNA

SMC1 and SMC3 belong to the family of ATP-binding cassette ATPases (*54*). They bind two ATP molecules at the dimer interface of their ATPase HDs (Fig. 2C). SMC1-SMC3 HD heterodimerization and ATP binding are interdependent events. NIPBL simultaneously contacts both SMC1 and SMC3 HDs (Fig. 4A), thus promoting their heterodimerization. Likewise, DNA also bridges both HDs to stabilize the heterodimer (Fig. 4B). Thus, one mechanism by which NIPBL and DNA activate the ATPase activity of cohesin is to promote and stabilize the ATP-binding, HD-engaged state of the SMC1-SMC3 heterodimer.

DNA and NIPBL trigger both localized and domain-specific conformational changes in SMC1 and SMC3 HDs. NIPBL and DNA contact and organize the loops involved in ATP binding. For example, DNA directly contacts the R-loop of SMC1 HD and organizes this loop for ATP binding at the SMC1 site (Fig. 4, B and F). NIPBL and DNA also induce a conformational change of the SMC3 HD and CC. Compared to yeast Smc3 bound to Scc1 NHD (*15*), the N-lobe of SMC3 HD in the intact complex undergoes a rotation relative to its C-lobe and CC (fig. S6D). This rotation positions the N-lobe closer to the SMC1 HD. The conformations of SMC1 and SMC3 HDs observed here resemble that of an engineered yeast cohesin subcomplex (fig. S6E) (*55*). The remodeling of the active sites by NIPBL and DNA may also stimulate the ATPase activity of cohesin.

Implications for cohesin loading

SMC complexes can adopt an open ring–like conformation with engaged HDs or a rod-shaped conformation with disengaged HDs (*39*, *40*, *56*–*58*) (Fig. 5A). Our structure captures the HD-engaged conformation with an open SMC compartment. DNA binds at the top of the engaged SMC1-SMC3 HDs and between the CCs—i.e., it is trapped inside the SMC compartment. The two HEAT repeat proteins NIPBL and STAG1 further wrap around DNA, creating a central tunnel. This conformational state of the cohesin-NIPBL complex does not permit long DNA (e.g., eukaryotic chromosomes) to enter or escape from this tunnel, without undergoing conformational rearrangements. Thus, our structure of the cohesin-NIPBL^{C}-DNA complex does not capture the initial docking complex but rather represents a postloading state. We propose that DNA is first recruited by NIPBL and STAG1 to cohesin in the HD-disengaged state (Fig. 5B). Upon ATP binding, SMC1 and SMC3 HDs engage to trap the DNA inside the SMC compartment. NIPBL and STAG1 then wrap around the DNA to form the central tunnel and further entrap the DNA.

The cohesin-NIPBL^{C}-DNA complex has a layered architecture with multiple DNA binding sites, which suggests an inchworm model for DNA translocation and loop extrusion. In such a model, the ATP-bound, engaged SMC HDs together with NIPBL constitute

a strong DNA binding site. Once ATP is hydrolyzed, the SMC HDs disengage and loosen their grips on DNA. SMC1 and SMC3 HDs redimerize upon ATP binding and, together with NIPBL, capture a downstream segment of DNA. The repeat of such a cycle could lead to processive translocation of cohesin on DNA. SMC3 acetylation, displacement of NIPBL by PDS5, and CTCF binding antagonize loop extrusion (*19*, *22*, *42*).

Sister chromatid cohesion requires the topological entrapment of both sister chromatids by cohesin (*14*). The topological loading of cohesin onto DNA involves the opening of any of the three interfaces: SMC1-RAD21, SMC3-RAD21, or the hinge. Our finding that the SMC1-SMC3 hinge can adopt two types of open conformations raises the possibility that the hinge might be an entry gate for DNA. Consistent with this notion, mutations of certain basic residues that line the central pore of the hinge disrupt sister chromatid cohesion in yeast (*14*, *30*).

Outlook

In this study, we determined the cryo-EM structure of the intact human cohesin in complex with its loader NIPBL and DNA, catching the first glimpse of the overall architecture of this SMC complex in the DNA-bound state. Our structure explains a large body of available biochemical findings, provides insight into the mechanisms of cohesin loading, and offers a blueprint for further functional and mechanistic dissection of cohesin and related molecular machines.

REFERENCES AND NOTES

1. S. Yatskevich, J. Rhodes, K. Nasmyth, *Annu. Rev. Genet.* **53**, 445–482 (2019).
2. F. Uhlmann, *Nat. Rev. Mol. Cell Biol.* **17**, 399–412 (2016).
3. J. H. Haarhuis, A. M. Elbatsh, B. D. Rowland, *Dev. Cell* **31**, 7–18 (2014).
4. G. Zheng, H. Yu, *Sci. China Life Sci.* **58**, 1089–1098 (2015).
5. E. Alipour, J. F. Marko, *Nucleic Acids Res.* **40**, 11202–11212 (2012).
6. J. H. I. Haarhuis *et al.*, *Cell* **169**, 693–707.e14 (2017).
7. S. S. P. Rao *et al.*, *Cell* **171**, 305–320.e24 (2017).
8. G. Wutz *et al.*, *EMBO J.* **36**, 3573–3599 (2017).
9. W. Schwarzer *et al.*, *Nature* **551**, 51–56 (2017).
10. L. Vian *et al.*, *Cell* **173**, 1165–1178.e20 (2018).
11. I. F. Davidson *et al.*, *Science* **366**, 1338–1345 (2019).
12. Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, H. Yu, *Science* **366**, 1345–1349 (2019).
13. C. H. Haering, A. M. Farcas, P. Arumugam, J. Metson, K. Nasmyth, *Nature* **454**, 297–301 (2008).
14. M. Srinivasan *et al.*, *Cell* **173**, 1508–1519.e18 (2018).
15. T. G. Gligoris *et al.*, *Science* **346**, 963–967 (2014).
16. P. J. Huis in 't Veld *et al.*, *Science* **346**, 968–972 (2014).
17. C. H. Haering *et al.*, *Mol. Cell* **15**, 951–964 (2004).
18. K. Hara *et al.*, *Nat. Struct. Mol. Biol.* **21**, 864–870 (2014).
19. Y. Li *et al.*, *Nature* **578**, 472–476 (2020).
20. J. Rhodes, D. Mazza, K. Nasmyth, S. Uphoff, *eLife* **6**, e30000 (2017).
21. Y. Murayama, F. Uhlmann, *Nature* **505**, 367–371 (2014).
22. N. J. Petela *et al.*, *Mol. Cell* **70**, 1134–1148.e7 (2018).
23. S. Kueng *et al.*, *Cell* **127**, 955–967 (2006).
24. T. Hartman, K. Stead, D. Koshland, V. Guacci, *J. Cell Biol.* **151**, 613–626 (2000).
25. T. Rolef Ben-Shahar *et al.*, *Science* **321**, 563–566 (2008).
26. E. Üal *et al.*, *Science* **321**, 566–569 (2008).
27. J. Zhang *et al.*, *Mol. Cell* **31**, 143–151 (2008).
28. B. D. Rowland *et al.*, *Mol. Cell* **33**, 763–774 (2009).
29. T. Nishiyama *et al.*, *Cell* **143**, 737–749 (2010).
30. A. Kurze *et al.*, *EMBO J.* **30**, 364–378 (2011).
31. M. B. Roig *et al.*, *FEBS Lett.* **588**, 3692–3702 (2014).
32. W. C. Chao *et al.*, *Cell Rep.* **12**, 719–725 (2015).
33. S. Kikuchi, D. M. Borek, Z. Otwinowski, D. R. Tomchick, H. Yu, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12444–12449 (2016).
34. B. G. Lee *et al.*, *Cell Rep.* **14**, 2108–2115 (2016).
35. K. W. Muir *et al.*, *Cell Rep.* **14**, 2116–2126 (2016).
36. Z. Ouyang, G. Zheng, D. R. Tomchick, X. Luo, H. Yu, *Mol. Cell* **62**, 248–259 (2016).
37. W. C. Chao *et al.*, *Nat. Commun.* **8**, 13952 (2017).
38. S. M. Hinshaw, V. Makrantoni, A. Kerr, A. L. Marston, S. C. Harrison, *eLife* **4**, e06057 (2015).
39. M. T. Hons *et al.*, *Nat. Commun.* **7**, 12523 (2016).
40. M. L. Diebold-Durand *et al.*, *Mol. Cell* **67**, 334–347.e5 (2017).
41. M. Hassler *et al.*, *Mol. Cell* **74**, 1175–1188.e9 (2019).
42. L. Dauban *et al.*, *Mol. Cell* **77**, 1279–1293.e4 (2020).
43. M. A. Deardorff *et al.*, *Nature* **489**, 313–317 (2012).
44. Y. Murayama, F. Uhlmann, *Cell* **163**, 1628–1640 (2015).
45. A. Chiu, E. Revenkova, R. Jessberger, *J. Biol. Chem.* **279**, 26233–26242 (2004).
46. J. J. Griese, G. Witte, K. P. Hopfner, *Nucleic Acids Res.* **38**, 3454–3465 (2010).
47. A. Alt *et al.*, *Nat. Commun.* **8**, 14011 (2017).
48. S. Gruber *et al.*, *Cell* **127**, 523–537 (2006).
49. Y. Murayama, C. P. Samora, Y. Kurokawa, H. Iwasaki, F. Uhlmann, *Cell* **172**, 465–477.e15 (2018).
50. A. Rojowska *et al.*, *EMBO J.* **33**, 2847–2859 (2014).
51. M. Kschonsak *et al.*, *Cell* **171**, 588–600.e24 (2017).
52. L. Mannini, F. Cucco, V. Quarantotti, I. D. Krantz, A. Musio, *Hum. Mutat.* **34**, 1589–1596 (2013).
53. Y. Li *et al.*, *eLife* **7**, e38356 (2018).
54. T. Hirano, *Genes Dev.* **16**, 399–414 (2002).
55. K. W. Muir, Y. Li, F. Weis, D. Panne, *Nat. Struct. Mol. Biol.* **27**, 233–239 (2020).
56. C. H. Haering, J. Löwe, A. Hochwagen, K. Nasmyth, *Mol. Cell* **9**, 773–788 (2002).
57. K. Kamada, M. Su'etsugu, H. Takada, M. Miyata, T. Hirano, *Structure* **25**, 603–616.e4 (2017).
58. D. E. Anderson, A. Losada, H. P. Erickson, T. Hirano, *J. Cell Biol.* **156**, 419–424 (2002).

ACKNOWLEDGMENTS

Single-particle cryo-EM data were collected at the University of Texas Southwestern Medical Center (UTSW) Cryo-Electron Microscopy Facility. X-ray diffraction data were collected at the beamline 19-ID of the Argonne National Laboratory, Structural Biology Center, at the Advanced Photon Source. SBC-CAT is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research, under contract DE-AC02-06CH11357. We thank D. Nicastro and D. Stoddard for cryo-EM facility access and data acquisition and D. Tomchick, Z. Chen, and Y. Li at the Structural Biology Laboratory for assistance with x-ray data collection and cryo-EM grid screening. **Funding:** This study was supported by the Cancer Prevention and Research Institute of Texas (CPRIT) (RR160082 to X.-c.B, and RP160667-P2 to H.Y.), National Institutes of Health (GM136976 to X.-c.B.), and the Welch Foundation (I-1944 to X.-c.B. and I-1441 to H.Y.). X.-c.B. is the Virginia Murchison Linthicum Scholar in Medical Research at UTSW. The UTSW Cryo-Electron Microscopy Facility is funded by a CPRIT Core Facility Support Award (RP170644). **Author contributions:** Z.S. prepared samples, performed x-ray crystallography studies, and built and refined structural models. H.G. assisted with cryo-EM grid screening. X.-c.B. collected and processed cryo-EM data, and determined the structure. X.-c.B. and H.Y. supervised the project. Z.S., X.-c.B, and H.Y. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data produced and materials used in this study are available. Cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession codes EMD-21658 and EMD-21663. Atomic coordinates have been deposited in the Protein Data Bank with IDs 6WG3, 6WG4, 6WG6, and 6WGE.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1454/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 and S2
References (*59*–*79*)
MDAR Reproducibility Checklist
Movies S1 to S3

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

29 January 2020; accepted 14 April 2020
Published online 14 May 2020
10.1126/science.abb0981

STUCTURAL BIOLOGY

Mechanisms of OCT4-SOX2 motif readout on nucleosomes

Alicia K. Michael[1]*, Ralph S. Grand[1]*, Luke Isbel[1]*, Simone Cavadini[1], Zuzanna Kozicka[1,2], Georg Kempf[1], Richard D. Bunker[1]†, Andreas D. Schenk[1], Alexandra Graff-Meyer[1], Ganesh R. Pathare[1], Joscha Weiss[1], Syota Matsumoto[1], Lukas Burger[1,3], Dirk Schübeler[1,2]‡, Nicolas H. Thomä[1]‡

Transcription factors (TFs) regulate gene expression through chromatin where nucleosomes restrict DNA access. To study how TFs bind nucleosome-occupied motifs, we focused on the reprogramming factors OCT4 and SOX2 in mouse embryonic stem cells. We determined TF engagement throughout a nucleosome at base-pair resolution in vitro, enabling structure determination by cryo–electron microscopy at two preferred positions. Depending on motif location, OCT4 and SOX2 differentially distort nucleosomal DNA. At one position, OCT4-SOX2 removes DNA from histone H2A and histone H3; however, at an inverted motif, the TFs only induce local DNA distortions. OCT4 uses one of its two DNA-binding domains to engage DNA in both structures, reading out a partial motif. These findings explain site-specific nucleosome engagement by the pluripotency factors OCT4 and SOX2, and they reveal how TFs distort nucleosomes to access chromatinized motifs.

Transcription factors (TFs) regulate gene expression and govern cell identity by interacting with specific sequence motifs. OCT4 and SOX2 serve as reprogramming TFs (*1*) that cooperate as critical mediators of pluripotency (*2*–*4*). Although they are not obligate heterodimers in solution, the combined OCT4-SOX2 motif drives OCT4-SOX2 complex formation (*5*, *6*).

Chromatin restricts DNA access (*7*, *8*), but a specialized subset of TFs, termed pioneer factors, can engage chromatinized motifs to trigger cell-fate changes (*9*). Several TFs have been shown to bind motifs embedded in nucleosomes in vitro (*1*, *10*, *11*); however, the nucleosome architecture, with histones H2A, H2B, H3, and H4 and its two DNA gyres (*12*), restricts TF access to >95% of nucleosomal DNA (*13*). Two extreme scenarios for nucleosomal TF-engagement have been put forward: TF binding without changing the nucleosomal architecture (*10*, *14*) or TF-mediated changes to the nucleosome by distorting the histone core, looping the DNA, or taking advantage of nucleosome unwrapping dynamics at the entry-exit sites (*15*, *16*).

OCT4 binding is predicted to be incompatible with the nucleosome architecture on the basis of its engagement with free DNA (*17*–*19*), although partial motifs have been identified where OCT4 engages only a portion of its binding site to maintain nucleosome integrity (*1*). Despite being critical for genome regulation, the structural and mechanistic principles governing nucleosome engagement by single or multiple TFs have yet to be determined.

SeEN-seq reveals preferential binding of OCT4-SOX2 to nucleosomal entry-exit sites

Each rotational and translational setting of a motif around a nucleosome places the TF in a different histone-DNA environment. To assess all potential TF-binding registers on a nucleosome, we built on approaches that examine select loci (*20*, *21*) and developed selected engagement on nucleosome sequencing (SeEN-seq), which measures the relative affinity of one or multiple TFs for each nucleosomal register in parallel (Fig. 1A). We focused on OCT4 and its dependence on SOX2 because both factors show strong co-occupancy at target genes in vivo (*22*, *23*), and the OCT4-SOX2 protein-protein interface is required for pluripotency (*24*). The canonical OCT4-SOX2 motif (matrix ID MA0142.1) (*25*) was tiled at 1–base pair (bp) intervals through a 601-nucleosome positioning sequence (*26*). The sequence library was then assembled into octameric nucleosome core particles (NCPs), incubated with the TF(s), and subjected to electrophoretic mobility shift assay (EMSA). TF-bound and unbound nucleosome complexes were excised and subjected to next-generation sequencing (NGS), which generated single molecule counts that approximate motif affinity as a function of position.

A range of protein concentrations of either OCT4, SOX2 (residues 37 to 118), or OCT4-SOX2 together were assessed, with a total of 4752

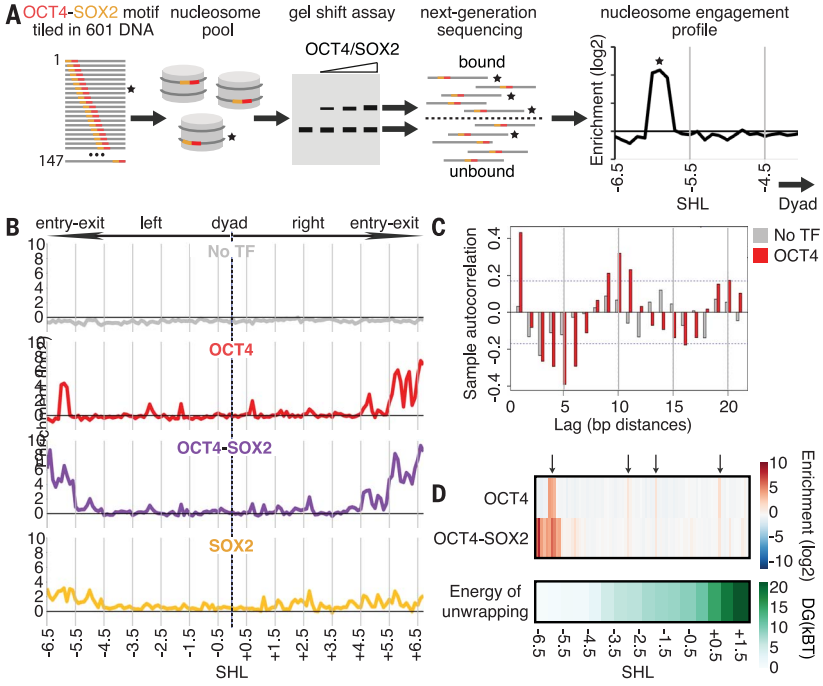


Fig. 1. SeEN-seq identifies preferred binding sites of OCT4 and OCT4-SOX2 across the nucleosome. (**A**) Principle of SeEN-seq. A library of TF motif–containing nucleosome positioning sequences is assembled into nucleosomes and incubated with TF(s). TF-bound and unbound nucleosome complexes are separated by EMSA and sequenced, revealing position-specific enrichments. Stars indicate a specific example sequence at each step of the assay. (**B**) SeEN-seq enrichment (n = 3 replicates) for nucleosome pool with or without TFs. Further details are provided in fig. S1C. Motif position is indicated by superhelix location (SHL) that describes where the minor groove faces away (SHLs ±1, ±2, etc.) or toward (SHLs ±1.5, ±2.5, etc.) the histone octamer. (**C**) Autocorrelation analysis of OCT4 enrichment; dashed lines indicate 95% confidence interval. (**D**) DNA nucleosome unwrapping energy (*15*) versus OCT4 and OCT4-SOX2 SeEN-seq enrichment profile. Arrows indicate enriched positions in OCT4. DG, delta Gibbs free energy; kBT, 0.6 kcal/mol, where T = 300 K.

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. [2]Faculty of Science, University of Basel, Petersplatz 1, 4003 Basel, Switzerland. [3]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland.
*These authors contributed equally to this work.
†Present address: Monte Rosa Therapeutics, Aeschenvorstadt 36, 4051 Basel, Switzerland.
‡Corresponding author. Email: nicolas.thoma@fmi.ch (N.H.T.); dirk@fmi.ch (D.S.)

conditions measured with high reproducibility (Fig. 1B and fig. S1, A to C). The trends in OCT4 or SOX2 binding at selected positions were validated by fluorescence polarization measurements (fig. S2, A to E). Previously, two OCT4 motif locations were tested and found to provide similar OCT4 access to the nucleosome-embedded motif (*27*). Although this was recapitulated in SeEN-seq, our comprehensive examination of all motif locations reveals clear OCT4 preference for nucleosomal DNA entry-exit sites as well as discrete preferential binding sites ~1 to 3 bp in width throughout the nucleosome (Fig. 1B and fig. S1C). SOX2 shows less differential enrichment in SeEN-seq, in line with published data (*10*), with some degree of preferred binding toward the entry-exit sites and near the dyad (Fig. 1B and fig. S1C). Given the small enrichment amplitudes for SOX2 alone, we focused on the robust and differential binding activity seen for OCT4 and how it is affected by SOX2. OCT4 and SOX2 cooperate strongly, which results in up to 650-fold increased binding compared with that caused by OCT4 alone (fig. S1C). This effect is most evident at entry-exit sites with weaker cooperativity observed at internal sites (fig. S1C).

Although OCT4-SOX2 binding appears roughly symmetrical across the dyad, for OCT4 alone, enrichment in the right half of the nucleosome [superhelix locations (SHLs) +4 to +6.5] is more pronounced compared with that in the left (SHLs −4 to −6.5) (Fig. 1B and fig. S1C). A notable difference is the motif orientation on either side of the dyad; on the left, the OCT4 portion of the motif is closest to the dyad, whereas on the right it is oriented toward the entry-exit site of the nucleosome. Positions enriched for OCT4 alone and OCT4-SOX2 exhibited stronger binding at discrete sites with 10-bp periodicity across the nucleosome (Fig. 1C and fig. S2, F and G). Both OCT4 and OCT4-SOX2 show a trend of stronger binding at the entry-exit site than at the dyad (Fig. 1D). This would be expected if OCT4 and OCT4-SOX2 binding was facilitated by nucleosomal breathing, which is more pronounced toward the nucleosomal entry-exit sites (*15*). However, binding is not simply a function of nucleosomal breathing, because in the presence of OCT4 alone, motifs on the left half of the nucleosome (SHLs −6.5 to −6.0) are not bound across 5 bp, whereas adjacent motifs more proximal to the dyad are tightly bound (Fig. 1, B and D, and fig. S2H). These entry-exit site loci are unbound by OCT4 alone, but they are bound cooperatively with SOX2 (Fig. 1D and fig. S2H). Thus, spatial orientation of the motif, cooperativity, and nucleosomal breathing dynamics all govern OCT4-SOX2 binding.

Structure of OCT4-SOX2 bound at SHL −6 shows DNA release from the histone core

To dissect mechanisms of nucleosome engagement, we performed structural studies with OCT4-SOX2, which co-bind key target genes in vivo and have a cooperative effect in vitro (fig. S2, I to L) (*22*). For structural studies, we focused on sites that show discrete OCT4 SeEN-seq enrichment and cooperative binding in the presence of SOX2 (fig. S3A). A site 57 bp from the dyad (SHL −6) enabled structure determination from ~90,000 particles at an overall resolution of 3.1 Å (Fig. 2, A and B, and fig. S3). The nucleosome core and discrete domains of OCT4 [POU-specific (POUS)] and SOX2 [high-mobility group (HMG)] were well resolved (see fig. S3F for local resolution), which allowed conservative refinement with reference restraints to existing high-resolution OCT4 and SOX2 DNA-bound crystal structures (Fig. 2, B and C; fig. S4; and fig. S5, A to D) (*18*, *19*).

In the OCT4-SOX2-NCP^{SHL-6} structure, OCT4 and SOX2 together remove the DNA from core histones (Fig. 2, C and D). OCT4 has a bipartite DNA-binding domain composed of a POUS and POU-homeodomain (POUHD) separated by 17 residues, whereas SOX2 utilizes an HMG domain (*18*). When nucleosome-bound, the OCT4-POUS and SOX2-HMG DNA-binding domains engage major and minor grooves,

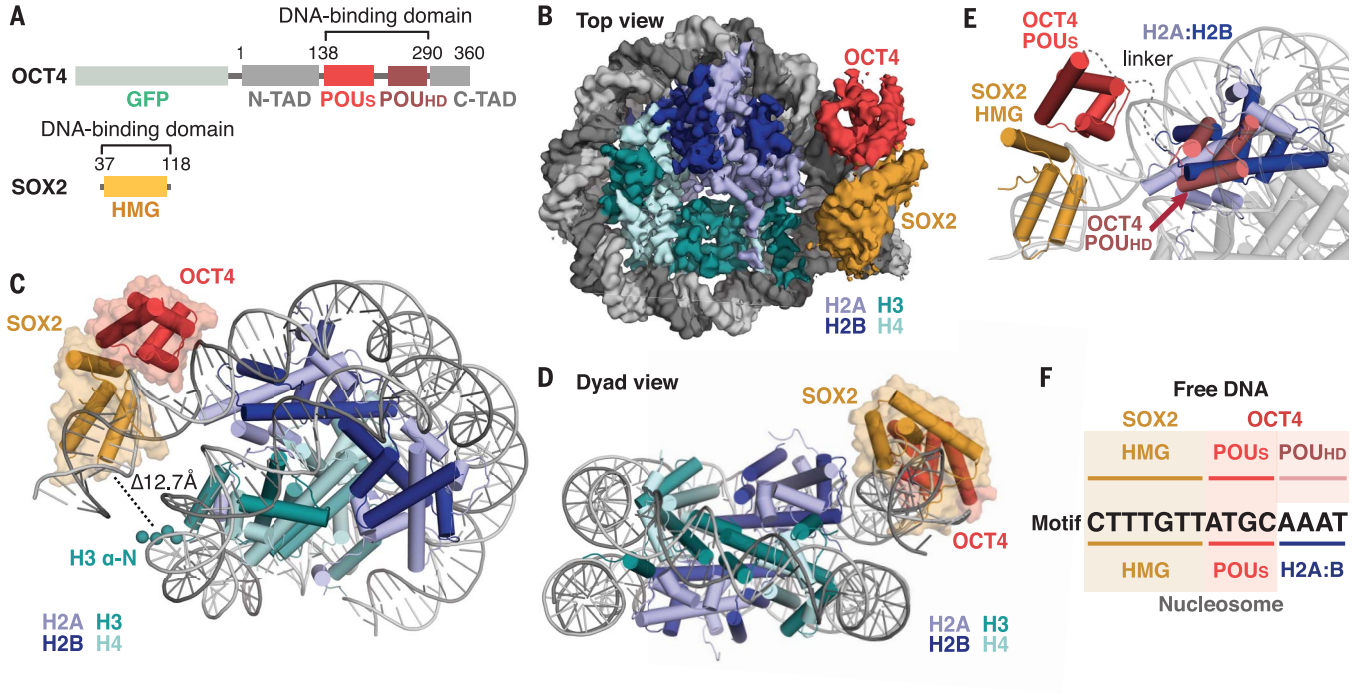


Fig. 2. Cryo-EM structure of OCT4-SOX2-NCP^{SHL-6} complex. (**A**) Domain schematic of OCT4 and SOX2 constructs. GFP, green fluorescent protein; TAD, transactivation domain. (**B**) Cryo-EM map of OCT4-SOX2-NCP^{SHL-6} at 3.1-Å resolution. (**C**) Model of the OCT4-SOX2-NCP^{SHL-6} complex. The H3 N-terminal helix and tail (shown as spheres) stabilize DNA at the entry-exit in a canonical nucleosome (*31*). The presence of OCT4-SOX2 increases the distance between H3 (Arg^{41}) and the nearest DNA base (T143) by 12.7 Å as compared with the unbound nucleosome. (**D**) Dyad view of the OCT4-SOX2-NCP^{SHL-6} structure. (**E**) In the depicted OCT4-SOX2 arrangement, a model of POUHD engagement with its motif shows significant clash with the H2A:H2B dimer. Overlay of free DNA–bound structure (PDB: 1O4X) with the nucleosome-bound structure, aligned on the DNA (see also fig. S5E). (**F**) Schematic of free DNA versus the observed nucleosome-binding mode of OCT4-SOX2.

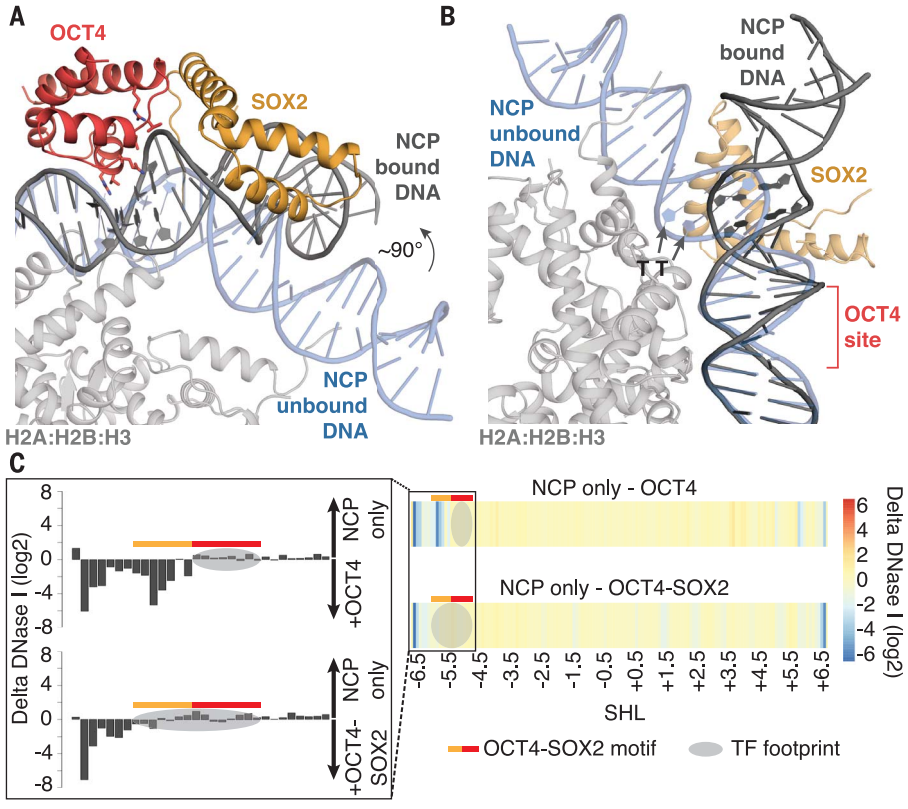


Fig. 3. OCT4-SOX2 lifts the entry-exit DNA away from the histone core. (**A**) Comparison of the unbound nucleosome DNA (blue) with the OCT4-SOX2 bound nucleosome structure (gray). The DNA is kinked ~90° away from the histones. Residues at the OCT4-DNA interface are shown as sticks. POUS motif nucleotides are shown as ribose and base rings. (**B**) SOX2 kinks the nucleosomal DNA away from the histones. SOX2 uses conserved (Phe[48] and Met[49]) residues to intercalate a TT base step, indicated by arrows (see also fig. S8A). OCT4 is removed for clarity. (**C**) Difference in DNaseI digestion across the nucleosome in the presence of OCT4 or OCT4-SOX2 (right). Zoom-in view around the OCT4-SOX2 motif (left).

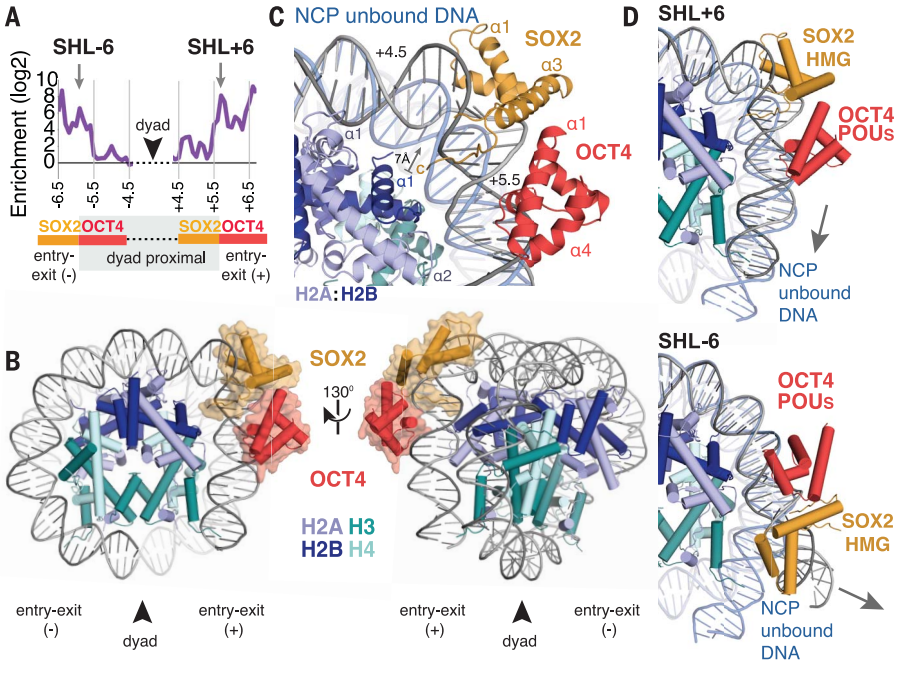


Fig. 4. OCT4-SOX2 bound at SHL +6 induces localized DNA distortion. (**A**) Depiction of OCT4-SOX2 SeEN-seq profile illustrating symmetric positions used for structure determination (arrows). As the OCT4-SOX2 motif is tiled across the nucleosome, the orientation of the proteins in relation to the dyad is inverted. (**B**) OCT4-SOX2-NCP^{SHL+6} model. (**C**) Details of SOX2-induced DNA kink. SOX2 binding locally distorts DNA near SHL +5. The C terminus of SOX2 approaches the histone core when bound in this orientation, distorting DNA away from the stabilizing H2B residues [N terminus (Arg[34]) and α1]. The 7-Å movement indicated is calculated between the phosphate backbone at Cys[123] in the bound and unbound DNA. SHLs are shown (SHLs +4.5 and +5.5). (**D**) Comparison of the DNA trajectory in the OCT4-SOX2-NCP^{SHL+6} (SHL +6) (top) versus OCT4-SOX2-NCP^{SHL-6} (SHL −6) (bottom) structures. The DNA trajectory at the entry-exit site of the TF-bound structures is indicated with an arrow.

respectively, with protein-DNA interactions consistent with those previously seen for the individual OCT4-POUS domain and SOX2 on free DNA (fig. S5E). The DNA remains attached and straightened around the OCT4 site but is detached around the SOX2 motif (Fig. 2 and fig. S5F). SOX2 kinks the DNA and, together with OCT4, synergistically releases DNA from the core histones (movie S1). OCT4-SOX2 forms no discernable histone contacts, with the DNA separating them from the nucleosome core (Fig. 2, C and D). Density for the OCT4 transactivation domains was not observed, which is consistent with similar nucleosome-affinity measurements for full-length OCT4 and OCT4 DNA-binding domain only (fig. S5G).

OCT4 recognizes a partial motif, engaging DNA with its POUS domain, whereas the

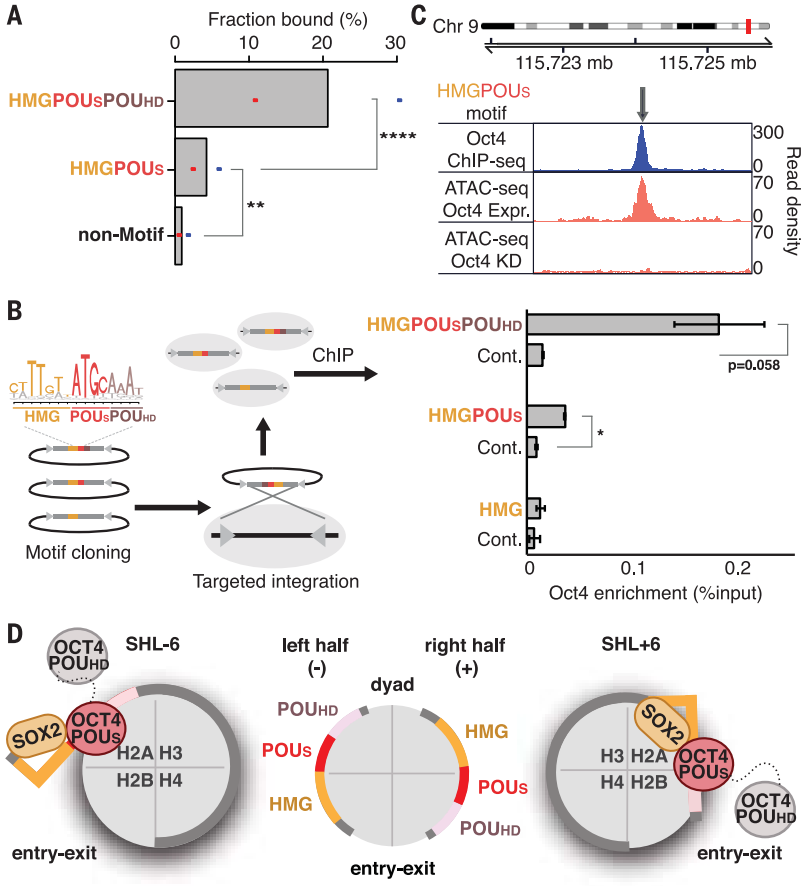


Fig. 5. The HMG-POUS partial motif is bound in vivo and opens chromatin. (**A**) Fraction of the top thousand motifs, ranked by match to the position weight motif, that are bound by Oct4 (2-fold ChIP-seq enrichment). Mean and individual values from two separate datasets are shown (*33*, *34*). Significant enrichment is detected for full and partial motifs compared with loci without a motif (lowest 1000 genomic matches to the Oct4-Sox2 motif). $^{**}P < 0.01$; $^{****}P < 0.0001$. (Pearson's chi-square test, indicated for dataset with least significance). (**B**) Schematic representation of targeted motif insertion and Oct4 ChIP-qPCR enrichment at the ectopic insertion locus and endogenous control (Cont.) locus. $^{*}P < 0.05$; error bars indicate SEM of at least two biological replicates. (**C**) Representative genomic region (chr9: 115723244 to 115725243) that shows binding of Oct4 to a partial HMG-POUS motif (top) and Oct4-dependent chromatin accessibility (bottom). Expr., Oct4-expressing cells; KD, Oct4-knockdown cells. (**D**) Graphical representation of the nucleosome distortion induced by OCT4-SOX2 when bound at dyad-symmetric sites.

POUHD is not engaged. On free DNA, both POU domains engage the major groove spanning 8 bp on opposite sides of the DNA and would clash with either the histones or DNA gyres at all locations on the nucleosome (*19*, *28*) (Fig. 2E and fig. S6A). To stabilize the OCT4-SOX2-NCP^{SHL-6} complex for imaging, GraFix cross-linking was necessary and a cross-link was evident between H2A and H2B (fig. S6B) (*29*). To test if POUHD access was blocked by the cross-linking of histones, we solved an ~4-Å map of the non–cross-linked OCT4-SOX2-NCP^{SHL-6} nucleosome (fig. S7, A to E), which resulted in a largely indistinguishable model (root mean square deviation, 1.3 Å; fig. S7, F and G). As the OCT4 POUHD motif is occluded by H2A-H2B, the binding mechanism involving only the OCT4 POUS is consistent with partial motif engagement (Fig. 2F) (*1*). Partial motif recognition, however, does not necessarily render TF binding compatible with the nucleosomal DNA structure (fig. S5F).

In the context of the nucleosome, SOX2 competes with histones for DNA binding and kinks DNA by ~90° at SHL −6.5 away from the histones, similar to HMG domains on free DNA (Fig. 3, A and B, and fig. S8A) (*18*). This is accomplished by intercalation of the SOX2 Phe^{48} and Met^{49} wedge at the TT base step (*18*). Variant SOX2 motifs that lessen distortion of the DNA induced by SOX2 have been described and may facilitate nucleosome-compatible dyad binding (*1*, *10*, *30*). However, with the canonical SOX2 motif used here, SOX2 facing the entry-exit site is not compatible with the canonical nucleosome architecture and triggers DNA release (*31*).

Despite disruption of histone-DNA contacts at SHL −6.5, no histone rearrangements were observed after OCT4-SOX2 binding and DNA release across 14 bp (fig. S8, B and C). To verify OCT4-SOX2 binding and DNA release using an orthogonal approach, we performed deoxyribonuclease I (DNaseI) footprinting in the absence and presence of OCT4 or OCT4-SOX2. This revealed increased digestion at the nucleosomal entry-exit site (SHLs −7 to −5.5) in the presence of OCT4-SOX2, in line with DNA detachment and partial motif binding (Fig. 3C and fig. S9, A and B). Notably, OCT4 alone triggers DNA release, and OCT4 and OCT4-SOX2 also destabilize the opposite nucleosomal entry-exit site (SHLs +5.5 to +7), which is also evident in the cryo–electron microscopy (cryo-EM) map (Fig. 2 and fig. S3). The structures, footprinting profiles (Fig. 3), and thermal stability assays (fig. S9C) together support the idea that OCT4 and OCT4-SOX2 release DNA from the histones and have a global effect on the DNA structure of the nucleosome.

OCT4-SOX2 bound at SHL +6 induces minimal distortion to nucleosomal DNA

The SeEN-seq profile suggests that OCT4-SOX2 engagement depends on motif orientation (Fig. 1B). To examine this structurally, we utilized the same position but with the motif inverted across the dyad axis, i.e., SHL +6 (Fig. 4A). Doing so places the ~90° kink-inducing SOX2 motif in a dyad-proximal orientation and OCT4 closer to the entry-exit site. The SHL +6 site was enriched for OCT4-SOX2 binding in SeEN-seq, and the use of this position enabled cryo-EM structure determination of an OCT4-SOX2-NCP^{SHL+6} complex at an overall resolution of 3.5 Å (Fig. 4B and fig. S10; see fig. S10E for local resolution). The map allowed unambiguous rigid body docking of the nucleosome and of SOX2 (figs. S11 and S12). OCT4 (POUS) density was less continuous but sufficient to dock a Cα model (fig. S11D). The resulting OCT4-SOX2 interface was consistent with previously determined structures (Fig. 2, fig. S11C, and materials and methods).

In the structure, OCT4 engages its motif with the POUS domain only (Fig. 4B), and the POUHD is unable to access its motif in the observed DNA configuration (fig. S13). OCT4-

SOX2 together give rise to an extended DNA-binding surface across 11 bp that further bends the nucleosomal DNA by ~90° at the SOX2 site (SHL +5) and straightens the DNA near the OCT4 site (SHL +6), producing an L-shaped DNA arrangement (Fig. 4B). This conformation locally lifts the duplex away from the histones and DNA gyre, but, in contrast to the reversed orientation, does not fully release the DNA from histones at the entry-exit site. At the SOX2 motif with the DNA locally detached, SOX2 helices 1 and 2 widen the minor groove, and the C terminus (residues 110 to 114) wedges between the DNA and histones (SHL +5) (Fig. 4C). Despite partial engagement of an internal-nucleosomal motif, the SOX2 and OCT4 (POUS) DNA interactions and induced DNA distortions are again similar to those previously seen on free DNA (fig. S11C). Within one helical turn, on either side of OCT4-SOX2, these DNA distortions are largely absorbed into the canonical DNA trajectory of the nucleosome (Fig. 4D). The histone core architecture again shows no substantial distortion. The DNA at the opposite nucleosomal entry-exit site (SHLs −7 to −5.5) appears to be disordered in the cryo-EM map (fig. S10E and fig. S14). In both structures (Figs. 2 and 4), OCT4 binds a partial DNA motif through its POUS domain and, along with SOX2, affects the entire nucleosomal DNA structure to varying extents.

The HMG-POUS partial motif is sufficient for TF engagement and the opening of chromatin in vivo

Previous work has identified that either *Oct4* or *Sox2* alone engage reduced-complexity motifs on nucleosomes during reprogramming (*1*). A recent study has also identified weaker-scoring motifs for SOX proteins on nucleosome-sized fragments (*32*). The structures now reveal that OCT4-SOX2 partial motif engagement is utilized in both orientations on the nucleosome. This led us to test whether OCT4-POUS and SOX2-HMG domains are sufficient to engage chromatin in vivo (Fig. 5A and fig. S15A). Through in-depth analysis of existing chromatin immunoprecipitation sequencing (ChIP-seq) datasets (*33*, *34*), we found that the partial HMG-POUS motif is sufficient to drive genomic binding, although the full motif was bound more frequently (Fig. 5A). To test this experimentally, we introduced full and partial motifs at a defined genomic position in mouse embryonic stem cells (mESCs) by recombination–mediated cassette exchange (*35*) (Fig. 5B). Motifs were introduced in the SHL −6 position of the 601 sequence (see materials and methods), and *Oct4* binding was determined. This revealed significant *Oct4* enrichment at both the full and partial motifs but not in the control (Fig. 5B and fig. S15B). Thus, single motifs recapitulate genome-wide *Oct4* binding to partial motifs.

Next, we asked if binding creates open chromatin. Comparing accessibility in mESCs (*36*) revealed that full and partial motifs can both generate accessible chromatin in an Oct4-dependent manner (Fig. 5C and fig. S15, C and D). Consistently, the same effect is evident upon knockdown of *Oct4* or *Sox2*, which shows that both TFs are required for full accessibility at these loci (*37*) (fig. S15, E to H). This confirms the cooperativity observed in SeEN-seq; however, the local effect is expected to be highly context dependent, as additional proteins contribute to accessibility in vivo (*36*). We interpret our structures to depict an initial encounter complex between OCT4-SOX2 and the nucleosome en route to open chromatin. Upon nucleosome removal, OCT4 is expected to engage a full motif with its two POU domains, thereby accounting for the stronger enrichment of the full versus the partial motif. Together, genome-wide binding, single-locus insertion, and genome-wide accessibility data demonstrate that the OCT4-POUS domain is sufficient to engage and open chromatin in conjunction with Sox2. This reveals the potential for such nucleosome-compatible motifs to function as bona fide binding sites beyond the ability to initially engage closed chromatin.

Discussion

The structures illustrate binding mechanisms of OCT4-SOX2 at two positions on the nucleosome. At SHL ±6, the structures depict OCT4-SOX2 near the entry-exit sites, where both TFs cooperate to access DNA. At the SHL −6 site, where SOX2 faces the entry-exit site, OCT4-SOX2 releases the DNA duplex from the histones. In the OCT4-SOX2-NCP^{SHL+6} structure, where SOX2 faces the dyad, the nucleosomal DNA assumes an L-shaped trajectory and is not fully released from the histones (Fig. 5D). The SHL −6 structure demonstrates that partial motif recognition and DNA release are not mutually exclusive (Figs. 2 and 3), whereas the SHL +6 structure depicts how more–internally bound sites can be accommodated without full removal of nucleosomal DNA ends (Fig. 4). We consistently find only the OCT4 POUS engaged, with the POUHD motif occluded by the nucleosome architecture. The OCT4 POUS and POUHD domains could, in principal, engage the full OCT4 motif if the DNA was further unwrapped from the histones, which we do not observe in our structures (Figs. 2 and 4) or DNaseI experiments (Fig. 3C). Thus, partial motif recognition allows TFs to minimize DNA unwrapping when engaging nucleosomal sites, although partial motifs do not fully preempt nucleosome distortions. We show that partial OCT4 motifs, in conjunction with SOX2, are recognized in vivo and create open chromatin (Fig. 5).

SeEN-seq binding profiles, combined with the structural data, allow us to rationalize OCT4-SOX2 engagement throughout the nucleosome. OCT4 and OCT4-SOX2 bind best at the nucleosomal entry-exit sites, where DNA breathing is expected to facilitate access. We observe distinct exceptions from such end-binding behavior for OCT4 in particular, where narrow regions of pronounced binding are juxtaposed to nonengaged regions. To correlate these accessibility profiles to the structure, we computationally translated isolated OCT4 POUS or POUHD domains along the DNA of the unbound nucleosome and calculated predicted atomic clash scores at each position (fig. S16). A comparison of OCT4 SeEN-seq data to the POU domain nucleosome-clash scores revealed that the solvent accessibility of the POUS—but not of the POUHD—correlates with OCT4 binding (fig. S16, A and B). The solvent accessibility for OCT4 POUS is also a good predictor for OCT4-SOX2 engagement (fig. S16, C and D). The presence of SOX2 enables tight OCT4-SOX2 binding at the nucleosome ends but has a limited effect on more-internal sites (Fig. 1). At nucleosome ends, SOX2 can also drive binding at motifs where the OCT4 POUS is inward-facing and OCT4 alone does not bind (fig. S2H). Our structural and functional findings are consistent with a model where cooperative binding between OCT4 and SOX2 not only strengthens DNA binding affinity but also triggers additional DNA distortions that must be accommodated. These distortions are better tolerated at the entry-exit sites, where nucleosomal DNA is more loosely bound (*38*). Whereas partial motifs delimit the TF footprint and DNA unwrapping, the available structures (*13*) show that protein domains bound to nucleosomal DNA retain their free DNA–binding mode (Figs. 2 and 4). TF-induced DNA distortions intrinsically destabilize the nucleosome (*10*), which likely facilitates the binding of additional factors and disrupts the internucleosomal interactions of higher-order chromatin (*15*, *39*) (fig. S17).

The OCT4-SOX2 structures and the accompanying in vitro and in vivo evidence provide a framework by which TFs use nucleosomal DNA distortion and not histone rearrangement to access parts of their motif. The degree of DNA distortion imposed on the nucleosome architecture depends on the position of the motif. Our structures suggest principal recognition mechanisms for nucleosome-incompatible TFs as well as for those TFs accommodated on the nucleosome without DNA release, illustrating how TFs can read out chromatinized binding sites.

REFERENCES AND NOTES

1. A. Soufi *et al.*, *Cell* **161**, 555–568 (2015).
2. D. J. Rodda *et al.*, *J. Biol. Chem.* **280**, 24731–24737 (2005).
3. K. Takahashi, S. Yamanaka, *Cell* **126**, 663–676 (2006).
4. V. Malik *et al.*, *Nat. Commun.* **10**, 3477 (2019).

5. D. C. Ambrosetti, C. Basilico, L. Dailey, *Mol. Cell. Biol.* **17**, 6321–6329 (1997).
6. T. Kumar Mistri *et al.*, *Biophys. J.* **100**, 74a (2011).
7. C. C. Adams, J. L. Workman, *Mol. Cell. Biol.* **15**, 1405–1421 (1995).
8. L. A. Mirny, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22534–22539 (2010).
9. K. S. Zaret, J. S. Carroll, *Genes Dev.* **25**, 2227–2241 (2011).
10. F. Zhu *et al.*, *Nature* **562**, 76–81 (2018).
11. M. Fernandez Garcia *et al.*, *Mol. Cell* **75**, 921–932.e6 (2019).
12. R. K. McGinty, S. Tan, *Chem. Rev.* **115**, 2255–2273 (2015).
13. S. Matsumoto *et al.*, *Nature* **571**, 79–84 (2019).
14. L. A. Cirillo *et al.*, *EMBO J.* **17**, 244–254 (1998).
15. B. Fierz, M. G. Poirier, *Annu. Rev. Biophys.* **48**, 321–345 (2019).
16. G. Li, M. Levitus, C. Bustamante, J. Widom, *Nat. Struct. Mol. Biol.* **12**, 46–53 (2005).
17. J. Huertas, C. M. MacCarthy, H. R. Schöler, V. Cojocaru, *Biophys. J.* S0006-3495(20)30032-1 (2020).
18. A. Reményi *et al.*, *Genes Dev.* **17**, 2048–2059 (2003).
19. D. Esch *et al.*, *Nat. Cell Biol.* **15**, 295–301 (2013).
20. X. Yu, M. J. Buck, *Genome Res.* **29**, 107–115 (2019).
21. G. D. Stormo, Z. Zuo, Y. K. Chang, *Brief. Funct. Genomics* **14**, 30–38 (2015).
22. L. A. Boyer *et al.*, *Cell* **122**, 947–956 (2005).
23. X. Chen *et al.*, *Cell* **133**, 1106–1117 (2008).
24. N. Tapia *et al.*, *Sci. Rep.* **5**, 13533 (2015).
25. A. Khan *et al.*, *Nucleic Acids Res.* **46**, D260–D266 (2018).
26. P. T. Lowary, J. Widom, *J. Mol. Biol.* **276**, 19–42 (1998).
27. S. Li, E. B. Zheng, L. Zhao, S. Liu, *Cell Rep.* **28**, 2689–2703.e4 (2019).
28. D. C. Williams Jr., M. Cai, G. M. Clore, *J. Biol. Chem.* **279**, 1449–1457 (2004).
29. H. Stark, *Methods Enzymol.* **481**, 109–126 (2010).
30. P. Scaffidi, M. E. Bianchi, *J. Biol. Chem.* **276**, 47296–47302 (2001).
31. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, *Nature* **389**, 251–260 (1997).
32. M. P. Meers, D. H. Janssens, S. Henikoff, *Mol. Cell* **75**, 562–575. e5 (2019).
33. C. Chronis *et al.*, *Cell* **168**, 442–459.e20 (2017).
34. Z. Liu, W. L. Kraus, *Mol. Cell* **65**, 589–603.e9 (2017).
35. F. Lienert *et al.*, *Nat. Genet.* **43**, 1091–1097 (2011).
36. H. W. King, R. J. Klose, *eLife* **6**, e22631 (2017).
37. E. T. Friman *et al.*, *eLife* **8**, e50087 (2019).
38. M. A. Hall *et al.*, *Nat. Struct. Mol. Biol.* **16**, 124–129 (2009).
39. T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, *Nature* **436**, 138–141 (2005).

ACKNOWLEDGMENTS

Funding: N.H.T. and D.S. acknowledge support from the Novartis Research Foundation, the European Research Council under the European Union's (EU) Horizon 2020 research and innovation program grant agreement (N.H.T., 666068; D.S., 667951), and the Swiss National Science Foundation (N.H.T., Sinergia-CRSII3_160734/1 and SNF 31003A_179541; D.S., 310030B_176394). A.K.M. acknowledges the Human Frontier Science Program. A.K.M. and R.S.G. acknowledge EMBO Long-Term Fellowships. R.S.G. and L.I. acknowledge the EU Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant (R.S.G., 705354; L.I., 748760). L.I. acknowledges National Health and Medical Research Council CJ Martin Fellowship APP1148380. Z.K. acknowledges EU Horizon 2020 Marie Skłodowska-Curie ESR grant 765445. **Author contributions:** A.K.M., R.S.G., and L.I. developed SeEN-seq. L.I. and R.S.G. performed NGS library preparation, data analysis, and recombinase-mediated cassette exchange (RMCE) insertions. A.K.M. purified proteins, assembled nucleosomes, and performed EMSA and biochemical assays. L.I. and L.B. performed analysis of periodicity and genomic datasets. A.K.M. prepared cryo-EM samples. A.K.M. and S.C. performed cryo-EM and analysis. Z.K. performed fluorescence polarization assays. G.K. and R.D.B. prepared atomic models with A.K.M. and S.C. R.S.G. performed DNaseI footprinting assay and analysis. A.D.S. and A.G.-M. provided technical support for cryo-EM. G.R.P., J.W., and S.M. contributed nucleosome preparations and reagents. A.K.M., R.S.G., L.I., D.S., and N.H.T. wrote the manuscript. The research was directed by D.S. and N.H.T. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Plasmids and cell lines generated by this study are available from the Friedrich Miescher Institute for Biomedical Research under a material transfer agreement. The electron density reconstructions and final models were deposited with the EM Database (accession codes: EMD-10406, EMD-10408, and EMD-10864) and with the Protein Data Bank (PDB) (accession codes: 6T90, 6T93, and 6YOV). All other data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1460/suppl/DC1
Materials and Methods
Figs. S1 to S17
Table S1
References (*40*–*83*)
MDAR Reproducibility Checklist
Movie S1

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

22 January 2020; accepted 16 April 2020
Published online 23 April 2020
10.1126/science.abb0074

REPORT

SPECTROSCOPY

Mass-resolved electronic circular dichroism ion spectroscopy

Steven Daly[1], Frédéric Rosu[2], Valérie Gabelica[1]*

DNA and proteins are chiral: Their three-dimensional structures cannot be superimposed with their mirror images. Circular dichroism spectroscopy is widely used to characterize chiral compounds, but data interpretation is difficult in the case of mixtures. We recorded the electronic circular dichroism spectra of DNA helices separated in a mass spectrometer. We studied guanine-rich strands having various secondary structures, electrosprayed them as negative ions, irradiated them with an ultraviolet nanosecond optical parametric oscillator laser, and measured the difference in electron photodetachment efficiency between left and right circularly polarized light. The reconstructed circular dichroism ion spectra resembled those of their solution-phase counterparts, thereby allowing us to assign the DNA helical topology. The ability to measure circular dichroism directly on biomolecular ions expands the capabilities of mass spectrometry for structural analysis.

Two centuries ago, the interaction of polarized light with crystals revealed that many molecules come in two forms, nonsuperimposable with their mirror images (*1*). The activity and toxicity of natural or synthetic molecules often depend on their chirality (*2*). Identifying which mirror image is present in a sample is thus crucial. Furthermore, because biomolecules such as DNA or proteins consist of repetitive chiral subunits, they can form helical structures, as in the iconic DNA double helix (*3*, *4*). Characterizing the types of helical structures formed by biomolecules is therefore also essential for structural biology.

Mass spectrometry (MS) is a widely used analytical method with expanding impact in structural biology (*5*, *6*). MS excels at separating and quantifying complex mixtures, and additional structural characterization can be obtained using tandem MS (*7*), ion mobility spectrometry (*8*), or infrared ion spectroscopy (*9*–*11*). However, MS-based measurements are typically blind to chirality. Characterizing chiral compounds by MS currently requires a physical interaction with other chiral molecules (*12*), either by separation on chiral phases in front of the mass spectrometer or by the formation of complexes with chiral auxiliaries in the gas phase (*13*).

Another way to characterize chiral biomolecules is through interaction with chiral light. In solution, DNA and proteins are conveniently characterized by circular dichroism (CD), which measures the difference of absorption between left and right circularly polarized light. Electronic CD is particularly useful to characterize the different types of helices formed by nucleic acids (*14*) or the secondary structures formed by proteins (*15*). The structural interpretation relies on spectral databases of known structures. However, the interpretation of CD spectra becomes difficult for samples wherein multiple species or structures coexist, because the resulting spectrum is the weighted average of all contributions.

Here, we report a method for recording electronic CD spectra on DNA helices separated in a mass spectrometer. Although CD and MS have been combined before (*16*–*18*), the previous methods were based on resonant multiphoton ionization (REMPI) and were thus applicable only to volatile neutral small molecules (<200 Da). We used electrospray to produce gas-phase polyanions of intact DNA multihelices (>5000 Da) and interrogated them with circularly polarized ultraviolet light directly inside the mass spectrometer.

Our mass analyzer is a quadrupole ion trap (Paul trap) with two opposite holes (diameter 1.7 mm) in the ring electrode. We reasoned that this configuration would minimize risks of reflection of the laser beam inside the mass spectrometer, which could possibly alter the circular polarization. The electrospray source was operated in the negative mode to produce multiply charged anions. We used gentle ion transfer conditions to prevent gas-phase restructuring and thus maximize the chances of preserving the solution-phase secondary structures.

We analyzed several DNA G-quadruplex tetrahelical structures—parallel right-handed tetramer TG4T [$(dTGGGGT)_4 \bullet (NH_4^+)_3$], parallel left-handed intramolecular ZG4 [$dT(GGT)_4TG(TGG)_3TGTT \bullet (NH_4^+)_3$] (*19*), antiparallel right-handed intramolecular 5YEY [$d(GGGTTA)_2GGGTTTGGG \bullet (K^+)_2$] (*20*)—and the silver-mediated parallel G-duplex GAgG [$(dG_{11})_2 \bullet (Ag^+)_{11}$] (*21*). These structures were chosen because they have remarkably different electronic CD spectra in solution in the wavelength range 220 to 300 nm (fig. S1). Also, these structures are stabilized by spines of central cations, which maximizes the likelihood of preserving the hydrogen-bonding and stacking arrangement in the gas phase while minimizing rearrangements due to self-solvation (*21*). The preservation of the known solution-phase shapes was confirmed by the good agreement between helium momentum transfer collision integrals measured in ion mobility spectrometry and those calculated for gas-phase structures obtained by molecular dynamics at the semi-empirical level (fig. S2). For TG4T, infrared ion spectroscopy also supported the preservation of G-quartets (*22*). We can thus safely assume that these molecular systems have the same base-stacking arrangement in the gas phase as in solution.

When coupled to MS, ion spectroscopy is an action spectroscopy (and not an absorption spectroscopy): One records the effect of the laser irradiation on the ions. Another reason for selecting DNA as the first test case is that, when irradiating DNA polyanions between 220 and 300 nm, the resulting action is electron photodetachment (ePD) (*23*, *24*). This action has two advantages over fragmentation: (i) There are few product ions to quantify, which in terms of statistics will increase the chances

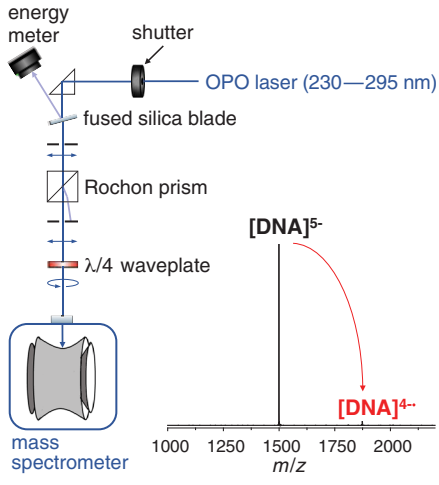


Fig. 1. Experimental setup for generating circularly polarized laser pulses. Also shown is the typical mass spectrum after isolation of D-$[(dTGGGGT)_4 \bullet (NH_4^+)_3]^{5-}$ ions and irradiation with left circularly polarized (LCP) light at 260 nm. See supplementary materials for setup details.

[1]Université de Bordeaux, Inserm & CNRS, Laboratoire Acides Nucléiques: Régulations Naturelle et Artificielle (ARNA, U1212, UMR5320), IECB, 33607 Pessac, France. [2]Université de Bordeaux, CNRS & Inserm, Institut Européen de Chimie et Biologie (IECB, UMS3033, US001), 33607 Pessac, France.
*Corresponding author. Email: v.gabelica@iecb.u-bordeaux.fr

of detecting differences in product ion yields between the two polarizations; and (ii) ePD is monophotonic (*25*), and thus the normalization for fluctuations of the laser power is simply linear for all wavelengths. We verified this linear relationship, and for all spectroscopy experiments we selected the pulse energy ranges in which only linear ePD was observed (fig. S3).

Nanosecond laser pulses were generated using a wavelength-tunable optical parametric oscillator laser (nanosecond pulses, ≤100 μJ transmitted to the trap). To generate circularly polarized laser pulses, the beam passes through an air-spaced Rochon prism, which gives pure linearly polarized light. Next, the laser passes through an achromatic broadband quarter-wave plate (Fig. 1, figs. S4 to S6, and supplementary materials). The angle of the fast axis of the quarter-wave plate relative to the polarization direction set by the Rochon prism determines the final polarization state of the laser pulse. The laser enters in the ion trap through a fused silica window in the vacuum manifold. The percentage of circular polarization was greater than 95% (fig. S7 and table S1). The ions are randomly oriented inside the Paul trap.

The CD of ions of a given mass-to-charge ratio m/z is measured by isolating the ions; irradiating them with a single laser pulse of the selected wavelength, polarization, and pulse energy; and then recording their mass spectrum (Fig. 1). Mass spectra and pulse energy are averaged for 90 s. The polarization state of the laser pulse is then changed by rotating the quarter-wave plate, and the mass spectrum and pulse energy are acquired for another 90 s. This process is repeated 10 times for each polarization. The relative electron detachment yield for each mass spectrum is calculated as

$$Y_{\mathrm{ePD}} = \frac{I_{\mathrm{ePD}}/I_{\mathrm{total}}}{\lambda E_{\mathrm{trans}}} \quad (1)$$

where I_{ePD} is the integrated intensity of the peak due to electron detachment, I_{total} is the total integrated intensity (precursor + ePD product), λ is the wavelength, and E_{trans} is the transmitted pulse energy determined from the reflected pulse energy by calibration (fig. S8). We then calculate the average value of the relative electron detachment yield for left- and right-handed circular polarizations. The CD monitored by ePD is expressed as an asymmetry factor, approximated by

$$g_{\mathrm{ePD}} = \frac{\Delta\mathrm{ePD}}{\overline{\mathrm{ePD}}} = 2 \times \left(\frac{\overline{Y_{\mathrm{ePD}}^{\mathrm{LCP}}} - \overline{Y_{\mathrm{ePD}}^{\mathrm{RCP}}}}{\overline{Y_{\mathrm{ePD}}^{\mathrm{LCP}}} + \overline{Y_{\mathrm{ePD}}^{\mathrm{RCP}}}} \right) \quad (2)$$

where LCP and RCP are left and right circularly polarized light, respectively. The gas-phase CD spectra are reconstructed by plotting the asymmetry factor g_{ePD} as a function of the wavelength. To facilitate the visual comparison, we plot the solution-phase CD spectra as $\Delta A/\bar{A}$ (where A is the absorbance), hence their shape differs from the ones traditionally displayed in molar CD.

To unambiguously prove that the gas-phase CD effect comes from the sample and not from an instrumental artifact, we performed the

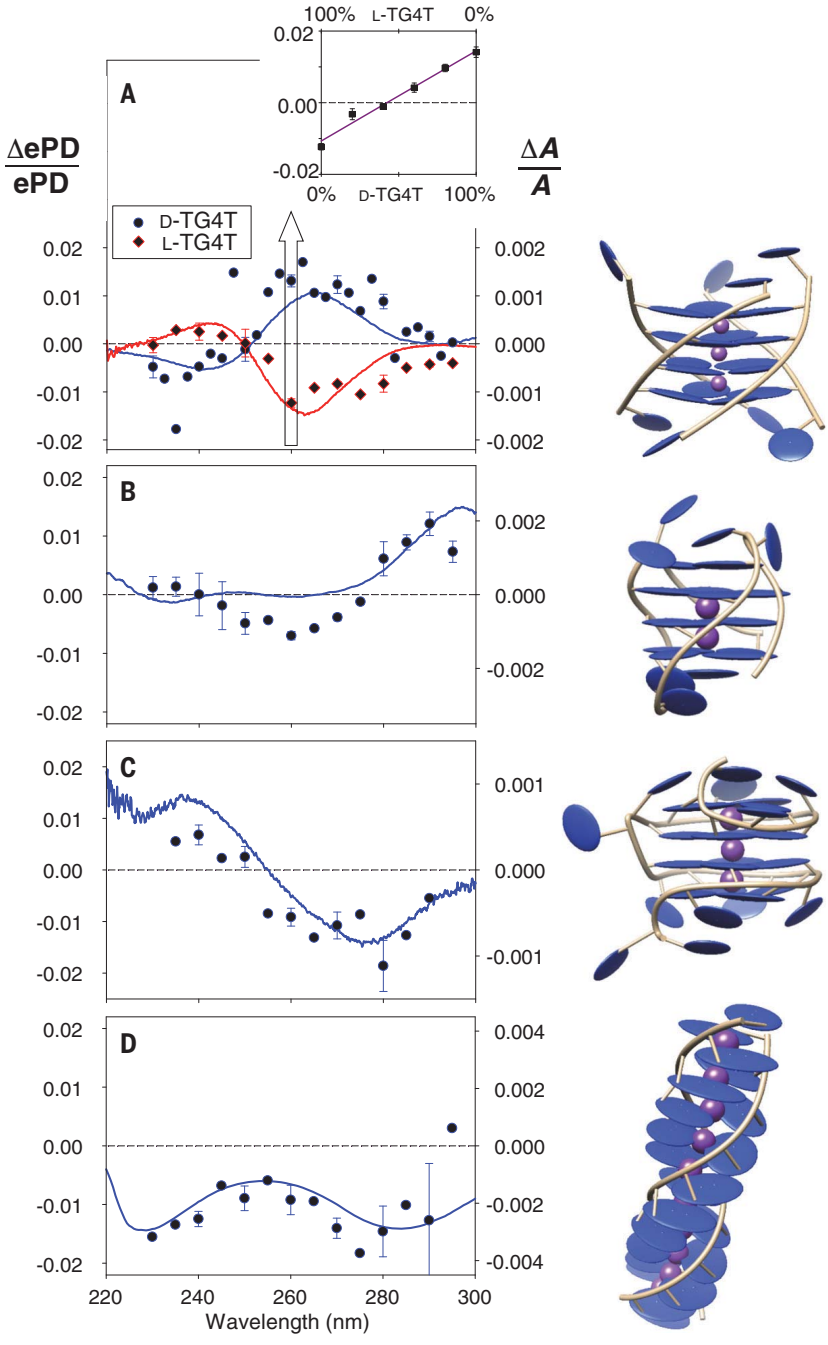


Fig. 2. Gas-phase circular dichroism spectra (ΔePD/ePD) compared to solution-phase spectra (Δ*A*/*A*). Symbols and lines denote gas-phase and solution-phase CD data, respectively. (**A**) TG4T: natural right-handed D-$[(TGGGGT)_4 \bullet (NH_4^+)_3]^{5-}$ (blue) and its mirror image L-$[(TGGGGT)_4 \bullet (NH_4^+)_3]^{5-}$ (red). Inset: Gas-phase CD measured at 260 nm for $[(TGGGGT)_4 \bullet (NH_4^+)_3]^{5-}$ ions prepared from 5 μM solutions with varying ratios of D-TG4T and L-TG4T. Error bars show the 95% confidence interval calculated from the standard error (three replicas, except at 260 nm with five replicas). (**B** to **D**) Gas-phase and solution-phase CD spectra for the antiparallel G-quadruplex $5YEY^{5-}$ (B), the left-handed G-quadruplex $ZG4^{6-}$ (C), and the G-duplex $GAgG^{5-}$ (D). The illustrations showing helicity were generated using Chimera from Protein Data Bank coordinates (entries 2O4F, 5YEY, 6GZ6, and 4JRD, respectively); the gas-phase structures are shown in fig. S2.

experiment on D-[(TGGGGT)$_4$•(NH_4^+)$_3$] formed from the natural DNA backbone (all D-sugars) and from its enantiomer (all L-sugars). The gas-phase CD signals have opposite signs (Fig. 2A). The magnitude is not exactly reversed, however. At 260 nm, we recorded a series of CD measurements by varying the relative concentration of the two enantiomers (D:L) in solution from only D to only L (Fig. 2A, inset). We obtained a straight line (r^2 = 0.975), which shows that the differences in ePD with left and right circularly polarized light are reporting on CD; this finding suggests that the technique could be suitable for quantification. The line did not pass exactly through 0 when the proportions were 50:50, indicating the presence of residual instrumental artifacts, likely because the polarization is only ~95% circular. In the future, use of an achiral internal standard could further reduce the uncertainty. Furthermore, if such imperfections vary with the wavelength, it would induce some distortion in the CD spectra that might hamper structural assignment. We thus compared the solution-phase and gas-phase CD spectra for other typical DNA helices (Fig. 2, B to D; symbols for gas-phase CD, lines for solution-phase CD). The solution- and gas-phase CD spectral shapes are similar in terms of sign and position of the maxima and minima. This suggests that the following conditions are met: (i) The base-stacking pattern existing in solution is preserved in the gas-phase ions. (ii) The gas-phase action reflects the absorption, in terms of sensitivity to the circular polarization. In other words, the electronic excited states that are responsible for the CD effect also trigger ePD. (iii) Possible distortions due to imperfections in the polarization do not preclude assigning the base-stacking arrangement on the basis of spectral shape. As a result, the gas-phase CD spectra can unambiguously discriminate between the different guanine-rich oligonucleotide structures.

The gas- and solution-phase asymmetry factors mainly differ in their magnitude, which is consistently larger in the gas phase (up to 2%) than in solution (<0.4%). This has been observed previously for small neutral molecules measured by REMPI (*17*, *26*–*28*), where magnitudes of up to 25% have been observed (*28*). High CD signals are also observed in photoelectron CD (PECD) (*29*, *30*), but in that case the measured property is the angular distribution of the photoelectrons. Here, we measure the electron detachment yield, and ePD is the result of a resonant excitation above the detachment threshold (*24*, *25*). Given that solution-phase spectra result from the CD in absorption, we conclude that the gas-phase CD effect in photodetachment yields is also due to the resonant absorption of circularly polarized light by a chiral molecule. It is thus puzzling that the CD spectral shapes are similar although the magnitudes differ.

Here we consider possible explanations. In small molecules, the solvent can affect the magnitude of CD (*31*), and thus a change in the dielectric environment from water to vacuum may change the magnitude of the CD spectrum. To test this hypothesis, we recorded solution-phase CD spectra of D-TG4T in mixtures of water up to 50% isopropanol (dielectric constant varying from 80.4 to 44.3). The mass spectra show that the G-quadruplex is intact, and no larger multimers are present (fig. S9). No change in the value of $\Delta A/A$ was observed (fig. S10), which suggests that the dielectric environment plays only a small role. We also recorded the gas-phase CD spectra for the 5-, 6-, and 7-ions of D-TG4T, reasoning that if intramolecular electric dipoles influence the electron ejection dynamics, the effect would increase with the charge state. We observed no significant change in either magnitude or shape of the CD spectra as a function of the charge state (fig. S11). We thus hypothesize that the origin of the larger gas-phase magnitude comes from the different definition of asymmetry factors. In the gas phase, absorption is revealed by electron detachment. If not all states that absorb result in electron detachment (*25*, *32*) while most states that are responsible for CD do, the denominator of the asymmetry factor is smaller in gas-phase CD action spectroscopy, and thus the asymmetry factor is larger. Because the electronic states responsible for the CD effect are most likely to be delocalized on the entire DNA helices, this may result in more efficient autodetachment after resonant excitation. Thus, although future work is needed to elucidate the origins and dynamics of ePD from DNA polyanions—for example, by measuring the photoelectron CD—the ePD action serendipitously revealed itself to be especially well suited to probe DNA higher-order structures by ion spectroscopy.

All experiments above were carried out on pure samples, and tandem MS was used to select a charge state. Next, we used mass separation to obtain the CD spectra of individual components from a mixture. Human telomeric sequences, consisting of TTAGGG repeats, can form several G-quadruplex topologies (*33*). At submillimolar potassium concentrations, complexes with 1 K^+ (two G-quartets) and 2 K^+ (three G-quartets) coexist (*34*). The 2-K^+ complex is fully formed at high KCl concentrations, but 1-K^+ complexes cannot be isolated in solution. We previously reconstructed solution-phase CD data for individual K^+ binding stoichiometries (Fig. 3A) by deconvoluting the solution-phase CD signal of a K^+ concentration series, having determined the amount of each complex in separate MS measurements. Here we applied mass-resolved CD ion spectroscopy directly on the mixture. Because the $[M+1K]^{5-}$ and $[M+2K]^{5-}$ complexes

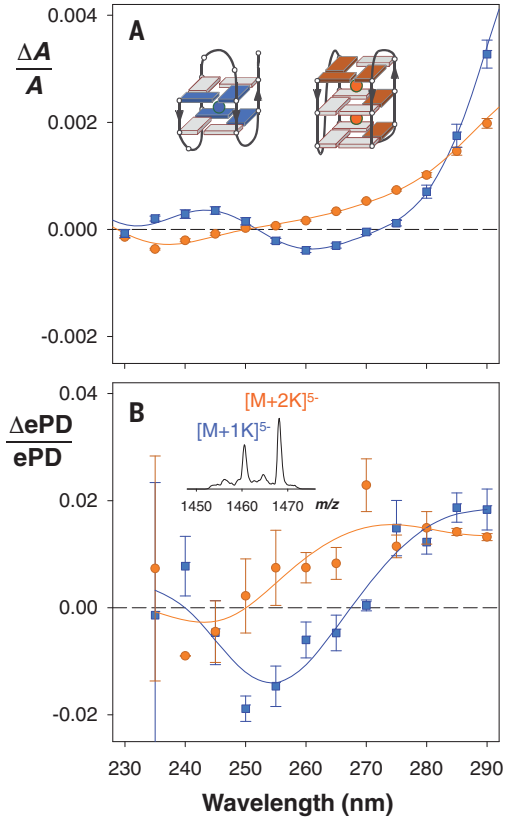


Fig. 3. Application of mass-resolved circular dichroism to a mixture of conformations adopted by the human telomeric DNA sequence dTAGGG(TTAGGG)$_3$. (**A**) Deconvoluted solution-phase CD spectra for the three-quartet (2 K^+, orange circles) and two-quartet (1 K^+, blue squares) complexes coexisting at KCl concentrations between 0 and 1 mM (*34*). The error bars are the standard uncertainties propagated from the original data; the lines were obtained by Savitzky-Golay smoothing. The insets show the stacking topologies deduced from the K^+ stoichiometry and the CD spectral shapes (light gray and colors denote anti and syn glycosidic bond angles, respectively). (**B**) Gas-phase CD spectra for the complexes $[M+1K]^{5-}$ (orange circles) and $[M+2K]^{5-}$ (blue squares) co-isolated from a solution containing 400 μM KCl. The error bars are the standard uncertainties from the replicates (listed in table S2); the lines were obtained by Savitzky-Golay smoothing. The inset shows the two co-isolated parent ions.

are close in *m/z*, they could be co-isolated (inset of Fig. 3B), and as their *m/z* differs, the ePD efficiencies of both species could be monitored in a single experiment (this is also how we envisage using internal standards in the future). Each spectrum was averaged over 180 s. Although the magnitudes of gas-phase CD signals (Fig. 3B) differ from those of their solution-phase counterparts, their position and sign allow us to infer the G-quadruplex stacking topology using the same rules as in solution (*35*): The 2-K^+ complex has positive CD signals at 290 and 260 nm and a negative signal at 240 nm (signature of coexisting homo- and heterostacking), whereas the 1-K^+ complex has positive signals at 290 and 240 nm and a negative signal at 260 nm (signature of heterostacking).

Our experiments show that it is possible to measure the CD of large DNA polyanions in the gas phase while exploiting the physical separation of components inside the mass spectrometer. The similarity between solution-phase and gas-phase spectra allowed us to distinguish the secondary structures of G-rich DNA, thereby extending the scope and capabilities of structural MS. In the future, proteins and their aggregates can be studied by the same approach, but overcoming signal-to-noise challenges will likely require using internal standards and alternative circularly polarized light sources giving access to shorter wavelengths. For positive ions and small molecules in general, cold mass tagging could provide alternative monophotonic action channels that may be less dependent on the nature of the excited state. Our demonstration of feasibility thus opens new avenues to study diverse classes of chiral molecules while leveraging the separation capabilities of contemporary mass spectrometry.

REFERENCES AND NOTES

1. F. Arago, *Mém. Class Sci. Math. Phys. Inst. Impérial France* **1**, 93–164 (1811).
2. S. W. Smith, *Toxicol. Sci.* **110**, 4–30 (2009).
3. J. D. Watson, F. H. C. Crick, *Nature* **171**, 737–738 (1953).
4. R. E. Franklin, R. G. Gosling, *Nature* **171**, 740–741 (1953).
5. I. Liko, T. M. Allison, J. T. Hopper, C. V. Robinson, *Curr. Opin. Struct. Biol.* **40**, 136–144 (2016).
6. P. Lössl, M. van de Waterbeemd, A. J. Heck, *EMBO J.* **35**, 2634–2657 (2016).
7. F. W. McLafferty, *Science* **214**, 280–287 (1981).
8. D. E. Clemmer, M. F. Jarrold, *J. Mass Spectrom.* **32**, 577–592 (1997).
9. E. Garand *et al*., *Science* **335**, 694–698 (2012).
10. J. Seo *et al*., *Nat. Chem.* **9**, 39–44 (2017).
11. M. Z. Kamrath, T. R. Rizzo, *Acc. Chem. Res.* **51**, 1487–1495 (2018).
12. H. Awad, A. El-Aneed, *Mass Spectrom. Rev.* **32**, 466–483 (2013).
13. W. A. Tao, F. C. Gozzo, R. G. Cooks, *Anal. Chem.* **73**, 1692–1698 (2001).
14. J. Kypr, I. Kejnovská, D. Renciuk, M. Vorlícková, *Nucleic Acids Res.* **37**, 1713–1725 (2009).
15. A. J. Miles, B. A. Wallace, *Chem. Soc. Rev.* **35**, 39–51 (2006).
16. R. Li, R. Sullivan, W. Al-Basheer, R. M. Pagni, R. N. Compton, *J. Chem. Phys.* **125**, 144304 (2006).
17. U. Boesl von Grafenstein, A. Bornschlegl, *ChemPhysChem* **7**, 2085–2087 (2006).
18. A. Hong *et al*., *Angew. Chem. Int. Ed.* **53**, 7805–7808 (2014).
19. W. J. Chung *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2729–2733 (2015).
20. C. Liu *et al*., *Chem. Sci.* **10**, 218–226 (2018).
21. S. M. Swasey, F. Rosu, S. M. Copp, V. Gabelica, E. G. Gwinn, *J. Phys. Chem. Lett.* **9**, 6605–6610 (2018).
22. V. Gabelica *et al*., *J. Am. Chem. Soc.* **130**, 1810–1811 (2008).
23. V. Gabelica *et al*., *Anal. Chem.* **78**, 6564–6572 (2006).
24. F. Rosu *et al*., *J. Phys. Chem. A* **116**, 5383–5391 (2012).
25. S. Daly, M. Porrini, F. Rosu, V. Gabelica, *Faraday Discuss.* **217**, 361–382 (2019).
26. P. Horsch, G. Urbasch, K. M. Weitzel, D. Kröner, *Phys. Chem. Chem. Phys.* **13**, 2378–2386 (2011).
27. H. G. Breunig *et al*., *ChemPhysChem* **10**, 1199–1202 (2009).
28. J. Lepelmeier, K. Titze, A. Kartouzian, U. Boesl, U. Heiz, *ChemPhysChem* **17**, 4052–4058 (2016).
29. L. Nahon, G. A. Garcia, C. J. Harding, E. Mikajlo, I. Powis, *J. Chem. Phys.* **125**, 114309 (2006).
30. M. H. Janssen, I. Powis, *Phys. Chem. Chem. Phys.* **16**, 856–871 (2014).
31. A. R. de Souza, V. F. Ximenes, N. H. Morgon, in *Stereochemistry and Global Connectivity: The Legacy of Ernest L. Eliel, Volume 2* (American Chemical Society, 2017), pp. 91–101.
32. V. Gabelica *et al*., *J. Am. Chem. Soc.* **129**, 4706–4713 (2007).
33. A. T. Phan, *FEBS J.* **277**, 1107–1117 (2010).
34. A. Marchand, V. Gabelica, *Nucleic Acids Res.* **44**, 10999–11012 (2016).
35. A. I. Karsisiotis *et al*., *Angew. Chem. Int. Ed.* **50**, 10645–10648 (2011).
36. S. Daly, F. Rosu, V. Gabelica, Zenodo, DOI: 10.5281/zenodo.3758200 (2020).

ACKNOWLEDGMENTS

We thank N. Khristenko for contributing to preliminary experiments. **Funding:** European Research Council, ERC-2013-CoG-616551-DNAFOLDIMS. **Author contributions:** F.R. and V.G. conceived the project; V.G. acquired the funding; F.R. and S.D. developed the methodology; S.D. and F.R. conducted the research and analyzed the data; S.D. and V.G. wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All materials are commercially available. All raw and processed data are provided (*36*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1465/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 and S2
References (*37*–*41*)

5 February 2020; accepted 27 April 2020
10.1126/science.abb1822

SOLAR PHYSICS

Meridional flow in the Sun's convection zone is a single cell in each hemisphere

Laurent Gizon[1,2,3]*, Robert H. Cameron[1], Majid Pourabdian[1,2], Zhi-Chao Liang[1], Damien Fournier[1], Aaron C. Birch[1], Chris S. Hanson[3]

The Sun's magnetic field is generated by subsurface motions of the convecting plasma. The latitude at which the magnetic field emerges through the solar surface (as sunspots) drifts toward the equator over the course of the 11-year solar cycle. We use helioseismology to infer the meridional flow (in the latitudinal and radial directions) over two solar cycles covering 1996–2019. Two data sources are used, which agree during their overlap period of 2001–2011. The time-averaged meridional flow is shown to be a single cell in each hemisphere, carrying plasma toward the equator at the base of the convection zone with a speed of ~4 meters per second at 45° latitude. Our results support the flux-transport dynamo model, which explains the drift of sunspot-emergence latitudes through the meridional flow.

Heat is transported by convective motions of the plasma in the outer 29% of the Sun (the solar convection zone). In this layer, convection interacts with rotation to drive global-scale axisymmetric flows (*1*). The longitudinal component of these flows is the solar differential rotation: The equator rotates once every 25 days, the poles once every 34 days. The latitudinal and radial components are the Sun's meridional flow. The differential rotation and the meridional flow both play a role in the solar dynamo (*2*). Differential rotation acts on latitudinal magnetic field to generate a longitudinal (toroidal) magnetic field. At the surface, the meridional flow transports the radial magnetic field from the equator toward the poles. The role of the deep meridional flow is less certain. In the class of models known as flux-transport dynamos (*3*), the meridional flow near the base of the convection zone is assumed to be equatorward and to transport the toroidal magnetic field to match the drift of 2° to 3° per year in the latitudes at which sunspots appear (*2*). Thus, these models provide testable predictions about the amplitude and sign of the deep meridional flow.

Testing these predictions is challenging. The geometry of the meridional flow is difficult to compute theoretically from first principles, as it results from a small imbalance between two large terms (*4*). Mass conservation implies that the plasma is carried around closed loops (cells). A poleward flow at the surface must return equatorward at some depth. There may be additional closed cells stacked on top of each other in the radial direction. Both one-cell and two-cell meridional flow geometries have been proposed on the basis of theory and numerical simulations (*5*). Observationally, the meridional flow can be constrained using helioseismology (*6*, *7*). This technique relies on measurements of the times taken by solar acoustic waves to travel between points on the surface through the interior. Suitable data are available for 1996–2011 from the Michelson Doppler Imager (MDI) on the Solar and Heliospheric Observatory (SOHO) spacecraft and for 2010 onward from the Helioseismic and Magnetic Imager (HMI) on the Solar Dynamics Observatory (SDO) spacecraft. Studies of the HMI data have reached differing conclusions on the geometry of the meridional flow: either one or two cells in the radial direction (*8*–*11*). The MDI data and the HMI data also show different flow structures (*12*). These differences may result from instrumental systematic errors, the calibration of the observations, and/or different assumptions made during the data analysis.

To confirm the validity of a helioseismic inference, it is necessary to compare results from two independent datasets that have an extended overlap period. For example, the Sun's internal rotation has been validated (*13*) by comparing observations from MDI to those of the ground stations operated by the Global Oscillation Network Group (GONG) (*14*).

Using multiple datasets, we studied the structure and time variability of the meridional flow in the convection zone. We used maps of the line-of-sight velocity at the Sun's surface (dopplergrams) at reduced spatial resolution, which provide information about sound waves propagating in the solar interior with spherical harmonic degrees up to 300. The reduced-resolution data are known to be less prone to instrumental errors (*15*). We considered all three main datasets: MDI, HMI, and GONG. The MDI observations consist of dopplergrams with 192 pixels by 192 pixels per frame for the period 1 May 1996 to 11 April 2011 (*16*). After April 2003, the SOHO spacecraft was rotated 180° every 3 months; we only used the observations when MDI was in the orientation it had before 2003 to ensure consistency of the data (*17*). The HMI data (*18*) that we used cover the period from 1 May 2011 to 30 April 2019; they were processed by the HMI team to have a format (204 pixels by 204 pixels) similar to that of the MDI data. We constructed the third dataset from the merged and calibrated high-resolution GONG dopplergrams (*19*) for 1 August 2001 to 30 April 2019 (839 pixels by 839 pixels) by applying Gaussian smoothing and downsampling to 200 pixels by 200 pixels.

We produced time series of MDI, HMI, and GONG images—remapped onto a heliographic coordinate system (scale of 0.6° per pixel)—by tracking areas on the solar surface at the Carrington rotation rate of 456.03 nHz (*20*). The MDI data were corrected to account for a misalignment of the instrument with respect to the spacecraft, corresponding to a 0.20° error in the solar P angle (one of the two angles describing the direction of the Sun's rotation axis with respect to the plane of the sky). This error was determined using HMI images as reference during the 2010 overlap period (*21*). The orientation of the HMI images is known to within 0.01°, based on the analysis of the transit of Venus from 5 to 6 June 2012 (*22*). The MDI, HMI, and GONG images were all corrected for a 0.08° difference between the inclination of the solar rotation axis to the ecliptic and the value of 7.25° measured by Carrington in 1863 (*20*).

The data were analyzed using time-distance helioseismology (*23*). Solar subsurface flows affect the time it takes for acoustic waves to travel between two points on the surface, A and B. The travel times were measured in both directions by cross-correlating the Doppler velocity data at A and those at B. The difference between the two travel times is sensitive to flows near the ray path that connects A and B through the interior. To determine the meridional flow, we considered points separated in latitude, using a quadrant geometry with arcs of 30° [(*12*) and fig. S1]. Points in locations of strong magnetic fields were excluded from the averages (*17*). By fitting a one-parameter model (*24*) to the cross-correlation functions computed daily, we measured the south-north travel-time differences $\tau(\Delta, \lambda)$, where $6° \le \Delta \le 42°$ is the angular distance between the arcs and $-54° \le \lambda \le 54°$ is the latitude of the midpoint between the arcs. We applied a center-to-limb correction using travel-time differences measured in the east-west direction (*25*) (fig. S2).

Figure 1, A to C, shows the measured travel times averaged over three ranges of travel distances at low latitudes in the north and south. The travel times are averaged in bins of 3 years

[1]Max-Planck-Institut für Sonnensystemforschung, 37077 Göttingen, Germany. [2]Institut für Astrophysik, Georg-August-Universität Göttingen, 37077 Göttingen, Germany. [3]Center for Space Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates.
*Corresponding author. Email: gizon@mps.mpg.de

starting from 1 May 1996 to aid in the comparison of the three datasets. For the period before 1 May 2002, we only show MDI data. For the periods thereafter, we included in the time averages only those days when travel times are available for two instruments: MDI (original orientation) and GONG until 30 April 2011 and HMI and GONG thereafter. The signs of the measured travel times are consistent with a poleward meridional flow near the surface. The time variations of the travel times are caused by surface inflows into magnetic regions (*26*) and therefore are related to the sunspot number (Fig. 1D). The MDI and GONG observations during the period from 2002 to 2011 differ by ~0.1 s (Fig. 1, A to C, and fig. S3), which is much smaller than the 1σ uncertainties caused by the random excitation of the acoustic waves. Taking noise correlations into account, MDI and GONG travel times are consistent for distances ≤30° [(*20*) and table S1]. However, the HMI travel times differ from the GONG travel times, with smaller values for HMI times in the northern hemisphere. We have been unable to identify the source of this discrepancy, so we do not use the HMI data for the rest of the analysis. We also considered data from the GONG network from 1 May 1996 to 31 May 2001, when it operated with lower-resolution cameras. The corresponding travel times are consistent with the MDI data for the same period but noisier [(*20*) and fig. S4]. We therefore combine the MDI data from 1 May 1996 to 30 April 2003 (before the first change in orientation of the SOHO spacecraft) and the GONG data from 1 May 2003 to 30 April 2019.

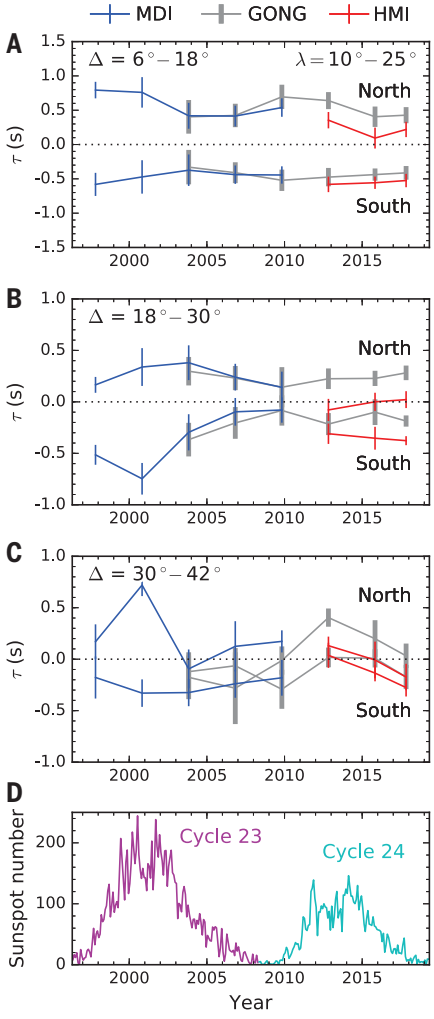


Fig. 1. Averages of the travel-time differences caused by the meridional flow. (**A** to **C**) The travel times were averaged in 3-year periods and over latitudes 10° ≤ λ ≤ 25° (North) and −25° ≤ λ ≤ −10° (South). Between 1 May 2002 and 30 April 2011, only the days when travel times for both MDI (original orientation) and GONG are available were included in the time averages. From 1 May 2011 onward, only the days when travel times for both HMI and GONG are available were included in the averages. Travel times were further averaged over angular distances from 6° to 18° (A), from 18° to 30° (B), and from 30° to 42° (C). The error bars indicate ±1σ. (**D**) Monthly sunspot number versus time, displaying solar cycles 23 and 24.

The south-north travel-time differences τ are linearly related to the meridional flow through functions of position known as sensitivity kernels. We computed the kernels using a finite element solver in the frequency domain (*27*). Denoting the colatitude by θ = 90° − λ, we represented the radial and colatitudinal components of the meridional flow, U_r and U_θ, as linear combinations of basis functions (*20*). The coefficients in these expansions are the parameters describing the flow, which were determined by inverting the travel times. We solved the linear inverse problem under the physical constraints of mass conservation and that the flow does not cross the convection zone boundaries (*20*). Inversions were validated using synthetic data (*20*). Eleven years of data allow us to distinguish between one- and two-cell flow profiles, with a noise level of ~1.5 m s^{-1} at the base of the convection zone. For 3 years of data, the noise increases to 2.5 m s^{-1}. These uncertainties are consistent with previous estimates (*28*).

Figure 2A shows the inferred U_θ in the convection zone, averaged over each solar cycle. The stream functions in Fig. 2B show that the flow takes the form of a single cell in each hemisphere. The flow is poleward at the surface (Fig. 3C) and equatorward at the base of the convection zone (Fig. 3A). The flow profile at the base of the convection zone approximately follows $U_\theta = U_b \sin 2\theta$, with $U_b = 4.8 \pm 1.0$ m s^{-1} for cycle 23 and $U_b = 3.6 \pm 1.0$ m s^{-1} for cycle 24 (the cycles are identified in Fig. 1D). The flow switches sign near $0.79R_\odot$, where $R_\odot$ is the solar radius (Fig. 3B). This is consistent with previous inversions that used the constraint of mass conservation (*9*, *11*).

The 3-year averages are shown in Fig. 2, C and D, and Fig. 3, D and E. The noise is higher than that for the 11-year averages. Except for the period between the two cycles, a single cell is evident in each hemisphere. At the

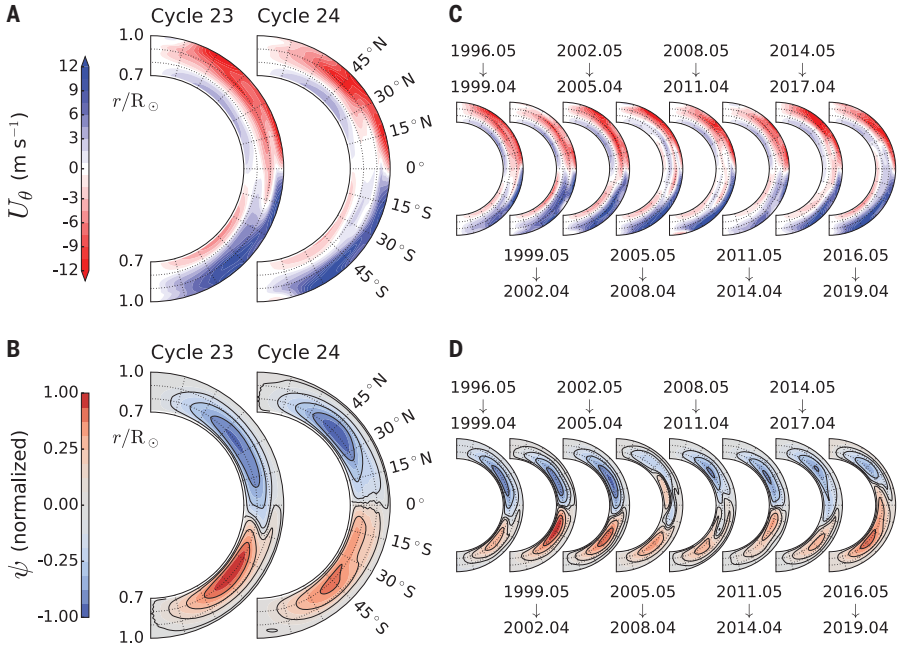


Fig. 2. Inferred meridional flow in the solar convection zone. (**A**) Flow U_θ for solar cycles 23 and 24 (U_r is shown in fig. S5). (**B**) Stream functions ψ for each solar cycle, defined by $\rho \boldsymbol{U} = \nabla \times [\hat{\boldsymbol{\phi}}\psi/(r\sin\theta)]$, where ρ is the density and $\hat{\boldsymbol{\phi}}$ is the unit vector in the longitudinal direction. The stream functions are normalized to their maximum absolute value. (**C** and **D**) Temporal variations with a cadence of 3 years.

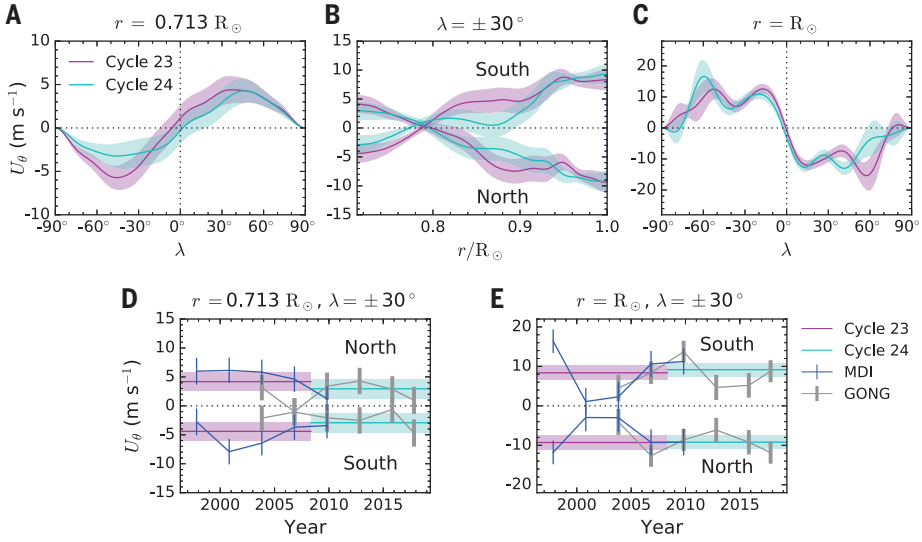


Fig. 3. Cuts through the solutions of the meridional flow. (**A**) Latitudinal dependence of U_θ at the base of the convection zone. The purple and light blue solid curves show U_θ for solar cycles 23 and 24, respectively, with the shaded regions indicating ±1σ. (**B**) Radial dependence of U_θ at latitudes of ±30° for each cycle. (**C**) Latitudinal dependence of U_θ at the surface for each cycle. (**D**) Time dependence of U_θ at latitudes of ±30° at the base of the convection zone. The purple and light blue horizontal lines show the averages over cycles 23 and 24 (Fig. 1D), respectively, with the shaded regions indicating ±1σ. The dark blue and gray lines connect the 3-year averages from MDI and GONG, respectively. Error bars indicate ±1σ. (**E**) Time dependence of U_θ at latitudes of ±30° at the surface.

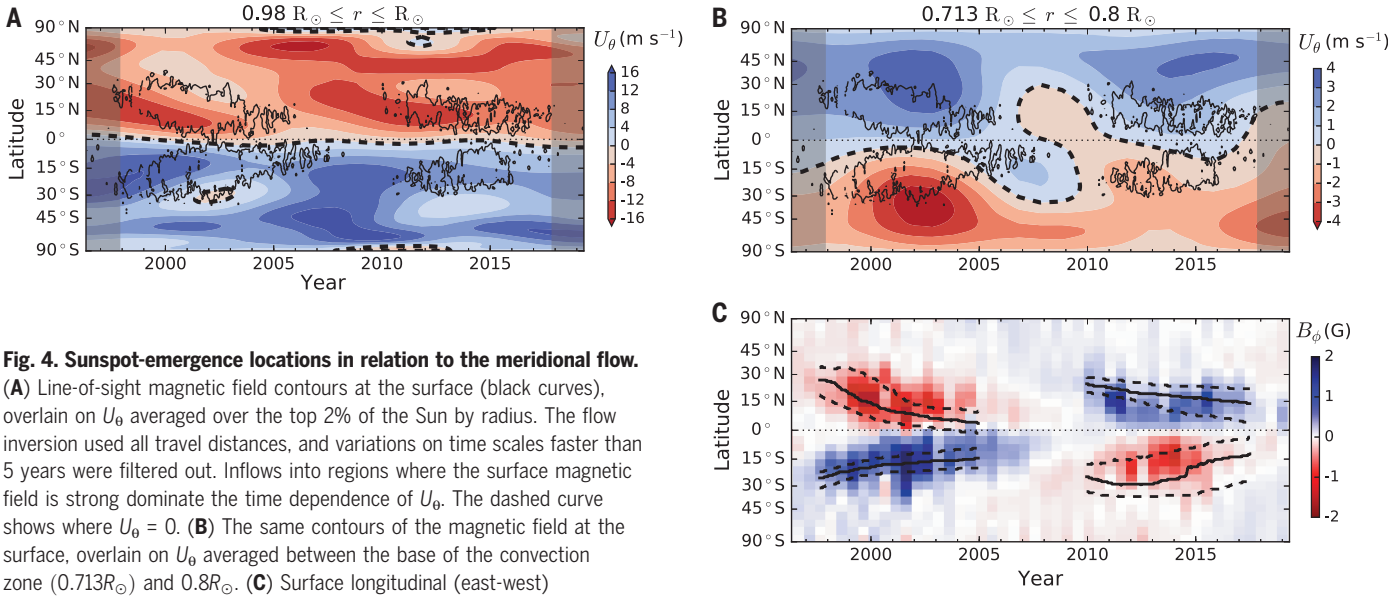


Fig. 4. Sunspot-emergence locations in relation to the meridional flow. (**A**) Line-of-sight magnetic field contours at the surface (black curves), overlain on U_θ averaged over the top 2% of the Sun by radius. The flow inversion used all travel distances, and variations on time scales faster than 5 years were filtered out. Inflows into regions where the surface magnetic field is strong dominate the time dependence of U_θ. The dashed curve shows where U_θ = 0. (**B**) The same contours of the magnetic field at the surface, overlain on U_θ averaged between the base of the convection zone ($0.713R_\odot$) and $0.8R_\odot$. (**C**) Surface longitudinal (east-west) magnetic field B_φ from Wilcox Solar Observatory observations (*30*) and, superimposed with black lines, latitudes at which the subsurface longitudinal field from the flux transport model—based on the measured U_θ—is maximum. The solid lines are the median and the dashed lines are the 16th and 84th percentiles, obtained from 300 realizations of the flow.

surface, the variations of U_θ at 30° latitude are substantial and anticorrelated with the sunspot number (Fig. 3E). Figure 4A shows the surface U_θ smoothed in time with a low-pass filter (5 years) and its relation to the latitudes where sunspots emerge. The changes with time and latitude can be understood as local inflows into magnetic regions, which are likely driven by horizontal pressure gradients caused by the surface magnetic field (*26*). In the middle of the convection zone (fig. S6), the meridional flow averaged over cycle 24 is poleward in the north and very weak in the south. The 3-year averages show a decrease of the amplitude of U_θ from about 2004 during the decaying phase of solar cycle 23. This decrease is seen in both the MDI and GONG data.

The agreement between the MDI (original orientation) and GONG data is better for angular distances of $\Delta \le 30°$ than for $\Delta > 30°$ (Fig. 1). Travel distances ≤30° are capable of distinguishing between single- and double-cell models (*12*), so we also performed inversions for $\Delta \le 30°$ only. The results (figs. S7 to S9) confirm the single-cell solution for each cycle for the MDI and GONG data. Additionally, the inversions for MDI and GONG restricted to the days in common and for separation distances $\Delta \le 30°$ are almost indistinguishable (fig. S10).

In flux-transport dynamo models, the amount of longitudinal magnetic field created in the convection zone is determined by the distribution of the radial magnetic field at the surface (*29*). The longitudinal magnetic field is then transported by U_θ near the base of the convection zone to produce the drift of latitudes at which sunspots emerge. To test this hypothesis, we used the one-dimensional mean-field equation governing the evolution of the longitudinal magnetic field (*20*), with U_θ from each solar cycle averaged from the bottom of the convection zone to $0.8R_\odot$. Using a cycle average for U_θ is appropriate because the temporal variations of U_θ deep in the convection zone (Fig. 4B) are consistent with noise (Fig. 3D). The latitude of the peak subsurface longitudinal magnetic field from the model shows an equatorial propagation that is consistent with the equatorward drift of locations of strong surface field (Fig. 4C). The one-cell meridional flow in

each hemisphere that we observed is thus consistent with the equatorial migration of the sunspots under a simple flux-transport model.

REFERENCES AND NOTES

1. G. Rüdiger, *Differential Rotation and Stellar Convection* (Gordon and Breach, 1989).
2. P. Charbonneau, *Solar and Stellar Dynamos* (Springer, 2013).
3. Y.-M. Wang, N. R. Sheeley, A. G. Nash, *Astrophys. J.* **383**, 431–442 (1991).
4. L. L. Kitchatinov, *Geomagn. Aeron.* **56**, 945–951 (2016).
5. N. A. Featherstone, M. S. Miesch, *Astrophys. J.* **804**, 67 (2015).
6. P. M. Giles, T. L. Duvall Jr., P. H. Scherrer, R. S. Bogart, *Nature* **390**, 52–54 (1997).
7. V. G. A. Böning, "Inferences of the deep solar meridional flow," thesis, Albert-Ludwigs-Universität Freiburg (2017); https://doi.org/10.6094/UNIFR/13652.
8. J. Zhao, R. S. Bogart, A. G. Kosovichev, T. L. Duvall Jr., T. Hartlep, *Astrophys. J.* **774**, L29 (2013).
9. S. P. Rajaguru, H. M. Antia, *Astrophys. J.* **813**, 114 (2015).
10. R. Chen, J. Zhao, *Astrophys. J.* **849**, 144 (2017).
11. K. Mandal, S. M. Hanasoge, S. P. Rajaguru, H. M. Antia, *Astrophys. J.* **863**, 39 (2018).
12. Z.-C. Liang, L. Gizon, A. C. Birch, T. L. Duvall Jr., S. P. Rajaguru, *Astron. Astrophys.* **619**, A99 (2018).
13. J. Schou *et al*., *Astrophys. J.* **567**, 1234–1249 (2002).
14. J. W. Harvey *et al*., *Science* **272**, 1284–1286 (1996).
15. P. M. Giles, "Time-distance measurements of large-scale flows in the solar convection zone," thesis, Stanford University (2000).
16. A. G. Kosovichev *et al*., *Sol. Phys.* **170**, 43–61 (1997).
17. Z.-C. Liang, D.-Y. Chou, *Astrophys. J.* **805**, 165 (2015).
18. P. H. Scherrer *et al*., *Sol. Phys.* **275**, 207–227 (2012).
19. J. Harvey, R. Tucker, L. Britanik, "High resolution upgrade of the GONG instruments" in *Structure and Dynamics of the Interior of the Sun and Sun-like Stars*, S. Korzennik, Ed. (ESA Special Publication, vol. 418, 1998), pp. 209–211.
20. Materials and methods are available as supplementary materials.
21. Z.-C. Liang, A. C. Birch, T. L. Duvall Jr., L. Gizon, J. Schou, *Astron. Astrophys.* **601**, A46 (2017).
22. S. Couvidat *et al*., *Sol. Phys.* **291**, 1887–1938 (2016).
23. T. L. Duvall Jr., S. M. Jeffferies, J. W. Harvey, M. A. Pomerantz, *Nature* **362**, 430–432 (1993).
24. L. Gizon, A. C. Birch, *Astrophys. J.* **614**, 472–489 (2004).
25. J. Zhao, K. Nagashima, R. S. Bogart, A. G. Kosovichev, T. L. Duvall Jr., *Astrophys. J.* **749**, L5 (2012).
26. L. Gizon, M. Rempel, *Sol. Phys.* **251**, 241–250 (2008).
27. L. Gizon *et al*., *Astron. Astrophys.* **600**, A35 (2017).
28. D. C. Braun, A. C. Birch, *Astrophys. J.* **689**, L161–L165 (2008).
29. R. Cameron, M. Schüssler, *Science* **347**, 1333–1335 (2015).
30. R. H. Cameron, T. L. Duvall Jr., M. Schüssler, H. Schunker, *Astron. Astrophys.* **609**, A56 (2018).

ACKNOWLEDGMENTS

We thank R. Burston for help with the netDRMS data management system. We thank H. Barucq and the Magique 3D team at Inria Bordeaux Sud-Ouest and Université de Pau et des Pays de l'Adour (UPPA) for making the finite element wave solver Montjoie available to us. SOHO is a project of international cooperation between the European Space Agency (ESA) and NASA. The MDI data are courtesy of the SOHO/MDI consortium. This work utilizes GONG and SOLIS (Synoptic Optical Long-term Investigations of the Sun) data obtained by the National Solar Observatory (NSO) Integrated Synoptic Program (NISP), managed by the NSO, the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation. The HMI data used are courtesy of NASA (SDO) and the HMI science team. The sunspot numbers are from WDC-SILSO (World Data Center Sunspot Index and Long-term Solar Observations), Royal Observatory of Belgium, Brussels. **Funding:** The data were processed at the German Data Center for SDO, funded by the German Aerospace Center under grant DLR 50OL1701. L.G. acknowledges support from European Research Council Synergy grant WHOLE SUN 810218. The Center for Space Science at New York University Abu Dhabi (NYUAD) is funded by the NYUAD Institute under grant G1502. M.P. was funded in part by the International Max Planck Research School (IMPRS) for Solar System Science at the University of Göttingen. **Author contributions:** L.G. and Z.-C.L. designed the research. Z.-C.L. measured the travel times. D.F. and C.S.H. computed the sensitivity kernels for flows. M.P. and D.F. inverted the travel times. R.H.C. developed the flux-transport dynamo model. L.G. and R.H.C. wrote the draft paper. All authors discussed the results and contributed to the final version of the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The MDI data were taken from the Joint Science Operations Center (JSOC) export tool at http://jsoc.stanford.edu/ajax/lookdata.html?ds=mdi.vw_v for the period May 1996 to April 2011. The HMI data were taken from the JSOC export tool at http://jsoc.stanford.edu/ajax/lookdata.html?ds=hmi.vw_v_45s for the period May 2010 to April 2019. The GONG merged velocity data ("mrvzi") were taken from https://gong2.nso.edu/archive/patch.pl?menutype=g for GONG dates 960501 to 190501. Our analysis software and the data necessary to reproduce our results are available on the Open Research Data Repository of the Max Planck Society (Edmond) at https://edmond.mpdl.mpg.de/imeji/collection/0MJjNql7GfpEl5Mb.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1469/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 and S2
References (*31*–*43*)

4 October 2019; accepted 4 May 2020
10.1126/science.aaz7119

PLASMONIC MATERIALS

Micelle-directed chiral seeded growth on anisotropic gold nanocrystals

Guillermo González-Rubio[1]*, Jesús Mosquera[1]*, Vished Kumar[1], Adrián Pedrazo-Tardajos[2], Pablo Llombart[3,4], Diego M. Solís[5], Ivan Lobato[2], Eva G. Noya[4], Andrés Guerrero-Martínez[3], José M. Taboada[6], Fernando Obelleiro[7], Luis G. MacDowell[3], Sara Bals[2]†, Luis M. Liz-Marzán[1,8,9]†

Surfactant-assisted seeded growth of metal nanoparticles (NPs) can be engineered to produce anisotropic gold nanocrystals with high chiroptical activity through the templating effect of chiral micelles formed in the presence of dissymmetric cosurfactants. Mixed micelles adsorb on gold nanorods, forming quasihelical patterns that direct seeded growth into NPs with pronounced morphological and optical handedness. Sharp chiral wrinkles lead to chiral plasmon modes with high dissymmetry factors (~0.20). Through variation of the dimensions of chiral wrinkles, the chiroptical properties can be tuned within the visible and near-infrared electromagnetic spectrum. The micelle-directed mechanism allows extension to other systems, such as the seeded growth of chiral platinum shells on gold nanorods. This approach provides a reproducible, simple, and scalable method toward the fabrication of NPs with high chiral optical activity.

Chirality is a key signature of nature that can be found across length scales, from subatomic particles, through molecules and biological systems, to galaxies (*1*–*3*). Imparting handedness to selected materials may provide important advantages in terms of their interaction with living organisms, allowing the development of enantioselective catalysts or devices with spin selectivity in electron transport (*4*–*6*). For example, the demonstration of chiral plasmon modes in noble metal nanocrystals (NCs) has drawn interest in the field of metamaterials and the design of enantioselective sensing probes (*7*–*9*). The fabrication of plasmonic nanostructures has thus become an active field of research (*10*–*16*). However, growing crystalline noble metals with dissymmetric morphology is challenging (*10*–*12*). The formation of chiral plasmonic nanomaterials is usually achieved by assembling achiral plasmonic gold and silver nanoparticles (NPs) with molecular templates such as DNA, proteins, or polymeric fibers (*13*–*15*). Although lithographic approaches have been successfully applied to fabricate chiral gold NPs, their scalability remains limited (*9*, *16*).

Advances in the colloidal synthesis of noble metal NPs achieved during the past decades suggest realistic prospects for the production of plasmonic NCs with diverse morphologies, including chiral ones (*17*–*19*). The presence of chiral amino acids during growth can guide the formation of NCs with distinct handedness. The most prominent example of this is the synthesis of helicoidal nanostructures assisted by cysteine and glutathione. The enantioselective interaction of amino acids with chiral geometrical elements naturally appearing at particular NC facets, has been claimed to induce the observed shape evolution into twisted geometries with high dissymmetry factors (*11*, *20*).

In addition to chiral additives, the synthesis of colloidal chiral NPs requires surface ligands that prevent undesired aggregation (*11*, *21*). However, most of these capping agents also play a role in the growth process and the resulting NC morphology (*17*, *22*). Among the wide variety of ligands used for the synthesis of noble metal NCs, quaternary alkylammonium halide (CTAX, where X = Cl or Br) surfactants are some of the most extensively investigated. Through the adsorption of CTAX micellar aggregates on certain crystallographic facets, a wide variety of NP morphologies have been obtained, including nanorods, nanotriangles, and Platonic geometries (*17*–*19*, *22*–*24*).

Recently, the addition of cosurfactants such as aromatic molecules, fatty acids, or long-chain alkyl alcohols has been shown to further improve the quality of colloidal gold NPs, and of nanorods in particular (*25*–*27*). The ability of cosurfactants to intercalate within CTAX surfactant aggregates has been proposed to increase the rigidity of the micelles, induce phase transitions, or both. The impact of cosurfactants on the aggregation behavior of surfactant molecules can also be extended to the formation of chiral micelles. For example, 1,1′-bi(2-naphthol) [BINOL], a cosurfactant with axial chirality (i.e., atropisomerism), can induce the formation of chiroptically active, giant, worm-like micelles (*28*).

Considering the role of surfactants on the synthesis of gold NCs and the possibility of producing chiral micelles in the presence of BINOL-like cosurfactants, we envisioned the surfactant-assisted seeded growth of chiral gold NCs. We hypothesized that elongated rod-like particles would have the ideal shape to guide the adsorption of twisted, worm-like micelles on the gold surface into helix-like chiral structures. To investigate both hypotheses, we first performed a computational analysis on an ideal system comprising cetyltrimethyl ammonium chloride (CTAC) and BINOL. Molecular dynamics (MD) simulations revealed that BINOL molecules induced the assembly of surfactants into chiral, worm-like aggregates. Such elongated micelles tend to coil around gold nanorods, which can thus be considered as templates for the seeded growth of anisotropic NPs with chiral features.

Although experimental evidence of the proposed idea was obtained using BINOL as a cosurfactant, the results of chiral growth were largely improved by using its derivative 1,1′-binaphthyl-2,2′-diamine (BINAMINE). The growth of plasmonic gold nanorods patterned with a complex chiral surface was indeed confirmed by high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) tomography. By tuning the dimensions of the final nanostructures, high anisotropy factors (g-factor ~0.20) were achieved within a wide wavelength range (from 500 to beyond 1350 nm). Theoretical modeling of the optical properties highlights the importance of growing well-defined chiral wrinkles to obtain intense circular dichroism (CD) responses.

MD simulations confirmed that in the presence of BINOL as a cosurfactant, CTAC can form giant cylindrical micelles that span across the simulation boxes (i.e., up to 60 nm in length; Fig. 1A and fig. S1). Regarding their chiral nature, we observed that unit vectors between the local micellar center and the principal micellar axis were distributed helically, with the handedness being dictated by the choice of cosurfactant stereoisomer. This effect can be shown by measuring the average angle formed between pairs of vectors (fig. S2). An achiral helix does not have any tendency to coil into a particular direction, so the angle formed between such vectors averages out to zero.

In our simulations, the average angles not only were finite but also appeared sinusoidally correlated and displayed a well-defined

[1]CIC biomaGUNE, Basque Research and Technology Alliance (BRTA), 20014 Donostia-San Sebastián, Spain. [2]Electron Microscopy for Materials Research (EMAT), University of Antwerp, 2020 Antwerp, Belgium. [3]Departamento de Química Física, Universidad Complutense de Madrid, 28040 Madrid, Spain. [4]Instituto de Química Física Rocasolano, CSIC, E-28006 Madrid, Spain. [5]Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. [6]Departamento de Tecnología de los Computadores y de las Comunicaciones, University of Extremadura, 10003 Cáceres, Spain. [7]Departamento de Teoría de la Señal y Comunicaciones, University of Vigo, 36310 Vigo, Spain. [8]Ikerbasque, Basque Foundation for Science, 48013 Bilbao, Spain. [9]CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), 20014 Donostia-San Sebastián, Spain.
*These authors contributed equally to this work.
†Corresponding author. Email: sara.bals@uantwerpen.be (S.B.); llizmarzan@cicbiomagune.es (L.M.L.-M.)

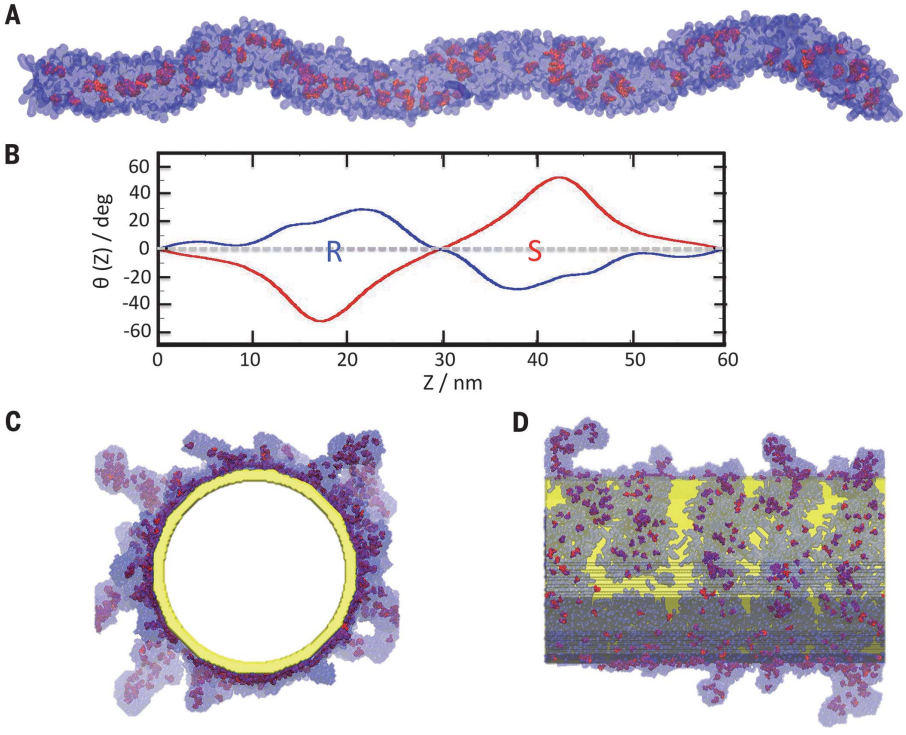


Fig. 1. MD simulations of BINOL-CTAC micelles and their assembly on gold nanorods. (**A**) MD Simulations for (*R*)-BINOL-CTAC in solution. The aggregation of CTA^+ (blue) in the presence of (*R*)-BINOL (red) results in chiral, worm-like aggregates spanning the whole simulation box 60 nm in length. (**B**) The shape of the micellar coil is described by a set of vectors along the principal micellar axis. The average angle between vectors is plotted as a function of their vertical distance for 60-nm-long rod-like micelles. The presence of (*R*)- or (*S*)-cosurfactant enantiomers in the micelles leads to coiling in opposite directions, as revealed by the sign of the average angle. (**C** and **D**) MD simulations for (*R*)-BINOL-CTAC adsorbed onto a gold nanorod in axial and lateral views, respectively.

handedness (Fig. 1B). Because of the marked atropisomerism of BINOL, we can assume that it could induce helical chirality onto the surfactant aggregates, likely by a "sergeants and soldiers" mechanism in which a chiral molecule directs many achiral molecules or atoms (*29*, *30*). In this case, the adsorbed chiral micelles would assist chirality transfer during further deposition of metal atoms onto the nanorod surface.

We tested this effect by simulating a gold nanorod immersed in an aqueous mixed solution of CTAC and BINOL. Visual inspection of the simulation results revealed that CTAC and BINOL adsorbed into worm-like micelles coiled around the gold nanorod surface (Fig. 1, C and D). Although no preferred orientation of the assembly was observed by the use of different BINOL enantiomers in this case, in the absence of cosurfactant, the coil-like morphology was largely lost. Instead, small aggregates and a high fraction of individual surfactant molecules adsorbed on the gold surface (fig. S3). The simulations indicated that chiral worm-like micelles likely formed first and were subsequently adsorbed onto the gold nanorod. Thus, a complex cooperative chiral transfer mechanism would convey the atropisomerism from BINOL into the giant micelles and ultimately to gold nanorods.

On the basis of the simulation results, we implemented experimentally the growth of chiral nanorods. (*R*)-BINOL was solubilized by surfactant micelles in aqueous solution at an optimum molar ratio of 40 and used to template the growth of preformed gold nanorods (130 ± 10 nm long, 29 ± 3 nm thick; for details, see Fig. 2A and the supplementary materials; note that (*R*)-BINOL is insoluble in water in the absence of surfactant). Preliminary characterization by CD spectroscopy revealed the appearance of moderate, chiral-localized surface plasmon resonance bands (g-factor ~0.002; figs. S4 and S5). HAADF-STEM analysis of the NPs grown in the (*R*)-BINOL and surfactant mixture revealed a complex morphology characterized by the presence of wrinkles on the nanorod surface (Fig. 2B and fig. S4). Previous theoretical and experimental studies demonstrated that intense CD responses are exhibited by NPs with distinct chiral distortions, stellated-like morphologies, or both (*9*, *11*, *31*).

Encouraged by the appearance of plasmonic bands in the CD spectra, our efforts focused on improving chiral seeded growth. We hypothesized that the use of stronger stabilizing agents would help to direct the growth of more structured chiral wrinkles (*32*). We investigated (*R*)-BINAMINE, an analog of (*R*)-BINOL in which the hydroxyl functionalities are replaced by amine moieties. Although the interaction of amines with gold surfaces is moderately strong (Au–N bond energy ~6 kcal/mol), the presence of two amines in a chelating configuration should ensure a greater affinity (*33*, *34*). MD simulations of the BINAMINE-CTAC system confirmed the formation of chiral, worm-like aggregates, similar to those for BINOL-CTAC, with a helical distribution of unit vectors (fig. S6). HAADF-STEM analysis of the particles obtained by seeded growth in (*R*)-BINAMINE-CTAC revealed the presence of gold nanorods displaying a highly complex surface morphology resembling an intricate network of wrinkles (Fig. 2C and figs. S7 to S10). Scanning electron microscopy (SEM) imaging of the same NPs also revealed a complex surface structure (fig. S11).

Notwithstanding, proper understanding of such complex structures demands a more detailed characterization than that provided by either HAADF-STEM or SEM, which only retrieve two-dimensional (2D) and surface information, respectively. We used HAADF-STEM tomography, which has been shown to be a powerful technique to investigate the 3D morphology of complex NCs. We prepared similar particles, but with increasing dimensions, using 130- × 29-nm gold nanorods as seeds and characterized them with HAADF-STEM (see the supplementary materials for synthetic details). Electron tomography reconstructions of nanorods with the smallest degree of growth, i.e., 165 nm long and 73 nm thick, revealed the presence of sharp wrinkles oriented in a seemingly radial direction and displaying tilt angles between 0° and 45° with respect to the short axis (Fig. 2D). Detailed analysis of the internal structure suggested that the wrinkles grew from the surface of the gold nanorod seed in a quasiradial direction up to ~15 to 20 nm in height and with a constant width of 3 to 4 nm. The average intergroove separation was measured to be 2 to 3 nm (Fig. 2E), which is near the 3.75-nm width of worm-like micelles obtained from MD simulations. This finding would suggest a micelle-directed chiral growth of the gold nanorods, as we initially hypothesized.

The width of the wrinkles was compatible with the diffusion of metal ion–loaded CTAC micelles from solution (*35*), leading to reduction at available gold sites between (*R*)-BINAMINE-surfactant micelles on the nanorod surface. By varying the concentration of nanorod seeds in the growth mixture, chiral nanorods were obtained with increased dimensions of 210 nm in length and 112 nm in thickness, and their surface was more entangled (Fig. 2F).

Nonetheless, the presence of tilted wrinkles was still visible and angles between 0° and 45° were again observed [between 0° and −45° in the case of (*S*)-BINAMINE; fig. S9]. Analysis of the internal structure revealed wrinkles of ~35 to 45 nm in height and a similar width to those in the smaller rods (Fig. 2G). Finally, even larger chiral nanorods (270 nm long and 175 nm thick) were investigated. Although the surface appeared more undefined in this case (Fig. 2H), analysis of the internal structure showed the characteristic features described for the previous samples (Fig. 2I). For completeness, 3D animations of all discussed systems are provided as supplementary materials (movies S1 to S48).

Diffraction of ideal helical structures results in X-shaped patterns (*36*, *37*). Thus, we applied fast Fourier transformations (FFTs) to our 3D reconstructions. We exemplify this analysis

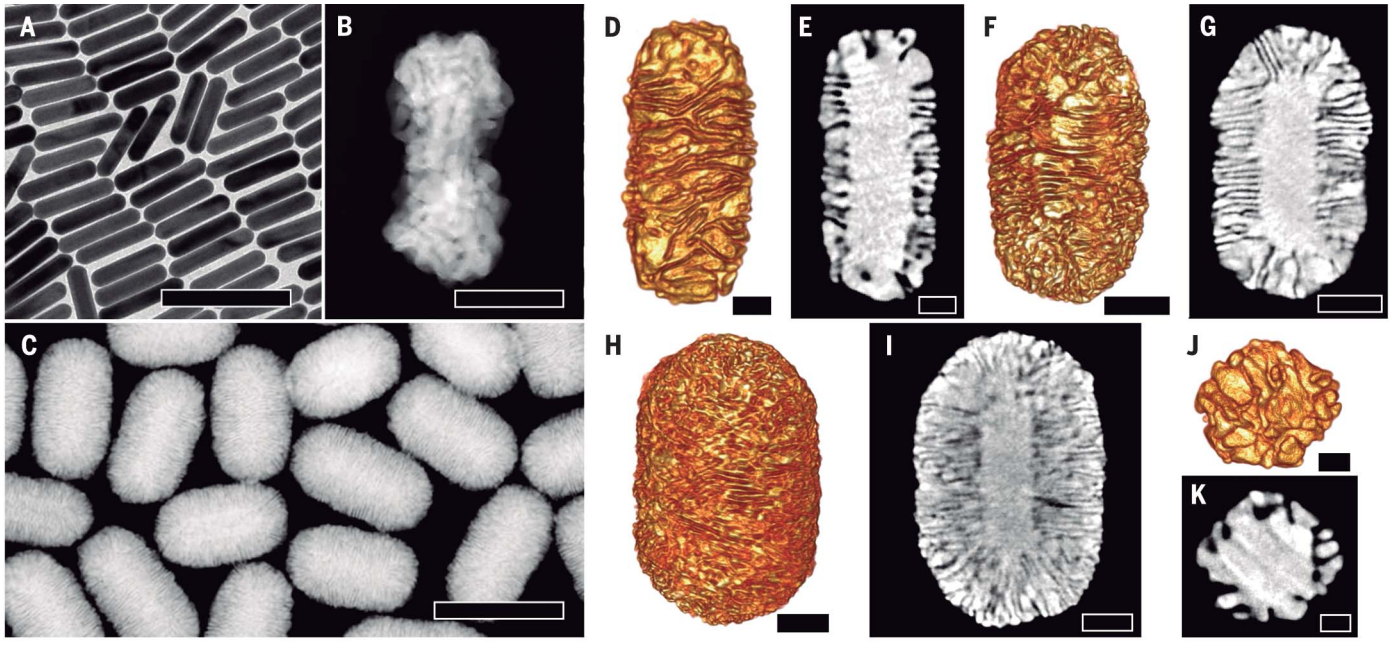
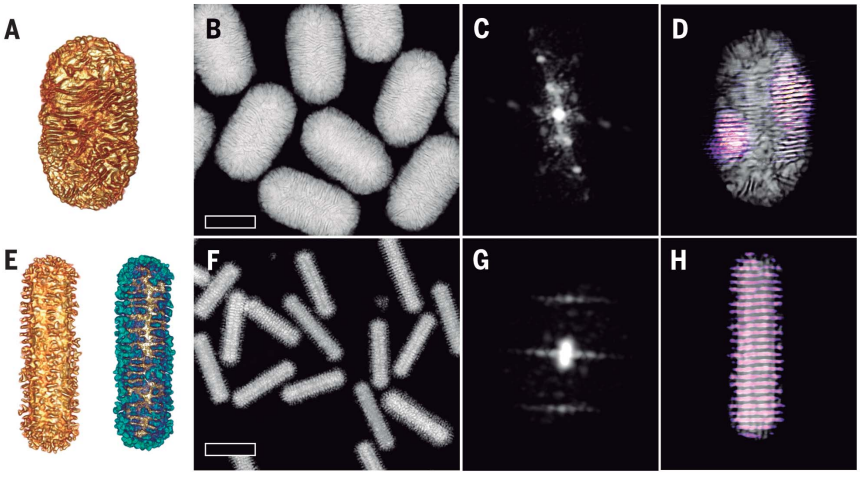


Fig. 2. Growth of chiral gold nanorods in micellar systems and effects of size and shape. (**A**) TEM image of gold nanorods 130 nm in length and 29 in width used as seeds. (**B**) High-magnification HAADF-STEM image of a chiral gold nanorod grown in (*R*)-BINOL displaying a complex surface containing wrinkles. (**C**) HAADF-STEM image at low magnification of gold nanorods obtained in the presence of (*R*)-BINAMINE displaying a complex surface containing sharp wrinkles. Scale bars, 200 nm [(A) and (C)] and 100 nm (B). (**D** to **I**) Gold nanorods of 165 × 73 nm [(D) and (E)], 210 × 112 nm [(F) and (G)], and 270 × 175 nm [(H) and (I)] grown in (*R*)-BINAMINE-surfactant mixtures were analyzed by HAADF-STEM. Scale bars, 20 nm [(D) and (E)], 50 nm [(F) to (I)], and 10 nm [(J) and (K)]. Tomography reconstructions [(D), (F), and (H)] reveal their surface topography, and selected orthoslices show the growth of wrinkles from the gold nanorod seeds and the internal structure of the wrinkle network [(E), (G) and (I)]. (**J** and **K**) Tomography reconstruction and selected orthoslice of an NP obtained by overgrowth of a 30-nm gold nanosphere in an (*R*)-BINAMINE-surfactant mixture. Scale bar, 100 nm. Corresponding animated reconstructions are provided in movies S1 (D), S3 (E), S7 (F), S9 (G), S18 (H), S20 (I), S43 (J), and S45 (K).

Fig. 3. Qualitative analysis of chiral features. (**A**) Tomography reconstruction of a gold nanorod grown in an (*R*)-BINAMINE-surfactant mixture (see also movie S7). (**B**) Low-resolution HAADF-STEM image of chiral gold nanorods obtained in the presence of (*R*)-BINAMINE. Scale bar, 100 nm. (**C**) 3D FFT from which the inverse FFT was computed after segmentation (see also movie S10). (**D**) Inverse FFT (see also movie S12) showing the areas of the particle with chiral features overlapped with the projection of the tomography, indicating a right-handed angle. (**E**) Tomography reconstruction in HAADF-STEM (left) and EDX (right, Pt displayed in green) modes of an Au@Pt nanorod grown in an (*R*)-BINOL-surfactant mixture (see also movies S24 and S25). (**F**) HAADF-STEM image at low magnification of Au@Pt nanorods. Scale bar, 100 nm. (**G**) 3D FFT from which the inverse FFT was computed after segmentation (see also movies S28 and S29). (**H**) Inverse FFT (see also movie S30) showing the areas of the particle with chiral features overlapped with the projection of the tomography, indicating a right-handed lower angle.



with chiral nanorods of 210 × 112 nm (Fig. 3, A and B), from which the 3D FFT indeed shows an X-shaped pattern (Fig. 3C); further details, as well as comparison with an idealized model (fig. S12), with a smooth nanorod (fig. S13), and with a (*R*)-BINOL-grown nanorod (fig. S13), are provided in the supplementary materials. Next, the spots in reciprocal space can be linked to the corresponding features in real space by manually segmenting the 3D FFT to minimize noise. By using the segmented 3D FFT as a mask (see details in the supplementary materials), an inverse FFT (*38*) was computed and overlayed (pink fringes) with the original reconstruction, as exemplified in Fig. 3D, so that the helical features are visually highlighted. Although the use of a 3D FFT does not provide a quantitative value such as, e.g., the Hausdorff chirality measure (*39*), it provides a qualitative and visual way to estimate the presence of chiral features. The inverse FFT enables direct location of specific periodicities at a local level.

Fringes with a right-handed angle can be seen in Fig. 3D, and the wrinkles growing on the sides of the gold nanorod seeds seemed to display a better-defined chiral arrangement than those located at the hemispherical tips. The observation of a curvature-dependent growth was confirmed by a control experiment, in which seeded growth was performed on 30-nm gold nanospheres (figs. S14 and S15). Although some chiral features could still be observed, the spheres yielded a more random wrinkle organization, as revealed by the corresponding HAADF-STEM tomography reconstruction (Fig. 2, J and K), and further supported by their 3D FFT and corresponding inverse FFT (fig. S14). Accordingly, more intense plasmonic CD bands were recorded when gold nanorods were used as seeds (g-factor ~0.2; Fig. 4A) compared with spheres (g-factor ~0.003; fig. S15).

Additional evidence behind the general validity of this chiral growth method and the proposed mechanism was provided by experiments in which we varied the nature of the metal deposited during seeded growth. We thus implemented the seeded growth of Pt on (*R*)-BINOL-CTAC–covered gold nanorods using identical dimensions and similar growth conditions. Representative results in Fig. 3, E to H, show that seeded growth resulted in very regular wrinkled platinum coating. The smaller dimensions of the wrinkles, as well as the different electron configuration between platinum and gold (see a 3D EDX reconstruction in Fig. 3E, right), allow a better distinction of the obtained pattern. Although the tilting angle is obviously lower than that measured for gold, a helical character is demonstrated by the 3D-FFT analysis (Fig. 3, G and H). Unfortunately, the lossy character of platinum does not allow recording of meaningful plasmonic optical activity for these samples.

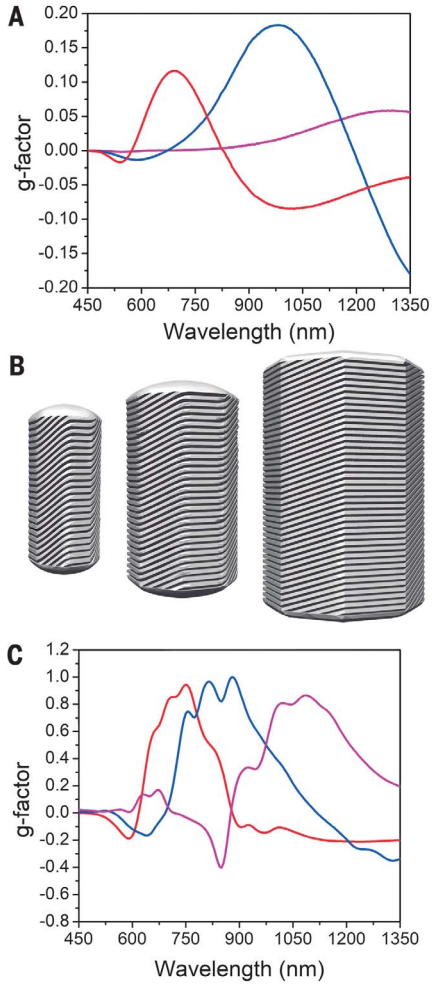


Fig. 4. Effect of nanorod dimensions on chiral plasmonic activity. (**A**) Spectral evolution of the anisotropy factor for chiral Au nanorods with increasing particle size: 165 × 73 nm (red), 210 × 112 nm (blue), and 270 × 175 nm (magenta). (**B**) Models of chiral gold nanorods used to simulate the chiral plasmonic properties (from left to right: 165 × 73, 210 × 112, and 270 × 175 nm). (**C**) Calculated anisotropy factor spectra [from left to right in (C): red, blue, and magenta].

We next demonstrate that the optical activity (CD bands) can be modulated by varying the dimensions of chiral nanorods (by varying the amount of seeds in the growth solution). In all cases, an intense negative Cotton effect was recorded for the (*R*)-enantiomer (*29*). As the size of the nanorods was increased from 165 × 73 up to 270 × 175 nm, the positive band redshifted from 700 up to 1300 nm, and the maximum of the negative band shifted from 1100 up to beyond 1350 nm (fig. S16). Analysis of the anisotropy factor showed values ranging from 0.1 to 0.2 (Fig. 4A), which are among the highest values reported for colloidal plasmonic NPs in the visible and near-infrared (NIR) ranges (*10*, *11*). By extending the analysis further into the NIR, up to the limit of our CD spectrometer (1600 nm), we recorded even higher dissymmetry factors of 0.25 to 0.28 (fig. S17). However, the noise in the 1400- to 1500-nm range precluded us from obtaining more compelling evidence to confirm the precise values.

When gold nanorods of different dimensions and degrees of anisotropy were used to seed the chiral growth (figs. S18 and S19), similar trends in optical activity were observed. However, the anisotropy factor was influenced by the nanorod dimensions. In general, slightly lower g-factor values were obtained when shorter and thinner gold nanorods were used as seeds (100 × 12 nm; i.e., higher aspect ratio), whereas a decrease was observed for seeds with similar length to the optimal one but thicker in diameter (135 × 52 nm; i.e., lower aspect ratio).

The relation between the chiral features observed by electron microscopy and the chiroptical activity measured by CD spectroscopy was further investigated by electromagnetic modeling. We used an accurate full-wave solver based on Maxwell's surface integral equations and the method of moments to model the plasmonic properties of chiral gold nanorods (see the supplementary materials for details) (*40*–*42*). Three-dimensional computer-aided design (CAD) models were constructed on the basis of the information provided by the experimental electron tomography reconstructions (Fig. 4B). Although suitable reproduction of the observed morphologies appeared challenging, given the intricate network of wrinkles covering the chiral gold nanorod surface, we created models for the three different dimensions characterized by electron microscopy: 165 × 73, 210 × 112, and 270 × 175 nm. To mimic the regions with wrinkles displaying different angles, each rod was built with 16 helices having four leveled and four inclined steps per pitch (pitch of 90 nm). The wrinkle width was fixed at 3.5 nm and the separation distance between wrinkles was 2.5 nm.

The simulations revealed chiral plasmonic bands that shifted toward longer wavelengths with increasing nanorod dimensions, in agreement with the experimental results (Fig. 4C and figs. S20 to S22). Differences in the anisotropy factors can be attributed to the polydispersity in size and shape of the real sample, as well as the difficulty in modeling the exact morphology of the NPs. Overall, our theoretical model confirmed the origin of the plasmonic chiroptical activity and the importance of sharp wrinkles for the emergence of strong CD responses. Replacement of BINAMINE by L-cysteine did not induce any clear modification on the CD signal obtained for chiral nanorods synthesized with either (*R*)-BINAMINE or (*S*)-BINAMINE. A substantial modification

would be expected if the chiroptical properties stemmed from the coupling of the chiral molecules with the NP plasmon rather than from the chiral structural features (fig. S23) (*43*).

We have experimentally demonstrated that chiral gold nanostructures can be readily obtained by seeded growth of gold nanorods in BINOL-surfactant mixtures, whereas NP colloids with intense CD responses were obtained using BINAMINE-surfactant mixed micelles. The chiroptical properties stem from the formation and stabilization of steep chiral wrinkles. These findings point to a dual role of mixed micelles: (i) templating the growth of steep wrinkles and (ii) the subsequent stabilization of such morphological features.

We propose that the formation of wrinkles during seeded growth can be explained by the presence of elongated micelles coiled on the gold nanorod seeds, which act as patterns directing the diffusion of micellar aggregates containing gold ions, from the aqueous solution toward the core NC at intermicellar regions. Indeed, the size of micelles in solution would fit the interwrinkle distance on the nanorod surface. However, wrinkle formation is only induced when the rate of gold ion deposition is faster than their diffusion on the surface (*44*). We observed that upon lowering the concentration of reducing agent (ascorbic acid, from 160 down to 1.6 mM), the CD intensity decreased by an order of magnitude and less-defined surface roughness was observed in HAADF-STEM (fig. S24).

An additional source of surface stabilization is required to preserve the wrinkles after growth. In this case, BINAMINE plays such a role during the growth process, likely because of its two amine moieties in chelating conformation (*33*). When one of the amine groups in BINAMINE was replaced by a hydroxyl moiety, the passivation effect was suppressed and the growth of chiral features was hindered, which resulted in a decrease of CD intensity (fig. S25). Once the growth is completed, addition of a stronger stabilizing agent (e.g., cysteine) is necessary because the adsorption of surfactant micelles is not sufficiently strong and the NPs may slowly reshape, as reflected by the loss of chiroptical properties (fig. S26).

The approach described here for the synthesis of chiral plasmonic anisotropic NCs is simple, reproducible, and holds great potential for the large-scale production intrinsic to colloid chemistry methods. Compared with the growth of inorganic chiral nanostructures based on the direct chemisorption of small additives on the NC surface to induce chiral growth, the micelle-directed growth method relies on the ability of the chiral cosurfactant to direct the formation of helical micelles. Such supramolecular assemblies have sufficient interaction points with the NPs to effectively transfer their chirality to them during the growth step through a multivalency effect.

REFERENCES AND NOTES

1. J. Bailey *et al*., *Science* **281**, 672–674 (1998).
2. L. A. Hodge, F. B. Dunning, G. K. Walters, R. H. White, G. J. Schroepfer Jr., *Nature* **280**, 250–252 (1979).
3. R. E. Franklin, R. G. Gosling, *Nature* **171**, 740–741 (1953).
4. D. Hanein, B. Geiger, L. Addadi, *Science* **263**, 1413–1416 (1994).
5. T. P. Yoon, E. N. Jacobsen, *Science* **299**, 1691–1693 (2003).
6. K. Ray, S. P. Ananthavel, D. H. Waldeck, R. Naaman, *Science* **283**, 814–816 (1999).
7. A. S. Karimullah *et al*., *Adv. Mater.* **27**, 5610–5616 (2015).
8. C. Hao, L. Xu, H. Kuang, C. Xu, *Adv. Mater.* **32**, e1802075 (2019).
9. J. K. Gansel *et al*., *Science* **325**, 1513–1515 (2009).
10. M. Hentschel, M. Schäferling, X. Duan, H. Giessen, N. Liu, *Sci. Adv.* **3**, e1602735 (2017).
11. H.-E. Lee *et al*., *Nature* **556**, 360–365 (2018).
12. G. Zheng *et al*., *Angew. Chem. Int. Ed.* **57**, 16452–16457 (2018).
13. A. Kuzyk *et al*., *Nature* **483**, 311–314 (2012).
14. C.-L. Chen, P. Zhang, N. L. Rosi, *J. Am. Chem. Soc.* **130**, 13555–13557 (2008).
15. A. Guerrero-Martínez *et al*., *Angew. Chem. Int. Ed.* **50**, 5499–5503 (2011).
16. A. G. Mark, J. G. Gibbs, T.-C. Lee, P. Fischer, *Nat. Mater.* **12**, 802–807 (2013).
17. Y. Xia, Y. Xiong, B. Lim, S. E. Skrabalak, *Angew. Chem. Int. Ed.* **48**, 60–103 (2009).
18. B. Nikoobakht, M. A. El-Sayed, *Chem. Mater.* **15**, 1957–1962 (2003).
19. S. Hong, K. L. Shuford, S. Park, *Chem. Mater.* **23**, 2011–2013 (2011).
20. H.-E. Lee *et al*., *Nat. Commun.* **11**, 263 (2020).
21. C. A. Silvera Batista, R. G. Larson, N. A. Kotov, *Science* **350**, 1242477 (2015).
22. M. L. Personick, C. A. Mirkin, *J. Am. Chem. Soc.* **135**, 18238–18247 (2013).
23. S. E. Lohse, N. D. Burrows, L. Scarabelli, L. M. Liz-Marzán, C. J. Murphy, *Chem. Mater.* **26**, 34–43 (2014).
24. S. K. Meena, M. Sulpizi, *Angew. Chem. Int. Ed.* **55**, 11960–11964 (2016).
25. L. Scarabelli, M. Grzelczak, L. M. Liz-Marzán, *Chem. Mater.* **25**, 4232–4238 (2013).
26. X. Ye, C. Zheng, J. Chen, Y. Gao, C. B. Murray, *Nano Lett.* **13**, 765–771 (2013).
27. G. González-Rubio *et al*., *ACS Nano* **13**, 4424–4435 (2019).
28. T. H. Ito *et al*., *Langmuir* **32**, 8461–8466 (2016).
29. M. M. Green *et al*., *J. Am. Chem. Soc.* **111**, 6452–6454 (1989).
30. E. E. Greciano *et al*., *J. Am. Chem. Soc.* **141**, 7463–7472 (2019).
31. Z. Fan, A. O. Govorov, *Nano Lett.* **12**, 3283–3289 (2012).
32. J. Zhang *et al*., *J. Am. Chem. Soc.* **132**, 14012–14014 (2010).
33. S. Engel, E.-C. Fritz, B. J. Ravoo, *Chem. Soc. Rev.* **46**, 2057–2075 (2017).
34. B.-K. Pong, J.-Y. Lee, B. L. Trout, *Langmuir* **21**, 11599–11603 (2005).
35. J. Pérez-Juste, L. M. Liz-Marzán, S. Carnie, D. Y. C. Chan, P. Mulvaney, *Adv. Funct. Mater.* **14**, 571–579 (2004).
36. H. Zhu *et al*., *Small* **1**, 1180–1183 (2005).
37. M. Gailhanou, J. M. Roussel, *J. Appl. Cryst.* **24**, 14012–14014 (2003).
38. M. C. Scott *et al*., *Nature* **483**, 444–447 (2012).
39. J.-Y. Kim *et al*., *J. Am. Chem. Soc.* **141**, 11739–11744 (2019).
40. D. M. Solís, J. M. Taboada, F. Obelleiro, L. M. Liz-Marzán, F. J. García de Abajo, *ACS Nano* **8**, 7559–7570 (2014).
41. J. M. Taboada, J. Rivero, F. Obelleiro, M. G. Araújo, L. Landesa, *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **28**, 1341–1348 (2011).
42. D. M. Solís, J. M. Taboada, F. Obelleiro, *IEEE Trans. Antenn. Propag.* **63**, 2141–2152 (2015).
43. A. O. Govorov, Z. Fan, P. Hernandez, J. M. Slocik, R. R. Naik, *Nano Lett.* **10**, 1374–1382 (2010).
44. Y. Xia, K. D. Gilroy, H.-C. Peng, X. Xia, *Angew. Chem. Int. Ed.* **56**, 60–95 (2017).

ACKNOWLEDGMENTS

Funding: L.M.L.-M. acknowledges funding from the European Research Council (ERC AdG grant no. 787510). G.G.-R. and J.M. thank the Spanish MICIU for FPI (BES-2014-068972) and Juan de la Cierva (FJCI-2015-25080) fellowships. S.B., L.M.L.-M., V.K., and A.P.-T. acknowledge financial support from the European Commission under the Horizon 2020 Programme by grant no. 731019 (EUSMI) and ERC Consolidator grant no. 815128 (REALNANO). J.M.T. and F.O. acknowledge financial support from the Spanish MICIU (grants TEC2017-85376-C2-1-R and TEC2017-85376-C2-2-R), as well as from the ERDF and the Galician Regional Government as part of the agreement for funding the Atlantic Research Center for Information and Communication Technologies (AtlantTIC). A.G.-M. acknowledges financial support from the Spanish MICIU (grant no. RTI2018-095844-B-I00). E.G.N. and L.G.M. acknowledge funds from the Spanish MICIU (grant no. FIS2017-89361-C3-2-P), as well as the use of the Mare-Nostrum supercomputer and the technical support provided by Barcelona Supercomputing Center from the Spanish Network of Supercomputing (grant nos. QCM-2018-3-0039 and QCM-2019-1-0038). This work was performed under the Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency (grant no. MDM-2017-0720). **Author contributions:** G.G.-R., J.M., and L.M.L.-M. conceived the project. G.G.-R., J.M., and V.K. performed NP synthesis and CD experiments. A.P.-T., I.L., and S.B. performed EM characterization and analysis. P.L., E.G.N., and L.G.M. executed and analyzed MD simulations. D.M.S., E.G.N. J.M.T., and F.O. performed electromagnetic modeling of plasmonic properties of chiral NPs. G.G.-R., J.M., A.G.-M., S.B., and L.M.L.-M. wrote the manuscript with comments from all authors. L.M.L.-M. supervised the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are presented in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1472/suppl/DC1
Materials and Methods
Figs. S1 to S26
Tables S1 to S6
References (*45*–*60*)
Movies S1 to S48

5 November 2019; resubmitted 22 January 2020
Accepted 1 May 2020
10.1126/science.aba0980

EXOPLANETS

A multiplanet system of super-Earths orbiting the brightest red dwarf star GJ 887

S. V. Jeffers[1]*, S. Dreizler[1], J. R. Barnes[2], C. A. Haswell[2], R. P. Nelson[3], E. Rodríguez[4], M. J. López-González[4], N. Morales[4], R. Luque[5,6], M. Zechmeister[1], S. S. Vogt[7], J. S. Jenkins[8,9], E. Palle[5,6], Z. M. Berdiñas[8], G. A. L. Coleman[3,10], M. R. Díaz[8], I. Ribas[11,12], H. R. A. Jones[13], R. P. Butler[14], C. G. Tinney[15], J. Bailey[15], B. D. Carter[16], S. O'Toole[17], R. A. Wittenmyer[18], J. D. Crane[19], F. Feng[14], S. A. Shectman[19], J. Teske[19], A. Reiners[1], P. J. Amado[4], G. Anglada-Escudé[3,11,12]

The closet exoplanets to the Sun provide opportunities for detailed characterization of planets outside the Solar System. We report the discovery, using radial velocity measurements, of a compact multiplanet system of super-Earth exoplanets orbiting the nearby red dwarf star GJ 887. The two planets have orbital periods of 9.3 and 21.8 days. Assuming an Earth-like albedo, the equilibrium temperature of the 21.8-day planet is ~350 kelvin. The planets are interior to, but close to the inner edge of, the liquid-water habitable zone. We also detect an unconfirmed signal with a period of ~50 days, which could correspond to a third super-Earth in a more temperate orbit. Our observations show that GJ 887 has photometric variability below 500 parts per million, which is unusually quiet for a red dwarf.

At visible wavelengths, GJ 887 (HD 217987) is the brightest red dwarf in the sky and, at a distance of 3.29 pc, is the 12th closest star system to the Sun. GJ 887 is the most massive red dwarf within 6 pc of the Sun, which is close enough for a direct stellar radius measurement using interferometry (*1*). GJ 887's stellar parameters are listed in Table 1. Red dwarfs are amenable to radial velocity (RV) searches for temperate Earth-mass exoplanets: their low luminosity means that temperate planets have short orbital periods, and their low stellar mass implies that Earth-mass planets can impart a reflex RV detectable with current instrumentation. Although the transit method of planet discovery efficiently detects planets because many stars can be simultaneously monitored, it will detect only planets that pass through the line of sight between Earth and the host star. Consequently, only 1 to 2% of habitable-zone planets (i.e., those with surfaces that can support liquid water) are detectable with the transit method. The RV method is necessary to achieve a complete census of the planets orbiting our closest stellar neighbors, especially red dwarfs.

We monitored GJ 887 as part of the Red Dots #2 project. Nightly observations were taken with the High Accuracy Radial Velocity Planet Searcher (HARPS) (*2*) for 3 months. We also obtained contemporaneous photometric observations (*3*). Regular nightly sampling combined with photometric observations mitigates against false-positive exoplanet detections caused by intrinsic stellar variability and other sources of correlated noise. We supplement our data with >200 archival observations, spanning nearly 20 years (*3*), from HARPS, the Planet Finder Spectrograph (PFS) (*4*), the High-Resolution Echelle Spectrometer (HIRES) (*5*), and the University College London Echelle Spectrograph (UCLES) (*6*). We used photometry from various ground-based observatories and the Transiting Exoplanet Survey Satellite (TESS) spacecraft (*7*). Tables S3 and S4 list all the data we used.

We searched for a candidate planet by adding a (circular) Keplerian orbit test signal to a base model and measuring the improvement in the logarithm of the likelihood statistic (*3*). The base model is composed of an offset and an instrumental jitter that are added to the measurement uncertainties for each dataset. We used this base model to generate log-likelihood periodograms for both the RV and photometric data and then searched for signals by plotting the increase in the log-likelihood statistic against test period (Fig. 1). The highest peaks were evaluated for statistical significance (*8*, *9*). We recursively added additional planet test signals, adjusting all the parameters to maximize the likelihood of all planet signals and the parameters of the base model. We continued this iterative process until no signals below a threshold of 0.1% false-alarm probability were found in the time series (fig. S1). We detected periodic signals at 9.3, 21.8, and 50.7 days (Fig. 1B) and verified them using several independent fitting procedures and algorithm implementations (*3*). The regular sampling of the Red Dots #2 dataset helped to disentangle the signals under investigation (Fig. 1A).

Stellar magnetic activity can induce an asymmetric distortion of the spectral lines, shifting the measured line center and consequently inducing an apparent RV shift, which may appear as a false-positive exoplanet at the stellar rotation period (*10*). The rotation period of GJ 887 is unknown, so we searched for periodicities in the photometric data (*3*). The archival data from 2002–2004 show an ~200-day period, but this period was undetectable in the 2018 quasi-simultaneous photometric observations because the time span is too short. Our analysis of the TESS photometry shows very low intrinsic variability, with a semi-amplitude of 240 parts per million. It is unclear whether the low photometric variability of the TESS photometry is caused by systematic effects, known to affect the TESS observations, but we use this value as an upper limit to the intrinsic variability of GJ 887. The TESS variability can be explained by one starspot, or a group of starspots, with a total diameter of 0.3% of the stellar surface, indicating that GJ 887 is slowly rotating with very few surface brightness inhomogeneities (*11*). GJ 887 is less magnetically active than most stars with the same effective temperature, as demonstrated by: the very low starspot coverage; low photometric variability; the activity metric derived from stellar Ca II H and K lines, $\log(R'_{\mathrm{HK}}) = -4.805$ (*12*); and the very low Hα activity (*13*).

The RV signals are detected in the Red Dots # 2 HARPS data alone (Fig. 1A), so we investigated additional spectral signatures of stellar magnetic activity using this dataset. We extracted a time series of the flux in the cores of the Na I D, Hα, and Hβ lines and calculated the S-index, which is the ratio of flux in the cores in the Ca II H and K lines compared with the continuum (*3*). The S-index and Na D lines both show a weak signal at about 55 days, whereas the Hα and Hβ lines show a weak signal at 38 days (fig. S5). These differing periodicities could reflect time scales of various stellar activity processes on the star, and despite being low in amplitude, they make a planetary origin for RV signals in the 30-to-60-day domain less certain. None of these activity periodicities are close to the RV signals at

[1]Institut für Astrophysik, Georg-August-Universität, 37077 Göttingen, Germany. [2]School of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK. [3]School of Physics and Astronomy, Queen Mary University of London, London E1 4NS, UK. [4]Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas, 18008 Granada, Spain. [5]Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain. [6]Departamento de Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain. [7]University of California/Lick Observatory, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. [8]Departamento de Astronomia, Universidad de Chile, Santiago, Chile. [9]Centro de Astrofísica y Tecnologías Afines, Santiago, Chile. [10]Physikalisches Institut, Universität Bern, 3012 Bern, Switzerland. [11]Institut de Ciències de l'Espai, Consejo Superior de Investigaciones Científicas, Campus Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain. [12]Istitut d'Estudis Espacials de Catalunya, E-08034 Barcelona, Spain. [13]Centre for Astrophysics Research, University of Hertfordshire, Hatfield AL10 9AB, UK. [14]Earth and Planets Laboratory, Carnegie Institution for Science, Washington, DC 20015, USA. [15]Exoplanetary Science at University of New South Wales, School of Physics, University of New South Wales, Sydney, NSW 2052, Australia. [16]Centre for Astrophysics, University of Southern Queensland, Springfield Central, QLD 4300, Australia. [17]Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia. [18]Centre for Astrophysics, University of Southern Queensland, Toowoomba, QLD 4350, Australia. [19]The Observatories of the Carnegie Institution for Science, Pasadena, CA 91101, USA.
*Corresponding author. Email: sandrajeffers.astro@gmail.com

9.3 and 21.8 days, but they do make the RV signal at 50.7 days questionable.

Correlated noise—for example, that caused by stellar activity—can be assessed using the covariances between observations. To further test the planetary origin of the detected RV signals, we fitted maximum likelihood model functions using two planet models with and without Gaussian processes (GPs) (*3*). All of the models including GPs improved the fit to the data compared with those without, and the amplitude of the signals with periods of 9.3 and 21.8 days remained unchanged within their 1σ uncertainty. The modeling of the correlated noise using GPs therefore does not affect these two signals. However, the significance of the third signal drops substantially when including a GP in the model, casting further doubts on its Keplerian nature. Table S4 lists the derived values and statistics from these models.

We conclude that the two signals with orbital periods of 9.3 and 21.8 days correspond to two exoplanets, GJ 887 b and GJ 887 c, respectively. The minimum masses of these planets ($m_p \sin i$)

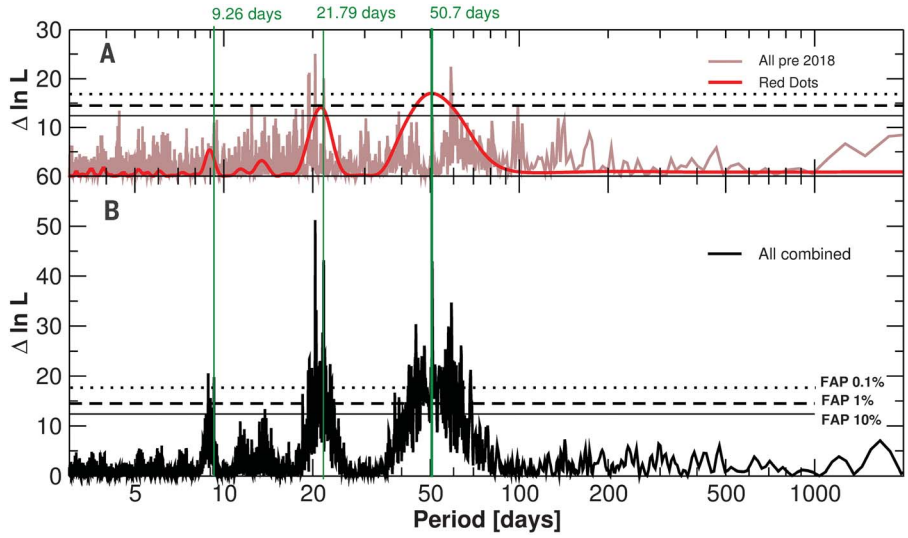


Fig. 1. Periodograms of the RV data. (**A**) The log-likelihood (Δln L) periodogram obtained separately for all RV data before 2018 (brown) and the Red Dots #2 campaign (red). (**B**) The same search for a first signal when combining all the RV observations together. The vertical green lines indicate our derived model periods for GJ 887 b and GJ 887 c and the third signal. The horizontal dashed lines in both panels indicate the false alarm probability (FAP) values.

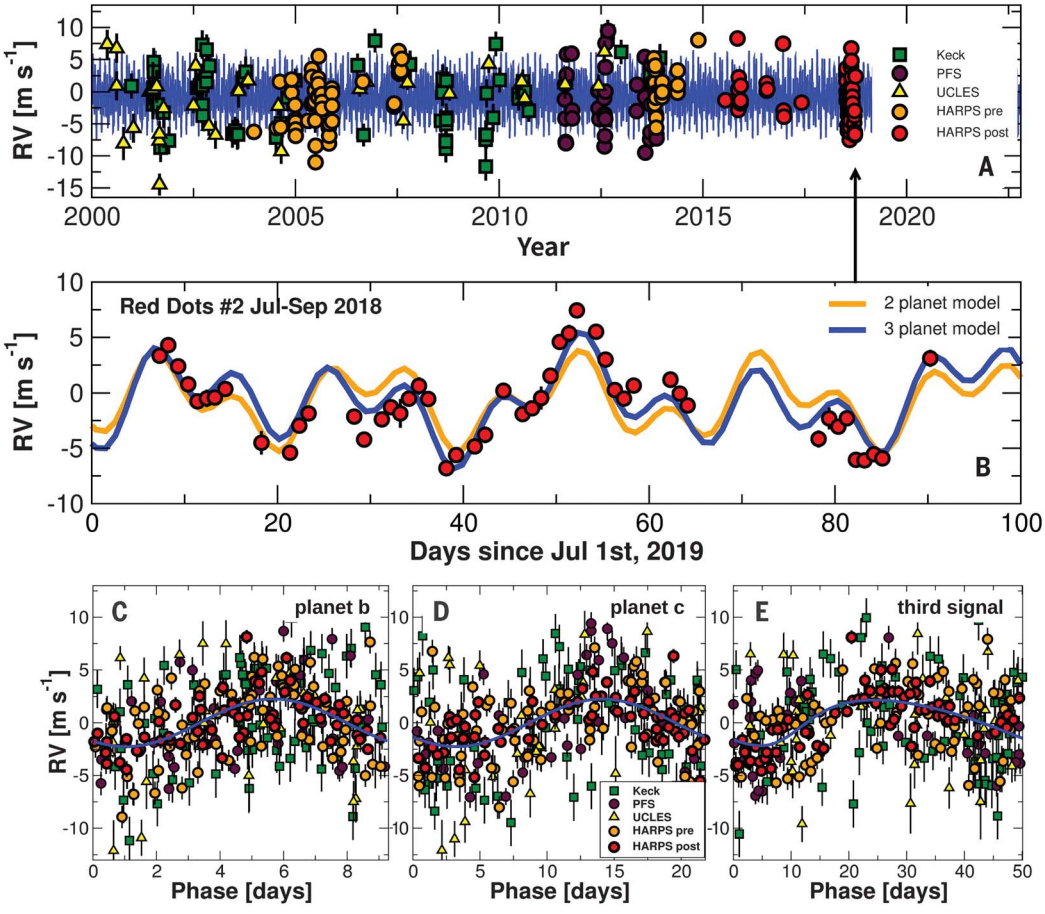


Fig. 2. Time series of radial velocity measurements. All RV measurements are shown, with the instruments used indicated in the legend. "HARPS pre" and "HARPS post" refer to data collected before and after an instrument upgrade, respectively. (**A**) Radial velocity measurements over 18 years using all instruments. The best-fitting model with three Keplerian signals is shown with a solid blue line. (**B**) Zoom in on area indicated with arrow in (A), showing the Red Dots #2 observations. The three-planet model is shown with a solid blue line, and the two-planet model with a solid orange line. It is uncertain whether the ~50-day signal has a planetary origin, but three periodic modulations are required to fit the observations. (**C** to **E**) The data from (A), folded on the periods of each candidate signal, after subtracting the other signals. The best-fitting model is shown as a solid blue line.

are 4.2 ± 0.6 and 7.6 ± 1.2 Earth masses ($M_\oplus$), which makes them two super-Earth exoplanets with orbital semimajor axes (a_p) of 0.068 and 0.120 astronomical units (au). The inner planet has an orbital eccentricity consistent with zero, but the outer planet is more likely to have a small nonzero eccentricity (Fig. 2). We regard the third signal at ~50 days as dubious and likely related to stellar activity. The fits to our two-planet model and the two-planet + third signal model are shown in Fig. 2.

The long-term dynamical stability of the orbits can also be used to test the physical plausibility of the system and investigate whether it contains unusual configurations such as dynamical resonances. We performed a dynamical stability study (*3*) using the software package MERCURY6 (*14*). We find that all two-planet solutions are stable even if eccentricities are allowed to vary. The ratio of periods of these two planets is close to 7:3, but the simulations do not support the presence of a dynamical resonance, as there is an absence of oscillating orbital alignment variations (*15*). However, we find that the system must be in a dynamically active state, driving oscillatory changes in the eccentricities of both planets. These interactions produce very regular variations, supporting the hypothesis that the two-planet configuration is dynamically stable on very long time scales. For a putative system with three planets, only ~25% of the 1000 best-fitting models would be dynamically stable over 10^5 years, but this is mostly due to the poorly constrained eccentricities. Given that the observations only provide upper limits on the eccentricities, we investigated orbits that are assumed to be initially circular (initial zero eccentricities). Even in this three-planet case, >99% of configurations were found to be stable, meaning that the presence of a third planet cannot be ruled out using dynamic stability considerations.

The separations between the planets, in units of their spheres of gravitational influence (Hill radii), are ~19.1 for GJ 887 b and GJ 887 c and ~17.2 for GJ 887 c and GJ 887 d (if GJ 887 d is real and has a mass of 8.3 $M_\oplus$); these values are consistent with the system having undergone dynamical relaxation (*16*). Dynamical relaxation in systems of super-Earths results in ~80% of planets having orbital eccentricities $e_p \leq 0.1$, with the remaining 20% having $e_p \leq 0.3$ (*17*). We examined the expected tidal evolution of GJ 887 b using analytical methods (*18*, *19*), finding that the tidal circularization time scale of GJ 887 b is a few billion years for an assumed tidal dissipation parameter $Q'_p = 1000$. The tidal evolution is consistent with our observation that GJ 887 b's orbit is almost circular.

The multiplanet super-Earth system around GJ 887 is consistent with recent planet formation models (*20*, *21*). These models typically form chains of multiple planets trapped in mean-motion resonances that then migrate into orbits close to the central star. Depending on where the initial planets formed in the protoplanetary disc, they could have accreted either large amounts of water ice or just dry rocky silicates. As such, the planets may be either water-rich or water-poor. At the end of the gas disc lifetime, the resonant chains of planets can remain stable, yielding systems similar to the seven-planet TRAPPIST-1 planetary system (*22*), or they can become unstable, leading to collisions between planets and thus a nonresonant configuration (*20*). The GJ 887 planetary system appears to be more consistent with the latter scenario of unstable evolution. The presence of dynamical resonances can be very sensitive to the existence or absence of additional planets. Consequently, if the third signal at 50.7 days is real or if additional planets exist, this may result in a more resonant system.

We used standard methods (*23*) to calculate the distances from GJ 887 within which planets could support liquid water on their surfaces [the star's habitable zone (HZ)] and found that it extends from ~0.19 to 0.38 au. With an a_p of 0.120 ± 0.004, GJ 887 c is closer to its host star than the HZ but near the inner edge. If the ~50-day signal is planetary in origin, it corresponds to a super-Earth in GJ 887's HZ. Assuming an albedo, α, similar to Earth's (α = 0.3), the equilibrium temperature, T_{eq}, of GJ 887 b and GJ 887 c would be 468 and 352 K, respectively. Their incident energy fluxes from the star (the insolation *S*) are 7.95 and 2.56 times the Sun's insolation on Earth. Figure 3 shows the insolation of known planets orbiting M-type

Table 1. Stellar parameters for GJ 887 and parameters for GJ 887 b and GJ 887 c. The top half of the table lists parameters for GJ 887: the parallax in milliarcseconds (mas), distance in parsecs (pc), *V*- and *G*-band magnitudes, stellar mass in solar masses ($M_\odot$), metallicity relative to the Sun [Fe/H], luminosity and radius in solar units, projected rotational velocity (*vsini*), and surface gravity log *g*. The stellar mass was computed using a mass-radius relation (*26*). The bottom half of the table lists parameters for GJ 887 b and GJ 887 c: K_p is this amplitude, P_p is the period, S_{eff} is the incident flux from GJ 887 in units of the incident flux on Earth from the Sun, and T_{eq} is the equilibrium temperature of the planet.

Parameter	Value	Reference
Spectral type	M1V	(*27*)
Parallax (mas)	304.2190 ± 0.0451	(*28*)
Distance (pc)	3.2871 ± 0.0005	–
Magnitude	V = 7.34, G = 6.522	(*28*)
Mass ($M_\odot$)	0.489 ± 0.05	–
[Fe/H]	−0.06 ± 0.08	(*27*)
T_{eff} (K)	3688 ± 86	(*27*)
Luminosity ($L_\odot$)	0.0368 ± 0.004	(*27*)
Radius ($R_\odot$)	0.4712 ± 0.086	(*1*, *29*, *30*)
vsini (km s^{-1})	2.5 ± 1.0	(*31*)
logR'_{HK} mean	−4.805 ± 0.023	(*12*)
log (age/years)	9.46 ± 0.58	(*27*)
log *g*	4.78	(*32*)
Parameter	**GJ 887 b**	**GJ 887 c**
K_p (m s^{-1})	$2.1^{+0.3}_{-0.2}$	2.8 ± 0.4
P_p (d)	9.262 ± 0.001	$21.789^{+0.004}_{-0.005}$
$m_p$sin*i* ($M_\oplus$)	4.2 ± 0.6	7.6 ± 1.2
a_p (au)	0.068 ± 0.002	0.120 ± 0.004
$S_{eff,p}$ ($S_{eff,\oplus}$)	7.95 ± 0.2	2.56 ± 0.2
T_{eq} (K)	468	352

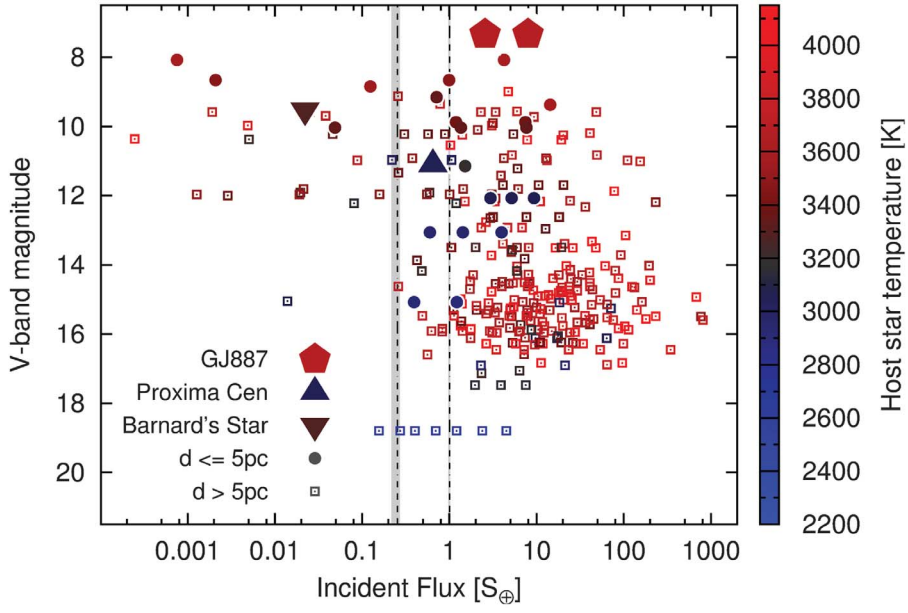


Fig. 3. The incident flux (insolation) of planets orbiting M dwarf stars. The dashed vertical lines delimit the habitable zone around GJ 887 for the maximum greenhouse planetary atmosphere (left) and the runaway greenhouse planetary atmosphere (right) (*23*). The solid vertical gray lines indicate the range of limits for the host stars of all planets plotted; these stars have T_{eff} ranging from 2400 to 4150 K (see color bar). GJ 887 b and GJ 887 c are indicated by the large red pentagons. Proxima Cen, Proxima Centauri; d, distance.

red dwarfs as a function of host star apparent magnitude. GJ 887 has the brightest apparent magnitude of any known M dwarf planet host. This brightness, combined with the high photometric stability of GJ 887, exhibited in the TESS data, and the high planet-star brightness and radius ratios, make these planets potential targets for phased-resolved photometric studies, especially in emission (*24*). Spectrally resolved phase photometry has been shown to be sensitive to the presence of an atmosphere and molecules such as CO_2 (*25*).

REFERENCES AND NOTES

1. T. S. Boyajian *et al.*, *Astrophys. J.* **757**, 112 (2012).
2. M. Mayor *et al.*, *Messenger* **114**, 20–24 (2003).
3. See supplementary materials.
4. J. D. Crane *et al.*, *Proc. SPIE* **7735**, 773553 (2010).
5. S. S. Vogt *et al.*, *Proc. SPIE* **2198**, 362–375 (1994).
6. F. Diego, A. Charalambous, A. C. Fish, D. D. Walker, *Proc. SPIE* **1235**, 562–576 (1990).
7. G. R. Ricker *et al.*, *J. Astron. Telesc. Instrum. Syst.* **1**, 014003 (2015).
8. R. V. Baluev, *Mon. Not. R. Astron. Soc.* **393**, 969–978 (2009).
9. I. Ribas *et al.*, *Nature* **563**, 365–368 (2018).
10. S. V. Jeffers *et al.*, *Mon. Not. R. Astron. Soc.* **438**, 2717–2731 (2014).
11. J. R. Barnes *et al.*, *Astrophys. J.* **812**, 42 (2015).
12. S. Boro Saikia *et al.*, *Astron. Astrophys.* **616**, A108 (2018).
13. S. V. Jeffers *et al.*, *Astron. Astrophys.* **614**, A76 (2018).
14. J. E. Chambers, *Mon. Not. R. Astron. Soc.* **304**, 793–799 (1999).
15. M. Perryman, *The Exoplanet Handbook* (Cambridge Univ. Press, ed. 2, 2018).
16. B. Pu, Y. Wu, *Astrophys. J.* **807**, 44 (2015).
17. S. T. S. Poon, R. P. Nelson, S. A. Jacobson, A. Morbidelli, *Mon. Not. R. Astron. Soc.* **491**, 5595–5620 (2020).
18. I. Dobbs-Dixon, D. N. C. Lin, R. A. Mardling, *Astrophys. J.* **610**, 464–476 (2004).
19. P. P. Eggleton, L. G. Kiseleva, P. Hut, *Astrophys. J.* **499**, 853–870 (1998).
20. G. A. L. Coleman, R. P. Nelson, *Mon. Not. R. Astron. Soc.* **457**, 2480–2500 (2016).
21. M. Lambrechts *et al.*, *Astron. Astrophys.* **627**, A83 (2019).
22. M. Gillon *et al.*, *Nature* **542**, 456–460 (2017).
23. R. K. Kopparapu *et al.*, *Astrophys. J.* **787**, L29 (2014).
24. E. M. R. Kempton *et al.*, *Publ. Astron. Soc. Pac.* **130**, 114401 (2018).
25. I. A. G. Snellen *et al.*, *Astron. J.* **154**, 77 (2017).
26. A. Schweitzer *et al.*, *Astron. Astrophys.* **625**, A68 (2019).
27. A. W. Mann, G. A. Feiden, E. Gaidos, T. Boyajian, K. von Braun, *Astrophys. J.* **804**, 64 (2015).
28. Gaia Collaboration, *Astron. Astrophys.* **616**, A1 (2018).
29. D. Ségransan, P. Kervella, T. Forveille, D. Queloz, *Astron. Astrophys.* **397**, L5–L8 (2003).
30. B.-O. Demory *et al.*, *Astron. Astrophys.* **505**, 205–215 (2009).
31. M. K. Browning, G. Basri, G. W. Marcy, A. A. West, J. Zhang, *Astron. J.* **139**, 504–518 (2010).
32. M. Rabus *et al.*, *Mon. Not. R. Astron. Soc.* **484**, 2674–2683 (2019).
33. R. P. Butler *et al.*, *Astron. J.* **153**, 208 (2017).

ACKNOWLEDGMENTS

We thank P. A. Peña Rojas for contributing results using the EMPEROR code. This work is based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under European Southern Observatory (ESO) programmes 0101.C-0516, 0101.C-0494, and 0102.C-0525. This Report includes data gathered with the 6.5-m Magellan Telescopes located at Las Campanas Observatory, Chile. Photometric data were collected with the robotic 40-cm telescope ASH2 at the SPACEOBS observatory (San Pedro de Atacama, Chile) operated by the Instituto de Astrofísica de Andalucía (IAA). This paper includes data collected with the TESS mission, obtained from the Mikulski Archive for Space Telescopes (MAST) data archive at the Space Telescope Science Institute (STScI). Funding for the TESS mission is provided by the NASA Explorer Program. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. **Funding:** S.V.J. acknowledges the support of the German Science Foundation (DFG) Research Unit FOR2544 "Blue Planets around Red Stars" project (JE 701/3-1) and DFG priority program SPP 1992 "Exploring the Diversity of Extrasolar Planets" JE 701/5-1. J.R.B. and C.A.H. acknowledge support from STFC Consolidated Grants ST/P000584/1 and ST/T000295/1. R.P.N. was supported by STFC Consolidated Grant ST/P000592/1. E.R., M.J.L.-G., N.M., and P.J.A. acknowledge support from the Spanish Agencia Estatal de Investigación through projects AYA2017-89637-R, AYA2016-79425-C3-3-P, ESP2017-87676-C5-2-R, and ESP2017-87143-R and the Centre of Excellence "Severo Ochoa" Instituto de Astrofísica de Andalucía (SEV-2017-0709). E.P. acknowledges support from the Spanish Agencia Estatal de Investigación PGC2018-098153-B-C31 and ESP2016-80435-C2-2-R. Z.M.B. acknowledges funds from CONICYT/FONDECYT Postdoctorado 3180405. G.A.L.C. acknowledges support from the Swiss National Science Foundation. M.R.D. acknowledges support of CONICYT/PFCHA-Doctorado Nacional 21140646, Chile. I.R. acknowledges support from the Spanish Ministry of Science and Innovation and the European Regional Development Fund through grants ESP2016-80435-C2-1-R and PGC2018-098153-B-C33 and the Generalitat de Catalunya/CERCA program. H.R.A.J. acknowledges support from the U.K. Science and Technology Facilities Council grant ST/M001008/1. C.G.T. is supported by Australian Research Council grants DP0774000, DP130102695, and DP170103491. J.T. was supported by NASA through Hubble Fellowship grant HST-HF2-51399.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. G.A.-E. is supported by the Ministerio de Ciencia, Innovación y Universidades Ramón y Cajal fellowship RYC-2017-22489 and by the Science and Technology Facilities Council grant number ST/P000592/1. **Author contributions:** S.V.J. led the observing proposal and team coordination, participated in the data analysis, and wrote the manuscript. S.D. led the data analysis and contributed to the writing of the manuscript. J.R.B. participated in the writing of the observing proposal, simulations, and manuscript review. C.A.H. participated in the writing of the observing proposal, final consistency checks, and writing of the manuscript. R.P.N.: Contributed the discussion of planetary dynamics and manuscript review. E.R.: ASH2 photometry, coordination of photometric observations, data analysis. M.J.L.G.: ASH2 photometry data reduction. N.M.: ASH2 photometry observer. R.L., M.Z., S.S.V., and J.S.J. ran the blind tests, data analysis, and manuscript review. E.P.: Data analysis and manuscript review. Z.M.B. and M.R.D.: Contribution of HARPS data and manuscript review. G.A.L.C.: Contributed to the discussion of planet formation and manuscript review. I.R., H.R.A.J., A.R., and P.J.A: Writing and review of manuscript. R.P.B.: AAT/PFS data collection and analysis. C.G.T., J.B., B.D.C., S.O., and R.A.W. were AAT/UCLES observers and reviewed the manuscript. J.D.C., F.F., S.A.S., and J.T. were PFS observers. G.A.E.: Writing of the observing proposal, data analysis, and writing of the manuscript. **Competing interests:** There are no competing interests to declare. **Data and materials availability:** The reduced RVs and photometric data are provided in data S1. Our HARPS raw data are available in the ESO archive (http://archive.eso.org) under the program IDs listed in table S1. Reduced HIRES RVs were taken from (*33*). The UCLES data are available from the AAT archive (https://datacentral.org.au/archives/aat/) by searching the coordinates RA '23:05:52h,' Dec '-35:51:11d,' a radius of 300 arcseconds, and dates 1998–2012. The PFS spectra, our dynamical stability simulations, and our Gaussian processes fitting code are available at https://figshare.com/s/d581c1a17536eeb813ea. The TESS photometry was retrieved from https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html, and the All Sky Automated Survey (ASAS) photometry was retrieved from www.astrouw.edu.pl/asas/?page=aasc.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1477/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S7
References (*34*–*70*)
Data S1

9 August 2019; accepted 12 May 2020
10.1126/science.aaz0795

CORONAVIRUS

Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China

Juanjuan Zhang[1], Maria Litvinova[2], Yuxia Liang[1], Yan Wang[1], Wei Wang[1], Shanlu Zhao[3], Qianhui Wu[1], Stefano Merler[4], Cécile Viboud[5], Alessandro Vespignani[6,2], Marco Ajelli[4]*†, Hongjie Yu[1]*†

Intense nonpharmaceutical interventions were put in place in China to stop transmission of the novel coronavirus disease 2019 (COVID-19). As transmission intensifies in other countries, the interplay between age, contact patterns, social distancing, susceptibility to infection, and COVID-19 dynamics remains unclear. To answer these questions, we analyze contact survey data for Wuhan and Shanghai before and during the outbreak and contact-tracing information from Hunan province. Daily contacts were reduced seven- to eightfold during the COVID-19 social distancing period, with most interactions restricted to the household. We find that children 0 to 14 years of age are less susceptible to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection than adults 15 to 64 years of age (odds ratio 0.34, 95% confidence interval 0.24 to 0.49), whereas individuals more than 65 years of age are more susceptible to infection (odds ratio 1.47, 95% confidence interval 1.12 to 1.92). Based on these data, we built a transmission model to study the impact of social distancing and school closure on transmission. We find that social distancing alone, as implemented in China during the outbreak, is sufficient to control COVID-19. Although proactive school closures cannot interrupt transmission on their own, they can reduce peak incidence by 40 to 60% and delay the epidemic.

The novel coronavirus disease 2019 (COVID-19) epidemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) began in Wuhan City, China, in December 2019 and quickly spread globally, with 2,063,161 cases reported in 185 countries or regions as of 16 April 2020 (*1*). A total of 82,692 cases of COVID-19, including 4632 deaths, have been reported in mainland China, including 50,333 cases in Wuhan City and 628 cases in Shanghai City (*2*). The epidemic in Wuhan and in the rest of China subsided after implementation of strict containment measures and movement restrictions, with recent cases originating from travel (*3*). However, key questions remain about the age profile of susceptibility to infection, how social distancing alters age-specific contact patterns, and how these factors interact to affect transmission. These questions are relevant to the choice of control policies for governments and policy-makers around the world. In this study, we evaluate changes in mixing patterns linked to social distancing by collecting contact data in the midst of the epidemic in Wuhan and Shanghai. We also estimate age differences in susceptibility to infection based on contact-tracing data gathered by the Hunan Provincial Center for Disease Control and Prevention (CDC), China. Based on these empirical data, we developed a mathematical disease transmission model to disentangle how transmission is affected by age differences in the biology of COVID-19 infection and altered mixing patterns owing to social distancing. Additionally, we project the impact of social distancing and school closure on COVID-19 transmission.

To estimate changes in age-mixing patterns associated with COVID-19 interventions, we performed contact surveys in two cities: Wuhan, the epicenter of the outbreak, and Shanghai, one of the largest and most densely populated cities in southeast China. Shanghai experienced extensive importation of COVID-19 cases from Wuhan as well as local transmission (*4*). The surveys were conducted from 1 February 2020 to 10 February 2020, as transmission of COVID-19 peaked across China and stringent interventions were put in place. Participants in Wuhan were asked to complete a questionnaire describing their contact behavior (*5*, *6*) on two different days: (i) a regular weekday between 24 December 2019 and 30 December 2019, before the COVID-19 outbreak was officially recognized by the Wuhan Municipal Health Commission (used as baseline); and (ii) the day before the interview (outbreak period). Participants in Shanghai were asked to complete the same questionnaire used for Wuhan but only report contacts for the outbreak period. For the baseline period in Shanghai, we relied on a survey conducted in 2017–2018 that followed the same design (*7*). In these surveys, a contact was defined as either a two-way conversation involving three or more words in the physical presence of another person or a direct physical contact (e.g., a handshake). Details are given in the supplementary materials (SM, sections 1 and 2).

We analyzed a total of 1245 contacts reported by 636 study participants in Wuhan and 1296 contacts reported by 557 participants in Shanghai. In Wuhan, the average daily number of contacts per participant was significantly reduced, from 14.6 for the baseline period (mean contacts weighted by age structure: 14.0) to 2.0 for the outbreak period (mean contacts weighted by age structure: 1.9) ($p < 0.001$). The reduction in contacts was significant for all stratifications by sex, age group, type of profession, and household size (Table 1). A larger reduction was observed in Shanghai, where the average daily number of contacts decreased from 18.8 (mean contacts weighted by age structure: 19.8) to 2.3 (mean contacts weighted by age structure: 2.1). Although an average individual in Shanghai reported more contacts than one in Wuhan on a regular weekday, this difference essentially disappeared during the COVID-19 outbreak period. A similar decrease in the number of contacts was found in the United Kingdom during the COVID-19 lockdown period (*8*).

The typical features of age-mixing patterns (*6*, *7*) emerge in Wuhan and Shanghai when we consider the baseline period (Fig. 1, A and D). These features can be illustrated in the form of age-stratified contact matrices (provided as ready-to-use tables in the SM, section 3.6), where each cell represents the average number of contacts that an individual has with other individuals, stratified by age groups. The bottom left corner of the matrix, corresponding to contacts between school-age children, is where the largest number of contacts is recorded. The contribution of contacts in the workplace is visible in the central part of the matrix, and the three diagonals (from bottom left to top right) represent contacts between household members. By contrast, for the outbreak period when strict social distancing policies were in place, many of the above-mentioned features disappear, essentially leaving the sole contribution of household mixing (Fig. 1, B and E). In particular, assortative contacts between school-age individuals are fully removed, as illustrated by differencing baseline and outbreak matrices (Fig. 1, C and F). Overall, contacts during the outbreak mostly occurred at home with household members (94.1% in Wuhan and 78.5% in Shanghai). Thus, the outbreak contact matrix nearly coincides with the within-household contact matrix in both study sites, and the pattern of assortativity by age observed for regular days almost entirely

[1]School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China. [2]ISI Foundation, Turin, Italy. [3]Hunan Provincial Center for Disease Control and Prevention, Changsha, China. [4]Bruno Kessler Foundation, Trento, Italy. [5]Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA. [6]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: marco.ajelli@gmail.com (M.A.); yhj@fudan.edu.cn (H.Y.)

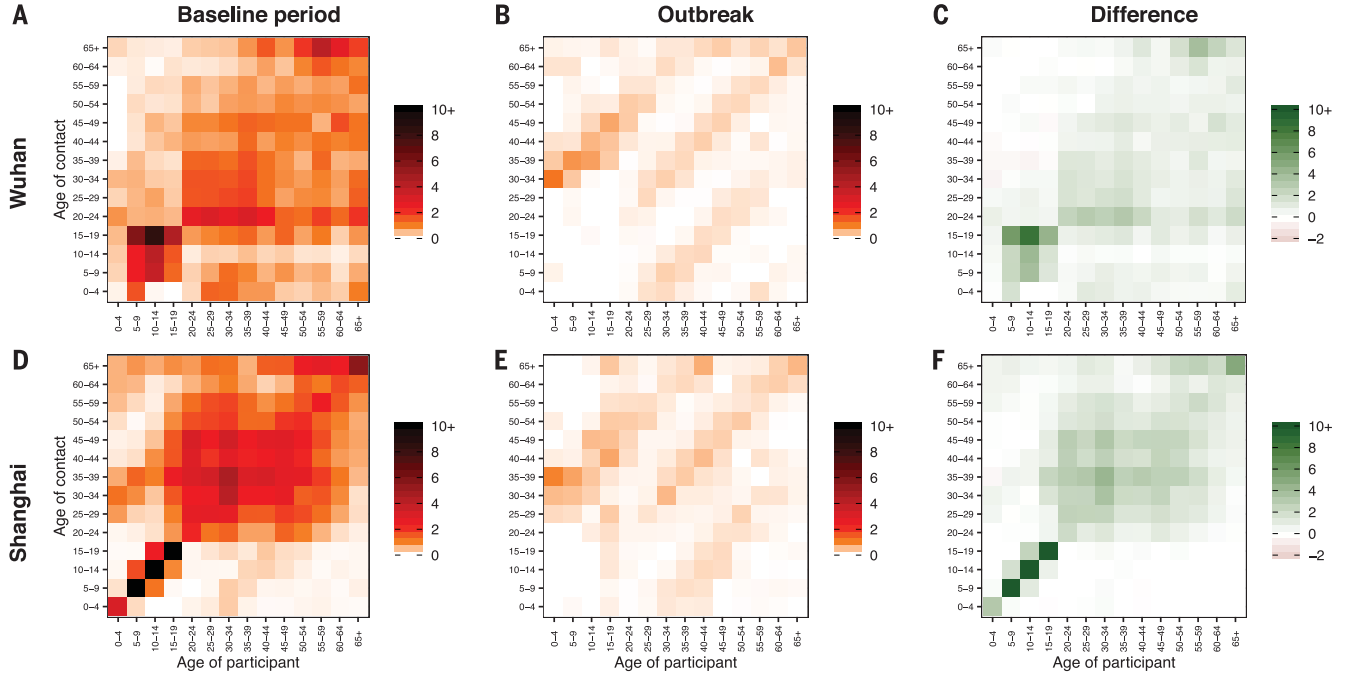


Fig. 1. Contact matrices by age. (**A**) Baseline period contact matrix for Wuhan (regular weekday only). Each cell of the matrix represents the mean number of contacts that an individual in a given age group has with other individuals, stratified by age groups. The color intensity represents the number of contacts. To construct the matrix, we performed bootstrap sampling with replacement of survey participants weighted by the age distribution of the actual population of Wuhan. Every cell of the matrix represents an average over 100 bootstrapped realizations. (**B**) Same as (A), but for the outbreak contact matrix for Wuhan. (**C**) Difference between the baseline period contact matrix and the outbreak contact matrix in Wuhan. (**D**) Same as (A), but for Shanghai. (**E** and **F**) Same as (B) and (C), but for Shanghai.

disappears (SM, section 3.6). These findings are consistent with trends in within-city mobility data, which indicate an 86.9% drop in Wuhan and 74.5% drop in Shanghai between early January and early February (see SM, section 4). Such a large decrease in internal mobility is consistent with most of the contacts occurring in the household during the outbreak period. Of note, the strict social distancing measures implemented in Wuhan and Shanghai did not entirely zero out contacts in the workplace, because essential workers continued to perform their activities (as observed in our data; see SM, section 3.5).

The estimated mixing patterns are based on self-reported contacts that can thus be affected by various biases. In particular, reported contacts for the baseline period in Wuhan may be prone to recall bias because contacts were assessed retrospectively. Further, because of the retrospective nature of the baseline survey in Wuhan, we were unable to account for the lower number of contacts during weekends. The more complete data from Shanghai did not suffer recall bias and allowed us to weight contacts for weekdays and weekends; sensitivity analyses suggest that this has little impact on results (SM, section 8.3). Another possible bias is that survey participants may have felt pressure to minimize reported contacts that occurred during the outbreak, given that social distancing was in place and strictly enforced by the government, even if the anonymity and confidentiality of the survey were emphasized. However, results are robust to inflating reported contacts outside of the home severalfold, suggesting that these compliance and social acceptability biases linked to the outbreak period do not affect our main findings (SM, section 8.2). Another caveat is that in parallel to population-level social distancing measures, case-based interventions were implemented and could have affected contacts, including rapid isolation of confirmed and suspected cases and quarantine of close contacts for 14 days. However, only a small portion of the population in the two study sites was affected by contact tracing and quarantine, thus having little to no effect on average contact patterns in the general population.

Next, to understand the interplay between social distancing interventions, changes in human mixing patterns, and outbreak dynamics, we need to consider potential age differences in susceptibility to infection. This is currently a topic of debate, because little information on the age profile of asymptomatic cases is available (*9*, *10*). To this aim, we analyzed COVID-19 contact-tracing information gleaned from detailed epidemiological field investigations conducted by the Hunan CDC (SM, section 5). Briefly, all close contacts of COVID-19 cases reported in Hunan province were placed under medical observation for 14 days and were tested using real-time reverse transcription polymerase chain reaction (RT-PCR). Those who tested positive were considered as SARS-CoV-2 infections. We estimated the odds ratios (ORs) for a contact of a certain age group to be infected, relative to a reference age group. We performed generalized linear mixed model regression to account for clustering and potential correlation structure of contacts exposed to the same index case (e.g., in the household). We included the age group and gender of a contact, type of contact, and whether the contact traveled to Hubei or Wuhan as regression covariates (SM, section 5). We found that susceptibility to SARS-CoV-2 infection increased with age. Young individuals (aged 0 to 14 years) had a lower risk of infection than individuals aged 15 to 64 years {OR = 0.34 [95% confidence interval (CI): 0.24 to 0.49], $p < 0.0001$}. By contrast, older individuals aged 65 years and older had a higher risk of infection than adults aged 15 to 64 years [OR = 1.47 (95% CI: 1.12 to 1.92), $p = 0.005$]. These findings are in contrast with a previous study in Shenzhen, where susceptibility to infection did not change with age (*9*).

Next, we explore how our data can inform control strategies for COVID-19. A key parameter regulating the dynamics of an epidemic is the basic reproduction number (R_0), which

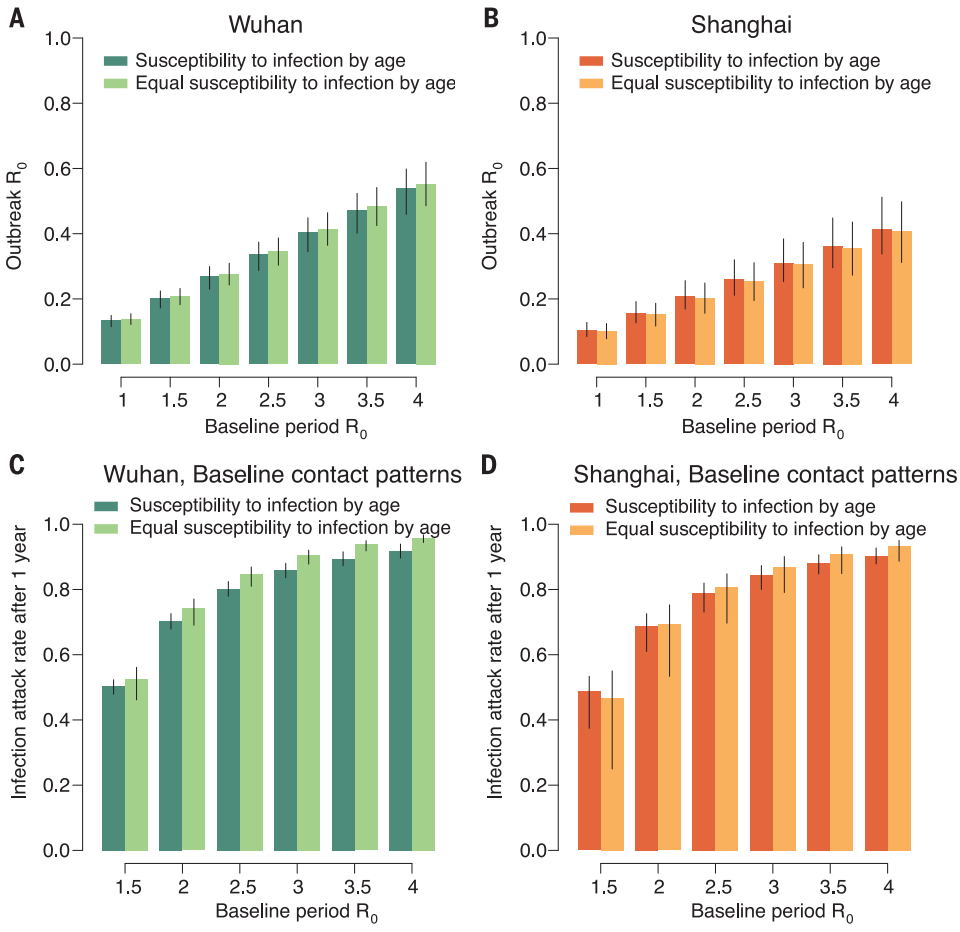


Fig. 2. Effect of contact patterns on the epidemic spread. (**A**) Estimated R_0 during the outbreak (mean and 95% CI), as a function of baseline R_0 (i.e., that derived by using the contact matrix estimated for the baseline period). The figure refers to Wuhan and includes both the scenario accounting for the estimated susceptibility to infection by age and the scenario where we assume that all individuals are equally susceptible to infection. The distribution of the transmission rate is estimated through the next-generation matrix approach by using 100 bootstrapped contact matrices for the baseline period to obtain the desired R_0 values. We then use the estimated distribution of the transmission rate and the bootstrapped outbreak contact matrices to estimate R_0 for the outbreak period. The 95% CIs account for the uncertainty on the distribution of the transmission rate, mixing patterns, and susceptibility to infection by age. (**B**) Same as (A), but for Shanghai. (**C**) Infection attack rate 1 year after the initial case of COVID-19 (mean and 95% CI) as a function of the baseline R_0. The estimates are made by simulating the SIR transmission model (see SM) using the contact matrix for the baseline period and considering the estimated susceptibility to infection by age and assuming that all individuals are equally susceptible to infection. The 95% CIs account for the uncertainty on the mixing patterns and susceptibility to infection by age. (**D**) Same as (C), but for Shanghai.

corresponds to the average number of secondary cases generated by an index case in a fully susceptible population. We estimated the impact of interventions on R_0, relying on our age-specific estimates of susceptibility to infection and contact patterns before and during interventions. We used the next-generation matrix approach to quantify changes in R_0 (*11*) (SM, section 6). Additionally, to illustrate the impact of age-mixing patterns on the dynamics of the epidemic, we developed a simple SIR model of SARS-CoV-2 transmission (SM, section 6). In the model, the population is divided into three epidemiological categories: susceptible, infectious, and removed (either recovered or deceased individuals), stratified by 14 age groups. Susceptible individuals can become infectious after contact with an infectious individual according to the estimated age-specific susceptibility to infection. The rate at which contacts occur is determined by the estimated mixing patterns of each age group. The mean time interval between two consecutive generations of cases was taken to be 5.1 days, assuming it aligns with the mean of the serial interval reported by Zhang *et al.* (*3*).

In the early phases of COVID-19 spread in Wuhan, before interventions were put in place, R_0 values were estimated to range between 2.0 and 3.5 (*12*–*18*). In this analysis, we extended this range from 1 to 4 for the baseline period (i.e., before interventions). We find that the considerable changes of mixing patterns observed in Wuhan and Shanghai during the social distancing period led to a drastic decrease in R_0 (Fig. 2). When we consider contact matrices representing the outbreak period, keeping the same baseline disease transmissibility as in the preintervention period, the reproductive number drops well below the epidemic threshold in Wuhan (Fig. 2A) and Shanghai (Fig. 2B). This finding is robust to relaxing assumptions about age differences in susceptibility to infection; the epidemic is still well controlled if SARS-CoV-2 infection is assumed to be equally likely in all age groups (Fig. 2, A and B). We also performed sensitivity analyses regarding possible recall and compliance biases of self-reported contacts as well as the definition of contact (i.e., considering only contacts lasting more than 5 min). The results are consistent with those reported here (SM, section 8).

In an uncontrolled epidemic (without intervention measures, travel restrictions, or spontaneous behavioral responses of the population) and for R_0 in the range of 2 to 3, we estimate the mean infection attack rate to be in the range 53 to 92% after a year of SARS-CoV-2 circulation, with slight variation between Wuhan (Fig. 2C) and Shanghai (Fig. 2D). These estimates should be considered as an upper bound of the infection attack rate because they are based on a compartmental model that does not account for high clustering of contacts (e.g., repeated contacts among household members). If we consider a scenario in which social distancing measures are implemented early on, as the new virus emerges, the estimated R_0 remains under the epidemic threshold and thus the epidemic cannot take off in either location. Furthermore, we estimate that the magnitude of interventions implemented in Wuhan and Shanghai would have been enough to block transmission for an R_0 before the interventions of up to ~6 in Wuhan and ~7.8 in Shanghai.

Next, we use the model to estimate the impact of preemptive mass school closure. We considered two different contact pattern scenarios, based on data from Shanghai: contacts estimated during vacation periods (*7*) and contacts estimated during regular weekdays, after all contacts occurring in school settings have been removed (*7*). Both scenarios represent a simplification of a school closure strategy. Indeed, school closures in response to the

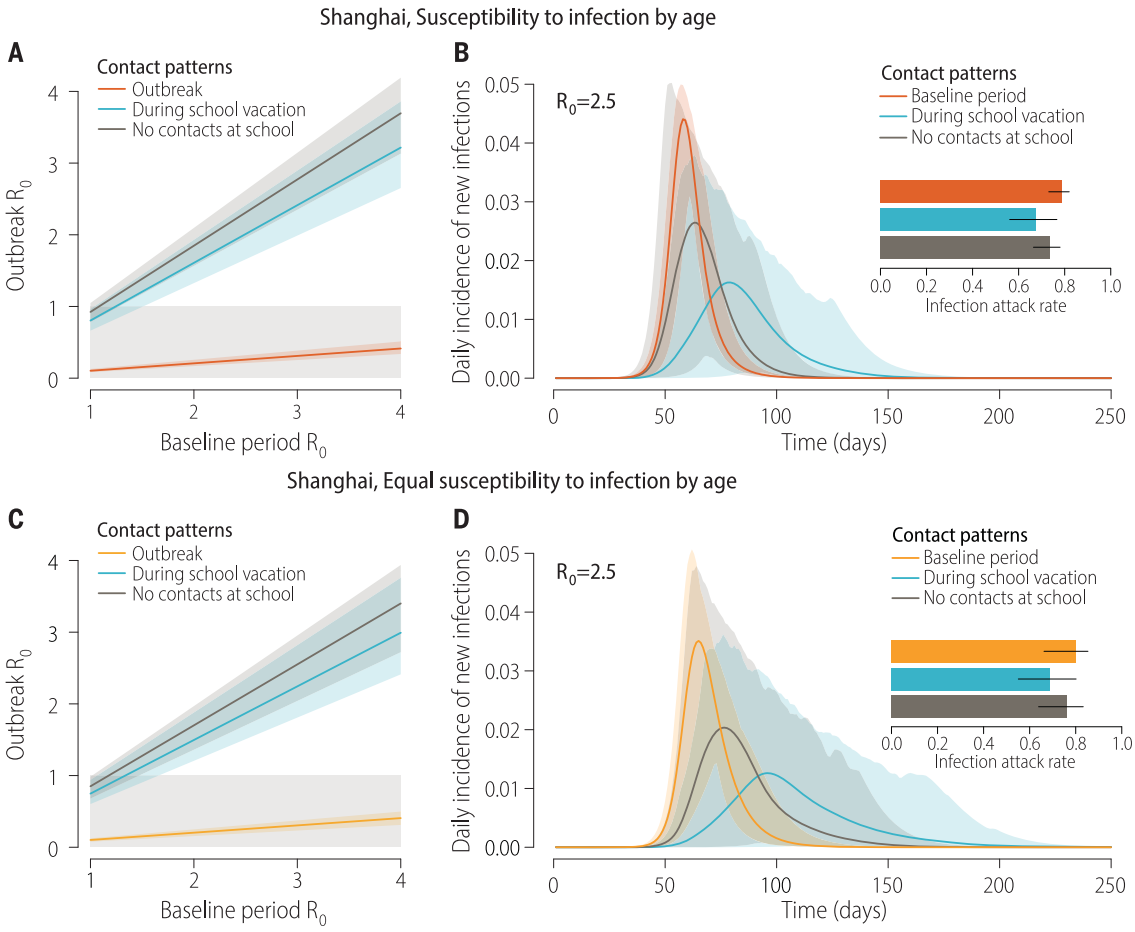


Fig. 3. Effect of limiting school contacts on the epidemic spread. (**A**) Estimated R_0 during the outbreak (mean and 95% CI), as a function of baseline R_0 (i.e., that derived by using the contact matrix estimated for the baseline period). The figure refers to Shanghai and the scenario accounting for the estimated susceptibility to infection by age. Three contact patterns are considered: (i) as estimated during the COVID-19 outbreak, (ii) as estimated during school vacations (*7*), and (iii) as estimated for the baseline period, but suppressing all contacts at school. (**B**) Daily incidence of new SARS-CoV-2 infections (mean and 95% CI), as estimated by the SIR model, assuming age-specific susceptibility to infection (see SM). Three mixing patterns are considered: (i) as estimated for the baseline period, (ii) as estimated during school vacations (*7*), and (iii) as estimated for the baseline period, but suppressing all contacts at school. The inset shows the infection attack rate 1 year after the introduction of the first COVID-19 case (mean and 95% CI). (**C**) Same as (A), but assuming equal susceptibility to infection by age. (**D**) Same as (B), but assuming equal susceptibility to infection by age.

COVID-19 pandemic in China have entailed interruption of all educational on-site services. However, mixing patterns measured during school vacations indicate that a fraction of children still attend additional educational activities, as is typical in Chinese cities. On the other hand, when removing all contacts in the school setting, we do not consider potential trickle-down effects on the mixing patterns of other age groups; for instance, parents may need to leave work to take care of school-age children. Our modeling approach indicates that limiting contact patterns to those observed during vacations would interrupt transmission for baseline R_0 up to 1.5 (Fig. 3, A and C). Removing all school contacts would do the same for baseline R_0 up to 1.2. If we apply these interventions to a COVID-19 scenario, assuming a baseline R_0 of 2 to 3.5, we can achieve a noticeable decrease in infection attack rate and peak incidence and a delay in the epidemic, but transmission is not interrupted (Fig. 3, B and D). For instance, for a baseline R_0 of 2.5 and assuming a vacation mixing pattern, the mean peak daily incidence is reduced by about 64%. In the corresponding scenario where school contacts are removed, we estimate a reduction of about 42%. Overall, school-based closure policies are not sufficient to entirely prevent a COVID-19 outbreak, but they can affect disease dynamics and hence hospital surge capacity. It is important to stress that individuals aged 5 to 19 years in Shanghai represent 9.5% of the population (*19*), markedly lower than the mean in China [16.8% (*19*)] and other countries [including Western countries; e.g., 19.7% in the United States (*20*)].

The results of this study should be considered in light of the following limitations. In our simulation model, we estimated the effect of social distancing alone; combining social distancing with other interventions would have a synergistic effect to even further reduce transmission. It is likely that population-wide social distancing, case-based strategies, and decontamination efforts all contributed to achieve control in Wuhan and Shanghai, and their effect is difficult to separate out in retrospective observational studies. Our estimates of age differences in susceptibility to infection are based on active testing of 7375 contacts of 136 confirmed index cases. These data suffer from the usual difficulties inherent to the reconstruction of epidemiological links and detection of index cases. Contact data are useful, but seroepidemiology studies will be essential to fully resolve population susceptibility profiles to SARS-CoV-2 infection and disease. Although the age patterns of contacts were similar in the two study locations during the COVID-19 outbreak period, these patterns may not be fully representative of other locations in China and abroad, where social distancing measures may differ. Because reliable estimates of the contribution of asymptomatic SARS-CoV-2 infections to transmission are still lacking, we did not explicitly model differences between symptomatic and asymptomatic individuals. We considered a serial interval of 5.1 days (*3*), based on a prior estimate from China, at a time when case-based and contact-tracing intervention measures were in place, which tends to shorten the interval between successive cases. However, this choice does not

Table 1. Number of contacts by demographic characteristics and location. *N* is the number of participants who provided non-missing contact data.

Characteristics	Wuhan					Shanghai				
	Baseline period		COVID-19 outbreak		Difference§	Baseline period		COVID-19 outbreak		Difference§
	N (%)†	Mean (95% CI‡)	*N* (%)†	Mean (95% CI‡)		*N* (%)	Mean (95% CI‡)	*N* (%)	Mean (95% CI‡)	
Overall	624 (100.0)	14.6 (12.9, 16.3)	627 (100.0)	2 (1.9, 2.1)	12.6***	965 (100.0)	18.8 (17.8, 19.8)	557 (100.0)	2.3 (2, 2.8)	16.4***
Sex										
Male	300 (48.1)	14.5 (12.2, 17.1)	301 (48)	1.8 (1.7, 2)	12.6***	474 (49.1)	19 (16.9, 21)	286 (51.3)	2.1 (1.9, 2.4)	16.9***
Female	324 (51.9)	14.7 (12.5, 17.1)	326 (52)	2.1 (2, 2.3)	12.5***	491 (50.9)	18.5 (16.8, 20.4)	271 (48.7)	2.6 (2.1, 3.6)	16***
Age group										
0–6 years	12 (1.9)	8.6 (3.4, 17.4)	12 (1.9)	2.2 (1.7, 2.8)	6.4***	88 (9.1)	11.6 (9.2, 14.3)	14 (2.5)	1.9 (1.7, 2.2)	9.7***
7–19 years	79 (12.7)	16.2 (12.7, 19.6)	79 (12.6)	2.1 (2, 2.2)	14.1***	141 (14.6)	27 (23.1, 30.7)	55 (9.9)	2.6 (2, 3.4)	24.5***
20–39 years	254 (40.7)	15.3 (12.8, 18)	256 (40.8)	2.1 (1.9, 2.2)	13.2***	236 (24.5)	22.4 (19.8, 25.9)	254 (45.6)	2.2 (2, 2.5)	20.2***
40–59 years	221 (35.4)	13.8 (11.4, 16.7)	220 (35.1)	2 (1.8, 2.2)	11.8***	233 (24.1)	19.9 (17.7, 23.3)	160 (28.7)	2.8 (2, 4.1)	17.1***
≥60 years	58 (9.3)	13.9 (7.9, 20.7)	60 (9.6)	1.4 (1.2, 1.7)	11.6***	267 (27.7)	12.6 (10.8, 14.7)	74 (13.3)	1.6 (1.3, 1.8)	11***
Type of profession										
Preschool	12 (1.9)	8.6 (3.4, 17.4)	12 (1.9)	2.2 (1.7, 2.8)	6.4***	79 (8.2)	10.4 (8, 13.3)	14 (2.5)	1.9 (1.7, 2.1)	8.5***
Student	107 (17.1)	14.6 (11.4, 18.2)	107 (17.1)	2.1 (2, 2.3)	12.5***	173 (17.9)	26.2 (23.1, 29.2)	71 (12.7)	2.5 (2, 3.4)	23.7***
Employed	391 (62.7)	15.4 (13.4, 17.4)	390 (62.2)	2.1 (1.9, 2.2)	13.2***	400 (41.5)	22.5 (20.7, 24.4)	354 (63.6)	2.5 (2.1, 3.2)	20***
Working-age not in the labor force	30 (4.8)	14.1 (5.7, 24.2)	31 (4.9)	1.8 (1.4, 2.4)	12.2***	29 (3)	14.5 (7.8, 24.2)	24 (4.3)	1.8 (1.3, 2.4)	12.6***
Retired	84 (13.5)	12.1 (7.2, 17.4)	87 (13.9)	1.5 (1.3, 1.7)	10.6***	278 (28.8)	11.8 (10.2, 13.2)	94 (16.9)	1.6 (1.3, 1.8)	10.2***
Household size										
1	45 (7.2)	10.5 (5.3, 17.2)	45 (7.2)	0.6 (0.1, 1.5)	9.9***	35 (3.6)	15.2 (10.1, 21.1)	61 (11)	0.3 (0.1, 0.5)	14.9***
2	73 (11.7)	12.6 (8.2, 18.3)	76 (12.1)	1.1 (1, 1.2)	11.5***	244 (25.3)	14.5 (12.7, 16.7)	138 (24.8)	1.4 (1.1, 1.7)	13.1***
3	282 (45.2)	14.8 (12.8, 17.3)	283 (45.1)	1.9 (1.8, 2)	13***	432 (44.8)	20.3 (17.7, 22.4)	216 (38.8)	2.2 (2, 2.3)	18.1***
4	133 (21.3)	11.9 (9.3, 15)	132 (21.1)	2.3 (2.2, 2.5)	9.6***	117 (12.1)	20.3 (16.5, 23.8)	78 (14)	3 (2.8, 3.3)	17.3***
≥5	91 (14.6)	21.5 (16.2, 27.3)	91 (14.5)	3.2 (2.9, 3.4)	17.8***	137 (14.2)	21.4 (18.2, 27)	64 (11.5)	5.9 (4, 9.9)	15.5***

†Can differ from total sample size (*N* = 636) because it also includes participants who had not recorded contacts during the baseline period or during the COVID-19 outbreak. Note that reduced denominators indicate missing data. Percentages may not total 100 because of rounding. ‡The 95% CIs on the mean are calculated by bootstrap sampling. §Difference is calculated by the subtraction of the number of contacts during the outbreak from the number of contacts during the baseline period. *p* values are taken from a negative binomial regression with a single binary variable distinguishing the baseline period from the outbreak. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

affect the estimated changes in reproduction number between the baseline and outbreak periods. Modeling results may underestimate the effect of social distancing interventions because our results concentrate on the number of contacts and ignore the type of social interactions (e.g., increased distance between individuals while in contact or use of a face mask), which may have changed owing to increased awareness of the population (*21*, *22*). Finally, it is worth noting that our school closure simulations are not meant to formulate a full intervention strategy, which would require identification of epidemic triggers to initiate closures and evaluation of different durations of intervention (*6*). Nonetheless, our modeling exercise provides an indication of the possible impact of a nationwide preemptive strategy on the infection attack rate and peak incidence. To generalize these findings to other contexts, location-specific age-mixing patterns and population structures should be considered. Perhaps most importantly, strict lockdown strategies of the kind implemented in Wuhan, Shanghai, and other regions of the world are extremely disruptive economically and mentally, and

more targeted approaches to block transmission are preferable in the long run. We do not necessarily endorse blunt lockdown policies here; we merely describe their impact on COVID-19 transmission based on the Chinese experience.

Our study provides evidence that the interventions put in place in Wuhan and Shanghai, and the resulting changes in human behavior, drastically decreased daily contacts, essentially reducing them to household interactions. This led to a dramatic reduction of SARS-CoV-2 transmission. As lockdown measures are put in place in other locations, human mixing patterns in the outbreak period could be captured by data on within-household contacts, which are available for several countries around the world (*5*–*7*, *23*–*25*). Moving forward, it will be particularly important to design targeted strategies for long-term control of COVID-19, including school- and work-based control strategies, along with large-scale testing and contact tracing (*26*–*28*). Research should concentrate on refining age-specific estimates of susceptibility to infection, disease, and infectiousness, which are instrumental to evaluating the impact of these strategies.

REFERENCES AND NOTES

1. Johns Hopkins University, COVID-19 Dashboard (2020); https://coronavirus.jhu.edu/map.html [accessed 16 April 2020].
2. Chinese Center for Disease Control and Prevention, Update on COVID-19 as of 24:00 on April 16, 2020 (2020); http://2019ncov.chinacdc.cn/2019-nCoV/ [accessed 17 April 2020].
3. J. Zhang *et al.*, *Lancet Infect. Dis.* S1473-3099(20)30230-9 (2020).
4. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, *China CDC Weekly* **2**, 113–122 (2020).
5. J. Mossong *et al.*, *PLOS Med.* **5**, e74 (2008).
6. M. Litvinova, Q. H. Liu, E. S. Kulikov, M. Ajelli, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13174–13181 (2019).
7. J. Zhang *et al.*, *Sci. Rep.* **9**, 15141 (2019).
8. C. I. Jarvis *et al.*, medRxiv 2020.03.31.20049023 [Preprint]. 3 April 2020; https://doi.org/10.1101/2020.03.31.20049023.
9. Q. Bi *et al.*, medRxiv 2020.03.03.20028423 [Preprint]. 27 March 2020; https://doi.org/10.1101/2020.03.03.20028423.
10. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, medRxiv 2020.02.20.20025866 [Preprint]. 6 March 2020; https://doi.org/10.1101/2020.02.20.20025866.
11. O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, *J. Math. Biol.* **28**, 365–382 (1990).
12. J. T. Wu, K. Leung, G. M. Leung, *Lancet* **395**, 689–697 (2020).
13. J. M. Read, J. R. Bridgen, D. A. Cummings, A. Ho, C. P. Jewell, medRxiv 2020.01.23.20018549 [Preprint]. 28 January 2020; https://doi.org/10.1101/2020.01.23.20018549.
14. Q. Li *et al.*, *N. Engl. J. Med.* **382**, 1199–1207 (2020).
15. S. Abbott, J. Hellewell, J. Munday, S. Funk; CMMID nCoVworking group, *Wellcome Open Res.* **5**, 17 (2020).
16. M. Chinazzi *et al.*, *Science* **368**, 395–400 (2020).
17. I. Natsuko *et al.*, "Report 3: Transmissibility of 2019-nCoV" (Imperial College London, 2020); www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-transmissibility-25-01-2020.pdf.
18. World Health Organization, "Report of the WHO-China Joint mission on coronavirus disease 2019 (COVID-19)," (WHO, 2020); www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf [accessed 11 March 2020].
19. National Bureau of Statistics, China census data (2020); www.stats.gov.cn/ [accessed 1 March 2020].
20. United Nations, World population prospects 2019 (2019); https://population.un.org/wpp/Download/Standard/Population/ [accessed 15 April 2019].
21. B. J. Cowling *et al.*, medRxiv 2020.03.12.20034660 [Preprint]. 16 March 2020; https://doi.org/10.1101/2020.03.12.20034660.
22. M. Qian *et al.*, medRxiv 2020.02.18.20024448 [Preprint]. 20 February 2020; https://doi.org/10.1101/2020.02.18.20024448.
23. T. Hoang *et al.*, *Epidemiology* **30**, 723–736 (2019).
24. D. Mistry *et al.*, arXiv:2003.01214 [q-bio.PE] (25 February 2020).
25. K. Prem, A. R. Cook, M. Jit, *PLOS Comput. Biol.* **13**, e1005697 (2017).
26. K. Prem *et al.*, *Lancet Public Health* S2468-2667(20)30073-6 (2020).
27. D. L. Heymann, N. Shindo; WHO Scientific and Technical Advisory Group for Infectious Hazards, *Lancet* **395**, 542–545 (2020).
28. S. Riley *et al.*, *Science* **300**, 1961–1966 (2003).
29. J. Zhang *et al.*, Zenodo (2020); doi:10.5281/zenodo.3775672.

ACKNOWLEDGMENTS

The authors acknowledge B. J. Cowling from the University of Hong Kong and C. L. Gilbert from the Johns Hopkins University for their helpful comments on the manuscript and N. Samay for her assistance in preparing the figures. This article does not necessarily represent the views of the NIH or the U.S. government. **Funding:** H.Y. acknowledges financial support from the National Science Fund for Distinguished Young Scholars (no. 81525023), the Key Emergency Project of Shanghai Science and Technology Committee (no. 20411950100), and the National Science and Technology Major Project of China (nos. 2018ZX10201001-010, 2018ZX10713001-007, and 2017ZX10103009-005). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. S.M. and M.A. acknowledge financial support from the European Commission H2020 MOOD project. **Author contributions:** M.A. and H.Y. are joint senior authors. H.Y. and M.A. designed the experiments. J.Z., Y.L., S.Z., and Q.W. collected data. J.Z., M.L., Y.W., W.W., Y.L., Q.W., and M.A. analyzed the data. J.Z., M.L., S.M., C.V., A.V., M.A., and H.Y. interpreted the results. J.Z., M.L., C.V., M.A., and H.Y. wrote the manuscript. A.V. edited the manuscript. **Competing interests:** A.V. has received funding from Metabiota Inc. H.Y. has received research funding from Sanofi Pasteur, GlaxoSmithKline, Yichang HEC Changjiang Pharmaceutical Company, and Shanghai Roche Pharmaceutical Company. **Ethics statement:** Ethics approval was obtained from the institutional review board of the School of Public Health, Fudan University (IRB#2020-01-0801). Verbal informed consent was obtained from all subjects (from a parent or guardian if the participant was under 18 years of age). **Data and materials availability:** All data and code are available in the main text or the supplementary materials or at Zenodo (*29*).
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/368/6498/1481/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S15
References (*30*–*41*)
MDAR Reproducibility Checklist

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

19 March 2020; accepted 27 April 2020
Published online 29 April 2020
10.1126/science.abb8001

QUANTUM OPTICS

Metalens-array–based high-dimensional and multiphoton quantum source

Lin Li[1,2,3,4]*, Zexuan Liu[1,5,6]*, Xifeng Ren[7,8]*, Shuming Wang[1,5,6]*†, Vin-Cent Su[9], Mu-Ku Chen[2,3,10], Cheng Hung Chu[3,10], Hsin Yu Kuo[3,10], Biheng Liu[7,8], Wenbo Zang[1,6], Guangcan Guo[7,8], Lijian Zhang[1,5,6]†, Zhenlin Wang[1,6]†, Shining Zhu[1,5,6]†, Din Ping Tsai[2,3,10]†

The development of two-dimensional metasurfaces has shown great potential in quantum-optical technologies because of the excellent flexibility in light-field manipulation. By integrating a metalens array with a nonlinear crystal, we demonstrate a 100-path spontaneous parametric down-conversion photon-pair source in a 10 × 10 array, which shows promise for high-dimensional entanglement and multiphoton-state generation. We demonstrate two-, three- and four-dimensional two-photon path entanglement with different phases encoded by metalenses with fidelities of 98.4, 96.6, and 95.0%, respectively. Furthermore, four-photon and six-photon generation is observed with high indistinguishability of photons generated from different metalenses. Our metalens-array–based quantum photon source is compact, stable, and controllable, indicating a new platform for integrated quantum devices.

Quantum optical systems, benefiting from the fast speed, long coherence time, versatile controllability, and large information capacity of photons, is one of the most attractive physical systems for the investigation of quantum information processing. They are widely used in quantum communication (*1*–*4*), quantum computation and simulation (*5*–*8*), and quantum metrology and sensing (*9*, *10*). With the development of quantum technologies, the requirement for both entanglement dimensionality and photon number increases, demanding large-scale, controllable, and stable quantum photonic sources (*11*, *12*). For example, quantum communication and imaging require photons with high-dimensional entanglement, which can be achieved by using different degrees of freedom of photons, including orbital angular momentum (OAM), time-bin, energy-time, frequency mode, and optical paths. However, none of these meets the requirement of high fidelity with large dimensionality for practical applications. Photonic quantum computation and metrology rely on multiphoton states, which can be synthesized with multiple single-photon sources from spontaneous parametric processes in nonlinear materials or by time-multiplexing the spontaneous emission of quantum dots (*13*), but the largest photon number is limited to ~20 (*14*). Although quantum dots show excellent performance in generating single photons, the spontaneous parametric process continues to be the main method for the generation of high-dimensional and/or multiphoton entangled states (*13*). Recent progress on integrated photonic systems based on such a process provides an ideal platform for large-scale quantum-optical sources (*15*).

Metasurfaces consist of a dense arrangement of dielectric or metallic subwavelength antennas in an ultrathin interface (*16*). By controlling the phase distribution, metasurfaces have been widely used to manipulate the wavefront of a light field (*17*–*23*). Metasurfaces have also found applications in the nonclassical region (*24*–*27*). However, quantum photonic sources based on metasurfaces have not yet been demonstrated.

Here, by using a metalens array, we demonstrate a 100-path spontaneous parametric down-conversion (SPDC) photon source. The 100-path SPDC photon source is realized by combining a metalens array with a 0.5-mm type II β-barium borate (BBO) crystal (Fig. 1, A to D). The metalens array consists of 100 designed metalenses (arranged in a 10 × 10 array), composed of GaN nanopillars fabricated by electron beam lithography (EBL), dry etching, and resist removing (*13*). Each metalens is designed with a uniform focal length of f = 1.1 mm at a working wavelength of 404 nm and an area of 100 μm by 100 μm. The period of the unit element is 200 nm and the height of the nanopillars is 800 nm [design details of the metalens are in the supplementary materials (*13*)]. The image of the focal spot array of the pump laser, with a wavelength of 404 nm (Fig. 1E), shows that the measured focusing efficiency of each metalens has a uniform value of 56.0 ± 6.6%, owing to the precise fabrication (fig. S1).

When a pump laser is incident on the metalens array, a 10 × 10 array of focal spots is formed inside the BBO crystal. Each of the spots can trigger a SPDC process and generate a pair of photons in a probabilistic manner. The phase-matching condition of the BBO crystal is designed to ensure that each photon pair has two well-defined beam-like spatial modes (*28*). The uniform intensity and spatial distributions of these focal spots enable the further realization of high-dimensional entangled and multiphoton quantum photonic sources. As a preliminary test, a vertically polarized diode laser (404 nm, 100 mW) is incident on the metalens–BBO system. An array of 10 × 10 SPDC photon pairs with nearly equal intensities is observed with an electron-multiplying charge-coupled device (EMCCD) (Fig. 1F).

First, the path-encoded quantum entanglement is demonstrated. Each metalens in the system generates a pair of photons in a pair of conjugate spatial modes (figs. S2 and S3). These modes are defined as s_0 to s_{99} for the signal photon (vertically polarized) and as i_0 to i_{99} for the idler photon (horizontally polarized). In the case that only one pair of photons is generated from the source, without knowing which metalens the photons are generated from, the two-photon state can be written as $\frac{1}{10}(|0,0\rangle + |1,1\rangle + |2,2\rangle + \ldots + |99,99\rangle)$, where the number represents the path defined previously. This state is a 100-dimensional path-encoded quantum entangled state. As a demonstration, we verify the quantum entanglement in two, three, and four dimensions by reconstructing the reduced quantum states with quantum state tomography (QST) measurements and discuss the high-dimensional entanglement measurement in the supplementary materials (*13*).

Two-dimensional (2D) entanglement states are analyzed using photon pairs generated by two metalenses, including the adjacent and nonadjacent ones. The QST measurements are performed (fig. S4) and the corresponding density matrix is accurately reconstructed by the maximum likelihood estimation (MLE) method. Figure 2A shows the experimental result of a maximally 2D path-entangled state of $(|00\rangle + |11\rangle)/\sqrt{2}$. The fidelity between the reconstructed state and the ideal maximally entangled state is as high as 0.985, proving that the produced state is very close to the ideal state. In addition to focusing the pump light,

[1]National Laboratory of Solid State Microstructures, School of Physics, College of Engineering and Applied Sciences, Nanjing University, Nanjing, 210093, China. [2]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. [3]Research Center for Applied Sciences, Academia Sinica, Taipei 11529, Taiwan. [4]State Key Laboratory of Precision Spectroscopy, School of Physics and Electronic Science, East China Normal University, Shanghai, 200062, China. [5]Key Laboratory of Intelligent Optical Sensing and Manipulation, Ministry of Education, Nanjing University, Nanjing, 210093, China. [6]Collaborative Innovation Center of Advanced Microstructures, Nanjing, 210093, China. [7]Key Laboratory of Quantum Information, CAS, University of Science and Technology of China, Hefei, 230026, China. [8]Synergetic Innovation Center of Quantum Information & Quantum Physics, University of Science and Technology of China, Hefei, 230026, China. [9]Department of Electrical Engineering, National United University, Miaoli 36003, Taiwan. [10]Department of Physics, National Taiwan University, Taipei 10617, Taiwan.
*These authors contributed equally to this work.
†Corresponding author. Email: wangshuming@nju.edu.cn (S.W.); lijian.zhang@nju.edu.cn (L.Z.); zlwang@nju.edu.cn (Z.W.); zhusn@nju.edu.cn (S.Z.); dinping.tsai@polyu.edu.hk (D.P.T.)

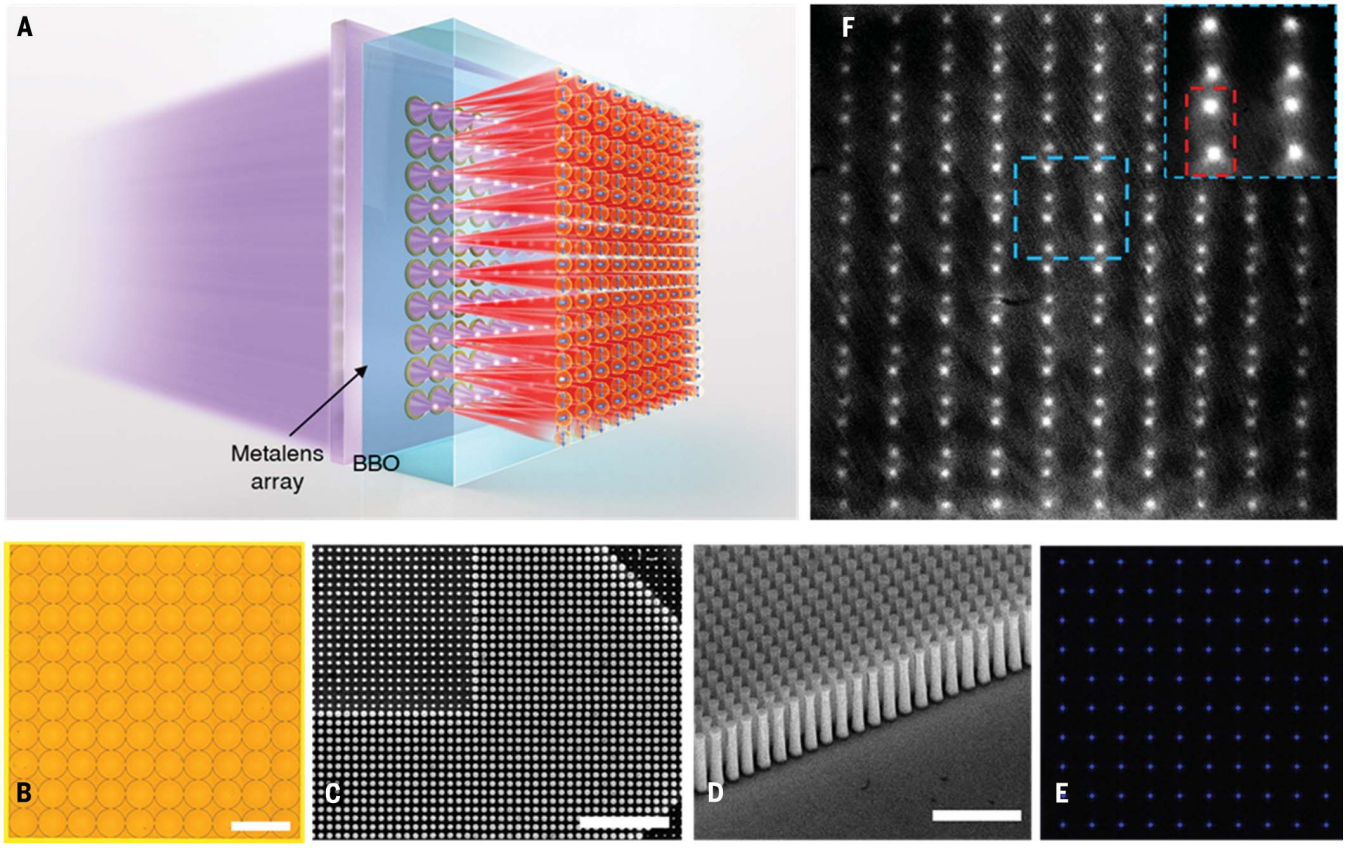


Fig. 1. Schematic and characterization of the quantum metalens array. (**A**) Schematic of the quantum source based on a metalens array. (**B**) Microscopy image of the metalens array. (**C** and **D**) Scanning electron microscopy image of the GaN metalens in (C) top view and (D) side view. (**E**) Microscopy image of the focal plane of the pump light. (**F**) Image of the SPDC photon pair array recorded by EMCCD. Inset: Magnified image of the area indicated by the blue dashed box. The red dashed box shows one pair of photons generated by one metalens. Scale bars: (B) 200 μm, (C) 2 μm, (D) 1 μm.

the metalens is able to encode its phase, which in turn controls the relative phase between different terms of the generated quantum state. We design the metalenses with different phase differences ($\Delta\varphi = \pi/2, \pi, 3\pi/2$) between two adjacent metalenses and analyze the 2D entanglement states generated with them (figs. S5 and S6). Figure 2, B to D, are the reconstructed density matrixes of the prepared states, with the phase differences of ($\Delta\varphi = \pi/2, \pi, 3\pi/2$), respectively. All these density matrixes show good agreement with the ideal corresponding states of $(|00\rangle + i|11\rangle)/\sqrt{2}$, $(|00\rangle - |11\rangle)/\sqrt{2}$ and $(|00\rangle - i|11\rangle)/\sqrt{2}$, respectively. Without loss of generality, we measure six 2D entangled states, with respect to adjacent metalenses pairs at random locations, for each phase difference and estimate them by the QST measurements. All the results show high fidelities, with an average value of 0.979, to the maximally entangled states and a fidelity of 0.984 to the closest pure entangled states (fig. S6). Other results with nonadjacent metalens pairs are shown in fig. S7. The fidelities to the maximally entangled states are all above 0.96 (*13*). These high fidelities confirm that we can reliably control the generated states by metananostructure.

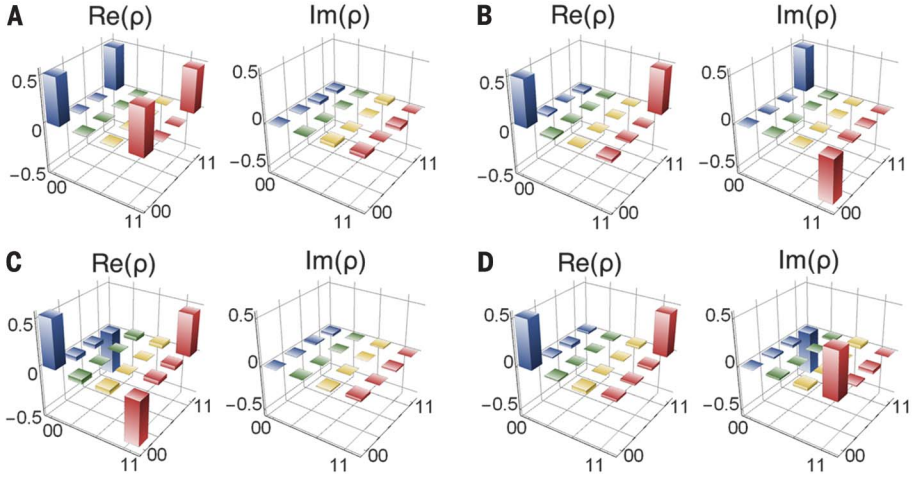


Fig. 2. Characterization of entangled quantum states. (**A** to **D**) Typical reconstructed density matrixes by QST measurements with metalens phase differences of $\Delta\varphi$ = 0, $\pi/2$, π, and $3\pi/2$, respectively. The quantum state fidelities are 0.985, 0.987, 0.987, and 0.984 to the maximally entangled states, respectively.

Quantum entangled states with higher dimensions are further characterized. Figure 3, B and D, are typical 3D and 4D density matrixes reconstructed by QST measurements and correspond to the configuration of the photon pair sources in the dashed boxes in Fig. 3, A and C, respectively. In these high-dimensional cases, the beam displacer (BD)–based measurement setup becomes considerably more complex, as shown in the inset of fig. S4. There are not

enough degrees of freedom to balance the additional accumulation of experimental fluctuations introduced by the BD-based spatial mode combing and analyzing systems. It is a challenge to reconstruct ideal density matrixes of the maximally entangled states. Here, the reconstructed results show very high fidelities to the pure quantum states. The fidelities to the estimated pure states are 0.966 for the 3D entangled state (Fig. 3B) and 0.950 for the 4D entangled state (Fig. 3D), and the fidelities to the maximally entangled states are 0.965 and 0.911, respectively. These results, in addition to the measured results in the supplementary materials (*13*), show entangled fidelities higher than 0.94 to the closest pure states (figs. S8 and S9), manifesting the high qualities of the high-dimensional entangled states from the metalens-array–based quantum source. The decreases in the measured fidelities are mainly caused by the long QST measurement time for high-dimensional entangled states, in which the drift of the measurement system cannot be ignored. Therefore, characterization of the entanglement in higher dimensions may need to resort to the development of integrated devices (*13*, *29*).

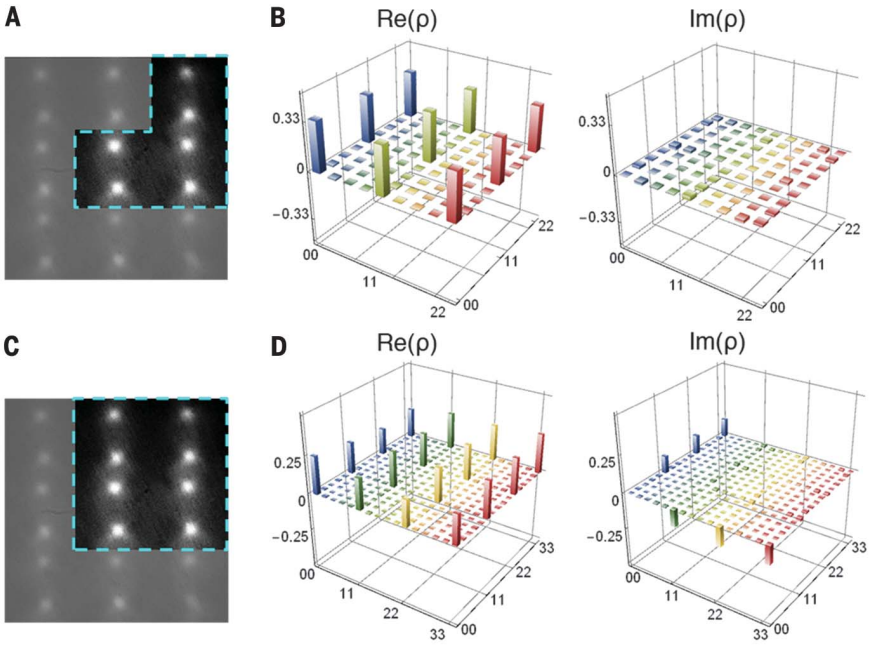


Fig. 3. High-dimensional quantum entanglement states characterization. (**A** to **D**) Reconstructed density matrixes for 3D (B) and 4D (D) quantum state by QST measurements corresponding to the photon pairs in the dashed blue boxes in (A) and (C), respectively.

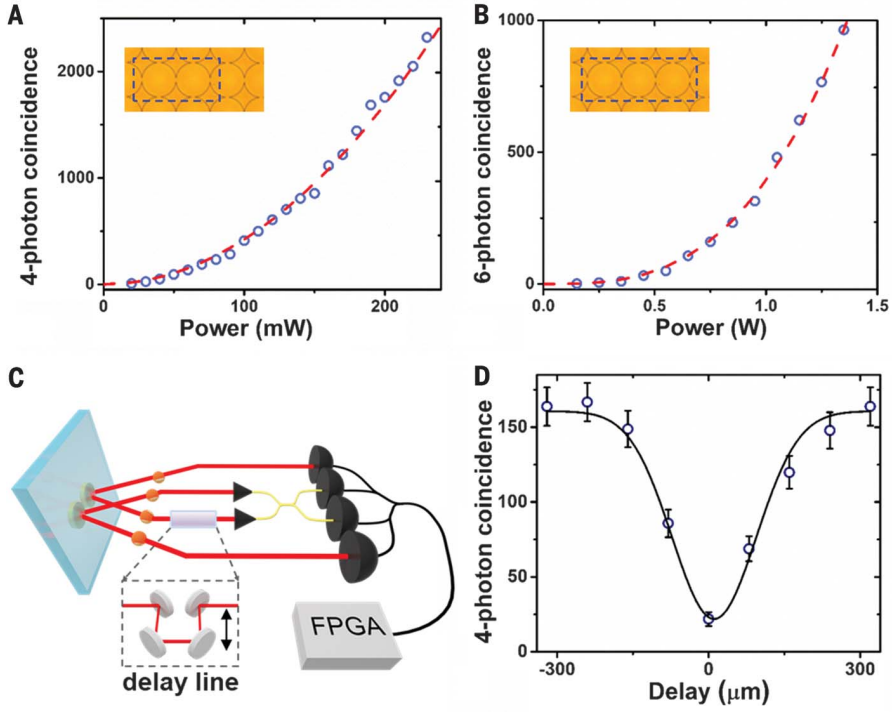


Fig. 4. Multiphoton quantum source based on a metalens array. (**A**) Four-photon and (**B**) six-photon coincidence dependence to the pump power. The red dashed lines are the theoretical estimation of quadratic (four-photon) and cubic (six-photon) trend. The insets show the metalenses involved in the corresponding tests. The schematic diagram (**C**) and the fringe (**D**) of the HOM interference measurement.

Furthermore, the multiphoton source based on the metalens array system is characterized. Unlike the general SPDC-based multiphoton sources, in which multiple nonlinear crystals and long-time stable complex optical setup are required (*14*), our source only needs one nonlinear crystal and the setup is substantially more compact and stable. The experimental setup is discussed in the supplementary materials (*13*), where a 415-nm femtosecond pulsed laser is introduced as the pump light to increase the possibility for multiphoton generation (fig. S10). Here, we use the same metalens array sample shown in Fig. 1 and the BBO crystal designed for the 415-nm pump laser (*13*). Because each metalens may generate one photon pair simultaneously, we can obtain multiphoton-pair sources within this 100-metalens array. As a demonstration, we characterize the performance of the four-photon and six-photon source from two and three adjacent metalenses, respectively. Figure 4, A and B, show the pump power dependencies of the four-photon and six-photon coincidence counts. The ideal power dependencies of the four-photon and six-photon coincidence counts follow a quadratic and cubic relationship, respectively, as shown in the red dashed lines in Fig. 4, A and B. The measured data, shown in the blue circles, agree well with the ideal trend, indicating a feasible multiphoton source.

We further carry out the Hong-Ou-Mandel (HOM) interference to test the purity and indistinguishability of photons generated from different metalenses. Two independent photon pairs are generated by two adjacent metalenses (Fig. 4C). One photon from each pair is used as the trigger (heralding), and two heralded photons interfere at a 50:50 fiber beam splitter. The HOM interference is observed by recording the fourfold coincidence as a function of the relative delay between the heralded photons. Figure 4D is the measured HOM interference result, and the visibility of the HOM dip is 86.3%. This clearly verifies the performance of the multiphoton quantum source and indicates

that such metalens array can successfully provide a compact platform for the multiphoton source preparation.

On the basis of the metalens array, we demonstrate a compact, stable, and controllable platform for a quantum optical source in both high-dimensional entanglement and multiphotons, which extends the spatial configuration and entanglement dimensionality of the integrated path-encoded quantum source. Our results indicate that metasurface structures can provide a route for the generation and control of complex quantum states, not only increasing the quantum system dimensionality but also allowing for the coherent control of multiple photons, thus providing a compact and practical platform for the development of advanced on-chip quantum photonic information processing.

REFERENCES AND NOTES

1. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, *Rev. Mod. Phys.* **74**, 145–195 (2002).
2. N. Gisin, R. Thew, *Nat. Photonics* **1**, 165–171 (2007).
3. V. Scarani *et al*., *Rev. Mod. Phys.* **81**, 1301–1350 (2009).
4. H. Lo, M. Curty, K. Tamaki, *Nat. Photonics* **8**, 595–604 (2014).
5. E. Knill, R. Laflamme, G. J. Milburn, *Nature* **409**, 46–52 (2001).
6. J. L. O'Brien, *Science* **318**, 1567–1570 (2007).
7. A. Aspuru-Guzik, P. Walther, *Nat. Phys.* **8**, 285–291 (2012).
8. I. M. Georgescu, S. Ashhab, F. Nori, *Rev. Mod. Phys.* **86**, 153–185 (2014).
9. V. Giovannetti, S. Lloyd, L. Maccone, *Nat. Photonics* **5**, 222–229 (2011).
10. S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, S. Lloyd, *Nat. Photonics* **12**, 724–733 (2018).
11. S. Khasminskaya *et al*., *Nat. Photonics* **10**, 727–732 (2016).
12. M. Malik *et al*., *Nat. Photonics* **10**, 248–252 (2016).
13. Materials, methods, and additional information are available as Supplementary Materials.
14. H. Wang *et al*., *Phys. Rev. Lett.* **123**, 250503 (2019).
15. J. Wang *et al*., *Science* **360**, 285–291 (2018).
16. N. Yu *et al*., *Science* **334**, 333–337 (2011).
17. N. Yu, F. Capasso, *Nat. Mater.* **13**, 139–150 (2014).
18. A. Arbabi, Y. Horie, M. Bagheri, A. Faraon, *Nat. Nanotechnol.* **10**, 937–943 (2015).
19. G. Zheng *et al*., *Nat. Nanotechnol.* **10**, 308–312 (2015).
20. S. Wang *et al*., *Nat. Commun.* **8**, 187 (2017).
21. S. Wang *et al*., *Nat. Nanotechnol.* **13**, 227–232 (2018).
22. R. J. Lin *et al*., *Nat. Nanotechnol.* **14**, 227–231 (2019).
23. M. Khorasaninejad *et al*., *Science* **352**, 1190–1194 (2016).
24. P. K. Jha, X. Ni, C. Wu, Y. Wang, X. Zhang, *Phys. Rev. Lett.* **115**, 025501 (2015).
25. K. Wang *et al*., *Science* **361**, 1104–1108 (2018).
26. T. Stav *et al*., *Science* **361**, 1101–1104 (2018).
27. P. Georgi *et al*., *Light Sci. Appl.* **8**, 70 (2019).
28. X. L. Niu, Y. F. Huang, G.-Y. Xiang, G.-C. Guo, Z. Y. Ou, *Opt. Lett.* **33**, 968–970 (2008).
29. H. Tang *et al*., *Sci. Adv.* **4**, eaat3174 (2018).

ACKNOWLEDGMENTS

Funding: The authors acknowledge financial support from the National Program on Key Basic Research Project of China (2017YFA0303700, 2016YFA0301700, 2018YFA030602, 2019YFA0308704), National Natural Science Foundation of China (no. 11621091, 11822406, 11834007, 11774164, 61590932, 11774333, 11690032, 91836303, 61490711), Anhui Initiative in Quantum Information Technologies (No. AHY130300), the Strategic Priority Research Program of the Chinese Academy of Sciences (no. XDB24030601), and Shenzhen Science and Technology Innovation Commission (no. SGDX2019081623281169).
Author contributions: L.L., Z.L., X.R., and S.W. contributed equally to this work. L.L., S.W., M.-K.C., C.H.C., H.Y.K., and D.P.T. conceived and performed the design, numerical calculation, and classical optical characterization of metalens array. V.-C.S., C.H.C., and M.K.C. fabricated the metalens array samples. L.L., Z.L., X.R., L.Z., S.W., B.L., and W.Z. performed the quantum measurement and the data analysis. S.W., L.Z., Z.W., G.G., S.Z., and D.P.T. organized and led the project. L.L., S.W., Z.L., X.R., L.Z., S.Z., and D.P.T. prepared the manuscripts; all authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1487/suppl/DC1
Materials and Methods
Figs. S1 to S10
Table S1

20 January 2020; accepted 1 May 2020
10.1126/science.aba9779

PHOTOSYNTHESIS

Quieting a noisy antenna reproduces photosynthetic light-harvesting spectra

Trevor B. Arp[1,2]*, Jed Kistner-Morris[1,2]*, Vivek Aji[2], Richard J. Cogdell[3,4]†, Rienk van Grondelle[4,5]†, Nathaniel M. Gabor[1,2,4]†

Photosynthesis achieves near unity light-harvesting quantum efficiency yet it remains unknown whether there exists a fundamental organizing principle giving rise to robust light harvesting in the presence of dynamic light conditions and noisy physiological environments. Here, we present a noise-canceling network model that relates noisy physiological conditions, power conversion efficiency, and the resulting absorption spectra of photosynthetic organisms. Using light conditions in full solar exposure, light filtered by oxygenic phototrophs, and light filtered under seawater, we derived optimal absorption characteristics for efficient solar power conversion. We show how light-harvesting antennae can be tuned to maximize power conversion efficiency by minimizing excitation noise, thus providing a unified theoretical basis for the observed wavelength dependence of absorption in green plants, purple bacteria, and green sulfur bacteria.

In photosynthesis, light energy harvesting begins with the absorption of sunlight. Photoexcitation energy is rapidly transferred through an antenna network before reaching the reaction center, where charge transfer converts excitation energy into an electrochemical potential gradient across the photosynthetic membrane (*1*). Even in the presence of dynamic light conditions, rapidly fluctuating molecular structure, and highly intricate energy transfer pathways (*1*–*5*), the light-to-electron conversion process exhibits near unity quantum efficiency. The delicate interplay of quantum effects with molecular mechanisms of energy management, e.g., nonphotochemical quenching (*6*–*8*), has been explored across diverse phototrophs (*9*–*12*) but the elementary connection between highly robust light energy harvesting and energetic fluctuations has not been established.

Transforming noisy inputs into quiet outputs represents a general design challenge in network architectures including multinational energy grids (*13*–*17*), auditory and visual neural networks (*18*–*21*), and nanoscale photocells for next-generation optoelectronics (*22*). Although network inputs exhibit statistical fluctuations (e.g., rapid changes of sunlight absorbed by a leaf or solar panel), network outputs may demand a steady rate of information or energy for optimal performance (e.g., constant power from the grid to maintain indoor lighting). Statistical fluctuations—arising from environmental variations and internal processes—fundamentally limit the throughput efficiency of any network. If the flow of energy (power) into a network is substantially larger or smaller than the flow out of the network required to optimally match the output demand, then the network must adapt or be structured in such a way as to reduce the sudden overflow or underflow of energy. When the network fails to manage fluctuations, the results may be extraordinary (e.g., photo-oxidative stress in photosynthetic light harvesting or explosive damage to transformers caused by fluctuations in the grid).

We constructed a model that uses generalizations of networks to extract essential aspects of photosynthetic light harvesting (Fig. 1). A simple network of nodes (points at which lines intercept) connected by links (connecting lines) represent physical objects: excitation energy levels and intermolecular transfer events within the antenna system, respectively. In photosynthesis, light enters the antenna through a large number of pigment molecules, each of which is a member of a small set of distinct molecular species (e.g., chlorophyll a and b). Our model considers the advantage in having light entering the network through two classes of absorbing excitation energy levels, nodes *A* and *B*, which can absorb powers $\mathcal{P}_A$ and $\mathcal{P}_B$.

After an absorption event, excitation energy moves between internal nodes of the antenna network, representing the excitation of intermediate states within the biological antenna complex (*2*, *11*, *23*). As an example, excitation energy absorbed by a chlorophyll b molecule in the light-harvesting complex LHC2 is transferred to chlorophyll a, and from there to another chlorophyll a in the same or in another complex. There are many such pathways through the antenna network that may share intermediate links, but each specific pathway (Fig. 1A, colored lines) eventually terminates on the output, *O*, which on average delivers power Ω. Rather than model each pathway, we modeled the average behavior from all inputs *A* and *B* through all pathways in the network. Thus, we define probabilities p_A and p_B, which are the total probabilities that any input *A* and *B* will absorb. Mathematically, the example pathway given above, along with all other

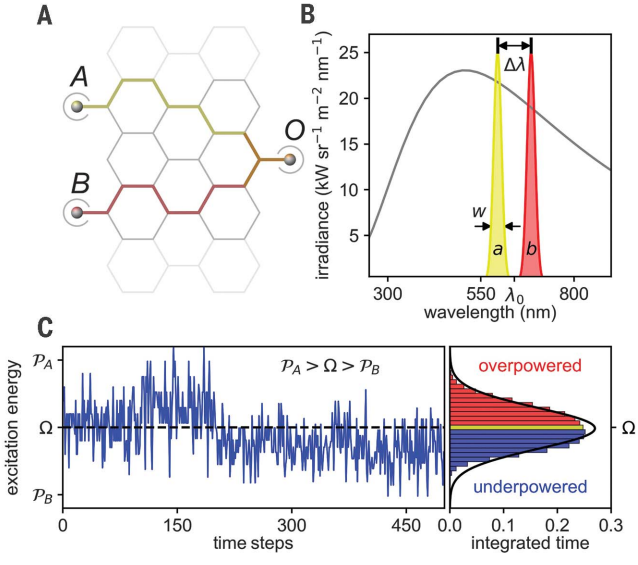


Fig. 1. Light-harvesting noisy antenna. (**A**), Schematic of a photosynthetic antenna reduced into a network with two input nodes *A* and *B* with input rates $\mathcal{P}_A$ and $\mathcal{P}_B$ and output *O* with rate Ω. Energy is absorbed by molecules *a* and *b* (at rates $\mathcal{P}_A$ and $\mathcal{P}_B$) and is transferred to the output as usable energy. (**B**) Schematic two-channel antenna absorption spectra (yellow and red) and incident blackbody light source (gray). The quantities λ_0, $\Delta\lambda$, and *w* are, respectively, the center wavelength, distance between peaks, and width of the absorption peaks. (**C**) (Left) Simulated average excitation energy as a function of time within a noisy antenna composed of 10 sets of *a* and *b* molecules. (Right) Time-averaged histogram of the internal energy (detailed in the supplementary materials, section S1.4). The antenna is subject to internal (fast) and external (slow) fluctuations. Over long time scales, the time-averaged histogram resembles a normal distribution (black line).

[1]Laboratory of Quantum Materials Optoelectronics, University of California, Riverside, CA 92521, USA. [2]Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA. [3]Institute of Molecular, Cell, and Systems Biology, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Glasgow G128QQ, UK. [4]Canadian Institute for Advanced Research, Toronto, Ontario M5G 1M1, Canada. [5]Department of Physics and Astronomy, Faculty of Sciences, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, Netherlands.

*These authors contributed equally to this work

†Corresponding author. Email: richard.cogdell@glasgow.ac.uk (R.C.); r.van.grondelle@vu.nl (R.v.G.); nathaniel.gabor@ucr.edu (N.M.G.)

pathways that originate on a node B, is contained in p_B. The average power (or rate of energy) coming from the B absorbers is therefore $p_B\mathcal{P}_B$. This is the average value; the actual flow of excitation energy at any given time is stochastic.

By analyzing the stochastic flow of excitation energy, we can characterize the antenna network by statistical averages (power throughput) and fluctuations in the rate of energy flow, which we will call noise (see the supplementary materials, section S1). The power throughput of the antenna system is determined by the external light conditions, the absorption characteristics of the absorbing pigment molecules (Fig. 1B) or input nodes, and the molecular dynamics of the network. The antenna inputs are described in the usual way: Light absorption by the pigment molecules is characterized by peak width w, separation $\Delta\lambda = |\lambda_B - \lambda_A|$, and the center wavelength (or average distance) between the peaks λ_0. The solar spectral irradiance (Fig. 1B, gray line), which varies as light propagates through air, the canopy, or seawater, gives the average power available within a given range of wavelengths. Choosing the wavelength of an absorption peak simultaneously specifies both the excitation energy and power entering the noisy antenna. Although the excitation energy is inversely proportional to wavelength, the absorbed power $\mathcal{P}_A$ or $\mathcal{P}_B$ is the integrated product of the spectral irradiance and the absorption characteristics of the light-harvesting antenna.

Because the absorbed solar power rarely matches exactly the rate of optimal output, a finely tuned network is one that most effectively balances minimizing the internal noise with robustness against external noise. Noise in the antenna arises from two main sources: inherent mismatch between inputs and output, which may arise because of fast dynamics in the protein structure and corresponding electronic properties, and dynamic external light conditions. In photosynthesis, an overpowered antenna will produce excess energy that can drive deleterious back reactions (*24*, *25*). Conversely, a light-harvesting network in an underpowered state produces nonoptimal output, because the rate of energy transfer out of the network is fixed by electrochemical processes (*26*). Over long periods of time, the degree to which the light-harvesting network is overpowered or underpowered is measured by the mean-squared deviation (i.e., noise) of the total input power (through $\mathcal{P}_A$ and $\mathcal{P}_B$) from the optimal output power at Ω, (Fig. 1C) (see the supplementary materials, section S1).

Tuning only the absorption characteristics, our goal was to find a network that spends the least amount of time in a state for which the input power is too large or too small compared with the output of the network, thus maximizing power conversion efficiency (Fig. 1C). Within our model, probabilities p_A and p_B couple the inputs of the network $\mathcal{P}_A$ and $\mathcal{P}_B$ to the output Ω: $p_A\mathcal{P}_A + p_B\mathcal{P}_B = \Omega$. From this expression, we first evaluate the variance of the average distribution $p_A\mathcal{P}_A + p_B\mathcal{P}_B$. Minimizing this variance yields the optimal values of $\mathcal{P}_A$ and $\mathcal{P}_B$ to quiet the antenna. We then input the local spectral irradiance to a model optimization function, the maxima of which determine the optimal absorption characteristics for noise cancellation (see the supplementary materials, sections S1.1 to S1.3).

Using this framework, we can predict the behavior of three noise regimes within the antenna network: overtuned, finely tuned, and poorly tuned (Fig. 2). For simplicity, these examples are where $\Omega = (\mathcal{P}_A + \mathcal{P}_B)/2$, whereas a broader parameter range is explored in detail in the supplementary materials, section S1.2. Although the light conditions are identical for all three cases (Fig. 2A, gray lines), we can examine how the noise changes with different absorption characteristics (details of this calculation can be found in the supplementary materials, section S1.4). When the absorbing peaks are spaced too closely (Fig. 2A, top), the inherent antenna noise can be strongly reduced, and in the limit that $\mathcal{P}_A = \Omega = \mathcal{P}_B$, there

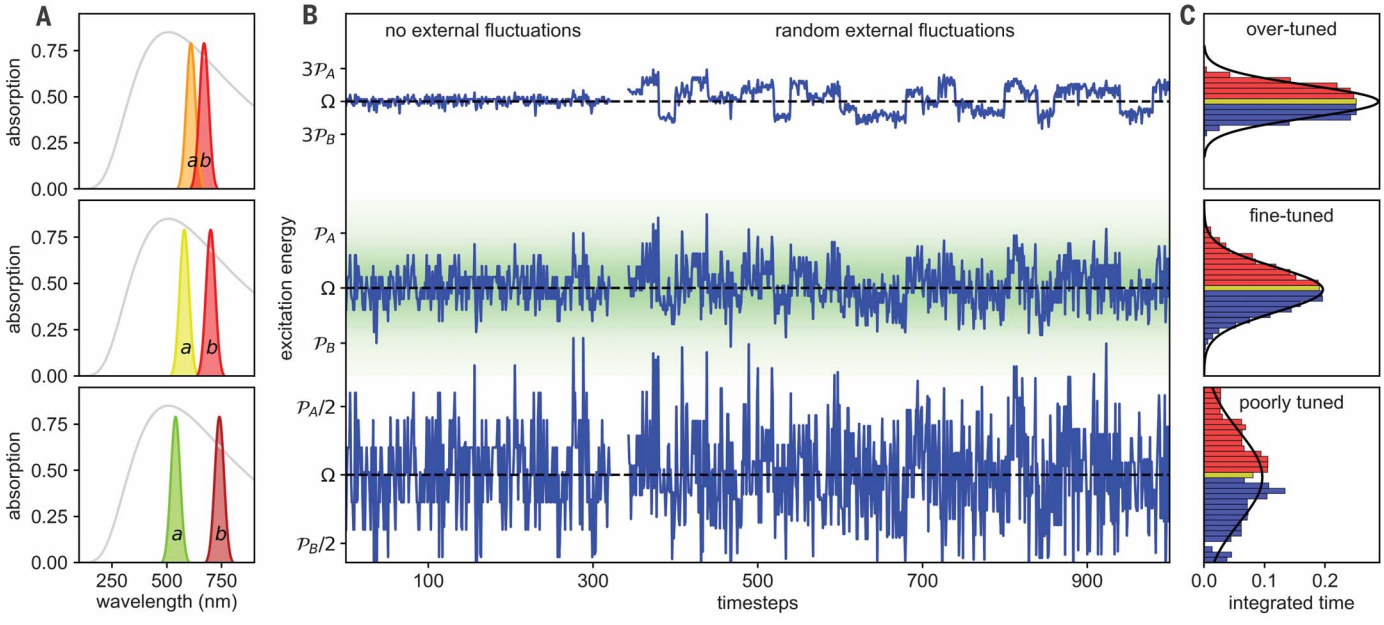


Fig. 2. Quieting a noisy antenna by tuning the absorption characteristics. (**A**) Absorption peaks for two absorbers *a* and *b* overlaid on an ideal blackbody solar spectrum (T = 5500 K, gray line) for three cases: two closely spaced absorbers (top), two absorbers separated to optimize the noisy antenna (middle), and two widely separated absorbers (bottom). (**B**) Simulated excitation energy versus time for a two-channel antenna with three different values of Δ, comparable to the cases shown in (A). The left side shows the excitation energy time traces without external fluctuations. The right side includes random external fluctuations. (**C**) Histograms of time spent in overpowered (red) and underpowered (blue) states for the three series in (B). In the top panel, the distribution is flat and favors no value. In the middle panel, the distribution is a sharply peaked normal distribution that favors Ω. In the bottom panel, the distribution is normal but wider than that in the middle panel.

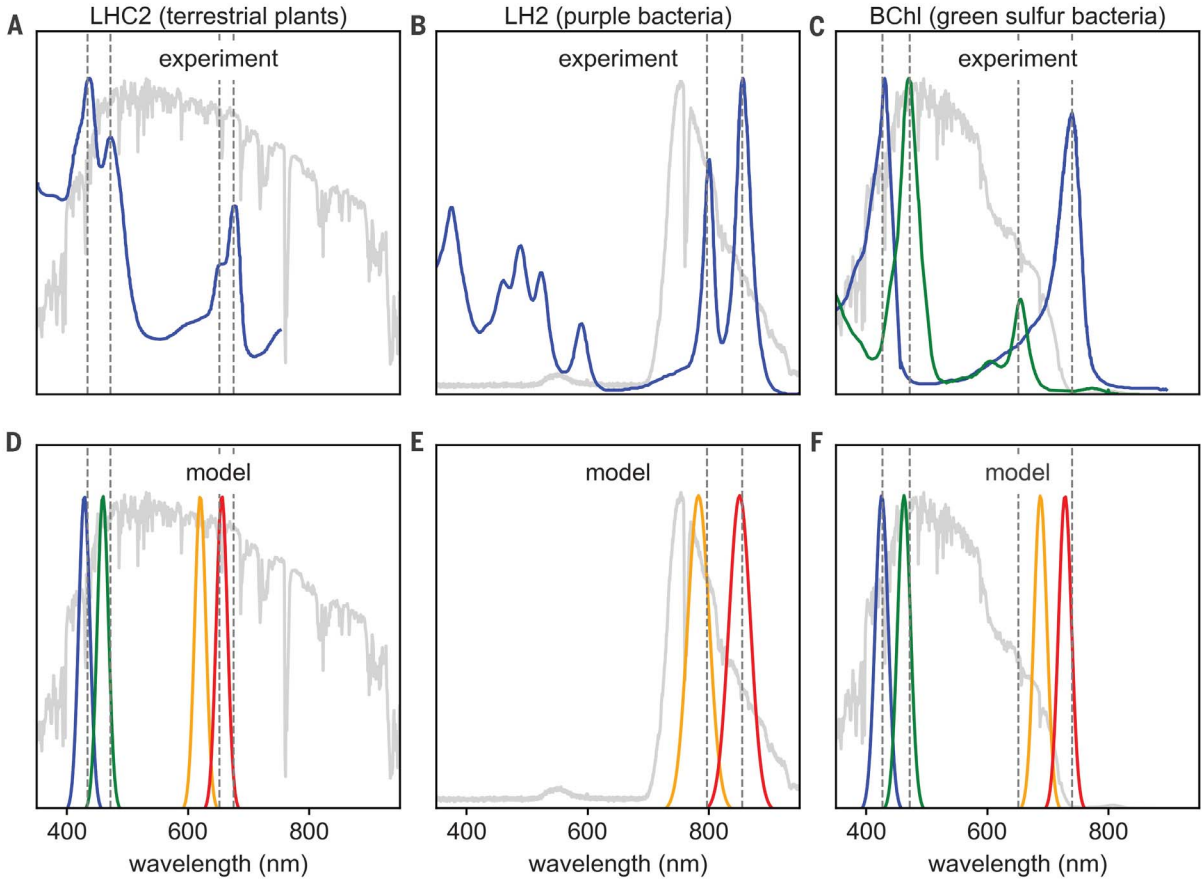


Fig. 3. Comparison of the noisy antenna model with photosynthetic absorption spectra in three distinct niches. (**A**) Absorption spectrum of LHC2 (blue) (*27*) overlaid on the terrestrial solar spectrum (light gray) (*31*). (**B**) Absorption spectrum of the LH2 complex (*28*) overlaid on the solar spectrum measured below a canopy of leaves (light gray) (see the supplementary materials, section S3, for details). (**C**) Absorption spectra of bacteriochlorophyll c (blue) and e (green) (*29*, *30*) compared with the calculated solar spectrum at a 2-m depth of water (light gray) (*32*, *33*). (**D** to **F**) Predicted ideal absorption peaks from optimizing $\Delta = \mathcal{P}_A - \mathcal{P}_B$ for the full solar spectrum, the solar spectrum attenuated through canopy, and the solar spectrum attenuated through seawater, respectively (see the supplementary materials, sections S2 and S3, for optimization and spectra details, respectively). Photosynthetic absorption peaks are identified with dashed lines.

are negligible fluctuations in the rate of energy flow (Fig. 2B, top left). This lower bound to the internal noise cannot be reached in natural photosynthetic antennae, where protein dynamics will always drive fluctuations of intermediate excitation energy transfer events. Rather, the overtuned antenna noise is directly proportional to, and thus dominated by, changes in the varying light spectrum (Fig. 2B, top right). In the presence of random external fluctuations, the distribution of time spent in an overpowered or underpowered state is flat (Fig. 2C, top). In the overtuned antenna, the average input rarely matches the optimal output.

A poorly tuned antenna (Fig. 2A, bottom) is similarly deficient. If the absorbing peaks are well separated, then the antenna spends most of the time overpowered or underpowered. When the power sources $\mathcal{P}_A$ or $\mathcal{P}_B$ are significantly greater or less than the power sink ($\mathcal{P}_A >> \Omega >> \mathcal{P}_B$), the noise (as evidenced by a histogram of the excitation energy) in the poorly tuned antenna becomes broader as the absorbing peaks become more separated (Fig. 2C, bottom). When viewed over long times, the poorly tuned antenna spends too little time outputting the optimal power Ω.

The finely tuned antenna absorbs at specific positions on the spectrum that give rise to robust light harvesting even in the presence of both varying light conditions and substantial internal noise. Compared with the overtuned and undertuned cases, the finely tuned antenna allows for intermediate internal noise levels (Fig. 2B, middle) yet delivers a narrow distribution of power centered at the optimal output Ω (Fig. 2C, middle). Robustness in light harvesting is thus the ability to output, on average, the optimal rate Ω yet simultaneously allow for internal noise.

To determine the optimal absorption spectrum for robust light harvesting, we computed the spectral positions for which the peaks are as close as possible on the light spectrum (favoring reduced internal noise) yet the difference in the absorbed power $\Delta = \mathcal{P}_A - \mathcal{P}_B$ is maximized (supporting robustness against external variations). This condition is equivalent to maximizing the derivative of the light spectrum with respect to wavelength, thus resulting in absorption peaks in regions of steepest slope (see the supplementary materials, section S1.3). The absorption spectra, and thus the excitation transitions, are tuned so that the time-averaged sum of input excitation energy is sharply peaked at the output rate (Fig. 2C, middle).

For three prototypical photosynthetic antennae, the light-harvesting complex of green plants (LHC2), the light-harvesting complex of purple bacteria (LH2), and the bacteriochlorophyll c and e pigments of green sulfur bacteria (BChl c and e), the natural absorption spectrum (Fig. 3, A to C) (*27*–*30*) can be compared with that predicted by our model

(Fig. 3, D to F) (see the supplementary materials, section S2, for full details), which takes as input the local irradiance spectrum, shown as solid gray lines in Fig. 3, A to F. The absorption peak positions and spectral separation predicted under light conditions in air (*31*), under canopy, or under seawater (*32*, *33*) (colored lines in Fig. 3, D, E, and F, respectively) show strong agreement with the absorption spectra of these three phototrophs. Using only the external light spectrum and the linewidth w, the predicted peak center position λ_0 and separation $\Delta\lambda$ reproduced the measured absorption peaks with an average error of 2.1% (Table 1).

The noisy antenna model also reproduced a general feature of photosynthetic light harvesting that was observed across all three prototypical phototrophs. Photosynthetic pigments do not absorb at the maximum solar power. Instead, all three phototrophs considered exhibit pairs of closely spaced peaks in regions where the spectrum shows a steep rate of change with respect to wavelength. Photosynthetic plants look green because their antenna complexes absorb light across the visible spectrum, including the blue and red portions, yet reflect green wavelengths (Fig. 3D). Purple bacteria are aquatic phototrophs (*34*). They have adapted to sunlight that is filtered through the canopy of trees and floating aerobic phototrophs (Fig. 3E, gray line, and see the supplementary materials, section S3) and use a light-harvesting complex in which bacteriochlorophyll dominates light absorption away from the visible, including green (Fig. 3E). Green sulfur bacteria are a geographically diverse group of bacteria that are adapted to solar light shining through seawater to depths where it is anaerobic (*35*). They do not absorb the peak intensity of this attenuated light spectrum and instead absorb in the region of steepest spectral rate of change.

Underwater phototrophs provide an excellent natural experiment with which to test the predictive strength of our model because the solar spectrum is highly variable as a function of depth (*36*). When considering the penetrating spectrum below the seawater surface, light intensity is attenuated as depth increases, particularly in the red and infrared zones, because of absorption and scattering in seawater (Fig. 4A). By comparing the absorption spectra of subsurface marine phototrophs, such as green sulfur bacteria, with those predicted by quieting a noisy antenna, we can explore whether the natural photosynthetic absorption spectrum matches our model predictions for the relevant phototroph's preferred depth.

From these attenuated solar spectra, we calculated an optimization parameter Δ^{op} as a function of $\Delta\lambda$ and λ_0. Δ^{op} is a function modified from the calculation of $\Delta = \mathcal{P}_A - \mathcal{P}_B$ such

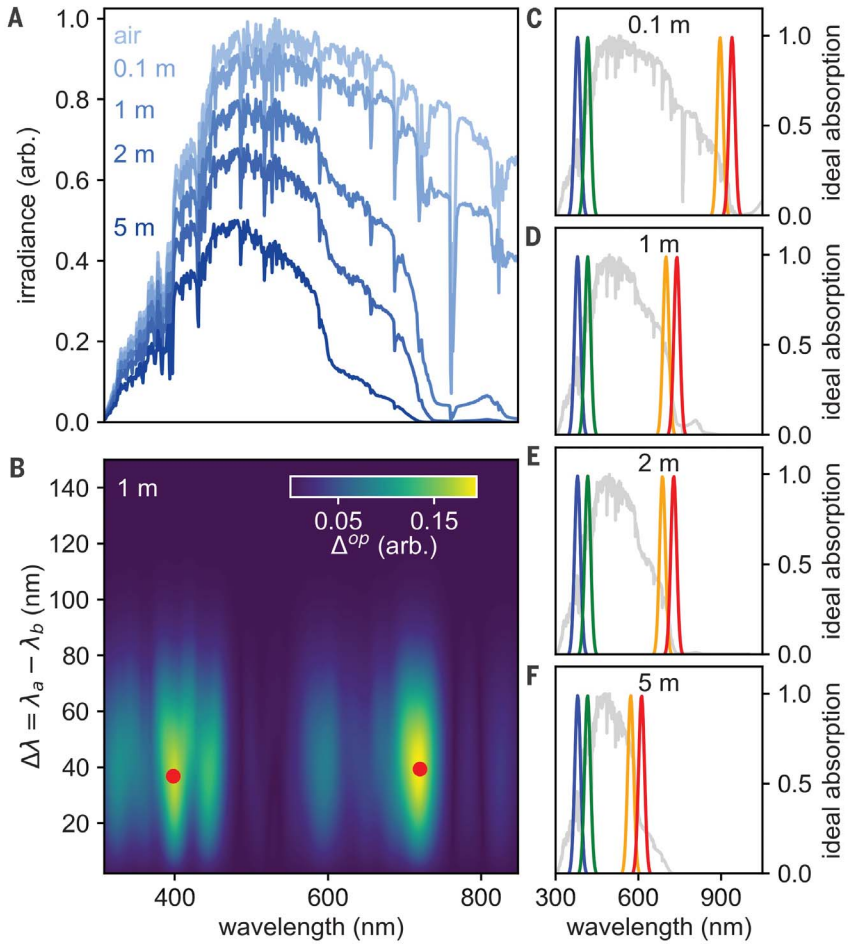


Fig. 4. Predicted absorption spectra of a finely tuned noisy antenna under seawater. (**A**) Solar spectrum in air (*31*) and attenuated by various depths of water (labeled) (*32*, *33*). (**B**) Optimization landscape calculation of Δ^{op} versus center wavelength λ_0 and peak separation $\Delta\lambda$ for the solar spectrum under 1 m of seawater (w = 15 nm). Red points identify two equally favorable maxima corresponding to a set of peaks on either side of the spectral maximum. (**C** to **F**) Ideal absorption peaks predicted from the solar spectrum at each depth. (D) shows the peaks extracted from the calculation in (B) color coded blue, green, orange, and red to track peak locations with depth.

Table 1. Absorption peak data versus model calculation.

Peak name	Actual value, nm (eV)	Calculated value, nm (eV)	Relative % error from nm (eV)	Reference
Chlorophyll a 1	428 (2.90)	429 (2.89)	0.23 (0.34)	(*27*)
Chlorophyll b 1	440 (2.82)	459 (2.70)	4.32 (4.26)	(*27*)
Chlorophyll b 2	652 (1.90)	620 (2.00)	4.91 (5.26)	(*27*)
Chlorophyll a 2	660 (1.88)	656 (1.89)	0.61 (0.53)	(*27*)
LH2 band 1	801 (1.55)	783 (1.58)	2.25 (1.94)	(*28*)
LH2 band 2	857 (1.45)	851 (1.46)	0.70 (0.69)	(*28*)
Bacteriochlorophyll c 1	431 (2.88)	426 (2.91)	1.16 (1.04)	(*29*)
Bacteriochlorophyll e 1	461 (2.69)	462 (2.68)	0.22 (0.37)	(*30*)
Bacteriochlorophyll e 2	655 (1.89)	688 (1.80)	5.04 (4.76)	(*30*)
Bacteriochlorophyll c 2	740 (1.68)	728 (1.70)	1.62 (1.19)	(*29*)

that its maxima quiet a noisy antenna (see the supplementary materials, section S1.3). An example color map of the magnitude of Δ^{op} at a depth of 1 m and w = 15 nm reveals two maxima in the color plot near λ_o = 400 and 750 nm (Fig. 4B). These maxima identify the wavelength characteristics of a finely tuned antenna under seawater. By extracting the values of $\Delta\lambda$ and λ_o at the maximum in Δ, we obtain the characteristic absorption spectra of the fine tuned antenna as a function of seawater depth (Fig. 4, C to F). We found that quieting a noisy antenna under 2 m of seawater accurately reproduces the absorption spectrum of green sulfur bacteria. Although highly adaptable, green sulfur bacteria are known to thrive at 1 to 2 m below the surface (*37*), coinciding with the conditions for which their light-harvesting antenna is finely tuned for solar power conversion.

The degree to which we were able to reproduce photosynthetic absorption spectra is a surprising result, suggesting an underlying organizing principle for light-harvesting systems: Fluctuations fundamentally limit the efficiency of networks and must be avoided. Phototrophs must balance environmental inputs to sustain steady production and storage of fuel under substantially different environmental conditions. Phototrophs across many photosynthetic niches may have adapted to build fluctuation-canceling light-harvesting antennae onto which other active mechanisms for reducing fluctuations can be added (e.g., nonphotochemical quenching) (*6*–*8*). Although the connection of our model to natural antenna systems requires detailed quantum models, our framework gives new insight into how extinction coefficients, delocalization lengths, and radiative rates conspire to reduce noise in natural antennae. Moreover, our findings will inspire comprehensive experiments in which the light environment is carefully controlled while the absorption spectrum of adaptable model organisms is monitored. By developing noise-canceling antennae as a technological foundation, natural and artificial energy-harvesting networks—from bacteria thriving underwater to extended power grids—could be adapted to efficiently convert noisy inputs into robust outputs.

REFERENCES AND NOTES

1. R. Blankenship, *Molecular Mechanisms of Photosynthesis* (Wiley, 2002).
2. R. van Grondelle, J. P. Dekker, T. Gillbro, V. Sundstrom, *Biochim. Biophys. Acta.* **1187**, 1–65 (1994).
3. G. P. van Nieuw Amerongen, S. van Delft, M. A. Vermeer, J. G. Collard, V. W. M. van Hinsbergh, *Circ. Res.* **87**, 335–340 (2000).
4. G. D. Scholes, G. R. Fleming, A. Olaya-Castro, R. van Grondelle, *Nat. Chem.* **3**, 763–774 (2011).
5. R. Croce, H. van Amerongen, *Nat. Chem. Biol.* **10**, 492–501 (2014).
6. C. Külheim, J. Agren, S. Jansson, *Science* **297**, 91–93 (2002).
7. A. A. Pascal *et al*., *Nature* **436**, 134–137 (2005).
8. A. V. Ruban *et al*., *Nature* **450**, 575–578 (2007).
9. R. G. Fleming, R. van Grondelle, *Phys. Today* **47**, 48–55 (1994).
10. E. Romero, V. I. Novoderezhkin, R. van Grondelle, *Nature* **543**, 355–365 (2017).
11. T. Mirkovic *et al*., *Chem. Rev.* **117**, 249–293 (2017).
12. G. D. Scholes *et al*., *Nature* **543**, 647–656 (2017).
13. K. Schmietendorf, J. Peinke, O. Kamps, *Eur. Phys. J. B* **90**, 222 (2017).
14. B. Schäfer, C. Beck, K. Aihara, D. Witthaut, M. Timme, *Nat. Energy* **3**, 119–126 (2018).
15. T. Nesti, A. Zocca, B. Zwart, *Phys. Rev. Lett.* **120**, 258301 (2018).
16. H. Haehne, J. Schottler, M. Waechter, J. Peinke, O. Kamps, *Europhys. Lett.* **121**, 30001 (2018).
17. T. Coletta, B. Bamieh, P. Jacquod, Transient performance of electric power networks under colored noise. arXiv:1807.09048v2 [math.OC] (31 January 2019).
18. D. De Ridder *et al*., *Neurosci. Biobehav. Rev.* **44**, 16–32 (2014).
19. A. M. Leaver *et al*., *Neuron* **69**, 33–43 (2011).
20. B. B. Averbeck, P. E. Latham, A. Pouget, *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
21. I. Kanitscheider, R. Coen-Cagli, A. Pouget, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6973–E6982 (2015).
22. T. B. Arp, Y. Barlas, V. Aji, N. M. Gabor, *Nano Lett.* **16**, 7461–7466 (2016).
23. R. van Grondelle, *Biochim. Biophys. Acta* **811**, 147–195 (1985).
24. P. Horton, A. V. Ruban, R. G. Walters, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 655–684 (1996).
25. A. V. Ruban, M. P. Johnson, C. D. P. Duffy, *Biochem. Biophys. Acta.* **1817**, 167–181 (2012).
26. G. D. Farquhar, S. von Caemmerer, J. A. Berry, *Planta* **149**, 78–90 (1980).
27. T. P. J. Krüger, R. van Grondelle, *Physica B* **480**, 7–13 (2016).
28. R. J. Cogdell *et al*., *FEBS Lett.* **555**, 35–39 (2003).
29. D. C. Brune, T. Nozawa, R. E. Blankenship, *Biochemistry* **26**, 8644–8652 (1987).
30. C. M. Borrego, J. B. Arellano, C. A. Abella, T. Gillbro, J. Garcia-Gil, *Photosynth. Res.* **60**, 257–264 (1999).
31. National Renewable Energy Laboratory (NREL), U.S. Department of Energy, *Reference Air Mass 1.5 Spectra* (2003); https://www.nrel.gov/grid/solar-resource/spectra-am1.5.html.
32. H. Buiteveld, J. H. M. Hakvoort, M. Donze, "Optical properties of pure water," in *Proceedings of SPIE 2258 Ocean Optics XII* (26 October 1994), pp. 174–183; https://doi.org/10.1117/12.190060.
33. L. Kou, D. Labrie, P. Chylek, *Appl. Opt.* **32**, 3531–3540 (1993).
34. C. N. Hunter, F. Daldal, M. C. Thurnauer, J. T. Beatty, *The Purple Phototropic Bacteria* (Springer, 2009).
35. J. M. Olsen, in *Green Photosynthetic Bacteria*, J. M. Olsen, J. G. Ormerod, J. Amesz, E. Stackebrandt, H. G. Trüper, Eds. (Springer, 1988), pp. 315–319.
36. P. Jenkins *et al*., *IEEE J. Photovoltaics* **4**, 202–207 (2014).
37. A. Kharcheva, A. Zhiltsova, in *WDS'16 Proceedings of Contributed Papers – Physics*, J. Šafránková and J. Pavlů, Eds. (Matfyzpress, 2016); pp. 214–218.
38. T. B. Arp, qmolabucr/qntenna: Publication Release, version v1.1, Zenodo (2020); https://doi.org/10.5281/zenodo.3765834.

ACKNOWLEDGMENTS

We thank V. Elser, J. Iezzi, P. McEuen, T. Sargent, and C. Varma for valuable discussions. **Funding:** T.B.A., J. K.-M., and N.M.G were supported by the Air Force Office of Scientific Research Young Investigator Program (YIP) award no. FA9550-16-1-0216, National Science Foundation Division of Materials Research CAREER award no. 1651247, and the U.S. Department of the Navy Historically Black Colleges, Universities and Minority Serving Institutions (HBCU/MI) award no. N00014-19-1-2574. N.M.G. acknowledges support from a Cottrell Scholar Award and a Canadian Institute for Advanced Research (CIFAR) Azrieli Global Scholar Award. T.B.A. acknowledges support from the Fellowships and Internships in Extremely Large Data Sets (FIELDS) program, a NASA MUREP Institutional Research Opportunity (MIRO) program grant no. NNX15AP99A. R.J.C. gratefully acknowledges support from the Photosynthetic Antenna Research Center, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under award no. DE-SC 0001035 and the Biotechnological and Biological Sciences Research Council (BBSRC). R.v.G. was supported by the Royal Netherlands Academy of Arts and Sciences and the Canadian Institute for Advanced Research (CIFAR). **Author contributions:** T.A. and J. K. M. performed detailed analysis and computational modelling. N.M.G. and V.A. conceived the conceptual model and supervised the analytical and computational modeling with additional input from R.J.C. and R.v.G. R.J.C., N.M.G., and R.v.G. chose which phototrophs should be used as exemplars. All authors contributed to the writing of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The code for the parameter search that generates the main results shown in Figs. 3 and 4, as well as the discrete simulation code used to generate Figs. 1C and 2B, are available on Zenodo (*38*). This repository includes the irradiance data shown in Fig. 3, D to F, and Fig. 4A such that all model results can be fully replicated.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1490/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
References (*39*–*59*)
MDAR Reproducibility Checklist

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

20 December 2019; accepted 4 May 2020
10.1126/science.aba6630

DOG GENOMICS

Arctic-adapted dogs emerged at the Pleistocene–Holocene transition

Mikkel-Holger S. Sinding[1,2,3,4,5]*†, Shyam Gopalakrishnan[1]*, Jazmín Ramos-Madrigal[1]*, Marc de Manuel[6]*, Vladimir V. Pitulko[7]*, Lukas Kuderna[6], Tatiana R. Feuerborn[1,3,8,9], Laurent A. F. Frantz[10,11], Filipe G. Vieira[1], Jonas Niemann[1,12], Jose A. Samaniego Castruita[1], Christian Carøe[1], Emilie U. Andersen-Ranberg[3,13], Peter D. Jordan[14], Elena Y. Pavlova[15], Pavel A. Nikolskiy[16], Aleksei K. Kasparov[7], Varvara V. Ivanova[17], Eske Willerslev[1,18,19,20], Pontus Skoglund[21,22], Merete Fredholm[23], Sanne Eline Wennerberg[24], Mads Peter Heide-Jørgensen[4], Rune Dietz[25], Christian Sonne[3,25,26], Morten Meldgaard[1,3], Love Dalén[8,27], Greger Larson[10], Bent Petersen[1,28], Thomas Sicheritz-Pontén[1,28], Lutz Bachmann[2], Øystein Wiig[2], Tomas Marques-Bonet[6,29,30,31]†‡, Anders J. Hansen[1,3]†‡, M. Thomas P. Gilbert[1,32]†‡

Although sled dogs are one of the most specialized groups of dogs, their origin and evolution has received much less attention than many other dog groups. We applied a genomic approach to investigate their spatiotemporal emergence by sequencing the genomes of 10 modern Greenland sled dogs, an ~9500-year-old Siberian dog associated with archaeological evidence for sled technology, and an ~33,000-year-old Siberian wolf. We found noteworthy genetic similarity between the ancient dog and modern sled dogs. We detected gene flow from Pleistocene Siberian wolves, but not modern American wolves, to present-day sled dogs. The results indicate that the major ancestry of modern sled dogs traces back to Siberia, where sled dog–specific haplotypes of genes that potentially relate to Arctic adaptation were established by 9500 years ago.

Despite decades of studies, consensus has yet to be reached on when and where dogs were first domesticated and when they were first deliberately used in many of the roles they exhibit today. In Siberia, late Upper Paleolithic artifacts of carved bone, antler, and ivory similar to tools used by modern Inuit for securing dog harness straps suggest ancient origins of dog sledding (*1*). Furthermore, archeological findings from Zhokhov Island provide evidence of sled technology and dogs by the Sumnagin Mesolithic culture ~9000 to 8000 years ago (*1*–*3*) (fig. S1), offering an opportunity to use genomics to further our understanding of early dog domestication and the origin of sled dogs.

We generated nuclear genomes from a dog mandible present at this site ("Zhokhov," 9.6× coverage), dated to 9524 calendar years before present (YBP) (Fig. 1A and fig. S2), and a Siberian Pleistocene wolf mandible ("Yana," 4.7× coverage), dated to 33,019.5 YBP (Fig. 1A and fig. S3). In addition, we sequenced 10 modern Greenland sled dog genomes, a dog best described as an indigenous landrace breed used for hunting and sledging by Inuit. Samples consisted of two individuals from each of five geographically diverse localities (Fig. 1A), thus providing a broad representation of the indigenous dog diversity.

We analyzed our data alongside genomes from 114 geographically and genetically diverse canids (table S1) using whole-genome pairwise distances, principal component analysis, TreeMix (*4*) admixture graphs, and *D* statistics (Fig. 1). Yana appeared alongside wolves (Fig. 1, B and C), whereas Zhokhov was found to be most closely related to dogs. Specifically, Zhokhov was most similar to modern sled dogs (Greenland sled dogs, Alaskan malamutes, and Alaskan and Siberian huskies) and American pre-European-contact dogs (PCDs), best illustrated by the ~2× Port au Choix dog from Maritime Archaic cultural context ~4000 YBP (*3*). Unsupervised clustering analyses with NGSadmix software (*4*) (fig. S6) grouped modern domestic dogs into four clusters: African, European, Asian, and sled dogs including Zhokhov. These relationships were confirmed by an admixture graph in which Yana was more closely related to a Pleistocene wolf from Taimyr Peninsula than to modern wolves, whereas Zhokhov represents a lineage that diverged from the ancestor of present-day sled dogs (Fig. 1C and figs. S8 and S9). This suggests genetic continuity in Arctic dog breeds for at least the past ~9500 years, setting a lower bound on the origin of the sled dog lineage.

Next, *D* statistics indicated an excess of allele sharing between Yana-Taimyr wolves and PCDs-Zhokhov-sled dogs (Fig. 1D and fig. S14), corroborating previous reports (*3*, *5*). This suggests that the admixture occurred between Pleistocene wolves and the ancestors of PCDs, Zhokhov, and sled dogs.

Previous studies have demonstrated an association between canine transmissible venereal tumors (CTVTs) and sled dogs, especially PCDs (*3*). Here, we evaluated the relationship among Zhokhov, two CTVT genomes (table S1), and dogs and wolves using f_3 statistics and phylogenetic analysis. Recent analyses of exome data suggested that CTVT expanded across Eurasia ~6000 years ago (*6*), thus reducing the likelihood that this transmissible cancer originated in the Americas. In our study, both the phylogenetic analysis (fig. S9) and f_3 statistics (fig. S10) placed the CTVT genomes closer to PCDs than to sled dogs or Zhokhov. These results suggest that the basal dog lineage that led to PCDs (*3*) occurred in Eurasia ~6000 years ago and/or there were multiple introductions of PCD-like dogs to the Americas.

We used NGSadmix, admixture analyses, and *D* statistics (figs. S6 to S8 and S11 to S15) to evaluate gene flow and shared ancestry between Zhokhov and modern dogs and wolves. We found no significant gene flow between any sled dog (including Zhokhov) and modern American–Arctic wolf populations compared

[1]The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. [2]Natural History Museum, University of Oslo, Oslo, Norway. [3]The Qimmeq Project, University of Greenland, Nuussuaq, Greenland. [4]Greenland Institute of Natural Resources, Nuuk, Greenland. [5]Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Ireland. [6]Institute of Evolutionary Biology (UPF-CSIC), Barcelona, Spain. [7]Institute for the History of Material Culture, Russian Academy of Sciences, St. Petersburg, Russia. [8]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden. [9]Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden. [10]The Palaeogenomics and Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, UK. [11]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK. [12]BioArch, Department of Archaeology, University of York, York, UK. [13]Department of Clinical Veterinary Sciences, University of Copenhagen, Frederiksberg C, Denmark. [14]Arctic Centre and Groningen Institute of Archaeology, University of Groningen, Netherlands. [15]Arctic and Antarctic Research Institute, St. Petersburg, Russia. [16]Geological Institute, Russian Academy of Sciences, Moscow, Russia. [17]VNIIOkeangeologia Research Institute (The All-Russian Research Institute of Geology and Mineral Resources of the World Ocean), St. Petersburg, Russia. [18]Danish Institute for Advanced Study (D-IAS), University of Southern Denmark, Odense, Denmark. [19]Department of Zoology, University of Cambridge, Cambridge, UK. [20]Wellcome Trust Sanger Institute, University of Cambridge, Cambridge, UK. [21]Department of Genetics, Harvard Medical School, Boston, MA, USA. [22]Francis Crick Institute, London, UK. [23]Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksberg C, Denmark. [24]Ministry of Fisheries, Hunting and Agriculture, Government of Greenland, Nuuk, Greenland. [25]Department of Bioscience, Arctic Research Centre, Aarhus University, Roskilde, Denmark. [26]Henan Province Engineering Research Center for Biomass Value-added Products, School of Forestry, Henan Agricultural University, Zhengzhou, Henan, China. [27]Centre for Palaeogenetics, Stockholm, Sweden. [28]Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), Faculty of Applied Sciences, AIMST University, Kedah, Malaysia. [29]Catalan Institution of Research and Advanced Studies, Barcelona, Spain. [30]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology, Barcelona, Spain. [31]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Barcelona, Spain. [32]University Museum, Norwegian University of Science and Technology, Trondheim, Norway.

*These authors contributed equally to this work.

†Corresponding author. Email: mhssinding@gmail.com (M.-H.S.S.); tomas.marques@upf.edu (T.M.-B.); ajhansen@sund.ku.dk (A.J.H.); tgilbert@sund.ku.dk (M.T.P.G.)

‡These authors cosupervised this work.

with the Eurasian wolf (fig. S15), suggesting that gene flow from modern wolves has not contributed to the sled dog gene pool within the past 9500 years. This result was surprising given genetic evidence for postdomestication admixture between other wolves and dog breeds (*5*, *7*). Furthermore, ethnographic evidence from Greenland indicates that, at least historically, dog-wolf matings were not uncommon (*8*). If true, then the lack of gene flow from modern American-Arctic wolves into sled dogs implies selection against hybrids.

The clustering and admixture results show gene flow between some sled dogs and other modern dog breeds (Fig. 1C and figs. S6 to S8). We further explored this by comparing pairs of sled dogs with Zhokhov using *D* statistics (Fig. 2A). Although pairs of Greenland sled dogs are symmetrically related to Zhokhov (*D*~0), indicating a lack of admixture, comparisons involving non-Greenland sled dogs were not always consistent with the null hypothesis of no admixture. *D*-statistics and admixture analyses (Fig. 2B and fig. S13) indicated that non-Greenland sled dogs carry ancestry from non-sled dogs and that Greenland sled dogs are the least admixed. These results imply that Greenland sled dogs have largely been kept isolated from contact with other dog breeds, and that their lineage traces more genomics ancestry to Zhokhov-like dogs relative to other dog breeds. Isolation of Greenland sled dogs was supported by inference of their historical effective population size (fig. S16), which showed that these dogs had a relatively stable population size until a severe bottleneck ~850 years ago. The timing of the bottleneck is consistent with the colonization of Greenland by Inuit (*9*), suggesting isolation in Greenland ever since.

Numerous generations of sled dogs living in the Arctic environment and being used as draft animals may have provided a unique selection pressure to these dogs. To detect putative signals of positive selection, we used population branch statistics (PBS) (*10*) to scan for genomic regions highly differentiated in modern sled dogs relative to non-sled dogs (hereafter referred to as "other dogs") and wolves. We computed these statistics on modern genomes of 17 sled dogs, 61 other dogs,

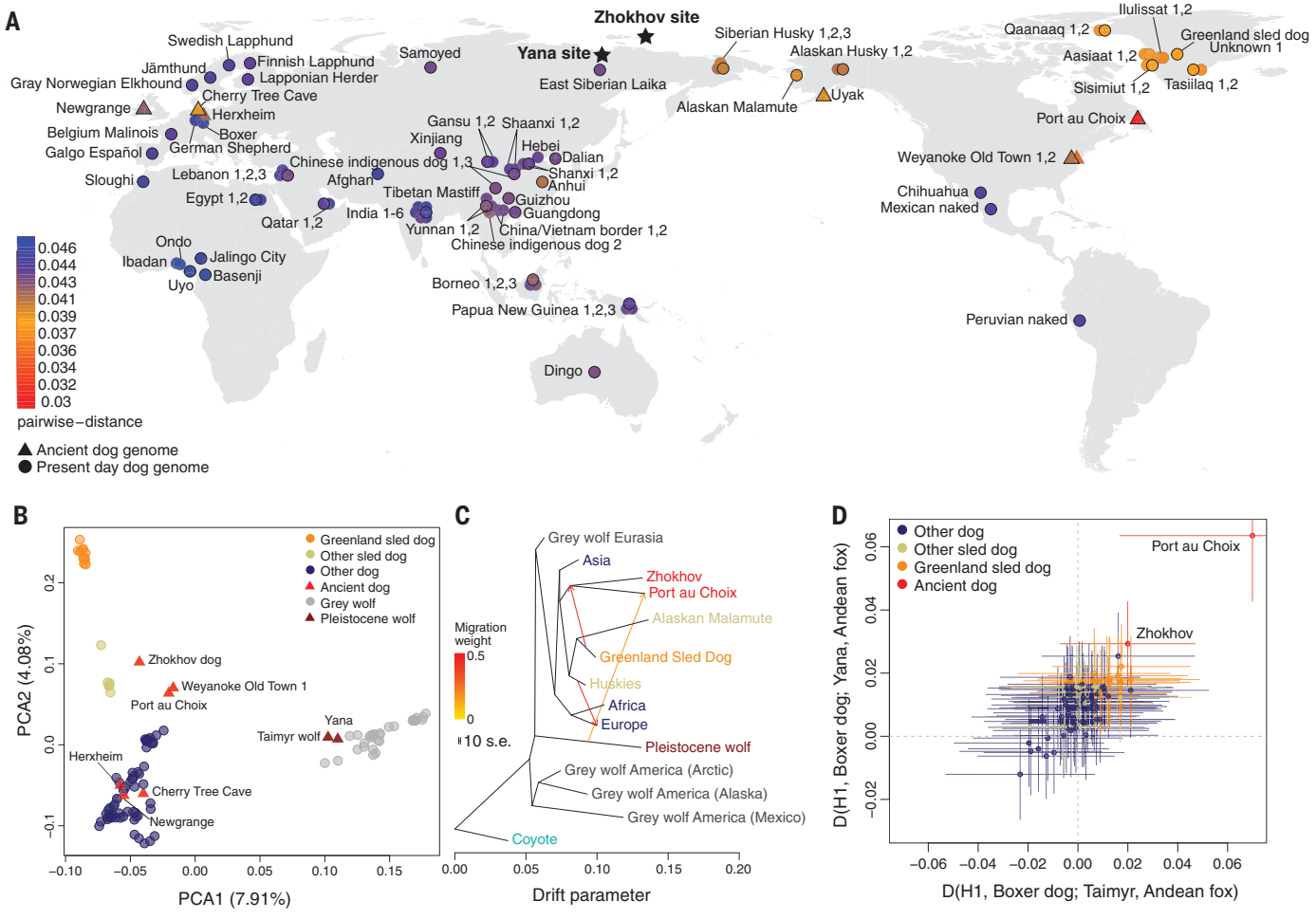


Fig. 1. Geographic location of the samples and overall genetic affinities. (**A**) Identity by state pairwise distances between Zhokhov and present-day dogs (table S1) of geographic affiliation of dogs and archaeological sites. Color scale indicates genetic distance between Zhokhov and each sample. Circles and triangles represent modern and ancient dogs, respectively. Stars show Zhokhov and Yana sites. (**B**) Principal component analysis (PCA) using whole-genome data (2,200,623 transversion sites) on all samples. (**C**) TreeMix admixture graph built using whole-genome data (766,082 transversion sites) on a dataset consisting of 66 canids merged into 15 groups according to their geographic location and admixture profile (table S1 and fig. S6). Colors indicate main groups as in (B). Arrows show inferred admixture edges colored by migration weight. (**D**) *D* statistic of the form *D*(H1, boxer dog; Taimyr or Yana, Andean fox) testing for Pleistocene wolf gene flow in ancient and modern dogs and whether samples share more alleles with Taimyr (*x*-axis) or Yana (*y*-axis) wolves when compared with the boxer dog. Color indicates the type of sample in H1. Points show the *D* statistic, and horizontal and vertical lines show 3 SEs for the test with the Taimyr (*x*-axis) and Yana (*y*-axis), respectively. The results obtained from both ancient wolves fall along the diagonal, suggesting that they are symmetrically related to all dogs.

and 30 wolves (table S1). A sliding window analysis revealed several genomic regions with high PBS values, hinting at selection in sled dogs (Fig. 3A). We took an outlier approach and focused on the most extreme values of the empirical distribution (above the 99.95th percentile). For each of these outlier regions (table S4), we identified overlapping genes and compared haplotypes across samples.

Enrichment analysis (*4*) on genomic regions with high PBS values (above the 99.95th percentile) identified three gene ontology (GO) terms that were overrepresented (table S6): γ-aminobutyric acid secretion (GO: 0014051, p = 0.119), calcium ion import (GO: 0070509, p = 0.119), and calcium ion transmembrane transport (GO: 0070588, p = 0.382). To investigate further, we focused on eight genomic regions that are highly differentiated in sled dogs and three regions where other dogs differ from sled dogs and wolves (Fig. 3A and fig. S18), and validated the autosomal regions with a cross-population composite likelihood ratio statistic (*5*) (fig. S21). In the differentiated regions, we focused on two sets of genes: those in which Zhokhov carries the same haplotype as modern sled dogs and those involved in adaptation to different diets.

TRPC4 is highly differentiated in sled dogs, and the putatively selected haplotype bears a marked similarity to Zhokhov (Fig. 3, A and B). *TRPC4* is a transient receptor potential (TRP) channel protein that plays an important role in vasorelaxation and lung microvascular permeability (*11*). It is also involved in a temperature sensitivity pathway (*12*, *13*), where it interacts with *TRPV2*, which is also highly differentiated in sled dogs (99.8th PBS percentile; table S4 and fig. S19A) and codes for temperature and potentially pain receptors (*14*). Several related thermo-TRP sensors in the same pathway, calcium ion transmembrane transport, have been previously reported to be under selection in cold-adapted woolly mammoths (*15*), which suggests convergent evolution in Arctic adaptation.

Another highly differentiated gene in sled dogs is *CACNA1A* (Fig. 3, A and C), a calcium channel subunit that plays an essential role in skeletal muscle contraction (*16*). Further, *CACNA1A* has been reported to be under positive selection in humans, specifically the Bajau sea nomads (*17*), where it is involved in hypoxia adaptation (*18*), indicating a possible role in managing exercise-induced hypoxia in sled dogs. We hypothesize that the *TRPC4*, *TRPV2*, and *CACNA1A* genes are involved in functions beneficial to physical activity in the Arctic. If so, given that the differentiated haplotypes are also found in Zhokhov (Fig. 3, A and B, and fig. S19A), any advantages that they confer would have been important to dogs in the Arctic ~9500 YBP.

Most domestic dogs are adapted to starch-rich diets through marked increases in *AMY2B* copy numbers and strong positive selection for a dog-specific *MGAM* haplotype (*19*). Consistent with previous findings (*20*), we observed that sled dogs carry substantially fewer *AMY2B* copies than other dog breeds (fig. S20). We also found that *MGAM* and *AMY2B* are the regions of the genome with the lowest PBS, suggesting high differentiation of other dogs relative to sled dogs and wolves (Fig. 3A). Because negative PBS can arise under different demographic scenarios, we confirmed these observations by computing PBS with other dogs as the focal population (fig. S18). Indeed, modern sled dogs and Zhokhov are among the only dogs in our dataset that carry the ancestral *MGAM* haplotype found at high frequency in wolves (Fig. 3C and fig. S18). Therefore, our observations suggest that sled dogs do not carry the genetic adaptations to starch-rich diets seen in other dog breeds.

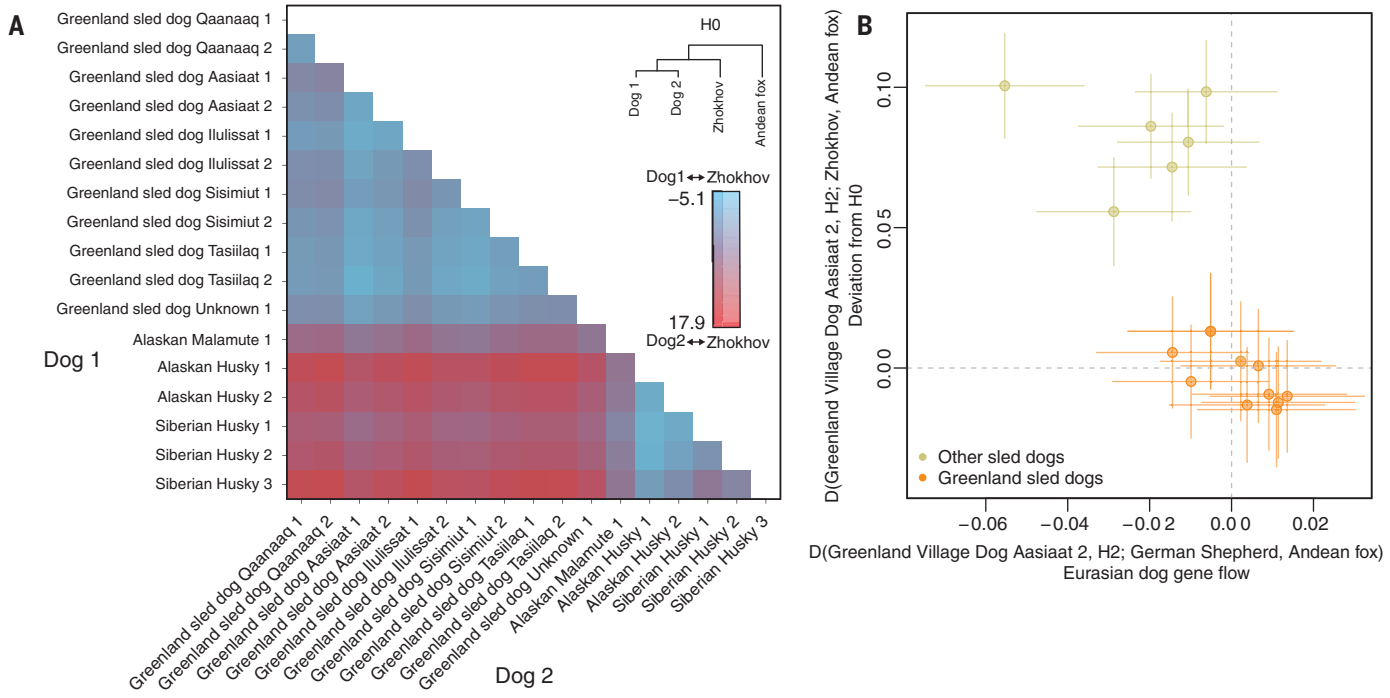


Fig. 2. Relationships between Zhokhov and present-day sled dogs. (**A**) *D* statistics testing the relationships between pairs of sled dogs and Zhokhov. Cell colors indicate the *Z* scores obtained from the test *D*(dog1, dog2; Zhokhov, Andean fox), where dog1 and dog2 are all possible pairs of sled dogs. Comparisons involving pairs of Greenland sled dogs and non-Greenland sled dogs resulted in significant deviations from H0 (|*Z*| > 3). (**B**) *D* statistics showing that sled dogs that are significantly further from Zhokhov compared with Greenland sled dog Aasiaat 2 [*y*-axis: *D*(Greenland sled dog Aasiaat 2, H2; Zhokhov, Andean fox)] also show evidence of significant gene flow from other dogs [*x*-axis: *D*(Greenland sled dog Aasiaat 2, H2; German shepherd dog, Andean fox)]. Points indicate the *D* statistic, and horizontal and vertical lines indicate 3 SEs for the *x*-axis and *y*-axis, respectively. We consider the test to be significant for gene flow when these lines do not overlap with the dotted line (|*Z*| > 3).

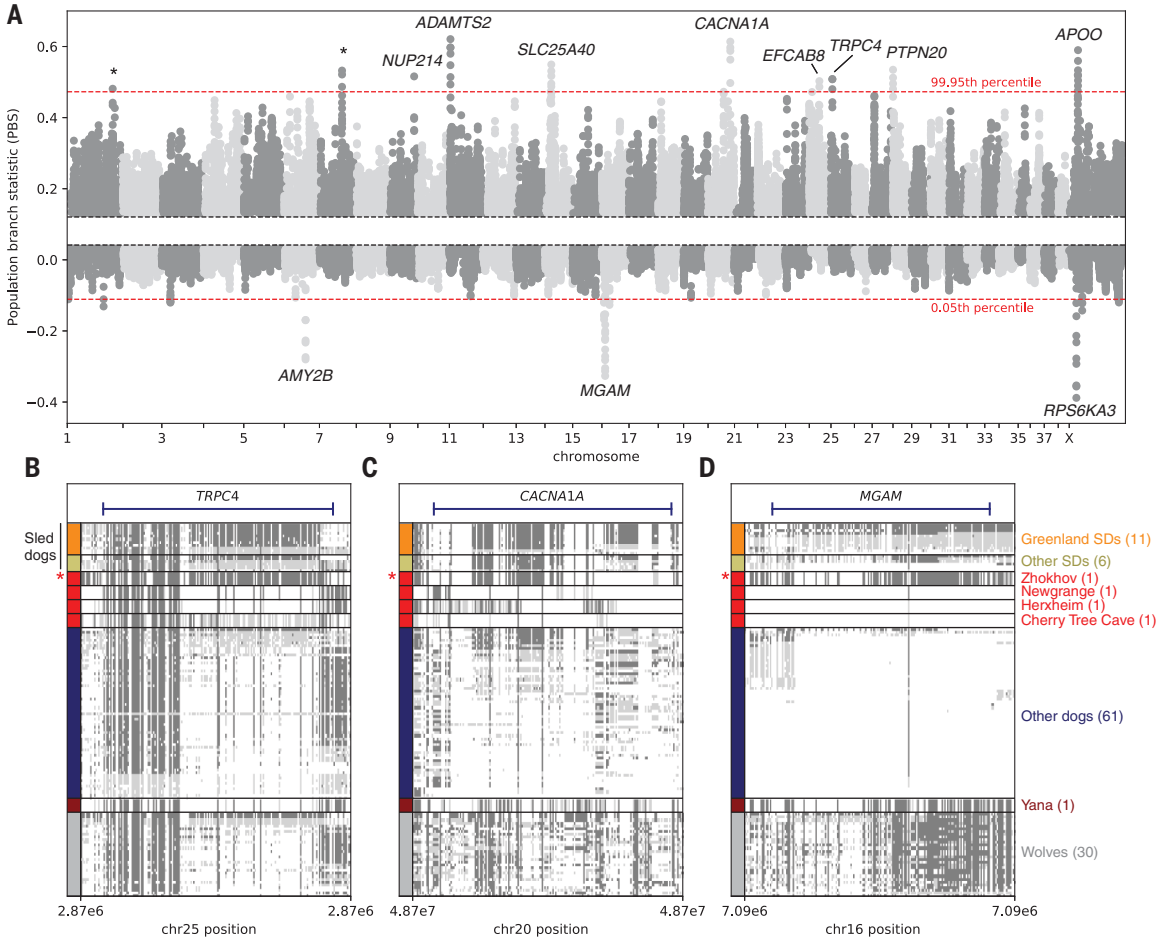


Fig. 3. Adaptation. (**A**) Manhattan plot of the PBS values (*y*-axis) in windows of 100 kg base pairs (kb) using a 20-kb slide across chromosomes (*x*-axis). Data points between the 20th and 80th percentile of the empirical distribution are not plotted and dashed red lines show the 99.95th and 0.05th percentiles. Names of genes within the highest peaks are shown, with asterisks representing no overlap with genes. We note that other genes not displayed in the figure can overlap the outlier regions; a full list can be found in tables S4 and 5. (**B** to **D**) Haplotype structures for TRPC4 (B), CACNA1A (C), and MGAM (D). Rows represent individuals, and columns represent polymorphic positions in the dog genome. Cells are colored by genotype: Dark gray indicates that the alternative allele is homozygous, light gray that it is heterozygous, and white that the reference allele is homozygous. The row height for ancient individuals was increased to facilitate visualization. Zhokhov is highlighted with a red asterisk. SDs, sled dogs.

By contrast, sled dogs harbor specific haplotypes of genes involved in coping with a high intake of fatty acids. *SLC25A40*, a mitochondrial carrier protein involved in clearing triglycerides from the blood (*21*), and *APOO*, an apolipoprotein gene involved in regulating high levels of fat and fatty acid metabolism (*22*), are both highly differentiated in sled dogs (Figs. 3A). The derived haplotypes of both genes are absent in Zhokhov, indicating that the haplotypes are specific to modern sled dogs and postdate their common ancestors with Zhokhov (fig. S19, B and E). As another example of convergent evolution, another gene of the apolipoprotein family, *APOB*, is reported to be under selection in polar bears, possibly as a result of adaptation to fat-rich diets and clearance of cholesterol from the blood (*23*). Overall, similar adaptations to high intake of fatty acids have been described in Inuit and other Arctic human populations (*24*, *25*), so our observations suggest that sled dogs adapted to a fat-rich and starch-poor diet, echoing the dietary adaptations of the Arctic human cultures with whom they coexisted.

Bone composition of polar bears and reindeer consumed at the Zhokhov site indicate an extensive hunting range and transport of large body parts back to camp (*26*). Further, abundant obsidian tools found at the site reveal movement of obsidian from ~1500 km away (*3*). Together, these findings indicate substantial long-distance travel and transportation of resources, in which dog sledding would have been highly advantageous—if not necessary. Putative sled remains and our genomic analyses of a 9500-year-old dog from the Zhokhov site indicate that the traditions and key genomic variations that define modern sled dogs were established in the northeast Asian Arctic >9500 years ago. Our results imply that the combination of these dogs with the innovation of sled technology facilitated human subsistence since the earliest Holocene in the Arctic.

REFERENCES AND NOTES

1. V. V. Pitulko, A. K. Kasparov, *Arctic Anthropol.* **33**, 1–36 (1996).
2. V. V. Pitulko, A. K. Kasparov, *J. Archaeol. Sci. Rep.* **13**, 491–515 (2017).
3. M. Ní Leathlobhair *et al.*, *Science* **361**, 81–85 (2018).
4. See the supplementary materials.
5. P. Skoglund, E. Ersmark, E. Palkopoulou, L. Dalén, *Curr. Biol.* **25**, 1515–1519 (2015).
6. A. Baez-Ortega *et al.*, *Science* **365**, eaau9923 (2019).
7. Z. Fan *et al.*, *Genome Res.* **26**, 163–173 (2016).
8. B. Muus, F. Salomonsen, C. Vibe, *Grønlands Fauna* [in Danish] (Gyldendal, 1981).
9. M. Raghavan *et al.*, *Science* **345**, 1255832 (2014).
10. X. Yi *et al.*, *Science* **329**, 75–78 (2010).
11. C. Tiruppathi *et al.*, *Circ. Res.* **91**, 70–76 (2002).
12. T. Hofmann, M. Schaefer, G. Schultz, T. Gudermann, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7461–7466 (2002).
13. K. Zimmermann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18114–18119 (2011).
14. N. Qin *et al.*, *J. Neurosci.* **28**, 6231–6238 (2008).
15. V. J. Lynch *et al.*, *Cell Rep.* **12**, 217–228 (2015).
16. S. Kaja *et al.*, *Eur. J. Neurosci.* **25**, 2009–2020 (2007).
17. M. A. Ilardo *et al.*, *Cell* **173**, 569–580.e15 (2018).
18. V. Wang, D. A. Davis, M. Haque, L. E. Huang, R. Yarchoan, *Cancer Res.* **65**, 3299–3306 (2005).
19. E. Axelsson *et al.*, *Nature* **495**, 360–364 (2013).
20. M. Arendt, K. M. Cairns, J. W. O. Ballard, P. Savolainen, E. Axelsson, *Heredity* **117**, 301–306 (2016).
21. E. A. Rosenthal *et al.*, *Am. J. Hum. Genet.* **93**, 1035–1045 (2013).
22. A. Turkieh *et al.*, *J. Clin. Invest.* **124**, 2277–2286 (2014).
23. S. Liu *et al.*, *Cell* **157**, 785–794 (2014).
24. A. Cardona *et al.*, *PLOS ONE* **9**, e98076 (2014).
25. M. Fumagalli *et al.*, *Science* **349**, 1343–1347 (2015).
26. V. V. Pitulko, V. V. Ivanova, A. K. Kasparov, E. Y. Pavlova, *Environ. Archaeol.* **20**, 120–157 (2015).

ACKNOWLEDGMENTS

We thank J. A. Leonard and B. von Holdt for input and comments in the conceptualization of this study, the Danish National High-Throughput Sequencing Centre and BGI-Europe for assistance in Illumina data generation, and the Danish National Supercomputer for Life Sciences – Computerome (https://computerome.dtu.dk) for the computational resources to perform the sequence analyses. **Funding:** This work is embedded in "The Qimmeq Project," funded by the Velux Foundations and Aage og Johanne Louis-Hansens Fond, and supported by ArchSci2020, funded by the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Marie Curie Actions grant no. 676154. We thank the Rock Foundation of New York for funding excavations at the Zhokhov and Yana sites in a 15-year-long effort starting in 2000. M.-H.S.S. was supported by the Independent Research Fund Denmark (8028-00005B) and NHM Oslo. S.G. was supported by Marie Skłodowska-Curie Actions (H2020 655732 - WhereWolf) and Carlsberg (CF14 - 0995). M.d.M.M. was supported by a Formació de Personal Investigador fellowship from Generalitat de Catalunya (FI_B01111). V.V.P., E.Y.P., and P.A.N. were supported by the Russian Science Foundation project no. 16-18-10265-RNF. T.M.B. was supported by BFU2017-86471-P (MINECO/FEDER, UE), Howard Hughes International Early Career, Obra Social "La Caixa" and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 880). M.T.P.G. was supported by a European Research Council grant (ERC-2015-CoG-681396–Extinction Genomics). G.L. and L.A.F. were supported by the ERC (Grant ERC-2013-StG-337574-UNDEAD) and the Natural Environmental Research Council (Grants NE/K005243/1 and NE/K003259/1). P.S. was supported by the Francis Crick Institute (FC001595).

Author contributions: M.-H.S.S., S.G., J.R.-M., M.d.M.M., and M.T.P.G. conceived of the project and designed the research. V.V.P., E.Y.P., P.A.N., A.K.K., V.V.I., and E.W. provided archaeological work, logistics, and/or ancient collected samples. M.-H.S.S., M.F., S.E.W., M.P.H.-J., R.D., and C.S. coordinated logistics of and/or provided modern samples. C.C. and M.-H.S.S. conducted the laboratory work. S.G., J.R.-M., M.d.M.M., L.K., L.A.F.F., F.G.V., J.N., and J.A.S.C. conducted the analyses of data with considerable input from M.-H.S.S., B.P., T.S.-P., T.M.-B., A.J.H., and M.T.P.G. S.G., J.R.-M., M.d.M.M., L.K., L.A.F.F., F.G.V., J.N., J.A.S.C., P.S., M.-H.S.S., T.M.-B., A.J.H., and M.T.P.G. interpreted results with considerable input from B.P., T.S.-P., V.V.P., T.R.F., E.U.A.-R., P.D.J., M.M., L.D., G.L., L.B., and Ø.W. M.-H.S.S., S.G., J.R.-M., M.d.M.M., and M.T.P.G. wrote the paper with input from all other authors.

Competing interests: The authors declare no competing interests.

Data and materials availability: Raw sequencing data can be accessed at the NCBI Short Read Archive under project number PRJNA608847.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6498/1495/suppl/DC1
Materials and Methods
Figs. S1 to S21
Tables S1 to S6
References (*27*–*70*)
MDAR Reproducibility Checklist

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

15 October 2019; accepted 6 May 2020
10.1126/science.aaz8599

Corrected 25 June 2020. See full text.

CORONAVIRUS

Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir

Wanchao Yin[1,2]*, Chunyou Mao[2]*, Xiaodong Luan[3,4,5]*, Dan-Dan Shen[2]*, Qingya Shen[2]*, Haixia Su[1,6]*, Xiaoxi Wang[1], Fulai Zhou[1], Wenfeng Zhao[1], Minqi Gao[7], Shenghai Chang[8,9], Yuan-Chao Xie[1], Guanghui Tian[1], He-Wei Jiang[10], Sheng-Ce Tao[10], Jingshan Shen[1,6], Yi Jiang[1,6], Hualiang Jiang[1,6], Yechun Xu[1,6]†, Shuyang Zhang[4,5,3]†, Yan Zhang[2,11]†, H. Eric Xu[1,6]†

The pandemic of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has become a global crisis. Replication of SARS-CoV-2 requires the viral RNA-dependent RNA polymerase (RdRp) enzyme, a target of the antiviral drug remdesivir. Here we report the cryo–electron microscopy structure of the SARS-CoV-2 RdRp, both in the apo form at 2.8-angstrom resolution and in complex with a 50-base template-primer RNA and remdesivir at 2.5-angstrom resolution. The complex structure reveals that the partial double-stranded RNA template is inserted into the central channel of the RdRp, where remdesivir is covalently incorporated into the primer strand at the first replicated base pair, and terminates chain elongation. Our structures provide insights into the mechanism of viral RNA replication and a rational template for drug design to combat the viral infection.

The coronavirus disease 2019 (COVID-19) pandemic that has arisen from widespread severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection has become a humanitarian crisis, with more than 1.5 million infections and 87,000 deaths reported on 8 April 2020 (*1*, *2*). These numbers have increased rapidly to more than 2.99 million infections and 207,000 deaths as of 27 April of 2020 (*2*). SARS-CoV-2 is closely related to severe acute respiratory syndrome coronavirus (SARS-CoV) and several members of the betacoronavirus family, including bat and pangolin coronaviruses (*3*–*5*). Compared with the binding behavior of other coronaviruses, however, the spike protein of SARS-CoV-2 has a stronger binding affinity for the host receptor (*6*–*10*), which may explain why SARS-CoV-2 has a much higher incidence of human-to-human transmission, resulting in infections throughout the world.

[1]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. [2]Department of Biophysics, and Department of Pathology of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China. [3]School of Medicine, Tsinghua University, Haidian District, Beijing 100084, China. [4]Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China. [5]Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China. [6]University of Chinese Academy of Sciences, Beijing 100049, China. [7]WuxiBiortus Biosciences Co. Ltd., Jiangyin 214437, China. [8]Center of Cryo-Electron Microscopy, Zhejiang University School of Medicine, Hangzhou 310058, China. [9]Center of Diagnostic Electron Microscopy, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China. [10]Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200240, China. [11]Key Laboratory of Immunity and Inflammatory Diseases of Zhejiang Province, Hangzhou 310058, China.
*These authors contributed equally to this work.
†Corresponding author. Email: eric.xu@simm.ac.cn (H.E.X.); zhang_yan@zju.edu.cn (Y.Z.); shuyangzhang103@nrdrs.org (S.Z.); ycxu@simm.ac.cn (Y.X.)

SARS-CoV-2 is a positive-strand RNA virus. Its replication is mediated by a multisubunit replication-and-transcription complex of viral nonstructural proteins (nsps) (*11*). The core component of this complex is the catalytic subunit (nsp12) of an RNA-dependent RNA polymerase (RdRp) (*12*, *13*). By itself, nsp12 has little activity and its functions require accessory factors, including nsp7 and nsp8 (*14*, *15*), that increase RdRp template binding and processivity. RdRp is also proposed to be the target of a class of antiviral drugs that are nucleotide analogs; this category includes remdesivir (*16*–*18*), which is a prodrug that is converted to the active drug in the triphosphate form [remdesivir triphosphate (RTP)] within cells (*19*). As such, RdRp has been the subject of intensive structural biology efforts. The structures of nsp7, nsp8, and the complex of nsp12-nsp7-nsp8 have been determined (*15*, *20*–*23*), providing the overall architecture of the RdRp complex. However, the drug discovery effort is hampered because the structures of the SARS-CoV-2 RdRp in complex with an RNA template and with nucleotide inhibitors are not known. In this study, we determined two cryo–electron microscopy (cryo-EM) structures of the SARS-CoV-2 RdRp complex: one in the apo form and the other in a complex with a template-primer RNA and the antiviral drug remdesivir.

For cryo-EM studies, we coexpressed nsp12 with nsp7 and nsp8 to form the core RdRp complex in insect cells (Fig. 1A and fig. S1, A to D). The stoichiometric amount of nsp7 and nsp8 appeared to be less than that of nsp12, and thus additional bacterially expressed nsp7 and nsp8 were supplemented before the final purification step to improve

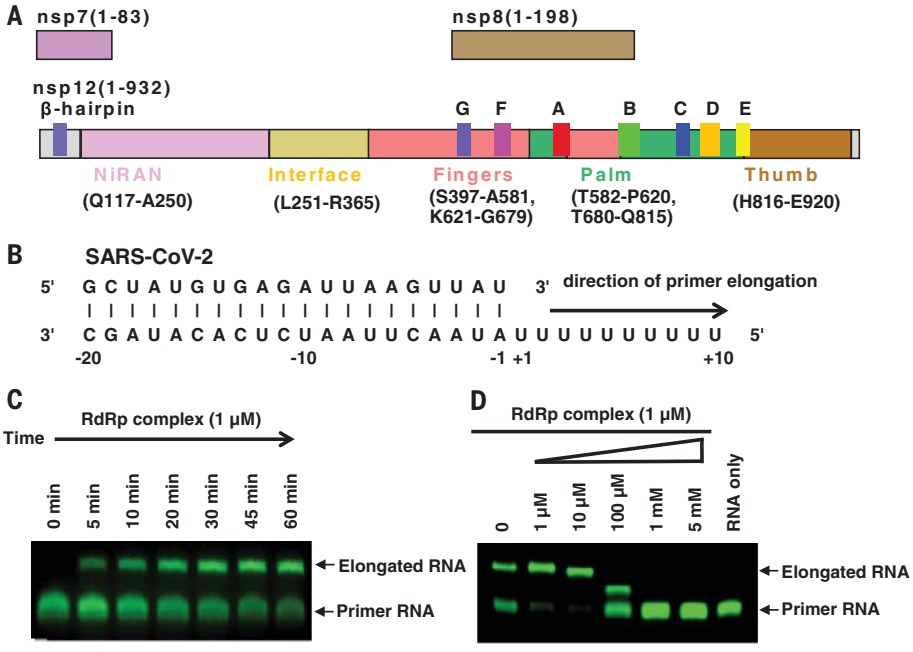


Fig. 1. Assembly of an active nsp12-nsp7-nsp8 RdRp complex and its inhibition by remdesivir. (**A**) Schematic diagram for the components of the RdRp complex, containing nsp12, nsp7, and nsp8. The polymerase motif (A to G) and the β hairpin specific to SARS-CoV-2 are highlighted. (**B**) Sequence of the RNA duplex with a 5′ U_{10} overhang as a template for primer extension and RdRp-RNA complex assembly. (**C**) The recombinant RdRp complex shows polymerase activity in vitro. The primer strand is labeled with fluorescence at the 5′ end. (**D**) Elongation of the partial RNA duplex by the purified RdRp complex and its inhibition by RTP.

Corrected 25 June 2020. See full text.

the yield of heterotrimeric complex. The purified nsp12 alone showed little activity in binding to a 50-base partial double-stranded template-primer RNA (Fig. 1B and fig. S1E), which is similar to the SARS-CoV nsp12 (*14*). The presence of nsp7 and nsp8 markedly increased nsp12 binding to the template-primer RNA (fig. S1E). The nsp12-nsp7-nsp8 complex also showed RNA polymerization activity on a poly-U template upon addition of adenosine triphosphate (ATP) (Fig. 1, B and C). This RNA polymerization activity was effectively inhibited by the addition of RTP (Fig. 1D). Even in the presence of 10 mM ATP, which is within the range of physiological concentrations of ATP, 1 mM RTP completely inhibited RdRp polymerization activity. In addition, 100 μM RTP completely blocked the full extension but allowed partial extension of the primer RNA (Fig. 1D), consistent with a delayed chain termination mechanism (*24*). However, this mechanism is dependent on low RTP concentrations and low RTP/ATP ratios. By contrast, 5 mM remdesivir (as a prodrug) had no inhibitory effect on the polymerization activity of the purified enzyme (fig. S1F), nor did remdesivir in its monophosphate form (RMP) (fig. S1G).

The purified RdRp complex is relatively thermostable, with a melting temperature of 53°C (fig. S1H). Negative-stain EM visualization of the nsp12-nsp7-nsp8 complex displayed monodispersed particles with excellent homogeneity (fig. S1I). For the apo nsp12-nsp7-nsp8 complex, we vitrified the sample in the presence of the detergent DDM. The initial attempt at image processing revealed that the particles are preferentially oriented (fig. S2A). Therefore, we collected >7400 micrograph movies of more than 5.7 million particle projections to increase the number of projections from the nonpreferential orientation. Of these, 81,494 particles were used to yield a density map of 2.8-Å resolution (fig. S2, B and E). Cryo-EM studies of the nsp12-nsp7-nsp8 complex bound with the template-primer RNA and RTP (termed the template-RTP RdRp complex) faced two challenges (fig. S3). First, most particles were adsorbed to cryo-EM grid bars rather than staying in the vitreous ice. Second, the RNA duplex was dissociated from the template-RTP RdRp complex, likely owing to conditions of cryo-EM specimen preparation. Eventually, we prepared the cryo-EM specimen of the template-RTP RdRp complex at 15 mg/ml, which is much higher than the normal concentrations used for EM studies of soluble protein complexes. The high concentration of the complex has a mass action effect to stabilize the RNA-protein complex and has an excess amount of the complex to escape the absorption of cryo-EM grid bars to enter the vitreous ice (fig. S3). We collected 2886 micrograph movies, which yielded a 2.5-Å resolution structure using 130,386 particle projections. Because of the relatively high resolution of our structure, the EM map is clear for key structural features across the complex (fig. S4, A to F).

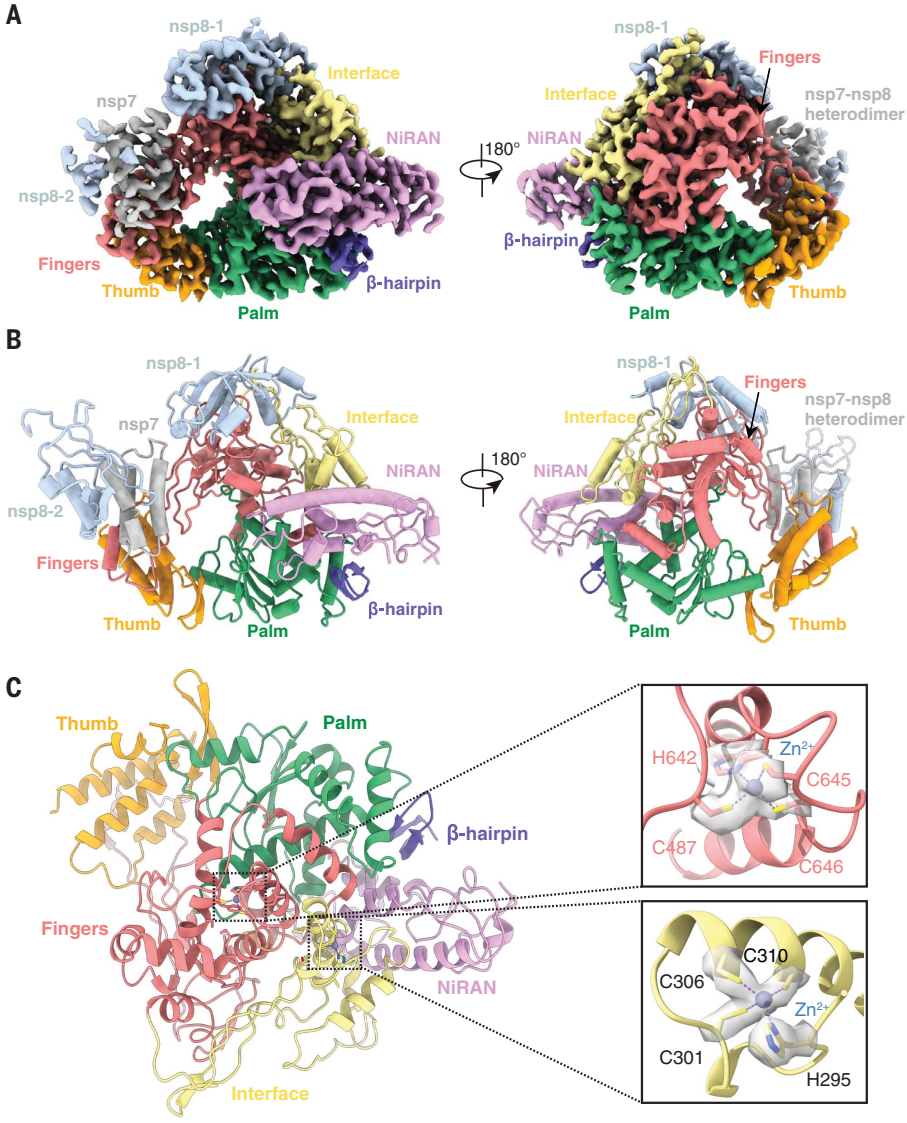


Fig. 2. Cryo-EM structure of the apo nsp12-nsp-7-nsp8 RdRp complex. (**A** and **B**) Two views of the cryo-EM map (A) and structure (B) of the apo nsp12-nsp7-nsp8 complex. The color scheme is according to Fig. 1A and is used throughout the figures. (**C**) The conserved zinc binding motifs are highlighted in the apo structure rendered in ribbon. The coordinate details of the zinc-binding residues are shown in stick representation, with the EM map in gray surface representation.

The structure of the apo RdRp complex contains one nsp12, one nsp7, and two nsp8 proteins, with an overall arrangement resembling those seen in SARS-CoV and the recently solved structure of SARS-CoV-2 (*15*, *23*) (Fig. 2, A and B). Our structure, which differs from the SARS-CoV RdRp structure but is similar to the recent SARS-CoV-2 RdRp structure, reveals that nsp12 also contains an N-terminal β hairpin (residues 31 to 50) and an extended nidovirus RdRp-associated nucleotidyl-transferase domain (NiRAN; residues 115 to 250) (*24*) with seven helices and three β strands (*15*, *23*). After the NiRAN domain is an interface domain (residues 251 to 365), composed of three helices and five β strands, that is connected to the RdRp domain (residues 366 to 920) (Figs. 1A and 2B). The nsp12 RdRp domain displays the canonical cupped right-handed configuration (*25*), with the finger subdomain (residues 397 to 581 and 621 to 679) forming a closed circle with the thumb subdomain (residues 819 to 920) (Fig. 2, A and B). The closed conformation is stabilized by the binding of

Corrected 25 June 2020. See full text.

nsp7 and nsp8, with one nsp8 molecule sitting atop the finger subdomain and interacting with the interface domain. The closed conformation of nsp12 is further stabilized by the nsp7-nsp8 heterodimer, which is packed against the thumb-finger interface (Fig. 2, A and B). In addition, we were able to assign two zinc ions in the conserved metal binding motifs composed of H295-C301-C306-C310 and C487-H642-C645-C646 (Fig. 2C), which are also observed in the SARS-CoV RdRp structure (*15*). These zinc ions likely function as conserved structural components in maintaining the integrity of the RdRp architecture.

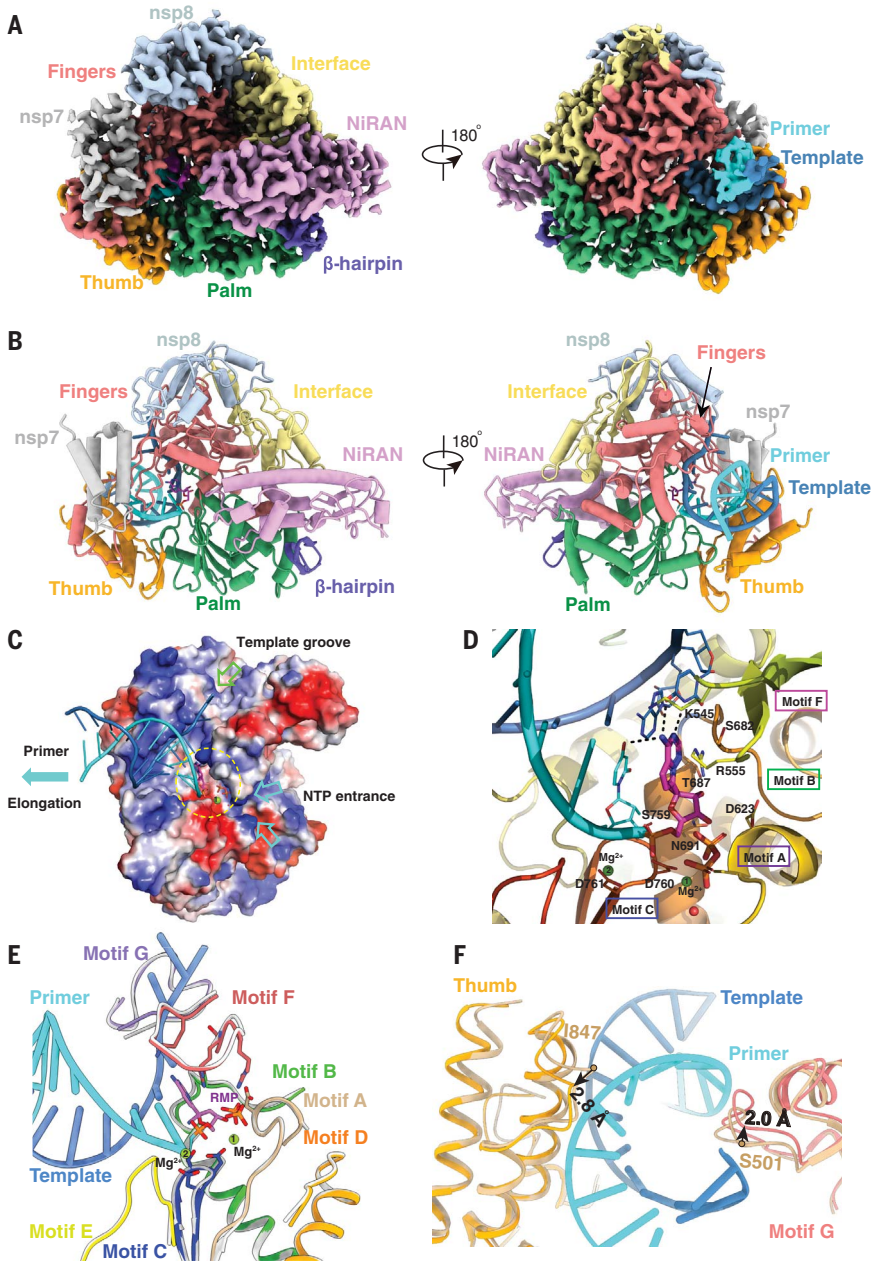


Fig. 3. Cryo-EM structure of the remdesivir- and RNA-bound RdRp complex. (**A** and **B**) Two views of the cryo-EM map (A) and structure (B) of nsp12-nsp7-nsp8 in complex with template-primer RNA and remdesivir. (**C**) Surface view of the RdRp active site with the electrostatic potential from red (negative) to blue (positive). For clarity, residues 410 to 442 and 834 to 919 of nsp12 and nsp8 are excluded from the figure. The covalently bound remdesivir in the monophosphate form and the product, pyrophosphate, are shown. The active site is emphasized with a yellow dashed circle. The template groove, the entrance for nucleotide triphosphate (NTP), and the elongation direction are annotated with different-colored arrows. (**D**) Close-up view of the RdRp active site, showing the covalently bound RMP, pyrophosphate, and magnesium ions. Key residues and bases that interact with remdesivir are shown. (**E** and **F**) Superposition of the conserved RdRp motifs (A to G) of the RNA-bound complex with the apo structure (colored in gray), with a close-up view at the active site (E) and at the exit of the template and primer strand (F).

The structure of the template-RTP RdRp complex contains one nsp12, one nsp7, and one nsp8 (Fig. 3, A and B). The second nsp8 was largely invisible in the EM map of the template-RTP complex (fig. S4C); therefore, it was not included in the final model. In addition, the template-RTP RdRp structure contains 14-base RNA in the template strand, 11-base RNA in the primer strand, and the inhibitor RMP (Fig. 3, C and D), which is covalently linked to the primer strand, as well as a pyrophosphate and two magnesium ions that may serve as catalytic ions near the active site (Fig. 3D and fig. S4, D to F) (*26*).

The overall structure of the template-RTP RdRp complex is similar to the apo RdRp structure, with nsp12 in a closed conformation (Figs. 2A and 3A). The double-stranded RNA helix, formed by 11 base pairs from the template-primer RNA (Figs. 3C and 4, A to E), is held by the finger-palm-thumb subdomains. Extensive protein-RNA interactions are observed between the template-primer RNA and nsp12, with a total of 41 residues from nsp12 directly participating in RNA binding (within 4.0 Å of RNA, 26 residues to the template strand and 15 residues to the primer strand; Fig. 4E). Surprisingly, no RNA interactions are mediated by nsp7 or nsp8, although these two proteins are required for RNA binding by RdRp. Most protein-RNA interactions involve the RNA phosphate-ribose backbones, with many interactions directly to 2′-OH groups (Fig. 4E), thus providing a basis to distinguish RNA from DNA. There are no contacts from nsp12 to any base pairs of the template-primer RNA, suggesting a sequence-independent binding of RNA by RdRp. This is consistent with the fact that no specific sequence is required for the enzymatic activity of RdRp at the elongation step.

At the 3′ end of the primer strand is RMP (Figs. 3D and 4, D and E, and fig. S4, E and F), which is covalently incorporated into the primer strand at the +1 position (Fig. 4E). Additional nucleotides at the +2 and +3 positions of the template strand interact with residues from the back of finger subdomain (Fig. 4, A and B). Despite the presence of excess RTP in complex assembly, only a single RMP is assembled into the primer strand, as observed in the structure. Consistent with the data from Fig. 1D, the primer extension is immediately terminated when RTP concentration is high and ATP/RTP ratio is low. Thus, remdesivir, like many nucleotide analog prodrugs, inhibits viral RdRp activity through nonobligate RNA chain termination, a mechanism that requires conversion of the parent drug to the active triphosphate form (*27*, *28*).

Corrected 25 June 2020. See full text.

The RMP position is at the center of the catalytic active site (Fig. 3D). As an adenosine monophosphate analog, RMP forms base-stacking interactions with an upstream base from the primer strand and two hydrogen bonds with the uridine base from the template strand (Fig. 3D and fig. S5). In addition, RMP forms interactions with side chains from K545 and R555. Near the bound RMP are two magnesium ions and a pyrophosphate. Both magnesium ions interact with the phosphate diester backbone and are part of the catalytic active site. The pyrophosphate is at the gate of the nucleotide entry channel to the active site and may block the entry of nucleotide triphosphate to the active side (Fig. 3, C and D).

The catalytic active site of the nsp12 RdRp consists of seven conserved motifs (A to G; Figs. 1A and 3E and fig. S6). Motifs A, B, C, and D are from the palm subdomain, with an SDD sequence (residues 759 to 761) in motif C forming the catalytic active center (Fig. 3D). Both D760 and D761 are involved in coordination of the two magnesium ions at the catalytic center. Motifs F and G are located within the finger subdomain; they interact with the template strand RNA and direct this strand into the active site (Fig. 3E). Motif F also interacts with the primer strand RNA, with the side chains of K545 and R555 contacting the +1 base (Fig. 3D) and thus stabilizing the incoming nucleotide in the correct position for catalysis. The orientation of template-primer RNA in the active site is similar to the orientation of template-primer RNA in the poliovirus RdRp elongation complex (*29*) and the hepatitis C virus NS5B RdRp inhibitor complex (*30*) (fig. S7). The residues involved in RNA binding and those that constitute the catalytic active site are highly conserved (*31*, *32*), highlighting the conserved mechanism of genome replication by RdRp in these diverse RNA viruses and suggesting that it may be possible to develop broad-spectrum antiviral inhibitors such as remdesivir (*18*) and galidesivir (BCX4430) (*33*).

Structural comparison reveals several notable differences between the apo and complex structures (Fig. 3, E and F, and fig. S8, A and B). First, nsp7 moves toward the RdRp core by 1.5 Å (as measured by nsp7 residue F49; fig. S8, A and B), leading to a rearrangement of the interface, which results in weaker association of the second nsp8 in the complex. Second, the loop that connects the first and second helices of the thumb subdomain moves outward by 2.8 Å (as measured by nsp12 residue I847) to accommodate the binding of the double-stranded RNA helix (Fig. 3F). Third, motif G residues K500 and S501 also move outward by 2.0 Å to accommodate the binding of the template-strand RNA. Outside of these changes, the apo nsp12 and the RNA complex nsp12 are very similar, with a root mean square deviation of 0.52 Å for all Cα atoms across the whole protein. In particular, the structural elements that make up the catalytic active site can be precisely superimposed (Fig. 3E), which suggests that the SARS-CoV-2 RdRp is a relatively stable enzyme that is ready to function as a replicase upon RNA template binding. Viral RdRp is a highly processive enzyme with a replication rate of up to 100 nucleotides per second (*34*). No substantial conformational changes between the apo and active enzyme structures are consistent with the high processivity of the viral RNA polymerase, which does not need to consume additional energy for conformational changes in the active site during the replication cycle.

Besides remdesivir, several nucleotide analog drugs—including favipiravir, ribavirin, galidesivir, and EIDD-2801—efficiently inhibit SARS-CoV-2 replication in cell-based assays (*35*, *36*). Like remdesivir, these nucleotide analogs are proposed to inhibit the viral RdRp through nonobligate RNA chain termination, a mechanism that requires conversion of the parent compound to the triphosphate active

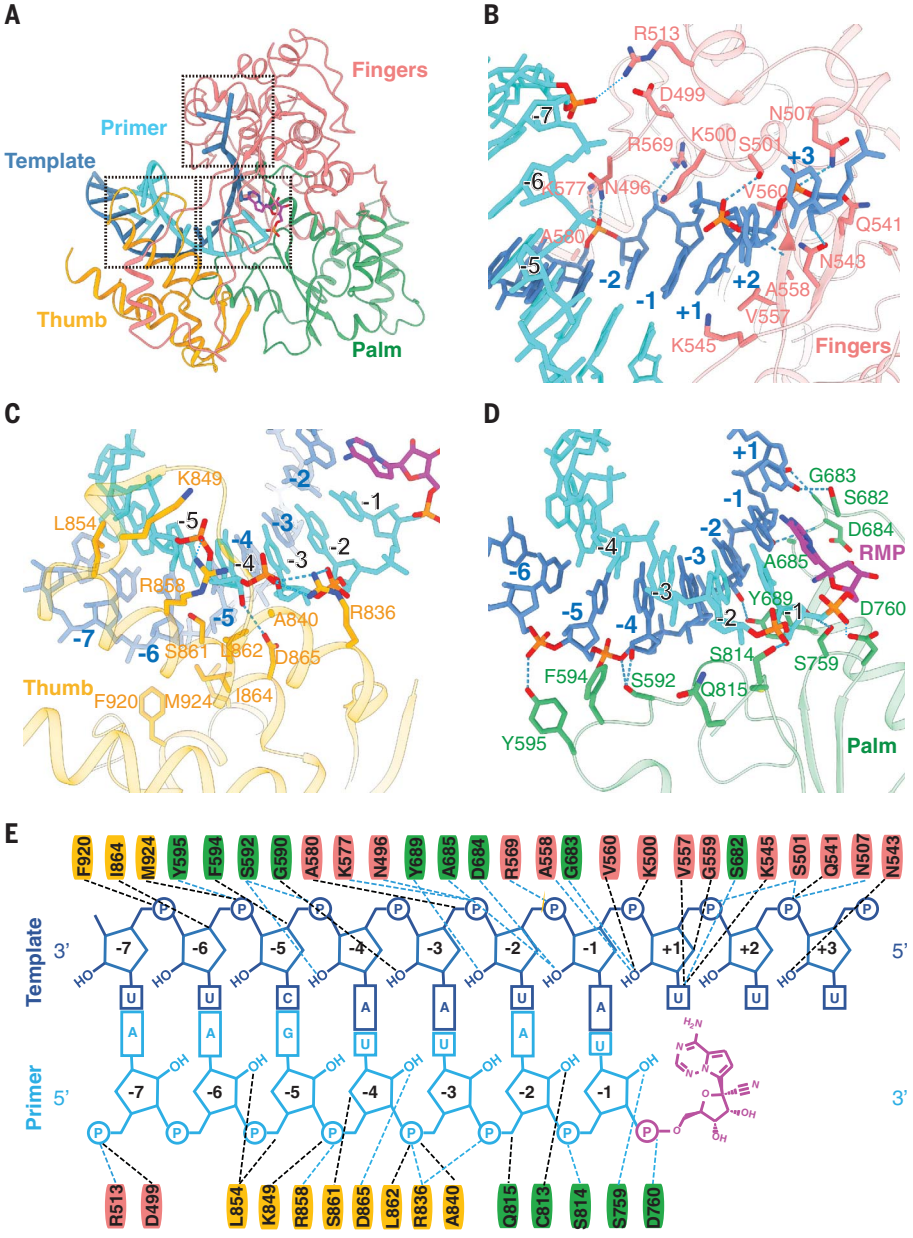


Fig. 4. RNA recognition by the RdRp complex. (**A** to **D**) Protein-RNA interactions in the RNA- and remdesivir-bound RdRp complex. (**E**) Diagram of detailed RNA interactions with RdRp. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Corrected 25 June 2020. See full text.

form (*33*). The structure of the template-RTP RdRp complex provides a useful model to rationalize how these drugs inhibit the activity of SARS-CoV-2 RdRp (fig. S8C). In particular, EIDD-2801 has been shown to be 3 to 10 times as potent as remdesivir in blocking SARS-CoV-2 replication (*36*). The N4 hydroxyl group off the cytidine ring forms an extra hydrogen bond with the side chain of K545, and the cytidine base also forms an extra hydrogen bond with the guanine base from the template strand. These two extra hydrogen bonds may explain the apparent higher potency of EIDD-2801 in inhibiting SARS-CoV-2 replication.

The COVID-19 pandemic has inflicted emotional pain and economic burden across the globe. Enzymes that are vital for the viral life cycle are suitable antiviral drug targets because they differ from host proteins. Among viral enzymes, RdRp is the primary target of many existing nucleotide drugs. In this paper, we report the structures of the SARS-CoV-2 RdRp complex in the apo form as well as in complex with a template-primer RNA and the active form of remdesivir. The structures reveal how the template-primer RNA is recognized by the enzyme and how chain elongation is inhibited by remdesivir. Structure comparison and sequence alignment suggest that the mode of substrate RNA recognition and remdesivir inhibition of RdRp is highly conserved in diverse RNA viruses, providing a foundation for designing broad-spectrum antiviral drugs based on nucleotide analogs. Moreover, our structures provide a solid template for modeling and modifying the existing nucleotide drugs, including the highly potent EIDD-2801. Together, these observations provide a rational basis to design even more potent inhibitors to combat SARS-CoV-2 infection.

REFERENCES AND NOTES

1. A. E. Gorbalenya *et al*., *Nat. Microbiol.* **5**, 536–544 (2020).
2. E. Dong, H. Du, L. Gardner, *Lancet Infect. Dis.* **20**, 533–534 (2020).
3. P. Zhou *et al*., *Nature* **579**, 270–273 (2020).
4. T. T.-Y. Lam *et al*., *Nature* 10.1038/s41586-020-2169-0 (2020).
5. T. Zhang, Q. Wu, Z. Zhang, *Curr. Biol.* **30**, 1346–1351.e2 (2020).
6. R. Yan *et al*., *Science* **367**, 1444–1448 (2020).
7. D. Wrapp *et al*., *Science* **367**, 1260–1263 (2020).
8. J. Lan *et al*., *Nature* **581**, 215–220 (2020).
9. J. Shang *et al*., *Nature* **581**, 221–224 (2020).
10. A. C. Walls *et al*., *Cell* **181**, 281–292.e6 (2020).
11. J. Ziebuhr, *Curr. Top. Microbiol. Immunol.* **287**, 57–94 (2005).
12. D. G. Ahn, J. K. Choi, D. R. Taylor, J. W. Oh, *Arch. Virol.* **157**, 2095–2104 (2012).
13. A. J. te Velthuis, J. J. Arnold, C. E. Cameron, S. H. van den Worm, E. J. Snijder, *Nucleic Acids Res.* **38**, 203–214 (2010).
14. L. Subissi *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3900–E3909 (2014).
15. R. N. Kirchdoerfer, A. B. Ward, *Nat. Commun.* **10**, 2342 (2019).
16. M. Wang *et al*., *Cell Res.* **30**, 269–271 (2020).
17. M. L. Holshue *et al*., *N. Engl. J. Med.* **382**, 929–936 (2020).
18. T. K. Warren *et al*., *Nature* **531**, 381–385 (2016).
19. D. Siegel *et al*., *J. Med. Chem.* **60**, 1648–1661 (2017).
20. Y. Zhai *et al*., *Nat. Struct. Mol. Biol.* **12**, 980–986 (2005).
21. W. Peti *et al*., *J. Virol.* **79**, 12905–12913 (2005).
22. M. A. Johnson, K. Jaudzems, K. Wüthrich, *J. Mol. Biol.* **402**, 619–628 (2010).
23. Y. Gao *et al*., *Science* **368**, 779–782 (2020).
24. K. C. Lehmann *et al*., *Nucleic Acids Res.* **43**, 8416–8434 (2015).
25. S. M. McDonald, *WIREs RNA* **4**, 351–367 (2013).
26. M. J. van Hemert *et al*., *PLOS Pathog.* **4**, e1000054 (2008).
27. E. P. Tchesnokov, J. Y. Feng, D. P. Porter, M. Götte, *Viruses* **11**, 326 (2019).
28. C. J. Gordon *et al*., *J. Biol. Chem.* **295**, 6785–6797 (2020).
29. P. Gong, O. B. Peersen, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22505–22510 (2010).
30. T. C. Appleby *et al*., *Science* **347**, 771–775 (2015).
31. A. J. te Velthuis, *Cell. Mol. Life Sci.* **71**, 4403–4420 (2014).
32. S. Venkataraman, B. V. L. S. Prasad, R. Selvarajan, *Viruses* **10**, 76 (2018).
33. T. K. Warren *et al*., *Nature* **508**, 402–405 (2014).
34. G. Campagnola, S. McDonald, S. Beaucourt, M. Vignuzzi, O. B. Peersen, *J. Virol.* **89**, 275–286 (2015).
35. C.-C. Lu, M.-Y. Chen, Y.-L. Chang, *J. Chin. Med. Assoc.* 10.1097/JCMA.0000000000000318 (2020).
36. T. P. Sheahan *et al*., *Sci. Transl. Med.* **12**, eabb5883 (2020).

ACKNOWLEDGMENTS

The cryo-EM data were collected at the Center of Cryo-Electron Microscopy, Zhejiang University. We thank MedChemExpress for making remdesivir. We also thank J. Richardson, D. Richardson, and T. Croll for help with validating the structures. **Funding:** This work was partially supported by the National Key R&D Programs of China 2018YFA0507002; Shanghai Municipal Science and Technology Major Project 2019SHZDZX02 and XDB08020303 to H.E.X.; Zhejiang University special scientific research fund for COVID-19 prevention and control E33 and the National Science Foundation of China 81922071 to Y.Z.; Science and Technology Commission of Shanghai Municipal 20431900100 and Jack Ma Foundation 2020-CMKYGG-05 to H.J. and J.S.; CAMS Innovation Fund for "13th Five-Year" National Science and Technology Major Project for New Drugs 2019ZX09734001-002, CAMS Innovation Fund for Medical Sciences no. 2020-I2M-CoV19-001, and Tsinghua University–Peking University Center for Life Sciences 045-160321001 to S.Z.; National Key Research and Development Program of China grant 2016YFA0500600 and National Natural Science Foundation of China 31970130 and 3167083 to S.-C.T.; National Key R&D Program of China 2016YFA0502301 to Y.X.; and National Natural Science Foundation 31770796 and National Science and Technology Major Project 2018ZX09711002 to Y.J. **Author contributions:** W.Y. designed the expression constructs, purified the RdRp complex, prepared samples for negative-stain EM and data collection toward the structures, and participated in figure and manuscript preparation. X.L. designed RdRp activity assays and remdesivir inhibition experiments as well as expression constructs of the RdRp complex. C.M. and D.-D.S. evaluated the specimen by negative-stain EM, screened the cryo-EM conditions, prepared the cryo-EM grids, and collected cryo-EM images with the help of S.C.; D.-D.S. and C.M. also performed density map calculations. Q.S., H.S., and W.Z. participated in model building and refined the final models. H.-W.J. and S.-C.T. provided nsp7 and nsp8 genes. X.W., F.Z., and M.G. participated in expression, purification, and functional assays of the RdRp. Y.-C.X., G.T., and J.S. made the remdesivir triphosphate form. Y.J. participated in experimental design and manuscript editing. H.J. conceived and coordinated the project. S.Z. conceived the project, initiated collaboration with H.E.X., and supervised X.L. Y.X. analyzed the structure and modeling and participated in figure preparation. Y.Z. supervised Q.S., C.M., and D.-D.S; analyzed the structures; and participated in manuscript writing. H.E.X. conceived and supervised the project, analyzed the structures, and wrote the manuscript, with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Density maps and structure coordinates have been deposited for immediate release. The Electron Microscopy Data Bank accession numbers and Protein Data Bank identifiers are EMD-30209 and PDB ID 7BV1 for the apo RdRp complex and EMD-30210 and PDB ID 7BV2 for the template RNA and remdesivir–bound RdRp complex. Materials are available upon 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/368/6498/1499/suppl/DC1
Figs. S1 to S8
Table S1
References (*37*–*46*)

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

8 April 2020; accepted 28 April 2020
Published online 1 May 2020
10.1126/science.abc1560

YOU'D SMILE TOO IF YOU JUST ADVANCED YOUR CAREER.



Find your next job at ScienceCareers.org

There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: sciencecareers.org

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

By **Montrai Spikes**

The pressure to assimilate

"Why are you dressed so nicely?" a fellow graduate student asked me in passing, after noticing my collared shirt, slacks, and dress shoes. "Oh, I have to teach today," I replied. He stopped and stared at me for a few moments, the confusion written plainly on his face. "As a Black man, students treat me with more respect when I dress up," I explained. What I did not say was, "Our society's current idea of professionalism is so intertwined with straight, white, masculinity that underrepresented people must go above and beyond or risk being seen as incompetent."

I love teaching. However, for a Black man in a STEM department, it can be taxing, frustrating, and daunting to figure out how to present myself to undergraduate students. Taxing because I have to learn to interact with students who may have preconceived notions of my abilities. Frustrating because the number of colleagues who share my experiences or can offer me guidance is essentially zero. Daunting because for some of these students, their experience with me will likely shape how they view all people of color.

During my first year as a teaching assistant (TA), I noticed some students didn't fully trust or respect me as an instructor. For example, one of my students emailed another TA to ask the very same question I'd answered for them earlier in the day. Another incident that got under my skin was when a student said I looked like rapper Wiz Khalifa, then proceeded to touch my dreadlocks. I felt embarrassed and violated, but I pretended their comment and action didn't bother me. I feared that correcting the student would only present me as an "angry Black man."



"The mask I crafted was disingenuous and was alienating students of color."

To garner more respect and confidence from my students, I started to diligently craft a mask that I thought would signal my professionalism. I wore nice clothes on teaching days and I reduced my use of Ebonics, an English dialect spoken by some Black Americans that I often used growing up. This mask seemed to work. I noticed my students were more engaged, asking questions and seeking me out for help. But the effort of maintaining my facade ultimately became unsustainable. Eventually, my mask cracked.

I was waiting in my office for a student, another Black man, to arrive to discuss his lab report. As I responded to emails, a rap album by Kendrick Lamar was playing. I was startled when the student exclaimed, "Yo! Trai, I didn't know you got down with Kendrick." This catalyzed an animated discussion about rappers, growing up in big cities, and people of color in STEM. My guard dropped, the Ebonics slipped out, and for a short period I was simply me. As I shifted our conversation to his assignment, he said offhandedly, "I wish you were more like this in class." He wasn't ill-intentioned, but his comment did prompt introspection.

I realized that although my "professional" mask appealed to my white students, it had unforeseen consequences. The mask I crafted was disingenuous and was alienating students of color while reinforcing the misconception that the only way to succeed is through cultural assimilation.

After this revelation, I discarded my mask. I continued to dress nicely because that's something I enjoy. But I started to speak Ebonics more. I also decided to take time, at the beginning of every semester, to tell my students about my educational and cultural background. This serves a dual purpose: It outlines my teaching qualifications and demonstrates, to underrepresented students especially, my pride in my culture and upbringing.

Admittedly, I'm not sure whether this is the best approach, and I do feel disheartened that I have to make a special effort to build credibility with my students. But I have noticed it has paid dividends. All my students have remained engaged, my underrepresented students seem empowered, and I no longer feel the need to put on an insincere mask.

I encourage my STEM colleagues to think seriously about who is more professional. Is it the man who wears button-down shirts or the person who is knowledgeable in their field? Is it the woman with straight hair or the person who leads productive discussions? A true professional doesn't have a prescribed appearance. I'm a young Black man with tattoos and dreadlocks who is fluent in Ebonics. I'm also a behavioral ecologist, activist, and educator. I am a professional. ■

Montrai Spikes is a Ph.D. student at the University of Oklahoma.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS