

**Experimental Huntington
drug brings rare hope** *p. 742*

Tomorrow's Earth
pp. 748, 751, & 766

**Gene linked to immune
dysfunction** *pp. 756 & 810*

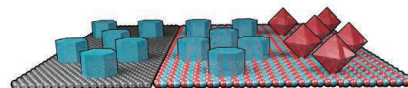
Science

\$15
24 AUGUST 2018
sciencemag.org

 AAAS

GRAPHENE QUASICRYSTALS

Twisted bilayer shows intriguing electronic properties *p. 782*



742

Michelle Dardengo is in a trial of a promising but unproven drug for Huntington disease.

NEWS

IN BRIEF

734 News at a glance

IN DEPTH

736 INTERPLANETARY SMALL SATELLITES COME OF AGE

With miniaturized guidance and propulsion systems, tiny craft aim for destinations beyond Earth *By E. Hand*

737 ANCIENT DNA REVEALS TRYST BETWEEN EXTINCT HUMAN SPECIES

Woman had a Neanderthal mother and a Denisovan father *By G. Vogel*

738 PROJECT LIFTS THE VEIL ON LIFE IN THE OCEAN'S TWILIGHT ZONE

New initiative aims to document midwater biodiversity *By E. Kintisch*

739 STICKY GLACIERS SLOWED TEMPO OF ICE AGES

Seafloor cores suggest thickening ice sheets triggered near-collapse of Atlantic currents *By P. Voosen*

740 MONTHLY SHOTS MAY REPLACE DAILY ANTI-HIV PILLS

Long-acting treatments now in clinical trials may work as preventives, too *By J. Cohen*

741 ARRIVAL OF DEADLY PIG DISEASE COULD SPELL DISASTER FOR CHINA

African swine fever threatens the world's largest pig herd *By D. Normile*

FEATURE

742 DARING TO HOPE

Patients thrill to reports of a promising antisense drug against Huntington disease, but no one is sure yet whether it works *By M. Wadman*

INSIGHTS

POLICY FORUM

748 THE PARADOX OF IRRIGATION EFFICIENCY

Higher efficiency rarely reduces water consumption *By R. Q. Grafton et al.*

► PODCAST

PERSPECTIVES

751 HOW AI CAN BE A FORCE FOR GOOD

An ethical framework will help to harness the potential of AI while keeping humans in control *By M. Taddeo and L. Floridi*

753 HOW TITANIUM DIOXIDE CLEANS ITSELF

Adsorption of hydrophobic molecules explains titanium dioxide self-cleaning properties *By J. Y. Park*

► REPORT P. 786

754 ORCHESTRATING CORTICAL BRAIN DEVELOPMENT

Signaling in the developing vasculature coordinates central nervous system morphology *By J.-L. Thomas*

► REPORT P. 767

755 ENTEROVIRUS OUTBREAK DYNAMICS

Predictability of outbreaks opens the door to model-guided public health planning

By B. Nikolay and S. Cauchemez

► REPORT P. 800

756 CONNECTING IMMUNE DEFICIENCY AND INFLAMMATION

RIPK1 mutations result in immune deficiency and autoinflammation in humans *By M. Pasparakis and M. Kelliher*

► REPORT P. 810

758 HOT LITHIUM-OXYGEN BATTERIES CHARGE AHEAD

Molten salt electrolytes and nickel oxide-based electrodes enable four-electron transfer *By S. Feng et al.*

► REPORT P. 777

BOOKS ET AL.

759 PUTTING SLEEP MYTHS TO BED

From groggy teenagers to fatal narcolepsy, two tomes tackle the science of slumber *By A. Woolfson*

760 CHANGING THE CONVERSATION

A pair of philosophers offer a new perspective on problematic dissent in science *By E. J. Nash*

LETTERS

761 SIDE EFFECTS OF ADDICTION TREATMENT

By G. Y. Tang and J. Parekh



761 EUROPE'S PLANS FOR RESPONSIBLE SCIENCE

By N. Mejlggaard et al.

762 ANTICIPATING DATA-INDUCED BIAS

By B. C. Schlick-Steiner et al.

762 TECHNICAL COMMENT ABSTRACTS

RESEARCH

IN BRIEF

763 From *Science* and other journals

REVIEW

766 HEALTH CARE

The promise and peril of universal health care *D. E. Bloom* et al.

REVIEW SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.aat9644](https://doi.org/10.1126/science.aat9644)

RESEARCH ARTICLES

767 NEUROSCIENCE

Endothelial Dab1 signaling orchestrates neuro-glia-vessel communication in the central nervous system *M. Segarra* et al.

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.aao2861](https://doi.org/10.1126/science.aao2861)

► PERSPECTIVE P. 754

768 CYTOSKELETON

Severing enzymes amplify microtubule arrays through lattice GTP-tubulin incorporation *A. Vemu* et al.

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.aau1504](https://doi.org/10.1126/science.aau1504)

769 PROTEOMICS

Co-regulatory networks of human serum proteins link genetics to disease *V. Emilsson* et al.

REPORTS

774 GAS GIANT PLANETS

Juno observations of spot structures and a split tail in Io-induced aurorae on Jupiter *A. Mura* et al.



777 BATTERIES

A high-energy-density lithium-oxygen battery based on a reversible four-electron conversion to lithium oxide *C. Xia* et al.

► PERSPECTIVE P. 758

782 QUASICRYSTALS

Dirac electrons in a dodecagonal graphene quasicrystal *S. J. Ahn* et al.

786 SURFACE CHEMISTRY

High-affinity adsorption leads to molecularly ordered interfaces on TiO₂ in air and solution *J. Balajka* et al.

► PERSPECTIVE P. 753

789 GRAPHENE

Interaction-driven quantum Hall wedding cake-like structures in graphene quantum dots *C. Gutiérrez* et al.

794 MAGNETISM

Observation of Dicke cooperativity in magnetic interactions *X. Li* et al.

797 ENVIRONMENTAL TOXINS

Diatom ooze—A large marine mercury sink *S. Zaferani* et al.

800 VIROLOGY

Serotype-specific immunity explains the incidence of diseases caused by human enteroviruses

M. Pons-Salort and *N. C. Grassly*

► PERSPECTIVE P. 755

804 OCEAN OXYGEN

Large-scale ocean deoxygenation during the Paleocene-Eocene Thermal Maximum *W. Yao* et al.

806 MOLECULAR BIOLOGY

An intrinsic S/G₂ checkpoint enforced by ATR *J. C. Saldivar* et al.

810 GENETICS

Biallelic *RIPK1* mutations in humans cause severe immunodeficiency, arthritis, and intestinal inflammation *D. Cuchet-Lourenço* et al.

► PERSPECTIVE P. 756

DEPARTMENTS

729 EDITORIAL

Retreat on economics at the EPA
By *Kevin Boyle* and *Matthew Kotchen*

818 WORKING LIFE

Inspiration from the outdoors
By *Jon Hellin*

ON THE COVER



Quasi-periodic tiling arises from bilayer graphene. Relativistic electrons called Dirac electrons are realized on a quasicrystal prepared by depositing one atomically thin carbon layer atop another at a 30° twist angle. Three basic building blocks—squares, rhombuses, and equilateral triangles—with different orientations can fill the entire space of the 12-fold rotationally symmetric quasicrystal without translational symmetry. See page 782. For more on the process behind the cover image, see <https://scim.ag/2MUHiqW>.
Illustration: *V. Altounian/Science*;
Code: *P. Stampfli*

Science Staff	726
New Products	814
Science Careers	815

Editor-in-Chief Jeremy Berg

Executive Editor Monica M. Bradford **News Editor** Tim Appenzeller

Deputy Editors Lisa D. Chong, Andrew M. Sugden(UK), Valda J. Vinson, Jake S. Yeston

Research and Insights

DEPUTY EDITOR, EMERITUS Barbara R. Jasny **SR. EDITORS** Gemma Alderton(UK), Caroline Ash(UK), Julia Fahrenkamp-Uppenbrink(UK), Pamela J. Hines, Stella M. Hurtle(UK), Paula A. Kiberstis, Marc S. Lavine(Canada), Steve Mao, Ian S. Osborne(UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Jelena Stajic, Peter Stern(UK), Phillip D. Szuroni, Sacha Vignieri, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Brent Grocholski, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon(UK), Keith T. Smith(UK) **ASSOCIATE BOOK REVIEW EDITOR** Valerie B. Thompson **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Amber Esplin, Chris Filiatreau, Cynthia Howe, Catherine Wolner **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Jerry Richardson, Alice Whaley(UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Ope Martins, Nida Masilius, Dona Mathieu, Ronmel Navas, Hilary Stewart(UK), Alana Warnke, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ADMINISTRATIVE SUPPORT** Janet Clements(UK), Jessica Waldo(UK), Ming Yang (UK)

News

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, David Grimm, Eric Hand, David Malakoff, Leslie Roberts **SR. CORRESPONDENTS** Daniel Clery(UK), Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Catherine Maticic **NEWS WRITERS** Adrian Cho, Jon Cohen, Jennifer Couzin-Frankel, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad(Cambridge, UK), Paul Voosen, Meredith Wadman **INTERNS** Katie Langin, Frankie Schembri, Matt Warren **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt(Berlin), Andrew Lawler, Matt Leslie, Eliot Marshall, Virginia Morell, Dennis Normile(Shanghai), Charles Pillier, Tania Rabesandratana(London), Emily Underwood, Gretchen Vogel(Berlin), Lizzie Wade(Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein(Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Executive Publisher Rush D. Holt

Publisher Bill Moran **Chief Digital Media Officer** Josh Freeman

DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT Sarah Whalen **DIRECTOR, PRODUCT AND CUSTOM PUBLISHING** Will Schweitzer **MANAGER, PRODUCT DEVELOPMENT** Hannah Heckner **BUSINESS SYSTEMS AND FINANCIAL ANALYSIS** Director Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYST** Eric Knott **ASSOCIATE DIRECTOR, PRODUCT MANAGEMENT** Kris Bishop **SENIOR SYSTEMS ANALYST** Nicole Mehmedovich **SENIOR BUSINESS ANALYST** Cory Lipman **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **BUSINESS ANALYSTS** Meron Kebede, Sandy Kim, Jourdan Stewart **FINANCIAL ANALYST** Julian Iriarte **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SALES COORDINATOR** Shirley Young **DIRECTOR, COPYRIGHT, LICENSING, SPECIAL PROJECTS** Emilie David **DIGITAL PRODUCT ASSOCIATE** Michael Hardesty **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **RIGHTS, CONTRACTS, AND LICENSING ASSOCIATE** Lili Catlett **RIGHTS & PERMISSIONS ASSISTANT** Alexander Lee **DIRECTOR, INSTITUTIONAL LICENSING** Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGERS** Marco Castellani, Chris Murawski **SENIOR OPERATIONS ANALYST** Lana Guz **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge

WEB TECHNOLOGIES TECHNICAL DIRECTOR David Levy **PORTFOLIO MANAGER** Trista Smith **PROJECT MANAGER** Dean Robbins **DEVELOPERS** Liana Birke, Elissa Heller, Ryan Jensen

DIGITAL MEDIA DIRECTOR OF ANALYTICS Enrique Gonzales **MULTIMEDIA MANAGER** Sarah Crespi **MANAGING WEB PRODUCER** Kara Estelle-Powers **DIGITAL PRODUCER** Jessica Hubbard **VIDEO PRODUCER** Chris Burns **SOCIAL MEDIA PRODUCER** Brice Russ

DIGITAL/PRINT STRATEGY MANAGER Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **DIGITAL PRODUCTION MANAGER** Lisa Stanford **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Antoinette Hodal, Lori Murphy, Anthony Rosen **CONTENT SPECIALISTS** Jacob Hedrick, Kimberley Oster

DESIGN DIRECTOR Beth Rakouskas **DESIGN MANAGING EDITOR** Marcy Atarod **SENIOR DESIGNER** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS MANAGING EDITOR** Alberto Cuadra **GRAPHICS EDITOR** Nirja Desai **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **INTERACTIVE GRAPHICS EDITOR** Jia You **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **PHOTO EDITOR** Emily Petersen **IMAGE RIGHTS AND FINANCIAL MANAGER** Jessica Adams

SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders: 202-326-6430 **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst: 202-326-6463 **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **SR. PRODUCTION SPECIALIST/GRAPHIC DESIGNER** Amy Hardcastle **SR. TRAFFIC ASSOCIATE** Christine Hall **DIRECTOR OF BUSINESS DEVELOPMENT AND ACADEMIC PUBLISHING RELATIONS**, ASIA Xiaoying Chu: +86-131 6136 3212, xchu@aaas.org **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu + 81532-81-5142 asandhu@aaas.org **EAST COAST/E. CANADA** Laurie Faraday: 508-747-9395, FAX 617-507-8189 **WEST COAST/W. CANADA** Lynne Stickrod: 415-931-9782, FAX 415-520-6940 **MIDWEST** Jeffrey Dembski: 847-498-4520 x3005, Steven Locher: 847-498-4520 x3006 **UK EUROPE/ASIA** Roger Gonçalves: TEL/FAX +41 43 243 1358 **JAPAN** Kaoru Sasaki (Tokyo): +81 (3) 6459 4174 ksasaki@aaas.org

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING ASSOCIATE** Aimee Aponte **MARKETING MANAGER, JOURNALS** Shawana Arnold **MARKETING ASSOCIATES** Mike Romano, Tori Velasquez **SENIOR DESIGNER** Kim Huynh **TRADE SHOW COORDINATOR** Andrew Clamp

GLOBAL SALES DIRECTOR ADVERTISING AND CUSTOM PUBLISHING Tracy Holmes: +44 (0) 1223 326525 **CLASSIFIED** advertise@sciencecareers.org **SALES MANAGER, US, CANADA AND LATIN AMERICA** SALES CAREERS Claudia Paulsen-Young: 202-326-6577 **EUROPE/ROW SALES** Sarah Lelarge **SALES ADMIN ASSISTANT** Kelly Grace +44 (0)1223 326528 **JAPAN** Miyuki Tani(Osaka): +81 (6) 6202 6272 mtani@aaas.org **CHINA/TAIWAN** Xiaoying Chu: +86-131 6136 3212, xchu@aaas.org

AAAS BOARD OF DIRECTORS, CHAIR Susan Hockfield **PRESIDENT** Margaret A. Hamburg **PRESIDENT-ELECT** Steven Chu **TREASURER** Carolyn N. Ainslie **CHIEF EXECUTIVE OFFICER** Rush D. Holt **BOARD** Cynthia M. Beall, May R. Berenbaum, Rosina M. Bierbaum, Kaye Husbands Fealing, Stephen P.A. Fodor, S. James Gates, Jr., Michael S. Gazzaniga, Laura H. Greene, Robert B. Millard, Mercedes Pascual, William D. Provine

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005 **INSTITUTIONAL SITE LICENSES** 202-326-6730 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Central Support** 866-434-2227 www.aaas.org/membercentral.

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.

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

BOARD OF REVIEWING EDITORS (Statistics board members indicated with \$)

Adriano Aguzzi, U. Hospital Zürich
Takuzo Aida, U. of Tokyo
Leslie Aiello, Wenner-Gren Foundation
Judith Allen, U. of Manchester
Sebastian Amigorena, Institut Curie
Meinrat O. Andrae, Max Planck Inst. Mainz
Paola Ariotti, 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
Carlo Beenakker, Leiden U.
Kamran Behnia, ESPCI
Yasmine Belkaid, NIAID, NIH
Philip Benfey, Duke U.
Gabriele Bergers, VIB
Bradley Bernstein, Massachusetts General Hospital
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 (\$) **\$**
Christian Büchel, UKE Hamburg
Dennis Burton, The Scripps Res. Inst.
Carter Tribley Butts, U. of California, Irvine
Gyorgy Buzsáki, New York U. School of Medicine
Blanche Capel, Duke U.
Mats Carlsson, U. of Oslo
Ib Chorkendorff, Denmark TU
James J. Collins, MIT
Robert Cook-Deegan, Arizona State U.
Lisa Coussens, Oregon Health & Science U.
Alan Cowman, Walter & Eliza Hall Inst.
Roberta Croce, VU Amsterdam
Jeff L. Dangl, U. of North Carolina
Tom Daniel, U. of Washington
Chiara Daraio, Caltech
Nicolas Daughas, U. of Chicago
Frans de Waal, Emory U.
Stanislas Dehaene, Collège de France
Robert Desimone, MIT
Claude Desplan, New York U.
Sandra Diaz, Universidad Nacional de Córdoba
Dennis Discher, U. of Penn.
Gerald W. Dorn II, 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 at Chapel Hill
Barry Everitt, U. of Cambridge
Vanessa Ezenwa, U. of Georgia
Ernst Fehr, U. of Zürich
Michael Feuer, The George Washington U.
Toren Finkel, U. of Pittsburgh Medical Ctr.
Kate Fitzgerald, U. of Massachusetts
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, Rice U.
Elizabeth Grove, U. of Chicago
Nicolas Gruber, ETH Zürich
Kip Guy, U. of Kentucky College of Pharmacy
Taekjip Ha, Johns Hopkins U.
Christian Haass, Ludwig Maximilians U.
Sharon Hammes-Schiffer, U. of Illinois at Urbana-Champaign
Wolf-Dietrich Hardt, ETH Zürich
Louise Harra, U. College London
Michael Hasselmo, Boston U.
Jian He, Clemson U.
Martin Heimann, Max Planck Inst. Jena
Carl-Philipp Heisenberg, IST Austria
Ykä Helariutta, U. of Cambridge
Janet G. Hering, Eawag
Kai-Uwe Hinrichs, U. of Bremen
David Hodell, U. of Cambridge
Lora Hooper, UT Southwestern Medical Ctr. at Dallas
Fred Hughson, Princeton U.
Randall Hulet, Rice U.
Auke Ijspeert, EPFL
Akiko Iwasaki, Yale U.
Stephen Jackson, USGS and U. of Arizona
Seema Jayachandran, Northwestern U.
Kai Johnsson, EPFL
Peter Jonas, Inst. of Science & Technology Austria
Matt Kaeblerlein, U. of Washington
William Kaelin Jr., Dana-Farber Cancer Inst.
Daniel Kammen, U. of California, Berkeley
Abby Kavner, U. of California, Los Angeles
Masashi Kawasaki, U. of Tokyo
V. Narry Kim, Seoul Nat. U.
Robert Kingston, Harvard Medical School
Etienne Kochlin, École Normale Supérieure
Alexander Kolodkin, Johns Hopkins U.
Thomas Langer, U. of Cologne
Mitchell A. Lazar, U. of Penn.

David Lazer, Harvard U.
Stanley Lemon, U. of North Carolina at Chapel Hill
Ottoline Leyser, U. of Cambridge
Wendell Lim, U. of California, San Francisco
Marcia C. Linn, U. of California, Berkeley
Jianguo Liu, Michigan State U.
Luis Liz-Marzán, CIC biomaGUNE
Jonathan Losos, Harvard U.
Ke Lu, Chinese Acad. of Sciences
Christian Lüscher, U. of Geneva
Fabienne Mackay, U. of Melbourne
Anne Magurran, U. of St. Andrews
Oscar Marín, King's College London
Charles Marshall, U. of California, Berkeley
Christopher Marx, U. of Idaho
C. Robertson McClung, Dartmouth College
Rodrigo Medellín, U. of Mexico
Graham Medley, London School of Hygiene & Tropical Med.
Jane Memmott, U. of Bristol
Tom Misteli, NCI, NIH
Yasushi Miyashita, U. of Tokyo
Richard Morris, U. of Edinburgh
Alison Motsinger-Reif, NC State U. (\$) **\$**
Daniel Neumark, U. of California, Berkeley
Kitty Nijmeijer, TU Eindhoven
Helga Nowotny, Austrian Council
Graham O'Reilly, U. of Warwick
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
Steve Palumbi, Stanford U.
Jane Parker, Max Planck Inst. Cologne
Giovanni Parmigiani, Dana-Farber Cancer Inst. (\$) **\$**
Samuel Pfaff, Salk Inst. for Biological Studies
Matthieu Piel, Institut Curie
Kathrin Plath, U. of California, Los Angeles
Martin Plenio, Ulm U.
Albert Polman, FOM Institute for AMOLF
Elvira Poloczanska, Alfred-Wegener-Inst.
Philippe Poulin, CNRS
Jonathan Pritchard, Stanford U.
David Randall, Colorado State U.
Sarah Reisman, Caltech
Félix A. Rey, Institut Pasteur
Trevor Robbins, U. of Cambridge
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, Medical U. of Vienna
Alexander Schier, Harvard U.
Wolfram Schlenker, Columbia U.
Susannah Scott, U. of California, Santa Barbara
Vladimir Shalae, Purdue U.
Beth Shapiro, U. of California, Santa Cruz
Jay Shendure, U. of Washington
Brian Shoichet, U. of California, San Francisco
Robert Siliciano, Johns Hopkins U. School of Medicine
Uri Simonsohn, U. of Penn.
Lucia Sivilotti, U. College London
Alison Smith, John Innes Centre
Richard Smith, U. of North Carolina at Chapel Hill (\$) **\$**
Mark Smyth, QIMR Berghofer
Pam Soltis, U. of Florida
John Speakman, U. of Aberdeen
Tara Spres-Jones, U. of Edinburgh
Allan C. Spradling, Carnegie Institution for Science
Eric Steig, U. of Washington
Paula Stephan, Georgia State U.
V. S. Subrahmanian, U. of Maryland
Ira Tabas, Columbia U.
Sarah Teichmann, U. of Cambridge
Shubha Tole, Tata Inst. of Fundamental Research
Wim van der Putten, Netherlands Inst. of Ecology
Bert Vogelstein, Johns Hopkins U.
David Wallach, Weizmann Inst. of Science
Jane-Ling Wang, U. of California, Davis (\$) **\$**
David Waxman, Fudan U.
Jonathan Weissman, U. of California, San Francisco
Chris Wickle, U. of Missouri (\$) **\$**
Terrie Williams, U. of California, Santa Cruz
Ian A. Wilson, The Scripps Res. Inst. (\$) **\$**
Timothy D. Wilson, U. of Virginia
Yu Xie, Princeton U.
Jan Zaanen, Leiden U.
Kenneth Zaret, U. of Penn. School of Medicine
Jonathan Zehr, U. of California, Santa Cruz
Maria Zuber, MIT

Retreat on economics at the EPA

Rigorous economic analysis has long been recognized as essential for sound, defensible decision-making by government agencies whose regulations affect human health and the environment. The acting administrator (since July 2018) of the U.S. Environmental Protection Agency (EPA) has emphasized the importance of transparency and public trust. These laudable goals are enhanced by external scientific review of the EPA's analytical procedures. Yet, in June 2018, the EPA's Science Advisory Board (SAB) eliminated its Environmental Economics Advisory Committee (EEAC). The agency should be calling for more—not less—external advice on economics, given the Trump administration's promotion of economic analyses that push the boundaries of well-established best practices. The pattern is clear: When environmental regulations are expected to provide substantial public benefits, assumptions are made to substantially diminish their valuations.

The EEAC, on which we served (K.B. since 2013, M.K. since 2015), consisted of nationally recognized economists appointed to provide independent advice to the EPA. The committee had been called upon by administrations of both political parties since the early 1990s. But today, many economic analyses that support the Trump administration's regulatory rollbacks conflict with the EPA's previous findings. The 2017 analysis for eliminating the Waters of the United States rule turned favorable only after excluding all benefits of protecting wetlands. Eliminating the Clean Power Plan is supported in another 2017 analysis only after changing assumptions about the scope of climate damages, the measurement of health effects, and the impact on future generations. Differing assumptions also underlie the economic justification of the administration's 2018 proposal to roll back automotive fuel economy standards.

At an institutional level, the EPA also issued a proposal in June to revamp its approach to benefit-cost

analysis. Many observers are concerned that this is an administrative move to institutionalize the agency's practices in the economic analyses noted above. This could result in the elimination of counting significant co-benefits. For example, a regulation that targets carbon dioxide emissions from power plants can simultaneously reduce other harmful pollutants, and the resulting co-benefit would not be counted.

EPA priorities always change between administrations,

and all economic analyses require assumptions. That is why external scientific review serves as a bulwark for separating political ideology from evidence-based decision-making. EEAC's elimination means that an important channel in this process has been lost. The leading explanation for its demise is that current expertise on the SAB is sufficient, and that ad hoc committees can address gaps. However, a lapse in appointing the EEAC chair, who would have been a SAB voting member, precluded an important voice in the discussion prior to the decision to shutter EEAC. Furthermore, the SAB's current membership reveals little environmental economics expertise, which is necessary for evaluating

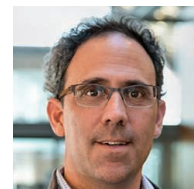
when and how economic reviews might be important. Other explanations include EEAC inactivity and costs. Inactivity was not a choice by the EEAC, which by design responded only to requests. Costs could be reduced by decreasing the number of appointed individuals, and it is unclear whether ad hoc committees offer cost savings. Because the EPA is often under pressure for timely analysis, a standing committee also has the advantage of being quicker to mobilize.

The EPA should reconsider its decision to eliminate the EEAC, and the SAB should consider how to better constitute an economic advisory committee to promote efficient and equitable outcomes from EPA policies.

—Kevin Boyle and Matthew Kotchen



Kevin Boyle
is a professor in the Department of Agricultural and Applied Economics at Virginia Tech, Blacksburg, VA, USA. kjboyle@vt.edu



Matthew Kotchen
is a professor of economics at Yale University, New Haven, CT, USA. kotchen@yale.edu



"The agency should be calling for more—not less—external advice..."

“We get this additional layer of hate mail, and people, I think, find it easier to put us down because we are women.”

Andrea Dutton of the University of Florida in Gainesville, a geologist, to E&E News, about the insults and harassment she and other female climate scientists receive.

IN BRIEF

Edited by
Jeffrey Brainard



India will use its Launch Vehicle Mark III rocket to send three “vyomanauts” into space.

SPACE EXPLORATION

India plans to launch crew by 2022

India last week announced plans to launch three astronauts into space using its own rocket by 2022—making it the fourth country to embark on human spaceflight, after Russia, the United States, and China. The Indian Space Research Organisation (ISRO) in Bengaluru has developed and tested a 640-ton rocket capable of sending a three-person crew into low-Earth orbit; the inaugural manned flight is estimated to cost about \$1.2 billion, including development and testing. ISRO has also test launched a crew module that returned to Earth. The three “vyomanauts”—after *vyom*, the Sanskrit word for space—will conduct scientific experiments, says ISRO Chairman Kailasavadivoo Sivan. India is not a partner in the International Space Station, and its only astronaut to reach space flew aboard a Soviet spacecraft in 1984. Sivan says the flight will not divert funding from India’s robotic space missions. However, India this month announced it will delay launch of Chandrayaan-2—its ambitious moon orbiter, lander, and rover—by 3 months to January 2019 to modify the lander’s design to minimize the dust cloud it may raise during its descent to the lunar south pole.

Measles epidemic hits Europe

PUBLIC HEALTH | Europe is battling a measles epidemic unprecedented this decade. More than 41,000 children and adults contracted the disease during the first 6 months of 2018 in the World Health Organization’s European Region, which includes 53 countries; at least 37 have died. Ukraine accounted for 23,000 cases, and France, Georgia, Greece, Italy, Russia, and Serbia had more than 1000 each. Measles is highly contagious and easily preventable, but complacency and unfounded fears about the safety of the measles vaccine have put a dent in immunization coverage. “Through an aggressive antivaccine lobby, we’ve allowed measles to sweep across Europe,” says Peter Hotez, dean of the National School of Tropical Medicine at Baylor College of Medicine in Houston, Texas. “This is a self-inflicted wound.” The Americas have seen an upsurge in the disease as well; last month, the Pan American Health Organization reported almost 2500 cases in 11 countries so far in 2018, two-thirds of them in Venezuela, whose health system is collapsing.

Gene therapy review pared back

BIOMEDICINE | The National Institutes of Health (NIH) in Bethesda, Maryland, proposed last week that an expert panel cease reviewing individual human gene therapy studies, leaving oversight in the hands of the Food and Drug Administration (FDA) in Silver Spring, Maryland. NIH created the Recombinant DNA Advisory Committee (RAC) in 1974 to gather advice about research that manipulates DNA and later expanded its purview to clinical trials using gene transfer to treat diseases. The gene therapy field had a rough start—an early trial resulted in a teenager’s death in 1999—but last year FDA approved the first treatments, two for cancer and one for an inherited blindness disease. Now, reviews of gene therapy protocols by the RAC and FDA are “duplicative,” FDA Commissioner Scott Gottlieb and NIH Director Francis Collins wrote in the 15 August issue of *The New England Journal of Medicine*. Since 2016, the RAC had already limited its reviews to gene therapy strategies that pose unusual risks, or just three of 275 submitted protocols. In a

plan published 17 August for public comment, NIH proposed that the RAC now focus only on new technologies such as gene editing and synthetic biology.

TB drug overhaul proposed

PUBLIC HEALTH | The World Health Organization (WHO) in Geneva, Switzerland, announced last week what it called “landmark changes” in the treatment of multidrug-resistant tuberculosis (MDR TB), recommending that current treatments be replaced with safer, faster-acting ones. Currently, WHO guidelines call for treating MDR TB with painful, injected drugs, which can have serious side effects, for up to 20 months. WHO says it now will encourage using less harmful drugs in regimens lasting as little as 9 months. Its new guidelines, to be released by the end of the year, also will clarify how to determine the best drugs to use, how to dose children, and how to treat the even more serious condition of extensively drug-resistant TB. WHO estimates there were 600,000 new cases of MDR TB in 2016.

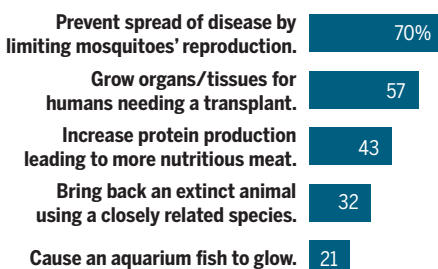
Accused bully loses grants

WORKPLACE | The Wellcome Trust, a London-based research charity, has pulled £3.5 million in grants from a prominent cancer researcher in the first use of its recently announced antibullying policy. In July, *The Guardian* revealed that Nazneen Rahman would leave the University of London’s Institute of Cancer Research (ICR) following an investigation into allegations that she had bullied and harassed staff at her laboratory. In a 17 August statement, the Wellcome Trust says it doesn’t know all the details in

SURVEY

When is it OK to genetically engineer an animal?

The Pew Research Center asked U.S. adults whether they would approve of genetically engineering animals for various purposes. Majorities backed such work to combat human illnesses. In every case, the level of approval was higher among people very knowledgeable about science than among those with low knowledge.



An antenna aboard the International Space Station can track this miniature transmitter, strapped to a scarlet macaw.

ANIMAL BEHAVIOR

Bird watching from space

A 16-year effort to detect from space the movements of small animals tagged with radio transmitters got a boost last week when Russian cosmonauts unfurled an antenna on the International Space Station. The 3-meter-by-2-meter device should become fully operational in 2019, giving animal researchers more comprehensive movement data over larger areas than they can monitor with handheld receivers in the field. They say the data will aid conservation and understanding animal behavior. Members of the International Cooperation for Animal Research Using Space (ICARUS) group have designed 5-gram tags that can send the space-based antenna many kinds of data, including an animal’s location and acceleration and the temperature and magnetic field in its surrounding environment. The antenna can scan 120 such tags over a 30- to 800-kilometer area every 3 seconds. Among many planned projects, the ICARUS team will test whether the behavior of goats on Italy’s Mount Etna predicts its eruptions and track African fruit bats to better understand Ebola’s spread.

the case, but ICR’s inquiry “deemed some of the allegations serious enough to warrant consideration at a disciplinary hearing” (ICR confirms an investigation occurred and says the disciplinary hearing didn’t take place because Rahman resigned.) Rahman declined to be interviewed.

Emissions reductions rolled back

CLIMATE SCIENCE | President Donald Trump’s administration on 21 August announced that it will replace former President Barack Obama’s Clean Power Plan with a much less aggressive effort to cut greenhouse gas emissions from U.S. power plants. The plan gives states greater power to set emissions limits and relaxes requirements on utilities. Analysts say the Trump plan, which the Environmental Protection Agency (EPA) is likely to finalize next year, would allow about 10 times more carbon emissions from the power sector than the 2015 Obama plan. Some states and environmental groups are certain to

challenge the new plan in court, arguing EPA is ducking its responsibility to fight climate change.

Brain megaproject head quits

NEUROSCIENCE | Chris Ebell, executive director of Europe’s controversial Human Brain Project (HBP), has decided to step down. His departure comes in the wake of a change of leadership at the Swiss Federal Institute of Technology in Lausanne, which coordinates the project, and amid “differences of opinion on governance and on strategic orientations for the HBP,” according to a 16 August statement. The €1 billion effort, which aims to simulate the entire human brain in a computer, has long been plagued by turmoil; it underwent a major reshuffle in 2015 after European neuroscientists attacked its scientific goals and governance model.

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

PLANETARY SCIENCE

Interplanetary small satellites come of age

With miniaturized guidance and propulsion systems, tiny craft aim for destinations beyond Earth

By Eric Hand

Cher, small satellites have swarmed into Earth orbit over the past decade, cutting the cost of studying our home planet from space. Now, these spacecraft, some no bigger than a briefcase, are becoming capable enough to venture into deep space—or at least the inner solar system. Two are halfway to Mars, more than a dozen planetary probes are in development, and scientists are coming up with ever more daring ideas for doing cheap, high-risk interplanetary science.

“Planetary is definitely getting excited,” Lori Glaze, head of NASA’s planetary science division, said last week at a symposium on small deep-space probes at Goddard Space Flight Center in Greenbelt, Maryland. Earlier this year, NASA began to accept proposals for a line of small planetary missions, with costs capped at \$55 million. Glaze says

12 teams have submitted proposals, and the agency plans to select several finalists in February 2019. Europe, too, has plans for small planetary probes, also known as CubeSats for the cube-shaped modules from which they are built. “We see now the potential for interplanetary CubeSats,” says Roger

Walker, the European Space Agency’s technology CubeSat manager in Noordwijk, the Netherlands.

Small satellites can be assembled from low-cost components and released by the dozen from a single rocket. But systems key to interplanetary flight, including propulsion, communication,

and navigation, have traditionally been too bulky to fit into a small package.

A mission called Mars Cube One (MarCO), twin craft launched in May along with the Mars InSight lander, is breaking that size barrier. Built from six standard, 10-centimeter cubes, they are meant to provide a communication relay for InSight as

The Mars-bound CubeSats “have flown farther than any ever before.”

Lori Glaze, NASA

The Mars Cube One mission—the first interplanetary CubeSats—will coast past the Red Planet this fall.

it descends to the surface. But Glaze says the craft, which passed the halfway point in their journey last week, are already pioneers. “These CubeSats have flown farther than any ever before,” she says. “They’ve already demonstrated the ability to do a comm relay.” An unfurled radio antenna panel, three times the size of the CubeSats themselves, transmits a trickle of data directly to Earth using the CubeSats’ limited solar power.

MarCO also showcases a miniature guidance, navigation, and control system developed by Blue Canyon Technologies in Boulder, Colorado. The technology has helped make CubeSats attractive for space science, says Dan Hegel, Blue Canyon’s director for advanced development. “CubeSats were tumbling around, not doing much,” he says. “There was no motivation before to try and shrink your instrument.” The company shrank reaction wheels, gyroscopes, and star trackers into a system that sells for less than \$150,000 and fits in half a cube.

Propulsion is a lingering concern. The small craft may need to change course, or slow down to orbit a planet, moon, or asteroid. Although MarCO’s propulsion system occupies half of the craft, it holds only enough fuel to make small trajectory adjustments en route to Mars, and it squirts pressurized gas like a fire extinguisher, an inefficient approach. As a result, the CubeSats will helplessly coast past the Red Planet after completing their mission.

CubeSats in Earth orbit have tested solar sails, thin mirrored foils that deliver a gentle push from the pressure of sunlight. Other developers are betting on solar electric propulsion systems. A device built by ExoTerra Resource in Littleton, Colorado, uses electricity from solar panels to bombard a xenon gas “fuel” with a beam of electrons, creating a charged plasma. An electric field shoots the plasma out the back, generating a feeble thrust. No bigger than a hockey puck, the device, called a Hall thruster, uses fuel much more efficiently than conventional rockets do, ExoTerra President Michael VanWoerkom says. “If you’re willing to wait longer to get there, you can package a lot of propellant into a very small space,” he says.

A big test of propulsion technologies will come at the end of 2019, when NASA’s heavy lift rocket, the Space Launch System, is due for its maiden voyage. It will carry 13 CubeSats, many of them focused on moon science. “Almost all are using different propulsion technologies,” says Goddard’s Barbara Cohen, principal investigator for

one of the missions, Lunar Flashlight, an effort to confirm the presence of ice in permanently shadowed regions of polar craters by shining lasers into them.

Better propulsion could help solve another problem facing planetary small satellites: a lack of rocket rides. CubeSats often piggyback on larger mission launches, but rideshares beyond low-Earth orbit are rare. Solar electric propulsion systems could help craft released into low-Earth orbits make an escape. A small satellite equipped with a Hall thruster could spiral out from Earth to the moon in a few months, VanWoerkom says. Reaching Mars would take a few years.

Scientists are starting to have big dreams for their small packages. Tilak Hewagama, a planetary scientist at the University of Maryland in College Park, wants to send a small satellite to intercept a comet on its first arrival in the solar system. Most comets have swung around the sun many times, and their once-pristine surfaces have grown weathered. But nearly every year, astronomers discover a few that are swooping in for the first time. By then, it is too late to develop a spacecraft to study them, Hewagama says. But a small satellite already parked in a stable orbit could maneuver in time to witness the comet's passage up close—a risky plan that Hewagama says NASA wouldn't be willing to pursue for a larger, more expensive craft.

Timothy Stubbs, a planetary scientist at Goddard, wants to use two 30-kilogram satellites to explore the origin of curious bright swirls on the surface of the moon. One idea is that weak magnetic fields in moon rocks—implanted by comet impacts or a long-extinct magnetic dynamo—might be repelling the solar wind particles that weather and darken the surrounding soil. But understanding the interactions between the particles and the fields requires skimming the moon in a close, unstable orbit that would require large amounts of fuel to maintain. Stubbs's solution: Orbit two small satellites in tandem, linked by a thin Kevlar tether 25 kilometers long, so that a satellite in a higher orbit can stabilize its mate a mere 2 kilometers above the surface.

Both teams plan to submit proposals to the new NASA funding program—if they can whittle costs down to fit the \$55 million cap. Small satellites may be cheap, but developing a deep-space mission traditionally requires a big team and lots of testing to pare down risk. Symposium organizer Geronimo Villanueva, a Goddard planetary scientist, says NASA officials are working on changing the rules for small satellites headed for deep space so that higher risk levels are acceptable. “We need to change the way we do business,” he says. ■

EVOLUTION

Ancient DNA reveals tryst between extinct human species

Woman had a Neanderthal mother and a Denisovan father

By Gretchen Vogel

The woman may have been just a teenager when she died more than 50,000 years ago, too young to have left much of a mark on her world. But a piece of one of her bones, unearthed in a cave in Russia's Denisova valley in 2012, may make her famous. Enough ancient DNA lingered within the 2-centimeter fragment to reveal her startling ancestry: She was the direct offspring of two different species of ancient humans—neither of them ours. An analysis of the woman's genome, reported in this week's issue of *Nature*, indicates her mother was Neanderthal and her father was Denisovan, the mysterious group of ancient humans discovered in the same Siberian cave in 2011. It is the most direct evidence yet that various ancient humans mated with each other and had offspring.

Based on other ancient genomes, researchers already had concluded that Denisovans, Neanderthals, and modern humans interbred in ice age Europe and Asia. The genes of both archaic human species are present in many people today. Other fossils found in the Siberian cave have shown that all three species lived there at different times. But the new finding “is sensational” just the same, says Johannes Krause, who studies ancient DNA at the Max Planck

Institute for the Science of Human History in Jena, Germany. “Now we have the love child of two different hominin groups, found where members of both groups have been found. It's quite a lot of things happening in one cave through time.”

Viviane Slon, a paleogeneticist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who did the ancient DNA analysis, says when she saw the results, her first reaction was disbelief. Only after repeating the experiment several times were she and her Leipzig colleagues—Svante Pääbo, Fabrizio Mafessoni, and Benjamin Vernot—convinced. That a direct

offspring of the two ancient humans was found among the first few fossil genomes recovered from the cave suggests, Pääbo says, “that when these groups met, they actually mixed quite freely with each other.”

The bone fragment's characteristics suggested it came from someone who was at least 13 years old. After pulverizing small samples, extracting DNA, and sequencing it, Slon and her colleagues found that its owner was female, and that her genome matched that of Denisovans and Neanderthals in roughly equal measure. Moreover, the proportion of genes in which her chromosome pairs harbored different variants—so-called heterozygous alleles—was close to 50% across all chromosomes, suggesting the maternal and paternal chromosomes came directly from different groups. And her mitochondrial DNA, which is inherited maternally, was uniformly Neanderthal, so the researchers concluded she was a first-generation hybrid of a Denisovan man and Neanderthal woman. The evidence “is so direct, we almost caught them in the act,” Pääbo says.

A closer look at the genome suggests her father also had some Neanderthal ancestry, possibly several hundred generations back. And the woman's Neanderthal genes are closer to those of a Neanderthal found in Croatia than those from remains found in the Siberian cave. That suggests distinct groups of Neanderthals migrated back and forth between western Europe and Siberia multiple times.

Along the way, apparently, they freely spread their genes to outsiders. That highlights the question, Krause says, of why Denisovans and Neanderthals nevertheless remained genetically distinct groups. “Why don't they come together as one population if they come together from time to time?” Geographic barriers probably played a role, he says, but researchers need more fossils with ancient DNA, from multiple sites, to understand the true legacy of these prehistoric couplings. ■



This bone fragment harbors the most direct evidence yet of ancient interspecies mating.



The dynamic midwater ocean teems with organisms such as this bristle worm.

MARINE SCIENCE

Project lifts the veil on life in the ocean's twilight zone

New initiative aims to document midwater biodiversity

By Eli Kintisch

The sea's murky depths might host more life than we thought. That's the preliminary conclusion of scientists who this week completed the inaugural cruise of the Ocean Twilight Zone (OTZ) initiative, a 6-year, \$35 million effort that is using innovative technologies—and an unusual funding model—to document the ocean's mysterious midwater layer.

The weeklong North Atlantic Ocean expedition was aimed primarily at testing the OTZ initiative's new workhorse: a 5-meter-long towed sled, dubbed Deep-See, that bristles with cameras, acoustic sensors, and samplers. But the trial also produced some eye-opening observations. At times when traditional surface instruments traced just a single, relatively dense layer of midwater organisms beneath the ship, for instance, Deep-See revealed a host of creatures distributed throughout the twilight zone, which extends from 200 meters to 1000 meters below the surface. "We kept seeing organisms all the way down," says Andone Lavery, a physicist with the Woods Hole Oceanographic Institution (WHOI) in Massachusetts, which is leading the project. "That was really surprising."

It was a promising start for OTZ scientists. They are targeting a little known ocean layer, between easily studied surface waters

and the dark abyss, which submersibles have explored. "The midwater zone has been severely neglected," says Heidi Sosik, a biological oceanographer with WHOI.

Researchers know it teems with life, including fish, crustaceans, jellies, worms, and squids. And they have speculated, based on acoustic and net surveys, that the total biomass of midwater fish might dwarf the current global catch of surface-dwelling fish by 100 times. But they have struggled to document this twilight ecosystem. "The big questions are who are the players [and] who's eating who," says biological oceanographer Mark Benfield of Louisiana State University in Baton Rouge. There are "species just waiting to be discovered," says WHOI physical oceanographer Gordon Zhang.

The project also aims to get a better grip on how twilight zone creatures influence the global carbon cycle. Midwater organisms perform perhaps the biggest daily migration on the planet, rising each night toward the surface to feast on a sunlight-fed bounty of plankton and fish. Then, as the sun rises, they sink back to the depths. That plunge prevents carbon captured at the surface from going "right back into the atmosphere," where it would amplify global warming, says marine ecologist Tracey Sutton of Nova Southeastern University in Dania Beach, Florida.

Traditional tools have proved inadequate for exploring midwater ecosystems. Ship-

mounted acoustical sensors—which use sound waves to locate objects—have trouble precisely detecting deep-swimming organisms. Towed nets can crush beyond recognition the gelatinous creatures found in the midwater. Because many are bioluminescent, the catch can resemble a ball of fire within the mesh, scaring away other animals. The pressure wave formed by an oncoming net can also warn off creatures.

Deep-See is engineered to overcome those challenges. "It's the equivalent of having the ship down at 600 meters," Benfield says. That allows the sled's acoustic sensors to deliver higher resolution data, because the sound waves don't have to pass through hundreds of meters of water. And because the sensors track seven different frequency bands, they might allow researchers to discern an animal's size and possibly even species. "It's like color TV versus black and white," Lavery says.

Deep-See's cameras, meanwhile, can image creatures as small as 50 microns in length, seven times per second. Other devices measure light and environmental variables. All told, 2 terabytes of data flow by cable from the rig to the ship each hour.

During Deep-See's recent trial, researchers compared the animals captured by the acoustic sensors and cameras to those caught in nets. The acoustic sensors imaged "big fish just a few meters away," says marine biologist Michael Jech of the National Oceanic and Atmospheric Administration in Woods Hole. But the large lights used by the cameras appeared to scare away creatures. The scientists are now brainstorming solutions, Jech says, which might include slowing the tow and trying different light colors.

WHOI scientists have already scheduled two more OTZ cruises and expect many more, across the globe, by the time the project ends in 2024. As a finale, they hope to establish permanent, tethered monitoring observatories in the midwater. "Who knows what we'll have in 20 years," Benfield says. "We may look at Deep-See as a primitive forerunner."

Funding for the project is coming from an unusual source: The Audacious Project, a new initiative by TED, the ideas-spreading nonprofit based in New York City. Audacious raises money from multiple private donors and vets proposals on their behalf, which cuts paperwork for grantees. Over the past 3 years, it has made seven awards, including the OTZ initiative, a methane-sensing satellite, and health care and hunger programs.

OTZ scientists are now combing through the more than 30 gigabytes of data they amassed. One goal: to see whether the data confirm that first impression of abundant midwater life. They are also looking ahead. "The big thing I want," Lavery says, "is to get Deep-See out in the ocean again soon." ■

ANCIENT CLIMATE

Sticky glaciers slowed tempo of ice ages

Seafloor cores suggest thickening ice sheets triggered near-collapse of Atlantic currents

By Paul Voosen, in Boston

About 1 million years ago, one of Earth's most important metronomes mysteriously shifted: Ice ages went from occurring every 40,000 years to every 100,000 years. At the same time, the "conveyor belt" of warming currents in the North Atlantic Ocean slowed sharply. Last week, scientists here at the Goldschmidt Conference presented a clue to these twin mysteries: evidence that glaciers in the Northern Hemisphere suddenly began to stick to their beds. Growing thicker, they might have triggered a cooling that disrupted the conveyor belt and allowed the 100,000-year cycle that we see today to take root.

"The system basically crashed," says Steve Goldstein, an ocean geochemist at Columbia University who led the study. Other scientists welcome the new clues to the transition. "This is really exciting new evidence," says Henrieka Detlef, a paleoclimatologist at Cardiff University in the United Kingdom. But she and others aren't sold yet on the long causal chain that Goldstein's team posits.

Scientists have long known that tiny changes in Earth's orbit around the sun, called Milankovitch cycles, drive the planet in and out of ice ages. But nothing changed in those orbital patterns 1 million years ago. Recently, Goldstein and his colleagues found signs of a possible contributor to the ice age transition: a near-collapse of the Atlantic meridional overturning circulation (AMOC). The AMOC shepherds shallow warm water to the North Atlantic, where it cools and sinks before returning south along the sea floor to the Southern Ocean to meet Pacific Ocean waters.

Goldstein's group deduces the overall strength of the AMOC from geochemical markers in ocean sediment cores. The researchers take advantage of a ratio between two isotopes of neodymium that varies with the age of their source rocks: ancient crust runs negative, whereas younger rocks are more positive. As it happens, the North

Atlantic is surrounded by ancient crust, whereas the Pacific, thanks to its volcanic Ring of Fire, tilts younger. The neodymium-carrying grit ends up incorporated into the shells of single-celled foraminifera or fish teeth, both of which accumulate over time on the sea floor. Changes in the isotope ratio record the wax and wane of intruding North Atlantic or Pacific waters.

Earlier this decade, the Columbia group tested its approach on two archived sediment cores from the South Atlantic. About 950,000 years ago, they saw the isotopic signals shoot up, reflecting an incursion of Pacific waters, with little evidence of returning North Atlantic waters—suggesting a stark "AMOC crisis." The slowdown could have sharply cooled the North Atlantic region—and might have lengthened the ice age rhythm.

had allowed a thick soil layer to build up on northern landmasses. At first, the soil acted as a grease that caused early ice sheets to collapse before they could thicken much. But repeated glaciations gradually scoured this grit away, and meltwater swept it into the ocean. As the glaciers dug deeper into older rock, the neodymium signal in ocean sediment became more negative. Eventually, the glaciers reached bedrock and began to stick to their base, allowing them to grow thicker—leading to a more profound and persistent cooling that somehow caused the AMOC to crash and the glacial cycle to lengthen. "We think we're seeing the trigger," says Maayan Yehudai, the Columbia graduate student who presented the work. (Scientists believe pronounced global warming—like the warming underway now—could also disrupt the AMOC.)

The neodymium evidence supports this geological story, Clark says. "It's a pretty clear signal that you should see," Detlef notes, however, that there is no conclusive evidence that northern ice sheets were increasing in thickness prior to the AMOC slowdown. But she accepts that something important happened in the North Atlantic leading up to the AMOC crisis.

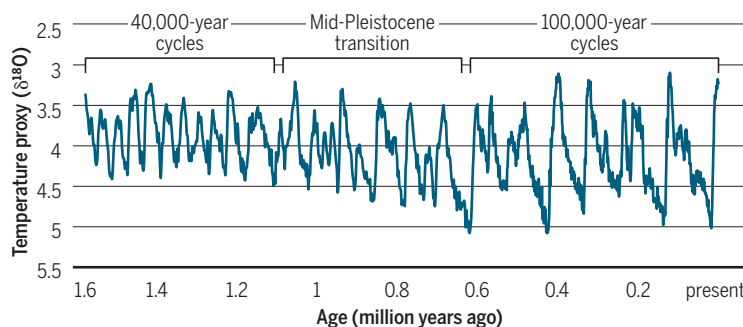
One hypothesis that does seem ruled out, however, is the notion that the growth of Antarctic ice sheets 900,000

years ago played a pivotal role in the tempo change. "Everything that's happening in the North Atlantic is happening before [that]," Yehudai says.

The AMOC and the glaciers may not have been the only factors in the transition, however. Some scientists have suggested that a small drawdown of carbon dioxide (CO₂), perhaps driven by a dust-fertilized plankton bloom in the Southern Ocean, would have been enough to shift the ice age rhythm. Yair Rosenthal, a paleo-oceanographer at Rutgers University in New Brunswick, New Jersey, thinks a CO₂ drop, thickening ice sheets, and a weak AMOC could have all played a role. "I'm not a fan of single triggers of anything." ■

A slowing pulse

Earth's orbital patterns cannot explain a shift in the ice ages' frequency that began about 1 million years ago. Thickening glaciers and crippled ocean currents could be to blame.



Now, the team has analyzed five other ocean cores that also show signs of a weak AMOC. Two of the cores, from the North Atlantic, suggest a possible trigger for the AMOC crisis. In the millennia leading up to it, the neodymium signal sharply trended negative before abating—a sign that an influx of older and older grit from the North Atlantic region had suddenly stopped.

The only plausible explanation, they say, is a long-standing hypothesis advanced by Peter Clark, a glaciologist at Oregon State University in Corvallis, and several others: that the northern ice sheets had finally ground their way to bedrock. Before Earth's current ice age cycles began 3 million years ago, a long warm period

BIOMEDICINE

Monthly shots may replace daily anti-HIV pills

Long-acting treatments now in clinical trials may work as preventives, too

By Jon Cohen

New results are raising hopes for easing one challenge of living with HIV: the need to take daily pills for life, both to ward off AIDS and to lower the risk of transmitting the virus to others. Missing doses can also foster the emergence of HIV strains with drug resistance, a danger both to the person receiving treatment and, if those strains spread, to entire populations. Now, a large-scale study has shown over 48 weeks that monthly injections of two long-acting anti-HIV drugs work just as well as taking daily pills.

ViiV Healthcare, a London-based collaboration between GlaxoSmith-Kline and Pfizer, revealed the highly anticipated findings in a press release on 15 August. This shareholder announcement, required by regulatory agencies to inform investors, offered scant data. But it was welcome news to other researchers studying long-acting anti-HIV medication schemes—and to clinicians who think they could transform both treatment and prevention of HIV infections.

Anton Pozniak, an HIV/AIDS clinician at Chelsea and Westminster Hospital in London, says long-acting injections could help HIV-infected people who have “pill fatigue,” difficulty swallowing the medication, or psychological issues that make it hard to cope with a daily reminder that they have a deadly virus. He calls the results “a fantastic development” but adds, “We still have a way to go.” Even if these injectables win regulatory approval, many practical questions remain about their cost, the impact of missing shots, inflammation at injection sites, and the burden on health care systems of providing monthly intramuscular injections.

The phase III study, called Antiretroviral Therapy as Long-Acting Suppression (ATLAS), is testing an experimental drug, cabotegravir, made by ViiV, and rilpivirine, a licensed medicine from Janssen Sciences Ireland UC in Dublin, in 618 HIV-infected people from 13 countries. All had fully suppressed the virus for at least 6 months with oral drugs. Half stayed on daily pills, while the others received an injection into the but-

tocks of each drug once a month. Viral suppression was the same in both groups, ViiV's statement says.

Kimberly Smith, who heads R&D for ViiV from Durham, North Carolina, says the injectables could make it easier for clinicians to know for certain that patients are adhering to their treatment. Studies have shown that about 30% of HIV-infected people have difficulty doing so at some point; even those taking just a single multidrug pill a day miss doses. Smith, an HIV/AIDS clinician before she came to ViiV, says, “I experienced having patients die ... because they

adherence in people on PrEP. Also underway is a phase III study of the ATLAS regimen in HIV-infected people who have never taken any antiretrovirals. It addresses the possibility that patients in the original trial, who were on oral drugs beforehand, might harbor resistant virus that could undermine the effectiveness of the injections. And a trial called ATLAS 2 will ask whether the injectables can effectively suppress HIV if given only once every 8 weeks.

Other long-acting HIV treatments at earlier phases of clinical testing include injections of a new Gilead Sciences drug that

cripples the making of HIV's capsid protein, a novel target. Researchers also are exploring long-acting alternatives to injections, including a Merck & Co. pill that inhibits HIV's reverse transcriptase enzyme for up to 10 days. In the future, researchers hope to extend the effect of antiretrovirals with slow-release skin implants or polymer-based pills that slowly dissolve in the stomach.

If both phase III studies have positive results, the drugmakers could apply for regulatory approval next year. But Landovitz cautions that many “nuanced questions” remain about how to use these long-acting drugs. What's the optimal dose and timing of injections? What resistance mutations might emerge and how would they hamper the effectiveness of conventional oral drugs?

Although a long-acting drug regimen can be a blessing, Smith says, “it has the potential to be a curse if a person disappears.” Because the drugs metabolize so slowly, they can have an unusually long pharmacological “tail,” their presence steadily declining in blood and tissues for a year or longer, which could allow drug-resistant strains of HIV to flourish.

It's anyone's guess what the long-acting drugs would cost and whether developing countries could afford them. It's also unclear how many people will opt for injections over pills. So only real-world experience will prove whether long-acting interventions will be “a niche or transformative,” Landovitz says. As NIAID head Anthony Fauci puts it, “This is not the end game, but it's an important first step.” ■



In a clinical trial in Mfekayi, South Africa, counselors counted antiretroviral pills to make sure participants took medications as directed.

just couldn't get over that hurdle of taking that pill every day.”

Long-duration anti-HIV drugs could also protect uninfected people at risk of getting the virus. Such people are even more reluctant to take daily pills, as required for so-called pre-exposure prophylaxis (PrEP). Raphael Landovitz, a clinician at the University of California, Los Angeles, who studies long-acting drugs for PrEP, says the ATLAS data are “incredibly encouraging and exciting” and are “certainly reassuring” to people doing similar prevention work.

Landovitz is collaborating on studies of how long-acting injectables, including cabotegravir and, separately, a monoclonal antibody developed by the U.S. National Institute of Allergy and Infectious Diseases (NIAID) in Bethesda, Maryland, improve



China has millions of pig farms, many of them small, such as this one on the outskirts of Beijing.

ANIMAL HEALTH

Arrival of deadly pig disease could spell disaster for China

African swine fever threatens the world's largest pig herd

By Dennis Normile

A nightmare is unfolding for animal health experts: African swine fever (ASF), a highly contagious, often fatal disease of domestic pigs and wild boars, has appeared in China, the world's largest pork producer. As of 21 August, ASF had been reported at sites in four provinces in China's northeast, thousands of kilometers apart. Containing the disease in a population of more than 430 million hogs, many raised in smallholder farmyards with minimal biosecurity, could be a monumental challenge.

"The entry of ASF into China is really a very serious issue," says Yang Hanchun, a swine viral disease scientist at China Agricultural University in Beijing. Given the scale of China's pork sector, the economic impact could be devastating, Yang says, and the outbreak puts a crucial protein source at risk. From China, the virus could also spread elsewhere; if it becomes endemic, "it will represent a major threat for the rest of the world, including the American continent," says François Roger, a veterinary epidemiologist at the Agricultural Research Center for International Development in Montpellier, France.

The virus that causes ASF does not harm humans, but it spreads rapidly among domestic pigs and wild boars through direct contact or exposure to farm workers' contaminated shoes, clothing, and equipment. The virus can survive heat and cold and persists for weeks in carcasses, feces, and fresh and semicured pork products, such as sausages.

Ticks can also spread it. Infection causes a high fever, internal bleeding, and, often, death. There is no ASF vaccine, and no treatment for infected animals.

Endemic in most African countries, ASF jumped to the nation of Georgia in 2007 and later spread through Russia; it has also been reported in Poland and the Czech Republic, and scientists worry about a jump to major pork producers such as Germany and Denmark (*Science*, 22 December 2017, p. 1516). East Asia's first confirmed outbreak occurred on 1 August in Shenyang, a city in Liaoning province, China's Ministry of Agriculture says. Investigators have traced the disease

A threat emerges

As of 21 August, African swine fever had been reported in four provinces in northeastern China.



back through sales of pigs and concluded the virus has been circulating in the area since at least March, says Wantanee Kalpravidh, a veterinarian at the Food and Agriculture Organization's (FAO's) Emergency Centre for Transboundary Animal Disease in Bangkok.

A genetic analysis suggests the virus is closely related to the strain circulating in Russia, scientists from the Institute of Military Veterinary Medicine in Changchun and other Chinese institutions reported on 13 August in *Transboundary and Emerging Diseases*. "The increasing demand for pork has resulted in a great increase in the volume of live pigs and pork products imported to China," heightening the risk of introduction, they wrote. The virus probably arrived in imported pork products, Kalpravidh says, which then infected pigs that were fed contaminated table and kitchen scraps.

A second outbreak occurred on 14 August at a slaughterhouse in Zhengzhou, the capital of Henan province; the afflicted pigs had been shipped from a market in Jiamusi, a town in Heilongjiang province, more than 2000 kilometers to the northeast (see map, below). The virus struck again on 15 August at a farm in Lianyungang, in Jiangsu province. The Chinese government has responded by culling sick and exposed animals—nearly 9000 were killed in Shenyang alone—blockading outbreak areas; disinfecting farms, markets, and processing facilities; controlling the movement of live pigs and pork products; screening animals; and conducting epidemiological surveys.

But there are serious challenges to containing the virus. Pig producers in China range from massive, sophisticated operations to small backyard farms; tailoring a response to suit them all "is the biggest challenge for China to control ASF," Yang says. The complexity of the production chain makes tracing paths of infection "an incredible effort to tackle," says Juan Lubroth, chief veterinarian at FAO's headquarters in Rome. Kalpravidh says China is trying to earn the cooperation of producers by immediately compensating them for culled animals, in hopes of stopping them from slaughtering sick pigs and selling their meat. But ticks and wild boar could also spread the disease, although their role is poorly understood.

So far, Lubroth says, China's "very sophisticated and knowledgeable veterinary workforce" has operated aggressively, and the government has been transparent about ASF's spread. But containing the disease "won't happen overnight," he warns. ■

With reporting by Bian Huihui.

FEATURES



DARING TO HOPE

Patients thrill to reports of a promising antisense drug against Huntington disease, but no one is sure yet whether it works

By **Meredith Wadman**, in Vancouver, Canada;
Photography by **John Lehmann**

The dark shadow of Huntington disease fell squarely over Michelle Dardengo's life on the day in 1986 that her 52-year-old father was found floating in the river in Tahsis, the remote Vancouver Island mill town where she grew up. Richard Varney had left his wedding ring, watch, and wallet on the bathroom counter; ridden his bike to a bridge that spans the rocky river; and jumped. The 4.5-meter drop broke his pelvis. The town doctor happened to be fishing below and pulled Varney out as he floated downstream, saving his life.

But his tailspin continued. The once funny man who read the *Encyclopedia Britannica* for pleasure; the good dancer who

loved ABBA, the Three Tenors, and AC/DC; the affable volunteer firefighter—that man was disappearing. He was being replaced by an erratic, raging misanthrope wedded to 40-ounce bottles of Bacardi whose legs would not stay still when he reclined in his La-Z-Boy.

In 1988, Varney was diagnosed with Huntington disease. That explained his transformation but offered little comfort. Huntington is a brutal brain malady caused by a mutant protein that inexorably robs victims of control of their movements and their minds. Patients are plagued by jerky, purposeless movements called chorea. They may become depressed, irritable, and impulsive. They inevitably suffer from progressive dementia. The slow decline typically

begins in midlife and lasts 15 to 20 years, as the toxic protein damages and finally kills neurons. For both families and the afflicted, the descent is agonizing, not least because each child of an affected person has a 50% chance of inheriting the fatal disease.

In the United States, about 30,000 people have symptomatic Huntington disease and more than 200,000 carry the mutation that ensures they will develop it. Globally, between three and 10 people in every 100,000 are affected, with those of European extraction at highest risk. In a large study published in the *Journal of Neurology* in June, Danish researchers found that people with Huntington disease were nearly nine times more likely to attempt suicide than the general population.





For the first 15 years after her father's diagnosis, while scientists discovered the genetic mutation that causes Huntington and embarked on a seemingly endless chase for a remedy, Dardengo avoided the genetic testing that would reveal whether her father's decline was a preview of her future. Already married when her dad was diagnosed, she gave birth to a son and a daughter and had a job she loved, doing software diagnostics for an insurance company. Then in 2003, the year her father died after being institutionalized for more than a decade, she got tested. "It was bothering me," she says. "You have to plan for the future." She was positive for the mutation.

About a decade later, Dardengo noticed that her handwriting was becoming "unmanageable." Her balance declined. She re-

turned from walking her dogs with bloody knees and scraped hands. In April 2015—at 52, the same age that her father had leapt from the bridge—she was forced to leave her job. "Cognitively, I don't think you are a good fit," a new boss told her.

Then an unfamiliar commodity entered her life: hope. Dardengo's physician, Blair Leavitt, a Huntington disease researcher at the University of British Columbia (UBC) here, asked whether she wanted to join a nascent clinical trial of a new medication made by a Carlsbad, California, biotechnology company, Ionis Pharmaceuticals. The drug, the product of 25 years of research, was an antisense molecule—a short stretch of synthetic DNA tailor-made to block production of the Huntington protein.

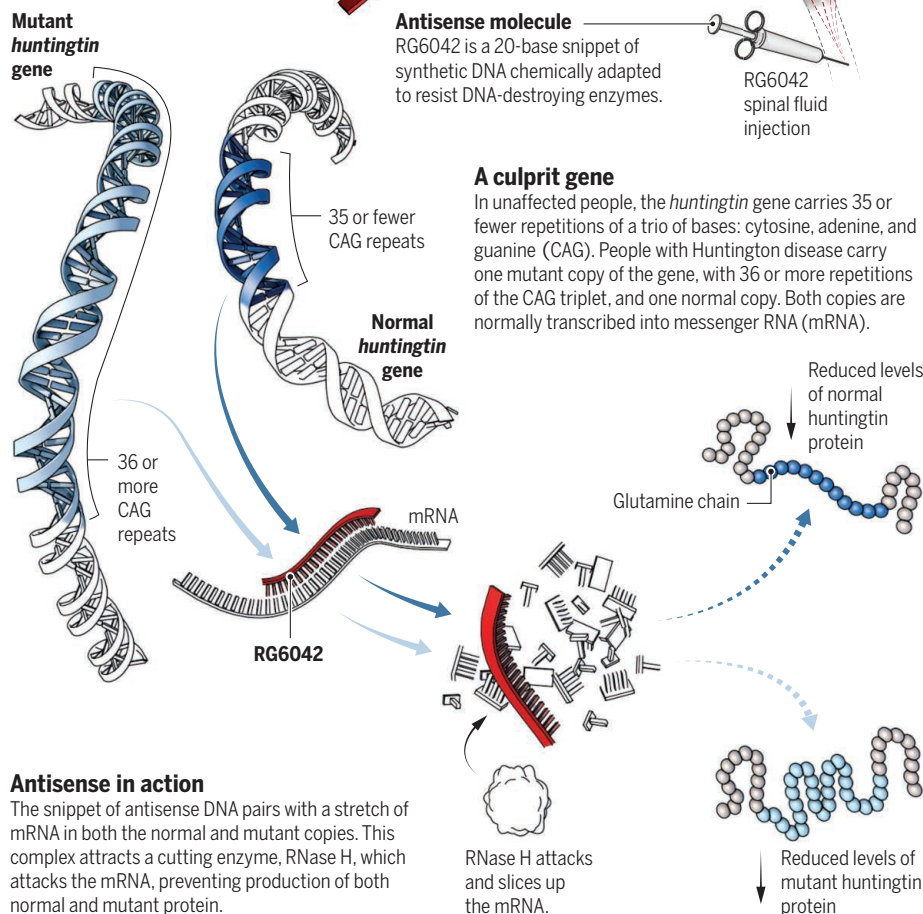
Michelle Dardengo walks her dog (above) near her home in Coquitlam, Canada. Dardengo's father, shown holding her son Joel in a 1989 photo (left), died of Huntington disease. Joel carries the mutation.

Joining the trial would be risky. Dozens of antisense drugs had entered clinical trials; few had become approved drugs. This particular medicine, injected into the fluid that surrounds the spinal cord, had never been put in humans. Dardengo might receive the active drug or she might get a placebo. In August 2015, she became "patient one."

Three years later, that trial is famous. Its results electrified the Huntington disease community in December 2017 by indicating that the drug, formerly IONIS-

Curbing a toxic protein

A new drug, a synthetic snippet of DNA called RG6042, is injected at the lower back into the fluid that bathes the brain and spinal cord; it arrives in key brain cells within 24 hours. Animal studies have shown that the drug, an antisense molecule, penetrates the areas that are most badly damaged in Huntington disease: the corpus striatum and the cerebral cortex.



HTTRx and now called RG6042, reduces the amount of the culprit protein in the human brain—the first time any medicine has done so. Preliminary data released in April even hint at improvements in a few clinical measures.

The Basel, Switzerland-based pharmaceutical company Roche, which licensed RG6042, is preparing to move it into a pivotal clinical trial, with details to be announced in the next few months. “We had emails and phone calls from patients saying, ‘I want to join this [upcoming] trial now. I’m willing to move house to do it,’” recalls Anne Rosser, a neuroscientist at Cardiff University, who oversaw one of nine sites in the first trial.

She and others who work with Hunting-

ton patients find themselves scrambling to manage expectations. “One feels for these patients because the clock is ticking for them,” Rosser says. But although the drug scored high for safety in the recent trial, whether it slows or stops the disease simply isn’t known. “We haven’t got evidence of efficacy,” she says. That will have to wait for the results of the next trial. And some scientists still worry about the drug’s long-term safety.

Given the desperation of patients and their families, Louise Vetter, president and CEO of the Huntington’s Disease Society of America in New York City, says her group and others “need to manage back the hope a little bit. And that’s really hard.”

THE MUTATION that causes Huntington—named after U.S. physician George Huntington, who described the disease in 1872—is located near the tip of the short arm of chromosome 4. Healthy people carry two copies of the *huntingtin* gene, each with up to 35 repetitions of a trio of nucleotides—cytosine, adenine, and guanine (CAG).

In the mutated gene, this CAG triplet is repeated 36 or more times, like a stutter. People who inherit a copy of the mutant gene from one parent produce a mutant huntingtin protein with an extra-long chain of glutamine amino acids.

No one is sure just how the mutant protein damages neurons. But chains of glutamine longer than 36 amino acids are much more likely to stick to each other, forming clumps that may disrupt membranes or bind and inactivate other molecules within neurons. Afflicted people produce mutant huntingtin throughout their lives, and the damage mounts until symptoms appear. The more CAG repeats a patient harbors, the sooner the inevitable day comes.

Regardless of how the destruction happens, scientists reasoned, stopping production of the mutant protein at its source should halt downstream damage. Back in 1992, *Science* declared antisense technology one of the 10 hottest areas of the year; 1 year later, the mutation that causes Huntington was published. Because just a single gene was responsible, shutting it down with antisense seemed an obvious approach for a therapy. The idea was that synthetic DNA snippets with a sequence complementary to part of the *huntingtin* gene’s messenger RNA (mRNA) would bind to the mRNA. In this case, the resulting DNA-mRNA complex summons a cutting enzyme, RNase H, that degrades the mRNA and prevents its translation into protein (see graphic, left).

But for 25 years the idea wasn’t practical. Early antisense oligonucleotides (ASOs) didn’t bind mRNA targets efficiently and were quickly degraded by enzymes. Several companies dropped out. By the early 2000s, having chemically improved the ASOs and done animal studies, Ionis began to tailor ASOs to specific neurological diseases, including one to curb spinal muscular atrophy (*Science*, 16 December 2016, p. 1359).

In 2006, Ionis researchers, led by the firm’s vice president of research, Frank Bennett, launched a collaboration with neuroscientist Don Cleveland and postdoc Holly Kordasiewicz of the University of California, San Diego, to develop an ASO to attack Huntington disease. (Kordasiewicz has since joined Ionis.) By 2012, they had a 20-letter ASO that reversed symptoms and slowed brain atrophy in mouse models of

Huntington. In macaques, dogs, and pigs, the researchers injected the ASO into the cerebrospinal fluid (CSF), which bathes the spinal cord and brain. The drug reduced the mutant protein in key brain regions, including the cortex and the corpus striatum, a deeper brain structure that is the disease's first target. And it appeared to be safe.

In early 2015, researchers at several institutions delivered the last tool needed for a human trial: new tests that could detect levels of mutant huntingtin protein in the CSF. Crucially, one group also showed that in mice, reductions of mutant protein in the CSF corresponded to reductions in the brain itself. Now, researchers could collect human CSF and see reflected there how their ASO likely affected levels of the toxic protein in patients' brains.

IN AUGUST 2015, Dardengo lay on her side at UBC Hospital, her face to the wall. She felt Leavitt's needle inject anesthetic in the small of her back. After that, she felt nothing while he withdrew 20 milliliters of her spinal fluid and replaced it with the medication, dissolved in liquid. That was the first of four monthly injections. The trial was double-blind: Neither she nor Leavitt knew whether she was receiving placebo or the lowest of the planned doses of active drug, the only dose given to the first subjects.

Dardengo was unfazed by being patient one. "I never felt afraid," she recalls. "Dr. Leavitt always asked: 'You know that you could die doing this?' I always said: 'Yes, but my kids aren't going to have to worry.'"

She thought she felt better after the last monthly injection, but she saw no improvement in her symptoms. Her balance was still off—she didn't dare wear high heels—and she still got stuck in the middle of sentences. Months earlier, she had quit clipping her cat's claws for fear of injuring the animal. Her husband Marc Dardengo had stopped asking her to run errands for his business because she so often got lost. These things didn't change.

Then in November 2016, she and Marc received news that left them numb. Joel Dardengo, age 27, told his parents he had been tested for the Huntington mutation. He was positive.

His mother's trial was cold comfort for Joel, a realist keenly aware of the pitfalls of science. "You can YouTube the disease and see what the endgame looks like," he says. "It's not the nicest."

Joel told his fiancée that she was free to leave him. She refused. The couple, who want to have children, began to discuss in

vitro fertilization, which would allow them to implant only Huntington-free embryos. But Joel, an electronics technician, and his fiancée, a prison guard, see no way to pay for the \$30,000 procedure.

Thirteen months later, at 12:01 a.m. on 11 December 2017, Michelle received an email: Ionis had announced the results of the trial in a press release. (A paper is still in the works.) After testing in 46 people, including her, RG6042 appeared to be safe, with minimal side effects. And after four monthly doses, levels of mutant huntingtin in patients' CSF had decreased in a dose-dependent manner: The antisense molecule appeared to be hitting its target.

The results triggered a happy uproar among Huntington families. Michelle and Marc Dardengo uncorked a bottle of chardonnay and toasted the possibility that their children would live Huntington-free lives. Vetter, 5000 kilometers away, got a text message from her scientific director

"I feel so much hope. Things are so different between my dad and me: I've got a life to live."

Michelle Dardengo

telling her to check her email immediately. She pulled over to the side of New Jersey's Route 23, read the Ionis press release, and wept. At Huntington disease clinics from Baltimore, Maryland, to Bochum, Germany, phones began buzzing, email inboxes overflowing, and appointment requests surging. At University College London, home to the trial's global leader, neurologist Sarah Tabrizi, patient referrals briefly quadrupled.

On the same day the results were revealed, Roche announced it had paid \$45 million to license the drug from Ionis and would move it into a pivotal clinical trial. That trial will enroll hundreds of patients in Europe, Canada, and the United States for up to 2 years, allowing the company to determine whether RG6042 slows or stops the disease.

Ionis presented more details about the first trial's results at two meetings earlier this year. In March, the company reported that levels of mutant protein in the two highest-dose patient groups had declined on average by 40% and in some cases up to 60%. (In lab animals, smaller reductions had reduced symptoms of the disease.) And in April, Tabrizi reported that aggregated data from all patients showed statistically significant improvement in three of five clinical measures of Huntington disease progression, including tests of both motor and cognitive function.

But researchers were quick to temper the resulting excitement with caution. Bennett notes that other clinical measures did not improve significantly. "You can always find some correlations if you look hard enough. ... The only way to really validate this is to do additional longer-term studies."

Tabrizi, who like Leavitt consults for Roche and competing companies, agrees. She points out that the first trial wasn't designed to gauge effectiveness. "The big question we have now is: Is the degree of mutant huntingtin lowering in the brain enough to slow the disease's progression?"

At the moment, no one can say whether reduced protein levels will help already damaged brains recover, how long any positive effects might last, or whether dangerous side effects will crop up down the road. Tabrizi says patients should not expect "some Lazarus-like effect."

One scientist not involved with the trial is even more circumspect. "People said in 1993:

'We have the gene—we have the cure,'" recalls Ole Isacson, a neurodegenerative disease expert at the Harvard Stem Cell Institute. In the mid-1990s, he ran a first, failed experiment trying to suppress huntingtin production in mice using an

ASO. "And 25 years later, there may be a slight chance. This is not a sure home run."

ALTHOUGH THE FIRST TRIAL appeared to show that RG6042 is safe, some researchers still have concerns because the drug suppresses production of both the normal and mutant forms of the huntingtin protein. Because mutant *huntingtin* genes vary in sequence, a drug targeted to one version of the mutant allele would not work for every patient. It also would be technically more challenging to develop. So Ionis opted for a nonselective ASO that targets a sequence found in all *huntingtin* genes, mutant and normal.

Doing so was a calculated risk. No one knows whether normal huntingtin is essential to brain health in adults. (It is essential to early development: Mice without it perish as embryos.) Bennett notes that RG6042's effects are less drastic than deleting the gene entirely: "We're not knocking out huntingtin [protein], we're reducing it." Leavitt shares his confidence, noting that in Ionis's in-house work, "Nothing from the preclinical animal studies suggested any toxicity."

But 1 year ago, scientists at The University of Tennessee Health Science Center (UTHSC) in Memphis found something different, as they reported in *PLOS Genetics*. They knocked out the huntingtin protein

in mice at 3, 6, and 9 months of age. At all ages, the mice developed severe motor and behavioral deficits and progressive brain pathology: The animals' thalami—sensory and motor signal relay stations—became calcified, and their brains atrophied.

"We end up sounding like a Cassandra," says Paula Dietrich, a neurobiologist at UTHSC who was the paper's first author. "We bring the bad news, and nobody wants to hear or believe it."

Dietrich concedes that mice findings may not apply to humans. But she notes that *huntingtin* is expressed throughout the brain and cautions that the levels of

says Chandra Vargeese, Wave's senior vice president and head of research. "The expanded allele is the causative factor. So why not leave the healthy allele intact?"

BY ABOUT 24 HOURS AFTER injection into the CSF, RG6042 has diffused into neurons in the brain, where it remains for 3 to 4 months. Bennett says animal studies tracing the drug's uptake and activity suggest that monthly injections of the highest trial dose might begin to measurably improve symptoms after 6 months—if the drug works.

In January, Michelle Dardengo began to receive monthly injections of 120 milligrams

Dardengo also reports that she and her dogs are walking farther, on rougher trails, than was her habit in January. She added: "My handwriting is becoming quite legible now." In July, Dardengo noted that she has resumed clipping the cat's claws and sent a video of herself doing so with apparently steady hands. Marc Dardengo confirms that his wife's walking stamina and handwriting have improved, and he notes that her violent leg movements while sleeping have virtually disappeared. And he thinks that the deterioration of her memory has stabilized. "Before the start of the [open label] trial, I could see her memory [go downhill] over the course of a year. Since January, [with her] taking this drug, I am not seeing that type of change."

But he cautioned that his wife's verbal fluency is a moving target. "She has easier days with conversation more often than before. But she also has days when the dots aren't connecting."

And Joel Dardengo believes his mother's memory is getting worse. In June, for example, she forgot to buy him a birthday cake, a long-standing tradition. Paying for a pricey meal at a restaurant, she uncharacteristically failed to tip the server.

Trial scientists Tabrizi and Leavitt strongly caution that one patient's experience is an anecdote, not a study. "An *n* of one is just that," Leavitt says, adding: "The placebo effect is very real. We don't want to put out the impression that [the drug] is working. Because we don't know yet." Tabrizi says she and her team are doing their all—by phone, email, and

in person—to communicate that to patients.

Today, Michelle Dardengo says her goal is to ride her bicycle in 6 months. To win Leavitt's permission to attempt this, she needs to pass the heel-to-toe walking test administered to suspected drunk drivers without losing her balance, a feat that eludes her today.

She is nonetheless buoyant. "I feel so much hope," she says. "Things are so different between my dad and me: I've got a life to live." For her son, her hopes are even higher: "Joel could definitely be a candidate to potentially never see or experience Huntington."

Joel himself is far more cautious. "I do wish for the best," he says. "At the same time, I do prepare for the worst." ■



Michelle Dardengo, on her way to a hiking trail with her dog, says her confidence behind the wheel has improved since last year.

RG6042 needed to penetrate the deeper structures where Huntington pathology begins could have unwanted effects. "To reach efficacy in the striatum, you may have to expose other regions of the brain to extremely high concentrations," she says.

Another company, Wave Life Sciences of Cambridge, Massachusetts, is developing two antisense drugs aimed at suppressing only the mutant gene, leaving the normal one untouched. Wave's ASOs target two sequence variants, one or both of which occur in 70% of Huntington patients. Both drugs are in initial clinical trials, with results expected next summer, and the firm thinks its approach may be safer than RG6042.

"No one knows the longer-term implications of knocking down the healthy allele,"

of RG6042—the highest dose dispensed in the first study. Along with 45 other patients from that study, she is taking part in an open-label extension of the trial. It is designed to produce long-term safety data and to chart all 46 patients' disease progression. But for Dardengo, it is a trial of the drug's promise.

In an in-person interview in March, shortly after her third such injection, Dardengo appeared tired. Her speech was sometimes halting and stopped midsentence.

Three months later, in a telephone interview in June, she spoke fluently, in complete sentences. She was still gamely answering questions after 2 hours. "I find that my speaking is a lot better. Like, I can actually finish a sentence," she volunteered.

WATER

The paradox of irrigation efficiency

Higher efficiency rarely reduces water consumption

By **R. Q. Grafton^{1,2}, J. Williams¹, C. J. Perry³, F. Molle⁴, C. Ringler⁵, P. Steduto⁶, B. Udall⁷, S. A. Wheeler⁸, Y. Wang⁹, D. Garrick¹⁰, R. G. Allen¹¹**

Reconciling higher freshwater demands with finite freshwater resources remains one of the great policy dilemmas. Given that crop irrigation constitutes 70% of global water extractions, which contributes up to 40% of globally available calories (1), governments often support increases in irrigation efficiency (IE), promoting advanced technologies to improve the “crop per drop.” This provides private benefits to irrigators and is justified, in part, on the premise that increases in IE “save” water for reallocation to other sectors, including cities and the environment. Yet substantial scientific evidence (2) has long shown that

increased IE rarely delivers the presumed public-good benefits of increased water availability. Decision-makers typically have not known or understood the importance of basin-scale water accounting or of the behavioral responses of irrigators to subsidies to increase IE. We show that to mitigate global water scarcity, increases in IE must be accompanied by robust water accounting and measurements, a cap on extractions, an assessment of uncertainties, the valuation of trade-offs, and a better understanding of the incentives and behavior of irrigators.

LOGIC AND LIMITS

Field IE is the ratio of the volume of all irrigation water beneficially used on a farmer's field [predominantly, evapotranspiration (ET) by crops and salt removal to maintain soil productivity] to the total volume of irrigation water applied (adjusted for changes in water

stored for irrigation in the soil) (2). Annually, governments spend billions of dollars subsidizing advanced irrigation technologies, such as sprinklers or drip systems (3). Sometimes their goal is to increase IE on the understanding that this will allow water to be reallocated from irrigation to cities (4), industry, or the environment, while maintaining or even increasing agricultural production.

But water saved at a farm scale typically does not reduce water consumption at a watershed or basin scale. Increases in IE for field crops are rarely associated with increased water availability at a larger scale (5), and an increase in IE that reduces water extractions may have a negligible effect on water consumption. This paradox, that an increase in IE at a farm scale fails to increase the water availability at a watershed and basin scale, is explained by the fact that previously nonconsumed water “losses” at a farm scale (for ex-

Sprinkler irrigation supports grape vines in the Okanagan Basin, British Columbia.

ample, runoff) are frequently recovered and reused at a watershed and basin scale.

Advanced irrigation technologies that increase IE may even increase on-farm water consumption, groundwater extractions (6), and water consumption per hectare (5). At a farm scale, this can arise from a switch to more water-intensive crops and, with the same crop, may occur when there is a strong marginal yield response from additional water. Moreover, the absence of an increase in water consumption per hectare because of a higher IE does not necessarily mean that the water potentially available for reallocation and reuse (see supplementary materials) at a watershed or basin scale increases. Subsidies for drip irrigation may reduce the water applied per hectare and increase water extractions because a higher IE can induce increases in the irrigated area, as shown for the Lower Rio Grande, New Mexico (7).

Although the hydrology related to IE has been known for decades, it is often overlooked or ignored. For example, the United Nations (UN) High-Level Panel on Water, comprising 11 sitting heads of state or government, recommends "...incentives for water users, including irrigators, to use water efficiently" (8) but fails to explicitly recognize that this may increase, rather than decrease, water consumption. Similar to IE, there is also confusion in policy circles about the effects of an increase in efficiency or water productivity (the biophysical or monetary output per volume of water inputs) on basin-scale water availability (see supplementary materials). The UN Sustainable Development Goal (SDG) 6.4, for instance, seeks to increase water use efficiency, but this does not necessarily mean reduced water extractions.

There are reasons why this evidence may be overlooked by policy-makers: Evidence resides in a specialized literature; subsidies for IE can promote rent-seeking behavior by beneficiaries who lobby to continue subsidies; and comprehensive water accounting from the scale of the field to that of the watershed or basin is necessary but frequently absent. Such accounting quantifies field water applications; ET by crops and weeds; evaporation from soil and water surfaces; and, particularly, surface and subsurface water flows returned to the environment or utilized elsewhere at the watershed or basin scale.

RESPONDING TO THE PARADOX

We respond to the paradox (2, 9) with two key insights and a research and policy agenda to deliver on SDG 6 ("ensure availability and sustainable management of water and sanitation for all"). First, irrigation systems are frequently managed to maximize irrigated crop production. This provides benefits but means more water is transpired locally and lost for other uses. Second, locally extracted, but not consumed, water flows to surface supplies and groundwater. Such volumes, perceived as losses to farmers and the irrigation system, do not disappear. They frequently have value and are typically recovered and reused elsewhere in a watershed or basin.

The figure visualizes the paradox within a watershed, showing three types of irrigation with different IEs: drip, sprinkler, and surface. Inflows are precipitation and interbasin transfers. Outflows are (i) beneficial water consumption from transpiration by crops; (ii) nonbeneficial water consumption through transpiration by weeds and evaporation from wet soil, foliage, and open water surfaces; (iii) locally recoverable return flows to surface water systems, from drains and surface runoff, and also to aquifers via subsurface recharge; and (iv) nonrecoverable flows to sinks, such as to saline groundwater and the ocean. Inflows less outflows over a given time period equals the change in water storage.

Conservation of mass requires that increased local beneficial water consumption, because of a higher IE, be fully offset by a decline in some combination of nonbeneficial water consumption, recoverable return flows (to surface or groundwater), and nonrecoverable flows to sinks. Thus, a higher IE (typically 90% for drip versus 50% for surface) is associated with lower rates of nonbeneficial water consumption, usually because of reduced soil evaporation (5% for drip and 20% for surface). These changes from a higher IE also result in a reduction in return flows, from 30% of water applied, in the case of surface irrigation, to 5%, for drip.

Studies in several locations confirm the effects of higher IE, including (i) Rajasthan, India, where subsidies for drip irrigation improved farm incomes but also increased the irrigated area and total volume of water applied by farmers (10); (ii) Snake River, Idaho, where farmers have increased their IE, but this has reduced groundwater recharge and led to a decline in the Eastern Snake Plain

Aquifer by about 30% since the mid-1970s, despite increased precipitation (11); (iii) the Rio Grande in the United States, where subsidies for drip irrigation increase crop yields and irrigators' net income but can reduce downstream flows and the water potentially available for other purposes (7); and (iv) the Souss and Tensift Basins of Morocco, where the adoption of drip irrigation, supported by subsidies, reduced recoverable return flows, principally to overexploited aquifers. This led to increased water consumption and exacerbated groundwater overexploitation in Morocco because of crop intensification, especially denser tree plantations; increased irrigated area owing to improved control of water; and a greater area of crops with higher water-use requirements (12).

These four cases, and others (5), show that increases in IE are typically associated with a reduction in recoverable return flows and an increase in crop yields and in crop transpiration. Contrary to the policy intent, however, a higher IE is not usually associated with a decline in water consumption. Only when a commensurate decrease in some combination of nonbeneficial water consumption and nonrecoverable flows is observed is it possible to reallocate water to other uses at a watershed or basin scale after an increase in IE (see supplementary materials).

Scientific understanding of the paradox highlights the importance of a comprehensive evaluation of the public costs of subsidizing increases in IE. This, in turn, requires that the estimated benefits (such as higher yields and farm net incomes) be compared to the external costs from induced reductions in recoverable return flows (such as groundwater degradation, losses to aquatic ecosystems, reduced environmental water volumes, removal of salts from watersheds and basins, and other water uses).

POLICY AND RESEARCH IMPLICATIONS

If increases in IE are to mitigate the global water crisis, then decisive actions, some of which have previously been highlighted (3, 5, 7, 9), are required. A key constraint to better decision-making is inadequate estimates of water inflows and outflows at watershed and basin scales. This analysis of water accounts is essential to demonstrate when IE policies are or are not in the public interest. Furthermore, successful integration of science into policy and practice requires several precon-



¹Crawford School of Public Policy, The Australian National University, ACT 2601, Australia. ²Groupe de Recherche en Économie Théorique et Appliquée (GREThA), UMR CNRS 5113, University of Bordeaux, 33608 Pessac, France. ³Consultant, London NW8 8QX, UK. ⁴Institut de Recherche pour le Développement (IRD) and G-Eau, University of Montpellier, 34196 Montpellier, France. ⁵Environment and Production Technology Division, International Food Policy Research Institute, Washington, DC 20005-3915, USA. ⁶Food and Agriculture Organization of the United Nations (FAO), Regional Office for the Near East and North Africa, 12311 Cairo, Egypt. ⁷Colorado Water Institute, Colorado State University, Fort Collins, CO 80523-1033, USA. ⁸Centre for Global Food and Resources, Faculty of Professions, University of Adelaide, Adelaide, SA 5001, Australia. ⁹School of Public Policy and Management—Institute for Contemporary China Studies, Tsinghua University, 100084 Beijing, China. ¹⁰Smith School of Enterprise and the Environment, University of Oxford, Oxford OX1 3QY, UK. ¹¹Kimberly Research and Extension Centre, University of Idaho, Kimberly, ID 83341, USA. Email: quentin.graffon@anu.edu.au

ditions. To avoid “regulatory capture,” there must be transparent and independent auditing of policy processes and data provision. There also needs to be public interest in the issue such that there is a cost to policy-makers who fail to act for the public good. And alignment of public interest-seeking actors, supported by transparent data and evidence, mitigates water misuse and misallocation.

We outline five steps, centered on water accounting and research advances, that promote more effective policy actions. First, physical water accounts need to be developed from the farm-scale to the basin scale to make transparent “who gets what and where” to support decision-making in the

extractions through a direct cap on water offtakes (9) or on the irrigated area. The need for such caps when promoting IE has been identified in the European Union and the western United States, where water rights have been denominated as net extractions that require the calculation of return flows. Water accounting in California, which includes ET, is providing decision-makers with the information needed to determine how much to reduce water consumption to ensure sustainable extractions. By contrast, in Australia, where water rights are denominated in gross extractions, actions to reduce extractions to reallocate water to the environment have,

tainty, better data quality and quantity, user-friendly software, and increased computing power all facilitate greater consideration of risks in future water planning.

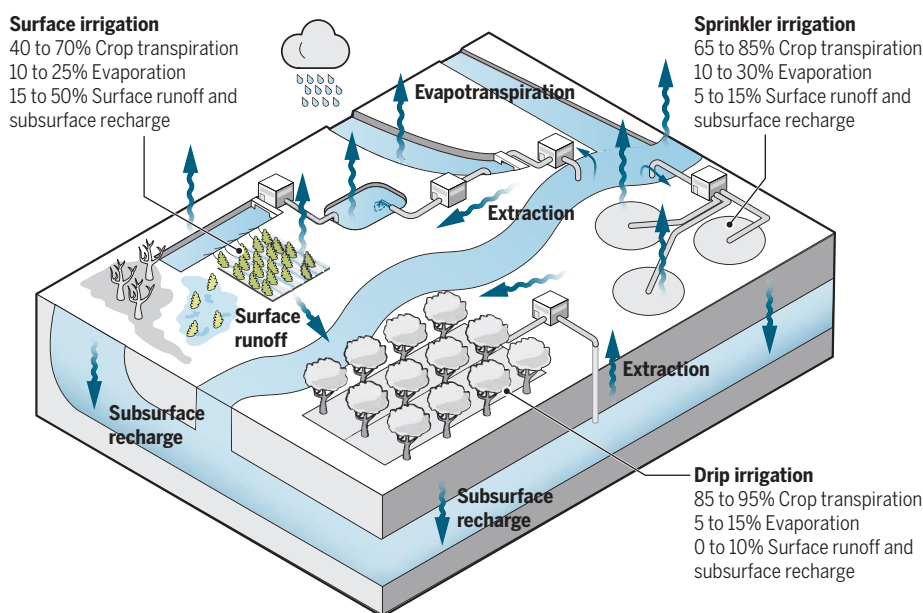
Fourth, although understanding water inflows and outflows is necessary, the payoff from subsidizing IE depends on whether the benefits exceed the costs, including those associated with reduced return flows. Comprehensive methods of valuation can make these trade-offs more explicit, as can advances in water accounting and measurements of changes in water quality.

Finally, the effects of policy actions (5) on the behavior of irrigators must be evaluated. Neither IE nor water extractions are constant: They vary by irrigator and differ by land and soil characteristics, crops grown, time of year, and weather conditions. Differences are more readily understood with developments in behavioral and experimental economics and by testing how irrigators’ actions change as IE increases. Such methods identify incentives for irrigators to maintain agricultural production with less water extracted.

Overcoming misunderstandings about the paradox of IE is required if SDG 6 is to be achieved. Our five-step reform of the current IE policy agenda—centered on water accounting and reductions in irrigation water extractions which are informed by advances in water valuation, risk assessment, and behavioral economics—offers a pathway to improved global water security. ■

Accounting for water

The paradox of irrigation efficiency (surface, sprinkler, and drip) and the water inflows and outflows can be seen in a watershed example. Ranges of crop transpiration, evaporation, runoff, and recharge are authors’ judgment of possible values. These values depend on crop and soil types, weather, and other factors.



public interest. This requires measurement or estimation of all inflows, water consumption, recoverable return flows, and nonrecoverable flows to sinks. Although a priority by the UN High-Level Panel on Water (13), robust and transparent water accounting is the exception. In some jurisdictions—such as Spain (9), Morocco (12), and the Murray-Darling Basin, Australia (14)—several billion U.S. dollars have been spent subsidizing IE, including canal lining and drip irrigation, without proper accounting of their effects on recoverable return flows, aquifers, and river ecology. Developments in remote sensing offer the possibility of estimates of water inflows and outflows at a much lower cost and a greater scale than previously available.

Second, reductions in water consumption are achievable by decreases in water

to date, been neither sufficient nor cost-effective (14). To meet environmental flow goals, incentives may be used to make irrigators account for return flows, such as water charges on the reductions in recoverable flows, or financial benefits to maintain such flows by reducing consumption. Incentive-based water reallocations, however, can be constrained by the funding needed to compensate users to facilitate transfers across competing water uses.

Third, to ensure desired outcomes are delivered, risk assessments are needed when evaluating the effects of increased IE, as are accurate measurements from on-the-ground monitoring of flows. Policy-makers must account for uncertainties in key water parameters when calculating water flows (15). Advances in decision-making under uncer-

REFERENCES AND NOTES

1. R. Q. Grafton et al., *Earth's Future* **5**, 449 (2017).
2. C. M. Burt et al., *ASCE J. Irrigation Drainage Eng.* **123**, 423 (1997).
3. S. M. Scheierling et al., *Water Resour. Res.* **41**, W03248 (2006).
4. M. Flörke et al., *Nature Sustainability* **1**, 51 (2018).
5. C. Perry et al., “Does improved irrigation technology save water? A review of the evidence” (Food and Agriculture Organization of the United Nations, Cairo, 2017), p. 42.
6. L. Pfeiffer, C.-Y. C. Lin, *J. Environ. Econ. Manage.* **67**, 189 (2014).
7. F. A. Ward, M. Pulido-Velazquez, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18215 (2008).
8. United Nations High-Level Panel on Water, “Making Every Drop Count. An Agenda for Water Action” (HPLW Outcome Report, United Nations, 2018).
9. C. A. Scott et al., *Hydrol. Earth Syst. Sci.* **18**, 1339 (2014).
10. T. Birkenholtz, *Water Int.* **42**, 663 (2017).
11. M. McVeigh, A. Wyllie, Memo on irrigation efficiency and ESPA storage changes, State of Idaho Department of Water Resources, 5 March 2018.
12. F. Molle, O. Tanouti, *Agric. Water Manag.* **192**, 170 (2017).
13. Australian Bureau of Meteorology, “Good practice guidelines for water data management policy: World Water Day Initiative” (Bureau of Meteorology, Melbourne, 2017).
14. R. Q. Grafton, S. A. Wheeler, *Ann. Rev. Resour. Economics* **10**, 487 (2018).
15. A. J. Clemmens, C. M. Burt, *J. Irrigation Drainage Eng.* **123**, 443 (1997).

ACKNOWLEDGMENTS

S.A.W. was supported by the Australian Research Council project FT140100773; C.R. was supported by the CGIAR Research Program on Water, Land, and Ecosystems; and F.M. was supported by the Agence Nationale de la Recherche (ANR) AMETHYST project (ANR-12-TMED-0006-01).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/page/suppl/DC1

10.1126/science.aat9314

ARTIFICIAL INTELLIGENCE

How AI can be a force for good

An ethical framework will help to harness the potential of AI while keeping humans in control

By **Mariarosaria Taddeo**^{1,2,3}
and **Luciano Floridi**^{1,2}

Artificial intelligence (AI) is not just a new technology that requires regulation. It is a powerful force that is reshaping daily practices, personal and professional interactions, and environments. For the well-being of humanity it is crucial that this power is used as a force of good. Ethics plays a key role in this process by ensuring that regulations of AI harness its potential while mitigating its risks.

AI may be defined in many ways. Get its definition wrong, and any assessment of the ethical challenges of AI becomes science fiction at best or an irresponsible distraction at worst, as in the case of the singularity debate. A scientifically sound approach is to draw on its classic definition (1) as a growing resource of interactive, autonomous, self-learning agency, which enables computational artifacts to perform tasks that otherwise would require human intelligence to be executed successfully (2). AI can then be further defined in terms of features such as the computational models on which it relies or the architecture of the technology. But when it comes to ethical and policy-related issues, the latter distinctions are unnecessary (3). On the one hand, AI is fueled by data and therefore faces ethical challenges related to data governance, including consent, ownership, and privacy. These data-related challenges may be exacerbated by AI, but would occur even without AI. On the other hand, AI is a distinct form of autonomous and self-learning agency and thus raises unique ethical challenges. The latter are the focus of this article.

The ethical debate on AI as a new form of agency dates to the 1960s (2, 4). Since then, many of the relevant problems have concerned delegation and responsibility. As AI is used in ever more contexts, from recruitment

to health care, understanding which tasks and decisions to entrust (delegate) to AI and how to ascribe responsibility for its performance are pressing ethical problems. At the same time, as AI becomes invisibly ubiquitous, new ethical challenges emerge. The protection of human self-determination is one of the most relevant and must be addressed urgently. The application of AI to profile users for targeted advertising, as in the case of online service providers, and in political campaigns, as unveiled by the Cambridge Analytica case, offer clear examples of the potential of AI to capture users' preferences and characteristics and hence shape their goals and nudge their behavior to an extent that may undermine their self-determination.

DELEGATION AND RESPONSIBILITY

AI applications are becoming pervasive. Users rely on them to deal with a variety of tasks, from delivering goods to ensuring national defense (5). Assigning these tasks to AI brings huge benefits to societies (see the photo). It lowers costs, reduces risks, increases consistency and reliability, and enables new solutions to complex problems. For example, AI applications can lower diagnostic errors by 85% in breast cancer patients (6), and AI cybersecurity systems can reduce the average time to identify and neutralize

cyberattacks from 101 days to a few hours (5). However, delegation may also lead to harmful, unintended consequences, especially when it involves sensitive decisions or tasks (7, 8) and excludes or even precludes human supervision (3). The case of COMPAS, an AI legal system that discriminated against African-American and Hispanic men when making decisions about granting parole (9), has become infamous. Robust procedures for human oversight are needed to minimize such unintended consequences and redress any unfair impacts of AI.

Still, human oversight is insufficient if it deals with problems only after they occur. Techniques to explain AI and predict its outcomes are also needed. The Explainable Artificial Intelligence program of DARPA (Defense Advanced Research Project Agency)

is an excellent example. The goal of this program is to define new techniques to explain the decision-making processes of AI systems. This will enable users to understand how AI systems work, and designers and developers to improve the systems to avoid mistakes and mitigate the risks of misuse. To be successful, similar projects must include an ethical impact analysis from the beginning, to assess AI's benefits and risks and define guiding principles for an ethically sound design and use of AI.

The effects of decisions or actions based on AI are often the result of countless interactions among many actors, including designers, developers, users, software, and hardware. This is known as distributed agency (10). With distributed agency comes distributed responsibility. Existing ethical frameworks address individual, human responsibility, with the goal of allocating punishment or reward based on the actions and intentions of an individual. They were not developed to deal with distributed responsibility.

Only recently have new ethical theories been defined to take distributed agency into account. The proposed theories rely on contractual and tort liability (11) or on strict liability (12) and adopt a faultless responsibility model. This model separates responsibility of an agent from their intentions to perform a given action or their ability to control its outcomes, and holds all agents of a distributed system, such as a company, responsible. This is key when considering the case of AI, because it distributes moral responsibility among designers, regulators, and users. In doing so, the model plays a central role in preventing evil and fostering good, because it nudges all involved agents to adopt responsible behaviors.

Establishing good practices for delegation and defining new models to ascribe moral responsibility are essential to seize the opportunities created by AI and address the related challenges, but they are still not enough. Ethical analyses must be extended to account for the invisible influence exercised by AI on human behavior.

INVISIBILITY AND INFLUENCE

AI supports services, platforms, and devices that are ubiquitous and used on a daily basis. In 2017, the International Federation of Robotics suggested that by 2020, more than 1.7 million new AI-powered robots will be installed in factories worldwide. In the same year, the company Juniper Networks issued a report estimating that, by 2022, 55% of households worldwide will have a voice assistant, like Amazon Alexa.

As it matures and disseminates, AI blends into our lives, experiences, and environ-



Read more articles
online at [scim.ag/
TomorrowsEarth](http://scim.ag/TomorrowsEarth)

¹Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3JS, UK. ²The Alan Turing Institute, 96 Euston Road, London NW1 2DB, UK. ³Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK. Email: mariarosaria.taddeo@oii.ox.ac.uk

ments and becomes an invisible facilitator that mediates our interactions in a convenient, barely noticeable way. While creating new opportunities, this invisible integration of AI into our environments poses further ethical issues. Some are domain-dependent. For example, trust and transparency are crucial when embedding AI solutions in homes, schools, or hospitals, whereas equality, fairness, and the protection of creativity and rights of employees are essential in the integration of AI in the workplace. But the integration of AI also poses another fundamental risk: the erosion of human self-determination due to the invisibility and influencing power of AI.

This invisibility enhances the influencing power of AI. With their predictive capabilities and relentless nudging, ubiquitous but imperceptible, AI systems can shape our choices and actions easily and quietly. This is not necessarily detrimental. For example, it may foster social interaction and cooperation (13). However, AI may also exert its influencing power beyond our wishes or understanding, undermining our control on the environment, societies, and ultimately on our choices, projects, identities, and lives. The improper design and use of invisible AI may threaten our fragile, and yet constitutive, ability to determine our own lives and identities and keep our choices open.

TRANSLATIONAL ETHICS

To deal with the risks posed by AI, it is imperative to identify the right set of fundamental ethical principles to inform the design, regulation, and use of AI and leverage it to benefit as well as respect individuals and societies. It is not an easy task, as ethical principles may vary depending on cultural contexts and the domain of analysis. This is a problem that the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (14) tackles with the aim of advancing public debate on the values and principles that should underpin ethical uses of AI.

More important, some agreement on the fundamental principles is emerging. A recent comparative analysis (15) of the main international initiatives focusing on AI ethics highlights substantive overlap of the principles endorsed by these initiatives and some of the key principles of bioethics, namely beneficence, nonmaleficence, autonomy, and justice. There is reason to be optimistic about further convergence, as other principles may be extracted from the Universal Declaration of Human Rights. This convergence will foster coherence, and hence compatibility, of different ethical frameworks for AI and provide overarching ethical guidance for the design,

regulations, and uses of this technology.

Once identified, ethical principles must be translated into viable guidelines to shape AI-based innovation. Such translation has precedents, especially in medicine, where translational research goes “from bench to bedside,” building on research advances in



The Avatar Kids project allows hospitalized children to be present in the classroom through a remote-controlled robot.

biology to develop new therapies and treatments. Likewise, translational ethics builds on academic advances to shape regulatory and governance approaches. This approach underpins the forthcoming recommendations for the ethical design and regulation of AI to be issued by the AI4People project.

Launched in the European Parliament in February 2018, AI4People was set up to help orient AI toward the good of society and everyone in it. The initiative combines efforts of a scientific committee of international experts and a forum of stakeholders, in consultation with the High-Level Expert Group on Artificial Intelligence of the European Commission, to propose a series of concrete and actionable recommendations for the ethical and socially preferable development of AI.

A translational ethics of AI needs to formulate foresight methodologies to indicate ethical risks and opportunities and prevent unwanted consequences. Impact assessment analyses are an example of this methodology. They provide a step-by-step evaluation of the impact of practices or technologies deployed in a given organization on aspects such as privacy, transparency, or liability.

Foresight methodologies can never map the entire spectrum of opportunities, risks, and unintended consequences of AI systems, but may identify preferable alternatives, valuable courses of action, likely risks, and mitigating strategies. This has a dual advantage. As an opportunity strategy, foresight methodologies can help leverage ethical solutions. As a form of risk management, they can help prevent or mitigate costly mistakes, by avoiding decisions or actions that are ethically unacceptable. This will lower the opportunity costs of choices not made or options not

seized for lack of clarity or fear of backlash.

Ethical regulation of the design and use of AI is a complex but necessary task. The alternative may lead to devaluation of individual rights and social values, rejection of AI-based innovation, and ultimately a missed opportunity to use AI to improve individual well-being and social welfare. Humanity learned this lesson the hard way when it did not regulate the impact of the industrial revolution on labor forces, and also when it recognized too late the environmental impact of massive industrialization and global consumerism. It has taken a very long time, social unrest, and even revolutions to protect workers' rights and establish sustainability frameworks.

The AI revolution is equally significant, and humanity must not make the same mistake again. It is imperative to address new questions about the nature of post-AI societies and the values that should underpin the design, regulation, and use of AI in these societies. This is why initiatives like the above-mentioned AI4People and IEEE projects, the European Union (EU) strategy for AI, the EU Declaration of Cooperation on Artificial Intelligence, and the Partnership on Artificial Intelligence to Benefit People and Society are so important (see the supplementary materials for suggested further reading). A coordinated effort by civil society, politics, business, and academia will help to identify and pursue the best strategies to make AI a force for good and unlock its potential to foster human flourishing while respecting human dignity. ■

REFERENCES AND NOTES

1. J. McCarthy *et al.*, *AI Mag.* **27**, 12 (2006).
2. A. L. Samuel, *Science* **132**, 741 (1960).
3. G.-Z. Yang *et al.*, *Sci. Robot.* **3**, eaar7650 (2018).
4. N. Wiener, *Science* **131**, 1355 (1960).
5. M. Taddeo, L. Floridi, *Nature* **556**, 296 (2018).
6. D. Wang *et al.*, *arXiv:1606.05718 [q-bio.QM]* (18 June 2016).
7. P. Asaro, *Int. Rev. Red Cross* **94**, 687 (2012).
8. S. Russell, *Nature* **521**, 415 (2015).
9. J. Larson *et al.*, “How We Analyzed the COMPAS Recidivism Algorithm” (May 2016); www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.
10. L. Floridi, *Sci. Eng. Ethics* **19**, 727 (2013).
11. U. Pagallo, in *Human Law and Computer Law: Comparative Perspectives*, M. Hildebrandt, J. Gaakeer, Eds. (Springer, Netherlands, 2013), pp. 47–65.
12. L. Floridi, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **374**, 20160112 (2016).
13. H. Shirado, N. A. Christakis, *Nature* **545**, 370 (2017).
14. IEEE Standards Association, *Ethically Aligned Design*, Version 2.
15. J. Cowls, L. Floridi, Prolegomena to a White Paper on an Ethical Framework for a Good AI Society, Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3198732, 19 June 2018.

ACKNOWLEDGMENTS

M.T. and L.F. are members of the Partnership on Artificial Intelligence to Benefit People and Society; L.F. is also chair of the scientific committee of AI4People.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/751/suppl/DC1

10.1126/science.aat5991

How titanium dioxide cleans itself

Adsorption of hydrophobic molecules explains titanium dioxide self-cleaning properties

By Jeong Young Park

Titanium dioxide (TiO_2) is a white pigment that is widely used in paints, medicine, paper, sunscreens, and toothpaste. Other applications include photocatalysts and solar cell coatings (1). TiO_2 coatings appear to have self-cleaning properties: When exposed to sunlight, TiO_2 reacts with water to generate hydroxyl radicals, which break down organic molecules and microbes adsorbed on the surface. When it rains, water spreads out on the hydrophilic TiO_2 coating and washes away dust and dirt. On page 786 of this issue, Balajka *et al.* (2) use atomic-scale probes and spectroscopy to characterize the molecular-scale self-cleaning mechanism of TiO_2 surfaces.

TiO_2 can both attract and repel water. When TiO_2 surfaces in air are irradiated with ultraviolet (UV) light, they become hydrophilic, but they slowly return to a hydrophobic state in the dark (3). The UV-induced transition can be explained by the facile oxidation of adsorbates, resulting in the formation of hydroxylated species (4). During the reversal to the hydrophobic state in the dark, formation of an ordered (2×1) surface structure has been reported, but the chemical nature of this structure remains debated (5–7). Balajka *et al.* now report the selective adsorption of atmospheric carboxylic acids, which are hydrophobic, on a $\text{TiO}_2(110)$ surface in contact with water or air. This selective adsorption is responsible for the hydrophobic state found when TiO_2 is in the dark.

The study was challenging because of the conditions under which the surface chemis-

try occurs. Most surface science studies are carried out on well-defined single-crystal surfaces under ultrahigh vacuum (UHV) conditions. Yet, catalytically active systems are typically used at high pressure or involve solid-liquid interfaces. This gap between UHV and application conditions is referred to as the “pressure gap” (8, 9).

To tackle the pressure gap, Balajka *et al.* constructed a small chamber that allows the introduction of an ultrapure water drop or a controlled gas, followed by characterization in a vacuum. The authors carefully avoided

potential contamination during vacuum pumping—for example, by removing water and any additional gases with a cryopump cooled with liquid nitrogen. They then characterized the atomic and electronic structures of the TiO_2 surface with scanning tunneling microscopy (STM) and x-ray photoelectron spectroscopy, respectively.

The results show that exposing TiO_2 surfaces to water or ambient air leads to the formation of a well-ordered (2×1) monolayer of carboxylic acid. Further analysis of STM heights on (2×1) structures and their comparison with theoretical values indicates that the (2×1) structures are formate/acetate monolayers (see the figure).

Earlier studies have also reported a (2×1) structure under a range of environmental conditions and in different geographic locations (5–7), indicating that the (2×1) monolayer is ubiquitous. To confirm the

chemical nature of this structure, Balajka *et al.* carried out additional experiments, exposing $\text{TiO}_2(110)$ to air in Ithaca, New York, about 4300 miles from the previous measurement in Vienna. The data taken in Ithaca show the same chemical signature and the same height difference, indicating that the chemical nature of the adsorbates is the same. The authors confirm that the adsorbates from

exposure to water or air are ubiquitous environmental species with a high affinity for binding to the TiO_2 surface.

These results provide insight into the self-cleaning properties of the TiO_2 surface. The self-assembled carboxylate monolayers are hydrophobic as well as highly water soluble because they have a hydrophobic tail and an acidic head. Because of the high water solubility, the surface can transition to a hydrophilic state during rinsing, which enables the water sheeting action for self-cleaning. These results are particularly important for understanding the photocatalytic properties of TiO_2 because the self-assembled carboxylate monolayer blocks undercoordinated surface cation sites that can serve as effective electron traps, impeding photocatalytic activity.

Balajka *et al.*'s study sheds light on the molecular-scale mechanism of TiO_2 self-cleaning. However, the surface analysis in this study was carried out in UHV, giving rise to the next challenge: understanding the surface chemistry of TiO_2 under ambient air or liquid by using in situ surface techniques. This approach requires the preparation of TiO_2 model systems by using wet chemical methods, followed by in situ surface characterization of the solid-liquid interfaces, for example, with ambient pressure scanning probe microscopy, x-ray photoelectron spectroscopy, and vibration spectroscopy. Recent improvements to in situ surface techniques for investigating solid-liquid interfaces have led to substantial advances in the molecular-level understanding of electrochemical processes (10, 11). Atomic-scale characterization of TiO_2 surfaces under electrochemical and photocatalytic reactions could provide important data for improving solar cells and other devices of importance for energy and environmental applications. ■

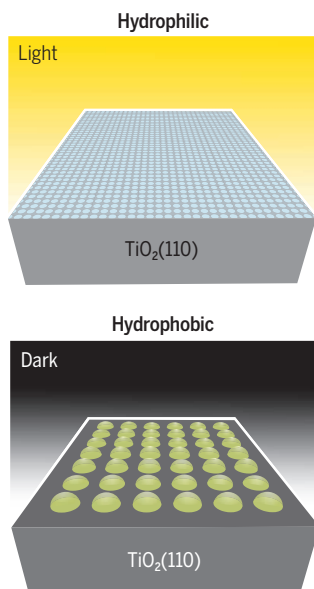
REFERENCES

1. A. Fujishima, K. Honda, *Nature* **238**, 37 (1972).
2. J. Balajka *et al.*, *Science* **361**, 786 (2018).
3. R. Wang *et al.*, *Nature* **388**, 431 (1997).
4. U. Diebold, *Surf. Sci. Rep.* **48**, 53 (2003).
5. H. Hussain *et al.*, *Nat. Mater.* **16**, 461 (2017).
6. G. Serrano *et al.*, *Adv. Mater. Interfaces* **2**, 1500246 (2015).
7. A. Song *et al.*, *J. Phys. Chem. C* **120**, 9326 (2016).
8. J. Y. Park, *Current Trends of Surface Science and Catalysis* (Springer, 2014).
9. G. A. Somorjai, J. Y. Park, *Chem. Soc. Rev.* **37**, 2155 (2008).
10. J.-J. Velasco-Velez *et al.*, *Science* **346**, 831 (2014).
11. D. E. Starr *et al.*, *Chem. Soc. Rev.* **42**, 5833 (2013).

10.1126/science.aau6016

Two main states of TiO_2 in an ambient environment

UV light decomposes organic molecules (such as carboxylic acid) on the TiO_2 surface, making it hydrophilic. In the dark, carboxylic acid assembles into an organized monolayer, making the surface hydrophobic.



Center for Nanomaterials and Chemical Reactions, Institute for Basic Science, and Graduate School of EWS and Department of Chemistry, KAIST, Daejeon 305-701, Republic of Korea. Email: jeongypark@kaist.ac.kr

DEVELOPMENT

Orchestrating cortical brain development

Signaling in the developing vasculature coordinates central nervous system morphology

By Jean-Leon Thomas^{1,2}

The developing central nervous system (CNS) acquires its own vascular network via ingression of blood vessels as the neural tissue expands. The relationship between the vasculature and the stromal (glial) and neuronal cell components has been investigated in different regions of the developing CNS (1). On page 767 of this issue, Segarra *et al.* (2) explore how signaling at the interface of neurons, endothelial cells (which line the vasculature), and glial cells is integrated for proper brain development. This provides important mechanistic understanding of the cross-talk between the developing CNS and endothelial cells whereby the vasculature does much more than deliver oxygen and nutrients.

Several ligand-receptor signaling systems are involved in neurovascular interactions, most commonly with ligands secreted by neural cells and corresponding receptors expressed by endothelial cells. The vascular endothelial growth factor A (VEGFA)-VEGF receptor 2 (VEGFR2) and neuropilin 1 pathway, the Wnt-Frizzled receptor- β -catenin signaling pathway, and G protein-coupled receptor 124 (GPR124) are regulators of CNS angiogenesis (3). Inactivating mutations in genes encoding members of these pathways impair CNS angiogenesis and lead to a reduction in CNS tissue growth and viability, likely due to inadequate delivery of oxygen and nutrients to developing neural tissues. However, endothelial cells can also instruct CNS development and provide morphogenic cues to neural progenitors and their daughter cells.

The vasculature contributes to neuronal navigation and positioning during development, through secretion of the inhibitory neurotransmitter γ -aminobutyric acid (GABA), which guides the tangential migration of GABAergic neurons in the embryonic forebrain (4), or through secretion of stromal cell-derived factor-1 (SDF-1), which controls the migration of Cajal-Retzius cells at the surface of the cerebral cortex (5). Cajal-Retzius cells are pioneer neurons that are critical for

the development of the cerebral cortex. They secrete the extracellular glycoprotein reelin (RELN), which plays a crucial role in neuronal migration and neocortical lamination. Inactivating mutations in *RELN* are associated with brain malformations and mental retardation in humans (6).

RELN is a ligand of apolipoprotein E receptor 2 (APOER2) and very low-density lipoprotein receptor (VLDLR), which leads to phosphorylation of the adaptor protein disabled homolog 1 (DAB1) (7). Segarra *et al.*

In agreement with the well-known role of RGCs functioning as a scaffold for cortical neuron migration (8), *Dab1*^{ΔEC} mice display abnormal lamination of the cortical layers. This phenotype partially recapitulates the cortical defects of mice with inactivation of RELN and mice with embryonic neuronal deletion of *Dab1* and also correlates with the cortical abnormalities observed in patients with RELN mutations. Thus, endothelial cells instruct RGCs for proper positioning of cortical neurons during development (see the figure).

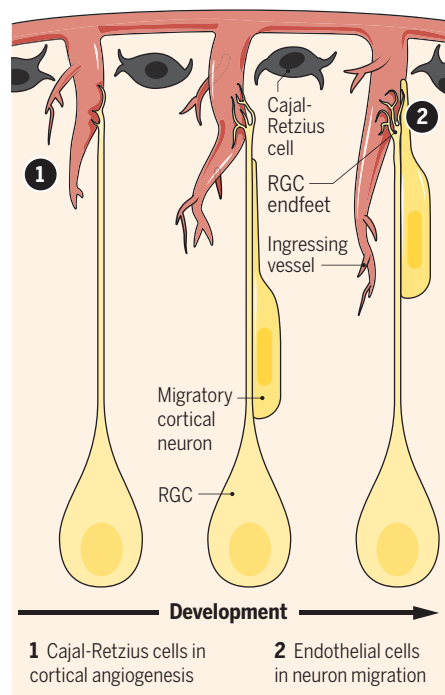
Later in development and in adulthood, endothelial cells stabilize neurovascular units that they form with astrocytes (a type of glial cell that is derived from RGC differentiation) to establish the blood-brain barrier (BBB). The BBB ensures CNS homeostasis by strictly controlling fluid and cell extravasation into the brain parenchyma (9). Segarra *et al.* showed that communication between endothelial cells and RGCs or astrocytes was reinforced by RELN-DAB1-mediated enrichment of laminin $\alpha 4$ (LAMA4) in the gliovascular basal lamina between these cells. LAMA4 is an extracellular matrix adhesion protein that stimulates RGC and astroglial cell attachment to the gliovascular basal lamina through integrin $\beta 1$ receptor signaling. This confirms the key role of laminins in stabilizing cell-to-cell interactions (10). *Dab1*^{ΔEC} mice exhibit altered BBB integrity, resulting in fluid extravasation from vessels, which persists in the adult.

The study of Segarra *et al.* reveals a dual and temporally regulated function of RELN-DAB1 signaling in CNS endothelial cells that affects both cortical neuron patterning and BBB etiology during development. Future work may address the distinct mechanisms of RELN signaling that are required for VEGFA-dependent angiogenesis, and whether RELN signaling eventually integrates with other CNS-specific endothelial cell regulators involved in CNS development and BBB formation, such as Wnt signaling and major facilitator domain containing protein 2A (MFSD2A) (11).

The findings of Segarra *et al.* exemplify two concepts. A molecular mechanism (RELN-DAB1-LAMA4 signaling) for how the vasculature orchestrates organogenesis is proposed. They also show that the extracellular matrix organizes tissue morpho-

The developing cortex

Cajal-Retzius cells promote endothelial cell proliferation and angiogenesis, which stimulates RGC attachment to the gliovascular basal lamina to ensure proper cortical neuron migration.



show that RELN-APOER2-DAB1 signaling in endothelial cells participates in the control of both VEGFA-VEGFR2-driven angiogenesis in the CNS and endothelial cell-radial glial cell (RGC) communication. They found a reduction in the number of the endfeet of RGC fibers around vessels located at the periphery of the developing cortex in mice with embryonic deletion of *Dab1* in endothelial cells (*Dab1*^{ΔEC} mice).

¹Department of Neurology, Yale University School of Medicine, New Haven, CT 06511, USA. ²Université Pierre et Marie Curie Paris 06 UMR51127, Sorbonne Université, Institut du Cerveau et de la Moelle Epinière, Paris, France. Email: jean-leon.thomas@yale.edu

genesis. Furthermore, the role for blood vessels ingressing the embryonic cerebral cortex is reminiscent of the role of the vasculature in cell specification and differentiation in other organs such as the liver and bones (12).

The secreted factors from endothelial cells, called angiocrine factors, instruct organ-specific cells especially by introducing new information within the extracellular matrix (13). Basement membranes (or basal lamina) are communication interfaces because they integrate molecules derived from vascular and organ-specific compartments and serve as ligand-presenting surfaces to receptors expressed by the cells in these compartments. At this interface, endothelial cells provide the instructions for organ development.

The discovery that endothelial cells are required for the scaffolding of RGC processes

“...the vasculature does much more than deliver oxygen and nutrients.”

could help improve the structure of brain organoids. These self-organizing structures derived from mouse or human neural stem cells are cultured in vitro to model brain development and disease (14). Brain organoids develop some aspects of cortical layering, but RGC scaffolding and proper cortical lamination have not been established. Thus, brain organoids cannot be used to study late stages of brain development or neurodegeneration. Cortical organization could be achieved by incorporation of a vascular network or overexpression of LAMA4.

DAB1 in endothelial cells could be a target to reinforce BBB integrity that is altered in pathological conditions such as infectious and inflammatory CNS diseases (15). However, additional studies are needed to determine the therapeutic potential of DAB1 in humans. ■

REFERENCES

1. C. Ruhrberg, V. L. Bautch, *Cell. Mol. Life Sci.* **70**, 1675 (2013).
2. M. Segarra et al., *Science* **361**, eaao2861 (2018).
3. A. Eichmann, J.-L. Thomas, *Cold Spring Harb. Perspect. Med.* **3**, a006551 (2013).
4. C. Won et al., *Nat. Commun.* **4**, 2149 (2013).
5. V. Borrelli, O. Marin, *Nat. Neurosci.* **9**, 1284 (2006).
6. S. E. Hong et al., *Nat. Genet.* **26**, 93 (2000).
7. H. H. Bock, P. May, *Front. Cell. Neurosci.* **10**, 166 (2016).
8. P. Rakic, *J. Comp. Neurol.* **145**, 61 (1972).
9. Z. Zhao et al., *Cell* **163**, 1064 (2015).
10. R. O. Hynes, *Nat. Rev. Mol. Cell Biol.* **15**, 761 (2014).
11. B. J. Andreone et al., *Annu. Rev. Neurosci.* **38**, 25 (2015).
12. S. Rafii et al., *Nature* **529**, 316 (2016).
13. H. G. Augustin, G. Y. Koh, *Science* **357**, eaal2379 (2017).
14. M. Heide et al., *Curr. Opin. Cell Biol.* **55**, 8 (2018).
15. A. Iwasaki, *Trends Mol. Med.* **23**, 227 (2017).

MICROBIOLOGY

Enterovirus outbreak dynamics

Predictability of outbreaks opens the door to model-guided public health planning

By Birgit Nikolay^{1,2,3} and Simon Cauchemez^{1,2,3}

Outbreaks of pathogens that cause acute immunizing infections are often highly predictable. The most studied example is measles, for which case incidence over time is robustly explained with simple mathematical models that account for variations in the number of susceptible individuals through infection and birth (1). By contrast, it is more challenging to predict outbreaks of infectious diseases that exhibit complex patterns of immunity, such as influenza, for which antigenic characteristics of circulating strains continuously change (2). Enteroviruses can cause a wide spectrum of clinical manifestations—including hand-foot-and-mouth disease (HFMD)—with potentially severe neurological complications (3). With more than 100 serotypes that may have varying immunological cross-protection (3), predicting the transmission dynamics of enteroviruses was expected to be difficult. On page 800 of this issue, Pons-Salort and Grassly (4) demonstrate that in contrast to this expectation, enteroviruses are highly predictable pathogens, with outbreaks largely driven by serotype-specific long-

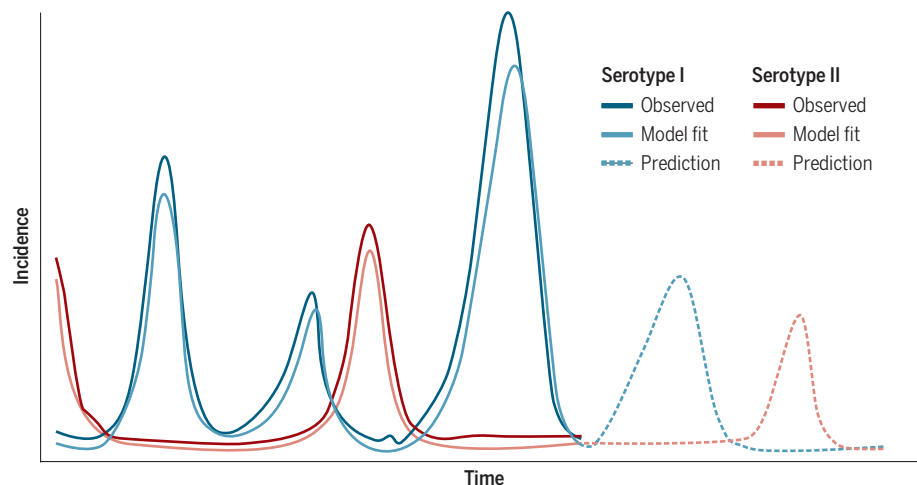
term immunity and birth rates. This opens the door to model-guided public health planning and outbreak preparedness.

Although it was known that simple mathematical models could reproduce epidemic patterns of two important enteroviruses (Enterovirus-A71 and Coxsackievirus-A16, the main causes of HFMD) (5), the study of Pons-Salort and Grassly demonstrates that such models are applicable to a wide range of enterovirus serotypes. They were able to explain infection dynamics for 18 of the 20 most prevalent serotypes in Japan with a single simple model. They also used this framework to investigate mechanisms that might have driven atypical incidence patterns observed for the two remaining serotypes.

The analysis indicates that predicting outbreaks for a specific enterovirus serotype may be possible without having to account for cross-protection conferred by other circulating serotypes. In the future, models could be used to anticipate when the next enterovirus outbreak might occur, how large it will be, and which serotype might cause it (see the figure). Because clinical form and severity of illness depend on the serotype (3), this information could help health care providers to better prepare for the detection of cases and the provision of appropriate care. Two Enterovirus-A71

Model-guided public health planning for enterovirus outbreaks

Based on enterovirus surveillance data, mathematical transmission models can be developed to explain serotype-specific incidence patterns. These models can be used to predict when the next enterovirus outbreak might occur and which serotype might be causing it, allowing for better outbreak preparedness.



10.1126/science.aau7155

vaccines have been licensed in China, and vaccines against other enteroviruses are being developed (6). Mathematical models could be used to assess the potential impact of vaccine introduction and to determine optimal vaccination strategies. Similar analyses were performed for measles virus and helped identify the optimal age for routine vaccination and assess how supplementary vaccination campaigns should be implemented in order to prevent epidemics (7–9). The analysis of Pons-Salort and Grassly is a step toward the development of a model-based approach to support and improve the public health response to enterovirus outbreaks.

A number of countries may want to implement such a powerful approach, but only those with good enterovirus surveillance will really benefit because the availability of detailed surveillance data is a prerequisite for accurate modeling. Besides Japan, few countries perform systematic serotyping of isolated enterovirus strains because of laboratory constraints. The potential to predict outbreaks should incentivize countries to improve enterovirus surveillance. We believe that the model presented by Pons-Salort and Grassly may help to achieve this objective by making it possible to derive the absolute number of infections and the associated reporting rate of the surveillance system, which are difficult to obtain otherwise. By comparing reporting rates by location, these models can be used to identify areas where surveillance needs to be strengthened. Moreover, analysis of reporting rates of the different serotypes should provide insight about their relative pathogenicity. Discrepancies between these estimates and the published literature could help refine surveillance to ensure that all key serotypes are correctly captured.

The study generated a number of hypotheses that are biologically or epidemiologically testable and should stimulate more research. For example, the analysis suggested that characteristics of two serotypes (pathogenicity of Coxsackievirus-A6 and antigenicity or transmissibility of Echovirus 18) changed at specific time points in the past. Genomic studies could help to further investigate these hypotheses. Moreover, the study suggested that although enterovirus immunity was in general of long duration (from 8 years to lifelong), it decreased over time at rates that varied by serotype. Cohort studies may clarify the dynamics of enter-

ovirus immune responses. Last, estimated differences in the transmissibility of serotypes can be validated by comparing the average age of detected cases, whereby higher transmissibility is indicated by a younger age of infection (10) and can be further investigated in transmission studies.

More developments are expected to consolidate the work by Pons-Salort and Grassly. Although cross-immunity between serotypes was not necessary to explain long-term trends in disease incidence, a formal assessment of the potential impact of short-lived cross-protection on enterovirus transmission dynamics [as reported for HFMD in China (5)] should be made. Even limited short-lived cross-protection may have an impact on overall dynamics in a disease system in which many serotypes cocirculate. This seems important to further validate the conclusion that vaccine introduction is unlikely to result in any form of serotype replacement, whereby the incidence of non-vaccine serotypes increases. Transmission models accounting for age or spatial structures, as previously developed for measles and other childhood diseases (7, 11), may more accurately reflect the transmission dynamics of enteroviruses and should be considered to assess vaccination strategies. Although the model satisfyingly explained enterovirus incidence patterns in Japan, it will be important to investigate whether it is equally successful when applied to data from outside Japan (5, 12, 13). Measles epidemic dynamics, for example, were considerably different between Europe and Africa but could be explained with the same model once differences in seasonality and birth rates were accounted for (11). Would the same be true for enteroviruses? Or are there other features, beyond birth rates and seasonality, that would need to be accounted for? Nonetheless, reducing complex incidence patterns to simple mathematical models that are applicable to a wide range of enterovirus serotypes represents a milestone in advancing our understanding of enterovirus transmission dynamics. ■

REFERENCES

1. D. J. Earn *et al.*, *Science* **287**, 667 (2000).
2. J. M. Fonville *et al.*, *Science* **346**, 996 (2014).
3. M. Pons-Salort *et al.*, *Curr. Opin. Infect. Dis.* **28**, 479 (2015).
4. M. Pons-Salort, N. C. Grassly, *Science* **361**, 800 (2018).
5. S. Takahashi *et al.*, *PLOS Med.* **13**, e1001958 (2016).
6. Q. Mao *et al.*, *Emerg. Microbes Infect.* **5**, e75 (2016).
7. C. J. E. Metcalf *et al.*, *Epidemiol. Infect.* **139**, 265 (2011).
8. J. Lessler *et al.*, *Epidemiol. Infect.* **139**, 1039 (2011).
9. J. M. Prada *et al.*, *Vaccine* **35**, 1488 (2017).
10. R. M. Anderson, R. M. May, *Infectious Diseases of Humans—Dynamics and Control* (Oxford Univ. Press, 1992).
11. M. J. Ferrari *et al.*, *Nature* **451**, 679 (2008).
12. Y. Podin *et al.*, *BMC Public Health* **6**, 180 (2006).
13. N. Khetsuriani *et al.*, *MMWR Surveill. Summ.* **55**, 1 (2006).

¹Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France. ²CNRS UMR2000: Génomique Evolutive, Modélisation et Santé, Institut Pasteur, Paris, France. ³Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, Paris, France. Email: birgit.nikolay@pasteur.fr; simon.cauchemez@pasteur.fr

IMMUNOLOGY

Connecting immune deficiency and inflammation

RIPK1 mutations result in immune deficiency and autoinflammation in humans

By Manolis Pasparakis¹ and Michelle Kelliher²

Advances in genomic technologies have revealed the genetic basis of an increasing list of human primary immune deficiencies (PID) as well as autoinflammatory diseases. The identification of a subset of patients suffering from recurrent infections combined with inflammatory diseases revealed a previously unappreciated connection between immunodeficiency and autoinflammation. The paradoxical combination of immune deficiency and increased inflammation was reported in patients who have mutations that perturb signaling to the inflammatory transcription factor, nuclear

“Although differences exist, RIPK1 functions as a critical regulator of immunity and inflammation in both mice and humans.”

factor- κ B (NF- κ B), including mutations in inhibitor of NF- κ B (I κ B α), NF- κ B essential modulator (NEMO), and in components of the linear ubiquitin chain assembly complex [LUBAC, heme-oxidized IRP2 ubiquitin ligase 1 (HOIL-1) and HOIL-1-interacting protein (HOIP)] (1–3). On page 810 of this issue, Cuchet-Lourenço *et al.* (4) describe four pediatric patients with PID and inflammatory disease that harbor loss-of-function mutations in the receptor-interacting serine/threonine-protein kinase 1 (RIPK1) gene. RIPK1 is involved in signal transduction to NF- κ B and is also a critical regulator of cell death, providing further evidence linking NF- κ B signaling

with immune deficiency, cell death, and inflammation.

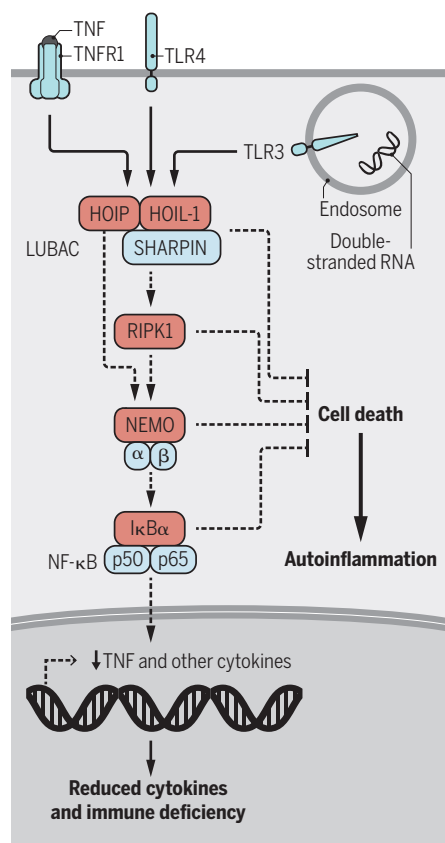
RIPK1 regulates proinflammatory and cell death signaling following activation of tumor necrosis factor receptor (TNFR), Toll-like receptor 3 (TLR3), and TLR4. RIPK1 acts as a scaffold, transducing signals to NF- κ B and to p38, c-Jun N-terminal kinase (JNK) and extracellular signal-regulated kinase (ERK), to drive proinflammatory and prosurvival gene expression. RIPK1 kinase activity induces apoptosis or necroptosis, which is an inflammatory form of cell death. Necroptotic cells release danger-associated molecular patterns (DAMPs) that activate immune responses in bystander cells, thereby inducing a feed-forward cycle of inflammation and cell death. Much of our knowledge about RIPK1 function comes from studies in mice—its role in humans has been extrapolated from these studies.

Cuchet-Lourenço *et al.*, describe four children with homozygous *RIPK1* mutations from three independent families. The RIPK1 protein was not detected in cells from the three patients examined, indicating that these homozygous mutations result in a complete RIPK1 deficiency. In contrast to RIPK1-deficient mice, which die shortly after birth (5), these patients survived for 3 to 13 years, indicating that in humans RIPK1 is not essential for survival. This is reminiscent of the findings that HOIP- and HOIL-1-deficiencies are lethal in mice but not in humans (2, 3, 6, 7). It is unclear to what extent the survival of humans lacking these proteins reflects fundamental differences in physiology or relates to genetic differences between inbred mouse strains and the highly heterogeneous human population. Studying the effect of these gene deficiencies in outbred mice, which are more genetically heterogeneous, should help resolve this question.

The RIPK1-deficient children suffered from recurrent viral, bacterial, and fungal infections, indicating severe immune deficiency similar to the broad infections observed in patients with inactivating mutations in the genes encoding NEMO, I κ B α , HOIP, or HOIL-1 (1–3). The RIPK1-deficient patients also developed early-onset inflammatory bowel disease (IBD) and progressive arthritis. These findings and further analysis reveal that RIPK1 regulates proinflammatory signaling by TNFR

Mutations in the NF- κ B pathway

In patients with HOIL-1, HOIP, NEMO, I κ B α , or RIPK1 deficiencies, cytokine production is impaired (dashed arrows), resulting in immune deficiency. Defects in this pathway also sensitize cells to inflammatory cell death, resulting in autoinflammation.



and TLR3 and suggest that immune deficiency in these patients results from reduced cytokine production in the absence of RIPK1. Impaired TLR3 signaling could also contribute to the susceptibility to viral infections in these patients.

The development of gastrointestinal inflammation in the RIPK1-deficient patients is reminiscent of the severe gut pathology of mice with an intestinal epithelial cell RIPK1-deficiency (*Ripk1^{IEC KO}* mice) (8, 9). The gut pathology in these mice was caused by intestinal epithelial cell apoptosis mediated largely by TNF. There was, however, no evidence of extensive intestinal cell death in gastrointestinal biopsies from the patients. However, they were regularly treated with anti-inflammatory drugs and antibiotics, which could suppress TNF production and subsequent intestinal epithelial cell apoptosis, as seen in *Ripk1^{IEC KO}* mice (9).

Although three of the patients developed IBD during the first months of life, the fourth patient developed intestinal inflammation at 4 years, suggesting that en-

vironmental factors such as the intestinal microbiota contribute to IBD onset. Interestingly, hematopoietic stem cell transplantation (HSCT) resolved the intestinal inflammation in one patient, suggesting that RIPK1-deficiency in hematopoietic cells contributes to intestinal inflammation in humans. Yet, mice in which *Ripk1* is deleted in the hematopoietic system do not develop intestinal inflammation or arthritis, potentially because these mice succumb to bone marrow failure in 30 days (10). HSCT was not successful in two patients and they died from multi-organ failure or viral infection. Thus, it may be premature to draw conclusions about whether RIPK1 primarily acts in hematopoietic and/or epithelial cells to regulate immune responses in humans. Interestingly, the patients did not develop skin inflammation, unlike *Ripk1*-deficient mice (8, 11).

Analysis of monocytes from one of the patients revealed increased amounts of the cytokine interleukin-1 β (IL-1 β) in response to immune stimulation. Monocyte necroptosis may promote inflammation by inducing the release of IL-1 β and other DAMPs, which could contribute to IBD and arthritis. However, mice with RIPK1-deficient hematopoietic cells exhibit elevated serum cytokines and chemokines, but increased serum IL-1 β was not observed (10). These mice are maintained under pathogen-free conditions, which limits exposure to infectious agents and inflammatory molecules, and likely accounts for the absence of recurrent infections and elevated IL-1 β .

Although differences exist, RIPK1 functions as a critical regulator of immunity and inflammation in both mice and humans. The study of Cuchet-Lourenço *et al.* adds RIPK1 deficiency to the list of genetic defects causing both PID and autoinflammation (see the figure). Future studies in humans and mice are needed to better understand the mechanisms balancing inflammatory and cell death signaling in order to harness this knowledge for the development of better treatments for infectious and inflammatory diseases. ■

REFERENCES

1. C. Picard *et al.*, *Clin. Microbiol. Rev.* **24**, 490 (2011).
2. B. Boisson *et al.*, *Nat. Immunol.* **13**, 1178 (2012).
3. B. Boisson *et al.*, *J. Exp. Med.* **212**, 939 (2015).
4. D. Cuchet-Lourenço *et al.*, *Science* **361**, 810 (2018).
5. M. A. Kelliher *et al.*, *Immunity* **8**, 297 (1998).
6. N. Peltzer *et al.*, *Cell Rep.* **9**, 153 (2014).
7. N. Peltzer *et al.*, *Nature* **557**, 112 (2018).
8. M. Dannappel *et al.*, *Nature* **513**, 90 (2014).
9. N. Takahashi *et al.*, *Nature* **513**, 95 (2014).
10. J. E. Roderick *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14436 (2014).
11. J. A. Rickard *et al.*, *Cell* **157**, 1175 (2014).

¹Institute for Genetics, Centre for Molecular Medicine (CMC), and Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, 50931 Cologne, Germany. ²Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, MA 01605, USA. Email: pasparakis@uni-koeln.de; michelle.kelliher@umassmed.edu

BATTERIES

Hot lithium-oxygen batteries charge ahead

Molten salt electrolytes and nickel oxide-based electrodes enable four-electron transfer

By Shuting Feng,^{1,2} Jaclyn R. Lunger,^{1,3}
Jeremiah A. Johnson,⁴ Yang Shao-Horn^{1,3}

The need to increase the energy storage per unit mass or volume and to decrease stored-energy cost from solar and wind (1) has motivated research efforts toward developing alternative battery chemistries. In particular, lithium-oxygen (Li-O₂) batteries offer great promise (2, 3). During discharge, oxygen can be reduced to form either peroxide (Li₂O₂ in a two-electron pathway) or oxide (Li₂O in a four-electron pathway). The estimated energy densities of lithium-oxygen batteries based on peroxide and oxide are two and four times higher than that of lithium-ion batteries, respectively (3), but degradation of organic electrolytes and of oxygen electrodes (typically made of carbon) by these reactive oxygen species has limited the reversibility of these systems. On page 777 of this issue, Xia *et al.* (4) address these issues by using inorganic components—a molten salt electrolyte and a nickel-based oxide supported by stainless steel mesh for the oxygen electrode—and demonstrate reversible operation for the four-electron-pathway Li-O₂ battery at 150°C.

The reversibility of Li-O₂ chemistry is hampered by the (electro)chemical instability of carbon electrodes (5) and of nonaqueous electrolytes (6–8) such as carbonates (9) used in lithium-ion batteries, as well as ethers (10) and sulfoxide (11, 12), which not only limit battery cycling lifetime but also battery voltage and faradaic efficiency (3, 8, 13). For example, the voltage hysteresis (the voltage gap between charge and discharge) of typical Li-O₂ batteries based on peroxide are on the order of ~1 V (3), corresponding to ~30% energy loss. One solution to this instability issue is to use a molten salt elec-

trolyte, such as a LiNO₃-KNO₃ eutectic, operated at moderate temperatures (~150°C). Giordani *et al.* (14) demonstrated a reversible Li-O₂ battery based on the two-electron pathway that can be cycled with a small voltage hysteresis of ~0.1 V for 50 times with a carbon-based oxygen electrode and a LiNO₃-KNO₃ electrolyte; this electrolyte was adopted by Xia *et al.*

Charging oxide in such batteries follows a direct four-electron pathway, as suggested by Xia *et al.* on the basis of online

This issue also highlights the opportunities to develop catalysts that could improve oxide formation kinetics and selectivity, which can be rate limiting. Moreover, a rational design of stability for the catalyst and electrode surfaces against peroxide and oxide is needed to achieve long-term cycling performance.

Although the impressive cycle life demonstrated in this work has been attributed in part to the lower reactivity of oxide than peroxide toward the electrode and elec-

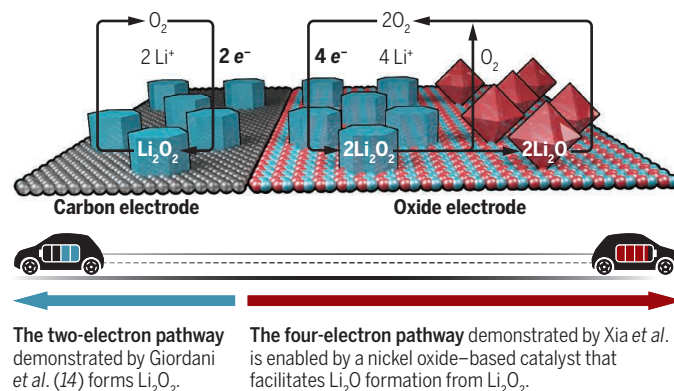
trolyte, systematic studies are needed to support the claim. Furthermore, to take full advantage of gravimetric energy gain by using Li-O₂ chemistry requires the development of light, transition metal-free, electronically and ionically conducting porous structures to host the catalyst and enable reversible oxidation and plating of lithium metal by stabilizing the lithium-electrolyte interface during cycling.

Xia *et al.* demonstrate the reversible operation of a Li-O₂ battery at 150°C via the four-electron pathway using a molten salt electrolyte and a nickel oxide-based oxygen electrode.

This discovery highlights immense opportunities in fundamental research to understand oxygen redox reactions from a chemical physics point of view and principle-guided designs for electrode materials as well as system-level optimization to enable new battery technologies that can potentially rival lithium-ion batteries and other storage technologies. ■

Two routes to lithium-oxygen batteries

Two types of rechargeable lithium-oxygen battery chemistries operate at high temperatures (150°C) and take advantage of molten salt electrolytes.



The two-electron pathway demonstrated by Giordani *et al.* (14) forms Li₂O₂.

The four-electron pathway demonstrated by Xia *et al.* is enabled by a nickel oxide-based catalyst that facilitates Li₂O formation from Li₂O₂.

electrochemical mass spectrometry measurements (see the figure). Remarkably, the reversible four-electron O₂ reduction and evolution could proceed with a small voltage hysteresis of ~0.3 V, considerably less than that of regenerative proton-exchange membrane fuel cells (~0.7 V), for 150 cycles. Also, the coulombic efficiency—the ratio of discharge to charge capacity—approached 100%. The authors have proposed that the electrochemical O₂ reduction gives rise to peroxide, which then disproportionates to oxide via a process that can be catalyzed by nickel-based oxide but not carbon, resulting in an apparent four-electron reduction during battery discharge.

Little is known about the reaction mechanism for the oxygen electrode for such environments. For example, the proposed nickel oxide-catalyzed oxide formation from peroxide upon discharge and the direct four-electron oxidation of oxide upon charge require atomistic understanding.

REFERENCES

1. B. Nykvist, M. Nilsson, *Nat. Clim. Chang.* **5**, 329 (2015).
2. J. W. Choi, D. Aurbach, *Nat. Rev. Mater.* **1**, 16013 (2016).
3. Y.-C. Lu *et al.*, *Energy Environ. Sci.* **6**, 750 (2013).
4. C. Xia, C. Y. Kwok, L. F. Nazar, *Science* **361**, 777 (2018).
5. D. M. Itkis *et al.*, *Nano Lett.* **13**, 4697 (2013).
6. S. Feng *et al.*, *J. Mater. Chem. A* **5**, 23987 (2017).
7. M. Huang *et al.*, *Energy Environ. Sci.* **11**, 1326 (2018).
8. K. Xu, *Chem. Rev.* **114**, 11503 (2014).
9. S. A. Freunberger *et al.*, *J. Am. Chem. Soc.* **133**, 8040 (2011).
10. B. D. McCloskey *et al.*, *J. Phys. Chem. Lett.* **3**, 3043 (2012).
11. N. Mozshukhina, L. P. Méndez De Leo, E. J. Calvo, *J. Phys. Chem. C* **117**, 18375 (2013).
12. D. G. Kwabi *et al.*, *J. Phys. Chem. Lett.* **5**, 2850 (2014).
13. J. Lu *et al.*, *Chem. Rev.* **114**, 5611 (2014).
14. V. Giordani *et al.*, *J. Am. Chem. Soc.* **138**, 2656 (2016).

10.1126/science.aau4792

¹Electrochemical Energy Laboratory, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. ²Department of Chemical Engineering, MIT, Cambridge, MA 02139, USA.

³Department of Materials Science and Engineering, MIT, Cambridge, MA 02139, USA. ⁴Department of Chemistry, MIT, Cambridge, MA 02139, USA. Email: shaohorn@mit.edu



Blue light emitted by electronics can disrupt circadian rhythms and compromise sleep.

BOOKS *et al.*

PHYSIOLOGY

Putting sleep myths to bed

From groggy teenagers to fatal narcolepsy, two tomes tackle the science of slumber

By **Adrian Woolfson**

Many of us have an ambivalent relationship with sleep. Although acknowledging its necessity, we begrudge these stolen hours of existence. Two new books provide a fresh perspective on this poorly understood phenomenon. Rather than being an “imperfection of our nature,” as extolled by physician Wilson Phillip in 1833, sleep emerges as critical to healthy bodily, mental, and emotional function.

In a charming analogy, Henry Nicholls in his book *Sleepyhead* likens the architecture of a perfect night's sleep to the aesthetic perfection of the Taj Mahal. When undermined, it's as though “the architectural plans have gone missing, and the resulting edifice is simply a miscellaneous jumble of marble.” Whereas Alice Gregory takes a developmental perspective of sleep in *Nodding Off*, charting its idiosyncrasies across the lifespan, Nicholls tackles the subject through the lens of pathology, being himself a lifetime sufferer of the sleep disorder narcolepsy.

In humans, the master timekeeper underwriting the architecture of sleep is a small region buried deep within the brain, known as the suprachiasmatic nucleus. The pacemaker

activity of this biological timepiece is controlled by several genes, which have names like *Period*, *Clock*, and *Timeless*. Mutations in these can transform us from night owls into morning larks or something in between. Fortunately, errant wanderings are typically adjusted by “zeitgebers”—literally, “time-givers”—principally in the form of blue light.

Nicholls meticulously details his experience of narcolepsy, from the first symptoms of excessive sleepiness at inappropriate moments, to the paroxysms of cataplexy that result in sudden, momentary paralysis. But he comes to realize that the reach of this disease is far deeper than he had ever realized, affecting his mood, dreaming, motivation, and body weight; causing hallucinations; and paradoxically resulting in disturbed nocturnal sleep. Along the way, he covers some fascinating material, from the rare prion-driven

disease known as fatal familial narcolepsy to the potentially atavistic nature of cataplexy.

In *Nodding off*, Gregory makes the point that although sleep studies typically focus on individuals, sleep is not always a solitary pastime and may be adversely affected by our choice of bedmate. Indeed, sleep researchers are now attempting to address the reality that adults often do not sleep in isolation.

The structure of sleep also changes across an individual's lifetime. The sleep patterns of young children, for example, are profoundly influenced by their belief systems. Gregory cautions parents against sending misbehaving children to bed early because this association between sleep and punishment may inadvertently condemn them to years of insomnia.

Similarly, she explains why the Sisyphean struggle to force teenagers to wake up early is invariably destined to fail. Their pattern of melatonin release differs to those of adults and children, and the recapitulation of this phenomenon in other mammals suggests that it has been hard-wired by evolution.

While extolling the virtues of sleep and its fundamental importance to our health, Gregory reveals some interesting tidbits, including the fact that dolphins sleep with just half of their brain at a time and that male armadillos have erections during non-REM (rapid eye movement) sleep, unlike their human counterparts, who experience this phenomenon only during REM sleep. The heterogeneity of sleep across different species indicates its essential function but also how it may be modified to perform different functions.

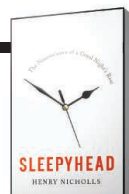
Although the precise function of sleep remains enigmatic, poor-quality sleep and sleep deprivation may have a profound impact on our health. In Denmark, Gregory reveals, the government has gone so far as to compensate long-term shift workers that develop breast cancer after mouse studies revealed the pro-carcinogenic effects of altered sleep.

It is noteworthy, and apparently contrary to the central thesis of these two tomes, that genius has sometimes emerged within the context of abnormal sleep patterns. The serial micronapper Leonardo da Vinci's remarkable canon of work, for example, was achieved on a sleeping pattern comprising naps of just 15 minutes taken every 4 hours. But it was, perhaps, Salvador Dali who, in tapping into the “secret of sleeping while awake,” demonstrated most idiosyncratically how unconventional sleeping habits can extract creative insights from this enigmatic shadowland. ■

Sleepyhead

The Neuroscience of a Good Night's Rest
Henry Nicholls

Basic Books, 2018. 368 pp.



Nodding Off

The Science of Sleep from Cradle to Grave

Alice Gregory
Bloomsbury Sigma, 2018.
304 pp.



The reviewer is the author of *Life Without Genes* (Flamingo, 2000). Email: adrianwoolfson@yahoo.com

SCIENCE COMMUNICATION

Changing the conversation

A pair of philosophers offer a new perspective on problematic dissent in science

By **Erin J. Nash**

Within Inmaculada de Melo-Martín and Kristen Intemann's *The Fight Against Doubt* is a vital observation: "As important as science is for sound public policy...a focus on the science can take us only so far."

I wholeheartedly agree. Since changing careers from science and public policy-making to philosophy, I've been on a quest to answer the question at the center of this book: How should we go about limiting the damage that can be done by problematic scientific dissent?

Scientific dissent—the act of challenging a widely held scientific position—often facilitates scientific and social progress. Recall, for example, that only decades ago the medical consensus was that homosexuality was a disorder. Without dissenting voices, homosexuality may never have been removed from the *Diagnostic and Statistical Manual of Mental Disorders*. But dissent can also be problematic, as Robert Proctor, Naomi Oreskes, and others have masterfully demonstrated (1, 2).

Many scientists and scholars take the very presence and circulation of misinformation to be a substantial part of the problem. However, philosophers de Melo-Martín and Intemann think that too much attention has been given to the proliferation of erroneous claims and not enough to the conditions within which dissent arises and takes hold.

In *The Fight Against Doubt*, the authors argue that problematic dissent is likely to be more damaging when public trust in scientists falters and when we fail to recognize the limits of scientific evidence in policy debates. They advocate moving the debate away from dissenting speech and addressing the ways in which research institutions and practices provide the public with good reasons to question the trustworthiness of scientists.

In the first half of the book, de Melo-Martín and Intemann survey and reject the

The reviewer is at the Practical Justice Initiative, Department of Philosophy, University of New South Wales, Sydney, NSW 2052, Australia and the Centre for Humanities Engaging Science & Society, Department of Philosophy, Durham University, Durham DH1, UK. Email: e.nash@unsw.edu.au

criteria established by other scholars for problematic dissent, arguing, rightly, that identifying it is not as straightforward as some would like to believe. Here, they also maintain that some responses to dissent are not only ineffective but can backfire, lending more credibility, rather than less, to dissenters' views—another observation with which I agree. However, I remain unconvinced that this is generally the case. As I argued earlier this year, there are many examples in which problematic scientific dissent has been accurately identified and when choices about how to deal with it have been both effective and ethically just.



Taking advantage of an opportunity to engage with the public, climate scientist Katharine Hayhoe participates on a panel with Barack Obama.

tifiable (3). Think about the restrictions many jurisdictions impose on the claims of the tobacco industry in advertising its products, for example.

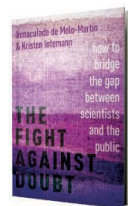
Moreover, finite resources demand that scientists, editors, journalists, and other communicators make judgments about which speech to prioritize. Where such decisions are inescapable, or shouldn't be avoided because of professional and moral obligations, the authors' suggestion that we shift away from a focus on dissenting speech acts and communicative ethics is not appropriate.

The second half of the book contains de Melo-Martín and Intemann's alternative proposal for limiting the negative consequences of problematic scientific dissent, which largely centers on building trust between scientists and the public. The au-

The Fight Against Doubt

*Inmaculada de
Melo-Martín and
Kristen Intemann*

Oxford University Press,
2018. 232 pp.



thors' attention to scientific practice, and to the social and institutional factors that affect the conduct of science, such as its increasing commercialization, are the book's greatest strengths. But the reader might rightly be skeptical of their claim that public doubt about the trustworthiness of scientists is often fair.

How many trust-building opportunities do most scientists have with the general public, especially when other actors mediate their testimony? Intermediaries, such as the news media, have profound control over how scientific communities are perceived. When gatekeepers don't communicate evidence of a scientific community's trustworthiness, or represent it with integrity, scientists' initiatives won't be effective. The key weakness of this book is that it doesn't fully account for such power dynamics or take into consideration the responsibilities of other actors at important nodes within public knowledge systems.

De Melo-Martín and Intemann's other main proposal to combat problematic dissent is to shift the focus of policy discussions from debating the truth of scientific claims to discussing the values that are at stake. This is good advice, but it's worth keeping in mind that some conversations will inevitably loop back to factual questions. As is true about the limits of science, values can only take us so far.

That de Melo-Martín and Intemann share the values and concerns of many of those engaged in this debate but disagree over the best way to mitigate problematic scientific dissent is an important reason to engage with *The Fight Against Doubt*. Its focus on a different aspect of the mechanisms at work—the qualities of the “soil,” rather than the mere existence of “seeds of doubt”—makes this a valuable book. ■

REFERENCES

1. R. N. Proctor, L. Schiebinger, *Agnostology: The Making and Unmaking of Ignorance* (Stanford Univ. Press, 2008).
2. N. Oreskes, E. M. Conway, *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues From Tobacco Smoke to Global Warming* (Bloomsbury Publishing, 2011).
3. E. J. Nash, *Persp. Sci.* **26**, 325 (2018).



Antiaddiction drugs could help curtail the opioid epidemic, but they may pose risks of their own.

3. H. Pétursson, *Addiction* **89**, 1455 (1994).
4. FDA, "Highlights of prescribing information" (2017), chap. 9.3; www.accessdata.fda.gov/drugsatfda_docs/label/2017/020427s016,022006s018lbl.pdf.
5. H. Petursson, M. H. Lader, *Br. Med. J. (Clin. Res. Ed.)* **283**, 643 (1981).

10.1126/science.aau6548

Europe's plans for responsible science

In the past, European framework programs for research and innovation have included funding for the integration of science and society (1). Collaborative projects have brought together diverse sets of actors to co-create and implement common agendas through citizen science, science communication, public engagement, and responsible research and innovation (RRI) and have built an evidence base about science-society interaction (2, 3). In the proposal for the upcoming Horizon Europe program, however, there is no sustained support for RRI, nor is there a program line dedicated to co-creating knowledge and agendas with civil society (4). These serious oversights must be corrected before the Horizon Europe program is adopted by the Council and the European Parliament.

Europe's commitment to RRI and citizen inclusion has led to substantial societal and scientific benefits (5, 6). Researchers who have participated in the framework program projects are more likely to be familiar with RRI principles and to incorporate responsible practices into their work (7). They are more attentive to democratic, social, and economic benefits for stakeholders and more likely to observe that being responsible improves their science (7). RRI is also transforming academic institutions. A growing share of European universities and research organizations are promoting open data sharing, expanding outreach and engagement activities, developing gender equality plans, or establishing research ethics committees and research integrity offices (8). Furthermore, European research and coordination activities have opened up avenues for learning about RRI between countries and have built new professional communities of researchers, administrators, policy-makers, and publics, who share core understandings about responsibility (9). This international mobilization across heterogeneous countries has promoted vital cohesion and cooperation that requires continued concerted support (10).

At a time of global distrust in science,

Edited by Jennifer Sills

Side effects of addiction treatment

Drug addiction is a major global health issue, and the opioid crisis is a notable example of its catastrophic effects (1). In his News In Depth story "Chemists seek antiaddiction drugs to battle hijacked brain" (13 April, p. 139), R. F. Service discusses some promising ways to treat drug addiction, including vigabatrin and a more effective version of that drug named OV329. However, the optimistic tone of the article should be tempered by the potential side effects of these treatments.

The gamma-aminobutyric acid aminotransferase (GABA-AT) enzyme plays a key role in brain signaling by inactivating GABA (2). Currently, the only licensed drug that targets this enzyme is vigabatrin, an anti-epileptic that is usually reserved for severe intractable seizures (2). OV329 is mechanistically similar to vigabatrin, but—as Service explains in the News story—it binds more tightly to GABA-AT.

Although more potent GABA-AT inhibitors may reduce the dopaminergic signaling that is responsible for reinforcing addiction, we should be aware of the possibility that these drugs may themselves cause physical and/or psychological dependence.

Benzodiazepines and barbiturates, drugs that increase GABA signaling by acting directly on the GABA-A receptor, cause a physical dependence with prolonged use, and rapid withdrawal may lead to fatal seizures (3). It is reassuring that vigabatrin has not been reported to have such effects (4), but the side-effect profile of this drug is such that it is seldom prescribed to those without a diagnosis of epilepsy, making it difficult to identify whether seizures that occur on withdrawal are due to the withdrawal of the drug or the underlying disorder. Furthermore, as OV329 is considerably more potent than vigabatrin, the effects of its withdrawal may be comparatively more prominent.

It should be noted that benzodiazepines were not found to be addictive until almost two decades after their introduction (5). When considering the role of GABAergic pharmacotherapy for opioid addiction, we should be cautiously optimistic, but it would be prudent to remember that opioid withdrawal, while highly unpleasant, is not lethal, whereas rapid decreases in GABA signaling may be.

George Yizhou Tang* and Jay Parekh

School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SP, UK.

*Corresponding author.

Email: george.tang@doctors.org.uk

REFERENCES

1. W. Ling, L. Mooney, M. Hillhouse, *Drug Alcohol Rev.* **30**, 300 (2011).
2. R. B. Silverman, *Chem. Rev.* **118**, 4037 (2018).

cynicism toward expertise, and growing concern with inequality (11, 12), efforts to align research and innovation with societal needs and to extend the positive trends generated through the framework programs should not be jeopardized. European citizens should maintain their stake in a future for which research and innovation will be increasingly important.

Niels Mejlgaard,^{1*} Richard Woolley,² Carter Bloch,¹ Susanne Bühner,³ Erich Griessler,⁴ Angela Jäger,³ Ralf Lindner,³ Emil Bargmann Madsen,¹ Frederic Maier,⁵ Ingeborg Meijer,⁶ Viola Peter,⁵ Jack Stilgoe,⁷ Milena Wuketich²

¹Aarhus University, 8000 Aarhus C, Denmark.

²INGENIO (CSIC-UPV), 46022 Valencia, Spain.

³Fraunhofer ISI, 76139 Karlsruhe, Germany.

⁴Institute for Advanced Studies, 1080 Vienna, Austria. ⁵Technopolis Group Belgium, 1150 Brussels, Belgium. ⁶Centre for Science and Technology Studies, Leiden University, 2333 BN, Leiden, Netherlands. ⁷University College London, London WC1E 6BT, UK.

*Corresponding author. Email: nm@ps.au.dk

REFERENCES

1. S. de Saille, *J. Responsible Innov.* **2**, 152 (2015).
2. Science with and for Society (<https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society>).
3. RRI Toolkit (www.rri-tools.eu/).
4. European Commission, "Horizon Europe—the next research and innovation framework programme" (European Commission, Brussels, 2018).
5. V. Peter *et al.*, "Monitoring the evolution and benefits of Responsible Research and Innovation in Europe: Summarising insights from the MoRRI project" (European Commission, Brussels, 2018).
6. U. Felt *et al.*, "Taking European knowledge society

seriously: Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate, Directorate-General for Research, European Commission" (European Commission, Brussels, 2007).

7. S. Bühner *et al.*, "Monitoring the evolution and benefits of RRI: Report on the researchers' survey" (European Commission, Brussels, 2017).
8. European Commission, "The evolution of responsible research and innovation in Europe: The MoRRI indicators report" (European Commission, Brussels, 2018).
9. European Commission, "Commission staff working document—Interim evaluation of Horizon 2020—Annex 2" (European Commission, Brussels, 2017).
10. A. Jónsdóttir, "Open letter on the need of a separate programme for funding science, society and citizens' initiatives" (SIS Network, 2018).
11. D. Sarewitz, *New Atlantis J. Technol. Soc.* **49**, 5 (2016).
12. P. Gluckman, J. Wilsdon, *Palgrave Commun.* **2**, 16077 (2016).

COMPETING INTERESTS

The authors in the past have been or currently are beneficiaries of funding schemes supporting science and society integration within the European framework programs.

10.1126/science.aav0400

Anticipating data-induced bias

In their Report "Prevalence-induced concept change in human judgment" (29 June, p. 1465), D. E. Levari and coauthors show that perception is influenced by contrast. In a sequence of experiments, they demonstrate that when people are accustomed to a certain number of blue dots, threatening faces, or examples of unethical behavior, they will expand their definition of blue, threatening, and unethical once the prevalence of each decreases. Levari *et al.*'s finding that judgment depends on context may be applicable to any field in which humans categorize continuous distributions.

Geologists, for example, identify rock structures based on satellite images. Individuals may vary by 10% in their identification of structures from multiple images of the same area, which differ in contrast (1). Biologists, in an effort to deliver prognoses in the face of climate change, score the heat tolerance of animals by classifying their behavior. The intraresearcher measurement error in such work may amount to 29% (2). Medical doctors diagnose millions of new skin cancer cases per year (3), starting from visual evaluation (4). Error rates of this evaluation may be as high as 50% (4). The contrast problem detected by Levari *et al.* could contribute to the variation inherent in these procedures across fields.

Identifying this researcher bias would open the door to addressing it. The problems in rock structure identification, behavior classification, and skin diagnostics could be solvable by advanced automated quantitative analysis, although current attempts need further refinement (1, 4). There

may also be ways to maintain contrast at a standardized frequency throughout analysis. This could be achieved based on prior knowledge about new data or by adding many known individuals to a few unknown before performing the identification task. Identifying counter-measurements and quality-control tools against this context dependency in science may be limited by lack of awareness rather than lack of options.

Birgit C. Schlick-Steiner, Wolfgang Arthofer, Florian M. Steiner*

Department of Ecology, University of Innsbruck, 6020 Innsbruck, Austria.

*Corresponding author.

Email: florian.m.steiner@uibk.ac.at

REFERENCES

1. T. Scheiber *et al.*, *GFF* **137**, 362 (2015).
2. L. E. Castañeda, G. Calabria, L. A. Betancourt, E. L. Rezende, M. Santos, *J. Therm. Biol.* **37**, 432 (2012).
3. WHO, "INTERSUN: The Global UV Project: A guide and compendium" (World Health Organization, Geneva, 2003).
4. J. Dinnes *et al.*, *Cochrane Database Syst. Rev.* **2015**, 1 (2015).

10.1126/science.aau9816

TECHNICAL COMMENT ABSTRACTS

Comment on "Tracking the global footprint of fisheries"

R. O. Amoroso, A. M. Parma, C. R. Pitcher, R. A. McConnaughey, S. Jennings

Kroodsma *et al.* (Reports, 23 February 2018, p. 904) mapped the global footprint of fisheries. Their estimates of footprint and resulting contrasts between the scale of fishing and agriculture are an artifact of the spatial scale of analysis. Reanalyses of their global (all vessels) and regional (trawling) data at higher resolution reduced footprint estimates by factors of >10 and >5, respectively.

Full text: [dx.doi.org/10.1126/science.aat6713](https://doi.org/10.1126/science.aat6713)

Response to Comment on "Tracking the global footprint of fisheries"

David A. Kroodsma, Juan Mayorga, Timothy Hochberg, Nathan A. Miller, Kristina Boerder, Francesco Ferretti, Alex Wilson, Bjorn Bergman, Timothy D. White, Barbara A. Block, Paul Woods, Brian Sullivan, Christopher Costello, Boris Worm

Amoroso *et al.* demonstrate the power of our data by estimating the high-resolution trawling footprint on seafloor habitat. Yet we argue that a coarser grid is required to understand full ecosystem impacts. Vessel tracking data allow us to estimate the footprint of human activities across a variety of scales, and the proper scale depends on the specific impact being investigated.

Full text: [dx.doi.org/10.1126/science.aat7789](https://doi.org/10.1126/science.aat7789)

NEXTGEN VOICES: SUBMIT NOW

Quality mentoring

Add your voice to *Science*! Our new NextGen VOICES survey is now open:

In her Working Life "Paying it forward as a mentor" (3 August, p. 522), B. Abderrahman describes how a mentor's encouragement can help shape a career. She then explains how her positive mentorship experience inspired her to mentor others. Describe one quality of a mentor you've had that you will try to emulate when you become a mentor yourself.

To submit, go to www.sciencemag.org/nextgen-voices

Deadline for submissions is 31 August. A selection of the best responses will be published in the 5 October issue of *Science*. Submissions should be 100 words or less. Anonymous submissions will not be considered.

TECHNICAL COMMENT

FISHERIES

Comment on “Tracking the global footprint of fisheries”

R. O. Amoroso^{1*}, A. M. Parma², C. R. Pitcher³, R. A. McConnaughey⁴, S. Jennings⁵

Kroodsma *et al.* (Reports, 23 February 2018, p. 904) mapped the global footprint of fisheries. Their estimates of footprint and resulting contrasts between the scale of fishing and agriculture are an artifact of the spatial scale of analysis. Reanalyses of their global (all vessels) and regional (trawling) data at higher resolution reduced footprint estimates by factors of >10 and >5, respectively.

Kroodsma *et al.* (1) used automatic identification system (AIS) data to track vessels they classified as “fishing” and estimated that fishing activities occurred in 55% of the world’s oceans in 2016. We show how strongly their results depend on the spatial scale of analysis. Their method gridded the ocean into large cells of 0.5° at the equator (~3100 km²) and counted every cell with any assumed fishing event of any duration in 2016 as fished, thus contributing its total area to fishing footprint.

We accessed the 0.01° grid fishing data made available by Global Fishing Watch (2) and re-analyzed these data at resolutions of ~3100, ~123, and ~1.23 km² (corresponding to 0.5°, 0.1°, and 0.01° at the equator), giving footprint estimates of 49%, 27%, and 4% of ocean area, respectively. Thus, higher-resolution analyses reduced their global fishing footprint estimates by a factor of >10. Our estimate of footprint at 0.5° (49%) differs from that reported by Kroodsma *et al.* (55%) because they improved their algorithm to identify fishing by squid jiggers after publication and updated data in the current release. Also, the method we used to reallocate fishing activity to grids differed slightly from that in Kroodsma *et al.*, leading to small differences in absolute footprint estimates, but these do not affect the relative relationships between footprints across spatial scales.

Kroodsma *et al.* also state that their 55% fishing footprint is larger than that of agriculture by a factor of 4. However, this comparison is strongly biased by the different scales of analysis and different criteria used to assign grid cells to

fishing or farming. The estimates of agricultural land-use footprint they use for comparison are gridded at higher resolution (5', ~86 km² versus ~3100 km²) and also account for the fraction of farmed or grazed area within each grid cell (3). Thus, the agricultural footprint describes only the area directly affected by farming, ignoring

any wider area subject to diffuse environmental impacts. Our more comparable high-resolution fishing footprint is less than the agriculture footprint by a factor of approximately 3.5.

All human activities have diffuse impacts that extend beyond the area of activity. However, for fishing activities, using a spatial grid of an arbitrary low resolution does not provide an appropriate or consistent quantitative assessment of diffuse impact. For example, some diffuse impacts would be assessed more effectively using catch and bycatch data and population or community analyses that account for the diverse movements and life histories of affected populations and species, as well as the different rates of mortality that result from their varied interactions with fishing activities (4–6).

We also quantified the effects of grid resolution on trawl fishing footprints with the Global Fishing Watch data (2). We focused on trawling because footprint is a consistent and well-defined concept for trawling vessels, which tow a net or nets directly behind the vessel(s) and for which gear dimensions are known or can be estimated more reliably. Further, high-resolution footprints for bottom trawling (although Kroodsma *et al.* did not distinguish bottom trawls from trawls

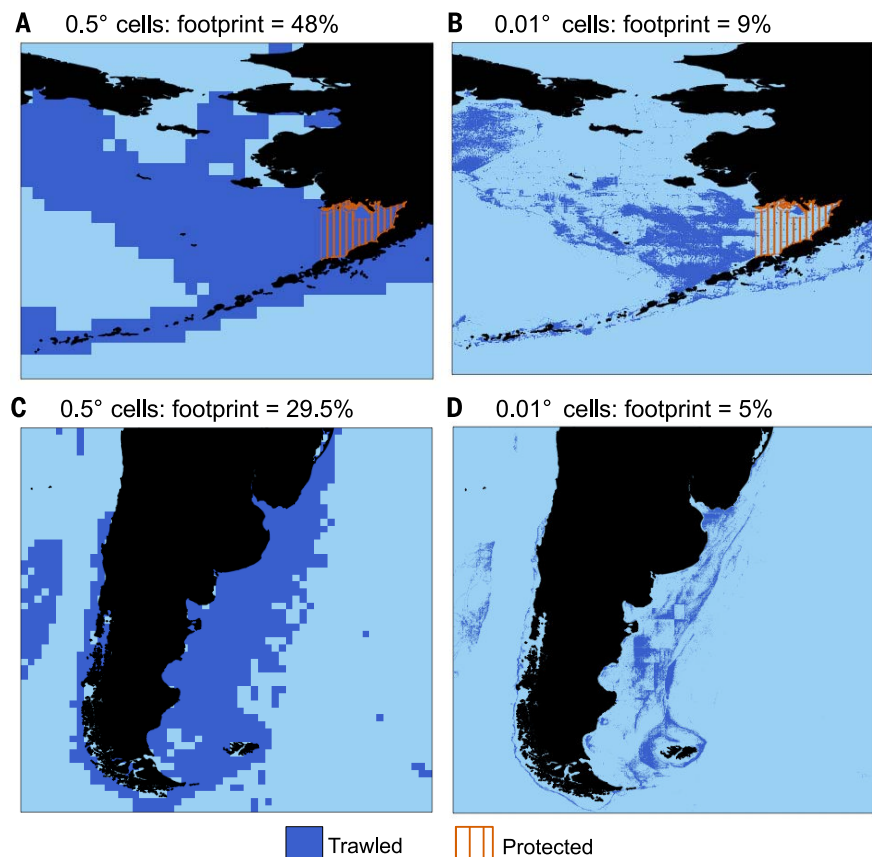


Fig. 1. Effect of grid resolution on the perception of fishing footprint. The areas in dark blue show the trawling footprints estimated for 2016 with (A and C) an equal-area grid with 0.5° resolution at the equator; (B and D) an equal-area grid with 0.01° resolution at the equator. The hatched area shows an example region of the North Pacific where all trawling was prohibited.

¹School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98103, USA. ²Centro para el Estudio de Sistemas Marinos, CENPAT, CONICET, 9120 Puerto Madryn, Chubut, Argentina. ³CSIRO Oceans and Atmosphere, Queensland Biosciences Precinct, St. Lucia, Queensland 4067, Australia. ⁴Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 7600 Sand Point Way N.E., Seattle, WA 98115, USA. ⁵International Council for the Exploration of the Sea, H. C. Andersens Boulevard 44-46, 1553 Copenhagen V, Denmark.

*Corresponding author. Email: ramoroso@uw.edu

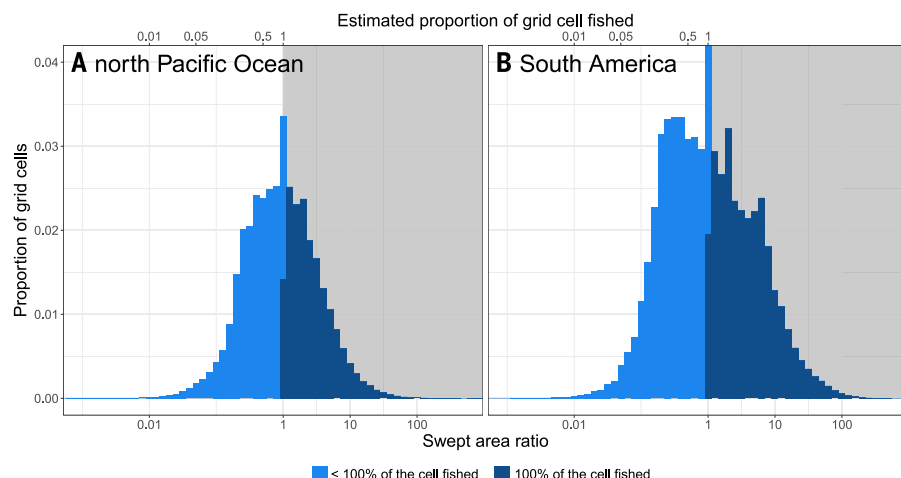


Fig. 2. Estimated distribution of the swept-area ratio within 0.01° grid cells contributing to the trawling footprint. (A) North Pacific and (B) South America regions during 2016. Light blue bars show estimated proportions of the grid cells where trawling covered less than 100% of the cell.

that do not contact the seabed) have long been used as metrics to assess fishing impacts on seabed habitats [e.g., (7–9)].

To illustrate the effects of grid resolution on trawling footprints, we considered regions of the north Pacific Ocean and off southern South America. For each region, trawling footprint (as proportion of the ocean area) was calculated using equal-area grids of 0.5° and 0.01° at the equator (Fig. 1). At the higher resolution of analysis, the estimated footprints in these regions fell by factors of 5.3 (48% to 9%) and 5.9 (29.5% to 5%), respectively. Further, if we take as an example a region of the north Pacific Ocean where trawling was banned in 2016 (10) (Fig. 1, A and B), then 100% of this area (59,000 km² of ocean) was incorrectly classified as trawled at 0.5° resolution. For such reasons, many published analyses of trawling footprints are conducted at higher resolution (11–13).

Even our highest-resolution regional analyses (0.01°) overestimate trawling footprint. This is because the grid-based method assumes that any trawling recorded in a cell justifies adding the entire cell area to the footprint. More sophisticated approaches for assessing footprint already account for trawling distributions within cells (14, 15). Untrawled area in a cell is a function of the swept-area ratio (SAR). SAR is defined as the total area swept by trawling in the cell divided

by the cell area. For the two example regions, we converted trawling effort in hours per cell into SAR, assuming conservatively high values for trawling speed (4 knots) and trawled path width (trawl door spread of 200 m). In existing analyses of trawling footprints, towing speed and door spread are usually allocated by vessel or by fleet to account for differences in gear type (8, 9), although such specifications were not available for the Global Fishing Watch data (2). Overall, 53% of 0.01° cells in the north Pacific and 52% of 0.01° cells off South America have SAR < 1 and could not have been fully trawled in 2016 (Fig. 2). Conservatively assuming that trawling activity was spread uniformly within each cell, the trawling footprint in each region fell further to 6.5% (factor of 7.4 reduction relative to 0.5° gridded approach) and 3% (factor of 9.8 reduction), respectively.

A coarse gridding of the positions of fishing vessels (globally or regionally) that ignores differences in catching power among vessels and gear, or ignores the scale of their direct and diffuse impacts, leads to footprint estimates that are primarily driven by the spatial resolution of analysis. Such analyses are unlikely to be a good proxy for the footprint of fishing or the status of species or ecosystems affected by fishing. The high temporal resolution of AIS data can provide valuable insight into the behavior of individual

vessels and allowed Kroodsmas *et al.* to classify different types and patterns of fishing activity. These analyses alone are an interesting achievement, but the footprint estimates and comparisons with agriculture highlighted in their report are misleading.

REFERENCES AND NOTES

1. D. A. Kroodsmas *et al.*, *Science* **359**, 904–908 (2018).
2. Global Fishing Watch; www.globalfishingwatch.org.
3. J. A. Foley *et al.*, *Nature* **478**, 337–342 (2011).
4. P. G. Fernandes *et al.*, *Nat. Ecol. Evol.* **1**, 0170 (2017).
5. N. K. Dulvy *et al.*, *eLife* **3**, e00590 (2014).
6. D. Ricard, C. Minto, O. P. Jensen, J. K. Baum, *Fish Fish.* **13**, 380–398 (2012).
7. G. J. Piet, F. J. Quirijns, *Can. J. Fish. Aquat. Sci.* **66**, 829–835 (2009).
8. O. R. Eigaard *et al.*, *ICES J. Mar. Sci.* **73** (suppl. 1), i27–i43 (2016).
9. O. R. Eigaard *et al.*, *ICES J. Mar. Sci.* **74**, 847–865 (2017).
10. North Pacific Fishery Management Council, *Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands* (North Pacific Fishery Management Council, Anchorage, AK, 2017).
11. J. Black, R. Wood, T. Berthelsen, R. Tinley, *Monitoring New Zealand's Trawl Footprint for Deepwater Fisheries: 1989–90 to 2009–2010*, New Zealand Aquatic Environment and Biodiversity Rep. 110 (2013).
12. S. A. Murawski, S. E. Wigley, M. J. Fogarty, P. J. Rago, D. G. Mountain, *ICES J. Mar. Sci.* **62**, 1150–1167 (2005).
13. S. J. Baird, J. Hewitt, B. A. Wood, *Benthic Habitat Classes and Trawl Fishing Disturbance in New Zealand Waters Shallower than 250 m*, New Zealand Aquatic Environment and Biodiversity Rep. 144 (2015).
14. N. Ellis, F. Pantus, C. R. Pitcher, *Can. J. Fish. Aquat. Sci.* **71**, 733–746 (2014).
15. H. D. Gerritsen, C. Minto, C. Lordan, *ICES J. Mar. Sci.* **70**, 523–531 (2013).

ACKNOWLEDGMENTS

D. Kroodsmas provided guidance and assistance with the interpretation and correct use of the database. **Funding:** Supported by the David and Lucile Packard Foundation, the Walton Family Foundation, the Alaska Seafood Cooperative, American Seafoods Group U.S., Blumar Seafoods Denmark, Clearwater Seafoods Inc., Espersen Group, Glacier Fish Company LLC U.S., Gorton's Seafood, Independent Fisheries Limited N.Z., Nippon Suisan Inc. (USA), Pesca Chile S.A., Pacific Andes International Holdings Ltd., San Arawa S.A., Sanford Ltd. N.Z., Sealord Group Ltd. N.Z., South African Trawling Association, and Trident Seafoods (R.O.A.); the Agencia Nacional de Promoción Científica y Tecnológica, Argentina (project PICT 2016-0218) (A.M.P.); the Commonwealth Scientific and Industrial Research Organization (C.R.P.); the National Oceanic and Atmospheric Administration (R.A.M.); and the International Council for the Exploration of the Sea (S.J.). **Author contributions:** R.O.A., A.M.P., C.R.P., R.A.M., and S.J. conceived the study; R.O.A., A.M.P., and S.J. wrote the manuscript with further input from all authors; and R.O.A. performed the data analyses and produced the figures. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All the data used for these analyses are available from the Global Fishing Watch webpage (2).

27 March 2018; accepted 25 July 2018
10.1126/science.aat6713

TECHNICAL RESPONSE

FISHERIES

Response to Comment on “Tracking the global footprint of fisheries”

David A. Kroodsma^{1*}, Juan Mayorga^{2,3}, Timothy Hochberg¹, Nathan A. Miller⁴, Kristina Boerder⁵, Francesco Ferretti⁶, Alex Wilson⁷, Bjorn Bergman⁴, Timothy D. White⁶, Barbara A. Block⁶, Paul Woods¹, Brian Sullivan⁷, Christopher Costello², Boris Worm⁵

Amoroso *et al.* demonstrate the power of our data by estimating the high-resolution trawling footprint on seafloor habitat. Yet we argue that a coarser grid is required to understand full ecosystem impacts. Vessel tracking data allow us to estimate the footprint of human activities across a variety of scales, and the proper scale depends on the specific impact being investigated.

We welcome Amoroso *et al.*’s comment (1), which demonstrates the power of vessel tracking data to estimate the environmental footprint of human activities on ocean ecosystems over a range of relevant scales. Their contribution also highlights the importance of making vessel tracking data freely available, allowing others to ask more detailed questions about the effects of fishing on ocean ecosystems. We disagree, though, that the environmental impacts of fishing can be easily divided into “diffuse” and “direct,” or that there is a “correct” scale of analysis; rather, the chosen scale depends both on the method of fishing and the question being asked. We further demonstrate that our estimate of the area of the ocean fished was conservative given the questions we addressed.

Consider the analogous challenge of calculating the global environmental footprint of motor vehicles from car tracking data. To estimate the immediate habitat loss, one would measure the area covered by roads (1- to 100-m scale). If the question pertained to air quality, the scale of inquiry would broaden to a range of 100 m to 100 km from the roadside (2). Ecosystem impacts manifest at various scales (1 to 100 km), such as through roadkill and fragmentation (3), whereas climate impacts are global (>1000 km); hence, the scale of analysis depends on the environmental damages in question. The determination of which impacts are diffuse versus direct is subjective; for example, a person with asthma would deem air pollution to be a direct impact.

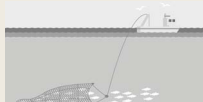

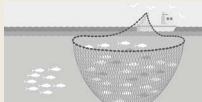
In fisheries, disturbance of seafloor habitat by bottom trawling is one of the best-known envi-

ronmental aspects of bottom fishing (4). Better estimates of this footprint, as shown by the analyses of Amoroso *et al.*, can now be achieved using fine-scale automatic identification system (AIS) vessel tracking data. But this scale of analysis is not universally applicable. Drifting longlines or purse seines, although they traverse more of the ocean than trawlers, have no contact with the seafloor, and thus no footprint by this assessment. Industrial longlines, for example, contain thousands of baited hooks and move with the local currents. A single set can “drift” many kilometers, intersecting the paths of mobile predators such as tuna and sharks along the way. Thus, the footprint of drifting longlines would be the area of the polygon defined by the start and end locations of the setting and hauling of gear (mean for a study of the Hawaii longline fleet: 224 km²) (5). For tuna purse seines, it

would be the area of the net, which can be almost half a kilometer in diameter (6) (Table 1).

A broader question relates to the spatial footprint of fishing on the abundance of target species. Catching fish in one location likely affects biomass across roughly the area that those fish travel. A review of commonly targeted species shows wide ranges for different species. Tagging data of European hake, which is one of the top species caught by trawling in southern Europe, show that individuals generally travel 20 to 40 km over several months, with some traveling more than 200 km (7). In the Bering Sea, Pacific cod and sablefish are two of the three most frequently landed target species by trawlers. Pacific cod were recaptured typically 20 to 370 km from where they were released, with some traveling more than 900 km (8), and sablefish were found to travel at a mean rate of 191 km/year (9). Many pelagic species travel much farther. In the Pacific, the top species targeted by longlines and purse seines are yellowfin, bigeye, skipjack, and albacore tuna. Tagging data show that all of these species travel hundreds to thousands of kilometers (10), and the average “activity space” of yellowfin tuna exceeds 250,000 km², the area of a grid cell just under 5° × 5° at the equator (11). All of these fish travel distances much larger than 0.01° (~1.1 km), and almost all are larger than 0.5° (~55 km). Figure 1 illustrates these much larger grid sizes that reflect typical species movements and compares them with the highest resolution of our public dataset.

The footprint calculation in our paper (12) served two goals: (i) to compare our dataset with previous global fisheries datasets, which are gridded at the same 0.5° scale (13, 14); and (ii) to make a general comparison with the area used for other forms of food production, namely agriculture. The majority of agricultural land is used for grazing (15), and the footprint of this

Table 1. Appropriate scales for different footprints. Examples of appropriate scales for calculating the footprint of fishing, given different fishing gear (columns) and questions asked (rows).			
	 Bottom trawl targeting European hake in the Adriatic Sea	 Drifting longline targeting yellowfin tuna in the Pacific	 Purse seine targeting yellowfin tuna in the Pacific
Seabed disturbed by gear	25–30m (width of door spread of trawl in Adriatic) (16) × vessel track while fishing	0	0
Area swept by gear while fishing	25–30m (width of door spread of trawl in Adriatic) × vessel track while fishing	Per set, length of longline × soak time × mean current (~224 km ² for study of Hawaii longline fleet) (5)	~0.3 km ² per set (assuming net length ~2 km) (6)
Typical movements of target species	~25 km	~500 km	~500 km

¹Global Fishing Watch, Washington, DC 20036, USA.
²University of California, Santa Barbara, CA 93106, USA.
³Pristine Seas, National Geographic Society, Washington, DC 20036, USA. ⁴SkyTruth, Shepherdstown, WV 25443, USA.
⁵Biology Department, Dalhousie University, Halifax, NS B3H 4R2, Canada. ⁶Biology Department, Stanford University, Stanford, CA 94305, USA. ⁷Google, Mountain View, CA 94043, USA.
*Corresponding author. Email: david@globalfishingwatch.org

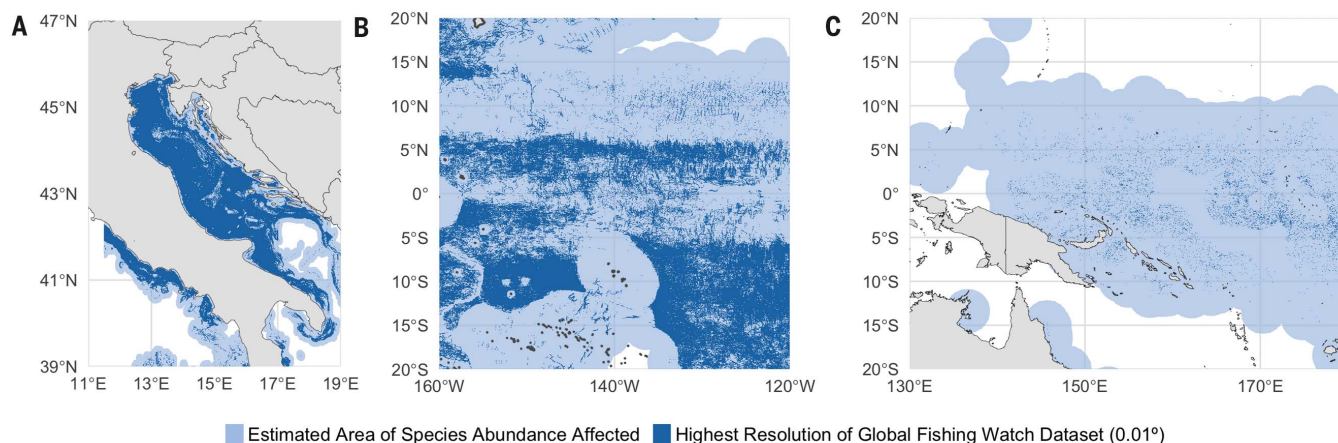


Fig. 1. Area swept by different fishing gear and area of fish biomass affected. Shown is fishing effort in 2016 from AIS data (11) for (A) trawlers in the Adriatic Sea, (B) drifting longlines in the central equatorial Pacific, and (C) purse seines in the western equatorial Pacific. Dark blue shows the highest resolution of the dataset, 0.01° resolution (~1.1 km at the equator), which may be a slight underestimate of area swept by drifting longlines and an overestimate of area swept by trawlers and purse seines (Table 1). Light blue shows target species' average movements mapped beyond fishing locations, using the range of European hake in the Adriatic (~25 km) and yellowfin tuna in the Pacific (~500 km).

food production system relates to the area that is grazed, not the area traversed by trucks hauling livestock to slaughter. In other words, the footprint of agriculture relates to the approximate area of ecosystems where a substantial portion of the net primary productivity is appropriated, directly or indirectly, for human consumption. In the ocean, the comparable “area fished” would be the area of the ecosystem that supports targeted fish, not the area swept by fishing gear. Our grid size, about 55 km on a side (0.5° at the equator), is conservative for most commonly targeted species (8–15). Using this scale, the area of marine ecosystems supporting fish caught by humans is more than four times that of terrestrial ecosystems in agriculture.

Our estimates and those of Amoroso *et al.* are just two of several footprint estimates that can be derived from our AIS vessel tracking data. For example, we can use engine power estimates to infer the fuel used and thus the carbon footprint of vessels in the database. We can also look

at the spatial overlap of fishing behavior with different species to understand risks of bycatch. With high temporal resolution, we can estimate how fishing pressure changes across and between years.

It is important to have these vessel tracking data freely available for such analyses, and we hope that other forms of tracking data, like those derived from vessel monitoring systems (VMS), are made more widely available for comparative scientific research. As demonstrated by both Amoroso *et al.*'s analyses and ours, these data can allow us to answer specific questions about the environmental impact of fishing on marine ecosystems, which in turn can contribute to improved, evidence-based management of the oceans.

REFERENCES AND NOTES

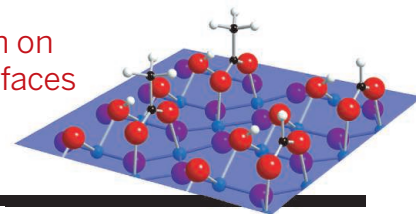
1. R. O. Amoroso, A. M. Parma, C. R. Pitcher, R. A. McConnaughey, S. Jennings, *Science* **361**, eaat6713 (2018).
2. C. Lamancusa, F. Parvez, K. Wagstrom, *Atmos. Environ.* **150**, 229–237 (2017).
3. S. C. Trombulak, C. A. Frissell, *Conserv. Biol.* **14**, 18–30 (2001).
4. M. J. Kaiser, J. S. Collie, S. J. Hall, S. Jennings, I. R. Poiner, *Fish. Fish.* **3**, 114–136 (2002).
5. D. C. Dunn, C. Y. Kot, P. N. Halpin, *Fish. Res.* **92**, 268–276 (2008).
6. FAO Fisheries and Aquaculture Department, “Fishing Techniques: Tuna Purse Seining”; www.fao.org/fishery/fishtech/40/en.
7. H. de Pontual *et al.*, *ICES J. Mar. Sci.* **60**, 1318–1327 (2003).
8. A. M. Shimada, D. K. Kimura, *U.S. Natl. Mar. Fish. Serv. Fish. Bull.* **92**, 800–816 (1994).
9. D. H. Hanselman, J. Heifetz, K. B. Echave, S. C. Dressel, *Can. J. Fish. Aquat. Sci.* **72**, 238–251 (2015).
10. B. A. Block *et al.*, *Nature* **475**, 86–90 (2011).
11. K. M. Schaefer, D. W. Fuller, B. A. Block, *Fish. Res.* **112**, 22–37 (2011).
12. D. A. Kroodsmas *et al.*, *Science* **359**, 904–908 (2018).
13. R. A. Watson, *Sci. Data* **4**, 170039 (2017).
14. D. Pauly, D. Zeller, Eds., *Catch Reconstruction: Concepts, Methods and Data Sources* (University of British Columbia, 2015); www.seaaroundus.org/catch-reconstruction-and-allocation-methods/.
15. J. A. Foley *et al.*, *Nature* **478**, 337–342 (2011).
16. A. Lucchetti, A. Sala, *Can. J. Fish. Aquat. Sci.* **69**, 1806–1816 (2012).

17 April 2018; accepted 25 July 2018
10.1126/science.aat7789

RESEARCH

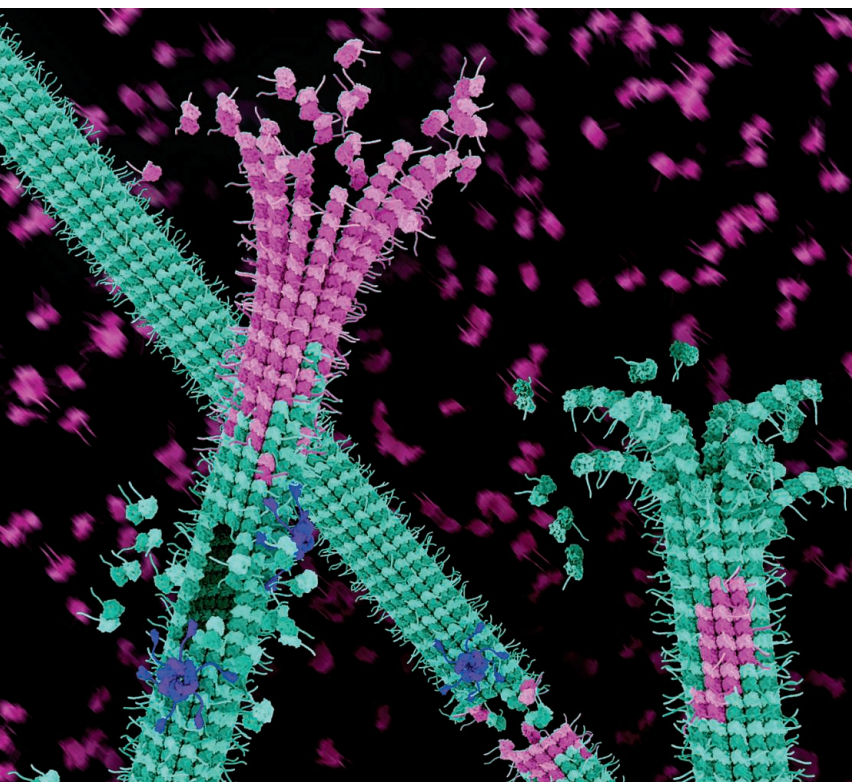
Overlayer formation on titanium dioxide surfaces

Balajka et al., p. 786



IN SCIENCE JOURNALS

Edited by **Stella Hurtley**



CYTOSKELETON

Severing to build microtubules

Microtubules are essential intracellular polymers, built from tubulin subunits, that establish cell shape, move organelles, and segregate chromosomes during cell division. Vemu *et al.* show that microtubule-severing enzymes extract tubulin subunits along the microtubule shaft. This nanoscale damage is repaired by the incorporation of free tubulin, which stabilizes the microtubule against depolymerization. When extraction outpaces repair, microtubules are severed, emerging with stabilized ends composed of fresh tubulin. The severed microtubules act as templates for new microtubule growth, leading to amplification of microtubule number and mass. Thus, seemingly paradoxically, severing enzymes can increase microtubule mass in processes such as neurogenesis and mitotic spindle assembly. —SMH

Science, this issue p. 768

New tubulin (pink) patches small holes in microtubule lattices made by severing enzymes (purple).

CANCER

Mutational signature sleuthing

Individuals with the inherited skin disease recessive dystrophic epidermolysis bullosa (RDEB) are predisposed to developing aggressive squamous cell carcinomas (SCCs). Cho *et al.* sequenced multiple RDEB SCC tumors and found that the mutation profile in these carcinomas was most consistent with APOBEC-associated mutagenesis, unlike other types of SCCs driven by ultraviolet light or exposure to tobacco smoke. This finding could help to target SCCs in RDEB patients. —CAC

Sci. Transl. Med. **10**, eaas9668 (2018).

OCEAN OXYGEN

Fishin' gone?

Because gas solubility decreases as temperatures increase, global warming is likely to cause oxygen loss from the oceans. This could have a detrimental impact on fish populations, the fishing industry, and global food availability. Have such impacts occurred before? Yao *et al.* report sulfur isotopic data from the Paleocene-Eocene Thermal Maximum, an interval around 55 million years ago when atmospheric carbon dioxide concentrations and global temperatures were also high. They found widespread anoxia and resulting high concentrations of hydrogen sulfide, which is toxic to marine organisms. Similar effects

could have severe negative effects on ocean ecosystems. —HJS

Science, this issue p. 804

VIROLOGY

Hidden truths of enteroviruses

Enteroviruses are important drivers of global health, but few countries undertake enterovirus surveillance. Pons-Salort and Grassly used Japanese surveillance data to model the interplay between the ratio of susceptible and immune individuals, accounting for declining birth and death rates, incomplete surveillance, and seasonality of infection (see the Perspective by Nikolay and Cauchemez). Enteroviruses have

highly predictable yet highly non-linear dynamics. The model also reveals signatures of increased pathogenicity and of antigenic change and transmissibility. —CA

Science, this issue p. 800;
see also p. 755

GENETICS

Humans as models of human disease

Mice are a convenient model for exploring the functions of cellular signaling pathways. Occasionally, however, an “experiment of nature” highlights the perils of overreliance on mice. RIPK1 is a well studied protein kinase that regulates cell death. Mice deficient in RIPK1 die soon after birth

because of the protein's widespread role in multiple tissues and organs. Cuchet-Lourenço *et al.* studied patients with inherited immunodeficiency of unknown cause (see the Perspective by Pasparakis and Keliher). They identified inactivating mutations in the *RIPK1* gene in four individuals. Unlike what has been seen in mice, the deleterious effects of *RIPK1* loss in humans were confined to the immune system, a finding with potential therapeutic implications. —PAK

Science, this issue p. 810;
see also p. 756

SOCIAL ROBOTICS

Robots help autistic kids interact with adults

Children with autism spectrum disorder (ASD) often struggle with social behaviors such as recognizing emotional responses in others and understanding gaze direction. Scassellati *et al.* put a fully autonomous, adaptive robot in the homes of 12 children with ASD for 1 month to help improve the children's social skills. The robot and a caregiver engaged with each child for 30 minutes every day, playing games that involved activities such as emotional storytelling and taking another's perspective. The robot autonomously adapted task difficulty to each child's performance and modeled appropriate gaze directions. The children showed improvements in attention skills and, crucially, demonstrated the improvements while interacting with adults even when the robot was not present. —RLK

Sci. Robot. **3**, eaat7544 (2018).

GAS GIANT PLANETS

Moons drive structure in Jupiter's aurorae

Like Earth, Jupiter has aurorae generated by energetic particles hitting its atmosphere. Those incoming particles can come from Jupiter's moons Io and Ganymede. Mura *et al.* used infrared observations from the Juno spacecraft to image the moon-generated aurorae. The

pattern induced by Io showed an alternating series of spots, reminiscent of vortices, and sometimes split into two arcs. Aurorae related to Ganymede could also show a double structure. Although the cause of these unexpected features remains unknown, they may provide a way to examine how the moons produce energetic particles or how the particles propagate to Jupiter. —KTS

Science, this issue p. 774

PROTEOMICS

The blood proteome in disease

Understanding the function of human blood serum proteins in disease has been limited by difficulties in monitoring their production, accumulation, and distribution. Emilsson *et al.* investigated human serum proteins of more than 5000 Icelanders over the age of 65. The composition of blood serum includes a complex regulatory network of proteins that are globally coordinated across most or all tissues. The authors identified modules and functional groups associated with disease and health outcomes and were able to link genetic variants to complex diseases. —LMZ

Science, this issue p. 769

GRAPHENE

An electronic wedding cake

In nanostructures such as quantum dots, spatial confinement forces electrons to assume discrete energy levels. Quantization can also occur in an external magnetic field, where electrons' energies group into so-called Landau levels (LLs). Gutiérrez *et al.* explored the interplay between these two mechanisms and electronic interactions in a circulator resonator made of graphene. As an external magnetic field was increased, the electron quantum states transformed from atomic-like states to LL-like states. Electronic interactions caused a characteristic wedding cake-like shape of electronic density at high fields. —JS

Science, this issue p. 789

IN OTHER JOURNALS

Edited by **Sacha Vignieri**
and **Jesse Smith**

STELLAR EVOLUTION

A population of merged white dwarfs

White dwarfs are the hot exposed cores left over when a dying low-mass star throws off its outer layers. The Gaia spacecraft recently provided accurate distances to more than a hundred times as many white dwarfs than were previously available, allowing detailed studies of the population. Kilic *et al.* investigated the mass and composition of these white dwarfs and reproduced them with stellar evolution models. They show that about 15% of the white dwarfs have higher-than-expected masses. This is a sign that they formed from mergers, either between two parent stars or two white dwarfs. —KTS

Mon. Not. R. Astron. Soc. Lett. **479**, L113 (2018).

Artist's rendering of the Gaia spacecraft

BLOOD

A turbulent way to make platelets

Donations from volunteers are the only source of blood for transfusions. But blood components such as platelets have a shelf life of only 5 days, and alternative sources of platelets are in demand. By visualizing fluorescently tagged megakaryocytes (precursor cells of platelets) in transgenic mice, Ito *et al.* demonstrated that highly turbulent blood flow is a determining factor of platelet production from megakaryocytes. Turbulence triggered the production of thrombopoietic factors from megakaryocytes, which, along with shear stress, stimulated platelet release. By using a turbulence-controllable bioreactor, functionally viable platelets could be generated from megakaryocytes derived from human-induced pluripotent stem cells at a quantity that

satisfies clinical-scale demand, suggesting the possibility of de novo platelet production as an alternative to acquiring platelets through blood donations. —MY
Cell **174**, 636 (2018).

INFECTIOUS DISEASE

Zika in the testes: A Trojan horse

Zika virus (ZIKV) is a mosquito-borne flavivirus that can also be sexually transmitted. Although people infected with ZIKV are often asymptomatic, there is an association between ZIKV infection in pregnant women and severe birth defects in their children. Matusali *et al.* showed that ZIKV can replicate for several days in testicular tissue explants. ZIKV infects testicular somatic cells, germ cells, and spermatozoa, and its presence has been detected in semen samples from ZIKV-infected patients. Despite



induction of antiviral genes, no overt inflammatory response was observed, and testicular morphology and hormone production remained unaffected. Apparently, ZIKV remains quiescent in the testes. This phenomenon may explain asymptomatic disease transmission and offer a possible target for antiviral drugs. —MY

J. Clin. Invest. 10.1172/JCI121735 (2018).

CELL BIOLOGY

Raman spectroscopic transcriptome

Raman spectroscopy records the vibrational Raman spectra of biomolecules, allowing the determination of chemical species in target samples. Although nondisruptive and label-free, its application in cell imaging is limited because of the complexity of cellular compositions. Kobayashi-Kirschvink *et al.* described

computational approaches to extract key information from cellular Raman spectra in yeast and bacteria. Dimension-reduction methods enabled a direct prediction of transcriptomes from cellular Raman spectra. A better correlation between the Raman spectra and noncoding transcriptome revealed a more linear correspondence of noncoding RNAs with cellular constituent biomolecules. Although it remains unknown if the linkage can be demonstrated at the single-cell level, spectroscopic live-cell omics provides new possibilities to monitor different cellular states. —SYM

Cell Syst. 7, 104 (2018).

EDUCATION

Teaching and research in synergy

The unspoken belief in graduate student training is that time

spent teaching, or learning about teaching, will decrease research productivity. Shortlidge and Eddy tested this trade-off using a national sample of life science Ph.D. students. Results show that graduate students who participated in evidence-based teaching (EBT) trainings did not exhibit a reduction in confidence in their research career, in their ability to communicate their research, or in their publication number. Contrary to current belief, the data trended toward a slight synergy between investing in EBT and research preparation, suggesting that institutions can integrate EBT training into graduate programs without reducing students' research success. Additionally, investment in EBT training can better prepare graduate students for the multifaceted role of a faculty member. —MMc

PLOS ONE 13, e0199576 (2018).

HIBERNATION

Snowy bat caves

Snow provides thermal protection from extreme temperatures, a phenomenon capitalized on by polar bears and people indigenous to parts of the Arctic. But snow does not provide a cozy environment, a likely reason why more mammals have not evolved to take advantage of its protection against extreme cold. Hirakawa and Nagasaka, however, report that Ussurian tube-nosed bats (*Murina ussuriensis*) appear to create tiny snow "caves" with their bodies, which then serve as opportunistic hibernacula. After coming

across anecdotal accounts of small bats being found curled up in the snow, the authors systematically searched for bats in such conditions, finding more than 30 near Sapporo, Japan. The animals displayed classic torpor positions, curled nose-to-tail, and decreased body temperatures. The authors model conditions under which the bats' bodies could create their small caves and set forth several hypotheses for hibernating conditions and scenarios. The finding that these bats use a snowy blanket for hibernation protection opens up the opportunity for many intriguing questions to be answered. —SNV

Sci. Rep. 8, 12047 (2018).

BIOMATERIALS

Layers of bone repair

The repair of osteochondral defects requires proper growth of subchondral bone, articular cartilage, and the interface between them, but current treatments have been more palliative than curative. Kang *et al.* developed a trilayer scaffold, with variations in the architecture and mineral environment along its depth. The bottom layer was enhanced with calcium phosphate to recruit endogenous cells, whereas the upper two layers were loaded with donor cells to support stratified cartilage formation. When implanted in vivo, these trilayer scaffolds formed osteochondral tissue with a lubricin-rich cartilage surface. —MSL

Acta Biomater. 10.1016/j.actbio.2018.07.039 (2018).



Ussurian tube-nosed bats hibernate in tiny snow caves.

ALSO IN SCIENCE JOURNALS

Edited by Stella Hurtley

HEALTH CARE

Population health and national development

Healthy populations translate into productive and stable nations. Universal health care (UHC) is a pragmatic and ethical ideal that, thanks to social and economic progress, seems almost achievable. However, UHC means different things in different contexts. The minimum ideal is that no individual or family should suffer financial hardship because of accessing good-quality medical assistance. Bloom *et al.* review health priorities around the world and what will be needed in terms of skills, funds, and technology to achieve health care access for all. —CA

Science, this issue p. 766

NEUROSCIENCE

Developing the blood-brain barrier

During development, signals need to be dynamically integrated by endothelial cells, neurons, and glia to achieve functional neuro-glia-vascular units in the central nervous system. During cortical development, neuronal Dab1 and ApoER2 receptors respond to a guidance cue called reelin. Studying mice, Segarra *et al.* found that Dab1 and ApoER2 are also expressed in endothelial cells (see the Perspective by Thomas). The integration of reelin signaling in endothelial cells and neurons facilitates the communication between vessels, glia, and neurons that is necessary for the correct positioning of neurons during cortical development. This integration is also important for correct communication at the neurovascular unit required for blood-brain barrier integrity in the mature brain. —PRS

Science, this issue p. 767;
see also p. 754

QUASICRYSTALS

Dirac fermions in quasicrystalline graphene

Quasicrystal lattices, which can have rotational order but lack translational symmetry, can be used to explore electronic properties of materials between crystals and disordered solids. Ahn *et al.* grew graphene bilayers rotated exactly 30° that have 12-fold rotational order. Electron diffraction and microscopy confirmed the formation of quasicrystals, and angle-resolved photoemission spectroscopy revealed anomalous interlayer electronic coupling that was quasi-periodic. The millimeter-scale layers can potentially be transferred to other substrates. —PDS

Science, this issue p. 782

SURFACE CHEMISTRY

A preference for acids

When titanium dioxide surfaces are exposed to water under ambient conditions, an ordered overlayer forms. Balajka *et al.* studied this process with scanning tunneling microscopy and x-ray photoelectron spectroscopy for water adsorption under vacuum conditions and in air (see the Perspective by Park). The ordered overlayer was only formed in air, the result of the adsorption of organic acids (formic and acetic acids). Although other species such as alcohols were present in much higher concentrations in air, the bidentate adsorption and entropic effects favored acid adsorption. —PDS

Science, this issue p. 786;
see also p. 753

MAGNETISM

Cooperative quantum magnetism

One of the earliest and most intensively studied problems in quantum optics is the interaction of a two-level system (an

atom) with a single photon. This simple system provides a rich platform for exploring exotic light-matter interactions and the emergence of more complex phenomena such as superradiance, which is a cooperative effect that emerges when the density of atoms is increased and coupling between them is enhanced. Going beyond the light-matter system, Li *et al.* observed analogous cooperative effects for coupled magnetic systems. The results suggest that ideas in quantum optics could be carried over and used to control and predict exotic phases in condensed matter systems. —ISO

Science, this issue p. 794

BATTERIES

An elevated lithium battery

Batteries based on lithium metal and oxygen could offer energy densities an order of magnitude larger than that of lithium ion cells. But, under normal operation conditions, the lithium oxidizes to form peroxide or superoxide. Xia *et al.* show that, at increased temperatures, the formation of lithium oxide is favored, through a process in which four electrons are transferred for each oxygen molecule (see the Perspective by Feng *et al.*). Reversible cycling is achieved through the use of a thermally stable inorganic electrolyte and a bifunctional catalyst for both oxygen reduction and evolution reactions. —MSL

Science, this issue p. 777;
see also p. 758

ENVIRONMENTAL TOXINS

Mercury sinking

Mercury is a highly toxic, globally ubiquitous pollutant that can seriously damage human health. Most mercury pollution enters the atmosphere from burning coal and other fossil fuels and

from industrial activity, but where does it all go? Zaferani *et al.* analyzed biogenic siliceous sediments (diatom ooze) from off the coast of Antarctica and found that they contained surprisingly large amounts of mercury. The results suggest that as much as 25% of mercury emissions over the past 150 years could be trapped in sediments like these, revealing the important role that the marine biological pump may play in the global mercury cycle. —HJS

Science, this issue p. 797

ARTIFICIAL INTELLIGENCE

An ethical way forward for AI

Artificial intelligence (AI) is becoming prevalent in everyday life. Within the next 5 years, an estimated 55% of households worldwide are expected to own a voice assistant. Furthermore, medical diagnostics, cybersecurity, and other applications are increasingly relying on AI. In a Perspective, Taddeo and Floridi discuss the ethical implications of these developments, particularly regarding the invisible influence on humans and the need to protect human self-determination. Initiatives around the world are beginning to develop fundamental ethical principles to inform the design, regulation, and use of AI. Ethical regulation of AI must include foresight methodologies that help to identify risks and avoid unwanted consequences to ensure that AI benefits and respects individuals and societies. —JFU

Science, this issue p. 751

GPCR SIGNALING

Determining signaling bias

G protein-coupled receptors (GPCRs) can mediate signaling through the G protein or β -arrestin pathways. Drugs that

selectively activate one of these pathways provide effective treatment without side effects. GPCRs must undergo phosphorylation mediated by the GPCR kinase (GRK) family of kinases to recruit β -arrestins. Choi *et al.* characterized a mutant GPCR with G protein–biased signaling. This type of signaling was driven more by an inability to recruit GRKs than by an inability to couple to β -arrestins, which may have implications for the design of biased drugs. —JFF

Sci. Signal. **11**, eaar7084 (2018).

PHYSIOLOGY

How sleep loss leads to weight gain

Chronic sleep loss can have negative health effects, including weight gain and type 2 diabetes. Underlying molecular processes in key metabolic tissues are thought to be to blame. Cedernaes *et al.* compared molecular changes such as DNA methylation in fat and skeletal muscle tissue samples taken from 15 young Caucasian males after a night of sleep loss and after a normal night's sleep. The two tissue types responded very differently. In muscle, sleep loss enhanced skeletal muscle breakdown by down-regulating a metabolic pathway. However, the same pathway was up-regulated in fat tissue after disrupted sleep. Thus, sleep loss may reprogram fat tissue to increase fat storage. —PJB

Sci. Adv. 10.1126/sciadv.aar8590 (2018).

HIV

Sweetening up antibodies

Certain individuals infected with HIV, so-called neutralizers, generate broadly neutralizing antibodies more efficiently than non-neutralizers. Lofano *et al.* compared HIV-specific antibodies isolated from neutralizers and non-neutralizers. They found sialylation of the Fc domain to be higher in neutralizers. The authors generated sialylated and nonsialylated isoforms of an HIV gp120-specific antibody.

Sialylation enhanced the deposition of antigen in B cell follicles in a complement-dependent manner. Besides stressing the importance of the Fc domain in regulating antibody functions, the study also highlights the role of the complement pathway in driving humoral immunity. —AB

Sci. Immunol. **3**, eaat7796 (2018).

MOLECULAR BIOLOGY

An additional cell cycle checkpoint

Cell division is controlled by checkpoints that regulate the temporal order of the cell cycle phases, including the G₁/S, G₂/M, and metaphase/anaphase transitions. Yet there are no known control mechanisms for a fourth fundamental transition—the S/G₂ transition. Saldivar *et al.* report a switchlike control mechanism that regulates the S/G₂ transition. The checkpoint kinase ATR senses ongoing DNA replication in S phase and represses the mitotic transcriptional network, ensuring that DNA replication in S phase is completed before mitosis. —SYM

Science, this issue p. 806

REVIEW SUMMARY

HEALTH CARE

The promise and peril of universal health care

David E. Bloom*, Alexander Khoury, Ramnath Subbaraman

BACKGROUND: The September 1978 Alma-Ata Declaration is a landmark event in the history of global health. The declaration raised awareness of “health for all” as a universal human right, whose fulfillment reduces human misery and suffering, advances equality, and safeguards human dignity. It also recognized economic and social development and international security as not only causes, but also consequences, of better health. In addition, it highlighted the power of primary health care and international co-operation to advance the protection and promotion of health in resource-constrained settings.

Building on the achievement of Alma-Ata and gaining further traction from the Millennium Development Goals and the Sustainable Development Goals set by the United Nations, universal health care (UHC) has emerged in recent years as a central imperative of the World Health Organization (WHO), the United Nations and most of its member states, and much of civil society. UHC characterizes national health systems in which all individuals can access quality health services without individual or familial financial hardship. More broadly, UHC covers social systems that provide medical and nonmedical services and infrastructure that are vital to promoting public health.

ADVANCES: Although there are numerous articulations of the UHC agenda, the WHO and World Bank offer a relatively simple UHC service-coverage index that is useful for inter-country comparison. This index focuses on four categories of health indicators: reproductive, maternal, and child health; infectious disease control; noncommunicable diseases; and service capacity and access. Comparison of UHC index values for 129 countries reveals that country index scores are positively correlated with income per capita, though there is considerable variation in scores among countries with similar incomes. These varia-

tions presumably reflect differentials in income inequality, commitment to public health infrastructure, and the quality and reach of human resources for health. The WHO and World Bank also offer multiple measures of health spending-related financial hardship in assessing UHC, which do not increase monotonically with increasing income, health spending per capita, or coverage of health services. Rather, catastrophic health expenditures



Bangkok, Thailand. A nurse keeps records of a blood donor at the Thai National Blood Centre on 14 June 2017.

tend to be lower in countries that channel health spending through public social security or insurance programs, rather than private insurance schemes.

OUTLOOK: The financial cost of massively expanding access to health care globally is a formidable barrier to achieving UHC. For example, the Disease Control Priorities Network estimates that low- and lower-middle-income countries would, on average, need to raise their respective annual per capita health expenditures by U.S.\$53 and U.S.\$61 per person to achieve coverage with the essential UHC package of 218 core interventions, a sizable burden in relation to average expenditure increases in recent years. Wealthy industrial countries are much further

along the path to achieving UHC, though they also face challenges involving rising costs of new health care technologies and the growing share of their populations at the older (and more health care-intensive) ages.

Technically and economically efficient approaches to the achievement of UHC may include the use of electronic medical records, telemedicine systems, digital monitors for drug adherence, and clinical decision-support ap-

ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aat9644>

lications; expansion of the quantity and quality of human resources for health at the physician, nurse, and community health worker levels; improvements in inventory systems and supply chains for the delivery of vaccines, drugs, diagnostics, and medical devices; screening for risk factors and early signs of disease; and focusing on the often neglected domains of surgical care, reproductive health, and mental health. Also key will

be efforts to ensure universal access to proven public health interventions that address social and environmental determinants of health, such as health education campaigns; access to safe water; regulation of excessive sugar and salt in the food supply; control of tobacco and the unsafe consumption of alcohol; road traffic safety; walkable city designs; expanding enrollment in high-quality primary and secondary schools; and more equitable distributions of income and wealth.

Achieving UHC is an ambitious aspiration and a powerful indicator of human progress. Fortunately, it may be expected to deliver myriad health, economic, and social welfare benefits along the way, helping to mobilize the substantial political and

financial resources needed for its continued future expansion. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: dbloom@hsph.harvard.edu (D.E.B.)
This is an open-access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Cite this article as D. E. Bloom *et al.*, *Science* **361**, eaat9644 (2018). DOI: [10.1126/science.aat9644](https://doi.org/10.1126/science.aat9644)



TOMORROW'S EARTH

Read more articles online at scim.ag/TomorrowsEarth

REVIEW

HEALTH CARE

The promise and peril of universal health care

David E. Bloom^{1*}, Alexander Khoury¹, Ramnath Subbaraman²

Universal health care (UHC) is garnering growing support throughout the world, a reflection of social and economic progress and of the recognition that population health is both an indicator and an instrument of national development. Substantial human and financial resources will be required to achieve UHC in any of the various ways it has been conceived and defined. Progress toward achieving UHC will be aided by new technologies, a willingness to shift medical tasks from highly trained to appropriately well-trained personnel, a judicious balance between the quantity and quality of health care services, and resource allocation decisions that acknowledge the important role of public health interventions and nonmedical influences on population health.

Universal health care (UHC) characterizes national health systems wherein all individuals can access quality health services without individual or familial financial hardship. More broadly, UHC covers social systems that provide medical and nonmedical services and infrastructure that are vital to promoting public health.

The notion of UHC dates to Otto von Bismarck, who established the world's first national social health insurance system in Germany in 1883 (1). More recently, the September 1978 Alma-Ata Declaration raised global awareness of "health for all" as a universal human right and of the power of primary health care to advance its achievement (2). During the 20th century, many industrialized countries extended UHC to their citizens. Although progress in expanding UHC slowed in the 1980s—mainly because of economic slowdowns, fiscal stress, and structural adjustment programs—achieving UHC in all countries is currently among the central imperatives of the World Health Organization (WHO), the United Nations (UN) and most of its member states, and much of civil society.

As the Alma-Ata Declaration's 40th anniversary approaches, we examine the rationale, progress, consequences, and prospects for achieving UHC globally. We first explore the rationale for UHC, the scope of what UHC encompasses, and its operational definitions at the international level. We then report statistics on current measures of UHC attainment, highlighting patterns by country income level. We go on to review evidence on three key premises of UHC: that it promotes longer, healthier lives; that it does so efficiently;

and that it confers social, economic, and political benefits above and beyond the utilitarian value of living healthier lives. Finally, we discuss prospects for further expansion of UHC.

We argue that UHC has considerable potential to improve the trajectory of human progress. To achieve UHC, however, governments and the public health community will have to mobilize substantial human, financial, and technological resources and avoid pitfalls in implementation.

Rationale and scope

Four sets of arguments are commonly advanced in support of UHC. The first set appeals to ethics and morality and the notion that safeguarding everyone's physical and mental health is just, fair, and consistent with principles of right conduct and distributive justice. The second argument, rooted in international law, relates to the acceptance of health as a fundamental human right (3). The third set of arguments is pragmatic, relating to the observation that healthy populations tend to be more socially cohesive and politically stable. The final set of arguments is economic in nature: UHC corrects health-related market failures, such as those related to the social benefits of disease prevention among individuals, and good health may promote economic well-being not just among healthy individuals but also at the macroeconomic level (4). These economic arguments are bolstered by evidence that committing resources to health care is associated with a high return on investment, rivaling, or even surpassing, other high-return investments like those in primary and secondary education (5–8).

Although there is a strong rationale for the possible benefits of UHC, there are also numerous challenges to its realization. A central challenge preceding any realization of UHC is defining its scope and boundaries. Although precise definitions of UHC vary widely among sources, the WHO's definition is a typical formulation of the

concept as a system in which "all individuals and communities receive the health services they need without suffering financial hardship. It includes the full spectrum of essential, quality health services, from health promotion to prevention, treatment, rehabilitation, and palliative care" (9).

This definition highlights many of the ambiguities involved in conceptualizing UHC. What levels of reduced mortality risk, increases to quality of living, or other thresholds must be crossed before a health service is considered needed or essential? Should financial hardship be defined by the amount of money spent relative to income, the amount of income that households retain after health spending, or some other criteria? Should these criteria shift or remain constant across settings? Given this definition's emphasis on health services, does UHC also imply a commitment to addressing social and environmental health determinants beyond the traditional purview of health service delivery?

As discussed below, the answer to this last question may have considerable implications for UHC's effectiveness in improving health outcomes. It is widely accepted that most health outcomes are associated with social and environmental factors, including wealth, income inequality, discrimination, education, occupation, diet, substance use, violence and conflict, air pollution, and water and sanitation access (10, 11). Addressing these factors is central to emerging public health agendas such as One Health (which views human, animal, and environmental health holistically) and Planetary Health (which focuses on the economic and social systems that shape human and environmental health). Deficiencies in the availability and quality of medical services are important but, nonetheless, contribute less to premature mortality than these nonmedical determinants (12). But even though UHC definitions that address nonmedical health determinants have greater potential to improve health outcomes, operationalizing a UHC agenda that addresses these determinants would require wide-ranging interventions in sectors outside of health care, which may be more politically, socially, and technically challenging.

Given the ambiguities in defining UHC, there are several possible approaches to put the concept into practice. These approaches vary according to intended use, such as making comparisons across countries, tracking progress over time, or delineating a roadmap for achieving UHC. The WHO and World Bank offer a relatively simple UHC service-coverage index (hereafter, "the WHO–World Bank index"), which is useful for inter-country comparisons. They define this index in terms of 16 indicators, grouped into four categories: reproductive, maternal, newborn, and child health; infectious-disease control; noncommunicable diseases; and service capacity and access (9).

This relatively small number of indicators allows 129 countries to be included in the UHC service-coverage index. The indicators are meant to serve as a proxy for the overall coverage of the health care system, which should ideally provide many more health services than those represented.

¹Department of Global Health and Population, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA.

²Center for Global Public Health and the Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA 02111, USA.

*Corresponding author. Email: dbloom@hsph.harvard.edu (D.E.B.)

However, although measurement of a handful of tracer conditions and services has often been used as a proxy for overall health system quality, many public health experts have concerns that only indicators that get measured actually get implemented in practice. In addition, the health services included in the index are fairly basic—in terms of the medical conditions covered, skill levels required by health care personnel, and technological capacity required—limiting this metric's value for comparing high-income countries with well-funded health systems. Some of the indicators, such as access to insecticide-treated bed nets for malaria prevention, have minimal relevance in most high-income countries. In addition, these indicators do not comprehensively capture many of the high-burden diseases that could be successfully addressed with health services in high-income countries, such as treatment for most types of cancer.

The WHO and World Bank also describe multiple approaches for measuring health spending-related financial hardship in assessing UHC. They suggest two thresholds for annual health spending—equal to 10 and 25% of total household expenditures—as alternative metrics for routinely measuring catastrophic health spending, which refers to out-of-pocket expenses exceeding a household's ability to pay without imposing considerable financial hardship. Two additional metrics aim to more directly assess impoverishment resulting from health expenditures, by measuring the percentage of households whose average daily nonhealth consumption expenditures would have placed its members above the U.S.\$1.90 and U.S.\$3.20 per capita poverty lines but for the household's spending on health care (12). Given the very low thresholds for impoverishing health expenditures, these metrics are primarily relevant in low- and middle-income countries (LMICs).

More comprehensive UHC priority descriptions exist. For example, the Disease Control Priorities (DCP) Network has compiled 218 distinct cost-effective interventions, which they argue should form a standard of essential services for LMICs because they address a substantial burden of disease. Unlike the indicators in the WHO–World Bank index, more than one-third of the DCP Network's essential interventions—including tobacco taxes, air pollution reduction, and road safety improvements—focus on broader social or environmental determinants and would require non-health care sector involvement (13). A subset of 108 interventions, termed the highest-priority package, avert death or disability while also scoring highly on a financial risk protection index. Comprehensive data are not available on population coverage for many interventions included in the DCP Network's UHC package, limiting its use in making comparisons among countries. Measurement and inclusion of many of these evidence-based services should be considered in future iterations of the global UHC agenda.

Progress toward achieving UHC

Despite the limitations of the WHO–World Bank index, analysis of its scores reveals some distinct patterns in UHC coverage. Country UHC index scores are positively correlated with the natural logarithm of gross domestic product (GDP) per capita (Fig. 1). Although this finding is consistent with models suggesting that spending on health and health care

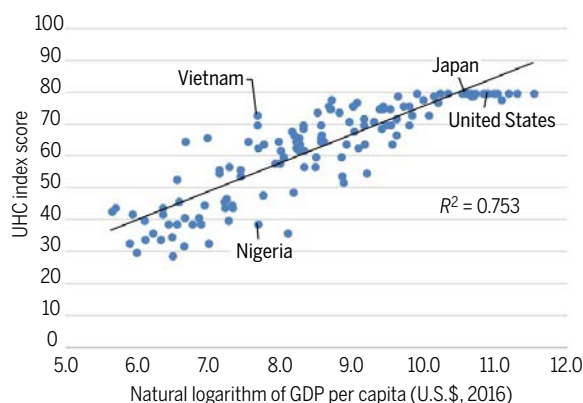


Fig. 1. WHO–World Bank UHC index score versus the natural logarithm of GDP per capita. The plot captures absolute changes in UHC index scores (maximum of 80) relative to percentage changes in GDP per capita. Source: UHC index scores from World Bank (2017) (9) and GDP per capita from World Bank (2018) (15). R^2 , coefficient of determination.

coverage increase with rising income levels (14), it is also likely that the higher country incomes are, at least in part, the result of better health care coverage and health (4). Disparities in UHC service coverage by income level are even more apparent when looking at groups of countries together: The average service-coverage score for low-income countries is roughly half that of high-income countries (Table 1). Sub-Saharan Africa and South Asia feature the lowest index scores, whereas the Latin American and Caribbean and the East Asia and Pacific regions have index scores comparable to those in North America and in Europe and Central Asia.

Also notable are the instances of similar-income countries having highly disparate index scores. For example, Nigeria and Vietnam both have per capita GDPs around U.S.\$2200, but Vietnam's UHC index score is 34 points higher than Nigeria's. This reflects the fact that Vietnam outperforms Nigeria on several indicators, including reported rates of three-dose diphtheria-tetanus-pertussis infant vaccination coverage (94 versus 42%), births attended by skilled professionals (94 versus 35%), and households with access to basic sanitation (78 versus 32%). Dissimilar income distributions in the two countries offer a plausible partial explanation for the coverage discrepancies. An estimated 78% of Nigeria's population lives on less than U.S.\$3.20 per day in 2017 dollars, compared with only 32% of Vietnam's population (15, 16). Poverty imposes constraints on accessing health services, particularly in LMICs (17). Furthermore, less-comprehensive health-

service coverage reinforces poverty by failing to protect individuals from illnesses that have high treatment costs or that limit their ability to work or learn (18).

Unlike the association between UHC service coverage and GDP per capita (Fig. 1), protection from catastrophic health expenditures is not clearly correlated with GDP per capita. In aggregate, middle-income countries have higher rates of catastrophic health expenditures than low- and high-income countries (Table 1). However, the variation in catastrophic expenditure rates within these income groups is greater than the variation among them. Furthermore, protection from catastrophic health expenditures does not systematically improve with increasing UHC service-coverage index score or with increasing percentage of GDP spent on health care (19). Thus, protection from health care-related financial ruin does not directly follow from GDP growth, improved essential health service coverage, or increased total health care spending. Rather, catastrophic health expenditures may be associated with the pathways through which health care spending occurs. Countries in which much of health spending is prepaid through public social security or insurance programs tend to have lower catastrophic health expenditure rates than countries that mostly rely on private insurance schemes (19).

Even among countries attaining the maximum UHC index score of 80, there is substantial heterogeneity in health outcomes, health spending, and the proportion of the population protected from catastrophic health spending. Comparing the UHC records of two high-income countries with perfect index scores, the United States and Japan, illustrates these disparities.

The United States is the only high-income country that does not explicitly provide UHC for its citizens, although its relative expenditures on health care—15% of GDP in 2008 and 17% of GDP in 2017—are much higher than those of any other Organisation for Economic Cooperation and Development country (20). Unlike the United States, Japan expanded health insurance coverage to its entire population in 1961. This change coincided with a massive improvement in the health of Japan's population, which, by 1983, had the highest life expectancy of any country (and also now, at 84 years) (15). Japan's health system has been lauded for its role in promoting a world-leading level of population health and for maintaining relatively low health care costs historically. However, these health expenditures have risen from 8% of GDP in 2008 to 11% in 2017 (20), and the Japanese health care system must adapt to a continually increasing elder share of the population while constrained by an economy that has performed relatively poorly since the 1990s.

Consequences of expanding UHC

The premise that UHC could lead to longer, healthier lives has a strong underlying rationale.

For most indicators in the WHO–World Bank index, achieving high coverage could benefit individuals through reduced disability, increased longevity, improved nutritional status, increased economic productivity, or decreased health-related financial hardship (Table 2).

Increased coverage of services can also have a population-level health impact, especially for leading infectious causes of death, such as tuberculosis (TB), HIV/AIDS, and malaria. For these diseases, early treatment of affected individuals can terminate the chain of transmission, thereby reducing disease incidence. For example, over a 7-year time period, HIV-uninfected individuals living in areas with high antiretroviral therapy coverage in KwaZulu Natal, South Africa, were 38% less likely to acquire HIV than those in areas with low coverage (21). Similarly, in China during the 1990s, TB prevalence declined only in provinces where the directly observed therapy short-course (DOTS) strategy—which involves provision of free or subsidized TB testing and treatment—was rolled out with high coverage (22).

Similarly, expanding vaccine coverage through the UHC agenda—especially for leading causes of child mortality such as *Streptococcus pneumoniae*, *Haemophilus influenzae*, and rotavirus—would have a population-level health impact in a highly cost-effective manner. The full societal benefits of disease prevention through vaccination include increased schooling and labor productivity, slowing of the pace at which antimicrobial resistance develops, and reductions in health and economic risk, all magnified by the value of improved health outcomes among nonvaccinated community members owing to herd effects (23).

Regarding the potential impacts of UHC on both health and financial hardship, some compelling evidence is found in the Oregon Health Insurance Experiment (24). In 2008, the U.S. state of Oregon randomly selected about 30,000 individuals to be eligible to apply for Medicaid from among the roughly 90,000 who had expressed interest in applying to the newly expanded program providing low-cost health coverage for low-income adults. Through comparisons of individuals who were not selected to those who applied and were accepted, researchers found that receiving Medicaid virtually eliminated catastrophic medical spending, reduced medical debt, increased use of preventive medical care, reduced depressive symptoms, and improved subjective perception of overall health status. Measures of physical health—including control of high blood pressure, high cholesterol, and diabetes—did not significantly improve among individuals who received Medicaid; however, the 2-year follow-up time for individuals may have been too short to detect meaningful improvements in these outcome indicators.

Other literature on the impact of increased coverage and density of primary care and hospital-based services on health outcomes is generally of weaker quality. Nevertheless, examples from Costa Rica and Cuba suggest a strong association between the universal expansion of public sector primary care services and rapid reductions in child and adult mortality and increases in life expectancy (25, 26). In addition, a systematic review highlights the consistency, across a variety of LMIC contexts, of the positive association between large-scale primary care ini-

tiatives and lower child mortality (27). In other LMIC settings, increased hospital access is associated with reduced maternal mortality (28). In high-income countries, a higher density of primary care providers is associated with lower all-cause mortality (29).

Beyond improving health, expanding UHC could potentially promote economic well-being, reduce economic inequalities, and bolster social and political stability (5, 30). Improving population health could accelerate economic growth by improving labor productivity, school attendance, educational attainment, cognitive function, capital accumulation, and fertility control (31, 32). Rigorous microeconomic evidence supports the impact of health improvements on individual or household economic circumstances. Interventions with demonstrated effects on education and earnings include iodine supplementation (33), iron supplementation (34), deworming (35, 36), and malaria eradication campaigns (37). These benefits may also have an appreciable macroeconomic impact (38): On average, a 10-year life expectancy gain is associated with up to a 1% increase in annual income per capita growth (5).

The impact of better health on economic growth may be particularly powerful in LMICs, where children, adolescents, and prime-age adults are the chief beneficiaries of health gains, leading to improvements in productivity across the life course (37). Ensuring access to basic health care, especially for the prevention and treatment of infectious diseases, may be essential for escaping poverty traps in settings where extreme poverty has historically been persistent (39, 40). But benefits of health on economic growth are also

Table 1. Population, income, health expenditure, and UHC index score by income group and geographic region. Figures are weighted according to population size. Source: World Bank (2018) (15), with UHC service-coverage index scores and catastrophic health expenditure data from World Bank (2017) (9). All data are for 2016, except for the health expenditure data, which are for 2015, and the catastrophic health spending data, which are for 2010.

	Number of countries	Percentage of world population (%)*	Income per capita (current U.S.\$)	Health expenditure per capita (current U.S.\$)	Health expenditure as a percentage of GDP (%)	Mean UHC index score (range)	Percentage of households experiencing catastrophic health spending (%)†
World	130	100	10,192	1,002	9.8	63 (29 to 80)	11.7
Income group							
Low income	21	8	616	35	5.7	39 (29 to 53)	8.1
Lower-middle income	39	42	2,078	83	4.0	53 (33 to 73)	12.4
Upper-middle income	35	36	7,994	470	5.9	74 (52 to 78)	13.8
High income	35	15	40,826	5,050	12.4	79 (64 to 80)	7.2
Region							
Sub-Saharan Africa	35	14	1,467	85	5.8	42 (29 to 67)	10.3
South Asia	7	26	1,638	58	3.5	53 (34 to 62)	13.5
Middle East and North Africa	11	4	7,200	416	5.8	64 (39 to 80)	13.4
East Asia and Pacific	14	32	9,783	626	6.4	72 (47 to 80)	12.9
Europe and Central Asia	44	12	22,238	2,089	9.4	72 (54 to 80)	7.0
Latin America and Caribbean	17	8	8,342	637	7.6	75 (57 to 79)	14.8
North America	2	5	56,102	9,031	16.1	80 (80 to 80)	4.6

*Percentage of world population refers to the entire income group or region, not just the countries included in the sample. †Catastrophic health spending refers to the proportion of individuals in the population who live in households that spend >10% of their consumption expenditure on out-of-pocket health care costs (9).

manifest in high-income countries, where gains in longevity tend to accrue disproportionately to older adults. The social and economic value resulting from these gains in longevity for older adults may not be well represented in national GDP because the value created is often related to the enabling effect of health on nonmarket activities such as child-rearing, caretaking of other individuals, and community volunteer work (41).

Expanding UHC also reduces health disparities because poor members of society are less likely to receive adequate health care than wealthier individuals where UHC systems are lacking. Increased access to primary care is associated with reduced wealth- and race-based mortality disparities in both LMICs (27) and high-income countries (29). As noted, decreasing health inequality may also reduce income, wealth, and education disparities. As with investments in education, expansion of health care coverage is one of the rare policies that simultaneously promotes equitable distribution of income while also increasing economic growth (6, 42). Reducing disparities through improved public health and social welfare systems may help to minimize the risk of political and social instability, though empirical evidence of this association is not especially robust (43). Through these various pathways, UHC serves important functions that support a healthy, prosperous, and cohesive society.

Although the potential benefits of UHC are numerous, possible pitfalls in implementation could undermine its impact and prevent UHC from fulfilling its promise. Rapid scale-up of UHC without sufficient concern for the quality of implementation could have unintended

adverse consequences, as delivery of health services will not be effective in improving health outcomes if the delivered care is not of reasonable quality. Deficiencies in quality of care such as medical errors, spread of infection in health care settings, and poor retention of patients across sequential steps of care (also known as the cascade of care) could undermine the benefits of expanded service coverage. Even though existing UHC frameworks allude to this problem, quality-related indicators can be hard to measure, and achieving high quality of care will be especially challenging with large-scale expansion of coverage (44).

The recent history of TB care delivery illustrates limitations of focusing on coverage of health services without ensuring that the services offered are of sufficient quality to be effective. In 1991, the World Health Assembly adopted the DOTS strategy, which included comprehensive coverage of free or subsidized TB testing and treatment as a key objective. Over the next two decades, high-burden countries such as India and China achieved high DOTS coverage nationally, leading to reductions in disease prevalence or TB-related mortality (22, 45). However, despite high global DOTS coverage, TB incidence is declining slowly (<1.5% per year); the disease remains the leading infectious cause of death, resulting in nearly 1.7 million deaths annually, one-third of which occur in India. Poor quality of care may in part explain these disappointing public health outcomes (46). For example, in India, considerable numbers of patients are lost across sequential steps of the care cascade; as a result, only about 39% of prevalent TB patients

were estimated to have achieved an optimal outcome in the government program in 2013 (Fig. 2) (47). Similarly, in Rwanda, improved rates of maternal institutional delivery have not translated into reductions in newborn mortality, likely owing to gaps in care quality (48).

As these examples suggest, poorly functioning health systems are a central challenge to realizing the benefits of UHC. Health systems in LMICs commonly suffer from a variety of weaknesses, including absenteeism and insufficient training among health care workers, mistreatment of patients by health care workers, corruption, poorly functioning inventory systems and supply chains, electricity cuts and outages, and lack of clean water. These shortcomings in health care delivery often reflect higher-level problems in governance and market failures. Achieving UHC will therefore require innovations in the structure and operation of health systems to ensure that rapid expansion in coverage is not undermined by shortcomings in delivery and quality of care.

With regard to the scope of UHC, it is entirely appropriate for countries to prioritize different health interventions in their UHC agendas to address local needs and constraints. It is also reasonable to expect the number of health services considered essential in each setting to undergo progressive expansion over time to reflect changing resource availability and to address new or emerging health concerns. For example, the WHO–World Bank index—perhaps the most prominent articulation of the UHC agenda—mostly focuses on health service coverage for conditions that have been long-standing global health priorities, such as maternal health, HIV,

Table 2. UHC essential services as defined by the WHO and World Bank and the rationale for their impact on health and social outcomes.	
Essential health service	Selected evidence for beneficial health, social, or economic outcomes
Family planning	Decreased maternal mortality (64), improved economic growth associated with reduced fertility (the “demographic dividend”) (31)
Antenatal and delivery care	Reduced infant and maternal mortality (65)
Child immunization	Reduced mortality for children less than 5 years old, improved educational attainment and economic productivity (23)
Pneumonia care	Reduced pneumonia-related morbidity and mortality (66)
Tuberculosis treatment	Improved tuberculosis treatment success, mortality, and prevalence (22, 45, 67)
HIV antiretroviral therapy	Increased life expectancy (68), reduced HIV transmission (21)
Insecticide-treated bed nets for malaria prevention	Reduced malaria episodes and child mortality (69)
Access to basic sanitation	Reduced mortality and stunting of children less than 5 years old (70)
Prevention and treatment of elevated blood pressure	Reduced cardiovascular and all-cause mortality in individuals more than 60 years old (71)
Prevention and treatment of elevated blood sugar	Reduced microvascular complications of diabetes, including kidney failure, loss of vision, and nerve damage (72)
Cervical cancer screening	Reduced cervical cancer incidence and mortality (73)
Tobacco (non)smoking	Reduction in lung cancer, obstructive pulmonary disease, cardiovascular, and all-cause mortality (74)
Basic hospital access	Lower maternal mortality (28), lower mortality from life-threatening emergencies (75)
Health care worker density	Reduced all-cause child and adult mortality and reduced health disparities among populations (27, 29)
Access to essential medicines	Reduction in the proportion of the population experiencing catastrophic health care costs (76, 77)
Compliance with international health regulations (health security)	Early detection of disease outbreaks (78), with benefits and limitations highlighted by the 2013–2014 Ebola outbreak (79)

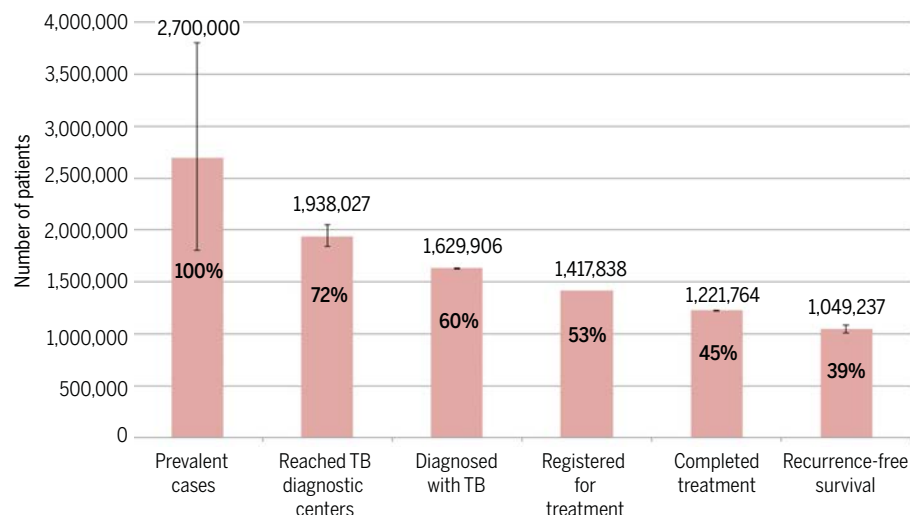


Fig. 2. Cascade of care for patients with any form of TB in India in 2013. Patient losses at each stage of care represent shortcomings in quality of care that undermine the effectiveness of TB services, despite a high level of population coverage. Source: Subbaraman *et al.* (47).

and TB. The index does not emphasize measurement of service coverage for other conditions that contribute substantially to the global burden of disability or death, such as depression and anxiety (the leading causes of disability globally) and conditions that require basic surgical care (inaccessible to about 5 billion people) (49, 50).

In addition, the WHO–World Bank index includes a measure of access to essential medicines but does not cover access to essential diagnostic tests, which are crucial to address population-level threats to health, such as antimicrobial resistance (51). Rising rates of antimicrobial resistance could be a major unintended consequence of UHC if increasing health care coverage does not go hand-in-hand with expanded access to diagnostic tests that facilitate judicious use of antibiotics. As suggested by these examples, if countries adhere to an overly narrow set of UHC priorities, they could miss out on opportunities to address conditions for which there is a dearth of health care providers and institutional capacity in LMICs.

The relatively limited inclusion of measures of nonmedical health determinants in most UHC frameworks represents another, more fundamental, limitation in scope. The WHO–World Bank index focuses on assessing delivery of medical services, with the exception of access to adequate sanitation and insecticide-treated bed nets. The UHC scope thus defined largely avoids the question of ensuring universal access to many public health interventions that could lead to healthier lives—including health education campaigns, in-home piped water supplies, regulation of excessive sugar and salt in the food supply, tobacco control, road traffic safety, construction of walkable cities, high-quality primary and secondary education, and equitable distribution of wealth.

Two examples illustrate the limitations of a UHC approach that avoids addressing underlying nonmedical health determinants. In the

United States, the dramatic rise in mortality among middle-aged white people in recent years occurred during a time of increasing health insurance coverage in the general population. These “deaths of despair”—largely attributable to mortality from substance use, suicide, and injuries—are thought to be driven by social determinants, such as lack of employment opportunities for blue-collar workers and increasing wealth inequality (52).

Another example is stunting owing to chronic child undernutrition, which is associated with poor health outcomes, cognitive development, and educational attainment. Most factors that contribute to stunting—poverty, lack of maternal education, poor maternal nutrition, lack of dietary diversity, and lack of sanitation—reflect failures to address nonmedical health determinants (53, 54). In India, which accounts for 40% of the world’s stunted children, social inequalities such as gender and caste discrimination drive deficiencies in maternal education and sanitation access, thereby impeding progress in reducing stunting (53, 54). As these examples suggest, UHC that narrowly focuses on health service delivery alone is necessary, but insufficient, to bring about wide-ranging health and social benefits. UHC will be implemented within the wider context of the Sustainable Development Goals (SDG) set by the UN, which includes targets related to some of these nonmedical determinants; however, embedding these SDG targets within a UHC-related public health framework could shape the approach and intensity with which these targets are achieved.

Prospects

The financial cost of massively expanding access to health care globally is a formidable barrier to achieving UHC. The cost of attaining UHC partly hinges on a population’s existing health, which is influenced by factors such as age structure, levels of physical activity, pollution, water and sanita-

tion infrastructure, vaccination coverage, and diet. Using their broad operationalization of UHC described above, the DCP Network estimates that low- and lower-middle-income countries would, on average, need to raise their respective annual per capita health expenditures by U.S.\$26 and U.S.\$31 per person to achieve coverage with the highest priority package (108 core interventions); achieving coverage with the essential UHC package (218 core interventions) would require an annual spending increase of U.S.\$53 and U.S.\$61 per person on average (13).

However, the authors caution that achievement of even the essential UHC package would not be sufficient to reach the SDG target of reducing deaths of individuals less than 70 years old by 40% by 2030. Achieving the highest-priority and essential UHC package would accomplish around half and two-thirds of this goal, respectively (13). Presumably, covering the essential health services in the WHO–World Bank index would require lower per capita health expenditure but would be expected to fall even shorter in reaching the SDG targets.

As Table 3 shows, the health expenditure growth needed to achieve essential UHC in LMICs by 2030 is comparable to the rate of health spending increases that these countries experienced in recent years. However, these raw estimates of recent growth in health spending could paint an overly optimistic picture. A recent study from the Global Burden of Disease Health Financing Collaborator Network uses data from a similar period (1995–2015) and an ensemble of models that include covariates associated with GDP and health expenditure growth (such as fertility rates and mean years of education) to project health expenditure growth through 2030 (55). The Network projects that the difference between the number of individuals covered by UHC in the “worst-case” and “best-case” health financing scenarios would be about 871 million people (55).

Given the sizable expenditure increases necessary to achieve UHC, rolling out UHC programs in stages will be necessary. The Lancet Commission on Investing in Health advocates a “progressive universalist” approach to funding these efforts, whereby selected health services are offered broadly and affordably to all citizens by the government, even if this necessitates offering a smaller package of interventions. The authors argue that this approach is more efficient and equitable than a system that covers more interventions but necessitates higher out-of-pocket expenditures or restricts coverage to fewer individuals (56).

In light of expected health expenditure increases required to achieve UHC, physicians and public health practitioners may have to radically rethink strategies for health care delivery to simultaneously improve efficiency and health outcomes. For example, lack of trained health care personnel, especially in LMICs, is arguably the most serious hurdle to scaling up UHC (57). In many countries—such as India, Bangladesh, and Uganda—most health care personnel are informal providers who lack formal medical training (58).

Informal providers are often assumed to deliver low-quality care; however, a recent randomized trial found that intensive training sessions with these providers can improve the quality of care that they deliver to a level that is, in some cases, on par with formal providers (59). Careful and constructive engagement with these informal providers may therefore be one strategy for bridging the substantial health care workforce gaps that threaten to undermine progress toward UHC in LMICs. Stemming outmigration of physicians from LMICs through bonding schemes (such as conditional scholarships) or enforcement of ethical recruitment policies in high-income countries may also help to reduce health care worker shortages (60).

Programs to recruit and train community health workers (CHWs) offer another, more widely accepted, strategy for expanding the health care workforce and increasing the coverage and effectiveness of primary health care. Growing evidence suggests that these programs can contribute to improved outcomes in child nutrition, maternal health, HIV, and TB (61). Moreover, CHW programs could potentially expand the reach of health care provision to the household level. This would be especially beneficial in the context of a rapidly increasing global burden of chronic disease. Primary and secondary prevention of chronic diseases requires early screening for risk factors and lifelong treatment of those risks (e.g., medications for hypertension), and many chronic diseases and risk factors cluster within households

(62). CHWs may also have an important role in tracking newborns at the household level from the first to the last vaccination during infancy. By extending screening, monitoring, and treatment of medical conditions to the household level, CHW programs could have substantial effects on preventing disease, increasing rates of health screening, and improving treatment outcomes, thus improving UHC coverage, efficiency, and impact.

Integrating innovative technologies into health systems—including electronic medical records, clinical decision-support applications, telemedicine, digital medication-adherence technologies, and point-of-care diagnostic tests—could also facilitate UHC by improving the reach, timeliness, efficiency, and quality of clinical care and public health monitoring. These technologies could improve the quality and coverage of longitudinal clinical records, facilitate health care providers' use of evidence-based clinical care algorithms, extend access to specialized medical knowledge to rural communities, reduce time delays for diagnosis and treatment, and enable real-time monitoring of medication adherence. Artificial intelligence and machine learning have the potential to perform some tasks—such as interpreting x-rays, electrocardiograms, and electroencephalograms—that currently require highly trained and specialized health care workers.

Technological innovations will not obviate the need to dramatically increase the health care workforce in LMICs, but they could still prove

to be game changers as the global community tries to rapidly scale up health service delivery to achieve UHC. The ambitious scope of the UHC agenda may provoke physicians and public health experts to reimagine how to deliver health services. New frontline health care personnel (such as CHWs and nonhealth professionals receiving appropriate training) and innovative technologies could help to move care provision into non-traditional spaces, such as homes or workplaces, extending the existing health system's effective reach.

The bottom line

Four decades after the Alma-Ata Declaration articulated primary care for all as being a most important worldwide social goal, the global community is striving to achieve UHC with renewed interest and ambition. A central motivation of the UHC agenda is the belief that access to health care—with the goals of extending longevity, minimizing disability, and diminishing suffering—is a fundamental human right that advances equality and safeguards human dignity. Achieving UHC would represent one of the most ambitious ventures in the area of human rights, even if UHC were defined narrowly as universal delivery of essential health services. In addition, evidence suggests that well-implemented universal coverage of essential health services could improve welfare more broadly, by reducing economic inequalities, promoting economic well-being, and, perhaps, improving social and political stability.

A broader UHC conception that aims to also address the nonmedical determinants that most strongly shape human health would have even greater implications for society and would require broader social transformations. Addressing cross-cutting social and environmental determinants that contribute to ill health—such as wealth inequality; race, gender, and caste discrimination; air pollution; and lack of water and sanitation facilities—could lead the UHC agenda to intersect more closely with the human rights, One Health, and Planetary Health agendas in the coming decades.

Numerous potential pitfalls could impede UHC expansion or undermine its positive impact on health and well-being. Most challenging, perhaps, is the need to increase health financing rapidly enough to facilitate universal coverage of essential health services among LMIC populations that are simultaneously growing in size and aging. For example, for the world's less-developed regions, an increase of roughly 1 billion people is projected from 2018 through 2030, with the percentage aged 60 years or older projected to increase from 10.6 to 14.2% (63). In addition, a UHC agenda that fails to address social determinants of health could limit its impact on health outcomes. Finally, focusing too much on coverage alone, rather than on ensuring the quality of health services, could undermine UHC effectiveness. Addressing these challenges may require radical transformations in the way that health services are delivered, potentially by expanding the use of frontline health personnel

Table 3. Health expenditures needed to attain the highest-priority package (HPP) and essential UHC (EUHC) package by income. Source: Watkins *et al.* (2017) (13), with public health expenditure data and average growth (2000–2015) calculated from WHO (2018) (15).

Health expenditure metric	Low-income countries	Lower-middle-income countries
Public health expenditures per capita (U.S.\$)*	18	28
Total (and incremental†) health expenditure per capita needed for HPP (U.S.\$)	42 (26†)	58 (31†)
Total (and incremental†) health expenditure per capita needed for EUHC (U.S.\$)	76 (53†)	110 (61†)
Average annual growth rate in public health expenditures needed to achieve HPP by 2030 (%)‡	6.6	5.3
Average annual growth rate in public health expenditures needed to achieve EUHC by 2030 (%)‡	10.3	8.4
Average annual growth rate in real public health expenditures per capita 2000–2015 (%)*	9.8	9.2

*Values provided refer to government and donor health expenditures per capita in 2012 U.S.\$; Table 1 provides total health expenditures for LMICs (including private expenditures). †Incremental health expenditures per capita refers to the amount health spending per person would have to increase from current levels to support the complete package of interventions. ‡The estimated growth in public health expenditures needed to achieve HPP and EUHC assumes that all additional coverage for these packages are met through government expenditure and that all additional government health care expenditure is spent on these intervention packages.

and incorporating innovative technologies into care delivery.

Ultimately, the path to UHC and the interventions prioritized in this process will be unique to each country pursuing universal coverage. Although achieving full UHC is a daunting task, incremental steps toward fulfilling this goal also offer myriad health, economic, and social welfare benefits. Recognizing these benefits should help mobilize the resources needed for continued future expansion of UHC.

REFERENCES AND NOTES

1. T. Barnighausen, R. Sauerborn, One hundred and eighteen years of the German health insurance system: Are there any lessons for middle- and low-income countries? *Soc. Sci. Med.* **54**, 1559–1587 (2002). doi: [10.1016/S0277-9536\(01\)00137-X](https://doi.org/10.1016/S0277-9536(01)00137-X); pmid: [12061488](https://pubmed.ncbi.nlm.nih.gov/12061488/)
2. International Conference on Primary Health Care, "Declaration of Alma-Ata" (Alma-Ata, USSR, 1978).
3. United Nations General Assembly, "Universal Declaration of Human Rights, Article 25" (United Nations, Paris, 1948).
4. D. E. Bloom, D. Canning, The health and welfare of nations. *Science* **287**, 1207–1209 (2000). doi: [10.1126/science.287.5456.1207](https://doi.org/10.1126/science.287.5456.1207)
5. D. E. Bloom, D. Canning, J. P. Sevilla, The effect of health on economic growth: A production function approach. *World Dev.* **32**, 1–13 (2004). doi: [10.1016/j.worlddev.2003.07.002](https://doi.org/10.1016/j.worlddev.2003.07.002)
6. G. Psacharopoulos, H. A. Patrinos, "Returns to investment in education: A decennial review of the global literature" (Policy Research Working Paper no. 8402, World Bank, Washington, DC, 2018).
7. R. W. Fogel, Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy. *Am. Econ. Rev.* **84**, 369–395 (1994).
8. R. W. Fogel, in *Handbook of Population and Family Economics*, M. Rosenzweig, O. Stark, Eds. (Elsevier, 1997), vol. 1, part A, pp. 433–481.
9. World Health Organization, World Bank, "Tracking universal health coverage: 2017 Global Monitoring Report" (World Bank, Washington, DC, 2017).
10. World Health Organization Commission on Social Determinants of Health, "Closing the gap in a generation: Health equity through action on social determinants of health. Final report of the Commission on Social Determinants of Health" (World Health Organization, Geneva, 2008).
11. J. M. McGinnis, P. Williams-Russo, J. R. Knickman, The case for more active policy attention to health promotion. *Health Aff. (Millwood)* **21**, 78–93 (2002). doi: [10.1377/hlthaff.21.2.78](https://doi.org/10.1377/hlthaff.21.2.78); pmid: [11900188](https://pubmed.ncbi.nlm.nih.gov/11900188/)
12. These figures are represented in 2011 purchasing power parity-adjusted U.S. dollars.
13. D. A. Watkins, D. T. Jamison, A. Mills, R. Atun, K. Danforth, A. Glassman, S. Horton, P. Jha, M. E. Kruk, O. F. Norheim, J. Qi, A. Soucat, S. Verguet, D. Wilson, A. Alwan, in *Disease Control Priorities*, D. T. Jamison, H. Gelband, S. Horton, P. Jha, R. Laxminarayan, C. N. Mock, R. Nugent, Eds. (World Bank, ed. 3, 2017), vol. 9, pp. 43–65.
14. R. E. Hall, C. I. Jones, The value of life and the rise in health spending. *Q. J. Econ.* **122**, 39–72 (2007). doi: [10.1162/qjec.122.1.39](https://doi.org/10.1162/qjec.122.1.39)
15. World Bank, World development indicators (2018); <https://data.worldbank.org/products/wdi>.
16. These estimates reflect the most recent data available for Nigeria (from 2009) and an average of Vietnamese data from 2008 and 2010. The most recent Vietnamese estimates (from 2014) suggest that only 11% of the country's population lives on less than U.S.\$3.20 per day.
17. D. H. Peters et al., Poverty and access to health care in developing countries. *Ann. N. Y. Acad. Sci.* **1136**, 161–171 (2008). doi: [10.1196/annals.1425.011](https://doi.org/10.1196/annals.1425.011); pmid: [17954679](https://pubmed.ncbi.nlm.nih.gov/17954679/)
18. J. P. Smith, Healthy bodies and thick wallets: The dual relation between health and economic status. *J. Econ. Perspect.* **13**, 145–166 (1999). doi: [10.1257/jep.13.2.145](https://doi.org/10.1257/jep.13.2.145); pmid: [15179962](https://pubmed.ncbi.nlm.nih.gov/15179962/)
19. A. Wagstaff et al., Progress on catastrophic health spending in 133 countries: A retrospective observational study. *Lancet Glob. Health* **6**, e169–e179 (2018). doi: [10.1016/S2214-109X\(17\)30429-1](https://doi.org/10.1016/S2214-109X(17)30429-1); pmid: [29248367](https://pubmed.ncbi.nlm.nih.gov/29248367/)
20. Organisation for Economic Co-operation and Development, OECD health statistics (2018); www.oecd-ilibrary.org/
21. F. Tanser, T. Barnighausen, E. Grapsa, J. Zaidi, M. L. Newell, High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science* **339**, 966–971 (2013). doi: [10.1126/science.1228160](https://doi.org/10.1126/science.1228160); pmid: [23430656](https://pubmed.ncbi.nlm.nih.gov/23430656/)
22. L. Wang et al., Tuberculosis prevalence in China, 1990–2010; a longitudinal analysis of national survey data. *Lancet* **383**, 2057–2064 (2014). doi: [10.1016/S0140-6736\(13\)62639-2](https://doi.org/10.1016/S0140-6736(13)62639-2); pmid: [24650955](https://pubmed.ncbi.nlm.nih.gov/24650955/)
23. D. E. Bloom, V. Y. Fan, J. P. Sevilla, The broad socioeconomic benefits of vaccination. *Sci. Transl. Med.* **10**, eaaj2345 (2018). doi: [10.1126/scitranslmed.aaj2345](https://doi.org/10.1126/scitranslmed.aaj2345); pmid: [29769285](https://pubmed.ncbi.nlm.nih.gov/29769285/)
24. K. Baicker et al., The Oregon experiment—Effects of Medicaid on clinical outcomes. *N. Engl. J. Med.* **368**, 1713–1722 (2013). doi: [10.1056/NEJMs1212321](https://doi.org/10.1056/NEJMs1212321); pmid: [23635051](https://pubmed.ncbi.nlm.nih.gov/23635051/)
25. M. Pesec et al., Primary health care that works: The Costa Rican experience. *Health Aff. (Millwood)* **36**, 531–538 (2017). doi: [10.1377/hlthaff.2016.1319](https://doi.org/10.1377/hlthaff.2016.1319); pmid: [28264956](https://pubmed.ncbi.nlm.nih.gov/28264956/)
26. P. K. Drain, M. Barry, Fifty years of U.S. embargo: Cuba's health outcomes and lessons. *Science* **328**, 572–573 (2010). doi: [10.1126/science.1189680](https://doi.org/10.1126/science.1189680); pmid: [20430999](https://pubmed.ncbi.nlm.nih.gov/20430999/)
27. M. E. Kruk, D. Porignon, P. C. Rockers, W. Van Lerberghe, The contribution of primary care to health and health systems in low- and middle-income countries: A critical review of major primary care initiatives. *Soc. Sci. Med.* **70**, 904–911 (2010). doi: [10.1016/j.socscimed.2009.11.025](https://doi.org/10.1016/j.socscimed.2009.11.025); pmid: [20089341](https://pubmed.ncbi.nlm.nih.gov/20089341/)
28. C. Hanson et al., Maternal mortality and distance to facility-based obstetric care in rural southern Tanzania: A secondary analysis of cross-sectional census data in 226 000 households. *Lancet Glob. Health* **3**, e387–e395 (2015). doi: [10.1016/S2214-109X\(15\)00048-0](https://doi.org/10.1016/S2214-109X(15)00048-0); pmid: [26004775](https://pubmed.ncbi.nlm.nih.gov/26004775/)
29. L. Shi, J. Macinko, B. Starfield, R. Politzer, J. Xu, Primary care, race, and mortality in US states. *Soc. Sci. Med.* **61**, 65–75 (2005). doi: [10.1016/j.socscimed.2004.11.056](https://doi.org/10.1016/j.socscimed.2004.11.056); pmid: [15847962](https://pubmed.ncbi.nlm.nih.gov/15847962/)
30. P. Lorentzen, J. McMillan, R. Wacziarg, Death and development. *J. Econ. Growth* **13**, 81–124 (2008). doi: [10.1007/s10887-008-9029-3](https://doi.org/10.1007/s10887-008-9029-3)
31. D. E. Bloom, 7 billion and counting. *Science* **333**, 562–569 (2011). doi: [10.1126/science.1209290](https://doi.org/10.1126/science.1209290); pmid: [21798935](https://pubmed.ncbi.nlm.nih.gov/21798935/)
32. D. E. Bloom, M. Kuhn, K. Prettnner, "Health and economic growth" (Working Paper no. 153, Program on the Global Demography of Aging, Harvard University, Cambridge, MA, February 2018).
33. E. Field, O. Robles, M. Torero, Iodine deficiency and schooling attainment in Tanzania. *Am. Econ. J. Appl. Econ.* **1**, 140–169 (2009). doi: [10.1257/app.1.4.140](https://doi.org/10.1257/app.1.4.140)
34. D. Thomas, E. Frankenberg, J. Friedman, J.-P. Habicht, M. Hakim, N. Ingwersen, X. Jaswadi, N. Jones, C. McKelvey, G. Peltó, B. Sikoki, T. Seeman, J. P. Smith, C. Sumantri, W. Suriastini, S. Wilopo, "Causal effect of health on labor market outcomes: Experimental evidence" (Working Paper, University of California, Los Angeles, 2006).
35. S. Baird, J. H. Hicks, M. Kremer, E. Miguel, Worms at work: Long-run impacts of a child health investment. *Q. J. Econ.* **131**, 1637–1680 (2016). doi: [10.1093/qje/qjw022](https://doi.org/10.1093/qje/qjw022); pmid: [27818531](https://pubmed.ncbi.nlm.nih.gov/27818531/)
36. H. Bleakley, Disease and development: Evidence from hookworm eradication in the American South. *Q. J. Econ.* **122**, 73–117 (2007). doi: [10.1162/qjec.121.1.73](https://doi.org/10.1162/qjec.121.1.73); pmid: [24146438](https://pubmed.ncbi.nlm.nih.gov/24146438/)
37. H. Bleakley, Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *Am. Econ. J. Appl. Econ.* **2**, 1–45 (2010).
38. D. E. Bloom, G. Fink, in *Manson's Tropical Diseases*, J. Farrar, P. Hotez, T. Jungthans, G. Kang, Eds. (Elsevier, 2013), pp. 23–30.
39. C. N. Nonghalala et al., General ecological models for human subsistence, health and poverty. *Nat. Ecol. Evol.* **1**, 1153–1159 (2017). doi: [10.1038/s41559-017-0221-8](https://doi.org/10.1038/s41559-017-0221-8); pmid: [29046570](https://pubmed.ncbi.nlm.nih.gov/29046570/)
40. M. Cervellati, U. Sunde, Human capital formation, life expectancy, and the process of development. *Am. Econ. Rev.* **95**, 1653–1672 (2005). doi: [10.1257/000282805775014380](https://doi.org/10.1257/000282805775014380); pmid: [29125727](https://pubmed.ncbi.nlm.nih.gov/29125727/)
41. D. Dosman, J. Fast, S. A. Chapman, N. Keating, Retirement and productive activity in later life. *J. Fam. Econ. Issues* **27**, 401–419 (2006). doi: [10.1007/s10834-006-9022-y](https://doi.org/10.1007/s10834-006-9022-y)
42. J. D. Ostry, A. Berg, C. G. Tsangaridis, "Redistribution, inequality, and growth" (SDN/14/02, International Monetary Fund, Washington, DC, 2014).
43. U.S. National Intelligence Council, National intelligence estimate: The global infectious disease threat and its implications for the United States. *Environ. Change Secur. Proj. Rep.* **2000 Summer**, 33–65 (2000).
44. J. Das, L. Woskie, R. Rajbhandari, K. Abbasi, A. Jha, Rethinking assumptions about delivery of health care: Implications for universal health coverage. *BMJ* **361**, k1716 (2018). doi: [10.1136/bmj.k1716](https://doi.org/10.1136/bmj.k1716); pmid: [29784870](https://pubmed.ncbi.nlm.nih.gov/29784870/)
45. G. R. Khatri, T. R. Frieden, Controlling tuberculosis in India. *N. Engl. J. Med.* **347**, 1420–1425 (2002). doi: [10.1056/NEJMs020098](https://doi.org/10.1056/NEJMs020098); pmid: [12409545](https://pubmed.ncbi.nlm.nih.gov/12409545/)
46. D. Cazonon et al., Quality of tuberculosis care in high burden countries: The urgent need to address gaps in the care cascade. *Int. J. Infect. Dis.* **56**, 111–116 (2017). doi: [10.1016/j.ijid.2016.10.016](https://doi.org/10.1016/j.ijid.2016.10.016); pmid: [27794468](https://pubmed.ncbi.nlm.nih.gov/27794468/)
47. R. Subbaraman et al., The tuberculosis cascade of care in India's public sector: A systematic review and meta-analysis. *PLOS Med.* **13**, e1002149 (2016). doi: [10.1371/journal.pmed.1002149](https://doi.org/10.1371/journal.pmed.1002149); pmid: [27780217](https://pubmed.ncbi.nlm.nih.gov/27780217/)
48. E. N. Okeke, A. V. Chari, "Can institutional deliveries reduce newborn mortality? Evidence from Rwanda" (Working Paper, RAND Corporation, Santa Monica, CA, 2015).
49. J. G. Meara et al., Global Surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. *Lancet* **386**, 569–624 (2015). doi: [10.1016/S0140-6736\(15\)60160-X](https://doi.org/10.1016/S0140-6736(15)60160-X); pmid: [25924834](https://pubmed.ncbi.nlm.nih.gov/25924834/)
50. World Health Organization, "Depression and other common mental disorders: Global health estimates" (Technical Document WHO/MSD/MER/2017.2, World Health Organization, Geneva, 2017).
51. L. F. Schroeder, J. Guarnar, A. Elbiree, P. E. Castle, T. K. Amukele, Time for a model list of essential diagnostics. *N. Engl. J. Med.* **374**, 2511–2514 (2016). doi: [10.1056/NEJMp1602825](https://doi.org/10.1056/NEJMp1602825); pmid: [27355530](https://pubmed.ncbi.nlm.nih.gov/27355530/)
52. A. Case, A. Deaton, Mortality and morbidity in the 21st century. *Brookings Pap. Econ. Act. Spring* **2017**, 397–476 (2017). pmid: [29033460](https://pubmed.ncbi.nlm.nih.gov/29033460/)
53. D. J. Corsi, I. Mejía-Guevara, S. V. Subramanian, Risk factors for chronic undernutrition among children in India: Estimating relative importance, population attributable risk and fractions. *Soc. Sci. Med.* **157**, 165–185 (2016). doi: [10.1016/j.socscimed.2015.11.014](https://doi.org/10.1016/j.socscimed.2015.11.014); pmid: [26625852](https://pubmed.ncbi.nlm.nih.gov/26625852/)
54. D. Coffey, D. Spears, *Where India Goes: Abandoned Toilets, Stunted Development, and the Costs of Caste* (Harper Collins Publishers India, New Delhi, 2017).
55. Global Burden of Disease Health Financing Collaborator Network, Trends in future health financing and coverage: Future health spending and universal health coverage in 188 countries, 2016–40. *Lancet* **391**, 1783–1798 (2018). doi: [10.1016/S0140-6736\(18\)30697-4](https://doi.org/10.1016/S0140-6736(18)30697-4); pmid: [29678341](https://pubmed.ncbi.nlm.nih.gov/29678341/)
56. D. T. Jamison et al., Global health 2035: A world converging within a generation. *Lancet* **382**, 1898–1955 (2013). doi: [10.1016/S0140-6736\(13\)62105-4](https://doi.org/10.1016/S0140-6736(13)62105-4); pmid: [24309475](https://pubmed.ncbi.nlm.nih.gov/24309475/)
57. T. Barnighausen, D. E. Bloom, in *Oxford Handbook of Health Economics*, S. Glied, P. C. Smith, Eds. (Oxford Univ. Press, 2011), pp. 486–519.
58. M. Sudhinaraset, M. Ingram, H. K. Lofthouse, D. Montagu, What is the role of informal health care providers in developing countries? A systematic review. *PLOS ONE* **8**, e54978 (2013). doi: [10.1371/journal.pone.0054978](https://doi.org/10.1371/journal.pone.0054978); pmid: [23405101](https://pubmed.ncbi.nlm.nih.gov/23405101/)
59. J. Das, A. Chowdhury, R. Hussam, A. V. Banerjee, The impact of training informal health care providers in India: A randomized controlled trial. *Science* **354**, 447–454 (2016). doi: [10.1126/science.aaf7384](https://doi.org/10.1126/science.aaf7384); pmid: [27846471](https://pubmed.ncbi.nlm.nih.gov/27846471/)
60. T. Barnighausen, D. E. Bloom, Designing financial-incentive programmes for return of medical service in underserved areas: Seven management functions. *Hum. Resour. Health* **7**, 52 (2009). doi: [10.1186/1478-4491-7-52](https://doi.org/10.1186/1478-4491-7-52); pmid: [19558682](https://pubmed.ncbi.nlm.nih.gov/19558682/)
61. H. B. Perry, R. Zulliger, M. M. Rogers, Community health workers in low-, middle-, and high-income countries: An overview of their history, recent evolution, and current effectiveness. *Annu. Rev. Public Health* **35**, 399–421 (2014). doi: [10.1146/annurev-pubhealth-032013-182354](https://doi.org/10.1146/annurev-pubhealth-032013-182354); pmid: [24387091](https://pubmed.ncbi.nlm.nih.gov/24387091/)
62. S. A. Patel et al., Chronic disease concordance within Indian households: A cross-sectional study. *PLOS Med.* **14**, e1002395 (2017). doi: [10.1371/journal.pmed.1002395](https://doi.org/10.1371/journal.pmed.1002395); pmid: [28961237](https://pubmed.ncbi.nlm.nih.gov/28961237/)
63. United Nations Department of Economic and Social Affairs, "World population prospects: The 2017 revision" (Report ESA/P/WP/248, United Nations, New York, 2017).
64. S. Ahmed, Q. Li, L. Liu, A. O. Tsui, Maternal deaths averted by contraceptive use: An analysis of 172 countries. *Lancet* **380**, 111–125 (2012). doi: [10.1016/S0140-6736\(12\)60478-4](https://doi.org/10.1016/S0140-6736(12)60478-4); pmid: [22784531](https://pubmed.ncbi.nlm.nih.gov/22784531/)
65. S. Hodgins et al., A new look at care in pregnancy: Simple, effective interventions for neglected populations. *PLOS ONE* **11**,

- e0160562 (2016). doi: [10.1371/journal.pone.0160562](https://doi.org/10.1371/journal.pone.0160562); pmid: [27537281](https://pubmed.ncbi.nlm.nih.gov/27537281/)
66. M. Singer, S. Nambiar, T. Valappil, K. Higgins, S. Gitterman, Historical and regulatory perspectives on the treatment effect of antibacterial drugs for community-acquired pneumonia. *Clin. Infect. Dis.* **47** (suppl. 3), S216–S224 (2008). doi: [10.1086/591407](https://doi.org/10.1086/591407); pmid: [18986293](https://pubmed.ncbi.nlm.nih.gov/18986293/)
 67. Z. Obermeyer, J. Abbott-Klafter, C. J. Murray, Has the DOTS strategy improved case finding or treatment success? An empirical assessment. *PLOS ONE* **3**, e1721 (2008). doi: [10.1371/journal.pone.0001721](https://doi.org/10.1371/journal.pone.0001721); pmid: [18320042](https://pubmed.ncbi.nlm.nih.gov/18320042/)
 68. J. Bor, A. J. Herbst, M. L. Newell, T. Barnighausen, Increases in adult life expectancy in rural South Africa: Valuing the scale-up of HIV treatment. *Science* **339**, 961–965 (2013). doi: [10.1126/science.1230413](https://doi.org/10.1126/science.1230413); pmid: [23430655](https://pubmed.ncbi.nlm.nih.gov/23430655/)
 69. C. Lengeler, Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst. Rev.* **2**, CD000363 (2004). pmid: [15106149](https://pubmed.ncbi.nlm.nih.gov/15106149/)
 70. G. Fink, I. Günther, K. Hill, The effect of water and sanitation on child health: Evidence from the demographic and health surveys 1986–2007. *Int. J. Epidemiol.* **40**, 1196–1204 (2011). doi: [10.1093/ije/dyr102](https://doi.org/10.1093/ije/dyr102); pmid: [21724576](https://pubmed.ncbi.nlm.nih.gov/21724576/)
 71. V. M. Musini, A. M. Tejjani, K. Bassett, J. M. Wright, Pharmacotherapy for hypertension in the elderly. *Cochrane Database Syst. Rev.* **4**, CD000028 (2009). pmid: [19821263](https://pubmed.ncbi.nlm.nih.gov/19821263/)
 72. R. R. Holman, S. K. Paul, M. A. Bethel, D. R. Matthews, H. A. Neil, 10-year follow-up of intensive glucose control in type 2 diabetes. *N. Engl. J. Med.* **359**, 1577–1589 (2008). doi: [10.1056/NEJMoa0806470](https://doi.org/10.1056/NEJMoa0806470); pmid: [18784090](https://pubmed.ncbi.nlm.nih.gov/18784090/)
 73. L. Peirson, D. Fitzpatrick-Lewis, D. Ciliska, R. Warren, Screening for cervical cancer: A systematic review and meta-analysis. *Syst. Rev.* **2**, 35 (2013). doi: [10.1186/2046-4053-2-35](https://doi.org/10.1186/2046-4053-2-35); pmid: [23706117](https://pubmed.ncbi.nlm.nih.gov/23706117/)
 74. T. R. Holford et al., Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964–2012. *JAMA* **311**, 164–171 (2014). doi: [10.1001/jama.2013.285112](https://doi.org/10.1001/jama.2013.285112); pmid: [24399555](https://pubmed.ncbi.nlm.nih.gov/24399555/)
 75. J. Nicholl, J. West, S. Goodacre, J. Turner, The relationship between distance to hospital and patient mortality in emergencies: An observational study. *Emerg. Med. J.* **24**, 665–668 (2007). doi: [10.1136/emj.2007.047654](https://doi.org/10.1136/emj.2007.047654); pmid: [17711952](https://pubmed.ncbi.nlm.nih.gov/17711952/)
 76. A. K. Wagner et al., Access to care and medicines, burden of health care expenditures, and risk protection: Results from the World Health Survey. *Health Policy* **100**, 151–158 (2011). doi: [10.1016/j.healthpol.2010.08.004](https://doi.org/10.1016/j.healthpol.2010.08.004); pmid: [20828854](https://pubmed.ncbi.nlm.nih.gov/20828854/)
 77. D. McIntyre, M. Thiede, G. Dahlgren, M. Whitehead, What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? *Soc. Sci. Med.* **62**, 858–865 (2006). doi: [10.1016/j.socscimed.2005.07.001](https://doi.org/10.1016/j.socscimed.2005.07.001); pmid: [16099574](https://pubmed.ncbi.nlm.nih.gov/16099574/)
 78. A. B. Suthar, L. G. Allen, S. Cifuentes, C. Dye, J. M. Nagata, Lessons learnt from implementation of the International Health Regulations: A systematic review. *Bull. World Health Organ.* **96**, 110–121E (2018). doi: [10.2471/BLT.16.189100](https://doi.org/10.2471/BLT.16.189100); pmid: [29403114](https://pubmed.ncbi.nlm.nih.gov/29403114/)
 79. World Health Organization, WHO leadership statement on the Ebola response and WHO reforms (2015); www.who.int/csr/disease/ebola/joint-statement-ebola/en/.

ACKNOWLEDGMENTS

We are grateful for feedback from K. Prettnner, T. Barnighausen, R. Steinglass, D. Cadarette, R. Glass, M. A. Pate, and two anonymous reviewers. **Funding:** Research by D.E.B. reported in this manuscript was supported by the National Institute on Aging of the National Institutes of Health under award number P30AG024409. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. D.E.B.'s research on this manuscript was also made possible by a grant from the Carnegie Corporation of New York. R.S. acknowledges support from a grant from the Bill and Melinda Gates Foundation via the Arcady Group (OPP1154665) and a Doris Duke Clinical Scientist Development Award. 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 <http://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. **Competing interests:** D.E.B. has been in receipt of grants, travel grants, and/or personal fees from Merck, Pfizer, GSK, Sanofi Pasteur, and Sanofi Pasteur–MSD (all related to his research on the value of vaccination) and Gilead Life Sciences (related to his work on the value of treatment for hepatitis C in India). The authors declare no other competing interests.

10.1126/science.aat9644

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Endothelial Dab1 signaling orchestrates neuro-glia-vessel communication in the central nervous system

Marta Segarra*, Maria R. Aburto*, Florian Cop*, Cecília Llaó-Cid, Ricarda Härtl, Miriam Damm, Ioanna Bethani, Marta Parrilla, Dewi Husainie, Anne Schänzer, Hannah Schlierbach, Till Acker, Laura Mohr, Laia Torres-Masjoan, Mathias Ritter, Amparo Acker-Palmer†

INTRODUCTION: The function of the brain relies on communication among the complex network of cells that constitute this organ. Vascularization of the central nervous system (CNS) ensures adequate delivery of oxygen and nutrients to build up and maintain homeostasis of neuronal networks. Thus, it is not surprising that blood vessels and neuronal cells share multiple parallelisms orchestrating their development in synchrony and in a mutually dependent manner in the CNS. Despite the essential role of the endothelium in brain function, the means by which signaling at the interface of endothelial cells, glial cells, and neurons is integrated temporally and spatially for proper brain development has remained largely unexplored.

RATIONALE: Integration of signaling pathways and cellular responses among endothelial cells, glial cells, and neurons is needed to ensure proper architecture of the brain. Reelin (Reln), a large secreted glycoprotein, induces Disabled 1 (Dab1)-dependent responses in neurons to guide their migration in all layered brain structures. Secretion of reelin by Cajal-Retzius cells in the marginal zone of the cortex timely coincides with active sprouting of pial vessels

ingrowing perpendicularly into the marginal zone and forming a complex vascular network needed to support brain development and function. Therefore, reelin might be in the perfect position to perform a bivalent function to timely and spatially orchestrate both neuronal migration and CNS vascularization. We reasoned that blood vessels might instruct the process of neuronal migration by a cell-autonomous function of Dab1 on endothelial cells. To investigate this, we deleted the expression of vascular Dab1 in mice and investigated the effects on CNS vascularization, neuroglial organization, and neurovascular unit function.

RESULTS: We found that reelin/Dab1 signaling is conserved in endothelial cells and exerts potent proangiogenic effects in the developing vasculature of the CNS by controlling endothelial cell proliferation and active filopodia extension of the vascular network. The interaction of the reelin receptor ApoER2 (apolipoprotein E receptor 2) and VEGFR2 (vascular endothelial growth factor receptor 2) mediated the proangiogenic roles of Dab1 in endothelial cells. Surprisingly, deletion of Dab1 exclusively in the vascular system induced changes in the

position of postmitotic pyramidal neurons in the cortical layers of the cerebral cortex. At the cellular level, depletion of vascular Dab1 reduced the docking of the radial glia processes to the pial surface at embryonic and postnatal stages and altered the differentiation of glial cells to astrocytes. The defects in neuronal migration persisted in adult mutant animals, where stereotypical attachment of the astrocytes to penetrating vessels in the glia limitans superficialis was also found to be aberrant. The functionality

ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aao2861>

of the neurovascular unit [blood-brain barrier (BBB) integrity] was also affected in reelin knockout animals, and we could attribute those defects to the lack of Dab1 signaling exclu-

sively in endothelial cells. The increased BBB permeability was again associated with an insufficient coverage of the brain vasculature by astrocytic endfeet. Mechanistically, we determined that the astroglial attachment to the vasculature is mediated by reelin-induced deposition of laminin- $\alpha 4$ by endothelial cells to the extracellular compartment, which in turn enables the binding of the glial processes to the CNS vasculature via the activation of integrin- $\beta 1$ in glial cells.

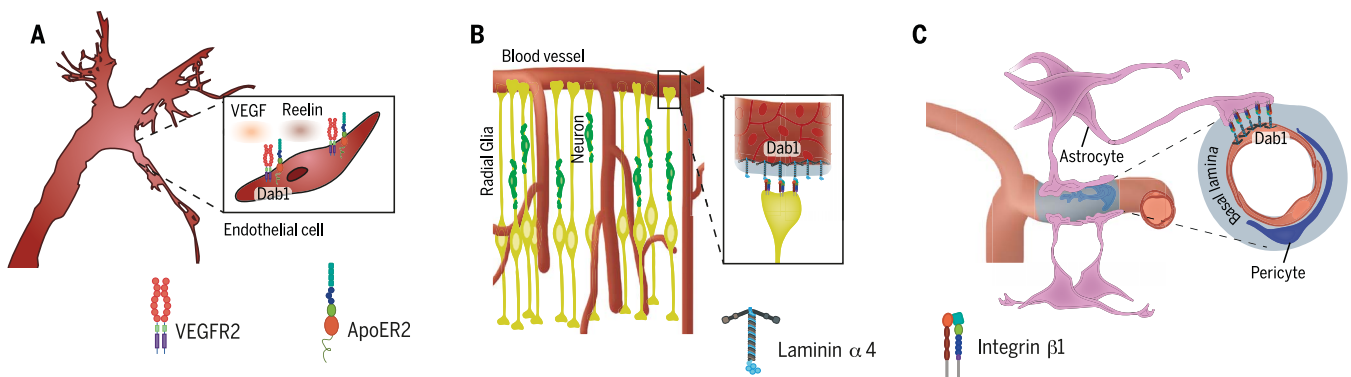
CONCLUSION: Our results shed new light on the function of the vasculature in CNS development and homeostasis—in particular, how signals from the endothelium orchestrate the communication among vessels, glial cells, and neurons, and how specific changes in the molecular signature of the endothelium affect a plethora of processes such as CNS vascularization, extracellular matrix composition, neuroglial cytoarchitecture, and BBB development. ■

The list of author affiliations is available in the full article online.

*These authors contributed equally to this work.

†Corresponding author. Email: acker-palmer@bio.uni-frankfurt.de

Cite this article as M. Segarra et al., *Science* 361, eaao2861 (2018). DOI: 10.1126/science.aao2861



Instructive functions of vascular Dab1 in the neurovascular interface. (A) ApoER2 and Dab1 control endothelial cell proliferation and tip cell filopodia extension during CNS vascularization by cross-talking to the VEGFR2 pathway. (B and C) Vascular Dab1 also instructs radial glia organization and neuronal migration in the developing cerebral cortex (B) and the development of the blood-brain barrier (C). In both cases, Dab1 signaling in the vasculature regulates the deposition of laminin- $\alpha 4$ and, in turn, the activation of integrin- $\beta 1$ in glial cells.

RESEARCH ARTICLE

NEUROSCIENCE

Endothelial Dab1 signaling orchestrates neuro-glia-vessel communication in the central nervous system

Marta Segarra^{1*}, Maria R. Aburto^{1,2*}, Florian Cop^{1*}, Cecília Llaó-Cid¹, Ricarda Härtl¹, Miriam Damm^{1,2}, Ioanna Bethani¹, Marta Parrilla^{1,3}, Dewi Husainie^{1,3}, Anne Schänzer⁴, Hannah Schlierbach⁴, Till Acker⁴, Laura Mohr¹, Laia Torres-Masjoan¹, Mathias Ritter¹, Amparo Acker-Palmer^{1,2,3†}

The architecture of the neurovascular unit (NVU) is controlled by the communication of neurons, glia, and vascular cells. We found that the neuronal guidance cue reelin possesses proangiogenic activities that ensure the communication of endothelial cells (ECs) with the glia to control neuronal migration and the establishment of the blood-brain barrier in the mouse brain. Apolipoprotein E receptor 2 (ApoER2) and Disabled1 (Dab1) expressed in ECs are required for vascularization of the retina and the cerebral cortex. Deletion of Dab1 in ECs leads to a reduced secretion of laminin- α 4 and decreased activation of integrin- β 1 in glial cells, which in turn control neuronal migration and barrier properties of the NVU. Thus, reelin signaling in the endothelium is an instructive and integrative cue essential for neuro-glia-vascular communication.

Vascularization of the central nervous system (CNS) ensures adequate delivery of oxygen and nutrients to build up and maintain homeostasis of neuronal networks (1). Apart from these metabolic functions, vessels have also been suggested to serve as niches and scaffolds for neuronal migration both during development and during adult neurogenesis (2). The orchestration of a perfect architecture of the neurovascular unit (NVU) is fundamental for brain function. Previous studies have examined the relation of the vasculature to the neuroglial components. It has been shown that the radial glia scaffold is necessary for angiogenesis because ablation of radial glia cells induces regression of cortical vessels (3), and that the expression of integrin- β 8 by glial cells regulates developmental vascularization of the brain (4). Additionally, besides the well-established support of neuroblast migration by the radial glia during corticogenesis (5), it has been reported that the vasculature also contributes to neuronal navigation and positioning; for example, periventricular endothelial cells (ECs), which release γ -aminobutyric acid (GABA), guide the tangential migration of GABAergic neurons during embryogenesis (6, 7). However, despite these emerging

studies, the means by which signaling at the interface of neurons, ECs, and glial cells is integrated for proper brain development remains largely unexplored.

Reelin, a well-known guidance cue for migrating neurons, regulates lamination of brain regions during development and synaptic plasticity in the adult brain (8). Binding of reelin to the apolipoprotein E receptor 2 (ApoER2) and to the very low density lipoprotein receptor (VLDLR) triggers a signaling cascade required for neuronal migration that involves the adaptor protein Disabled1 (Dab1) (9–13). Reelin knockout or Dab1 knockout mice exhibit aberrant cortical, hippocampal, and cerebellar architecture (14, 15). Notably, reelin signaling also has an impact on the blood and lymphatic vasculature (16–18). In humans, reelin mutations cause severe developmental defects (19) and are associated with several neurological diseases (20–22). While neurons are migrating and colonizing the layers in the mouse cortex guided by the expression of reelin in the marginal zone (8), pial vessels start to sprout perpendicular to the cortical marginal zone and grow into the cortex to form the complex vascular network needed to support brain development and function (23).

Reelin induces angiogenic responses

Our ex vivo culture technique allows quantification of acute EC responses and the guidance of tip cells during angiogenic sprouting in the mouse retina (24). Stimulation of explanted retinas with exogenous reelin (Reln) resulted in an increase of filopodia extensions per vessel length (Fig. 1, A and B). In vitro, stimulation of ECs with exogenous reelin promoted EC tube formation

(fig. S1, A to C) and Dab1 phosphorylation (Fig. 1, C to E, and fig. S1, D and E). In the mouse retina vascularization model (25), Reln^{-/-} and ApoER2^{-/-} mice both showed defective extension of the superficial vascular network at postnatal day 2 (P2) (Fig. 1, F and G, and fig. S1, F and G) and a strong decrease in filopodia extensions at the vascular front at P7 (Fig. 1, H and I, and fig. S1, H and I). Reelin's proangiogenic effects are mediated by its receptor ApoER2, because retinas explanted from ApoER2^{-/-} mice failed to increase filopodia extensions after reelin stimulation (Fig. 1, J and K).

Early postnatal retinal vascularization and endothelial tip cell formation are regulated by astrocytes within the ganglion cell layer, which produce the vascular chemotactic factor VEGF (vascular endothelial growth factor) (26, 27). Moreover, in cortical neurons in culture, VEGF and reelin pathways cross-talk to regulate the phosphorylation of the NR2 subunit of the N-methyl-D-aspartate (NMDA) receptor (28). We found that ApoER2 coimmunoprecipitated and coclustered with VEGF receptor 2 (VEGFR2) in ECs, and that such interaction was induced by VEGF-A and by reelin (Fig. 2, A to C, and fig. S2, A and B). VEGF-A activated VEGFR2 and induced Dab1 phosphorylation in ECs (Fig. 2, D to F). VEGF-C, another ligand for VEGFR2 (29), also induced Dab1 phosphorylation (fig. S2C), supporting the idea that VEGFR2 activation occurs upstream of Dab1 signaling. Reciprocally, reelin also activated VEGFR2 (fig. S2D). Costimulation with VEGF and reelin synergistically increased Dab1 phosphorylation (fig. S2, E and F) and filopodia extension in the ex vivo retina assay (Fig. 2G). Moreover, ApoER2^{-/-} vessels were unable to respond to VEGF-mediated tip cell sprouting (Fig. 2, H and I), indicating the presence of functional cross-talk between the two receptors.

Dab1 regulates vascular morphogenesis cell-autonomously

The retinal ganglion cell layer, underneath the astrocytes and the vessel bed, appeared as the major source of reelin (fig. S3, A and B). Both ApoER2 and Dab1 are expressed in retinal vessels (fig. S3, C and D). To uncouple the function of Dab1 in the vasculature from its function in neuronal cells, we generated inducible endothelial-specific Dab1 loss-of-function mice (Dab1^{ΔEC}) by crossing Dab1^{lox/lox} mice to the tamoxifen-inducible vascular-specific Cre deleter line Cdh5(PAC) creERT2. Tamoxifen injection for three consecutive days efficiently induced the vascular-specific genetic recombination and loss of Dab1 expression in ECs (fig. S4, A to C). Dab1 deletion from P1 to P3 impaired vascular radial growth in Dab1^{ΔEC} retinas at P3 to P4 (Fig. 3, A and B). This reduction in vessel growth was persistent at P7 (fig. S4, D and E) and recovered at adult stages (fig. S4, F and G). Moreover, EC proliferation and vascular network complexity were also decreased in the Dab1^{ΔEC} postnatal retinas (Fig. 3, C and D, and fig. S4, H and I). A reduced number of filopodia per vessel length and a reduced tip/stalk cell ratio were also observed in Dab1^{ΔEC}

¹Institute of Cell Biology and Neuroscience and Buchmann Institute for Molecular Life Sciences, University of Frankfurt, D-60438 Frankfurt am Main, Germany. ²Focus Program Translational Neurosciences, University of Mainz, D-55131 Mainz, Germany. ³Max Planck Institute for Brain Research, D-60438 Frankfurt am Main, Germany. ⁴Institute of Neuropathology, University of Giessen, D-35392 Giessen, Germany.

*These authors contributed equally to this work.

†Corresponding author. Email: acker-palmer@bio.uni-frankfurt.de

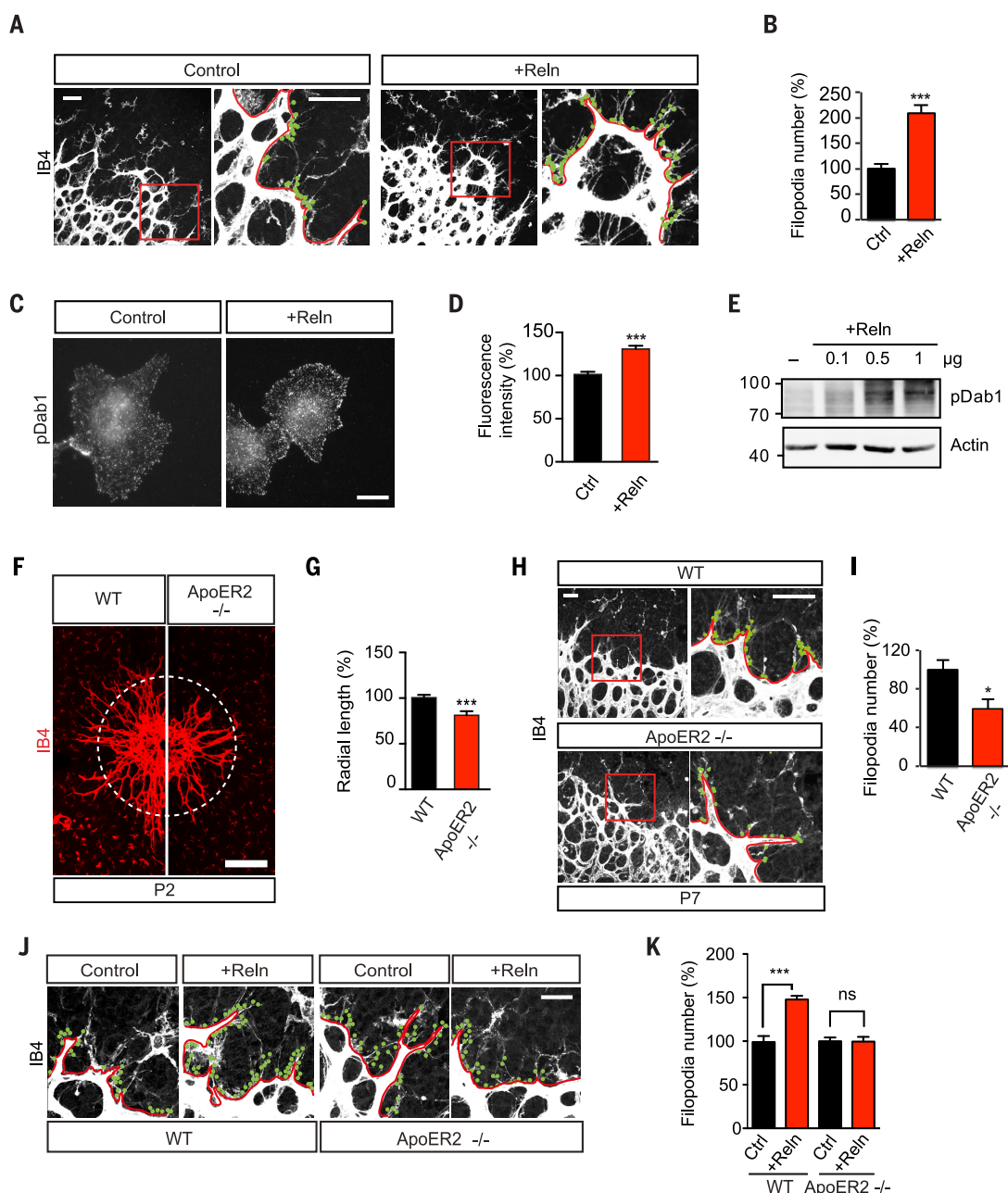
mice (Fig. 3, E to H). Expression of VEGF or reelin was not changed in $Dab1^{1ABC}$ or ApoER2 mutants (fig. S4J).

In the cerebral cortex, reelin is secreted by Cajal-Retzius cells during developmental stages (30) and regulates the migration of neurons during their correct positioning in the different cortical layers (8, 31). Cortical vessels expressed both ApoER2 and Dab1, and brain ECs also responded to reelin by phosphorylating Dab1 (fig. S5, A to C). The embryonic brain is vascularized by the pial and periventricular vessels (32), with sprouts from the pial vessels penetrating into the neural tube from the perineural vascular plexus around

embryonic day 10.5 (E10.5) by sprouting angiogenesis and forming columnar vascular structures perpendicular to the meninges (23, 33). In the absence of reelin, vascularization of the embryonic neocortex was aberrant (Fig. 3, I and J). From E11.5 on, we found fewer sprouts penetrating from the pial vessels into the neocortex and connecting to the periventricular vessels (Fig. 3I, arrows). In control animals, penetrating vessels from the pia at E14.5 started branching directly below the marginal zone, but this pattern was lost in the $Reln^{-/-}$ mutants (Fig. 3I, dashed line). Vessels were also less branched at E13.5 and E14.5, which later resulted in a reduction of

vessel density as quantified at E17.5 (Fig. 3J). The same aberrant phenotype was observed when deleting $Dab1$ exclusively from the vessels at E10.5 to E12.5 and analyzing the cortical vasculature at E17.5 (Fig. 3, K and L); this finding supported the cell-autonomous function of $Dab1$ observed in the retina. $Dab1$ deletion at early postnatal stages from P1 to P3 and analysis at P7 to P8 also unraveled a defective vessel architecture (Fig. 3M) reflected by a reduction in vessel density (Fig. 3N), vessel length (Fig. 3O), and number of branch points (Fig. 3P) as well as a preferential orientation at an 80° to 90° angle with respect to the pial surface (Fig. 3Q). These morphological defects

Fig. 1. Reelin pathway is proangiogenic in endothelial cells. (A) Isolectin B4 (IB4) whole-mount staining of wild-type retinal explants after 4 hours of reelin (Reln) stimulation. (B) Quantification of the number of filopodia (green dots) at the vascular front (areas outlined in red) depicted in (A) ($n = 9$ to 15 images, 3 explants per condition). (C) Immunofluorescence of phospho-Dab1 (pDab1) in human umbilical vein endothelial cells (HUVECs) after 30 min of Reln stimulation. (D) Quantification of (C) ($n = 56$ to 66 images, 6 experiments). (E) Western blot analysis of Dab1 phosphorylation in HUVECs after Reln stimulation. Actin was used as a loading control. (F) IB4 whole-mount staining of wild-type (WT) and ApoER2 $^{-/-}$ mutant vascular networks at P2. (G) Quantification of the vessel radial length in (F) ($n = 20$ or 21 retinas, 10 or 11 animals per genotype). (H) ApoER2 $^{-/-}$ retinas at P7. (I) Quantification of filopodia extensions in (H) ($n = 5$ retinas, 3 to 5 animals per genotype). (J) Reln stimulation of retinal explant cultures for 4 hours increases filopodia sprouting at the vascular front in WT explants but not in ApoER2 $^{-/-}$ explants. (K) Quantification of the number of filopodia related to (J) ($n = 10$ to 18 images, 3 explants per condition). Data were normalized to corresponding controls. Scale bars, 30 μ m [(A), (H), and (J)], 10 μ m (C), 200 μ m (F). Data are means \pm SEM. * $P < 0.05$, *** $P < 0.001$; ns, not significant.



did not impair vessel perfusion (fig. S5, D and E). As in the retina, the morphological vascular defects in the cortex were compensated at adult stages (fig. S5, F to J).

Dab1 in the endothelium is required for proper migration of neurons during cortical layering

The absence of reelin expression by Cajal-Retzius cells in the marginal zone (MZ) as well as the lack of Dab1 signaling in neurons leads to many defects in cortical lamination (8, 15, 34, 35), such as projection neurons invading the MZ or layer I. We hypothesized that the cell-autonomous function of reelin signaling in the vasculature could also contribute to the proper migration of neurons during cortical lamination. Examination of the general cortical cytoarchitecture of the early embryonic deleted (E10.5 to E12.5) *Dab1*^{ΔEC} mutants analyzed at E17.5 revealed a poorly defined separation of cortical layers as well as a remarkable invasion of cells in the marginal zone (Fig. 4A) recapitulating some of the defects observed in *Reln*^{-/-} mice (35). Markers for deeper and upper layers—*Tbr1*⁺ and *Cux1*⁺, respectively—showed

aberrant positioning of early- and late-born neurons at E17.5 (Fig. 4, B to E). Positioning and numbers of *Pax6*⁺ apical progenitors and *Tbr2*⁺ intermediate (basal) neuronal progenitors were not affected by embryonic *Dab1* vascular deletion (fig. S6, A to D), indicating that the defects in neuronal positioning in the *Dab1*^{ΔEC} mutants are not a consequence of a deficient neurogenic pool. We then focused on analyzing the neuronal migration defects in more detail. At perinatal days, later-born neurons, which are destined to upper cortical layers, are still navigating along the radial glia processes before translocating into layers II/III (36). Neuronal distribution analysis in the *Dab1*^{ΔEC} mice at P7–P8 revealed invasion of postmitotic neurons in layer I after perinatal *Dab1* deletion (Fig. 4, F and G). Immunostaining for *Cux1* confirmed that later-born neurons entered into layer I (Fig. 4, H and I). Additionally, we found an increased number of *Cux1*⁺ cells below layer IV (Fig. 4, H and J). Birth-dating neurons with bromodeoxyuridine (BrdU) injected at E15.5 and analyzed at P8 after perinatal deletion of *Dab1* confirmed the presence of later-born neurons in the lower cortical layers as well

as a delayed migration of E15-born neurons into the upper layers (fig. S6, E to G). General defects in brain size were not observed in the *Dab1*^{ΔEC} mice (fig. S7, A to F).

Endothelial Dab1 is required for vessel-glia communication

Apical radial glial cells (RGCs) extend a single basal process (radial fiber) that reaches the meningeal basement membrane (BM) and anchors via bulb endfeet structures making frequent contact with the pial vessels (5, 37). Such endfeet-BM interactions are thought to contribute to the final placement of migrating neurons in the cortex (38–40). Analysis of E17.5 *Dab1*^{ΔEC} mice (tamoxifen administration E10.5 to E12.5) and E17.5 *Reln*^{-/-} mice revealed a reduction of docking endfeet of RGCs to the BM (Fig. 5, A and B, and fig. S8, A and B). The same results were observed after perinatal removal of vascular *Dab1* (Fig. 5, C and D). Deletion of *VEGFR2* in the vessels reduced vascularization but did not recapitulate the same defects in glia anchoring (fig. S8, C to F), which suggests that the effects of *Dab1* in glia-EC communication are independent

Fig. 2. Reelin pathway cross-talks with VEGF signaling in endothelial cells.

(A) Immunoprecipitation (IP) of *VEGFR2* from HUVEC lysates and Western blot for *VEGFR2* and *ApoER2*. TL, total lysates. (B) *ApoER2*/*VEGFR2* clusters (white punctae) upon 10 min of stimulation with *VEGF-A* in HUVECs. PLA, proximity ligation assay. (C) Quantification of *ApoER2*/*VEGFR2* clusters in (B) ($n = 108$ to 110 cells, 3 experiments). (D) Western blot of phosphorylation of *Dab1* (p*Dab1*) and *VEGFR2* (p*VEGFR2*) after *VEGF-A* stimulation of HUVECs. (E) Immunofluorescence staining with a phospho-specific *Dab1* antibody in HUVECs showing increased p*Dab1* after *VEGF-A* stimulation. (F) Quantification of p*Dab1* in (E) ($n = 26$ to 30 images, 3 experiments). (G) WT retinal explants exposed to a combination of *Reln* and *VEGF-A* for 4 hours show increased number of filopodia when compared to single stimulation conditions and to the control ($n = 13$ to 21 images, 4 or 5 explants per condition). (H) *VEGF-A* stimulation of WT and *ApoER2*^{-/-} retinal explants. (I) Quantification of (H) ($n = 9$ or 10 images, 3 explants per condition). Stimulation is normalized to each control condition. Scale bars, $30\ \mu\text{m}$ [(B) and (H)], $10\ \mu\text{m}$ (E). Data are means \pm SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

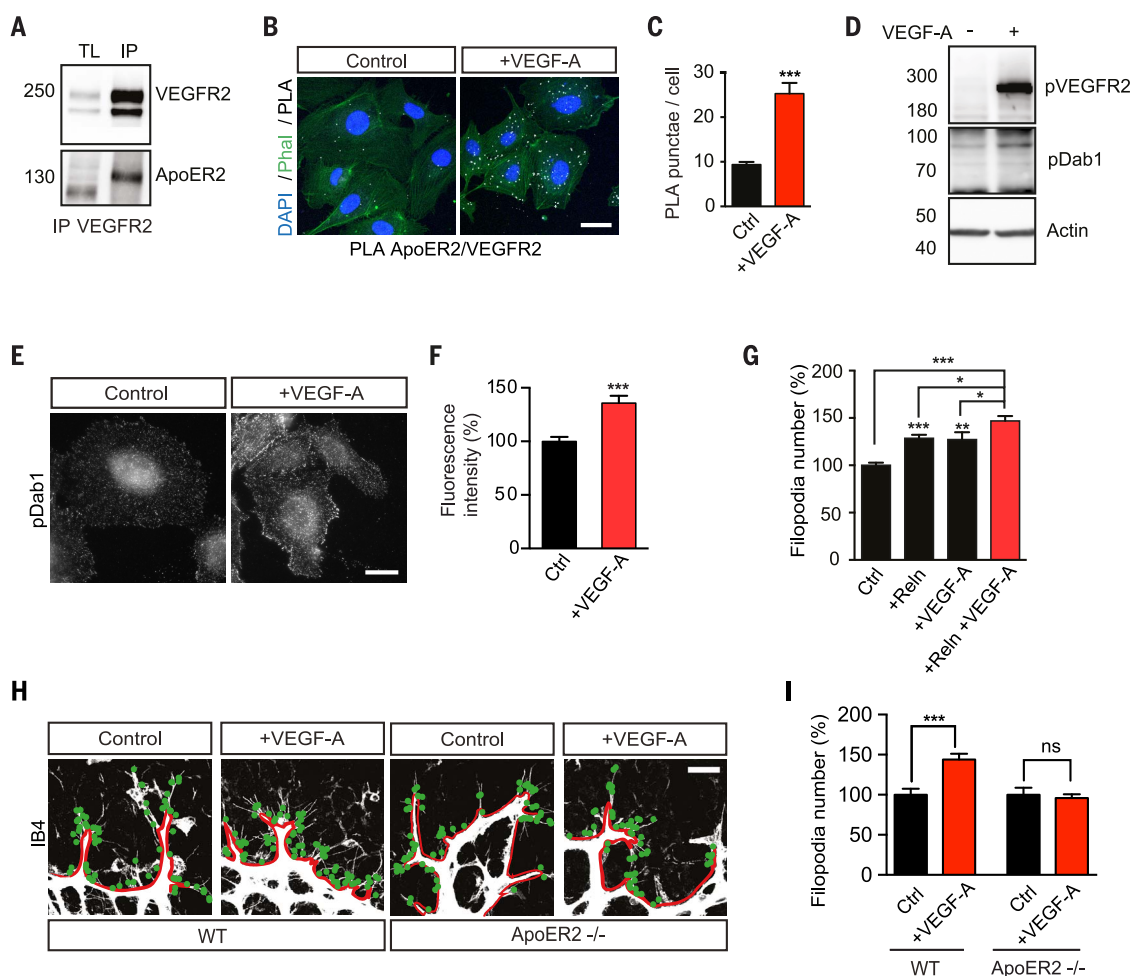


Fig. 3. Endothelial Dab1 is essential for retina and cortex vascularization.

(A) IB4 staining of *Dab1*^{ΔEC} retinas at P3–P4 after tamoxifen (TMX) administration from P1 to P3.

(B) Quantification of vessel radial length shown in (A) ($n = 6$ to 12 retinas, 3 to 7 animals per genotype).

(C) Vascular proliferation (phospho-histone H3, pHH3) in *Dab1*^{ΔEC} animals at P3–P4 after TMX administration from P1 to P3.

(D) Quantification of the number of pHH3⁺ ECs in (C) ($n = 9$ to 15 retinas, 5 to 9 animals per genotype).

(E) Filopodia extensions at the vascular front of *Dab1*^{ΔEC} retinas at P7.

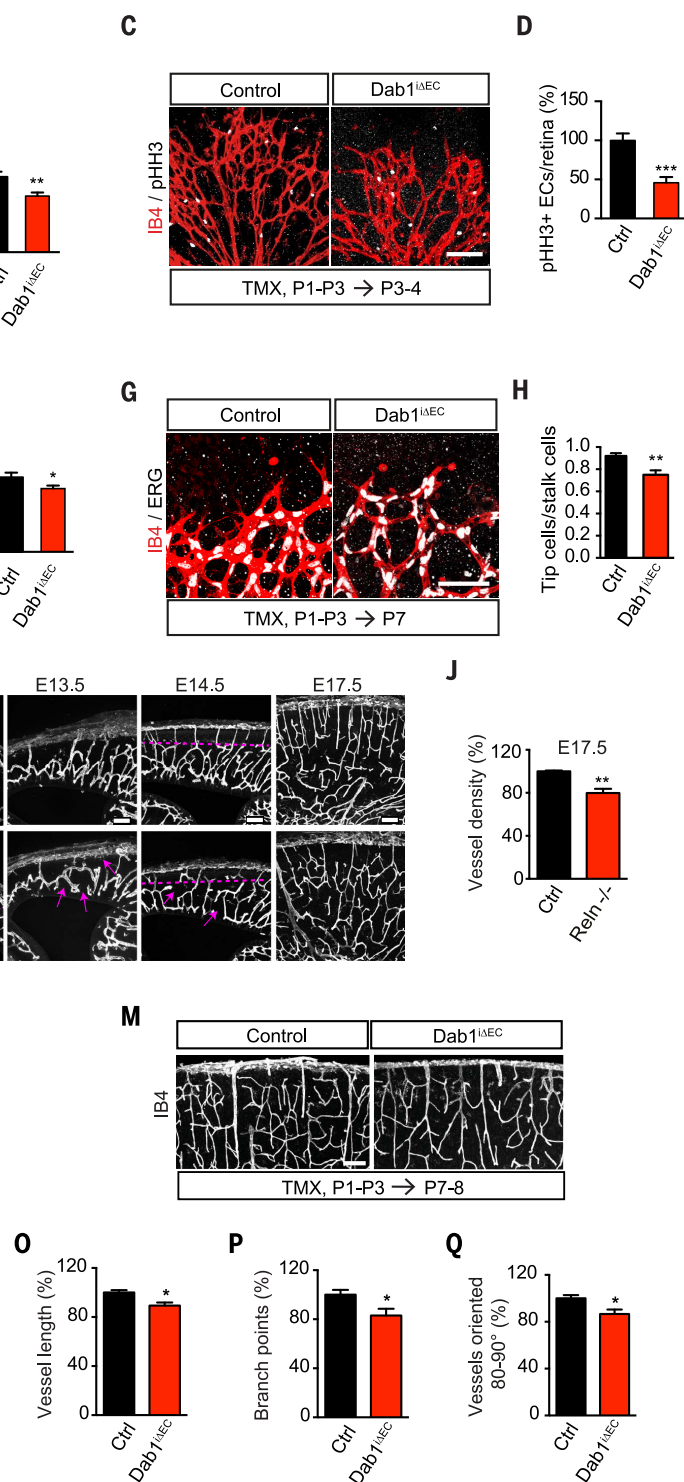
(F) Quantification of (E) ($n = 5$ retinas, 3 animals per genotype).

(G) ETS-related gene (ERG) and IB4 staining in *Dab1*^{ΔEC} retinas at P7 after TMX administration from P1 to P3.

(H) Quantification of the relative number of tip cells versus stalk cells at the vascular front shown in (G) ($n = 6$ to 10 retinas, 5 or 6 animals per genotype).

(I) Development of the vasculature of control and *Reln*^{-/-} embryos stained with IB4. Arrows point to defects in ingrowing sprouts from the pial vessel and branching defects in the neo-cortex below the marginal zone. Dashed line delineates the marginal zone. (J) Quantification of vessel density in E17.5 cortices in (I) ($n = 3$ animals per genotype).

(K and M) *Dab1*^{ΔEC} cortices stained with IB4 at E17.5 (K) or P7–P8 (M) after TMX administration at the indicated time points. The vascularization of the upper cortex is reduced in the mutant animals, with decreased area covered by vessels and fewer vascular intersections. (L) Quantification of vessel density in (K) ($n = 5$ or 6 animals



per genotype). (N to Q) Quantification of vessel density (N) ($n = 8$ or 9 animals per genotype), vessel length (O) ($n = 5$ animals per genotype), branch points (P) ($n = 5$ animals per genotype), and vessel orientation in relation to the pial surface (Q) ($n = 5$ or 6 animals per genotype) shown in (M). Scale bars, 200 μm (A), 100 μm [(C), (I), (K), (M)], 20 μm (E), 75 μm (G). Data are means ± SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

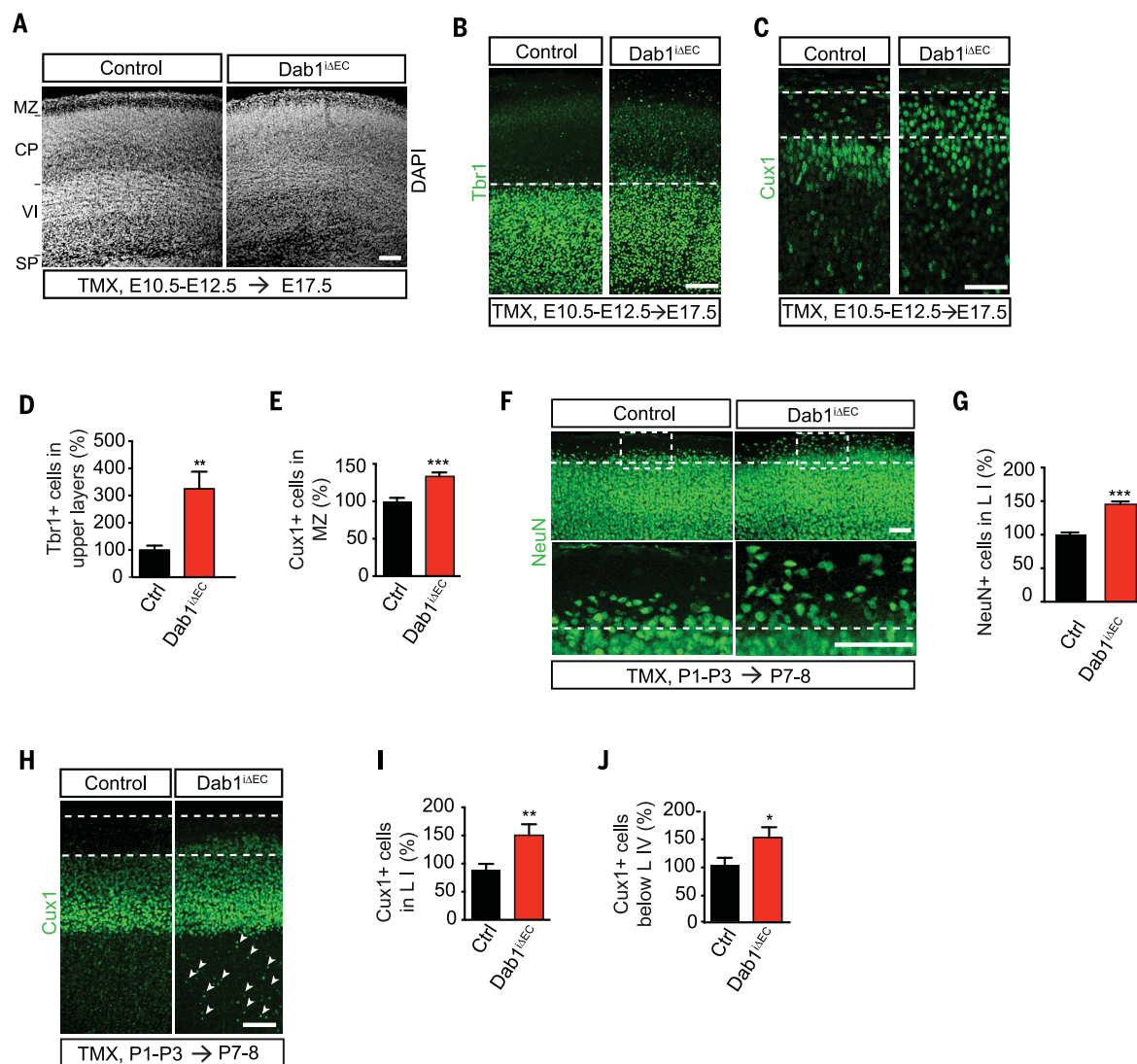
Fig. 4. Vascular reelin signaling mediates neuronal positioning in the neocortex.

(A to C) DAPI (A), Tbr1 (B), and Cux1 (C) staining of *Dab1^{ΔEC}* cortices at E17.5 after TMX administration from E10.5 to E12.5.

(D) Quantification of the number of Tbr1⁺ cells in upper layers in (B) ($n = 10$ to 13 images, 2 or 3 animals per genotype). (E) Quantification of Cux1⁺ cells in the marginal zone in (C) ($n = 30$ to 37 images, 5 animals per genotype). (F) NeuN staining of *Dab1^{ΔEC}* cortices at P7–P8 after TMX administration from P1 to P3. (G) Quantification of NeuN⁺ cells in layer I in (F) ($n = 29$ to 34 images, 5 or 6 animals per genotype). (H) Cux1 staining of *Dab1^{ΔEC}* cortices at P7–P8 after TMX administration from P1 to P3. (I and J) Quantification of Cux1⁺

neurons in layer I ($n = 19$ images, 4 animals per genotype) (I) and below layer IV ($n = 40$ to 57 images, 7 to 10 animals per genotype) (J). Scale bars, 100 μm [(A) and (F)], 50 μm [(B), (C), and (H)]. MZ, marginal zone; CP, cortical plate; VI, layer VI; SP, subplate; L, layer. Data are means \pm SEM.

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



of VEGF/VEGFR2 signaling and distinct from the effects of reelin/Dab1 signaling on vessel growth (Figs. 1 to 3).

Later in postnatal stages, RGCs differentiate into mature astrocytes after retracting their contacts from the pial and ventricular surfaces (41). We also observed a distortion of the columnar morphology of RGCs in vascular *Dab1* mutants at P4 (Fig. 5E, left panels). This was accompanied by a premature morphological change of the bipolar radial glial cells toward stellate-shaped cells characteristic of mature astrocytes (GFAP⁺) (Fig. 5E, right panels). Moreover, analysis of *Dab1^{ΔEC}* mice at 4 to 7 weeks of age after perinatal deletion revealed that the invasion of neurons in layer I persisted in adult stages (Fig. 5, F and G). The astrocytic organization at the glia limitans superficialis, with astrocytes closely sealing the meningeal surface, is characterized by GFAP⁺ fibrillary processes radially distributed

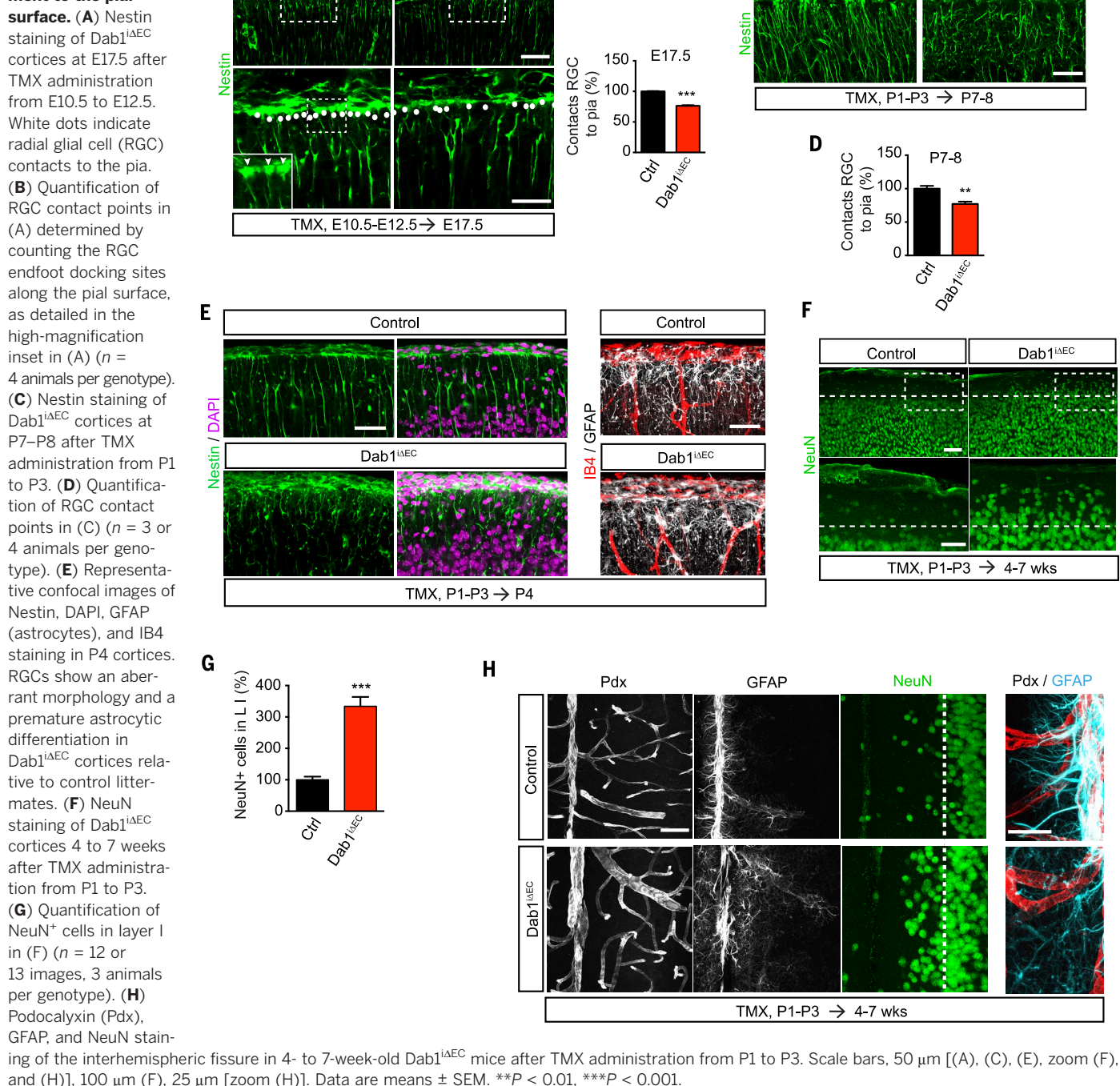
along the interhemispheric vasculature and the penetrating vessels. This organization was disturbed in the *Dab1^{ΔEC}* mice denoted by GFAP⁺ astrocytic processes detaching from the pial surface and its penetrating vessels (Fig. 5H).

Vascular Dab1 is necessary to form a functional blood-brain barrier

In the CNS, blood vessels are central components of the NVU that regulate homeostatic functions of the brain: the blood-brain barrier (BBB) and the neurovascular coupling (42, 43). Extravasation of the BBB-impermeable fluorescent tracer Alexa Fluor 555 cadaverine was observed in postnatal and adult mice after perinatal deletion of vascular *Dab1* (Fig. 6, A to F) and in the *Reln^{-/-}* mutants (Fig. 6, G to I), indicating compromised BBB integrity. Electron microscopy analysis revealed anatomically defective tight junctions as well as a functionally augmented rate of trans-

cytosis (Fig. 6, J to L). *Reln^{-/-}* vessels exposed unsealed tight junctions with an altered alignment in relation to the vessel lumen (Fig. 6, J and K). Moreover, a significantly increased number of cytoplasmic vesicles filled with the tracer horseradish peroxidase was observed in the *Reln^{-/-}* endothelium (Fig. 6, J and L). The blood-retinal barrier was also impaired at P7 in *Dab1^{ΔEC}* mice (Fig. 6, M to O), and defects were compensated after 4 weeks of age (fig. S9, A to C). We performed a late acute removal of *Dab1* by using the *Cdh5* (PAC)creERT2 deleter mice and injecting tamoxifen for three consecutive days starting at P25 and evaluating the BBB permeability at P30. In this setting, removal of *Dab1* in mature vessels did not seem to significantly affect the BBB maintenance (fig. S9, D to H). We also found a down-regulation of *Dab1* expression at P30 relative to postnatal stages (fig. S9I). Overall, this suggests that *Dab1* function is important

Fig. 5. Vascular reelin signaling mediates radial glial attachment to the pial surface. (A) Nestin staining of *Dab1^{ΔEC}* cortices at E17.5 after TMX administration from E10.5 to E12.5. White dots indicate radial glial cell (RGC) contacts to the pia.



for BBB etiology but not for physiological barrier maintenance.

Vascular *Dab1* is necessary for laminin- $\alpha 4$ secretion and astrocytic integrin- $\beta 1$ activation

Analysis of the NVU cellular components in *Dab1^{ΔEC}* and *Reln^{-/-}* mice revealed a prominent reduction of the coverage of cerebral vessels by the endfeet water channel aquaporin4 (Aqp4) (Fig. 7, A and B, and fig. S10, A and B). Total

levels of Aqp4 were not changed, which suggested an uncoupling of astrocytic endfeet to the NVU (Fig. 7C and fig. S10C). We confirmed that pericyte ensheathment, also important for BBB integrity (44, 45), was not significantly affected in our mutants (fig. S11, A to D).

The extracellular matrix (ECM) network between perivascular astrocytic endfeet and ECs regulates barrier properties of the cerebral vasculature (46). Laminins are major components of the gliovascular lamina (47). Albumin extra-

vasation from vessels in the *Reln^{-/-}* mice coincided with an impaired deposition of laminin- $\alpha 4$ (Lama-4, produced by endothelium) and a decrease in laminin- $\alpha 2$ [Lama-2, expressed at the endfeet (48)] (Fig. 7D). Moreover, deposition of Lama-4 coincided with sites of accumulation of Aqp4 staining in control animals (Fig. 7E, upper panels, arrows), which suggests that Aqp4-enriched endfeet dock on sites of endothelial Lama-4 accumulation. In agreement with this, *Reln^{-/-}* mutants showed decreased Lama-4 deposition in

correlation with the absence of Aqp4 staining (Fig. 7E, lower panels, arrowheads). Correlative accumulation of Aqp4 and Lama-4 can be appreciated in higher-magnification pictures and in the corresponding line intensity profile of the staining (Fig. 7F).

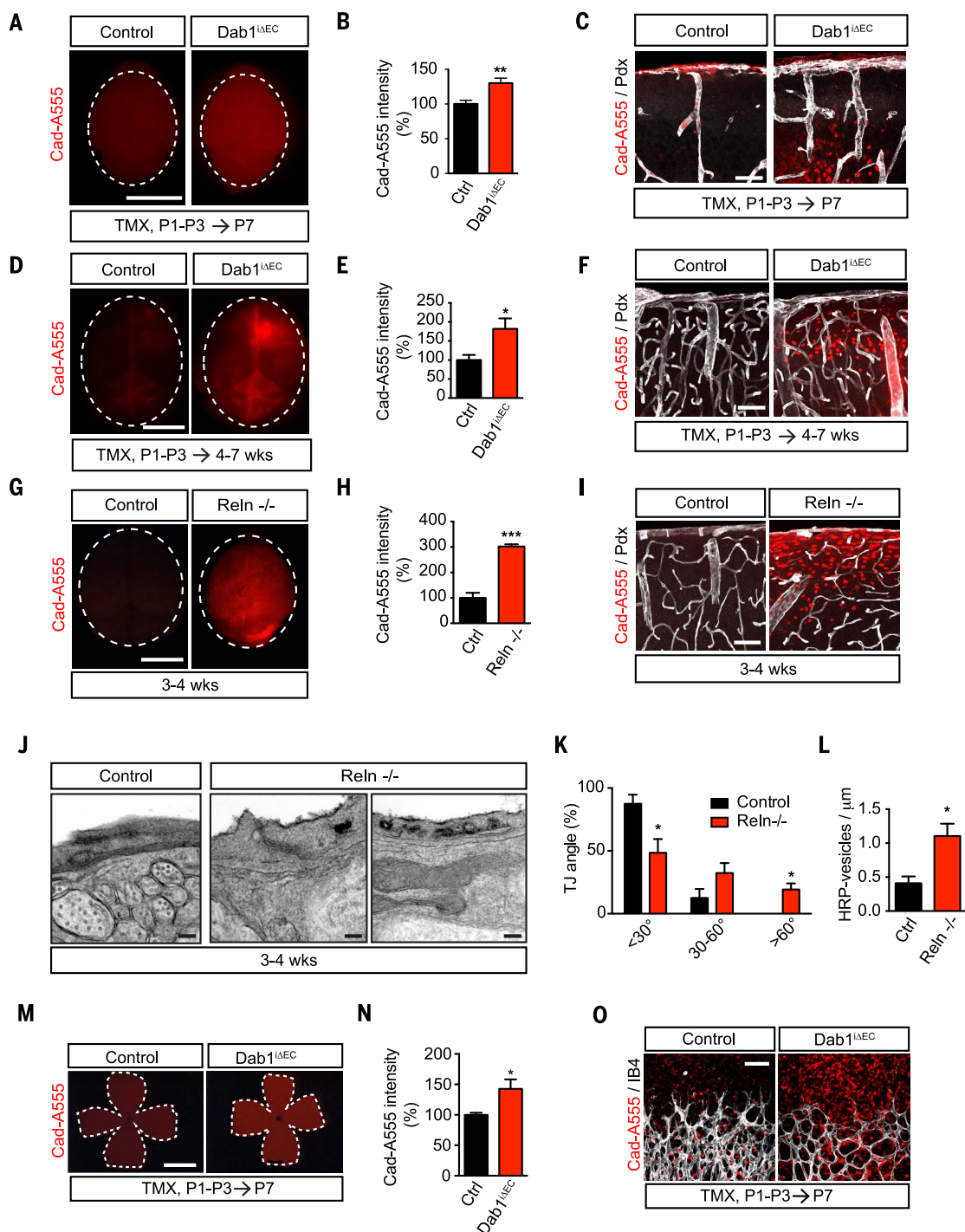
Lama-4 in the ECM binds to its cognate receptors integrin- $\alpha 3\beta 1$ and integrin- $\alpha 6\beta 1$, which

are expressed by astrocytes and whose activation is required for endfeet anchorage (49, 50). Disrupted deposition of Lama-4 from ECs in vivo in *Reln*^{-/-} and vascular *Dab1* mutants also had an effect on the functional activation of integrin- $\beta 1$. Analysis of the ratio of staining at the vessel wall versus the intracellular compartment of the endothelium revealed that Lama-4 accumulated

at the vessel wall, and that such accumulation correlated with sites of integrin- $\beta 1$ activation in control but not in *Reln*^{-/-} or *Dab1*^{iAEC} mice (Fig. 8, A and B). Integrin- $\beta 1$ is expressed by multiple cell types of the CNS, including vessels and astrocytes (51). Integrin- $\beta 1$ activation also accumulated at the astrocytic endfeet, as evidenced by colocalization with the endfeet marker Aqp4

Fig. 6. Endothelial *Dab1* regulates NVU integrity.

(A) Fluorescent whole-brain images of *Dab1*^{iAEC} and control littermates injected with Alexa Fluor 555 cadaverine (Cad-A555) at P7. (B) Quantification of fluorescence intensity in (A) ($n = 6$ animals per genotype). (C) Cad-A555 and podocalyxin (Pdx) staining of P7 *Dab1*^{iAEC} cortices. (D) Fluorescent whole-brain images of adult *Dab1*^{iAEC} and control littermates injected with Cad-A555. (E) Quantification of fluorescence intensity in (D) ($n = 5$ to 7 animals per genotype). (F) Cad-A555 and Pdx staining of 4- to 7-week-old *Dab1*^{iAEC} cortices. (G) Fluorescent whole-brain images of 3- to 4-week-old *Reln*^{-/-} and control littermates injected with Cad-A555. (H) Quantification of fluorescence intensity in (G) ($n = 3$ animals per genotype). (I) Cad-A555 and Pdx staining of *Reln*^{-/-} cortices. (J) Transmission electron microscopy images of control and *Reln*^{-/-} cortical vessels. (K) Quantification of the distribution of the tight junction (TJ) angle in (J) ($n = 3$ animals per genotype). (L) Quantification of the number of horseradish peroxidase (HRP)-filled vesicles relative to the lumen length in (J) ($n = 3$ animals per genotype). (M) Fluorescent whole-retina images of P7 *Dab1*^{iAEC} and control littermates after intravenous injection of Cad-A555. (N) Quantification of fluorescence intensity corresponding to (M) ($n = 6$ to 8 retinas, 5 animals per genotype). (O) Representative confocal images showing an elevated cadaverine extravasation in *Dab1*^{iAEC} retinas from the blood vessels (IB4). Scale bars, 5 mm [(A), (D), and (G)], 50 μ m [(C), (F), and (I)], 100 nm (J), 2 mm (M), 75 μ m (O). Data are means \pm SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



in control animals, and its coexpression was dimmer in the mouse mutants (Fig. 8C). Vascular signaling downstream of endothelial integrin- $\beta 1$ has been shown to be important for the proper localization of VE-cadherin, and inactivation of endothelial integrin- $\beta 1$ signaling leads to aberrant VE-cadherin distribution and extensive hemorrhaging in retinas (52). Reelin $^{-/-}$ and Dab1 iAEC did not show any alteration in VE-cadherin arrangements in the vessels (fig. S11E), which suggests that the lack of Lama-4 deposition in the reelin signaling mutants preferentially affected the endfeet's integrin- $\beta 1$ activation.

Direct stimulation of brain ECs with exogenous reelin led to a specific increase in Lama-4 secre-

tion (Fig. 8, D and E) without changes in total mRNA transcript and protein levels (fig. S12, A and B). Conversely, Lama-5, which is also expressed by brain ECs, was not affected by reelin stimulation (fig. S12, C to F). Also, total protein levels for endothelial laminins were not affected in vivo (fig. S12, G to J). Reelin-induced secretion of Lama-4 had a direct impact on the attachment of primary astrocytes to monolayers of brain ECs. Stimulation of brain ECs with reelin led to a significant increase in the attachment of seeded astrocytes; however, pretreating primary astrocytes with a blocking antibody against integrin- $\beta 1$ (53) impaired the adhesion of the astrocytes onto the EC monolayer in a dose-dependent man-

ner (Fig. 8F and fig. S13). Furthermore, an in vitro BBB model using a coculture of ECs apposed to astrocytes showed that reelin stimulation of the endothelial compartment significantly increased the tightness of the barrier and such effect was blocked when astrocytic integrin- $\beta 1$ was inhibited (Fig. 8G), mimicking the leakiness we observed in the reelin signaling mutants in vivo. These results indicate that Lama-4 is required for integrin- $\beta 1$ -dependent astrocytic adhesion and barrierogenesis.

The same defects in Lama-4 secretion were recapitulated at the missing contacts of the radial glia endfeet with the pial vessels during embryonic development (Fig. 8, H to J) that resulted in impaired neuronal migration (Figs. 4

Fig. 7. Astrocytic endfeet attachment is reduced in reelin signaling mutants.

(A) Aquaporin4 (Aqp4), Pdx, and Cad-A555 staining of 4- to 7-week-old Dab1 iAEC cortices.

Defects in Aqp4 coverage are detected in larger vessels (arrow) as well as smaller capillaries (arrowhead).

(B) Aqp4 vessel coverage quantification as percentage of vessel area covered by Aqp4 staining in (A).

Vessels of wide-ranging diameter were included in the quantifications ($n = 3$ to 6 animals per genotype).

(C) Western blot for Aqp4 in brain lysates from Dab1 iAEC mutants. Pan-cadherin (pan-Cadh) was used as a loading control.

(D) Laminin- $\alpha 4$ (Lama-4), laminin- $\alpha 2$ (Lama-2), and albumin staining of 3- to 4-week-old Reelin $^{-/-}$ cortices.

(E) Lama-4 and Aqp4 staining of Reelin $^{-/-}$ cortices. Arrows indicate sites of colocalization of Lama-4 with Aqp4, which are decreased in the mutant (arrowheads).

Boxes indicate magnification areas in (F).

(F) Magnified pictures from (E) and line intensity profiles of staining intensity for Lama-4 (black lines) and Aqp4 (purple lines). Note the co-incident distribution of staining in the control vessels, which is lost in Reelin $^{-/-}$ vessels. Scale bars, 50 μ m [(A), (D), (E), (F)]. Data are means \pm SEM. *** $P < 0.001$.

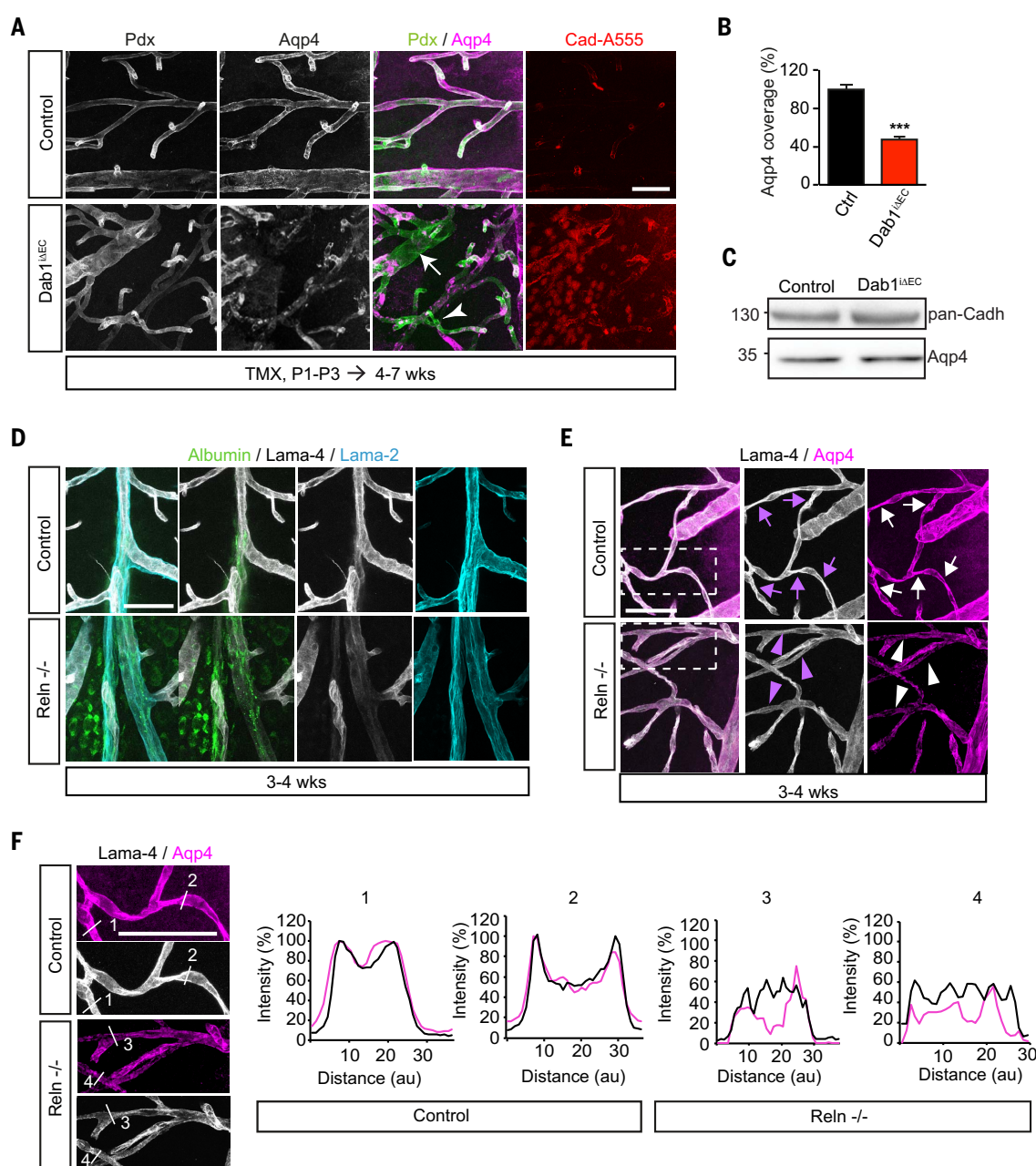
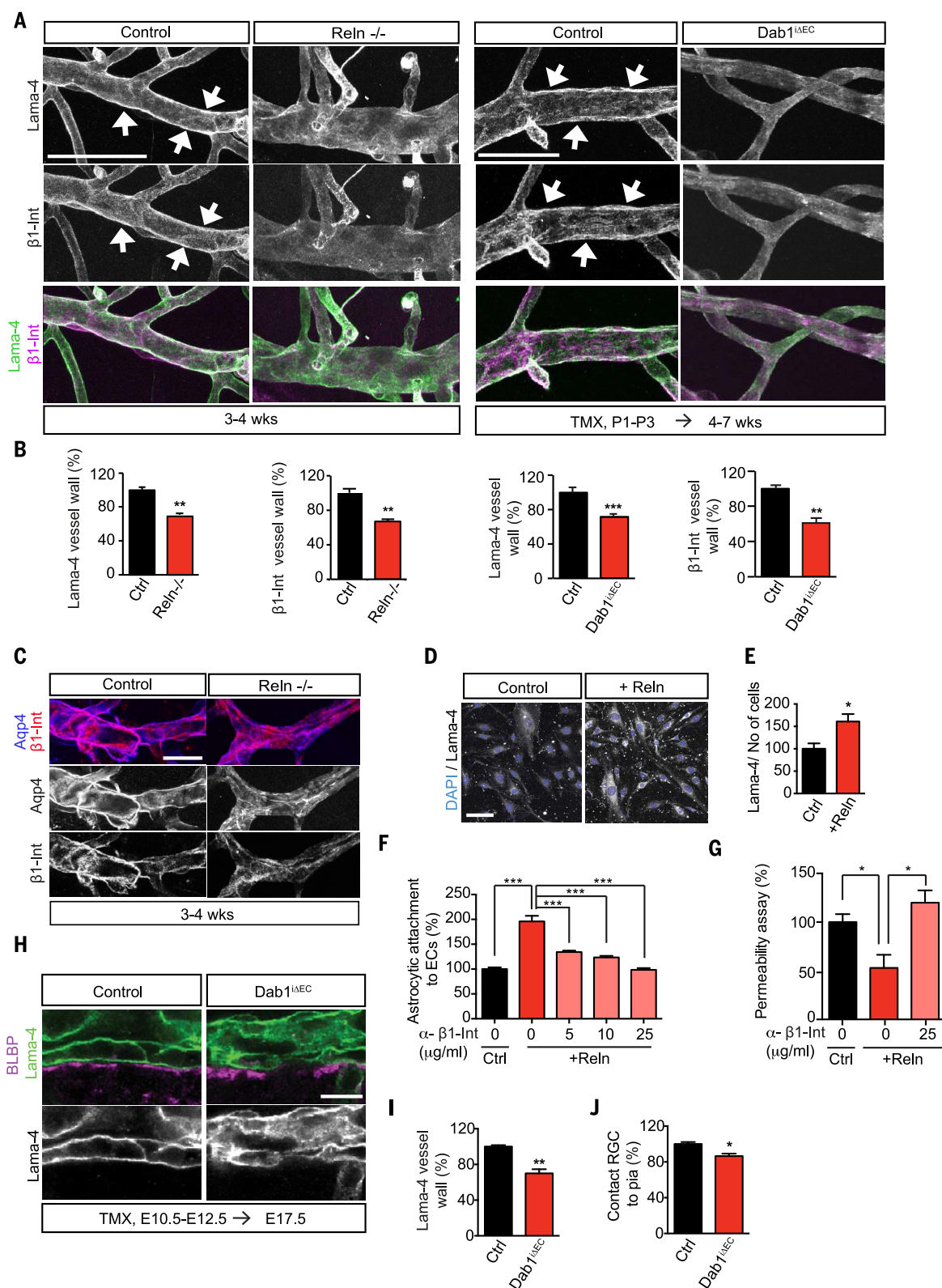


Fig. 8. Dab1 instructive roles on astrocytic endfeet attachment are mediated by vascular laminin- α 4 secretion and astrocytic integrin- β 1 activation. (A) Lama-4 and activated integrin- β 1 (β 1-Int) staining of Reln $^{-/-}$ and Dab1 $^{\Delta EEC}$ cortices. Arrows indicate areas of Lama-4 and β 1-Int enhanced colocalization. **(B)** Quantification of Lama-4 and β 1-Int relative signal at the vessel wall in (A) [$n = 3$ (Reln $^{-/-}$), $n = 3$ or 4 (Dab1 $^{\Delta EEC}$) animals per genotype]. **(C)** Representative pictures of β 1-Int and Aqp4 co-staining, showing the activation of β 1-Int in the astrocytic endfeet. **(D)** Lama-4 staining in reelin-stimulated brain ECs (bEND.3). **(E)** Representative quantification of Lama-4 fluorescent signal in (D) ($n = 7$ or 8 images per condition, 3 experiments). **(F)** Representative experiment of attachment of primary astrocytes to reelin-stimulated bEND.3 cells. The adhesion of astrocytes to the EC monolayer is impaired after blocking astrocytic β 1-Int ($n = 5$ measurements per condition, 3 experiments). **(G)** Permeability assay based on an in vitro BBB model of cocultured ECs and astrocytes ($n = 3$ experiments). bEND.3 cells in the luminal compartment were stimulated with reelin 24 hours before astrocytic seeding in the abluminal compartment of the insert. **(H)** Lama-4 and brain lipid binding protein (BLBP) staining of the pial region of Dab1 $^{\Delta EEC}$ embryos at E17.5 after TMX administration from E10.5 to E12.5. **(I)** Quantification of Lama-4 relative signal at the vessel wall in (H) ($n = 3$ animals per genotype). **(J)** Quantification of BLBP $^{+}$ cells contacting pial vessels in (H) ($n = 3$ animals per genotype). Scale bars, 50 μ m (A), 25 μ m [(C) and (D)], 10 μ m (H). Data are means \pm SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



Scale bars, 50 μ m (A), 25 μ m [(C) and (D)], 10 μ m (H). Data are means \pm SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

and 5). These findings suggest that the same mechanism applies to the instructive role of vascular Dab1 in neuronal migration.

Discussion

Our study provides precise mechanistic insights into an important function of the vasculature that goes beyond the supply of oxygen and nutrients and extends into an instructive structural role in building up functional architecture in the brain. Dab1 expressed in ECs is necessary to instruct the communication of vessels with the glia for the proper positioning of neurons during cortical development as well as for the correct communication at the NVU. We described how different neurovascular processes are synchronically regulated by the convergence on vascular Dab1 signaling. On one hand, reelin/ApoER2/Dab1 signaling regulates angiogenesis in the CNS by cross-talking to the VEGF/VEGFR2 pathway; on the other hand, vascular Dab1 regulates the deposition of Lama-4 and the docking of the radial glia/astrocytes on the vasculature, consequently affecting the neuronal migration at developmental stages and the etiology of BBB integrity.

Reelin in the vasculature exerts proangiogenic functions during development via the interaction with VEGF/VEGFR2 pathway. This function of Dab1 in ECs regulates proliferation and tip/stalk cell fate. The vascular growth defects observed during embryonic and postnatal stages are recovered during adulthood both in the retinal and in the cortical vessels. This suggests that other compensatory mechanisms are in place to correct for such defects in the vasculature. Dab2, a relative protein to Dab1, has been shown to be involved in developmental angiogenesis by controlling VEGFR2 endocytosis (54) and therefore could be a good candidate to compensate for the lack of Dab1 in the vessels. In agreement with this, we observed a down-regulation of Dab1 in the adult vessels, whereas Dab2 continues to be expressed during adulthood (55).

However, the neuronal defects provoked by the loss of function of Dab1 in the embryonic and early postnatal vasculature persist to adulthood and have a severe impact on neuronal positioning and BBB function. Dab1 in neurons is essential to guide correct neuronal positioning in the cortex during development (36, 56). Dab1 cell-autonomous function in neurons seems to be involved specifically in the somal translocation process during radial neuronal migration but not in glia-mediated locomotion (36). We have shown that Dab1 in the vessels is essential for both the maintenance of the RGC scaffold supporting migration of later-born neurons and their proper navigation toward the correct cortical layer during somal translocation. Dab1 deficiency in the vessels leads to detachment of the glial processes observed during both embryonic and early postnatal developmental stages, when the endfeet of the RGCs are anchored to the pial basement membrane and later when the mature cortical astrocytes extend their endfeet to surround the vessels and form a functional BBB. In agreement

with our results, several studies have shown that RGC detachment of the pial surface induces deficiencies in neuronal positioning (39, 40, 57). Notably, neuroglial integrin- β 1 conditional mutant mice showed important layering defects, including neuronal invasion to the marginal zone, originated by the deficient anchorage of radial glia processes to the meningeal surface (39), which we have shown to be recapitulated by the Dab1 endothelial-specific mutant mice. Furthermore, it has also been shown that proper neuronal migration in the cortex requires the selective expression of integrin- β 1 by RGCs but not by neurons (58), hence the activation of integrin- β 1 in glial cells is crucial for proper corticogenesis.

The basement membrane components Lama-2 and Lama-4 are involved in the attachment of the radial glia processes to the meningeal surface (59). In line with these studies, we found that vascular reelin signaling regulates the secretion of Lama-4 by ECs, which in turn has an impact on radial glia attachment to the pial surface and regulates neuronal migration. Although the regulatory mechanisms of Lama-4 deposition are still largely unknown, it has been reported that Lama-4 is accumulated in focal adhesions (60), which might also be influenced by reelin signaling (61).

Defects in Lama-4 deposition and integrin- β 1 activation also altered the assembly of the astrocytic endfeet ensheathment on the vasculature and consequently resulted in increased BBB permeability. Astrocytes exert a crucial function in maintaining the BBB in different physiological and pathological conditions (62) through their endfeet anchorage to the basal lamina proteins, which are relevant elements regulating BBB function in health and disease (46). Interestingly, Lama-4 null mice suffer from multiple hemorrhages starting at embryonic stages (63). Although the severe defects derived from the complete abrogation of Lama-4 expression are not comparable to the local defects in Lama-4 deposition observed in our mutants, it is remarkable that both animal models show a loss of BBB integrity.

Changes in Lama-4 secretion could regulate both endothelial and glial integrin- β 1 activation. However, our vascular Dab1 loss-of-function mutants phenocopy the defects shown by several reports after neuroglial deletion of integrin- β 1 (39, 58, 59). Conversely, vascular loss of integrin- β 1 has rather divergent phenotypes relative to those seen with vascular Dab1 deletion, such as hyperproliferation, sprouting, or altered VE-cadherin patterning (52, 64). In addition, we have shown that astrocytic adhesion to endothelial cells and astrocytic-induced barrier properties are dependent on reelin signaling on ECs and that astroglial integrin- β 1 activation is indispensable for such processes. Hence, on the basis of previous literature and our observations, we deduce that astroglial integrin- β 1 activation plays a relevant role in endothelial-astroglial interaction.

Together, our findings constitute a fundamental example of the integration of the neurovascular signaling that converges in the organization of the

neuronal-(astro)glial-vascular elements that regulate the development and homeostasis of the CNS. Such integrative signaling might explain, at the molecular level, the comorbidity of neuropathological conditions and vascular dysfunction.

Materials and methods

Genetically modified mice and treatments

Dab1^{fllox/flox} mice carrying loxP sites flanking exon 2 for the *Dab1* gene (36) (kindly provided by U. Mueller) were crossed with Cdh5(PAC)-CreERT2 (65) (kindly provided by R. Adams) to generate endothelial cell-specific Dab1 knockout mice (Dab1^{iAEC}). VEGFR2^{fllox/flox} mice (66) (kindly provided by E. Wagner) were crossed with Cdh5(PAC)-CreERT2 (65) to generate endothelial cell-specific VEGFR2 knockout mice (VEGFR2^{iAEC}). Cre activity was induced by intraperitoneal injection of 0.1 ml of tamoxifen (1 mg/ml) or 4-hydroxytamoxifen (4-OHT) (1 mg/ml) each day for 3 days (P1 to P3) for postnatal analysis at P3–P4, P7–P8, and postnatal weeks 4 to 7. For adult induction, 0.1 ml of 4-OHT (5 mg/ml) was intraperitoneally injected for 3 consecutive days (P25 to P27) for analysis at P30. For embryonic induction, pregnant females were injected intraperitoneally with 0.2 ml of 4-OHT (10 mg/ml) each day for 3 days (E10.5 to E12.5). Tamoxifen injectable solution was prepared in ethanol and peanut oil as described (67). Cre-negative animals were used as controls. Additionally, the Cdh5(PAC)-CreERT2 line was crossed with the ROSA26R(EYFP) reporter strain and genetic recombination was induced as described for the Dab1^{iAEC} mice. Reelin knockout (Reln^{-/-}) and ROSA26R(EYFP) mice were obtained from Jackson Laboratories; ApoER2 knockout (ApoER2^{-/-}) mice were kindly provided by J. Herz (68). For Reln^{-/-} mice, both wild-type and heterozygous littermates were used as controls, except as mentioned. Wild-type C57BL/6J animals were used for fluorescent in situ hybridization (FISH) analysis, isolation of primary cell cultures, and retinal explant cultures. Both males and females were used for all experiments indistinctively.

All animals were genotyped by PCR. Protocols and primer sequences were used as described by the distributor or donating investigator.

For intravenous Alexa Fluor 555 cadaverine tracer injection and detection, mice were deeply anesthetized by injection of ketamine and xylazine at 180 mg/kg and 10 mg/kg of body weight, respectively; 50 μ l of Alexa Fluor 555 cadaverine (1 mg/ml) was injected into the retro-orbital venous sinus in 4- to 7-week-old (Dab1^{iAEC}) or 3- to 4-week-old (Reln^{-/-}) mice as described (69). For postnatal mice (P7), 40 μ l of Alexa Fluor 555 cadaverine (1 mg/ml) was injected intraperitoneally. Cadaverine was allowed to circulate for 20 min in adult mice and for 2 hours in postnatal mice, respectively. Mice were then perfused intracardially with 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Eyes from P7 animals were fixed in 4% PFA for 2 hours at room temperature (RT) before retina isolation in PBS. Eyes from 4- to 7-week-old adult mice were fixed in 4% PFA

for 10 min prior to retina isolation in 4% PFA and post-fixed overnight (4% PFA, 4°C). Brains were post-fixed at RT in 4% PFA for 3 hours. For cadaverine detection, whole brains and isolated retinas (mounted in PBS) were imaged using a dissecting microscope with an attached fluorescent lamp and a Texas Red filter and subsequently processed for immunostaining. Cadaverine leakage was quantified by measuring mean fluorescence intensity of whole-brain pictures.

For BrdU experiments, pregnant females were injected intraperitoneally with 0.2 ml of BrdU solution (10 mg/ml) dissolved in sterile PBS at E15.5. Brains from the offspring were collected and analyzed at P8.

For intravenous peroxidase injection, horse-radish peroxidase Type II (HRP) was dissolved at 100 mg/ml in sterile PBS, injected at 10 mg per 20 g body weight in deeply anaesthetized mice, and let to circulate for 30 min. Mice were perfused with 0.3 M HEPES, 1.5% PFA, and 1.5% glutaraldehyde (GDA) and processed for histochemistry.

For intracardiac isolectin B4 (IB4), mice were processed the same way as for regular perfusion. Prior to the PBS/4% PFA perfusion, 1 ml of IB4 (20 ng/μl) was intracardially injected.

All animal experiments were approved by the Regierungspräsidium of Darmstadt and the Veterinäramt of Frankfurt am Main.

Cell culture and treatment

For isolation of primary mouse brain endothelial cells, mouse brain microvascular fragments were processed as described (70). Briefly, capillary fragments were seeded on collagen I-coated wells and cultured in DMEM (Dulbecco's modified Eagle's medium) with 20% fetal bovine serum (FBS) supplemented with heparin (100 μg/ml) and EC growth supplement (ECGS; 5 μg/ml). After 2 days of puromycin selection (4 μg/ml), cells were cultured for two more days without puromycin before their experimental use.

Mouse primary lung endothelial cells (MLECs) were isolated as described (27). Lungs were isolated into dissection buffer (HBSS supplemented with 10% FBS) and treated with collagenase type II. Filtered tissue pellets were resuspended in dissection buffer and incubated with anti-rat IgG-coated magnetic beads, pre-coupled with rat anti-mouse CD31. Beads were resuspended in endothelial cell medium consisting of high-glucose DMEM GlutaMAX-I, penicillin (100 U/ml), streptomycin (100 μg/ml), 20% FBS, 0.4% endothelial cell growth supplement with heparin and plated onto gelatin-coated plates.

For isolation of mouse primary astrocytes, brains of P2–P5 mice were isolated into dissection buffer (DMEM) and subsequently incubated in 0.25% trypsin/EDTA for 20 to 30 min at 37°C. Digested tissue was homogenized with a fire-polished Pasteur pipette and centrifuged. Pelleted cells were resuspended in astrocyte medium consisting of high-glucose DMEM GlutaMAX-I, penicillin (100 U/ml), streptomycin (100 μg/ml), 10% FBS, 1% MITO⁺ Serum Extender and plated onto flasks. After the cultures were confluent, the

astrocytes were purified by shaking the flasks on a rotator at 250 rpm at 37°C for 2 days to detach all other cell types.

Pooled human umbilical vein endothelial cells (HUVECs) were cultured in endothelial basal medium (EGM) supplemented with hydrocortisone (1 μg/ml), bovine brain extract (3 μg/ml), gentamicin sulfate (30 μg/ml), amphotericin B (50 μg/ml), EGF (10 μg/ml), and 10% FBS on gelatin-coated culture dishes at 37°C, 5% CO₂.

For the detection of phospho-Dab1 (pDab1) by immunostaining upon reelin stimulation, HUVECs were starved in DMEM containing 2% FBS overnight. Stimulation was performed by adding 150 μl of reelin supernatant or green fluorescent protein (GFP) supernatant for 45 min. For the detection of pDab1 by immunostaining upon VEGF-A stimulation, HUVECs were starved in serum-free medium (EGM) for 4 hours prior to stimulation with recombinant VEGF-A (50 ng/ml) for 30 min. For costimulation with VEGF-A and reelin, HUVECs were stimulated with 40-fold concentrated reelin supernatant and recombinant VEGF-A (50 ng/ml) for 30 min. Mouse primary brain endothelial cell (MBEC) cultures were starved with Opti-MEM serum reduced medium at 37°C, 2 hours prior to stimulation. For stimulation, MBECs were treated with 40-fold concentrated reelin supernatant or GFP supernatant. For the detection of pDab1, cultures were stimulated for 15 min at 37°C. Dab1 phosphorylation was assessed by measuring pDab1 fluorescence intensity (HUVECs and MLECs) or by measuring the area covered by pDab1 fluorescence (MBECs).

For Western blot analysis, cells were starved in EGM for 1 hour prior to stimulation. Stimulation was performed by adding recombinant VEGF-A (50 ng/ml), recombinant VEGF-C (100 ng/ml), recombinant reelin fragment (100 ng/ml unless mentioned otherwise) for 15 min or reelin supernatant for 30 min.

The immortalized mouse brain microvascular endothelial cell line bEND.3 was grown in DMEM supplemented with 10% FBS and antibiotics. For the detection of Lama-4 and Lama-5 by immunostaining upon reelin stimulation, bEND.3 cells were starved in minimum essential medium (MEM) containing penicillin/streptomycin (100 μg/ml) for 2 hours. Stimulation was performed by adding recombinant reelin fragment (100 ng/ml) overnight. Cells were fixed with 4% TCA for 10 min at RT and processed for immunostaining. For quantifications, % of area covered by positive pixels of Lama-4 staining was measured using ImageJ and normalized to the number of cells (number of DAPI⁺ nuclei) per field. Fisher's compared *t*-values test was used to estimate the significance among the biological replicates.

Preparation of reelin-containing and control supernatants

To obtain reelin-enriched supernatants and GFP control supernatants, incubation medium [DMEM, penicillin (100 U/ml), streptomycin (100 μg/ml), G418 (0.360 g/liter), 10% FBS] from reelin-transfected 293-HEK cells or GFP-transfected control 293-HEK cells (a gift from M. Goetz) was

replaced by serum-free medium containing penicillin/streptomycin (100 μg/ml) and cells were incubated for 2 days at 37°C, 5% CO₂. The conditioned medium was collected and concentrated 40-fold by centrifugation using filter units. Reelin content, as well as its absence in control cell supernatants, was confirmed by Western blotting using mouse anti-reelin antibody 1:1000 (Millipore; MAB5364).

Tube formation assay

Tube formation assays were performed using μ-slides angiogenesis. Experimental settings were set up as suggested by the manufacturer with minor modifications. Briefly, 2 × 10⁴ HUVECs were cultured on growth factor-reduced Matrigel at 37°C in a humidified incubator supplied with 5% CO₂ for 6 to 8 hours in the presence of concentrated reelin supernatant or control GFP supernatant. The tubular network was quantified by counting the number of branch points and by measuring total tube length.

Mouse retinal organotypic explants

Retinal explant experiments were performed as described (24). In short, eyes were enucleated from newborn pups (P3–P5) and transferred to FBS-free DMEM medium. Retinas were dissected from eyecups and vitreous bodies were removed. Retinas were flat-mounted onto the hydrophilic polytetrafluoroethylene (PTFE) membrane of culture plate inserts with the nerve fiber layer facing the membrane. DMEM with 10% FBS was layered underneath the membrane and dropped on the retinas to prevent dryness. Retina explants were incubated at 35°C in a humidified incubator with 5% CO₂ for 2 to 4 hours before stimulation. For stimulation of retina endothelial tip cells, VEGF-A was diluted in DMEM or Opti-MEM I and 3% FBS to a final concentration of 1 μg/ml and layered underneath the insert membrane and dropped on the explants. 40-fold concentrated reelin and control GFP supernatants were used for reelin stimulation experiments. For the combined VEGF-A/reelin stimulation experiments, VEGF-A (100 ng/ml) was diluted in the concentrated reelin or GFP supernatants. Stimulation was carried out at 35°C in a humidified incubator with 5% CO₂ for 4 hours. Explants were fixed with 4% PFA at RT for 30 min and stained with IB4 (1:200). For quantifications, numbers of filopodia at the vascular front were counted and normalized against 100 μm vessel length.

Proximity ligation assay

Endothelial cells were fixed with 4% PFA in PBS for 10 min at RT and permeabilized using 0.1% Triton X-100 in PBS for 4 min on ice. Subsequently, the cells were washed with PBS, and blocking solution (2% BSA, 4% NDS in PBS) was applied for 30 min at 37°C in a humidified chamber. Proximity ligation assay was performed as suggested by the manufacturer. All incubation periods were performed at 37°C. In brief, primary antibodies against ApoER2 1:80 (Abcam; ab86548) and VEGFR2 1:80 (R&D Systems; AF644) in blocking solution were added for 30 min. The

cells were incubated with the corresponding PLA probes for 60 min. Phalloidin-FITC 1:500 was added to visualize actin filaments. Ligation and amplification of the probes was performed for 30 min and 100 min, respectively. Cells were incubated with DAPI, washed, and mounted. Incubation times with antibodies occurred at 4°C and overnight for the primary antibodies and 1 hour at 37°C for the PLA probe antibodies. For quantification, the number of PLA probe punctae per cell was counted.

Immunohistochemistry, immunocytochemistry, and FISH

All samples from mutant and littermate controls in an experiment were always processed and stained at the same conditions.

Brains were dissected and post-fixed for 3 hours in 4% PFA or 10% TCA at RT for immunostainings or overnight in 4% PFA at 4°C for FISH. For immunostainings, brains were sectioned coronally at 80 µm using a vibratome. For FISH, retinas and brains were cryoprotected by consecutive immersions in 15% and 30% sucrose in PBS at 4°C. Samples were then embedded in Tissue-Tek O.C.T. compound and frozen on dry ice. Coronal sections with a thickness of 16 µm were generated using a cryostat microtome. For retinal whole-mount staining, eyes were collected from mutant mice and their control littermates and fixed in 4% PFA solution overnight at 4°C or for 2 hours at RT.

Embryonic, postnatal, adult brain, and retina sections were incubated with primary antibodies after 30 min or 1 hour of blocking and permeabilization with 5 to 10% normal donkey serum (NDS), 0.5% Triton X-100 in PBS, respectively. For BrdU immunodetection, 80 µm vibratome sections were pretreated with 2 N HCl for 30 min and subsequently neutralized with sodium tetraborate (Na₂B₄O₇, 0.1 M).

For immunostaining of whole retinas, retinas were isolated, blocked, and permeabilized in 5% NDS and 0.5% Triton X-100 in PBS at RT for 1 hour. Primary antibodies were diluted in 5% NDS in PBS. Incubations were performed overnight at 4°C. For retinal vessel visualization, retinas were incubated with IB4 (1:200) in 1% Triton X-100 in PBS.

Primary brain endothelial and bEND.3 cells were fixed in cold 10% TCA for 10 min and incubated with primary antibodies in 2% NDS in PBS for 2 hours at RT after 15 min of blocking and permeabilizing with 5% NDS and 0.1% Triton X-100 in PBS.

HUVEC and MLEC cultures were fixed with 4% PFA for 20 min, incubated with NH₄Cl for 10 min at RT, blocked and permeabilized with 4% NDS, 0.2% Triton X-100 and 2% bovine serum albumin (BSA) in PBS for 30 min. Cells were incubated with the primary antibody in blocking solution for 1 hour at RT.

The following primary antibodies were used: rabbit anti-Cux1 1:100 (Santa Cruz; SC-13024), rabbit anti-aquaporin4 1:100 (Millipore; AB2218), mouse anti-BrdU 1:200 (Millipore; MAB3424), rabbit anti-ERG 1:200 (Abcam; ab92513), rabbit anti-phospho-Histone H3 1:200 (Millipore; 06-570),

rat anti-mouse CD29 (activated integrin-β1) 1:100 (BD Pharmingen; 550531), rabbit anti-brain lipid binding protein (BLBP) 1:100 (Millipore; ABN14), rabbit anti-Glut1 1:200 (07-1401; Millipore), rat anti-laminin α-2 1:100 (abcam; ab11576), goat anti-laminin α-4 1:100 (R&D; AF3837), rabbit anti-laminin α-5 1:100 (Novus Biologicals; NBP1-18714), mouse anti-NeuN 1:200 (Millipore; MAB377B), rabbit anti-NeuN 1:200 (Millipore; ABN78), rabbit anti-Pax6 1:200 (Covance; PRB-278P), rabbit anti-Tbr1 1:200 (Abcam; ab31940), rabbit anti-Tbr2 1:200 (Abcam; ab183991), goat anti-PDGFRβ 1:200 (Neuromics; GT15065-100), rabbit anti-phospho-Dab1 (Y232) 1:200 (Cell Signaling Technology; 3325S), rat anti-VE-cadherin 1:100 (BD Pharmingen; 555289) and goat anti-reelin 1:100 (R&D Systems; AF3820). After primary antibody incubation, samples were washed with PBS (cells and retinas) or TBS-T (150 mM NaCl, 25 mM Tris base, 0.1% Tween 20; pH 7.6; brain sections) for three times and incubated with the appropriate fluorophore-coupled secondary antibodies for 1 hour at RT (cells and postnatal retinas) or overnight (brain sections and adult retinas). The secondary antibodies were Alexa Fluor 488-, 555-, 568- and 647-conjugated donkey anti-rabbit/mouse/goat 1:200 (Life Technologies) or anti-rabbit-Cy3 secondary antibody 1:200 (Jackson ImmunoResearch). Nuclei were counterstained with DAPI. Sections were washed with PBS or TBS-T before mounting them using fluorescence mounting medium. Proliferating endothelial cells were quantified as described (71). The tip cell/stalk cell ratio was assessed as suggested by others (72, 73). Retinal tip cell filopodia quantification was performed as described (27). The number of branch points in the retina vasculature was assessed by counting the number of branch points between artery and vein and normalizing against the selected area. Vascular density in adult retinas was assessed by thresholding and measuring the area covered by IB4 signal in all three layers of the adult retina. Brain vascular parameters (vessel density, total vessel length, vessel orientation, and number of branch points) were analyzed in the cortical area using AngioTool (74) and ImageJ software (75). Aqp4 coverage was quantified by measuring the area of the vessel covered with staining. The area of Aqp4 staining was normalized to the vessel area (podocalyxin staining). Lama-4 and integrin-β1 signals were quantified by measuring the immunofluorescence intensity at the vascular walls normalized to the intracellular compartment. Metamorph software was used for these morphometric analyses.

For FISH, whole C57BL/6 adult mice brains were dissected, and RNA was extracted using TRIzol reagent. RNA was reverse-transcribed into cDNA using High Capacity cDNA Reverse Transcription Kit and the resulting cDNA used to obtain the PCR products. The primer sequences used for each probe are: *Dab1*-probe1-Fw AAC-CTGTTATCTCGACTTGA, ISH: *Dab1*-probe1-Rv TGAACAAGGGGCTGCTGGCC, ISH: *Dab1*-probe2-Fw, GTCCATAAATCATGGGACTGGT, ISH: *Dab1*-probe2-Rv, TGGAGAGACTCAGATAGCCACA, ISH: ApoER2-Fw, TCTACTGGACAGACTCAGGCAA and

ISH: ApoER2-Rv, CCGTAGCATCTCTTCATGTCTG. Probes for *Dab1*-probe2 and *ApoER2* were obtained from Allen Brain Atlas: <http://developing-mouse.brain-map.org/experiment/show/79762299>; <http://developingmouse.brain-map.org/experiment/show/100045438>. PCR conditions, PCR product purification, cloning, transformation, and plasmid amplification were done as described (76). Plasmids with the right sequence were linearized using restriction enzymes and purified with Wizard SV Gel and PCR Clean-Up System. Finally, linear plasmids were transcribed into RNA probes labeled with digoxigenin (DIG). FISH was performed as described (76) with minor modifications: (i) slices of P7-8 brains and retinas were incubated in proteinase K solution (12 µg/ml) at 37°C for 12 min, (ii) *Dab1* detection was performed by mixing two different riboprobes mixed together, (iii) anti-DIG-alkaline phosphatase was incubated together with goat anti-podocalyxin 1:200 (R&D Systems; AF1556), (iv) signal of DIG was detected using HNPP/Fast Red and podocalyxin with Alexa Fluor 647-conjugated donkey anti-goat, 1:250 (Life Technologies).

Images were taken using a laser scanning confocal spectral microscope. Brightness and contrast of the images were adjusted using the software Adobe Photoshop CS6 or ImageJ. Figures were prepared using Adobe Illustrator CS5.1.

Transmission electron microscopy (TEM)

For assessing the rate of transcytosis, HRP-diaminobenzidine histochemistry was performed on brain vibratome sections from animals injected with HRP (see above). Sections were incubated with 3,3'-diaminobenzidine tetrahydrochloride (DAB) for 30 min at RT and washed with PBS. Subsequently, samples were processed for TEM imaging.

Tissue was post-fixed with 6% glutaraldehyde/0.4 M PBS for 24 hours at RT and subsequently washed 5 times in 0.1 M Epon-PBS. Small tissue samples were cut out from the parasagittal area and processed with a tissue processor with 1% osmium tetroxide. Dehydration steps were followed by using increasing ethanol concentrations (25%, 35%, 50%, 70%, 75%, 85%, 100%). Prior to embedding, tissue combined with resin (Agar 100 Resin Kit) was dehydrated in an exsiccator for 24 hours. After embedding, the samples were kept in a steaming cabinet at 60°C for minimum of 4 days.

From the resin-embedded tissue, ultrathin sections (0.23 µm) were cut with a microtome and placed at 200 mesh copper grids (3.05 mm). Ultrathin sections were then contrasted with EM AC20 (0.5% uranyl acetate/Ultrastain I and 3% lead citrate/Ultrastain II). Samples were examined with a transmission electron microscope equipped with a Slowscan-2K-CCD-digital camera (2K-wide-angle). Morphometric analyses were performed with ImageSp software. For quantifications, numbers of HRP-filled vesicles were normalized per length of vascular lumen, and the angle of tight junction in relation to the lumen surface was measured with ImageJ software.

Quantitative real-time PCR

Whole C57BL/6 adult mice brains were dissected, and RNA was extracted using TRIzol reagent. RNA was reverse-transcribed into cDNA using High Capacity cDNA Reverse Transcription Kit. Quantitative PCR assays were performed using an ABI 7500 Fast Real-Time PCR System using TaqMan Fast Universal PCR master mix and TaqMan Gene Expression probes for mouse *Dab1* (Mm01256039_m1), mouse *Reelin* (Mm00465200_m1), mouse *Vegf* (Mm00437306_m1), mouse *Laminin-α4* (Mm01193660_m1), mouse *Laminin-α5* (Mm01222029_m1), and mouse $\beta 2m$ (Mm00437762_m1), which served as an endogenous control.

PCR for *Dab1* gene excision

Genomic DNA was extracted from the tail of the animals and PCR was performed using 5 μ l of the DNA extract. The primers used to amplify the excised *Dab1* fragment were: 5'GGTTCAGTGCC-TATCATGTATC3'(Fwd) 5'CTTACTTTCTAGAGAATAGGAAC3'(Rv). PCR was performed with Promega GoTaq Green Master Mix using T_m = 54°C and 38 cycles. PCR product was loaded in a 3% agarose gel and DNA was stained with 0.01% ethidium bromide. After electrophoresis, gel was imaged in a transilluminator using a UV light source.

Attachment assay

bEND.3 cells were seeded in quintuplicates on a 96-well plate. bEND.3 cells were grown overnight at 37°C and were then starved with MEM containing penicillin/streptomycin (100 μ g/ml) for 2 hours at 37°C. After starvation, bEND.3 cells were stimulated with recombinant reelin at 100 ng/ml overnight at 37°C in starving medium. Primary astrocytes were incubated with 6 μ M Texas Red Hydrazide in order to label them fluorescently and with rat anti-mouse integrin- $\beta 1$ blocking antibody (BD Biosciences; 553715) at the indicated concentrations (53) for 30 min at 37°C. After the stimulation of the bEND.3 cells, astrocytes were pelleted, resuspended in medium and seeded (40,000/well) onto the bEND.3 cell monolayer for 3 hours at 37°C. After the attachment incubation, non-attached astrocytes were washed gently 3 times with PBS. Finally, 100 μ l of 1% SDS solution was added per well and fluorescent intensity was read (λ_{ex} 590 nm; λ_{em} 620 nm).

Permeability assay

We used an in vitro model of BBB based in a coculture of endothelial cells and astrocytes seeded in the opposite sides of a prehydrated 24-well membrane insert. 40,000 bEND.3 cells were seeded onto the luminal side of the insert and cultured in bEND.3 medium for 24 hours in a 37°C, 5% CO₂ cell culture incubator until monolayer was formed. Prior to reelin stimulation, bEND.3 cells were washed with PBS and starved in MEM for 2 hours in an incubator. bEND.3 cells were then stimulated with recombinant reelin (100 ng/ml) for 24 hours. Shortly prior to adding astrocytes, reelin was removed and bEND.3 cells were kept in bEND.3 medium.

Primary cortical astrocytes were pre-incubated with blocking integrin- $\beta 1$ antibody for 30 min. Pellet of blocked astrocytes was resuspended in bEND.3 medium. 40,000 astrocyte cells were seeded on the abluminal side of the insert to let them attach for 3 hours in the incubator. Prior to permeability assay, both luminal and abluminal sides of the membrane insert were washed with PBS to remove unattached astrocytes and the excess medium. Insert was transferred to a new 24-well plate filled with 500 μ l of PBS. In vitro BBB permeability was assessed by adding 200 μ l of sodium fluorescein (100 μ g/ml) to the luminal side of the insert. After permeation time of 15 min, the insert was transferred to a new well and another permeation was repeated (in total 4 times). The liquid in the well plate (now containing sodium fluorescein that crossed the cell layers) was thoroughly mixed and 100 μ l of this was transferred to 96-well plate. The plate was read in a fluorescence plate reader (λ_{ex} 460 nm; λ_{em} 515 nm).

Western blot

Tissue and cells were lysed in lysis buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 1% Triton X-100; 1 mM sodium orthovanadate; 10 mM NaPPi; 20 mM NaF) or RIPA buffer (150 mM sodium chloride; 1% Triton X-100; 0.5% sodium deoxycholate; 0.1% SDS; 50 mM Tris, pH 8.0) and 1% complete protease inhibitor cocktail (Complete EDTA-free Proteinase inhibitor cocktail tablets). Protein content was determined using Pierce BCA Protein Assay Reagent according to manufacturer's instructions and samples were separated by SDS-PAGE.

For immunoprecipitation of VEGFR2, endothelial cells were lysed with NET lysis buffer (50 mM Tris HCl buffer, pH 7.4, 15 mM EDTA pH 7.4, 1% NP-40, 150 mM NaCl, 10 mM sodium pyrophosphate, 20 mM NaF, 1 mM sodium orthovanadate, and 1% complete protease inhibitor cocktail) for 30 min at 4°C, centrifuged at 21,000g for 15 min and supernatants were collected. Samples were pre-incubated with protein G-Sepharose beads for 1 hour at 4°C. Beads were removed by centrifugation at 400g. Supernatants were incubated with protein G-Sepharose beads, rabbit anti-VEGFR2 antibody (Cell Signaling; 2479) and TBS/0.1% NP40 for 2 hours at 4°C. Samples were washed with NENT 300 washing buffer (20 mM Tris pH 7.4, 300 mM NaCl, 1 mM EDTA pH 7.4, 0.1% NP40, 25% glycerol) and TBS/0.1% NP40. For immunoprecipitation of *Dab1*, a similar protocol was applied with minor changes. Lysis buffer was used for lysis of MLEC, incubation of lysates with the antibody (goat anti-*Dab1* antibody, Abcam, Ab16674) coupled sepharose beads and washing of the beads prior to Western blot analysis.

Protein samples from total lysates or IP were boiled with sample buffer (8% SDS, 200 mM Tris-HCl pH 6.8, 400 mM DTT, 0.4% Bromophenol blue, 40% Glycerol) prior to separation by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were blocked in TBS-T with skimmed milk powder (3% or 5%) or BSA (5%), depending

on the antibody manufacturer's recommendation. The following antibodies were used: rabbit anti-phospho-Dab1 Y232 1:500 (Cell Signaling Technology; 3325), goat anti-*Dab1* 1:1000 (Abcam; Ab16674), rabbit anti-VEGFR2 1:1000 (Cell Signaling Technology; 2479), rabbit anti-phospho-VEGFR2 Y1175 1:1000 (Cell Signaling Technology; 2478), rabbit anti-ApoER2 1:1000 (Sigma-Aldrich; A3481), goat anti-laminin- $\alpha 4$ 1:1000 (R&D; AF3837), rabbit anti-laminin- $\alpha 5$ 1:1000 (Novus Biologicals; NBPI-18714), rabbit anti-aquaporin4 1:1000 (Millipore; AB2218). Goat anti-actin, 1:1000 (Santa Cruz; sc-1615) and mouse anti-pan-cadherin 1:1000 (Sigma; C1821) were used as a loading controls. Primary antibodies were incubated overnight at 4°C, and membranes were subsequently incubated with HRP-conjugated secondary antibodies goat anti-rabbit HRP, donkey anti-goat HRP and goat anti-mouse HRP 1:1000 (Jackson Immuno Research Laboratories) in blocking solution 2 hours at RT. HRP activity was detected using enhanced chemiluminescence detection reagent (ECL) and the ImageQuant LAS 4000 system.

Statistical analysis

Quantifications were normalized to control and represented as percentage of control (%), unless otherwise indicated. Statistical significance was determined using 2-tailed unpaired Student's *t*-test when comparing 2 variables, unless otherwise indicated. Statistical analysis was performed with Prism version 5. One-way analysis of variance (ANOVA) was performed in Prism Version 5 to assess statistical significance of the differences between multiple measurements. All animal experiments included animals from at least two litters. Statistical significance was defined as $P < 0.05$ (*), $P < 0.01$ (**) and $P < 0.001$ (***). All values indicate mean \pm SEM.

REFERENCES AND NOTES

1. D. Attwell et al., Glial and neuronal control of brain blood flow. *Nature* **468**, 232–243 (2010). doi: [10.1038/nature09613](https://doi.org/10.1038/nature09613); pmid: [21068832](https://pubmed.ncbi.nlm.nih.gov/21068832/)
2. M. Segarra, B. C. Kirchmaier, A. Acker-Palmer, A vascular perspective on neuronal migration. *Mech. Dev.* **138**, 17–25 (2015). doi: [10.1016/j.mod.2015.07.004](https://doi.org/10.1016/j.mod.2015.07.004); pmid: [26192337](https://pubmed.ncbi.nlm.nih.gov/26192337/)
3. S. Ma, H. J. Kwon, H. Johng, K. Zang, Z. Huang, Radial glial neural progenitors regulate nascent brain vascular network stabilization via inhibition of Wnt signaling. *PLOS Biol.* **11**, e1001469 (2013). doi: [10.1371/journal.pbio.1001469](https://doi.org/10.1371/journal.pbio.1001469); pmid: [23349620](https://pubmed.ncbi.nlm.nih.gov/23349620/)
4. J. M. Proctor, K. Zang, D. Wang, R. Wang, L. F. Reichardt, Vascular development of the brain requires $\beta 8$ integrin expression in the neuroepithelium. *J. Neurosci.* **25**, 9940–9948 (2005). doi: [10.1523/JNEUROSCI.3467-05.2005](https://doi.org/10.1523/JNEUROSCI.3467-05.2005); pmid: [16251442](https://pubmed.ncbi.nlm.nih.gov/16251442/)
5. P. Rakic, Mode of cell migration to the superficial layers of fetal monkey neocortex. *J. Comp. Neurol.* **145**, 61–83 (1972). doi: [10.1002/cne.901450105](https://doi.org/10.1002/cne.901450105); pmid: [4624784](https://pubmed.ncbi.nlm.nih.gov/4624784/)
6. C. Won et al., Autonomous vascular networks synchronize GABA neuron migration in the embryonic forebrain. *Nat. Commun.* **4**, 2149 (2013). doi: [10.1038/ncomms3149](https://doi.org/10.1038/ncomms3149); pmid: [23857367](https://pubmed.ncbi.nlm.nih.gov/23857367/)
7. S. Li et al., Endothelial cell-derived GABA signaling modulates neuronal migration and postnatal behavior. *Cell Res.* **28**, 221–248 (2018). doi: [10.1038/cr.2017.135](https://doi.org/10.1038/cr.2017.135); pmid: [29086765](https://pubmed.ncbi.nlm.nih.gov/29086765/)
8. F. Tissir, A. M. Goffinet, Reelin and brain development. *Nat. Rev. Neurosci.* **4**, 496–505 (2003). doi: [10.1038/nrn1113](https://doi.org/10.1038/nrn1113); pmid: [12778121](https://pubmed.ncbi.nlm.nih.gov/12778121/)
9. B. W. Howell, R. Hawkes, P. Soriano, J. A. Cooper, Neuronal position in the developing brain is regulated by mouse

- disabled-1. *Nature* **389**, 733–737 (1997). doi: [10.1038/39607](https://doi.org/10.1038/39607); pmid: [9338785](https://pubmed.ncbi.nlm.nih.gov/9338785/)
10. T. Hiesberger et al., Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces tyrosine phosphorylation of disabled-1 and modulates tau phosphorylation. *Neuron* **24**, 481–489 (1999). doi: [10.1016/S0896-6273\(00\)80861-2](https://doi.org/10.1016/S0896-6273(00)80861-2); pmid: [10571241](https://pubmed.ncbi.nlm.nih.gov/10571241/)
 11. H. H. Bock, J. Herz, Reelin activates SRC family tyrosine kinases in neurons. *Curr. Biol.* **13**, 18–26 (2003). doi: [10.1016/S0960-9822\(02\)01403-3](https://doi.org/10.1016/S0960-9822(02)01403-3); pmid: [12526740](https://pubmed.ncbi.nlm.nih.gov/12526740/)
 12. A. Sentürk, S. Pfennig, A. Weiss, K. Burk, A. Acker-Palmer, Ephrin Bs are essential components of the Reelin pathway to regulate neuronal migration. *Nature* **472**, 356–360 (2011). doi: [10.1038/nature09874](https://doi.org/10.1038/nature09874); pmid: [21460838](https://pubmed.ncbi.nlm.nih.gov/21460838/)
 13. D. S. Rice et al., Disabled-1 acts downstream of Reelin in a signaling pathway that controls laminar organization in the mammalian brain. *Development* **125**, 3719–3729 (1998). pmid: [9716537](https://pubmed.ncbi.nlm.nih.gov/9716537/)
 14. D. S. Falconer, Two new mutants, 'trembler' and 'reeler', with neurological actions in the house mouse (*Mus musculus* L.). *J. Genet.* **50**, 192–201 (1951). doi: [10.1007/BF02996215](https://doi.org/10.1007/BF02996215); pmid: [24539699](https://pubmed.ncbi.nlm.nih.gov/24539699/)
 15. M. Sheldon et al., Scrambler and yotari disrupt the disabled gene and produce a reeler-like phenotype in mice. *Nature* **389**, 730–733 (1997). doi: [10.1038/39601](https://doi.org/10.1038/39601); pmid: [9338784](https://pubmed.ncbi.nlm.nih.gov/9338784/)
 16. S. Lutter, S. Xie, F. Tatin, T. Mäkinen, Smooth muscle-endothelial cell communication activates Reelin signaling and regulates lymphatic vessel formation. *J. Cell Biol.* **197**, 837–849 (2012). doi: [10.1083/jcb.201110132](https://doi.org/10.1083/jcb.201110132); pmid: [22665518](https://pubmed.ncbi.nlm.nih.gov/22665518/)
 17. Y. Ding et al., Loss of Reelin protects against atherosclerosis by reducing leukocyte-endothelial cell adhesion and lesion macrophage accumulation. *Sci. Signal.* **9**, ra29 (2016). doi: [10.1126/scisignal.aad5578](https://doi.org/10.1126/scisignal.aad5578); pmid: [26980442](https://pubmed.ncbi.nlm.nih.gov/26980442/)
 18. J. R. Heckenlively et al., Mouse model of subretinal neovascularization with choroidal anastomosis. *Retina* **23**, 518–522 (2003). doi: [10.1097/00006982-200308000-00012](https://doi.org/10.1097/00006982-200308000-00012); pmid: [12972764](https://pubmed.ncbi.nlm.nih.gov/12972764/)
 19. S. E. Hong et al., Autosomal recessive lissencephaly with cerebellar hypoplasia is associated with human RELN mutations. *Nat. Genet.* **26**, 93–96 (2000). doi: [10.1038/79246](https://doi.org/10.1038/79246); pmid: [10973257](https://pubmed.ncbi.nlm.nih.gov/10973257/)
 20. F. Impagnatiello et al., A decrease of reelin expression as a putative vulnerability factor in schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15718–15723 (1998). doi: [10.1073/pnas.95.26.15718](https://doi.org/10.1073/pnas.95.26.15718); pmid: [9861036](https://pubmed.ncbi.nlm.nih.gov/9861036/)
 21. A. Botella-López et al., Reelin expression and glycosylation patterns are altered in Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5573–5578 (2006). doi: [10.1073/pnas.0601279103](https://doi.org/10.1073/pnas.0601279103); pmid: [16567613](https://pubmed.ncbi.nlm.nih.gov/16567613/)
 22. C. A. Haas et al., Role for reelin in the development of granule cell dispersion in temporal lobe epilepsy. *J. Neurosci.* **22**, 5797–5802 (2002). doi: [10.1523/JNEUROSCI.22-14-05797.2002](https://doi.org/10.1523/JNEUROSCI.22-14-05797.2002); pmid: [12122039](https://pubmed.ncbi.nlm.nih.gov/12122039/)
 23. J. M. James, Y. S. Mukoyama, Neuronal action on the developing blood vessel pattern. *Semin. Cell Dev. Biol.* **22**, 1019–1027 (2011). doi: [10.1016/j.semcdb.2011.09.010](https://doi.org/10.1016/j.semcdb.2011.09.010); pmid: [21978864](https://pubmed.ncbi.nlm.nih.gov/21978864/)
 24. S. Sawamiphak, M. Ritter, A. Acker-Palmer, Preparation of retinal explant cultures to study ex vivo tip endothelial cell responses. *Nat. Protoc.* **5**, 1659–1665 (2010). doi: [10.1038/nprot.2010.130](https://doi.org/10.1038/nprot.2010.130); pmid: [20885378](https://pubmed.ncbi.nlm.nih.gov/20885378/)
 25. A. Stahl et al., The mouse retina as an angiogenesis model. *Invest. Ophthalmol. Vis. Sci.* **51**, 2813–2826 (2010). doi: [10.1167/iovs.10-5176](https://doi.org/10.1167/iovs.10-5176); pmid: [20484600](https://pubmed.ncbi.nlm.nih.gov/20484600/)
 26. H. Gerhardt et al., VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia. *J. Cell Biol.* **161**, 1163–1177 (2003). doi: [10.1083/jcb.200302047](https://doi.org/10.1083/jcb.200302047); pmid: [12810700](https://pubmed.ncbi.nlm.nih.gov/12810700/)
 27. S. Sawamiphak et al., Ephrin-B2 regulates VEGFR2 function in developmental and tumour angiogenesis. *Nature* **465**, 487–491 (2010). doi: [10.1038/nature08995](https://doi.org/10.1038/nature08995); pmid: [20445540](https://pubmed.ncbi.nlm.nih.gov/20445540/)
 28. K. R. Howell, M. N. Hoda, A. Pillai, VEGF activates NR2B phosphorylation through Dab1 pathway. *Neurosci. Lett.* **552**, 30–34 (2013). doi: [10.1016/j.neulet.2013.07.032](https://doi.org/10.1016/j.neulet.2013.07.032); pmid: [23916658](https://pubmed.ncbi.nlm.nih.gov/23916658/)
 29. V. Joukov et al., A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases. *EMBO J.* **15**, 1751 (1996). pmid: [8612600](https://pubmed.ncbi.nlm.nih.gov/8612600/)
 30. M. Ogawa et al., The reeler gene-associated antigen on Cajal-Retzius neurons is a crucial molecule for laminar organization of cortical neurons. *Neuron* **14**, 899–912 (1995). doi: [10.1016/0896-6273\(95\)90329-1](https://doi.org/10.1016/0896-6273(95)90329-1); pmid: [7748558](https://pubmed.ncbi.nlm.nih.gov/7748558/)
 31. Y. Hirota, K. Nakajima, Control of Neuronal Migration and Aggregation by Reelin Signaling in the Developing Cerebral Cortex. *Front. Cell Dev. Biol.* **5**, 40 (2017). doi: [10.3389/fcell.2017.00040](https://doi.org/10.3389/fcell.2017.00040); pmid: [28507985](https://pubmed.ncbi.nlm.nih.gov/28507985/)
 32. A. Vasudevan, J. E. Long, J. E. Crandall, J. L. Rubenstein, P. G. Bhidé, Compartment-specific transcription factors orchestrate angiogenesis gradients in the embryonic brain. *Nat. Neurosci.* **11**, 429–439 (2008). doi: [10.1038/nn2074](https://doi.org/10.1038/nn2074); pmid: [18344991](https://pubmed.ncbi.nlm.nih.gov/18344991/)
 33. T. Walchli et al., Wiring the Vascular Network with Neural Cues: A CNS Perspective. *Neuron* **87**, 271–296 (2015). doi: [10.1016/j.neuron.2015.06.038](https://doi.org/10.1016/j.neuron.2015.06.038); pmid: [26182414](https://pubmed.ncbi.nlm.nih.gov/26182414/)
 34. V. S. Caviness Jr., Neocortical histogenesis in normal and reeler mice: A developmental study based upon [³H]thymidine autoradiography. *Brain Res.* **256**, 293–302 (1982). doi: [10.1016/0165-3806\(82\)90141-9](https://doi.org/10.1016/0165-3806(82)90141-9); pmid: [7104762](https://pubmed.ncbi.nlm.nih.gov/7104762/)
 35. C. Lambert de Rouvroit, A. M. Goffinet, *The Reeler Mouse as a Model of Brain Development* (Springer, 1998). doi: [10.1007/978-3-642-72257-8](https://doi.org/10.1007/978-3-642-72257-8)
 36. S. J. Franco, I. Martínez-Garay, C. Gil-Sanz, S. R. Harkins-Perry, U. Müller, Reelin regulates cadherin function via Dab1/Rap1 to control neuronal migration and lamination in the neocortex. *Neuron* **69**, 482–497 (2011). doi: [10.1016/j.neuron.2011.01.003](https://doi.org/10.1016/j.neuron.2011.01.003); pmid: [21315259](https://pubmed.ncbi.nlm.nih.gov/21315259/)
 37. T. Takahashi, J. P. Misson, V. S. Caviness Jr., Glial process elongation and branching in the developing murine neocortex: A qualitative and quantitative immunohistochemical analysis. *J. Comp. Neurol.* **302**, 15–28 (1990). doi: [10.1002/cne.903020103](https://doi.org/10.1002/cne.903020103); pmid: [2086612](https://pubmed.ncbi.nlm.nih.gov/2086612/)
 38. H. E. Beggs et al., FAK deficiency in cells contributing to the basal lamina results in cortical abnormalities resembling congenital muscular dystrophies. *Neuron* **40**, 501–514 (2003). doi: [10.1016/S0896-6273\(03\)00666-4](https://doi.org/10.1016/S0896-6273(03)00666-4); pmid: [14642275](https://pubmed.ncbi.nlm.nih.gov/14642275/)
 39. D. Graus-Porta et al., β 1-class integrins regulate the development of laminae and folia in the cerebral and cerebellar cortex. *Neuron* **31**, 367–379 (2001). doi: [10.1016/S0896-6273\(01\)00374-9](https://doi.org/10.1016/S0896-6273(01)00374-9); pmid: [11516395](https://pubmed.ncbi.nlm.nih.gov/11516395/)
 40. W. Halfter, S. Dong, Y. P. Yip, M. Willem, U. Mayer, A critical function of the pial basement membrane in cortical histogenesis. *J. Neurosci.* **22**, 6029–6040 (2002). doi: [10.1523/JNEUROSCI.22-14-06029.2002](https://doi.org/10.1523/JNEUROSCI.22-14-06029.2002); pmid: [12122064](https://pubmed.ncbi.nlm.nih.gov/12122064/)
 41. D. E. Schmechel, P. Rakic, A Golgi study of radial glial cells in developing monkey telencephalon: Morphogenesis and transformation into astrocytes. *Anat. Embryol.* **156**, 115–152 (1979). doi: [10.1007/BF00300010](https://doi.org/10.1007/BF00300010); pmid: [111580](https://pubmed.ncbi.nlm.nih.gov/111580/)
 42. S. J. Tam, R. J. Watts, Connecting vascular and nervous system development: Angiogenesis and the blood-brain barrier. *Annu. Rev. Neurosci.* **33**, 379–408 (2010). doi: [10.1146/annurev-neuro-060909-152829](https://doi.org/10.1146/annurev-neuro-060909-152829); pmid: [20367445](https://pubmed.ncbi.nlm.nih.gov/20367445/)
 43. B. J. Andreone, B. Lacoste, C. Gu, Neuronal and vascular interactions. *Annu. Rev. Neurosci.* **38**, 25–46 (2015). doi: [10.1146/annurev-neuro-071714-033835](https://doi.org/10.1146/annurev-neuro-071714-033835); pmid: [25782970](https://pubmed.ncbi.nlm.nih.gov/25782970/)
 44. A. Armulik et al., Pericytes regulate the blood-brain barrier. *Nature* **468**, 557–561 (2010). doi: [10.1038/nature09522](https://doi.org/10.1038/nature09522); pmid: [20944627](https://pubmed.ncbi.nlm.nih.gov/20944627/)
 45. R. Daneman, L. Zhou, A. A. Kebede, B. A. Barres, Pericytes are required for blood-brain barrier integrity during embryogenesis. *Nature* **468**, 562–566 (2010). doi: [10.1038/nature09513](https://doi.org/10.1038/nature09513); pmid: [20944625](https://pubmed.ncbi.nlm.nih.gov/20944625/)
 46. K. M. Baeten, K. Akassoglou, Extracellular matrix and matrix receptors in blood-brain barrier formation and stroke. *Dev. Neurobiol.* **71**, 1018–1039 (2011). doi: [10.1002/dneu.20954](https://doi.org/10.1002/dneu.20954); pmid: [21780303](https://pubmed.ncbi.nlm.nih.gov/21780303/)
 47. R. Hallmann et al., Expression and function of laminins in the embryonic and mature vasculature. *Physiol. Rev.* **85**, 979–1000 (2005). doi: [10.1152/physrev.00014.2004](https://doi.org/10.1152/physrev.00014.2004); pmid: [15987800](https://pubmed.ncbi.nlm.nih.gov/15987800/)
 48. M. Sixt et al., Endothelial cell laminin isoforms, laminins 8 and 10, play decisive roles in T cell recruitment across the blood-brain barrier in experimental autoimmune encephalomyelitis. *J. Cell Biol.* **153**, 933–946 (2001). doi: [10.1083/jcb.153.5.933](https://doi.org/10.1083/jcb.153.5.933); pmid: [11381080](https://pubmed.ncbi.nlm.nih.gov/11381080/)
 49. H. Fujiwara, Y. Kikkawa, N. Sanzen, K. Sekiguchi, Purification and characterization of human laminin-8. Laminin-8 stimulates cell adhesion and migration through α 3 β 1 and α 6 β 1 integrins. *J. Biol. Chem.* **276**, 17550–17558 (2001). doi: [10.1074/jbc.M010155200](https://doi.org/10.1074/jbc.M010155200); pmid: [11278628](https://pubmed.ncbi.nlm.nih.gov/11278628/)
 50. G. J. del Zoppo, R. Milner, Integrin-matrix interactions in the cerebral microvasculature. *Arterioscler. Thromb. Vasc. Biol.* **26**, 1966–1975 (2006). doi: [10.1161/01.ATV.0000232525.65682.a2](https://doi.org/10.1161/01.ATV.0000232525.65682.a2); pmid: [16778120](https://pubmed.ncbi.nlm.nih.gov/16778120/)
 51. R. Milner, I. L. Campbell, The integrin family of cell adhesion molecules has multiple functions within the CNS. *J. Neurosci. Res.* **69**, 286–291 (2002). doi: [10.1002/jnr.10321](https://doi.org/10.1002/jnr.10321); pmid: [1225070](https://pubmed.ncbi.nlm.nih.gov/1225070/)
 52. H. Yamamoto et al., Integrin β 1 controls VE-cadherin localization and blood vessel stability. *Nat. Commun.* **6**, 6429 (2015). doi: [10.1038/ncomms7429](https://doi.org/10.1038/ncomms7429); pmid: [25752958](https://pubmed.ncbi.nlm.nih.gov/25752958/)
 53. M. Lenter et al., A monoclonal antibody against an activation epitope on mouse integrin chain β 1 blocks adhesion of lymphocytes to the endothelial integrin α 6 β 1. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9051–9055 (1993). doi: [10.1073/pnas.90.19.9051](https://doi.org/10.1073/pnas.90.19.9051); pmid: [7692444](https://pubmed.ncbi.nlm.nih.gov/7692444/)
 54. M. Nakayama et al., Spatial regulation of VEGF receptor endocytosis in angiogenesis. *Nat. Cell Biol.* **15**, 249–260 (2013). doi: [10.1038/ncb2679](https://doi.org/10.1038/ncb2679); pmid: [23354168](https://pubmed.ncbi.nlm.nih.gov/23354168/)
 55. M. Vanlandewijck et al., A molecular atlas of cell types and zonation in the brain vasculature. *Nature* **554**, 475–480 (2018). doi: [10.1038/nature25739](https://doi.org/10.1038/nature25739); pmid: [29443965](https://pubmed.ncbi.nlm.nih.gov/29443965/)
 56. K. Sekine et al., Reelin controls neuronal positioning by promoting cell-matrix adhesion via inside-out activation of integrin α 5 β 1. *Neuron* **76**, 353–369 (2012). doi: [10.1016/j.neuron.2012.07.020](https://doi.org/10.1016/j.neuron.2012.07.020); pmid: [23083738](https://pubmed.ncbi.nlm.nih.gov/23083738/)
 57. N. Haubst, E. Georges-Labouesse, A. De Arcangelis, U. Mayer, M. Götz, Basement membrane attachment is dispensable for radial glial cell fate and for proliferation, but affects positioning of neuronal subtypes. *Development* **133**, 3245–3254 (2006). doi: [10.1242/dev.02486](https://doi.org/10.1242/dev.02486); pmid: [16873583](https://pubmed.ncbi.nlm.nih.gov/16873583/)
 58. R. Belvindrah, D. Graus-Porta, S. Goebbels, K. A. Nave, U. Müller, β 1 integrins in radial glia but not in migrating neurons are essential for the formation of cell layers in the cerebral cortex. *J. Neurosci.* **27**, 13854–13865 (2007). doi: [10.1523/JNEUROSCI.4494-07.2007](https://doi.org/10.1523/JNEUROSCI.4494-07.2007); pmid: [18077697](https://pubmed.ncbi.nlm.nih.gov/18077697/)
 59. R. Radakovits, C. S. Barros, R. Belvindrah, B. Patton, U. Müller, Regulation of radial glial survival by signals from the meninges. *J. Neurosci.* **29**, 7694–7705 (2009). doi: [10.1523/JNEUROSCI.5537-08.2009](https://doi.org/10.1523/JNEUROSCI.5537-08.2009); pmid: [19535581](https://pubmed.ncbi.nlm.nih.gov/19535581/)
 60. M. Gonzales et al., Structure and function of a vimentin-associated matrix adhesion in endothelial cells. *Mol. Biol. Cell* **12**, 85–100 (2001). doi: [10.1091/mbc.12.1.85](https://doi.org/10.1091/mbc.12.1.85); pmid: [11160825](https://pubmed.ncbi.nlm.nih.gov/11160825/)
 61. M. Frotscher, S. Zhao, S. Wang, X. Chai, Reelin Signaling Inactivates Cofilin to Stabilize the Cytoskeleton of Migrating Cortical Neurons. *Front. Cell. Neurosci.* **11**, 148 (2017). doi: [10.3389/fncel.2017.00148](https://doi.org/10.3389/fncel.2017.00148); pmid: [28588454](https://pubmed.ncbi.nlm.nih.gov/28588454/)
 62. N. J. Abbott, L. Rönnebeck, E. Hansson, Astrocyte-endothelial interactions at the blood-brain barrier. *Nat. Rev. Neurosci.* **7**, 41–53 (2006). doi: [10.1038/nrn1824](https://doi.org/10.1038/nrn1824); pmid: [16371949](https://pubmed.ncbi.nlm.nih.gov/16371949/)
 63. J. Thyböll et al., Deletion of the laminin α 4 chain leads to impaired microvessel maturation. *Mol. Cell. Biol.* **22**, 1194–1202 (2002). doi: [10.1128/MCB.22.4.1194-1202.2002](https://doi.org/10.1128/MCB.22.4.1194-1202.2002); pmid: [11809810](https://pubmed.ncbi.nlm.nih.gov/11809810/)
 64. D. Stenzel et al., Endothelial basement membrane limits tip cell formation by inducing Dll4/Notch signalling in vivo. *EMBO Rep.* **12**, 1135–1143 (2011). doi: [10.1038/embor.2011.194](https://doi.org/10.1038/embor.2011.194); pmid: [21979816](https://pubmed.ncbi.nlm.nih.gov/21979816/)
 65. I. Sørensen, R. H. Adams, A. Gossler, DLL1-mediated Notch activation regulates endothelial identity in mouse fetal arteries. *Blood* **113**, 5680–5688 (2009). doi: [10.1182/blood-2008-08-174508](https://doi.org/10.1182/blood-2008-08-174508); pmid: [19144989](https://pubmed.ncbi.nlm.nih.gov/19144989/)
 66. U. Mühler et al., Formation of transformed endothelial cells in the absence of VEGFR-2/Flk-1 by Polyoma middle T oncogene. *Oncogene* **18**, 4200–4210 (1999). doi: [10.1038/sj.onc.1203014](https://doi.org/10.1038/sj.onc.1203014); pmid: [10435633](https://pubmed.ncbi.nlm.nih.gov/10435633/)
 67. M. E. Pitulescu, I. Schmidt, R. Bénédict, R. H. Adams, Inducible gene targeting in the neonatal vasculature and analysis of retinal angiogenesis in mice. *Nat. Protoc.* **5**, 1518–1534 (2010). doi: [10.1038/nprot.2010.113](https://doi.org/10.1038/nprot.2010.113); pmid: [20725067](https://pubmed.ncbi.nlm.nih.gov/20725067/)
 68. M. Trommsdorff et al., Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97**, 689–701 (1999). doi: [10.1016/S0092-8674\(00\)80782-5](https://doi.org/10.1016/S0092-8674(00)80782-5); pmid: [10380922](https://pubmed.ncbi.nlm.nih.gov/10380922/)
 69. T. Yardeni, M. Eckhaus, H. D. Morris, M. Huizing, S. Hoogstraten-Miller, Retro-orbital injections in mice. *Lab Anim.* **40**, 155–160 (2011). doi: [10.1038/labana051155](https://doi.org/10.1038/labana051155); pmid: [21508954](https://pubmed.ncbi.nlm.nih.gov/21508954/)
 70. S. Liebnér, U. Kniessel, H. Kalbacher, H. Wolburg, Correlation of tight junction morphology with the expression of tight junction proteins in blood-brain barrier endothelial cells. *Eur. J. Cell Biol.* **79**, 707–717 (2000). doi: [10.1078/0171-9335-00101](https://doi.org/10.1078/0171-9335-00101); pmid: [11089919](https://pubmed.ncbi.nlm.nih.gov/11089919/)
 71. K. Wilhelm et al., FOXO1 couples metabolic activity and growth state in the vascular endothelium. *Nature* **529**, 216–220 (2016). doi: [10.1038/nature16498](https://doi.org/10.1038/nature16498); pmid: [26735015](https://pubmed.ncbi.nlm.nih.gov/26735015/)
 72. E. Kur et al., Temporal modulation of collective cell behavior controls vascular network topology. *eLife* **5**, e13212 (2016). doi: [10.7554/eLife.13212](https://doi.org/10.7554/eLife.13212); pmid: [26910011](https://pubmed.ncbi.nlm.nih.gov/26910011/)

73. M. Hellström *et al.*, Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* **445**, 776–780 (2007). doi: [10.1038/nature05571](https://doi.org/10.1038/nature05571); pmid: [17259973](https://pubmed.ncbi.nlm.nih.gov/17259973/)
74. E. Zudaire, L. Gambardella, C. Kurcz, S. Vermeren, A computational tool for quantitative analysis of vascular networks. *PLOS ONE* **6**, e27385 (2011). doi: [10.1371/journal.pone.0027385](https://doi.org/10.1371/journal.pone.0027385); pmid: [22110636](https://pubmed.ncbi.nlm.nih.gov/22110636/)
75. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012). doi: [10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089); pmid: [22930834](https://pubmed.ncbi.nlm.nih.gov/22930834/)
76. M. Parrilla, I. Chang, A. Degl'Innocenti, M. Omura, Expression of homeobox genes in the mouse olfactory epithelium. *J. Comp. Neurol.* **524**, 2713–2739 (2016). doi: [10.1002/cne.24051](https://doi.org/10.1002/cne.24051); pmid: [27243442](https://pubmed.ncbi.nlm.nih.gov/27243442/)
- U. Bauer, K. Hammer, S. Seidel, D. Schmelzer, and T. Belefkih for technical support; and A. Frangakis and H. Zimmermann for helpful discussions. **Funding:** Supported by ERC_AdG_Neurovessel (project 669742); Deutsche Forschungsgemeinschaft grants SFB 834, SFB1080, FOR2325, EXC 115, and EXC 147; the Max Planck Fellow Program and Gutenberg Research College (GRC) at Johannes Gutenberg University Mainz (A.A.-P.); and the Marie Curie-CIG 293902 (M.S.). **Author contributions:** M.S. performed initial experiments and designed and supervised the project; M.R.A. designed, performed, and supervised experiments in the vascular-glia and BBB parts; F.C. designed, performed, and supervised experiments in the vascular part; C.L.-C. performed experiments in the vascular-glia and BBB parts; R.H. and I.B. performed experiments in the vascular part; M.D. performed PLA assays and VEGFR2 cortex analysis; M.P. performed in situ hybridizations; D.H. performed astrocyte and BBB permeability experiments and designed the model cartoons; A.S., H.S., and T.A. designed, performed, and supervised the electron microscopy; L.M. and M.R. performed initial experiments in the vascular part; L.T.-M. performed astrocyte binding assays; A.A.-P. analyzed data and designed and supervised all stages of the project; M.S. and A.A.-P. wrote the manuscript with contributions from M.R.A. and F.C.; and all authors discussed and interpreted the data and gave input to the written manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this paper are present either in the main text or the supplementary materials. The Dab1 floxed mice are available from U. Müller under a material agreement with The Scripps Research Institute.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/eaao2861/suppl/DC1
Figs. S1 to S13

4 July 2017; accepted 5 July 2018
10.1126/science.aao2861

ACKNOWLEDGMENTS

We thank U. Müller, J. Herz, E. Wagner, and R. Adams for providing mouse strains; S. Sawamiphak for initial help in the project;

RESEARCH ARTICLE SUMMARY

CYTOSKELETON

Severing enzymes amplify microtubule arrays through lattice GTP-tubulin incorporation

Annapurna Vemu, Ewa Szczesna, Elena A. Zehr, Jeffrey O. Spector, Nikolaus Grigorieff, Alexandra M. Deaconescu, Antonina Roll-Mecak*

INTRODUCTION: The microtubule cytoskeleton is continuously sculpted by polymerization, depolymerization, cross-linking, and microtubule severing. Microtubule severing generates internal breaks in microtubules through a poorly understood mechanism. It is mediated by the AAA [adenosine triphosphatases (ATPases) associated with various cellular activities] ATPases katanin, spastin, and fidgetin. Paradoxically, despite the destructive activity of severing enzymes, loss of severing enzyme activity leads to a decrease rather than an increase in microtubule mass. It was hypothesized that this severing enzyme-dependent increase in microtubule mass results from templated nucleation from the severed ends. This is an attractive hypothesis for a mechanism to generate microtubule mass, especially in the absence of centrosome-based nucleation as in neurons or meiotic spindles. However, for this amplification to operate, the guanosine diphosphate (GDP)-tubulin lattice exposed through

severing would have to be stabilized. The GDP-microtubule lattice is the product of guanosine triphosphate (GTP)-tubulin polymerization and depolymerizes spontaneously when exposed in the absence of a stabilizing GTP cap. We examined this paradox by examining the effects of the severing enzymes spastin and katanin on microtubule structure and dynamics in vitro.

RATIONALE: Because light microscopy-based severing assays fail to capture ultrastructural features of severing intermediates due to resolution limitations, we used negative-stain transmission electron microscopy (TEM) to capture and image spastin- and katanin-mediated microtubule severing in vitro. We combined these experiments with quantitative analyses of tubulin and microtubule polymer dynamics by using total internal reflection fluorescence (TIRF) microscopy to understand the effects of severing on microtubule networks.

RESULTS: Our electron microscopy analyses coupled with TIRF microscopy revealed that spastin and katanin actively extract tubulin dimers out of the microtubule, introducing nanoscale damage along the microtubule, and that this action is counteracted by spontaneous, de novo incorporation of GTP-tubulin dimers from the soluble pool. Depending on the local balance between the rates of active tubulin extraction and passive repair, there are two

ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aau1504>

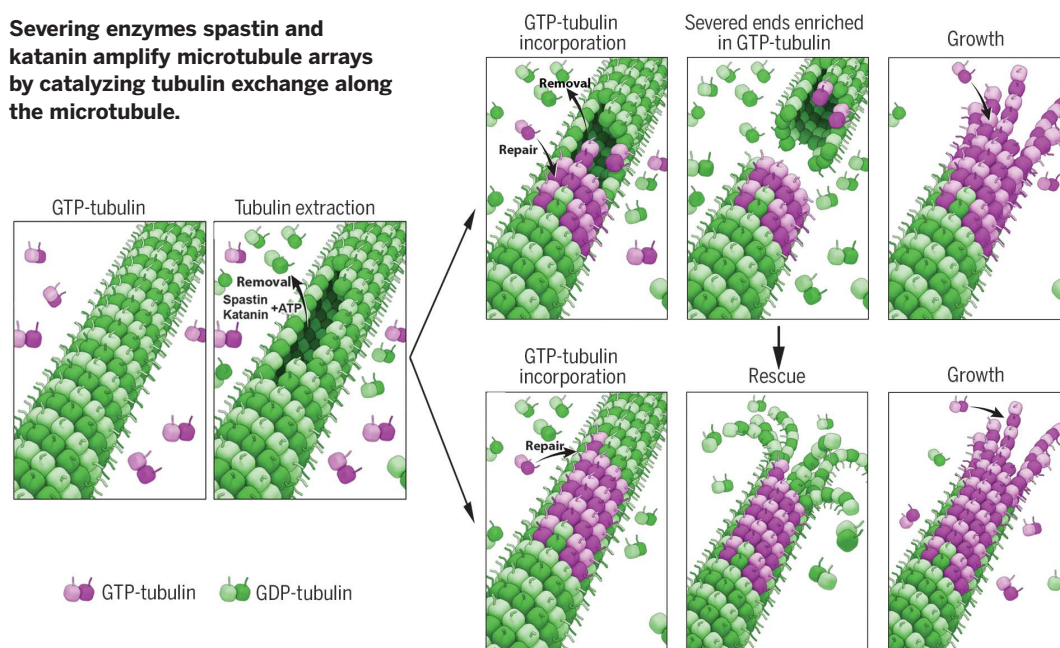
non-mutually exclusive consequences: The microtubule is rejuvenated with GTP-tubulin islands that stabilize it against depolymerization, or severing proceeds to completion

and the newly severed microtubule ends emerge with a high density of stabilizing GTP-tubulin. Consistent with this, we found that spastin and katanin activities increase rates of microtubule rescue and that rescues occur preferentially at sites of enzyme-dependent GTP-tubulin incorporation. Lastly and unexpectedly, we found that the incorporation of GTP-tubulin at severing sites ensures that the newly severed plus ends are stable because they emerge with a high density of GTP-tubulin that protects them against spontaneous depolymerization and promotes elongation. The synergy between the increased rescue rates and the stabilization of the newly severed ends leads to microtubule amplification.

CONCLUSION: Our study identifies the microtubule-severing enzymes spastin and katanin as biological agents that introduce GTP-tubulin islands within microtubules and demonstrates that microtubule-severing enzymes alone can amplify microtubule number and mass

by promoting GTP-tubulin incorporation into the microtubule shaft, away from the dynamic ends long thought to be the sole locus of tubulin exchange. This microtubule-based amplification mechanism in the absence of a nucleating factor helps explain why the loss of spastin and katanin results in the loss of microtubule mass in systems that are dependent on noncentrosomal microtubule generation. ■

Severing enzymes spastin and katanin amplify microtubule arrays by catalyzing tubulin exchange along the microtubule.



The list of author affiliations is available in the full article online.

*Corresponding author. Email:

antonina@mail.nih.gov

Cite this article as A. Vemu et al.,

Science 361, eaau1504 (2018).

DOI: 10.1126/science.aau1504

RESEARCH ARTICLE

CYTOSKELETON

Severing enzymes amplify microtubule arrays through lattice GTP-tubulin incorporation

Annapurna Vemu^{1*}, Ewa Szczesna^{1*}, Elena A. Zehr¹, Jeffrey O. Spector¹, Nikolaus Grigorieff^{2,†}, Alexandra M. Deaconescu³, Antonina Roll-Mecak^{1,4,‡}

Spastin and katanin sever and destabilize microtubules. Paradoxically, despite their destructive activity they increase microtubule mass in vivo. We combined single-molecule total internal reflection fluorescence microscopy and electron microscopy to show that the elemental step in microtubule severing is the generation of nanoscale damage throughout the microtubule by active extraction of tubulin heterodimers. These damage sites are repaired spontaneously by guanosine triphosphate (GTP)-tubulin incorporation, which rejuvenates and stabilizes the microtubule shaft. Consequently, spastin and katanin increase microtubule rescue rates. Furthermore, newly severed ends emerge with a high density of GTP-tubulin that protects them against depolymerization. The stabilization of the newly severed plus ends and the higher rescue frequency synergize to amplify microtubule number and mass. Thus, severing enzymes regulate microtubule architecture and dynamics by promoting GTP-tubulin incorporation within the microtubule shaft.

The plasticity of the microtubule cytoskeleton follows from multiple levels of regulation through microtubule-end polymerization and depolymerization, cross-linking, and microtubule severing. Microtubule severing generates internal breaks in microtubules. It is mediated by three enzymes of the AAA [adenosine triphosphatases (ATPases) associated with various cellular activities] ATPase family—katanin, spastin, and fidgetin [reviewed in (1)]—that are widely conserved in animals and plants. They are critical for the generation and maintenance of complex noncentrosomal microtubule arrays in neurons (2–5) and the plant cortex (6–8) and regulate meiotic and mitotic spindle morphology and length (9–12), cilia biogenesis (13, 14), centriole duplication (14, 15), cytokinesis (16, 17), axonal growth (18), wound healing (19), and plant phototropism (7, 8). Both spastin and katanin are associated with debilitating diseases. Spastin is mutated in hereditary spastic paraplegias, neurodegenerative disorders characterized by lower-extremity weakness due to axonopathy [reviewed in (1)]. Katanin mutations

cause microcephaly, seizures, and severe developmental defects (14, 15, 20). Disease mutations impair microtubule severing (21, 22).

Paradoxically, in many of these systems, the loss of the microtubule-severing enzyme leads to a decrease in microtubule mass [reviewed in (1)]. Spastin loss causes sparse disorganized microtubule arrays at *Drosophila* synaptic boutons (2) and impaired axonal outgrowth and sparse microtubule arrays in zebra fish axons (23). Similarly, katanin loss leads to sparse cortical microtubule arrays in *Arabidopsis* (8, 24), whereas in *Caenorhabditis elegans* meiotic spindles, it results in the loss of microtubule mass and number (25). It was hypothesized that the observed increase in microtubule number and mass results from templated nucleation from the severed ends (26, 27). This is an attractive mechanism for rapidly generating microtubule mass, especially in the absence of centrosome-based nucleation as in neurons or meiotic spindles. This severing-dependent microtubule amplification has been directly observed in plant cortical microtubule arrays (8). However, for this amplification to operate, the guanosine diphosphate (GDP)-tubulin lattice exposed through severing would have to be stabilized because GDP-microtubules depolymerize spontaneously in the absence of a stabilizing guanosine triphosphate (GTP) cap (28–31). To study this paradox, we combined time-resolved transmission electron microscopy (TEM) and total internal reflection fluorescence (TIRF) microscopy to directly observe the effects of the severing enzymes spastin and katanin on microtubule structure and dynamics in vitro.

Severing enzymes cause nanoscale damage to microtubules

Because light microscopy-based severing assays fail to capture ultrastructural features of severing intermediates due to resolution limitations, we used negative-stain TEM to capture and image spastin-mediated microtubule-severing intermediates in vitro with purified, recombinant spastin. To minimize severing-intermediate breakage, we performed severing reactions directly on electron microscopy (EM) grids. These on-grid reactions revealed a high density of “bites” into the protofilament structure (Fig. 1) that resulted in the removal of tubulin dimers. Severing reactions performed in a test tube with Taxol-stabilized microtubules that were then transferred to EM grids by pipetting produced many short microtubules with blunt ends (fig. S1A), similar to those previously reported in vitro with katanin (32), indicating that the fragile nanoscale-damaged severing intermediates are lost during pipetting. Thus, in our on-grid severing setup, we were able to capture intermediates that were otherwise disrupted by shear forces introduced by pipetting. Upon prolonged incubation (>5 min), severing was driven to completion on the EM grid, with severe destruction of the microtubule structure indicating that the intermediates observed were on pathway (fig. S1B). The nanoscale damage sites were observed with GDP-microtubules regardless of whether they were nonstabilized or stabilized with Taxol (Fig. 1, A and B). Similar observations were made with microtubules polymerized with the nonhydrolyzable analog guanylyl (α,β -methylene diphosphonate (GMPCPP) (Fig. 1C). The nanoscale damage we observed in vitro is reminiscent of that observed by electron tomography in freeze-substituted *C. elegans* meiotic spindles (25). The same extraction of tubulin dimers and protofilament fraying were observed if reactions were performed in solution and then microtubules were deposited on an EM grid without pipetting to avoid shear (Fig. 1D and Materials and methods). In control reactions without the enzyme, the integrity of the lattice was preserved (fig. S1, C and D), whereas in the spastin-treated samples, nanoscale damage sites were detected every ~2.2 μm (fig. S1D). Time-course experiments revealed a gradual increase in nanoscale damage, as well as in the number of shorter microtubules (Fig. 1E). We extended our TEM analyses to the microtubule-severing enzyme katanin (Fig. 1, F and G, and fig. S1, E to H). As with spastin, TEM revealed that katanin microtubule severing proceeds through progressive extraction of tubulin dimers out of the microtubule.

Tubulin incorporation repairs nanoscale damage

Our TEM analysis showed that GMPCPP-microtubules, Taxol-stabilized microtubules, or nonstabilized microtubules do not sever even when peppered with spastin- and katanin-induced nanoscale damage and do not catastrophically depolymerize upon removal of the initial tubulin subunits. This raised the possibility that this damage could be repaired by incorporation of

¹Cell Biology and Biophysics Unit, Porter Neuroscience Research Center, National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA. ²Howard Hughes Medical Institute, Brandeis University, Waltham, MA 02454, USA. ³Department of Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, RI 02903, USA. ⁴Biochemistry and Biophysics Center, National Heart, Lung, and Blood Institute, Bethesda, MD 20892, USA.

*These authors contributed equally to this work.

[†]Present address: Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA.

[‡]Corresponding author. Email: antonina@mail.nih.gov

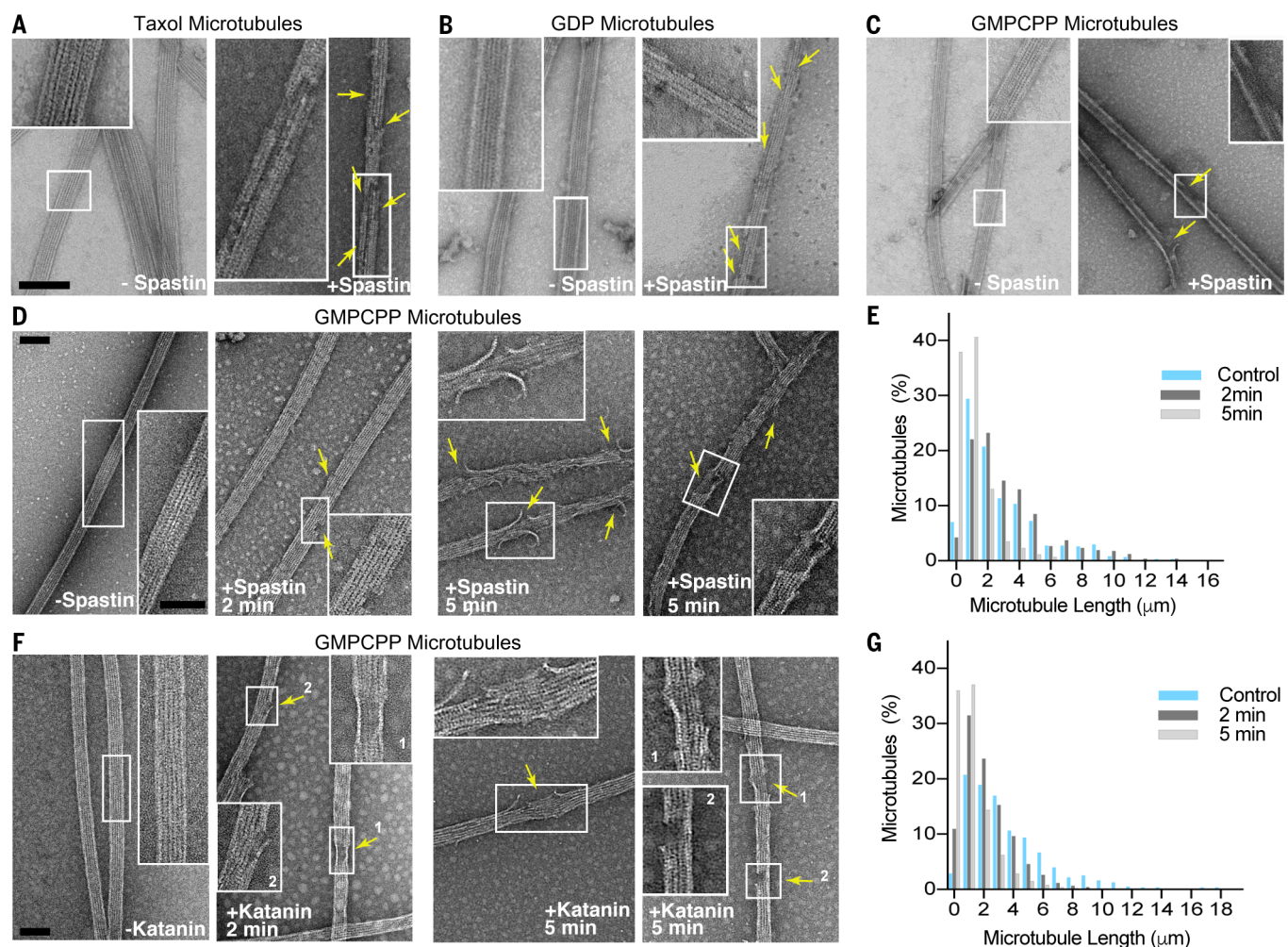


Fig. 1. Spastin and katanin extract tubulin out of the microtubule.

(A to C) Microtubules in the absence or presence of 33 nM spastin. The reaction proceeded on an EM grid for 1 min and was imaged by using negative-stain TEM (Materials and methods). Boxed regions are shown at 2× magnification in insets. Microtubules were imaged at 30,000× magnification. Arrows indicate nanoscale damage sites. Scale bar, 50 nm. (D) Fields of GMPCPP-microtubules incubated with buffer or 25 nM spastin. Severing proceeded in solution, and reaction mixtures

were passively deposited onto EM grids, negatively stained, and visualized by TEM (Materials and methods). Arrows indicate nanoscale damage. Microtubules were imaged at 13,000× magnification; boxed regions are shown at 30,000× magnification in insets. Scale bar, 50 nm. (E) Microtubule length distribution after incubation with spastin. (F) Fields of GMPCPP-microtubules incubated with buffer or 100 nM katanin and imaged as in (D). Scale bar, 50 nm. (G) Microtubule length distribution after incubation with katanin.

tubulin subunits from the soluble pool, as recently observed with mechanically damaged or photo-damaged microtubules in vitro (33, 34). To test this hypothesis, we preassembled GMPCPP-microtubules fluorescently labeled with HiLyte 647 and incubated them with spastin (or katanin) and adenosine triphosphate (ATP) to initiate severing (Materials and methods). Under these conditions, we observed rare severing events (Fig. 2). Upon perfusion with soluble HiLyte 488-labeled tubulin and GTP, we observed tubulin incorporation in discrete patches along microtubules. These patches were numerous, far exceeding the number of severing events. Mock-treated microtubules showed no incorporation of tubulin into microtubules (Fig. 2, A to C). The tubulin concentration used was below the critical concentration for tubulin polymerization. Sim-

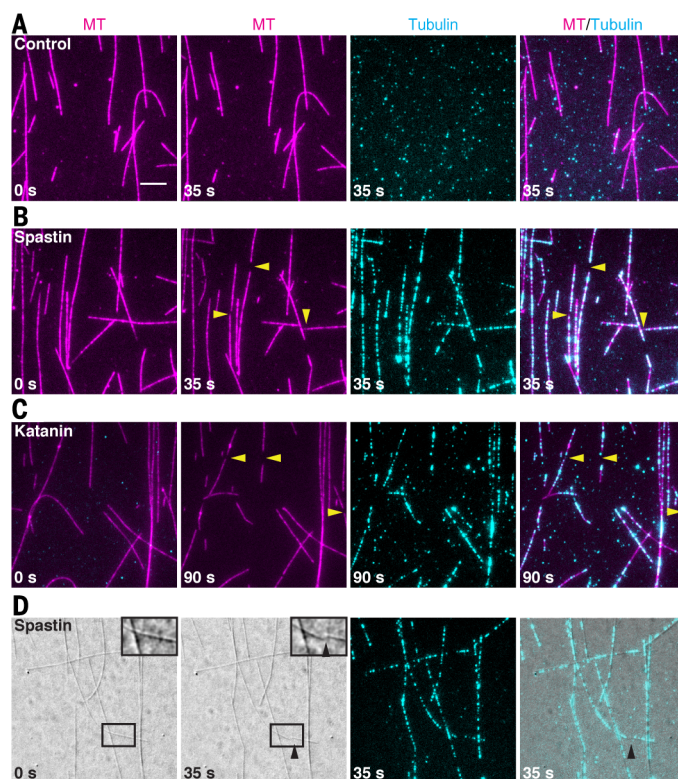
ilar results were obtained with Taxol-stabilized microtubules (fig. S2). Because photodamage can induce lattice defects in fluorescently labeled microtubules (34), we also performed experiments with unlabeled microtubules visualized by differential interference contrast (DIC) microscopy and also observed incorporation of tubulin into spastin-treated microtubules but not into controls (Fig. 2D).

In time-course experiments, both the number of repaired nanoscale damage sites and the mean fluorescence along repaired microtubules increased over time (figs. S3, A and B, and S4, A and B). The size of the repair sites [full width at half maximum (FWHM)] (figs. S3C and S4C) was initially diffraction limited and shifted toward larger values at longer incubation times, indicating an expansion of the damage as de-

ected by soluble GTP-tubulin incorporation. Frequent nanoscale damage events were visible when severing events were extremely sparse: As early as 35 s, the density of spastin-induced nanoscale damage sites was $0.35 \pm 0.01 \mu\text{m}^{-1}$, compared with $0.0008 \pm 0.0004 \mu\text{m}^{-1}$ for severing events (fig. S3, A and D). Thus, most nanoscale damage events did not lead to macroscopic severing events. Once a sufficient number of tubulin dimers was removed from the lattice, the microtubule unraveled and a macroscopic severing event was visible. Consistent with this, we observed an abrupt increase in mesoscale severing at 120 and 90 s for spastin and katanin, respectively (figs. S3D and S4D).

Next, we probed the effect of soluble tubulin on spastin microtubule severing by performing severing assays in the presence of fluorescently

Fig. 2. Spastin- and katanin-catalyzed nanoscale damage is repaired by spontaneous tubulin incorporation. (A and B) HiLyte 647-labeled GMPCPP-microtubules (MT) incubated with buffer (A) or 10 nM spastin (B) for 35 s and then incubated with 1 μ M HiLyte 488-labeled GTP-tubulin and washed to remove excess tubulin (Materials and methods). Arrowheads indicate severing events. Scale bar, 5 μ m. (C) HiLyte 647-labeled GMPCPP-microtubules incubated with 2 nM katanin for 90 s and then incubated with 1 μ M HiLyte 488-labeled GTP-tubulin and washed to remove excess tubulin (Materials and methods). Arrowheads indicate severing events. (D) DIC-imaged unlabeled GMPCPP-microtubules incubated with 10 nM spastin and then with 1 μ M HiLyte 488-labeled GTP-tubulin (cyan) and washed to remove excess tubulin (Materials and methods). Insets correspond to boxed areas and show the progression to a severing event. Arrowheads indicate the severing site.



labeled soluble tubulin (fig. S5). This allowed us to detect microtubule nanoscale damage and severing simultaneously. Spastin-induced severing was not significantly affected with 100 nM tubulin, even though we observed incorporation of HiLyte 488-labeled tubulin into microtubules (Fig. 3A, fig. S5, and Materials and methods). However, severing was considerably reduced in the presence of 2 μ M soluble tubulin (Fig. 3A), and in this case, tubulin fluorescence intensity at repair sites was also significantly higher (Fig. 3B). Thus, the tubulin extraction activity of the enzyme was not significantly inhibited by soluble tubulin as proposed previously for katanin (35), but the rate of tubulin incorporation at nanoscale damage sites increased with tubulin concentration. This higher rate of tubulin incorporation at damage sites delays (and can even prevent) the completion of a severing event. Consistent with this, the time between the incorporation of HiLyte 488-tubulin at a nanoscale damage site and the completion of a severing event was longer in the presence of 2 μ M tubulin than in the presence of 100 nM tubulin (Fig. 3C). Thus, whereas almost all nanoscale damage sites detectable under our experimental conditions proceeded to complete severing within 65 s after tubulin incorporation in the presence of 100 nM soluble tubulin, only 47% did so at 2 μ M tubulin (Fig. 3D). We also monitored live the addition of single fluorescently labeled tubulin dimers by

TIRF microscopy (Fig. 3E and Materials and methods). Fluorescence intensity analyses revealed that repair proceeded mainly through the incorporation of tubulin heterodimers and not through the addition of larger tubulin polymers or aggregates because the fluorescence intensity distribution of incorporated tubulin was similar to that of single tubulin subunits immobilized to glass (Fig. 3F).

Severing enzymes introduce GTP-tubulin islands

To rule out repair as an artifact of working with stabilized microtubules (either Taxol or GMPCPP stabilized), we extended our experiments to non-stabilized GDP-microtubules. We polymerized GDP-microtubules from axonemes and stabilized their ends with a GMPCPP cap to avoid spontaneous depolymerization (Materials and methods). We then introduced spastin in the absence or presence of fluorescently labeled soluble GTP-tubulin. Within 50 s of introducing 5 nM spastin and 5 μ M soluble tubulin [tubulin concentrations in vivo are 5 to 20 μ M (36, 37)], we observed the incorporation of tubulin as puncta along microtubules (Fig. 3G and movie S1). At these enzyme and tubulin concentrations, most tubulin incorporation sites did not progress to a severing event, and the severing rate was considerably lower than in the absence of soluble tubulin (Fig. 3H). However, tubulin incorpora-

tion always preceded microtubule severing. No repair sites were observed in the absence of spastin. Thus, the local balance between active tubulin removal catalyzed by the enzyme and passive tubulin incorporation determines whether a nanoscale damage site progresses to a mesoscale severing event or fails to do so because of the repair with GTP-tubulin from the soluble pool.

We also visualized the lattice-incorporated tubulin at a higher resolution by using TEM. We generated recombinant human α 1 β III tubulin with an engineered FLAG tag at the β -tubulin C terminus (38). We then used this recombinant tubulin to repair brain microtubules damaged at the nanoscale by spastin. The presence of the FLAG tag on the recombinant tubulin allowed specific detection of recombinant tubulin both in fluorescence and TEM images with the use of fluorescent or gold-conjugated secondary antibodies against FLAG antibodies (Materials and methods). Fluorescence microscopy revealed that the recombinant tubulin robustly incorporates along microtubules with nanoscale damage by spastin with ATP. No incorporation was detected with spastin and adenosine 5'-O-(3-thiotriphosphate) (ATP- γ -S) (fig. S6). TEM showed the discrete, productive incorporation of recombinant α 1 β III tubulin in islands along microtubules and the absence of tubulin aggregates at nanoscale damage sites (fig. S7). The FLAG primary and secondary gold-conjugated antibodies are specific for the recombinant tubulin, as brain microtubules showed only background antibody decoration (fig. S7, C and D). In the absence of recombinant soluble tubulin in the reaction, microtubules were robustly damaged at the nanoscale under these conditions (fig. S7E). Moreover, neither recombinant tubulin incorporation nor association with the microtubule lattice was observed by fluorescence and TEM assays with the slow-hydrolyzing analog ATP- γ -S (figs. S6 and S7, A and C). Thus, soluble tubulin was incorporated productively into the microtubule lattice at nanoscale damage sites created by spastin in an ATP hydrolysis-dependent manner.

Severing enzymes promote rescues

Because spastin and katanin catalyze GTP-tubulin incorporation along microtubules, we next examined their effects on microtubule dynamics. It has been recognized for 30 years that tubulin incorporation into a growing microtubule stimulates hydrolysis of the bound GTP. The resulting GDP-tubulin lattice is unstable but is protected from depolymerization by a layer of GTP-tubulin. This GTP cap at the microtubule end results from a lag between the GTP hydrolysis rate on the incorporated tubulin and the microtubule growth speed (30, 31, 39–42). More recently, islands of GTP-tubulin were detected along microtubules in cells and were correlated with rescue (34, 43)—the transition from depolymerization to growth, one of the parameters of microtubule dynamic instability. As in stabilized GMPCPP-microtubules and GMPCPP-capped GDP-microtubules, the newly perfused GTP-tubulin was

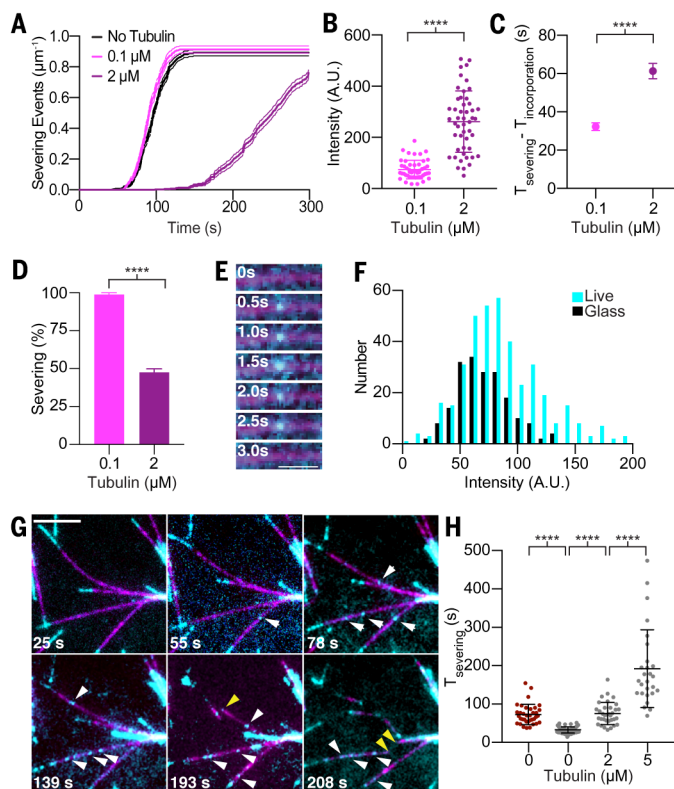
rapidly incorporated along the GDP-microtubule lattice of dynamic microtubules in the presence of spastin and katanin with ATP; by contrast, in the control without ATP, addition was visible only at microtubule ends (Fig. 4, A to D, and movies S2 and S3). We then characterized microtubule dynamics in the presence of spastin or katanin at physiological concentrations [25 nM; spastin and katanin concentrations in HeLa cells are 46 and 28 nM, respectively (37)]. At these enzyme concentrations, we observed robust microtubule severing and internal GTP-tubulin incorporation. Spastin and katanin increased rescue frequencies by factors of ~13- and 9, respectively (with $0.5 \pm 0.2 \text{ min}^{-1}$ for the control versus $6.6 \pm 1.6 \text{ min}^{-1}$ and $4.5 \pm 0.7 \text{ min}^{-1}$ for spastin and katanin, respectively) (Fig. 4E). Whereas only 13%

of the depolymerization events were rescued in the control, 61% were rescued in the presence of spastin or katanin (Fig. 4F). Consistent with their promotion of tubulin exchange along the microtubule shaft, spastin and katanin had no significant effect on rates of microtubule growth and catastrophe (Fig. 4, G and H). This is in contrast to other rescue-promoting factors, such as cytoplasmic linker-associated proteins, which promote rescue by increasing the on rate of tubulin dimers at microtubule ends and thus decrease catastrophe and increase growth rates (44), or conventional microtubule-associated proteins (MAPs) such as MAP2, which promote rescue by stabilizing the microtubule lattice (45).

In our dynamics assays, tubulin was continually extracted by the enzyme, while at the same

Fig. 3. Incorporation of soluble tubulin into spastin-induced nanoscale damage sites inhibits microtubule severing.

(A) Severing rates in the presence of soluble tubulin ($n = 31$, 28, and 36 microtubules from multiple chambers for no tubulin, 100 nM tubulin, and 2 μM tubulin, respectively). Thin lines indicate SEM. (B) Intensity distribution of fluorescent tubulin puncta incorporated at spastin-induced nanoscale damage sites ($n = 50$ and 49 puncta from multiple chambers for 100 nM and 2 μM tubulin, respectively). Bars indicate the means and SD. A.U., arbitrary units. (C) Repair at



damage sites delays severing ($n = 81$ and 83 severing events from multiple chambers for 100 nM and 2 μM tubulin, respectively). T_{severing} and $T_{\text{incorporation}}$, time required for a severing event and for initial incorporation of tubulin, respectively. (D) Fraction of GMPCPP-microtubules severed by 20 nM spastin within 65 s of initial tubulin incorporation in the presence of 100 nM and 2 μM HiLyte 488-labeled soluble tubulin. Error bars indicate SEM in (C) and (D). (E) Live imaging of Alexa 488-labeled GTP-tubulin (cyan) incorporation into HiLyte 647-labeled GMPCPP-microtubules (magenta) after spastin-induced damage. Scale bar, 1.5 μm . (F) Fluorescence intensity distribution of Alexa 488-labeled tubulin (labeling ratio, ~1.0) immobilized on glass or incorporated into spastin-induced nanoscale damage sites ($n = 188$ and 398 for glass-immobilized and microtubule-incorporated particles, respectively). (G) Spastin-induced nanoscale damage and spontaneous tubulin repair of GDP-microtubules (magenta) grown from axonemes and stabilized with a GMPCPP cap (bright cyan) in the presence of spastin (5 nM) and 5 μM soluble HiLyte 488-labeled GTP-tubulin (cyan). Images were bleach corrected. White arrowheads, tubulin incorporation sites; yellow arrowheads, severing events. Scale bar, 5 μm . (H) Average completion time of a severing event after spastin perfusion. Brown, GMPCPP-microtubules; gray, GMPCPP-capped GDP-microtubules in the absence or presence of soluble tubulin ($n = 36$, 63, 34, and 27 microtubules from multiple chambers for GMPCPP-microtubules and GMPCPP-capped GDP-microtubules with 0, 2 μM , and 5 μM soluble GTP-tubulin, respectively). Bars indicate the means and SD. **** $P < 0.0001$, determined by a two-tailed t test, for (B), (C), (D), and (H).

time the lattice was healed with newly incorporated GTP-tubulin that was gradually converted into GDP-tubulin. To decouple these processes and establish directly whether the GTP-tubulin islands introduced by these enzymes can act as microtubule rescue sites, we introduced non-hydrolyzable GTP-tubulin islands into the microtubule. We induced nanoscale damage to a GMPCPP-capped GDP-microtubule with spastin or katanin and healed it with GMPCPP-tubulin, removed the enzyme and GMPCPP-tubulin from the chamber, and initiated microtubule depolymerization through laser ablation close to the GMPCPP cap (Fig. 5; fig. S8, A to C; and Materials and methods). No GMPCPP-tubulin incorporation was detected in the control performed in the presence of enzyme without ATP. These microtubules depolymerized all the way to the seed upon ablation (Fig. 5B and fig. S8A). In contrast, microtubules with GMPCPP-tubulin islands incorporated along their lengths through the ATP hydrolysis-dependent activity of spastin or katanin were stabilized against depolymerization at the location of the island (Fig. 5C and fig. S8B), despite the absence of soluble tubulin in the chamber: 75% and 76% paused when they encountered a GMPCPP island introduced by spastin and katanin, respectively (Fig. 5D; fig. S8, A to C; and movie S4). Those that depolymerized through the island showed a decrease in the depolymerization speed (Fig. 5E and fig. S8D). Moreover, fluorescence intensity analysis revealed that GMPCPP islands that paused depolymerization were statistically significantly brighter than those that did not (Fig. 5F and fig. S8E). Next, we wanted to establish whether these enzyme-generated GMPCPP islands were competent to support microtubule regrowth. We again performed the above-described experiment, but during the last step we introduced 7 μM soluble GTP-tubulin into the chamber (Fig. 5, A, G, and H, and movie S5). Whereas at these tubulin concentrations rescue events were very rare in the control, we saw a higher probability of rescue of microtubules with spastin-incorporated GMPCPP islands (Fig. 5I). When the GMPCPP island did not support a rescue, it did slow down depolymerization (Fig. 5J). Moreover, fluorescence intensity analysis revealed that GMPCPP islands that supported rescues were significantly brighter than those that did not (Fig. 5K). Thus, microtubule dynamics measurements and experiments with GMPCPP-tubulin islands indicate that GTP islands introduced in a microtubule severing enzyme-dependent manner promote microtubule rescue and that there is a minimal local GTP-tubulin density required to robustly support rescue at that site. Because the microtubule is rescued when the balance shifts from net tubulin loss to net tubulin addition, it is likely that the correlation between the size of the GTP-tubulin island and rescue probability will vary with the tubulin concentration or the presence of MAPs. Thus, smaller GTP-tubulin islands may still be effective as rescue sites at higher tubulin concentrations or in the presence of MAPs that increase the tubulin on rate.

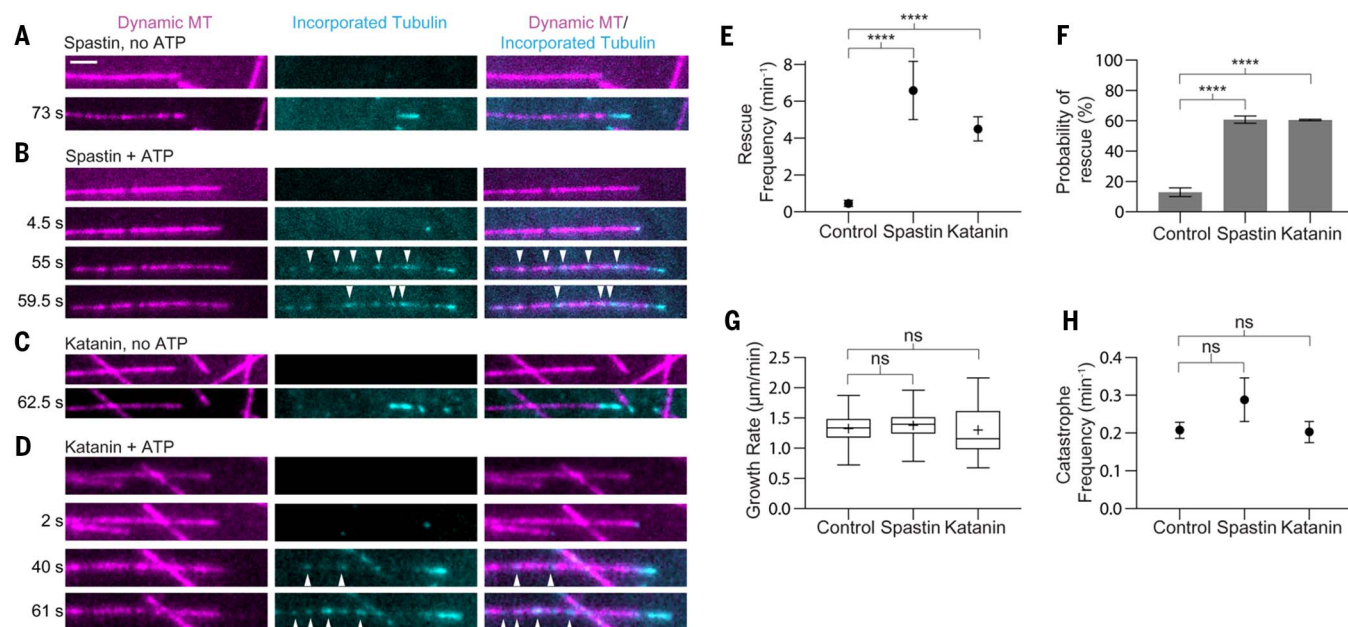


Fig. 4. Spastin and katanin promote GTP-tubulin island formation and increase rescues. (A and B) Time course of a dynamic 10% HiLyte 647-labeled microtubule at 12 μ M tubulin in the presence of 25 nM spastin without (A) or with (B) ATP showing HiLyte 488-labeled tubulin incorporation at the microtubule tip (A) or incorporation (arrowheads) along the microtubule in addition to the tip (B). The first micrograph for each condition was recorded just before the perfusion of the chamber with 12 μ M 10% HiLyte 488-labeled tubulin. Scale bar, 2 μ m. (C and D) Time course of a dynamic 10% HiLyte 647-labeled microtubule at 12 μ M tubulin in the presence of 25 nM katanin without (C) or with (D) ATP showing HiLyte 488-labeled tubulin incorporation at the microtubule tip (C) or incorporation (arrowheads) along the microtubule in addition to the tip (D). The first micrograph for each condition was recorded just before the perfusion of the chamber with 12 μ M 10% HiLyte

488-labeled tubulin. (E) Rescue frequency at 10 μ M tubulin in the absence or presence of 25 nM spastin and 25 nM katanin with ATP ($n = 47, 45$, and 61 microtubules from multiple chambers for the control without enzyme, spastin, and katanin, respectively). **** $P < 0.0001$, determined by the Mann-Whitney test. (F) Probability of rescue of a depolymerizing microtubule in the absence or presence of spastin and katanin with ATP ($n = 68, 57$, and 78 depolymerization events for the control, spastin, and katanin, respectively). **** $P < 0.0001$, determined by a two-tailed t test. (G and H) Growth rates (G) and catastrophe frequency (H) in the absence or presence of spastin and katanin with ATP [$n = 56, 37$, and 34 growth events for the control, spastin, and katanin, respectively, in (G) and $n = 62, 70$, and 71 microtubules for the control, spastin, and katanin, respectively, in (H)]. Plus signs in (G) indicate the means. ns, not significant. Error bars indicate SEM throughout.

Severing enzyme-generated GTP islands recruit EB1

The GTP state of tubulin is recognized by MAPs belonging to the end-binding (EB) protein family. EB1 preferentially binds to growing microtubule ends by sensing the GTP [or GDP-inorganic phosphate (P_i)] state of tubulin (46, 47). Consistent with the creation of GTP-tubulin islands, in the presence of spastin or katanin and ATP we observed EB1 not only at the growing ends as in the control but also as distinct puncta along microtubules (Fig. 6, A to D). These puncta are reminiscent of the EB3 puncta observed at sites of tubulin repair after laser-induced damage (34). Of the newly incorporated GTP-tubulin islands, 89% colocalized with EB1 (Fig. 6, E and F). These EB1 puncta were transient, consistent with the dynamic removal and incorporation of new tubulin into the lattice and the gradual GTP hydrolysis of the incorporated tubulin (Fig. 6, A and C; fig. S9; and movie S6). Consistent with a protective effect of the GTP islands, microtubule dynamics assays in the presence of spastin and EB1 revealed that 74% of rescues were associated with the presence of EB1 at the rescue site (fig. S10A). This number is significantly higher than the pre-

diction given by the random superposition of EB1 puncta and rescue events (74% versus 14%; $P < 0.0001$ by Fisher's exact test) (Materials and methods). Similarly, 63% of rescues in the presence of katanin occurred at the site of an EB1 spot (fig. S10B), compared with 0% when the distribution was randomized ($P < 0.00001$ by Fisher's exact test) (Materials and methods). Laser ablation of microtubules peppered with EB1 puncta also revealed a marked increase in rescue frequency. Whereas microtubules were rescued from 100% of ablation-induced depolymerization events within 4 s, they were rescued from only 15% of events in the presence of spastin and ATP- γ -S (Fig. 6, G and H). Similar results were obtained with katanin (Fig. 6I and movie S7). Thus, the ATP-dependent action of the enzyme that promotes tubulin exchange within the lattice is required for the observed increase in rescue frequency.

Severing amplifies microtubule mass and number

The GDP-tubulin lattice is unstable and, when exposed by laser ablation, is rapidly depolymerized at the plus ends, even in the presence of soluble tubulin (Fig. 7A), a result consistent

with those of classic experiments performed with laser-ablated or mechanically cut microtubules (28, 29, 48–51). Surprisingly, at 12 μ M tubulin, the majority of new plus ends generated by spastin or katanin were stable and rapidly reinitiated growth (Fig. 7, B and C). By contrast, in the absence of either enzyme or in the presence of spastin or katanin and ATP- γ -S, new plus ends generated through laser ablation rapidly depolymerized (Figs. 6, H and I, and 7A). This indicates that it is not the passive binding of the protein that stabilizes the new plus ends against spontaneous depolymerization but the ATP-dependent incorporation of GTP-tubulin at severing sites. The minus ends were stable regardless of whether they were generated through enzyme action or laser ablation, consistent with results of earlier experiments using laser ablation (28, 29). Thus, when local tubulin extraction by spastin or katanin outpaces the rate of tubulin incorporation, a severing event occurs and the newly severed microtubule ends emerge with a high density of GTP-tubulin that is protective (Fig. 7, B and C). Moreover, the plus ends that depolymerize immediately after severing resume growth after a lower net loss of polymer mass (fig. S8, F and G).

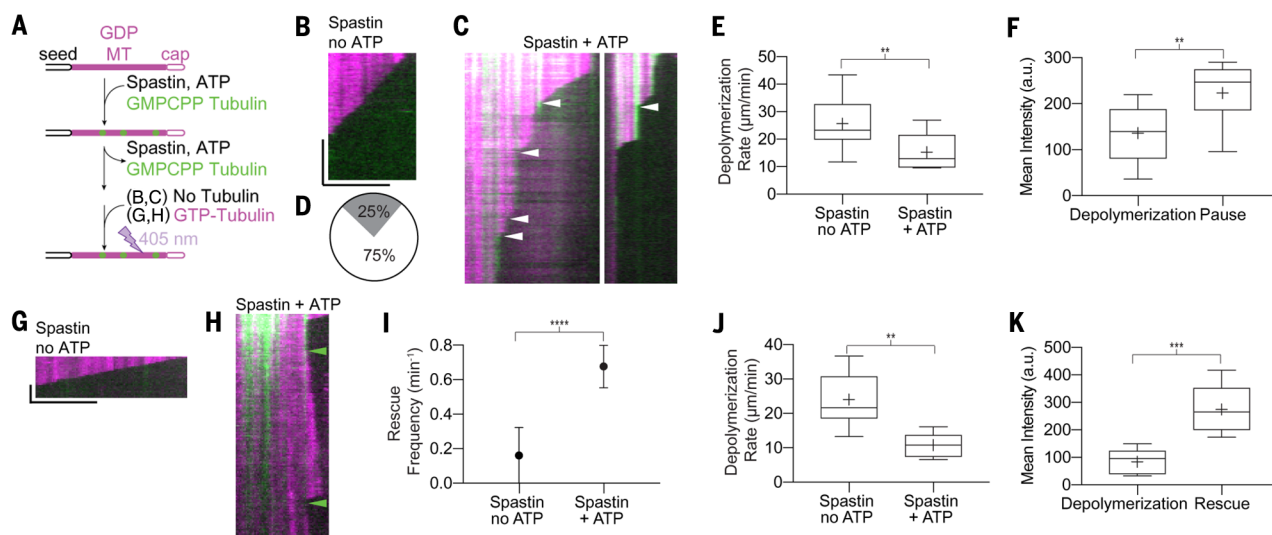


Fig. 5. Enzyme-generated GMPCPP islands protect against depolymerization and act as rescue sites. (A) Experiment schematic. GDP-microtubules (solid magenta) were polymerized from seeds and capped with GMPCPP-tubulin (magenta outline). Spastin, ATP, and GMPCPP-tubulin (green) were added and washed out of the chamber. Microtubules were laser ablated in the absence (B to F) or presence (G to K) of GTP-tubulin (Materials and methods). (B) Kymograph of a depolymerizing laser-ablated microtubule (magenta) preincubated with spastin and no ATP. Horizontal scale bar, 5 μm ; vertical bar, 10 s. (C) Kymographs of depolymerizing laser-ablated microtubules pausing at GMPCPP-tubulin islands (green) introduced by spastin with 1 mM ATP. Arrowheads, pauses. (D) Pie chart showing the proportion of depolymerization events that paused at GMPCPP islands (white) or did not (gray) ($n = 44$ events). (E) Depolymerization rates of microtubules without GMPCPP islands preincubated with spastin and no ATP or of microtubules depolymerized through GMPCPP islands introduced by spastin with 1 mM ATP ($n = 17$ and 7 microtubules for no ATP and ATP, respectively). (F) Fluorescence intensity of GMPCPP islands through which microtubules depolymerized or paused ($n = 9$ and 14 islands, respectively). (G and H) Kymographs of laser-ablated microtubules in the presence of 7 μM soluble GTP-tubulin after preincubation with spastin and no ATP showing complete depolymerization (G) or rescue (arrowheads) at a GMPCPP island introduced by spastin with ATP (H). Horizontal scale bar, 5 μm ; vertical bar, 20 s. (I) Rescue frequency for laser-ablated microtubules incubated with spastin with or without ATP ($n = 23$ and 24 microtubules with and without ATP, respectively). (J) Depolymerization rates in the presence of 7 μM GTP-tubulin for microtubules preincubated with spastin and no ATP or for microtubules that depolymerized through GMPCPP islands introduced by spastin with ATP ($n = 9$ and 6 microtubules without and with ATP, respectively). (K) Fluorescence intensity of GMPCPP islands that did not stop depolymerization ($n = 6$) or at which microtubules were rescued in the presence of spastin and ATP ($n = 9$). $^{**}P < 0.01$, $^{***}P < 0.001$, determined by the Mann-Whitney test. Plus signs in (E), (F), (J), and (K) indicate means. Error bars indicate SEM throughout.

Thus, the increase in microtubule number with each severing event (Fig. 7, D and E) synergizes with the higher rescue frequency to produce a rapid amplification of total microtubule number and mass (Fig. 7, F to J).

Discussion

The classical view of microtubule dynamics has been that tubulin dimer exchange occurs exclusively at microtubule ends through polymerization and depolymerization (30, 52). By visualizing a severing reaction at the ultrastructural level, we have shown that spastin and katanin extract tubulin subunits from the microtubule (Fig. 1) and that this ATP hydrolysis-dependent tubulin removal is counteracted by spontaneous lattice incorporation of soluble GTP-tubulin (Figs. 2 to 4 and figs. S5 to S7). The nanoscale-damaged microtubules do not immediately unravel but are long-lived enough to have a chance to heal through the productive incorporation of tubulin into the lattice. Because longitudinal lattice contacts are stronger than lateral ones (42), we speculate that tubulin dimer loss from the microtubule wall has a slight longitudinal bias that proceeds along the protofilament. This would give the microtubule a chance to heal before it is severed across and

generate GTP-tubulin islands that consist of several tubulin dimers in the longitudinal direction. The geometry of the nanoscale damage sites and the mechanism of tubulin incorporation and conformational changes at these sites will be exciting and fundamental areas for future exploration.

This mechanism of lattice repair can explain the earlier observation of the inhibition of katanin severing by soluble tubulin (53, 54). The ragged, Swiss cheese nature of the nanoscale-damaged microtubules is conducive to healing, as the incoming tubulin dimers can make stabilizing lateral interactions. Thus, depending on the local rates of the severing enzyme-catalyzed tubulin removal and the spontaneous incorporation of new GTP-tubulin into the lattice, the action of a microtubule-severing enzyme results in a severing event where the newly emerging ends have a high density of GTP-tubulin or a microtubule that preserves integrity but acquires a GTP island at the site of enzyme action. The higher GTP density at the newly severed ends can also act to quickly recruit molecular motors and MAPs that can modulate the fate of the newly generated end.

Although in vitro microtubule repair after the introduction of defects through laser-induced photodamage (34) or mechanical stress (33, 55)

has been reported previously, our study identifies a family of enzymes as biological agents that promote the ATP-dependent incorporation of GTP-tubulin islands into microtubules. Microtubule repair has a high incidence in vivo at microtubule crossovers or bundles (34), where microtubule-severing enzymes have been shown to act (7, 8, 17, 56). Our findings thus suggest that the high incidence of repair at these sites is due not exclusively to mechanical damage (34) but also to the action of microtubule-severing enzymes. As spastin and katanin preferentially target glutamylated microtubules (13, 57, 58), they may also selectively rejuvenate aging microtubules with accumulated glutamylation marks through GTP-tubulin incorporation. GTP-tubulin islands have been identified along axonal microtubules (59), a neuronal compartment where severing enzymes act. This finding raises the possibility that severing enzymes are also used as quality control and maintenance factors in hyperstable microtubule arrays, such as those in axons, centrioles, and cilia, where spastin and katanin are important for biogenesis and maintenance (2, 5, 13, 14) and where spastin and katanin may serve to remove and replace old, possibly damaged tubulin subunits without affecting overall microtubule

organization. Future work should establish how impaired lattice repair contributes to the disease phenotypes seen in patients with spastin and katanin mutations.

Our study shows that the severing enzyme-catalyzed incorporation of GTP-tubulin along microtubules has two physiological consequences: It increases the frequency at which microtubules are rescued (Figs. 4 to 6), and it stabilizes newly severed plus ends that emerge against de-

polymerization with a high density of GTP-tubulin (Fig. 7). Thus, microtubule dynamics can be modulated not only by factors that affect tubulin incorporation at microtubule ends but also by severing enzymes that promote the exchange of tubulin subunits within the microtubule shaft. The synergy between the increased rescue rates and the stabilization of the newly severed ends leads to microtubule amplification in the absence of a nucleating factor, explaining why, paradox-

ically, the loss of spastin and katanin results in the loss of microtubule mass in many systems (2, 23, 25, 27). Such a mechanism of polymer amplification has parallels to the actin cytoskeleton, where severed filaments are used for templated actin polymerization [(26, 60); reviewed in (61)]. When severing enzymes are expressed at high levels or are positively regulated, tubulin extraction outpaces repair and the microtubule array disassembles. Cells likely modulate severing activity and the rate of tubulin lattice incorporation through the action of MAPs to elicit these two different outcomes. This regulation will be a notable area of future exploration.

Materials and methods

Protein expression and purification

Drosophila melanogaster full-length spastin was purified by affinity chromatography and ion exchange as previously described (62). *Caenorhabditis elegans* MBP-tagged katanin Meil/Mei2 (12) was purified on amylose resin. The affinity tag was removed by tobacco etch virus protease, and the protein was further purified on an ion exchange MonoS column (GE Healthcare) as previously described (63). Peak fractions were concentrated, buffer was exchanged into 20 mM HEPES (pH 7.0), 300 mM KCl, 10 mM MgCl₂, and 1 mM TCEP, and fractions were flash frozen in small aliquots in liquid nitrogen. *Homo sapiens* EB1-green fluorescent protein (GFP) was expressed and purified as previously described (64). Human α 1A β III tubulin with an engineered FLAG tag at the β -tubulin C terminus was expressed by using baculovirus and purified as described previously (38).

Transmission electron microscopy of microtubule-severing reactions

Taxol-stabilized GDP-microtubules were prepared by polymerizing 10 μ l of 100 μ M glycerol-free porcine tubulin (Cytoskeleton, Denver, CO) in 80 mM K-PIPES (pH 6.8), 1 mM MgCl₂, 1 mM EGTA, 10% DMSO, and 1 mM GTP for 1 hour in a 37°C water bath. Taxol was added to 20 μ M final concentration, and the reaction was incubated on the bench top for 1 to 2 hours. Microtubules were loaded onto a 60% glycerol cushion [BRB80, 60% (v/v) glycerol, and 20 μ M Taxol] at 37°C by using a pipette tip with the tip cut off. Non-polymerized tubulin was removed by centrifugation in a TLAI00 rotor at 35,000 rpm for 15 min at 37°C. The pellet was gently resuspended to 2.5 μ M tubulin in BRB80 supplemented with 20 μ M Taxol and 1 mM GTP at 37°C by using a pipette tip with the tip cut off.

For GDP-microtubules, all polymerization and severing reactions were performed at 37°C. Twenty microliters of 100 μ M glycerol-free porcine tubulin (Cytoskeleton) was polymerized in 10% DMSO, 1 mM GTP, and 10 mM MgCl₂ for 1 hour at 37°C in a water bath. The microtubules were passed through a 60% glycerol cushion [BRB80, 60% (v/v) glycerol, and 1 mM GTP] by using a TLAI00 rotor at 53,000 \times g for 15 min to remove nonpolymerized tubulin. The pellet was washed twice using 50 μ l of buffer (BRB80, 10% DMSO, 1 mM GTP)

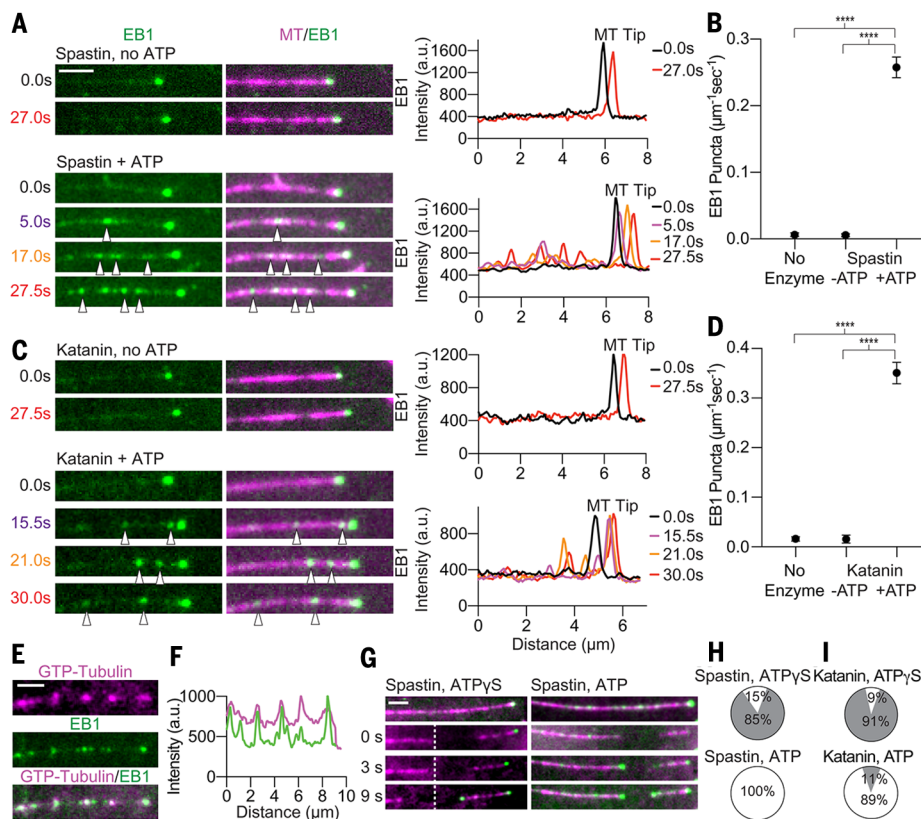


Fig. 6. Spastin- and katanin-generated GTP-tubulin islands recruit EB1. (A) Time course of EB1-GFP on a dynamic microtubule in the presence of 25 nM spastin without or with ATP. Scale bar, 2 μ m. Line scans on the right show EB1-GFP intensity profiles along the microtubule at the indicated times. Intensity profiles start on the microtubule lattice and end at the microtubule tip. Arrowheads show lattice EB1 puncta. (B) Density of EB1-GFP puncta on microtubules incubated without spastin or with spastin without and with ATP. Error bars indicate SEM. **** $P < 0.0001$. (C) Time course of EB1-GFP on a dynamic microtubule in the presence of 25 nM katanin without and with ATP. Intensity profiles are as in (A). Arrowheads show lattice EB1 puncta. (D) Density of EB1-GFP puncta on microtubules incubated without katanin or with katanin without or with ATP. Error bars indicate SEM. **** $P < 0.0001$. (E) Colocalization of newly incorporated GTP-tubulin (top) and EB1-GFP (middle) in the presence of spastin and ATP. (Bottom) Overlay. Images were acquired immediately after the perfusion of the chamber with enzyme and EB1-GFP. Scale bar, 2 μ m. (F) Fluorescence intensity of incorporated tubulin (magenta) and EB1-GFP (green) along the microtubule lattice in (E) showing their colocalization. Eighty-nine percent of tubulin islands colocalize with EB1-GFP ($n = 38$ puncta from 22 microtubules from multiple chambers measured immediately after perfusion with 10% HyLite 647-tubulin). (G) Time course of laser-ablated dynamic microtubules (magenta) incubated with 25 nM spastin with ATP- γ -S or spastin with ATP in the presence of 50 nM EB1-GFP (green) (Materials and methods). The dotted line marks the ablated region and the start of depolymerization. Scale bar, 2 μ m. (H and I) Pie charts show the fates of plus ends generated through laser ablation of microtubules incubated with spastin (H) or katanin (I) with ATP- γ -S or ATP. The percentage of plus ends that depolymerized (gray) or were rescued (white) within 4 s after ablation is shown ($n = 13$ and 13 microtubules from multiple chambers for spastin with ATP- γ -S and ATP, respectively; $n = 54$ and 9 microtubules from multiple chambers for katanin with ATP- γ -S and ATP, respectively).

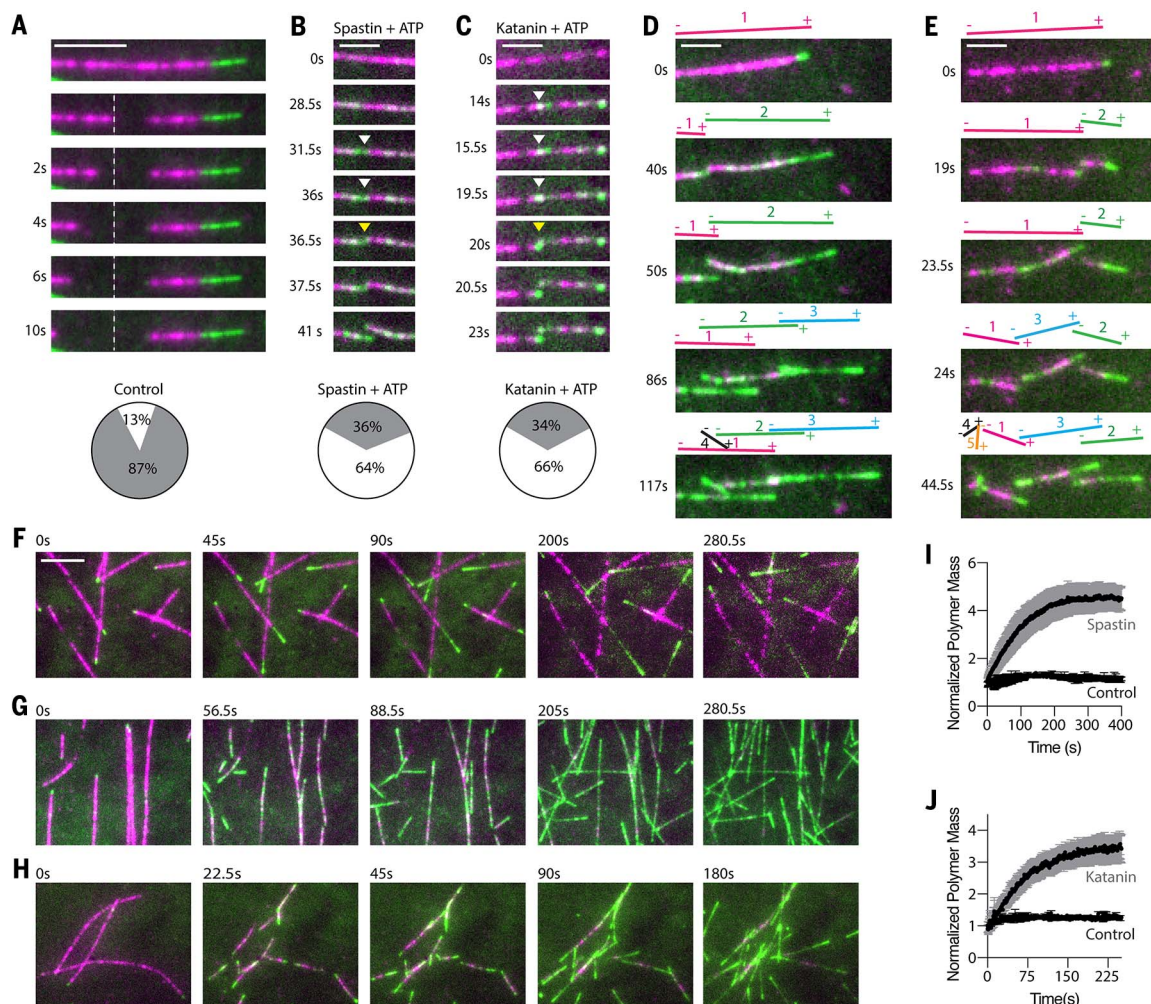
Fig. 7. Severing enzyme-based microtubule number and mass amplification.

(A) Plus ends generated through laser ablation depolymerize. The pie chart shows the percentages of plus ends that are stable (white) or depolymerize (gray) ($n = 32$ microtubules from multiple chambers). Scale bar, 5 μm . The dashed line marks the ablated region and the start of depolymerization.

(B and C) Spastin (B)- or katanin (C)-severed ends emerge with newly incorporated GTP-tubulin and are stable. The pie charts show the percentages of plus ends that are stable (white) or depolymerize (gray) ($n = 96$ and 94 microtubules from multiple chambers for spastin and katanin, respectively). White and yellow arrowheads indicate tubulin incorporation and a severing event, respectively.

Scale bars, 2 μm . (D and E) Time-lapse images showing consecutive spastin (D)- or katanin (E)-induced severing events on a microtubule. Lines and numbers indicate individual microtubules after severing. + and - mark microtubule ends. Magenta, microtubule; green, incorporated tubulin. Scale bars, 2 μm . (F) Time-lapse images showing microtubule dynamics at 12 μM tubulin in the absence of a severing enzyme. Green, newly incorporated tubulin at the growing

ends. The last two frames are bleach corrected. Scale bar, 5 μm . (G and H) Time-lapse images showing microtubule number and mass amplification through spastin (G) and katanin (H) severing. Green, newly incorporated HiLyte 488-tubulin perfused into the chambers together with the severing enzymes. (I and J) Microtubule mass as a function of time ($n = 4, 5$, and 4 chambers for the control, spastin, and katanin, respectively). Error bars indicate SEM.



and gently resuspended to 30 μM in the same buffer by using a pipette tip with the tip cut off.

GMPCPP-microtubules were prepared by polymerizing 20 μl of 100 μM glycerol-free porcine tubulin (Cytoskeleton) in 1 mM GMPCPP in BRB80 [80 mM PIPES-KOH (pH 6.8), 1 mM MgCl_2 , 1 mM EGTA, 1 mM DTT] on ice for 5 min and then in a water bath at 37°C for 1 hour. Nonpolymerized tubulin was removed by centrifugation in a TLA100 rotor at 126,000 $\times g$ for 5 min at 37°C. The pellet was washed twice with 50 μl of BRB80 at 37°C and resuspended in 50 μl of ice-cold BRB80. The reaction mixture was kept on ice for 30 min and periodically mixed up and down to fully depolymerize microtubules. GMPCPP was added to 1 mM, and the polymerization reaction mixture was kept on ice for 10 min and then transferred to 37°C for 2 to 4 hours or overnight. Nonpolymerized tubulin was removed by centrifugation

and washed as described above. The microtubule pellet was gently resuspended to 2.5 μM tubulin in BRB80 by using a pipette tip with the tip cut off.

We found that performing severing reactions in the tube followed by pipetting onto EM grids resulted in microtubule breakage. We therefore first carried out severing reactions on the EM grid. Briefly, 2 μl of microtubule solution (at 1 to 3 μM) in BRB80 [80 mM PIPES (pH 6.8), 1 mM MgCl_2 , 1 mM EGTA] was applied to a glow-discharged Cu grid, followed by pipetting of 2 μl of ATP solution (10 mM ATP in BRB80 supplemented with 20 mM Taxol for Taxol-stabilized microtubules) and 2 μl of spastin (at 100 nM). The reaction was allowed to proceed on the grid for 1 min or as specified, after which the liquid was wicked off with calcium-free filter paper and the grid was stained with 0.75% (w/v) uranyl formate

and air-dried. Images were collected on a FEI Morgagni 286 electron microscope operated at 80 kV and equipped with an AMT lens-coupled 1k \times 1k CCD camera. For the solution severing reaction time courses, 20 μl of GMPCPP or Taxol-stabilized microtubules in BRB80 buffer at 2.5 and 1.0 μM was applied to parafilm, followed by the addition of 20 μl of 50 nM spastin or 200 nM katanin in 20 mM HEPES (pH 7.5), 300 mM KCl, 10 mM MgCl_2 , 1 mM TCEP, and 1 mM ATP to a final concentration of 25 nM spastin and 100 nM katanin. For the solution severing reaction time courses of nonstabilized GDP-microtubules, 20 μl of 30 μM GDP-microtubules in the presence of 10% DMSO was incubated with 2 μl of 20 nM katanin. Buffer without severing enzymes was added to microtubules as a negative control. The severing reaction mixtures were incubated for 30 s or 2 or 5 min, and carbon-coated grids (carbon film

only on 400 mesh; Ted Pella) were dipped into the reaction mixtures. Excess liquid was blotted with filter paper. Grids were washed three times with 40 μ l BRB80, stained with 0.75% (w/v) uranyl formate, and air-dried. Images were collected on a T12 Technai electron microscope (FEI) equipped with a 2k \times 2k Gatan US1000 CCD camera. Images were collected at nominal magnifications of 550 \times , 13,000 \times , or 30,000 \times , corresponding to pixel sizes of 84 $\text{\AA}/\text{pix}$, 3.55 $\text{\AA}/\text{pix}$, or 1.54 $\text{\AA}/\text{pix}$, respectively.

TIRF-based assays of tubulin incorporation into stabilized microtubules damaged by spastin and katanin

Double-cycled, GMPCPP-stabilized microtubules (65) were polymerized from 2-mg/ml porcine brain tubulin (Cytoskeleton). The first polymerization was 1 hour, and the second polymerization step was at least 4 hours to obtain long microtubules. Then microtubules were centrifuged, resuspended in warm BRB80 [80 mM K-PIPES (pH 6.8), 1 mM MgCl_2 , 1 mM EGTA], and stored at 37°C or room temperature (RT) before use. The same results were obtained regardless of whether the storage temperature was 37°C or RT. Taxol-stabilized microtubules (62) were polymerized from 5-mg/ml porcine brain tubulin containing 1% biotinylated and 20% HiLyte 647-labeled tubulin (Cytoskeleton) in BRB80 with 10% DMSO, 0.5 mM GTP, and 10 mM MgCl_2 . After 1 hour of incubation at 37°C, 20 μ M Taxol was added and the mixture was further incubated overnight. Microtubules were then centrifuged through a 60% glycerol cushion for 12 min at 109,000 \times g at 35°C. The microtubule pellet was washed with warm BRB80 supplemented with 14.3 mM 2-mercaptoethanol and 20 μ M Taxol and was resuspended gently in the same buffer.

Chambers for TIRF microscopy were assembled as previously described (62). Double-cycled GMPCPP-microtubules containing 1% biotinylated tubulin and 20% HiLyte 647-labeled tubulin (or unlabeled tubulin for the DIC assays) assembled as described above were immobilized in the chamber with 2-mg/ml NeutrAvidin (Thermo Fisher Scientific) and imaged by TIRF or DIC microscopy in severing buffer [BRB80 buffer with 2-mg/ml casein, 14.3 mM 2-mercaptoethanol, 2.5% glycerol, 50 mM KCl, 2.5 mM MgCl_2 , 1 mM ATP, 1% Pluronic F127 (Life Technologies), and oxygen scavengers]. To introduce and detect nanoscale damage in microtubules (Fig. 2), immobilized microtubules were then incubated with 10 nM spastin or 2 nM katanin in severing buffer for 35 or 90 s, respectively. Microtubules in control experiments were incubated without severing enzyme. The enzyme mixture was then replaced with 1 μ M HiLyte 488-labeled tubulin (Cytoskeleton), 1 mM ADP, 0.5 mM GTP, 1% Pluronic F127, and 2.5-mg/ml casein in BRB80 and left to incubate for 5 min. The tubulin-containing solution was then washed out with 45 μ l of BRB80 supplemented with oxygen scavengers, 1.5-mg/ml casein, 10 mM 2-mercaptoethanol, and 1% Pluronic F127. Microtubules and HiLyte 488-labeled

tubulin were imaged by TIRF. Multiple fields of view were imaged. The same assay was performed for Taxol-stabilized microtubules, but in this case the repair step was performed with 0.1 μ M soluble tubulin to prevent microtubule nucleation in the presence of Taxol. For time-course experiments, the same protocol was used except that microtubules were incubated with 2 nM spastin (fig. S3) or 2 nM katanin (fig. S4) for 35 to 120 s. Control microtubules were incubated without severing enzyme for 120 s. HiLyte 488-labeled tubulin (1 μ M) was used for the repair step. For repair with 1 μ M recombinant human tubulin (fig. S6), nanoscale-damaged microtubules were incubated for 5 min with recombinant tubulin. Unincorporated tubulin was washed away, and tubulin incorporated into microtubules was detected by anti-FLAG M2 antibodies (Sigma-Aldrich; diluted 1:500) and goat anti-mouse antibodies conjugated with Alexa Fluor 488 (Invitrogen; diluted 1:1000). All assays were performed at RT. Details regarding image acquisition and analysis are described in the subsection below.

Image acquisition and analysis of tubulin incorporation in GMPCPP- and Taxol-stabilized microtubules by TIRF microscopy

Images were acquired by using a Nikon Ti-E microscope equipped with a 100 \times 1.49 NA oil objective and a TI-TIRF adapter (Nikon). The 488 excitation laser (Coherent) was set at 20 mW, and the 647 nm laser (Coherent) was set to 2 mW before being coupled into the Ti-TIRF optical fiber (Nikon). Two-color simultaneous imaging was performed by using a TuCAM (Andor) device that splits the emission onto two separate EMCCD cameras (Andor iXon 897). The excitation and emission were split by a quad band dichroic (Semrock), and the emission was further split by an FF640 filter (Semrock) and further filtered with an FF01-550/88 (Semrock) for the 488 channel and an FF01-642/LP (Semrock) for the 640 channel. The TuCAM imaging system introduces an extra 2 \times magnification, yielding a final pixel size of 77 nm. The images from the two cameras were aligned by first imaging a grid of spots (Nanogrid MiralomaTech) on each camera and using the GridAligner plug-in for ImageJ.

DIC illumination was provided by a SOLA-SE-II (Lumencor) coupled to the microscope by a liquid light guide. A standard set of polarizer and analyzer (Nikon 100 X-II High NA/Oil) prisms was used, and the image was captured on a CoolSNAP (Photometrics) camera. The final pixel size for DIC images was 65 nm. Raw DIC images were processed using an FFT band-pass filter. DIC images were scaled and transformed to overlay with fluorescent images by imaging fluorescent microtubules in both channels for image registration. The entire imaging setup was controlled by Micro-Manager (66).

For data shown in figs. S3 and S4, images were analyzed using scripts in ImageJ and MATLAB. First, the offset between 640 and 488 channels was corrected with the GridAligner plug-in. Then microtubules were selected with 7 px-wide line

selection, and line scans were generated. These line scans were imported into a MATLAB script that identified the peaks in the 488 channel and recorded the number, intensity, and FWHM of the repair sites. The FWHM for a diffraction-limited spot was obtained by using 100-nm TetraSpeck beads (Thermo Fischer Scientific). Data were exported to Prism software for graphing.

Transmission electron microscopy of microtubules repaired with recombinant tubulin

GMPCPP-microtubules at 1 μ M concentration in 1 \times BRB80 were applied to parafilm in a humidity chamber and incubated with 20 nM spastin in enzyme buffer [20 mM HEPES (pH 7.5), 300 mM KCl, 10 mM MgCl_2 , 1 mM TCEP, and 0.5 mM ATP]. Buffer containing 0.5 mM ATP- γ -S instead of ATP was used as a control. Severing was allowed to proceed for 30 s, followed by the addition of 0.6 μ M soluble FLAG-tagged single-isoform recombinant neuronal human α 1A β III tubulin to repair the microtubule lattice in the presence of 1 mM GTP and 5 mM ADP to inactivate the enzyme. The repair reaction was carried out for 5 min. Microtubules were then stabilized by the addition of 5 volumes of 0.2% glutaraldehyde in 1 \times BRB80 (80 mM PIPES, 1 mM MgCl_2 , 1 mM EGTA). After 3 min, cross-linking was quenched by the addition of Tris-HCl (pH 7.5) to a 20 mM final concentration and cross-linked microtubules were transferred into a 10-ml centrifuge tube (Beckman Coulter). The microtubule severing and healing procedure was repeated three more times, reaction mixtures were pooled into the same centrifuge tube, and microtubules were then spun down in an MLA-80 rotor at 100,000 \times g for 15 min at 30°C. The microtubule pellet was gently washed with 200 μ l of 1 \times BRB80 at 37°C twice and resuspended in 50 μ l of warm 1 \times BRB80. Five microliters of 6.7 μ M monoclonal mouse-raised anti-FLAG M2 antibody (Sigma-Aldrich) and 5 μ l of 11.45 μ M goat anti-mouse antibody conjugated to 4-nm spherical gold nanoparticles, C11-4-TGAMG-50 (Nanopartz), were added to microtubules to label repaired sites. Antibody labeling was allowed to proceed for 5 min, and the reaction was mixed with 10 volumes of 30% glycerol in 1 \times BRB80. Microtubules in 30% glycerol were loaded onto a 1 \times BRB80 cushion containing 40% glycerol and spun down onto glow-discharged carbon-coated grids (carbon film only on 400 mesh; Ted Pella) at 4200 \times g for 20 min at 30°C. Excess liquid was blotted with filter paper. Grids were washed three times with 30 μ l of BRB80, stained with 0.75% (w/v) uranyl formate, and air-dried. Images were collected on a T12 Technai electron microscope (FEI) equipped with a 2k \times 2k Gatan US1000 CCD camera. Images were collected at nominal magnifications of 6800 \times and 18,500 \times , corresponding to pixel sizes of 6.8 $\text{\AA}/\text{pix}$ and 2.5 $\text{\AA}/\text{pix}$, respectively. Images in fig. S7F were collected on a TF20 electron microscope (FEI) equipped with a K2 camera (Gatan). Images were collected at 50,000 \times and 9600 \times magnifications, corresponding to pixel sizes of 0.73 $\text{\AA}/\text{pix}$ and 3.65 $\text{\AA}/\text{pix}$, respectively.

Live imaging of severing and tubulin incorporation into nanoscale-damaged GMPCPP-microtubules and GMPCPP-capped GDP-microtubules

To observe microtubule severing and tubulin incorporation at damage sites simultaneously (Fig. 3 and fig. S5, A to C), GMPCPP-stabilized double-cycled microtubules labeled with 1% biotin and 20% HiLyte 647-tubulin were immobilized in imaging chambers. Image acquisition was started by using 100-ms continuous exposure in the 647 and 488 channels simultaneously, and the chamber was perfused with severing buffer containing 0.5 mM GTP, 20 nM spastin, and 0, 0.1, or 2 μ M HiLyte 488-labeled tubulin. Severing rates were calculated by manual counting of severing events (microtubule breaks) as a function of time. Tubulin incorporation sites were readily visible in the 488 channel. To observe the live incorporation of single tubulin dimers into microtubules damaged by spastin (Fig. 3, F and G), double-cycled GMPCPP-microtubules composed of 20% HiLyte 647-labeled and 1% biotinylated tubulin were immobilized in imaging chambers as described above. The chamber was then perfused with severing buffer, and images of microtubules were acquired. Microtubules were then incubated for 30 s with 20 nM spastin in severing buffer. Image acquisition was started during the spastin incubation step, and a solution containing fluorescently labeled tubulin [50 nM Alexa 488-labeled tubulin (PurSolutions) in BRB80 with 2-mg/ml casein, 14.3 mM 2-mercaptoethanol, 50 mM KCl, 2.5 mM $MgCl_2$, 1 mM ADP, 0.5 mM GTP, 1% Pluronic F127, and oxygen scavengers] was flushed in. Images were acquired for 5 min at 10 Hz in the 488-nm channel. After tubulin perfusion, the 640 laser was turned off to prevent photobleaching and microtubule photodamage. Images of fluorescent tubulin molecules landing on the microtubule were analyzed by using a 7×7 pixel box, and the intensity of tubulin molecules incorporated into the microtubule was calibrated against the intensity of single tubulin dimers obtained by immobilizing 0.5 nM Alexa 488-tubulin on glass with an anti- β -tubulin antibody (SAP.4G5; Sigma-Aldrich) and imaging under the same conditions.

For imaging of nonstabilized GDP-microtubules with a GMPCPP cap, sea urchin axonemes purified as described previously (67) were nonspecifically adhered to the coverslip, and 15 μ M tubulin containing 20% HiLyte 647-tubulin and 1 mM GTP were added to start microtubule growth from the axonemes. After the desired microtubule length (10 to 20 μ m) was achieved, the solution was exchanged quickly to introduce HiLyte 488-tubulin (20%) and 0.5 mM GMPCPP. After the growth of the GMPCPP cap, tubulin and nucleotide were washed out and spastin (5 nM) was introduced into the chamber with 1 mM ATP in the absence or presence of soluble tubulin at 2 μ M (500 nM HiLyte 488-tubulin + 1.5 μ M unlabeled tubulin) or 5 μ M (500 nM HiLyte 488-tubulin + 4.5 μ M unlabeled tubulin) and 0.5 mM GTP. Polymerization and imaging were performed at 30°C.

Microtubule dynamics measurements and EB1 recognition of lattice-incorporated GTP-tubulin

TIRF microscopy chambers were prepared as described above. HiLyte 647 (10%)–labeled microtubules were polymerized at 30°C at 10 μ M tubulin. The chamber was perfused with 25 nM spastin or katanin and 10 μ M porcine brain tubulin containing 10% HiLyte 647-labeled tubulin in severing assay buffer (50 mM KCl, 1% Pluronic F127, 0.2-mg/ml casein, 6.2 mM 2-mercaptoethanol, 1.5% glycerol, 0.1% methylcellulose 4000cP, and oxygen scavengers in $1 \times$ BRB80) with 1 mM GTP and 1 mM ATP together with 50 nM EB1-GFP. Images were acquired in the 647 and 488 channels simultaneously at 2 Hz. Microtubule rescues are defined as the transition of microtubules from shrinkage to growth. Rescue frequency was calculated as the number of rescues divided by the time spent depolymerizing. Catastrophes are defined as the transition of microtubules from growth to shrinkage. Catastrophe frequency was calculated as the number of catastrophes divided by the time spent in the polymerization state. The EB1 puncta and the microtubule rescue site were considered colocalized when the distance between the EB1 spot and the end of the depolymerizing microtubule was less than two pixels. The cutoff for an EB1 punctum was defined as having a mean intensity in a 5×5 pixel box that is at least 3 standard deviations above the mean background EB1 lattice intensity. Background EB1 lattice intensity was determined from control chambers without severing enzymes. Background EB1 lattice intensity was the same in the absence of severing enzymes or the presence of severing enzymes but in the absence of ATP. For statistical significance calculation, rescue site analysis was also performed by using synthetic data generated by shifting the position of the EB1 spots by 7 pixels on the microtubule (alternatively, both toward the plus and the minus ends).

For the GTP-tubulin and EB1-GFP colocalization experiments shown in Fig. 6E, microtubule extensions were grown in the absence of fluorescent tubulin for 8 min at 30°C at 12 μ M porcine brain tubulin (Cytoskeleton) in severing assay buffer. The chamber was perfused with 20 nM spastin, 50 nM EB1-GFP, and 12 μ M porcine brain tubulin containing 10% HiLyte 647-labeled tubulin in severing assay buffer. Image acquisition was started during perfusion in the 640 and 488 channels simultaneously at 5 Hz. The offset between the 640 and 488 channels was corrected by using a nanogrid (Nanogrid Miraloma Tech) and the GridAligner plug-in in ImageJ.

Laser ablation of microtubules with spastin- or katanin-generated GMPCPP islands

GMPCPP-stabilized unmodified microtubule seeds were immobilized on glass. To pregrow microtubules, 16 μ M tubulin containing 12.5% HiLyte 647-labeled tubulin with 1 mM GTP was perfused into the chamber and incubated for 10 min at 30°C. Microtubules were then capped using 6 μ M tubulin with 10% HiLyte 647 and 0.5 mM

GMPCPP. The chamber was washed after 2 min with severing assay buffer without GTP and then incubated with 4 nM spastin and 6 μ M tubulin containing 25% HiLyte 488-labeled tubulin in the presence of 200 μ M GMPCPP in severing assay buffer (50 mM KCl, 1% Pluronic F127, 0.2-mg/ml casein, 6.2 mM 2-mercaptoethanol, 2.5% glycerol, 0.1% methylcellulose 4000cP, and oxygen scavengers in $1 \times$ BRB80) with or without 1 mM ATP for 3 min. The chamber was washed with buffer containing severing assay buffer. Microtubules were ablated with a 405-nm laser at 40% power using the iLas laser illuminator (BioVision). Images in the 488 and 647 channels were acquired sequentially with 100-ms exposure. For the rescue frequency measurements, 15% HiLyte 647-labeled tubulin at 7 μ M in severing assay buffer containing 1 mM GTP was perfused into the chamber. For the katanin experiments, the chamber was washed after microtubule capping with severing assay buffer without GTP and then incubated with 20 nM katanin and 8 μ M tubulin containing 25% HiLyte 488-labeled tubulin in the presence of 200 μ M GMPCPP in severing assay buffer with or without ATP for 45 s. Microtubule depolymerization rates through the GMPCPP islands were determined by dividing the length of the island by the time it takes to depolymerize through it.

Laser ablation of dynamic microtubules with enzyme-generated GTP islands

TIRF microscopy chambers were prepared as described above. HiLyte 647-labeled microtubule extensions were polymerized for 8 min at 30°C at 12 μ M porcine brain tubulin (Cytoskeleton) containing 20% HiLyte 647-labeled tubulin in severing assay buffer. The chamber was perfused with 25 nM spastin or katanin, 50 nM EB1-GFP, and 12 μ M porcine brain tubulin containing 20% HiLyte 647-labeled tubulin in severing assay buffer with ATP or ATP- γ -S. Microtubules were ablated by using a DeltaVision OMX with the 405-nm laser at 100% power for 1 s or with a 405-nm laser at 40% power using an iLas laser illuminator (BioVision). Images were acquired in the 647 and 488 channels at 5 Hz on the DeltaVision OMX and 2.9 Hz on the iLas system.

Live imaging of tubulin incorporation and severing into dynamic microtubules

Chambers for TIRF microscopy were prepared as described above. GMPCPP-stabilized, unmodified microtubules containing 2% biotinylated tubulin were immobilized with 0.1-mg/ml NeutrAvidin (Thermo Fisher Scientific). Microtubule extensions were polymerized for 12 min at 30°C at 10 or 12 μ M porcine brain tubulin (Cytoskeleton) containing 10% HiLyte 647-tubulin in severing assay buffer (50 mM KCl, 1% Pluronic F127, 1 mM ATP, 1 mM GTP, 0.2-mg/ml casein, 6.2 mM 2-mercaptoethanol, 1.5% glycerol, 0.1% methylcellulose 4000cP, and oxygen scavengers in $1 \times$ BRB80). Then, 25 nM katanin or spastin with 12 μ M porcine brain tubulin containing 10% HiLyte 488-labeled tubulin was perfused into the chamber in severing assay buffer. Images were acquired with 488 and 640 lasers simultaneously

at 2 Hz at 100-ms exposure. The incorporation of the HiLyte 488-tubulin was immediately visible upon perfusion only at microtubule tips in the control and along the microtubules and the dynamic tips in the enzyme and ATP conditions. Total polymer mass was obtained by measuring the background-corrected total integrated fluorescence in both the 488 and 640 channels. The laser ablation controls were performed at the same enzyme and tubulin concentrations but with 1 mM ATP- γ -S. Microtubules were ablated with a 405-nm laser at 40% power using an iLas laser illuminator (BioVision) for the katanin experiments and the DeltaVision OMX for spastin.

Quantification and data analysis

n numbers and statistical tests are reported for all experiments in the figure legends. All experiments were performed multiple times, and only representative images are shown. ImageJ was used for image analysis. Prism (GraphPad) was used for graphing and statistical analysis.

REFERENCES AND NOTES

1. A. Roll-Mecak, F. J. McNally, Microtubule-severing enzymes. *Curr. Opin. Cell Biol.* **22**, 96–103 (2010). doi: [10.1016/j.cub.2009.11.001](#); pmid: [19963362](#)
2. N. T. Sherwood, Q. Sun, M. Xue, B. Zhang, K. Zinn, *Drosophila* spastin regulates synaptic microtubule networks and is required for normal motor function. *PLoS Biol.* **2**, e429 (2004). doi: [10.1371/journal.pbio.0020429](#); pmid: [15562320](#)
3. N. Trotta, G. Orso, M. G. Rossetto, A. Daga, K. Broadie, The hereditary spastic paraplegia gene, spastin, regulates microtubule stability to modulate synaptic structure and function. *Curr. Biol.* **14**, 1135–1147 (2004). doi: [10.1016/j.cub.2004.06.058](#); pmid: [15242610](#)
4. M. C. Stone et al., Normal spastin gene dosage is specifically required for axon regeneration. *Cell Rep.* **2**, 1340–1350 (2012). doi: [10.1016/j.celrep.2012.09.032](#); pmid: [23122959](#)
5. F. J. Ahmad, W. Yu, F. J. McNally, P. W. Baas, An essential role for katanin in severing microtubules in the neuron. *J. Cell Biol.* **145**, 305–315 (1999). doi: [10.1083/jcb.145.2.305](#); pmid: [10209026](#)
6. V. Stoppin-Mellet, J. Gaillard, M. Vantard, Katanin's severing activity favors bundling of cortical microtubules in plants. *Plant J.* **46**, 1009–1017 (2006). doi: [10.1111/j.1365-3113.2006.02761.x](#); pmid: [16805733](#)
7. Q. Zhang, E. Fishel, T. Bertroche, R. Dixit, Microtubule severing at crossover sites by katanin generates ordered cortical microtubule arrays in *Arabidopsis*. *Curr. Biol.* **23**, 2191–2195 (2013). doi: [10.1016/j.cub.2013.09.018](#); pmid: [24206847](#)
8. J. J. Lindeboom et al., A mechanism for reorientation of cortical microtubule arrays driven by microtubule severing. *Science* **342**, 1245533 (2013). doi: [10.1126/science.1245533](#); pmid: [24200811](#)
9. D. Zhang, G. C. Rogers, D. W. Buster, D. J. Sharp, Three microtubule severing enzymes contribute to the "Pacman-flux" machinery that moves chromosomes. *J. Cell Biol.* **175**, 881–891 (2006). doi: [10.1083/jcb.200608117](#); pmid: [17178907](#)
10. R. Loughlin, J. D. Wilbur, F. J. McNally, F. J. Nédélec, R. Heald, Katanin contributes to interspecies spindle length scaling in *Xenopus*. *Cell* **147**, 1397–1407 (2011). doi: [10.1016/j.cell.2011.11.014](#); pmid: [22153081](#)
11. K. McNally et al., Katanin maintains meiotic metaphase chromosome alignment and spindle structure in vivo and has multiple effects on microtubules in vitro. *Mol. Biol. Cell* **25**, 1037–1049 (2014). doi: [10.1091/mbc.e13-12-0764](#); pmid: [24501424](#)
12. N. Sharma et al., Katanin regulates dynamics of microtubules and biogenesis of motile cilia. *J. Cell Biol.* **178**, 1065–1079 (2007). doi: [10.1083/jcb.200704021](#); pmid: [17846175](#)
13. W. F. Hu et al., Katanin p80 regulates human cortical development by limiting centriole and cilia number. *Neuron* **84**, 1240–1257 (2014). doi: [10.1016/j.neuron.2014.12.017](#); pmid: [25521379](#)
14. K. Mishra-Gorur et al., Mutations in KATNB1 cause complex cerebral malformations by disrupting asymmetrically dividing neural progenitors. *Neuron* **84**, 1226–1239 (2014). doi: [10.1016/j.neuron.2014.12.014](#); pmid: [25521378](#)
15. J. W. Connell, C. Lindon, J. P. Luzio, E. Reid, Spastin couples microtubule severing to membrane traffic in completion of cytokinesis and secretion. *Traffic* **10**, 42–56 (2009). doi: [10.1111/j.1600-0854.2008.00847.x](#); pmid: [19000169](#)
16. J. Guizetti et al., Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* **331**, 1616–1620 (2011). doi: [10.1126/science.1201847](#); pmid: [21310966](#)
17. A. Karabay, W. Yu, J. M. Solowska, D. H. Baird, P. W. Baas, Axonal growth is sensitive to the levels of katanin, a protein that severs microtubules. *J. Neurosci.* **24**, 5778–5788 (2004). doi: [10.1523/JNEUROSCI.1382-04.2004](#); pmid: [15215300](#)
18. R. A. Charafeddine et al., Fidgetin-Like 2: A microtubule-based regulator of wound healing. *J. Invest. Dermatol.* **135**, 2309–2318 (2015). doi: [10.1038/jid.2015.94](#); pmid: [25756798](#)
19. G. Yigit et al., A syndrome of microcephaly, short stature, polysyndactyly, and dental anomalies caused by a homozygous KATNB1 mutation. *Am. J. Med. Genet. A* **170**, 728–733 (2016). doi: [10.1002/ajmg.a.37484](#); pmid: [26640080](#)
20. A. Roll-Mecak, R. D. Vale, The *Drosophila* homologue of the hereditary spastic paraplegia protein, spastin, severs and disassembles microtubules. *Curr. Biol.* **15**, 650–655 (2005). doi: [10.1016/j.cub.2005.02.029](#); pmid: [15823537](#)
21. K. Jiang et al., Microtubule minus-end regulation at spindle poles by an ASPM-katanin complex. *Nat. Cell Biol.* **19**, 480–492 (2017). doi: [10.1038/ncb3511](#); pmid: [28436967](#)
22. J. D. Wood et al., The microtubule-severing protein Spastin is essential for axon outgrowth in the zebrafish embryo. *Hum. Mol. Genet.* **15**, 2763–2771 (2006). doi: [10.1093/hmg/ddl212](#); pmid: [16893913](#)
23. D. H. Burk, Z.-H. Ye, Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein. *Plant Cell* **14**, 2145–2160 (2002). doi: [10.1105/tpc.003947](#); pmid: [12215512](#)
24. M. Srayko, E. T. O'toole, A. A. Hyman, T. Müller-Reichert, Katanin disrupts the microtubule lattice and increases polymer number in *C. elegans* meiosis. *Curr. Biol.* **16**, 1944–1949 (2006). doi: [10.1016/j.cub.2006.08.029](#); pmid: [17027492](#)
25. K. Ribbeck, T. J. Mitchison, Meiotic spindle: Sculpted by severing. *Curr. Biol.* **16**, R923–R925 (2006). doi: [10.1016/j.cub.2006.09.048](#); pmid: [17084690](#)
26. J. A. Roll-Mecak, R. D. Vale, Making more microtubules by severing: A common theme of noncentrosomal microtubule arrays? *J. Cell Biol.* **175**, 849–851 (2006). doi: [10.1083/jcb.200611449](#); pmid: [17178905](#)
27. R. A. Walker, S. Inoué, E. D. Salmon, Asymmetric behavior of severed microtubule ends after ultraviolet-microbeam irradiation of individual microtubules in vitro. *J. Cell Biol.* **108**, 931–937 (1989). doi: [10.1083/jcb.108.3.931](#); pmid: [2921286](#)
28. P. T. Tran, R. A. Walker, E. D. Salmon, A metastable intermediate state of microtubule dynamic instability that differs significantly between plus and minus ends. *J. Cell Biol.* **138**, 105–117 (1997). doi: [10.1083/jcb.138.1.105](#); pmid: [9214385](#)
29. T. Mitchison, M. Kirschner, Dynamic instability of microtubule growth. *Nature* **312**, 237–242 (1984). doi: [10.1038/312237a0](#); pmid: [6504138](#)
30. M. F. Carlier, D. Pantaloni, Kinetic analysis of guanosine 5'-triphosphate hydrolysis associated with tubulin polymerization. *Biochemistry* **20**, 1918–1924 (1981). doi: [10.1021/bi00510a030](#); pmid: [7225365](#)
31. D. Zhang et al., *Drosophila* katanin is a microtubule depolymerase that regulates cortical-microtubule plus-end interactions and cell migration. *Nat. Cell Biol.* **13**, 361–370 (2011). doi: [10.1038/ncb2206](#); pmid: [21378981](#)
32. L. Schaedel et al., Microtubules self-repair in response to mechanical stress. *Nat. Mater.* **14**, 1156–1163 (2015). doi: [10.1038/nmat4396](#); pmid: [26343914](#)
33. C. Aumeier et al., Self-repair promotes microtubule rescue. *Nat. Cell Biol.* **18**, 1054–1064 (2016). doi: [10.1038/ncb3406](#); pmid: [27617929](#)
34. M. E. Bailey, D. L. Sackett, J. L. Ross, Katanin severing and binding microtubules are inhibited by tubulin carboxy tails. *Biophys. J.* **109**, 2546–2561 (2015). doi: [10.1016/j.bpj.2015.11.011](#); pmid: [26682813](#)
35. G. Hiller, K. Weber, Radioimmunoassay for tubulin: A quantitative comparison of the tubulin content of different established tissue culture cells and tissues. *Cell* **14**, 795–804 (1978). doi: [10.1016/0092-8674\(78\)90335-5](#); pmid: [688394](#)
36. D. N. Itzhak, S. Tyanova, J. Cox, G. H. Bomer, Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* **5**, e16950 (2016). doi: [10.7554/eLife.16950](#); pmid: [27278775](#)
37. A. Vemu et al., Structure and dynamics of single-isoform recombinant neuronal human tubulin. *J. Biol. Chem.* **291**, 12907–12915 (2016). doi: [10.1074/jbc.C116.731133](#); pmid: [27129203](#)
38. M.-F. Carlier, Guanosine-5'-triphosphate hydrolysis and tubulin polymerization. *Mol. Cell. Biochem.* **47**, 97–113 (1982). doi: [10.1007/BF00234410](#); pmid: [6755216](#)
39. M. F. Carlier, D. Pantaloni, Assembly of microtubule protein: Role of guanosine di- and triphosphate nucleotides. *Biochemistry* **21**, 1215–1224 (1982). doi: [10.1021/bi00535a017](#); pmid: [7074077](#)
40. M. F. Carlier, D. Didry, D. Pantaloni, Microtubule elongation and guanosine 5'-triphosphate hydrolysis. Role of guanine nucleotides in microtubule dynamics. *Biochemistry* **26**, 4428–4437 (1987). doi: [10.1021/bi00388a036](#); pmid: [3663597](#)
41. V. VanBuren, D. J. Odde, L. Cassimeris, Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6035–6040 (2002). doi: [10.1073/pnas.092504999](#); pmid: [11983898](#)
42. A. Dimitrov et al., Detection of GTP-tubulin conformation in vivo reveals a role for GTP remnants in microtubule rescues. *Science* **322**, 1353–1356 (2008). doi: [10.1126/science.1165401](#); pmid: [18927356](#)
43. J. Al-Bassam et al., CLASP promotes microtubule rescue by recruiting tubulin dimers to the microtubule. *Dev. Cell* **19**, 245–258 (2010). doi: [10.1016/j.devcel.2010.07.016](#); pmid: [20708587](#)
44. K. Ichihara, H. Kitazawa, Y. Iguchi, H. Hotani, T. J. Itoh, Visualization of the stop of microtubule depolymerization that occurs at the high-density region of microtubule-associated protein 2 (MAP2). *J. Mol. Biol.* **312**, 107–118 (2001). doi: [10.1006/jmbi.2001.4934](#); pmid: [11545589](#)
45. M. Zanic, J. H. Stear, A. A. Hyman, J. Howard, EB1 recognizes the nucleotide state of tubulin in the microtubule lattice. *PLoS ONE* **4**, e7585 (2009). doi: [10.1371/journal.pone.0007585](#); pmid: [19851462](#)
46. S. P. Maurer, F. J. Fourniol, G. Bohner, C. A. Moores, T. Surrey, EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell* **149**, 371–382 (2012). doi: [10.1016/j.cell.2012.02.049](#); pmid: [22500803](#)
47. T. P. Spurrck et al., UV microbeam irradiations of the mitotic spindle. II. Spindle fiber dynamics and force production. *J. Cell Biol.* **111**, 1505–1518 (1990). doi: [10.1083/jcb.111.4.1505](#); pmid: [211823](#)
48. J. Brugués, V. Nuzzo, E. Mazur, D. J. Needleman, Nucleation and transport organize microtubules in metaphase spindles. *Cell* **149**, 554–564 (2012). doi: [10.1016/j.cell.2012.03.027](#); pmid: [22541427](#)
49. R. B. Nicklas, G. M. Lee, C. L. Rieder, G. Rupp, Mechanically cut mitotic spindles: Clean cuts and stable microtubules. *J. Cell Sci.* **94**, 415–423 (1989). pmid: [2698889](#)
50. J. S. Tirnauer, E. D. Salmon, T. J. Mitchison, Microtubule plus-end dynamics in *Xenopus* egg extract spindles. *Mol. Biol. Cell* **15**, 1776–1784 (2004). doi: [10.1091/mbc.e03-11-0824](#); pmid: [14767058](#)
51. J. Howard, A. A. Hyman, Microtubule polymerases and depolymerases. *Curr. Opin. Cell Biol.* **19**, 31–35 (2007). doi: [10.1016/j.cub.2006.12.009](#); pmid: [1718486](#)
52. R. D. Vale, Severing of stable microtubules by a mitotically activated protein in *Xenopus* egg extracts. *Cell* **64**, 827–839 (1991). doi: [10.1016/0092-8674\(91\)90511-V](#); pmid: [1671762](#)
53. M. E. Bailey, M. M. Morelli, J. D. Diaz, J. L. Ross, Katanin P60 targets microtubules with defects. *Biophys. J.* **102**, 701a (2012). doi: [10.1016/j.bpj.2011.11.3806](#)
54. H. de Forges et al., Localized mechanical stress promotes microtubule rescue. *Curr. Biol.* **26**, 3399–3406 (2016). doi: [10.1016/j.cub.2016.10.048](#); pmid: [27916523](#)
55. C. Wang et al., KTN80 confers precision to microtubule severing by specific targeting of katanin complexes in plant cells. *EMBO J.* **36**, 3435–3447 (2017). doi: [10.1525/emboj.201796823](#); pmid: [28978669](#)
56. M. L. Valenstein, A. Roll-Mecak, Graded control of microtubule severing by tubulin glutamylation. *Cell* **164**, 911–921 (2016). doi: [10.1016/j.cell.2016.01.019](#); pmid: [26875866](#)
57. B. Lacroix et al., Tubulin polyglutamylation stimulates spastin-mediated microtubule severing. *J. Cell Biol.* **189**, 945–954 (2010). doi: [10.1083/jcb.201001024](#); pmid: [20530212](#)
58. T. Nakata, S. Niwa, Y. Okada, F. Perez, N. Hirokawa, Preferential binding of a kinesin-1 motor to GTP-tubulin-rich

- microtubules underlies polarized vesicle transport. *J. Cell Biol.* **194**, 245–255 (2011). doi: [10.1083/jcb.201104034](https://doi.org/10.1083/jcb.201104034); pmid: [21768290](https://pubmed.ncbi.nlm.nih.gov/21768290/)
60. A. Y. Chan, M. Bailly, N. Zebda, J. E. Segall, J. S. Condeelis, Role of cofilin in epidermal growth factor-stimulated actin polymerization and lamellipod protrusion. *J. Cell Biol.* **148**, 531–542 (2000). doi: [10.1083/jcb.148.3.531](https://doi.org/10.1083/jcb.148.3.531); pmid: [10662778](https://pubmed.ncbi.nlm.nih.gov/10662778/)
 61. V. DesMarais, M. Ghosh, R. Eddy, J. Condeelis, Cofilin takes the lead. *J. Cell Sci.* **118**, 19–26 (2005). doi: [10.1242/jcs.01631](https://doi.org/10.1242/jcs.01631); pmid: [15615780](https://pubmed.ncbi.nlm.nih.gov/15615780/)
 62. N. E. Ziółkowska, A. Roll-Mecak, In vitro microtubule severing assays. *Methods Mol. Biol.* **1046**, 323–334 (2013). doi: [10.1007/978-1-62703-538-5_19](https://doi.org/10.1007/978-1-62703-538-5_19); pmid: [23868597](https://pubmed.ncbi.nlm.nih.gov/23868597/)
 63. E. Zehr *et al.*, Katanin spiral and ring structures shed light on power stroke for microtubule severing. *Nat. Struct. Mol. Biol.* **24**, 717–725 (2017). doi: [10.1038/nsmb.3448](https://doi.org/10.1038/nsmb.3448); pmid: [28783150](https://pubmed.ncbi.nlm.nih.gov/28783150/)
 64. A. J. Albee, C. Wiese, Xenopus TACC3/maskin is not required for microtubule stability but is required for anchoring microtubules at the centrosome. *Mol. Biol. Cell* **19**, 3347–3356 (2008). doi: [10.1091/mbc.e07-11-1204](https://doi.org/10.1091/mbc.e07-11-1204); pmid: [18508920](https://pubmed.ncbi.nlm.nih.gov/18508920/)
 65. C. Gell, C. T. Friel, B. Borghonovo, D. N. Drechsel, A. A. Hyman, J. Howard, “Purification of tubulin from porcine brain,” in *Microtubule Dynamics: Methods and Protocols*, A. Straube, Ed. (Springer, 2011), pp. 15–28.
 66. A. D. Edelstein *et al.*, Advanced methods of microscope control using µManager software. *J. Biol. Methods* **1**, e10 (2014). doi: [10.14440/jbm.2014.36](https://doi.org/10.14440/jbm.2014.36); pmid: [25606571](https://pubmed.ncbi.nlm.nih.gov/25606571/)
 67. I. R. Gibbons, E. Fronk, A latent adenosine triphosphatase form of dynein 1 from sea urchin sperm flagella. *J. Biol. Chem.* **254**, 187–196 (1979). pmid: [214440](https://pubmed.ncbi.nlm.nih.gov/214440/)

ACKNOWLEDGMENTS

We thank A. Szyk, National Institute of Neurological Disorders and Stroke (NINDS), for purified spastin and katanin, and X. Wu, National Heart, Lung, and Blood Institute (NHLBI), for help in the Light Microscopy Core. **Funding:** N.G. is a Howard Hughes Medical Institute investigator. A.M.D. is supported by NIH grant R01GM121975. A.R.-M. is supported by the intramural programs of NINDS and NHLBI. **Author contributions:** A.V. performed TIRF experiments with dynamic microtubules and laser ablation. E.S. performed TIRF experiments with stabilized and dynamic microtubules. A.V., E.S., and J.O.S. performed image analysis. E.A.Z. performed EM time

courses with spastin and all EM experiments with katanin and recombinant tubulin. A.M.D. performed initial in-the-tube and on-grid EM experiments with spastin. A.M.D. and N.G. discussed EM experiments. A.R.-M. wrote the manuscript with input from A.M.D., E.S., A.V., and E.A.Z. All authors planned experiments and reviewed the manuscript. A.R.-M. conceived of the project, conceptualized the manuscript, and supervised research. **Competing interests:** The authors declare no competing financial interests. **Data and material availability:** All data needed to understand and assess the conclusions of this research are available in the main text and supplementary materials. Information requests and requests for reagents should be directed to the corresponding author.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/eaau1504/suppl/DC1
Figs. S1 to S10
Movies S1 to S7

11 May 2018; accepted 18 July 2018
10.1126/science.aau1504

RESEARCH ARTICLE

PROTEOMICS

Co-regulatory networks of human serum proteins link genetics to disease

Valur Emilsson^{1,2*†}, Marjan Ilkov^{1*}, John R. Lamb^{3*†}, Nancy Finkel⁴, Elias F. Gudmundsson¹, Rebecca Pitts⁴, Heather Hoover⁴, Valborg Gudmundsdottir¹, Shane R. Horman³, Thor Aspelund^{1,5}, Le Shu⁶, Vladimir Trifonov³, Sigurdur Sigurdsson¹, Andrei Manolescu⁷, Jun Zhu⁸, Örn Olafsson¹, Johanna Jakobsdottir¹, Scott A. Lesley³, Jeremy To³, Jia Zhang³, Tamara B. Harris⁹, Lenore J. Launer⁹, Bin Zhang⁸, Gudny Eiriksdottir¹, Xia Yang⁶, Anthony P. Orth³, Lori L. Jennings^{4†}, Vilundur Gudnason^{1,10††}

Proteins circulating in the blood are critical for age-related disease processes; however, the serum proteome has remained largely unexplored. To this end, 4137 proteins covering most predicted extracellular proteins were measured in the serum of 5457 Icelanders over 65 years of age. Pairwise correlation between proteins as they varied across individuals revealed 27 different network modules of serum proteins, many of which were associated with cardiovascular and metabolic disease states, as well as overall survival. The protein modules were controlled by cis- and trans-acting genetic variants, which in many cases were also associated with complex disease. This revealed co-regulated groups of circulating proteins that incorporated regulatory control between tissues and demonstrated close relationships to past, current, and future disease states.

Human serum contains a dynamic flux of proteins synthesized by tissues and cells of the body (1). The secretome is complex and likely involves 15% or more of all proteins (2). Secreted proteins and circulating blood cells mediate global homeostasis via intercellular communication, immune responses, vascular and endothelial cell function, tissue remodeling, fluid exchange, and nutrient assimilation (3). Defined functional roles for many individual proteins in circulation remains to be ascribed owing to our limited ability to monitor their production, accumulation, and distribution in both model systems and humans.

Heterochronic parabiosis experiments that surgically joined the circulation of young and old mice showed a system-wide effect on the re-

generative capacity of organs (4, 5). Thus, serum proteins and other circulating factors may directly regulate complex processes such as aging and the development of common chronic diseases. In contrast to monogenic diseases, complex diseases are caused not by proteins acting alone but instead by highly interacting protein networks that may result from genetic and environmental perturbations and ultimately drive physiological states toward disease (6–9). Because blood mediates coordination between nonadjacent tissues, it is of the highest interest to understand if and how this regulation occurs via serum proteins and their networks.

A custom-designed aptamer-based multiplex proteomic platform

To date, high-throughput detection and quantification of serum proteins in a large human population have been hampered by the limitations of proteomic profiling technologies. The Slow-Off rate Modified Aptamer (SOMAmer)-based technology has emerged as a proteomic profiling platform (SOMAscan) with high sample throughput and sensitivity of detection (10, 11). We designed an expanded custom version of this platform to include proteins known or predicted to be found in the extracellular milieu, including the predicted extracellular domains of single- and certain multipass transmembrane proteins. This resulted in an updated array of 5034 SOMAmers, 4783 of which recognize 4137 individual human proteins (table S1)—i.e., some proteins are targeted by more than one aptamer, whereas the rest recognize nonhuman targets. We applied the platform to a large population-based sampling of

5457 participants in the AGES Reykjavik study (12), a prospective study of deeply phenotyped and genotyped individuals older than 65 years of age. Table S2 reports baseline characteristics of the study population, while fig. S1 shows the workflow of the present study. For direct and inferential measures of aptamer specificity, see tables S3 to S6 and figs. S2 and S3.

Identification and characterization of serum protein networks

Biological networks are characterized by a non-random distribution of links between objects that are scale-free in nature (13). We reconstructed the protein co-regulation network using weighted gene-to-gene coexpression analysis (14). Co-regulation networks describe functional relationships that can reflect both physical and nonphysical interactions between objects, including proteins. The parameters used for constructing the serum protein network were chosen on the basis of the topological scale-free criterion (14) and tuned to enhance signal-to-noise ratio in the protein adjacency matrix (Fig. 1A and fig. S4, A and B). This analysis established that serum proteins cluster into 27 highly structured co-regulatory modules ranging in size between 20 and 921 proteins (table S7), whereas 15% fell outside of these modules. We examined the preservation of the network architecture by randomly splitting the AGES cohort into training and test sets, and applying a suite of statistics as previously described (15). We observed strong preservation of the overall structure of the network (Fig. 1B). Permutation testing of the data indicated that these modules were unlikely to have occurred by chance (fig. S4C). Pairwise correlation of proteins could be controlled at any of the many steps from transcription to secretion and clearance, and here we use the term “co-regulation” to encompass all such regulation between proteins.

We applied various annotation tools and found that the modules were enriched for distinct functional and tissue-specific signatures (table S8). This suggests that, to some degree, peripheral proteins cluster according to their function and/or tissue of origin. Individual modules frequently contained proteins from many distinct tissues, indicating that individual tissues were not the sole contributor to any module (tables S9 and S10). Indeed, comparison of the serum protein modules to 2672 coexpression mRNA modules constructed from solid tissues (16) revealed some, but largely an insignificant, agreement between the two (fig. S5), suggesting that serum protein co-regulation is distinct from that of most tissues. Thus, the human serum proteome appeared as functionally distinct modules of proteins produced by many tissues of the body.

We characterized each module's eigenprotein [$E^{(q)}$, where q denotes a module] through a singular value decomposition and transformation of the variable protein levels for any given module. Each $E^{(q)}$ is a unique representation that most closely reflects the collective behavior of that module, explaining on average 40% of the protein

¹Icelandic Heart Association, Holtasmar 1, IS-201 Kopavogur, Iceland. ²Faculty of Pharmacology, University of Iceland, 101 Reykjavik, Iceland. ³Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121, USA. ⁴Novartis Institutes for Biomedical Research, 22 Windsor Street, Cambridge, MA 02139, USA. ⁵Centre of Public Health Sciences, University of Iceland, 101 Reykjavik, Iceland. ⁶Department of Integrative Biology and Physiology, University of California, Los Angeles CA, USA. ⁷School of Science and Engineering, Mentavegur 1, IS-101, Reykjavik University, 101 Reykjavik, Iceland. ⁸Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁹Laboratory of Epidemiology and Population Sciences, Intramural Research Program, National Institute on Aging, Bethesda, MD 20892-9205, USA. ¹⁰Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland.

*These authors contributed equally to this work.

†Corresponding author. Email: valur@hjarta.is (V.E.); v.gudnason@hjarta.is (V.G.); jlamb@gnf.org (J.R.L.)

††These authors contributed equally to this work.

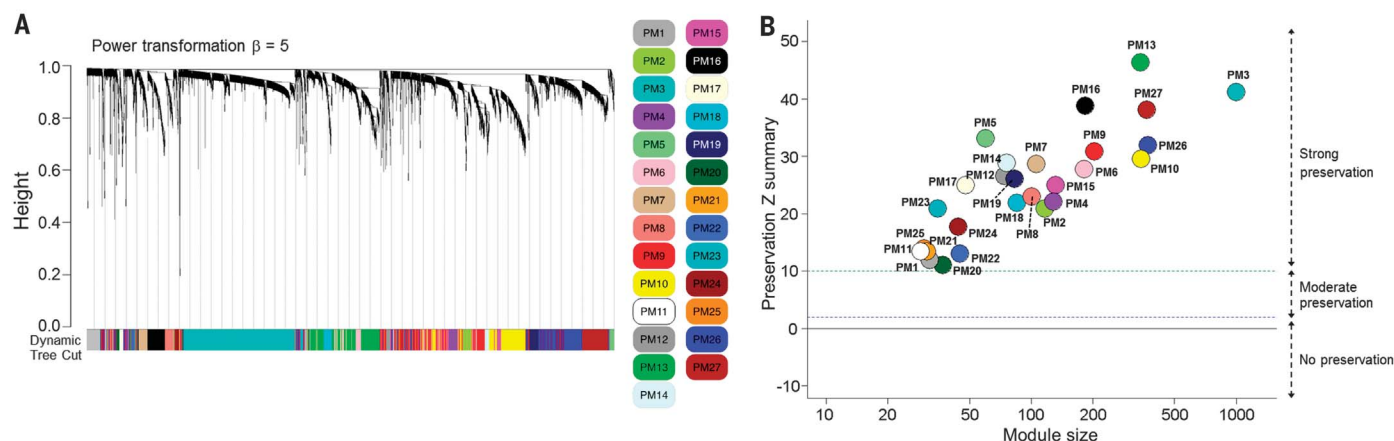
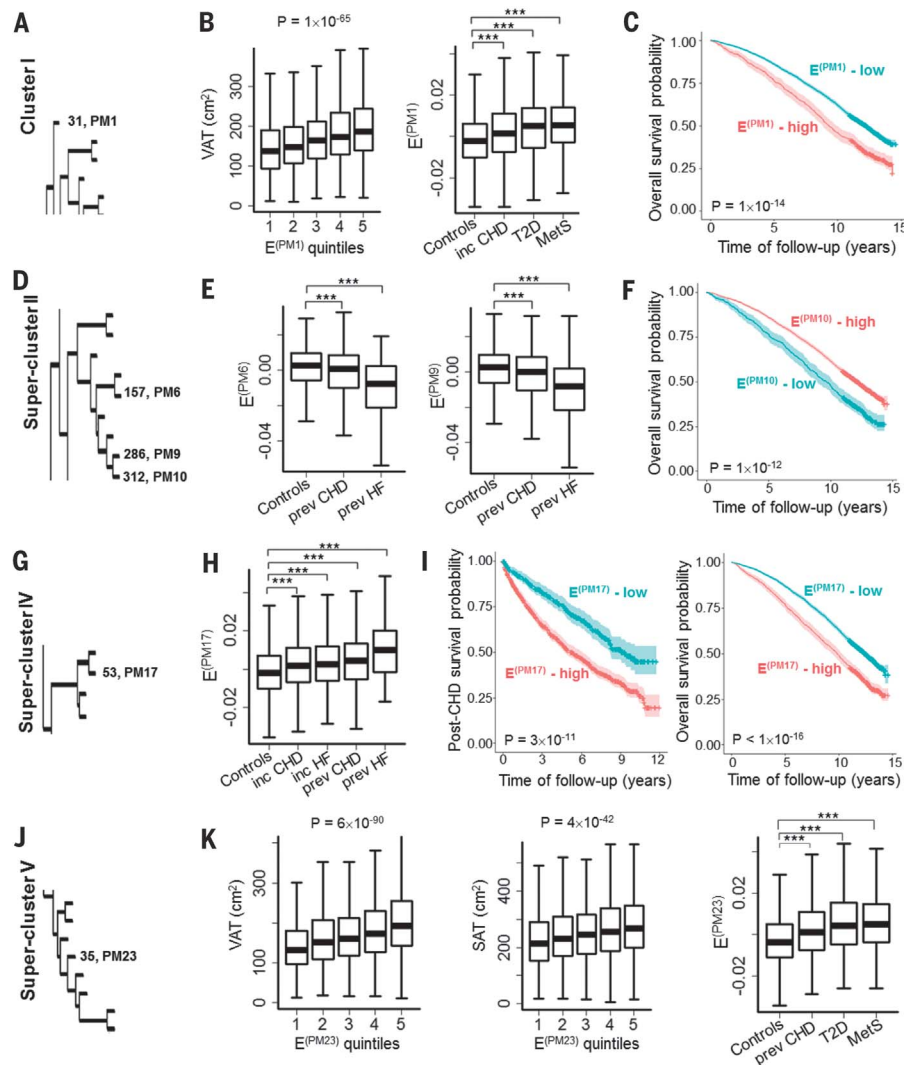


Fig. 1. The serum protein network structure. (A) Hierarchical clustering dendrogram using dynamic tree cut (29), revealing 27 serum protein modules. Each branch of the dendrogram represents a single protein, and the colored bar below denotes its corresponding protein module, as annotated in the legend to the right. The dendrogram height is the distance between proteins (14). (B) The cohort was randomly split into two equal parts, one for a training set and another for the test

set, and a summary Z score statistics (15), plotted for each module presented as colored data points. The summary Z score <2 (blue dotted line) indicates no preservation; 2 < summary Z score <10 (between the blue and green dotted lines) indicates moderate evidence of preservation; and a summary Z score >10 (green dotted line) indicates strong evidence of preservation. See Fig. S10 for the $\frac{2}{3}$ versus $\frac{1}{3}$ split of the cohort.

Fig. 2. The relationship between module's $E^{(n)}$ to disease-related measures. (A) The module PM1 is a single cluster of 31 proteins.

(B) Positive associations of $E^{(PM1)}$ quintiles to variation (cm^2) in visceral adipose tissue (VAT), incident coronary heart disease (inc CHD), type 2 diabetes (T2D), and the metabolic syndrome (MetS), $***P < 1 \times 10^{-10}$. (C) Overall survival, i.e., with respect to all-cause mortality, was reduced for high $E^{(PM1)}$ levels (red curve) compared to low $E^{(PM1)}$ levels (cyan curve). (D) The modules PM6, PM9, and PM10 are members of supercluster II. (E) Inverse association of the $E^{(PM6)}$ and $E^{(PM9)}$ to prevalent CHD (prev CHD) and prevalent heart failure (prev HF), $***P \leq 1 \times 10^{-9}$. (F) Reduced overall survival for low $E^{(PM10)}$ levels (cyan curve) compared to high $E^{(PM10)}$ levels (red curve). (G) The PM17 is in supercluster IV. (H) Positive association of module's $E^{(PM17)}$ to incident CHD and HF as well as prevalent CHD and HF, $***P < 1 \times 10^{-17}$. (I) Reduced postincident CHD survival as well as overall survival for high $E^{(PM17)}$ levels (red curve) compared to low $E^{(PM17)}$ levels (cyan curve). (J) The module PM23 is a member of supercluster V. (K) Positive associations of the module $E^{(PM23)}$ quintiles to VAT and subcutaneous adipose tissue (SAT), and prevalent CHD, T2D, and MetS, $***P < 1 \times 10^{-13}$. Data were analyzed using forward linear or logistic regression or Cox proportional hazards regression, depending on the outcome being continuous, binary, or a time to an event. Kaplan-Meier plots were used to display survival probabilities. The number of proteins per module is denoted at the branches of the dendrogram. Controls are individuals free of the disease in question.



covariance. Some modules' $E^{(q)}$ s were more correlated to others and formed several superclusters of modules with shared functional categories (table S11 and fig. S6). We assessed if the modules were related to disease status of the AGES donors. Association of module $E^{(q)}$ s to various outcomes—

including coronary heart disease (CHD), heart failure (HF), type 2 diabetes (T2D), visceral and subcutaneous adipose tissue (VAT and SAT, respectively), and metabolic syndrome (MetS)—were found (table S12). Whereas some modules showed no association to any disease measure,

others were associated to some or all of them (table S12; Fig. 2, A to K; and fig. S7, A to I). Notably, superclusters, although internally consistent, often showed opposing relationships between superclusters to CHD-related outcomes (table S12; Fig. 2, E, H, and K; and fig. S7, B, E, and H), which may reflect differential roles of the protein modules in disease. The two modules PM1 and PM23 associated with T2D were also associated with MetS, VAT-, SAT-, and CHD-related outcomes in a directionally consistent manner (table S12 and Fig. 2, B and K). Finally, modules associated with incident disease and to all-cause or post-CHD mortality (table S12; Fig. 2, C, F, and I; and fig. S7, C, F, and I) indicate that the protein network predicted future events and disease outcome. Time between diagnosis and sample collection had no effect on the association of individual proteins to prevalent disease (fig. S8). Thus, individual disease associations of the protein modules link the shared functions of many proteins to common diseases.

The networks were scale-free in nature where a few protein nodes were highly connected. These are often referred to as network hubs, which organize network connectivity and information flow (17). Studies of gene networks from solid tissues suggest that hub nodes tend to be essential and evolutionarily conserved (18). The scaled intramodule connectivity ($K_i = \frac{k_i}{k_{max}}$) of proteins was compared to various outcomes, where protein hubs showed stronger association to phenotypic measures than less well connected proteins (Fig. 3, A to L, and fig. S9, A to D). We observed strong reproducibility of the hub status of proteins through the network preservation analysis (fig. S10). In summary, the hub proteins consistently showed stronger association to various disease-related outcomes compared to proteins located in the periphery of the network. Thus, the structure of the protein network derived solely from measured protein variation was aligned with disease and phenotype variation across donors.

Genetic variants influence serum protein levels

Genome-wide association studies (GWAS) have identified thousands of common DNA sequence variants affecting human diseases (19). By integrating GWAS signals with genetic variants that affect intermediate traits like mRNA and/or protein, the identification of the causal candidates and pathways can be enhanced (6–9). Protein single-nucleotide polymorphisms (pSNPs) are DNA sequence variants associated with allelic imbalance in protein levels. Using Bonferroni correction, we identified 1046 significant cis pSNP–protein associations within a 300-kb window across the corresponding protein-coding sequences (table S13 and fig. S11A). Cross-referencing cis pSNP–protein pairs to cis eSNPs (expression SNPs)—transcript data from more than 30 tissues and cells (20) revealed an overlap of 37.3% (table S14). This suggests that ~60% of the genetic effects on serum protein levels are mediated either by an as yet unknown transcriptional effect and/or by

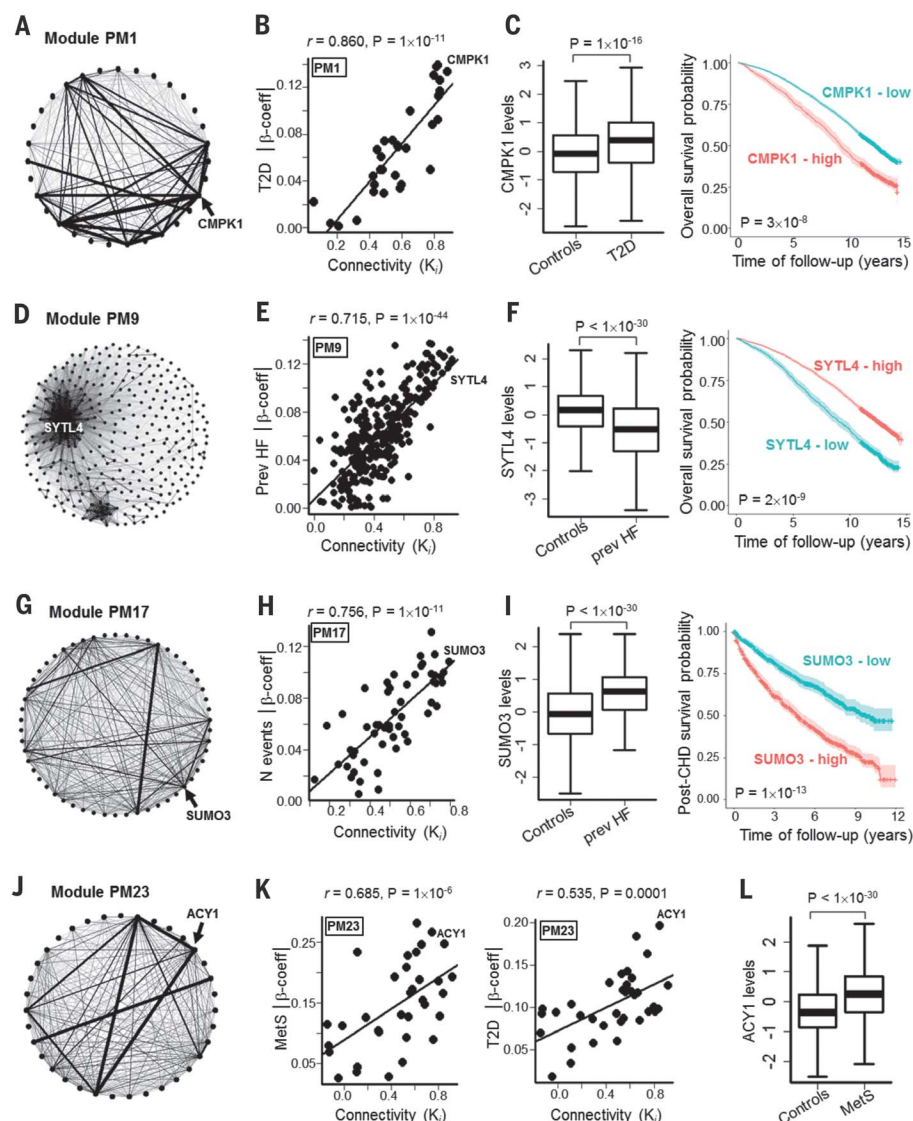


Fig. 3. The relationship between connectivity of proteins and disease-related measures.

(A) Circle graph of PM1 highlighting the hub protein CMPK1. (B) Positive correlation between within module connectivity (K_i) (x axis) and the absolute value of the effect size of the association of proteins to type 2 diabetes (T2D) (y axis). (C) Positive association of CMPK1 to T2D, and reduced overall survival associated with high serum CMPK1 levels (red curve). (D) Spring graph of PM9, highlighting the hub SYTL4. (E) Positive correlation between K_i and the association of proteins to prevalent heart failure (prev HF). (F) Inverse association of SYTL4 to HF, $P < 1 \times 10^{-30}$, and reduced overall survival associated with low serum SYTL4 levels (cyan curve). (G) A circle graph of PM17 highlighting the hub protein SUMO3. (H) Positive correlation between the K_i of proteins and their association to prevalent HF. (I) SUMO3 is positively associated with prevalent HF, and high levels of SUMO3 (red curve) predict reduced survival postincident CHD. (J) A circle graph of the PM23 highlighting the hub ACY1. (K) Positive correlation between the K_i of proteins and their association to both MetS and T2D. (L) Strong positive association of ACY1 to MetS. Network visualization was performed with the igraph package in R (30). Pearson's r was estimated for correlation between K_i of proteins and their strength of association to disease measures. See Fig. 2 for other relevant statistics.

posttranscriptional mechanisms. The cis pSNP-proteins were underrepresented among highly connected protein nodes (fig. S11, B and C), consistent with the observation that there was reduced selective pressure on non-hub proteins (27). We observed examples of cis protein effects likely underlying reported GWAS risk loci for T2D, adiposity, and/or CHD (fig. S12A and tables S15 and S16).

We also found that various lead GWAS SNPs mediate distal trans effects on serum proteins (figs. S12, B to D, and S13A to E, and table S16). Many of these risk loci influenced proteins in both cis and trans, including the risk locus rs579459 for CHD upstream of *ABO* (fig. S12D and table S16), of which many were overrepresented in the PM27 module (fig. S12E). Genetic pleiotropy is a well-established phenomenon at the *ABO* locus (fig. S13G) (22). Indeed, the *ABO* blood groups have been linked to many diseases, which is in part mediated through their effect on the prothrombotic risk factor vWF (von Willebrand factor) (23, 24), a protein also affected by rs579459 in trans (table S16). In light of these

results, all cis-acting pSNPs were tested directly for trans effects on proteins, revealing that 16% of all cis pSNPs affected in total 911 proteins in trans (table S17). Notably, twice as many, or 40.7% of cis pSNPs affecting proteins in trans, matched reported GWAS loci, compared to 20.7% among all cis pSNPs (tables S15 and S17), suggesting a link between trans protein regulation and disease variation. Finally, we confirmed on average 80% of previously reported cis and 74% of trans effects on plasma proteins in our dataset (table S18 and fig. S14). The finding that GWAS risk loci tend to regulate many proteins in trans and that these proteins cluster in the same module raised the possibility that common DNA sequence variants determine the architecture of the serum protein network.

Linking genetic variants to serum protein networks

Concerns have been expressed regarding the adequacy of GWAS to identify useful disease links, resulting in a call for greater efforts in identifying

the gene regulatory networks that integrate the numerous GWAS signals (25). Given the hierarchical organization of the serum protein network and our observation that trans co-regulated proteins cluster in modules, it is possible that common cis- and trans-acting pSNPs underlie the structure of the serum protein network. We adopted a standard single-point GWAS test of linear regression against all module $E^{(q)}$ s using a threshold of $P \leq 5 \times 10^{-8}$ for detection of genome-wide significant effects. Even though the $E^{(q)}$ represented many proteins as a complex trait, we often observed strong single SNP associations to the modules $E^{(q)}$ s (table S19), which we have termed network-associated protein SNPs (npSNPs), of which many were associated with more than one module consistent with their relationship. Given that npSNPs were associated with $E^{(q)}$ s of distinct protein modules, npSNPs may act via individual proteins in circulation (tables S19 and S20). For instance, the npSNP rs704 (NP_000629.3: p.Thr400Met) in *VTN* affected VTN in cis (Fig. 4, A to C, and fig. S13H), and 698 proteins in trans that were significant components of the modules enriched for immune-related pathways (Fig. 4, D and E, and table S20). Furthermore, distinct loci at *APOE* or *BCHE* were associated with the lipoprotein-enriched module PM11 (Fig. 5A, fig. S12I, and tables S8, S19, and S20), exerting cis effects on *APOE* and *BCHE* (Fig. 5B) and affecting 64 proteins in trans or 89% of all proteins in PM11 (Fig. 5, C and D, and tables S19 and S20). These results highlight the genetic architecture of the serum protein network and show that the modules and disease variation are intimately connected.

The discoveries of npSNPs and cis-trans protein pairs allowed us to assess whether the serum protein networks resulted from cross-tissue regulatory control by comparing them to gene expression data from 53 different human tissues and involving only the top tissue-specific proteins (table S21 and S22). Many cis-acting pSNPs affected serum levels of tissue-specific proteins and subsequently affected variable serum levels of other proteins synthesized in distinct tissues (table S21 and S22). These data suggest that the serum protein network arose at least in part via systemic cross-tissue regulation.

Discussion

Deep protein profiling in a large well-characterized population of the elderly revealed the higher-order topology and modularity of the human serum proteome. The study cohort used was population-based and thus contained nondiseased donors and donors suffering from various diseases and complications (12). As such, the relationship between proteins is the consensus of the whole population. Recent studies have revealed a pronounced effect of aging on transcriptional changes that are both tissue and organismal specific (26–28). It is possible that the observed network structure of the serum proteins reflects to some extent those changes associated with the old age of the population and may indeed reflect a certain set-point of homeostasis or even a decline

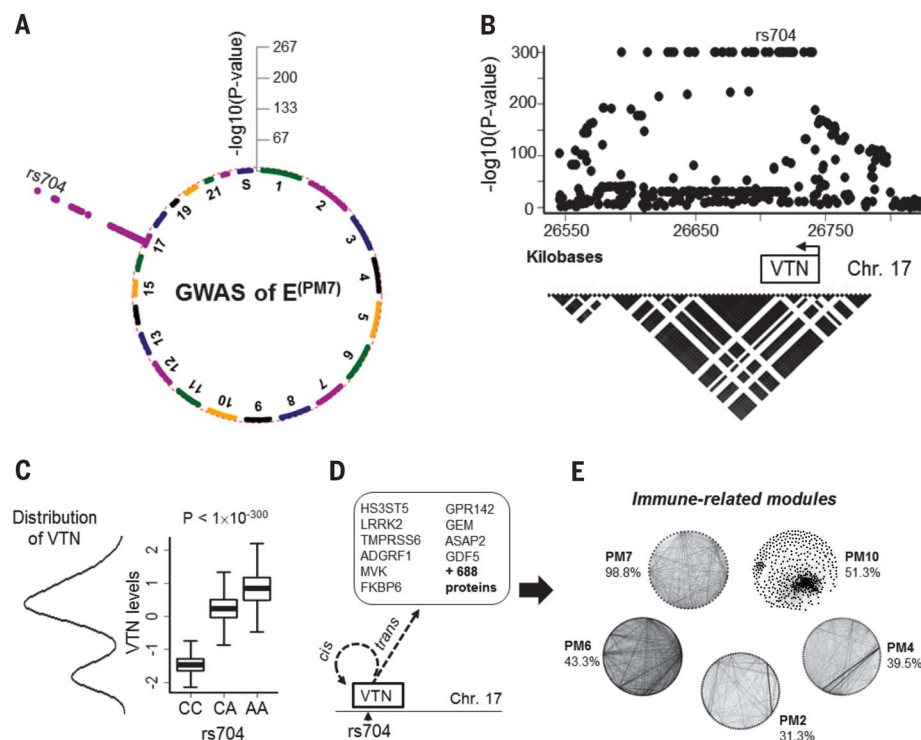
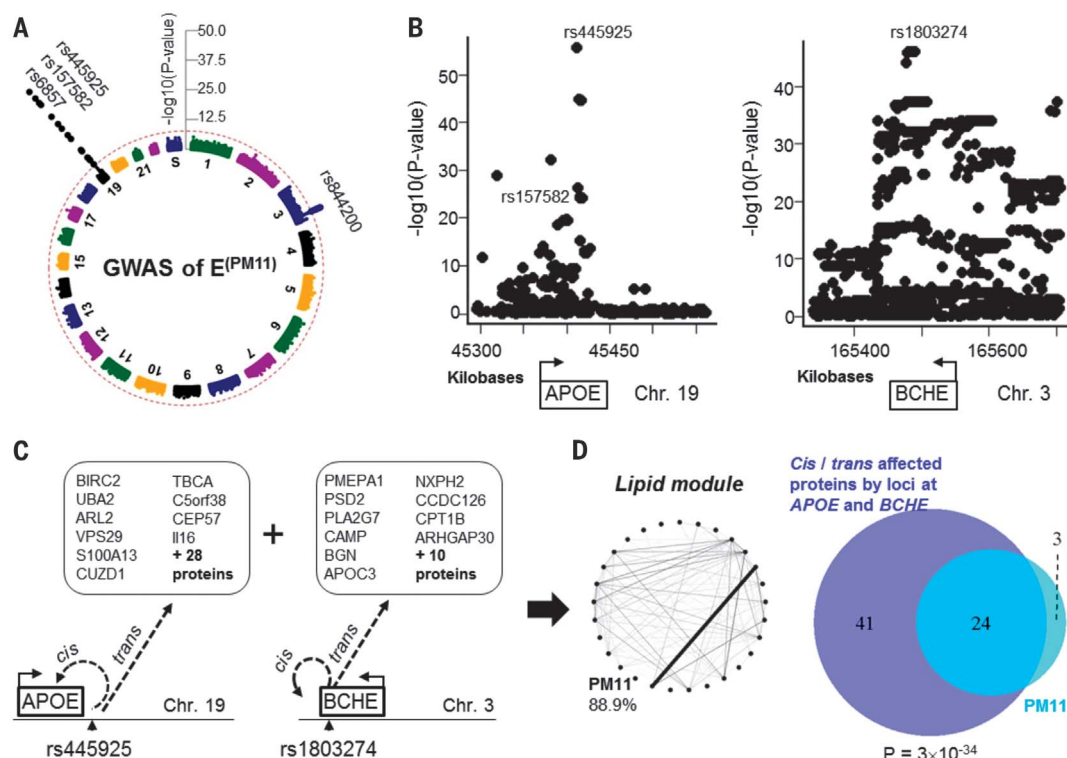


Fig. 4. The network-associated pSNP rs704 regulates modules related to immune functions. (A) A circular Manhattan plot highlights the GWAS results for $E^{(PM7)}$, revealing a single highly significant association at rs704, a missense variant (NP_000629.3: p.Thr400Met) in *VTN*. Four other modules within supercluster II were affected by rs704 (table S20). (B) The rs704 variant, and many other linked SNPs in the region, exert a strong cis-acting effect on *VTN* within the 300-kb genomic region across the *VTN* gene. The black triangle demonstrates linkage disequilibrium (r^2) patterns in the region derived from the AGES cohort data. (C) The bimodal population distribution of the *VTN* protein is explained by the drastic reduction in *VTN* levels in individuals homozygous for the rs704 C minor allele. (D) A schematic presentation of the single cis and many trans effects mediated by the rs704 in *VTN* on serum proteins. (E) Proteins affected by rs704 in trans (and *VTN* in cis) cluster in modules of immune-related functions (table S20). The percentage denotes the fraction of proteins within a given module regulated by the rs704 locus.

Fig. 5. Network-associated pSNPs at the APOE and BCHE loci regulate proteins of module PM11. (A) A circular Manhattan plot for the $E^{(PM11)}$, revealing two distinct genomic loci at chromosomes 3 (BCHE) and 19 (APOE / TOMM40). The three npSNPs at the APOE / TOMM40 locus are not correlated ($r^2 = 0$). (B) The npSNPs at the two genomic regions exert a strong cis-acting effect on the serum levels of APOE (left) and BCHE (right). (C) The npSNPs at the APOE locus affected APOE in cis and mediated trans effects on 38 proteins, whereas the npSNP rs1803274, a missense variant (NP_000046.1: pAla567Thr) in BCHE, affected BCHE in cis and 20 other proteins in trans (table S20). (D) The distinct npSNPs at the APOE and BCHE loci regulate 88.9% of all proteins that constitute PM11, Fisher's exact test $P = 3 \times 10^{-34}$, as demonstrated in the Venn diagram to the right. The number 3 refers to proteins in PM11 that are not regulated by the npSNPs.



in adaptive homeostasis in comparison to younger populations. Note that, for all the analyses of the present study, we adjusted for the confounder age.

Structural features of the serum protein network resembled those of regulatory networks constructed in solid tissues (7, 8), including association of hub proteins to diseases. However, comparison of cis-trans protein pairs and cognate protein modules to gene expression data and coexpression networks derived from solid tissues suggests that the serum protein network arose in part via systemic cross-tissue regulation. A follow-up replication testing of the many findings from these primary data, including the strong connection of protein hubs to disease traits, is warranted.

We anticipate that additional npSNPs can be identified with this technique, comparable to those found with conventional GWAS analyses of common diseases (25). Given that the effect of established GWAS loci is more complex than previously anticipated, this underscores the role of protein networks as the sensors and integrators of complex disease. The strong association of individual proteins and networks to disease states observed in this study indicates that the serum proteome may be a rich and accessible setting to mine for biomarkers of disease and disease responses to integrate information from tissues in a global regulatory network. As such, coordinated variance of serum proteins may offer unrecognized opportunities for target and biomarker identification in human disease.

REFERENCES AND NOTES

1. J. M. Schwenk *et al.*, *J. Proteome Res.* **16**, 4299–4310 (2017).
2. M. Uhlen *et al.*, *Science* **347**, 1260419 (2015).
3. M. Stastna, J. E. Van Eyk, *Proteomics* **12**, 722–735 (2012).
4. I. M. Conboy *et al.*, *Nature* **433**, 760–764 (2005).
5. S. A. Villeda *et al.*, *Nature* **477**, 90–94 (2011).
6. E. E. Schadt, *Nature* **461**, 218–223 (2009).
7. B. Zhang *et al.*, *Cell* **153**, 707–720 (2013).
8. V. Ennissson *et al.*, *Nature* **452**, 423–428 (2008).
9. Y. Chen *et al.*, *Nature* **452**, 429–435 (2008).
10. D. R. Davies *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19971–19976 (2012).
11. L. Gold *et al.*, *PLOS ONE* **5**, e15004 (2010).
12. T. B. Harris *et al.*, *Am. J. Epidemiol.* **165**, 1076–1087 (2007).
13. A. L. Barabási, R. Albert, *Science* **286**, 509–512 (1999).
14. B. Zhang, S. Horvath, *Stat. Appl. Genet. Mol. Biol.* **4**, e17 (2005).
15. P. Langfelder, R. Luo, M. C. Oldham, S. Horvath, *PLOS Comput. Biol.* **7**, e1001057 (2011).
16. L. Shu *et al.*, *PLOS Genet.* **13**, e1007040 (2017).
17. H. Jeong, S. P. Mason, A. L. Barabási, Z. N. Oltvai, *Nature* **411**, 41–42 (2001).
18. A. L. Barabási, N. Gulbahce, J. Loscalzo, *Nat. Rev. Genet.* **12**, 56–68 (2011).
19. M. Muñoz *et al.*, *Nat. Genet.* **48**, 980–983 (2016).
20. J. R. Staley *et al.*, *Bioinformatics* **32**, 3207–3209 (2016).
21. N. Mahler *et al.*, *PLOS Genet.* **13**, e1006402 (2017).
22. J. K. Pickrell *et al.*, *Nat. Genet.* **48**, 709–717 (2016).
23. M. Franchini, G. Lippi, *BMC Med.* **13**, 7 (2015).
24. M. Franchini, F. Capra, G. Targher, M. Montagnana, G. Lippi, *Thromb. J.* **5**, 14 (2007).
25. E. A. Boyle, Y. I. Li, J. K. Pritchard, *Cell* **169**, 1177–1186 (2017).
26. D. Alfego, U. Rodeck, A. Kriete, *PLOS ONE* **13**, e0190457 (2018).
27. J. Yang *et al.*, *Sci. Rep.* **5**, 15145 (2015).
28. J. M. Zahn *et al.*, *PLOS Genet.* **3**, e201 (2007).
29. P. Langfelder, B. Zhang, S. Horvath, *Bioinformatics* **24**, 719–720 (2008).
30. G. Csardi, T. Nepusz, *Complex Syst.* **1695**, 1 (2006).

ACKNOWLEDGMENTS

We thank the staff at SomaLogic (CO) for performing the assays to measure protein levels, and P. MacNamara and G. Joyce of the Genomics Institute of the Novartis Research Foundation (GNF) for their leadership in supporting this work. We thank J. Loureiro, V. Swaroop, S. Abubucker, and F. Mapa of NIBR Cambridge for their contributions

in support of the mass spectrometry workflow. We thank K. Bjarnadóttir, S. Gunnarsdóttir, and A. Hauksdóttir at the Icelandic Heart Association (IHA) for all specimen handling. **Funding:** Supported by Novartis Institute for Biomedical Research (NIBR), IHA, and in part by the intramural research program at the National Institute of Aging (N01-AG-12100 and HHSN271201200022C), the Althingi (the Icelandic Parliament), and the Icelandic Centre for Research (RANNIS) grant 141101-051. **Author contributions:** Conception and design: V.E., J.R.L., and V.G.; Data curation: E.F.G., T.A., S.R.H., V.T., N.F., S.S., G.E., T.B.H., L.L.J.; Methodology and analysis: V.E., M.J., J.R.L., T.A., E.F.G., V.G., N.F., L.L.J., R.P., H.H., L.S., X.Y., S.A.L., J.T., J.Z., B.Z., J.Z., A.P.O., A.M., V.A.G., Ö.O., and J.J.; Supervision: V.E., V.G., J.R.L., and L.L.J.; Writing—original draft: V.E.; Writing—review and editing: V.E., J.R.L., V.G., N.F., L.L.J., M.J., V.A.G., L.S., and X.Y. **Competing interests:** J.T., S.R.H., V.T., N.F., R.P., L.L.J., A.P.O., and J.R.L. are all employees and stockholders of Novartis. All other authors declare no competing interests. **Data and materials availability:** The raw mass spectrometry data (DDA or MRM) were deposited to the ProteomeXchange Consortium with the five dataset identifiers PXD008819 to PXD008823, and the dataset identifier PASS01145. The custom-design Novartis SOMAScan is available through a collaboration agreement with the Novartis Institutes for Biomedical Research (lori.jennings@novartis.com). Data from the AGES Reykjavik study are available through collaboration (AGES_data_request@harta.is) under a data usage agreement with the IHA. All data supporting the conclusions of the paper are presented in the main text; supplementary materials; and Excel tables S1, S3, S4, S6, S7, S9, S10, S13 to S15, S17, S21, and S22 hosted online. A description of methods and materials is available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/769/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S22
References (31–68)

10 October 2017; resubmitted 7 February 2018
Accepted 13 July 2018
Published online 2 August 2018
10.1126/science.aag1327

REPORT

GAS GIANT PLANETS

Juno observations of spot structures and a split tail in Io-induced aurorae on Jupiter

A. Mura^{1*}, A. Adriani¹, J. E. P. Connerney^{2,3}, S. Bolton⁴, F. Altieri¹, F. Bagenal⁵, B. Bonfond⁶, B. M. Dinelli⁷, J.-C. Gérard⁶, T. Greathouse⁴, D. Grodent⁶, S. Levin⁸, B. Mauk⁹, M. L. Moriconi⁷, J. Saur¹⁰, J. H. Waite Jr.^{4,11}, M. Amoroso¹², A. Cicchetti¹, F. Fabiano⁷, G. Filacchione¹, D. Grassi¹, A. Migliorini¹, R. Noschese¹, A. Olivieri¹², G. Piccioni¹, C. Plainaki^{1,12}, G. Sindoni¹², R. Sordini¹, F. Tosi¹, D. Turrini¹

Jupiter's aurorae are produced in its upper atmosphere when incoming high-energy electrons precipitate along the planet's magnetic field lines. A northern and a southern main auroral oval are visible, surrounded by small emission features associated with the Galilean moons. We present infrared observations, obtained with the Juno spacecraft, showing that in the case of Io, this emission exhibits a swirling pattern that is similar in appearance to a von Kármán vortex street. Well downstream of the main auroral spots, the extended tail is split in two. Both of Ganymede's footprints also appear as a pair of emission features, which may provide a remote measure of Ganymede's magnetosphere. These features suggest that the magnetohydrodynamic interaction between Jupiter and its moon is more complex than previously anticipated.

Like Earth, Jupiter has aurorae. These permanent electromagnetic emissions are associated with the precipitation of magnetospheric plasma onto the planet's ionosphere. Jupiter's aurorae have emission features, absent at Earth, that are associated with its moons; bright spots appear in the ionosphere at the base of the magnetic field lines that sweep past the Galilean moons Io, Europa, and Ganymede. The discovery of auroral features at the magnetic field footprints of Io (1, 2) showed the electromagnetic interaction between Jupiter and its moons. In situ observations by the Galileo spacecraft of intense bidirectional, magnetic field-aligned electron beams, with energies up to 200 keV, within the wake and over the poles of Io showed that these interactions can lead to strong electron acceleration (3–5).

Infrared (IR) and ultraviolet (UV) observations of Jupiter's auroral footprints led to theoretical

understanding that the equatorial magnetospheric plasma roughly corotates with the planet, but the moons, which orbit more slowly, form an obstacle to this plasma (6, 7). This interaction leads to the formation of magnetohydrodynamic

waves—specifically, Alfvén waves (8, 9)—that accelerate electrons toward the planet's atmosphere, where they excite auroral emissions (9). Whereas electrons travel rapidly along the magnetic field lines, the slower Alfvén waves are convected downstream, resulting in a standing Alfvén wing that is slightly tilted away from the undisturbed magnetic field line (8). The speeds of the waves are highly variable, depending on the path toward Jupiter through different plasma environments (10). Waves generated at Io first travel a dense plasma torus around the moon and then into the low-density magnetosphere lobe regions (Fig. 1). Successive traversals of plasma discontinuities may lead to multiply reflected Alfvén waves, which has been suggested as a possible cause of the observed multiple footprint features (11–13).

The observation of auroral H_3^+ emission in relation to the positions of the moons, where the interaction occurred, is a powerful tool for remote diagnostic of the jovian magnetosphere (9).

The Jovian Infrared Auroral Mapper (JIRAM) (14) on NASA's Juno spacecraft (15, 16) has been observing the jovian aurora since 2016 (17) from a polar, highly elliptical orbit. From its position above the poles, JIRAM has obtained views of Jupiter's auroral regions (18). JIRAM is an IR imager and spectrometer; one of the imaging channels operates in the range of 3.3 to 3.6 μm , at which some of the most intense H_3^+ auroral emission lines are observed. During Juno's eighth close pass of the planet (1 September 2017), JIRAM targeted the predicted location of the Io footprint (IFP).

In Fig. 2, we show the IFP peak emissions observed by JIRAM on that occasion at a spatial resolution of ~ 20 km. Io orbits Jupiter rapidly, so its longitude in the reference frame of the planet was $\sim 80^\circ\text{W}$ during the northern observations and $\sim 135^\circ\text{W}$ during the southern observations. In both the northern and southern hemispheres, the main IFP is preceded by a ~ 1000 -km-long, thin emission feature, interpreted as the IR signature of electron beams accelerated in both hemispheres [trans-hemispheric electron beam (TEB)] (13). However, on this occasion, the precursor feature is not detached from the rest of the footprint. The characteristic decay time of H_3^+ is a few hundred seconds (19)—longer than the almost instantaneous UV hydrogen-related time scale (12), which may explain the blurring and extent of the TEB signature in the IR, unlike previous UV observations (13).

The southern main spot is followed by a series of regularly spaced secondary spots, which are alternately displaced poleward and equatorward of the median track (Fig. 2A). The transverse displacement is 100 ± 10 km. By tracing back to the orbit of Io by use of a magnetic field model (20), this distance corresponds to 4000 ± 1000 km, which is roughly the size of Io itself.

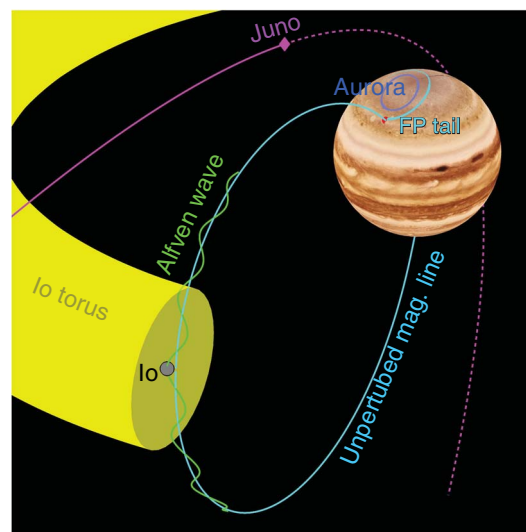


Fig. 1. Schematic illustration of the interaction between Io and the magnetosphere of Jupiter. The figure shows Io, the Io plasma torus (yellow; simplified and not to scale), an unperturbed magnetic field line (cyan), an Alfvén wave (green; shape is simplified and not to scale), the main IFP (red), the footprint tail (cyan), the main auroral arc (blue), and the trajectory and position of the Juno spacecraft.

¹Institute for Space Astrophysics and Planetology, National Institute for Astrophysics, Rome, Italy. ²Space Research Corporation, Annapolis, MD, USA. ³NASA Goddard Space Flight Center, Greenbelt, MD, USA. ⁴Southwest Research Institute, San Antonio, TX, USA. ⁵Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, CO, USA. ⁶Space Science, Technologies and Astrophysical Research Institute, Laboratory for Planetary and Atmospheric Physics, University of Liège, Liège, Belgium. ⁷Institute of Atmospheric Sciences and Climate, National Research Council, Italy. ⁸Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. ⁹Applied Physics Laboratory, Johns Hopkins University, Laurel, MD, USA. ¹⁰Institut für Geophysik und Meteorologie, University of Cologne, Köln, Germany. ¹¹Department of Physics and Astronomy, University of Texas at San Antonio, San Antonio, TX, USA. ¹²Agenzia Spaziale Italiana, Rome, Italy. *Corresponding author. Email: alessandro.mura@iaps.inaf.it

The main spot location is exactly along the track predicted by the model (20), and the repetition step of the features is 350 ± 10 km.

Given the relative velocity of the main spot over the surface of Jupiter, the separation of secondary spots corresponds to ~ 60 s. The chemical lifetime of H_3^+ is longer (up to 1000 s) (19) and is expected to produce extensive smearing, which should hide the observed small-scale structure of the spots, as proposed for the leading TEB spot. A possible explanation is that the secondary spots are generated by higher-energy electrons. A model of inertial Alfvén waves acceleration (21) indicates that the electron energy spectrum might be different for the direct, main Alfvén wing and the conjugate (TEB) electron beams. High-energy electrons would precipitate deeper into the atmosphere, producing H_3^+ ions where their lifetime is shorter because of the high methane density (H_3^+ destruction occurs via the reaction $\text{H}_3^+ + \text{CH}_4 \rightarrow \text{CH}_5^+ + \text{H}_2$) (22). An alternative explanation, for the blurring of the precursor feature, could be that the electron beam that creates the TEB spot, formed by up-going electrons from the opposite hemisphere, broadened on its way along the field line through the whole magnetosphere.

The 60-s separation between spots, and the similar intensities of subsequent spots, are difficult to explain with multiple specular reflections, neither inside nor outside the dense plasma torus (10). However, high-amplitude nonlinear waves are generally not reflected specularly. The multiplicity and shape of the spots can be explained by models (23), in which strong nonlinearities trigger wave-interference patterns that lead to the occurrence of multiple spots, depending on Io's location in the torus. The observations suggest that the Alfvén waves might be broken up into smaller waves with different travel paths, leading to a turbulent emission pattern. The angular offset of the elongated main spot and the alternating displacements of the tail spots might be influenced by the Hall effect in Io's ionosphere, which is expected to rotate the Alfvén waves (24). Alternatively, because this feature superficially resembles a von Kármán vortex street (a hydrodynamically observed phenomena) (25), this may result from Io-produced flow shears that generate vortices downstream in the plasma-atmosphere interaction region.

In the northern hemisphere (Fig. 2, C to E), the shape of the IFP is generally consistent with that observed in the south: An array of secondary spots is visible in the first part of the tail. Away from the main IFP, the tail becomes gradually more turbulent, likely because of a continuation of the Alfvénic interaction, similar to the main spot (26).

Each spot is ~ 200 km wide. Because JIRAM was observing from an angle of $\sim 35^\circ$ between the vertical direction and the line of sight, this means that the height of the main H_3^+ emission region cannot be taller than $200 \text{ km} / \tan(35^\circ) \approx 300$ km. This short column height is consistent with generation by high-energy electrons (27).

Extrapolating the radiance collected by JIRAM filter to the full H_3^+ spectrum, and integrating it

over the surface, the total radiated energy can be estimated at ~ 6 GW in the south and ~ 3 GW in the north, which is ~ 3 times less than in the UV (28). The north-south asymmetry of the emission, both in shape but especially in intensity, is a recurrent feature of Jupiter's aurorae (29). Some hypotheses have been proposed (30) to explain this as a consequence of the north-south magnetic field asymmetries: The magnetic field controls

the electron acceleration efficiency and the probability for an electron to reach the atmosphere. In our case, the difference may also be due to differing observation times.

The far tail of the IFP as observed on 27 August 2016 in the northern hemisphere is shown in Fig. 3. The main spot is $\sim 100^\circ$ upstream in longitude from this region. In this sector, we found that the tail consists of two separate parallel arcs. The

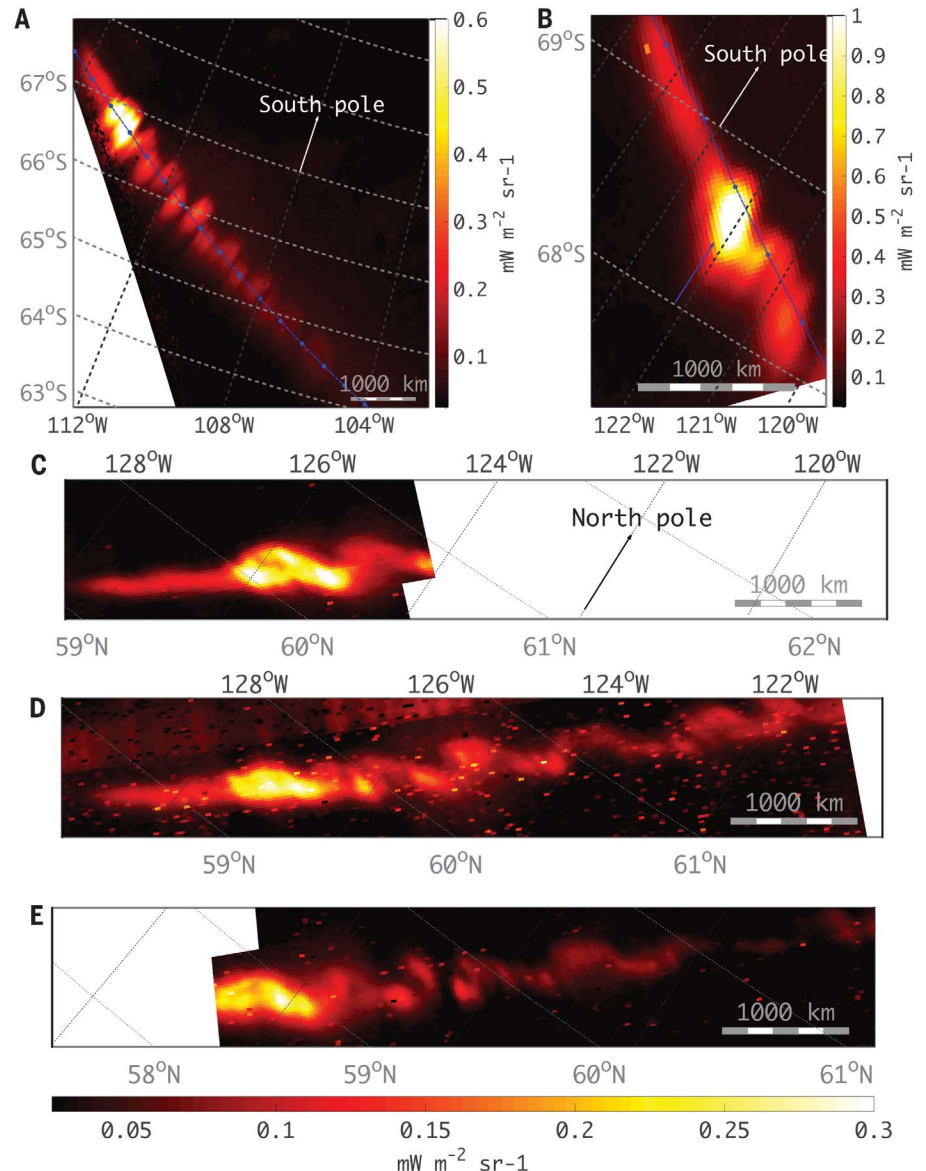


Fig. 2. Images of the IFP (radiance) in polar orthographic projections with parallels and meridians overlotted. (A) IFP on the south pole of Jupiter, as seen on 1 September 2017 at 22:50. The main spot (top left corner of the image) is followed by a series of regularly spaced secondary spots, which alternate in displacement above and below the median track (blue). The thin blue arc with dots is the sequence of the predicted IFP positions (20), with dots showing its position every 100 s. Radiance is in $\text{mW m}^{-2} \text{sr}^{-1}$. **(B)** As in (A), but showing a later image taken at 23:09. The main spot (indicated with a blue arrow) has an oval shape, with an angular offset from the average direction of the tail (median track). **(C to E)** The IFP in the north aurora, observed at 20:43, 20:48, and 20:53, respectively, on 1 September 2017. The model (20) is not shown in (C), (D), and (E) because it lies outside of the visible region. The black and white pixels [especially in (D)] are the effect of penetrating radiation affecting the detector. The results of Jacobsen *et al.*'s model is available in (23), figure 3, for comparison.

Fig. 3. The IFP tail, observed $\sim 100^\circ$ behind the main spot. Image was taken on 27 August 2016 (contour plot). The tail consists of two separate arcs. The blue line is the prediction from a magnetic field model (20). The vertical artefacts on this image are produced by the noise-reduction routine. Local time is approximately dawn.

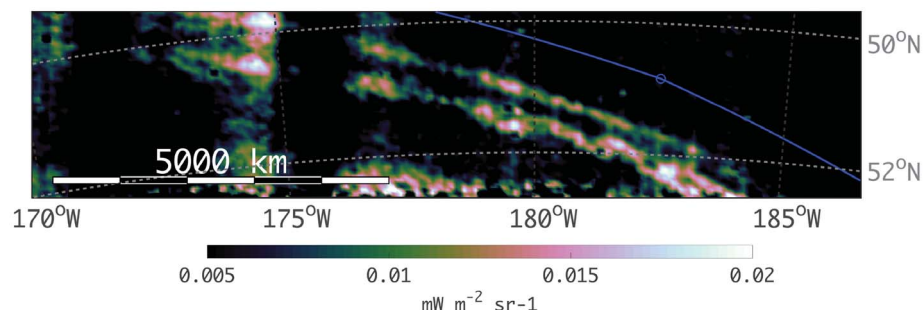
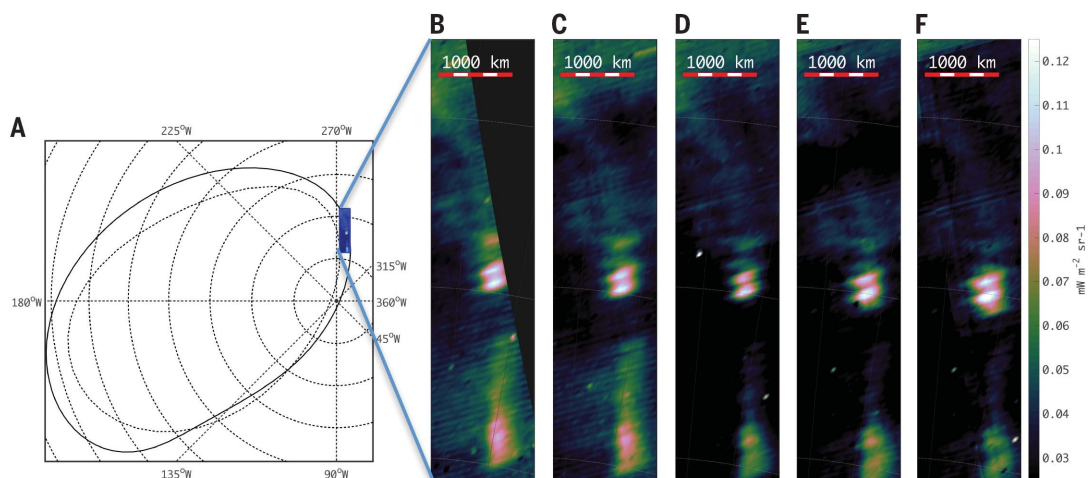


Fig. 4. Five consecutive JIRAM images of Ganymede footprints. (A) Diagram illustrating the main oval (dotted oval), Ganymede footprint (solid line oval) as predicted from a recent magnetic field model (33) and the approximate location of the images in (B) to (F). (B to F) Images are taken from 00:07:45 to 00:09:48 on 11 July 2017 on the north hemisphere (contour plots). The main Ganymede footprint (in the middle of each image, and moving down) is preceded by a dimmer precursor spot.



poleward arc is thicker, more intense, and more turbulent, whereas the equatorward arc is fainter, thinner, and less turbulent. These features are visible in previous and subsequent images (taken within 1 hour) and extend for at least 70° along the tail. This double footprint tail has not been observed in UV images, probably because of insufficient spatial resolution. However, other JIRAM images with similar spatial resolution [(18), their figures 2 and 3] show a single tail, indicating that the double feature is not persistent or not present at all longitudes.

Compared with the cross-track displacement of the IFP tail close to Io shown in Fig. 2 (~ 100 km), the separation in Fig. 3 is larger: The parallel tail components are separated by 500 to 1000 km. The viewing angle in Fig. 3 is less than 10° , ruling out the possibility that JIRAM was observing a vertical (rather than horizontal) structure. The difference in the shape and intensity of the two arcs could be due to the Hall effect generating brighter aurorae in Io's atmosphere on the anti-jovian side than on the jovian side (24). The poleward part of the corresponding footprint on Jupiter should then be brighter than the equatorward part, as observed. We suggest that the smooth tail might arise from trans-hemispheric electrons similarly to those of the TEB (23) feature in Fig. 2A, whereas the other tail might be related to the wave, which would produce a patchier tail. The splitting could then be due to a northern anomaly in Jupiter's magnetic field (31), which might also explain why this bifurcation is not observed in the southern aurora.

The comparison of the IFP with the footprint of Ganymede (GFP) can provide insight into the mechanisms that generate these morphologies. Shown in Fig. 4 are several images of the GFP (taken 30 s apart), as observed on 11 July 2017 on Juno's seventh close pass of the planet (north hemisphere); at this time, Ganymede was well south of the plasma sheet. The secondary, upstream spot could be a TEB spot, and the main spot-TEB distance of 1500 km is consistent with similar measurements in the far UV at the same longitude (28). However, the JIRAM images show that each spot consists of two identical peaks ~ 170 km apart. This separation corresponds to a tangential separation of $\sim 10,000$ km at the orbit of Ganymede, if magnetically mapped with a recent model (32) in a way similar to that of the IFP. Each spot is also followed by a very faint tail (26), which appears to be twisted.

As with the IFP, the spatial separations of features in the Ganymede footprint are not consistent with the simple specular reflection of Alfvén waves. In the Ganymede case, however, there is no trailing array of spots, with only two visible. A possible explanation is that both the leading and trailing magnetic reconnection regions in Ganymede magnetosphere [(33), their figure 1] generate locally enhanced Alfvén waves. If so, the separation between the peaks is related to the size of Ganymede's magnetosphere. Both the TEB and the main GFP display this double structure, perhaps suggesting that the cause is located at the origin of both the TEB and the main GFP Alfvén waves. However, other scenar-

ios, such as reflections in the ionospheric resonator (34), could be invoked as well.

Auroral footprints of jovian moons are peculiar, but not unique; a similar footprint of Enceladus is observed at Saturn (35), and even tentative evidence of footprints of extrasolar planets on their central stars exists (36, 37). These auroral processes apply to any electrically conducting moon orbiting within a planet's magnetosphere. The Juno/JIRAM high-resolution images of the satellite footprints in the IR show substructures, both in the spots and in the extended tail. The close-spaced, multiple spots require a process that occurs at smaller scales than would be produced by a single pair of Alfvén wings, in linear regime, bouncing between hemispheres, suggesting that they are produced by interactions occurring close to the moon, such as a train of vortices generated in Io's wake. Alternatively, interference between Alfvén waves, rather than reflections, could be responsible. Our images show that the extended tail can occasionally form two parallel arcs. Because this bifurcation is not observed often, it is possible that it is triggered by the northern anomaly in Jupiter's magnetic field. The Ganymede aurora expresses as twin spots, which may probe Ganymede's distant magnetosphere using the auroral emission.

REFERENCES AND NOTES

1. J. E. P. Connerney, R. Baron, T. Satoh, T. Owen, *Science* **262**, 1035–1038 (1993).
2. J. T. Clarke *et al.*, *Science* **274**, 404–409 (1996).
3. D. J. Williams *et al.*, *Science* **274**, 401–403 (1996).
4. D. J. Williams, R. M. Thorne, *J. Geophys. Res.* **108**, 1397 (2003).

5. L. A. Frank, W. R. Paterson, *J. Geophys. Res.* **104**, 28657–28669 (1999).
6. P. Goldreich, D. Lynden-Bell, *Astrophys. J.* **156**, 59–78 (1969).
7. M. H. Acuña, F. M. Neubauer, N. F. Ness, *J. Geophys. Res.* **86**, 8513–8521 (1981).
8. F. M. Neubauer, *J. Geophys. Res.* **85**, 1171–1178 (1980).
9. J. E. P. Connerney, T. Satoh, *Philos. Trans. R. Soc. London Ser. A* **358**, 2471–2483 (2000).
10. F. Bagenal, *J. Geophys. Res.* **88**, 3013–3025 (1983).
11. J. T. Clarke *et al.*, *Nature* **415**, 997–1000 (2002).
12. J.-C. Gérard, A. Saglam, D. Grodent, J. T. Clarke, *J. Geophys. Res.* **111**, A04202 (2006).
13. B. Bonfond *et al.*, *J. Geophys. Res.* **114**, A07224 (2009).
14. A. Adriani *et al.*, *Space Sci. Rev.* **213**, 393–446 (2014).
15. S. J. Bolton *et al.*, *Science* **356**, 821–825 (2017).
16. F. Bagenal *et al.*, *Space Sci. Rev.* **213**, 219–287 (2017).
17. J. E. P. Connerney *et al.*, *Science* **356**, 826–832 (2017).
18. A. Mura *et al.*, *Geophys. Res. Lett.* **44**, 5308–5316 (2017).
19. T. Stallard, S. Miller, G. Millward, R. D. Joseph, *Icarus* **156**, 498–514 (2002).
20. J. E. P. Connerney, M. H. Acuna, N. Ness, T. Satoh, *J. Geophys. Res.* **103**, 11,929–11,939 (1998).
21. S. L. G. Hess, P. A. Delamere, V. Dols, B. Bonfond, D. Swift, *J. Geophys. Res.* **115**, A06205 (2010).
22. J.-C. Gérard *et al.*, *Icarus* **312**, 145–156 (2018).
23. S. Jacobsen, F. M. Neubauer, J. Saur, N. Schilling, *Geophys. Res. Lett.* **34**, L10202 (2007).
24. J. Saur, F. M. Neubauer, D. F. Strobel, M. E. Summers, *Geophys. Res. Lett.* **27**, 2893–2896 (2000).
25. T. von Kármán, *Aerodynamics* (McGraw-Hill, 1963).
26. B. Bonfond *et al.*, *J. Geophys. Res.* **122**, 7985–7996 (2017).
27. C. Tao, S. V. Badman, M. Fujimoto, *Icarus* **213**, 581–592 (2011).
28. B. Bonfond *et al.*, *Geophys. Res. Lett.* **40**, 4977–4981 (2013).
29. A. Adriani *et al.*, *Geophys. Res. Lett.* **44**, 4633–4640 (2017).
30. S. L. G. Hess *et al.*, *Planet. Space Sci.* **88**, 76–85 (2013).
31. D. Grodent *et al.*, *J. Geophys. Res.* **113**, A09201 (2008).
32. J. E. P. Connerney *et al.*, *Geophys. Res. Lett.* **45**, 2590–2596 (2018).
33. M. G. Kivelson *et al.*, *J. Geophys. Res.* **103**, 19963–19972 (1998).
34. Y. Su *et al.*, *J. Geophys. Res.* **111**, A06211 (2006).
35. W. R. Pryor *et al.*, *Nature* **472**, 331–333 (2011).
36. E. Shkolnik, D. A. Bohlender, G. A. H. Walker, A. Collier Cameron, *Astrophys. J.* **676**, 628–638 (2008).
37. J. Saur, T. Grambusch, S. Duling, F. M. Neubauer, S. Simon, *Astron. Astrophys.* **552**, A119 (2013).

ACKNOWLEDGMENTS

We thank Agenzia Spaziale Italiana (ASI) for support of the JIRAM contribution to the Juno mission. **Funding:** JIRAM is funded with ASI contract 2016-353 23-H.O. B.B., D.Gro., and J.-C.G. are supported by the PRODEX program managed by ESA in collaboration with the Belgian Federal Science Policy Office. T.G. is supported by the

Juno Project under NASA grant NMO6AA75c to Southwest Research Institute. B.B. is funded by the Fund for Scientific Research (F.R.S.-FNRS). **Authors contributions:** Preparation of the manuscript, figures, calculations, data analysis was performed by A.Mu. and A.A.; data interpretation was performed by A.Mu., A.A., J.E.P.C., S.B., F.A., F.B., B.B., B.M.D., J.-C.G., T.G., D.G., S.L., B.M., M.L.M., J.S., J.H.W., and D.T.; revisions of the manuscript were made by A.Mu., A.A., J.E.P.C., S.B., F.A., F.B., B.B., B.M.D., J.-C.G., T.G., D.Gro., S.L., B.M., M.L.M., J.S., and J.H.W.; S.B. is the Juno principal investigator; A.A. is responsible for the JIRAM instrument; preparation for measurements and data acquisition was performed by M.A., A.C., F.F., G.F., D.Gra., A.Mi., R.N., A.O., G.P., C.P., G.S., R.S., F.T., and D.T.; calibration was done by A.A., G.F., A.Mu., and G.P.; and georeferencing was performed by F.T., R.N., and A.Mu. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data in this study is publicly available on the Planetary Data System (<http://pds.nasa.gov>) and can be downloaded from <http://atmos.nmsu.edu:8080/pds>. The individual datasets are available at https://atmos.nmsu.edu/PDS/data/jnojr_xxxx, where xxxx is 1001, 1002, or 1003 for EDR (Experiment Data Record; raw data) and 2001, 2002, or 2003 for RDR (Reduced Data Record; calibrated data) volumes. Noise reduction was performed by using the algorithm in (18); a 2- by 1-pixel filter was applied to remove the odd-even effect.

30 January 2018; accepted 15 June 2018

Published online 5 July 2018

10.1126/science.aat1450

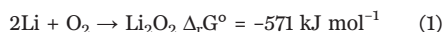
BATTERIES

A high-energy-density lithium-oxygen battery based on a reversible four-electron conversion to lithium oxide

C. Xia, C. Y. Kwok, L. F. Nazar*

Lithium-oxygen (Li-O_2) batteries have attracted much attention owing to the high theoretical energy density afforded by the two-electron reduction of O_2 to lithium peroxide (Li_2O_2). We report an inorganic-electrolyte Li-O_2 cell that cycles at an elevated temperature via highly reversible four-electron redox to form crystalline lithium oxide (Li_2O). It relies on a bifunctional metal oxide host that catalyzes O–O bond cleavage on discharge, yielding a high capacity of 11 milliamperes-hours per square centimeter, and O_2 evolution on charge with very low overpotential. Online mass spectrometry and chemical quantification confirm that oxidation of Li_2O involves transfer of exactly $4 e^-/\text{O}_2$. This work shows that Li-O_2 electrochemistry is not intrinsically limited once problems of electrolyte, superoxide, and cathode host are overcome and that coulombic efficiency close to 100% can be achieved.

Lithium-oxygen (Li-O_2) batteries have attracted interest because of their energy density being at least one magnitude higher than that of conventional Li-ion batteries (1). A typical Li-O_2 cell is composed of a Li anode and a porous carbon cathode, separated by a Li^+ -ion conducting organic electrolyte (2). During discharge, oxygen is reduced on the carbon cathode, forming insoluble lithium peroxide (Li_2O_2) (Eq. 1), and this product is oxidized upon charge.



The commercialization of the Li-O_2 battery and its Na-O_2 battery cousin (3) have been impeded by the decomposition of organic electrolytes as well as by the corrosion of the porous carbon cathode hosts (4, 5). These drawbacks are directly correlated to the high reactivity of peroxide and superoxide in the Li-O_2 and Na-O_2 batteries, respectively, and superoxide has also been reported as the sole product of the Li-O_2 cell (6). Although effort has been made to improve the cycling performance of Li-O_2 cells—albeit by means of a complex and debated pathway (7, 8)—the issues remain.

In comparison with the peroxide and superoxide, lithium oxide (Li_2O) is much less chemically reactive with organic solvents. Both peroxide (9) and superoxide (10) react with organic electrolytes and with carbon in the cell. Owing to its oxide anion, Li_2O is benign as an oxidizing agent; for example, it does not react with dimethyl sulfoxide (DMSO), whereas Li_2O_2 oxidizes DMSO to form dimethyl sulfone (fig. S1). More importantly, Li-O_2 cells based on Li_2O as the discharge product (Eq. 2) can theoretically deliver a high specific energy and energy density of

5.2 kilowatt-hour (kWh) kg^{-1} and 10.5 kWh L^{-1} , respectively, exceeding that of fossil fuels [gasoline (9.5 kWh L^{-1})] (2)



Therefore, it is of much interest to seek a pathway for reversible oxygen reduction to the oxide. Consideration of the thermodynamics of the oxygen reduction reaction (ORR) shows that not only is the standard Gibbs reaction energy ($\Delta_r G^\circ$) of Eq. 1 [-571 kJ mol^{-1} (Li_2O_2)] lower than that of Eq. 2 [-561 kJ mol^{-1} (Li_2O)] (11), but the formation of oxide requires O–O bond cleavage of oxygen molecules, whereas peroxide does not. Therefore, the formation of Li_2O , not Li_2O_2 , is thermodynamically and kinetically favored at ambient conditions.

We demonstrate that by increasing the operating temperature and exploiting stable inorganic electrolytes and ORR catalysts, the reversible for-

mation of Li_2O leads to a highly rechargeable Li-O_2 cell with high capacity, low overpotential with transfer of $4 e^-/\text{O}_2$, and excellent cycling performance. Our cells were cycled at 150°C because the thermodynamic driving force favors Li_2O as the product above this temperature rather than Li_2O_2 (Fig. 1A), as described by the Gibbs-Helmholtz equation. The cell design is depicted in Fig. 1B. A lithium nitrate/potassium nitrate ($\text{LiNO}_3/\text{KNO}_3$) eutectic molten salt operates as the liquid electrolyte because of its good chemical stability and high conductivity (12), and a solid electrolyte [$\text{Li}_{1.5}\text{Al}_{0.5}\text{Ge}_{1.5}(\text{PO}_4)_3$ (LAGP)] membrane at the Li anode inhibits the crossover of soluble products (fig. S2). A noncarbonaceous composite cathode composed of nickel (Ni) nanoparticles coated in situ to form Li_2NiO_2 supplies the vital electrocatalyst that reversibly catalyzes O–O bond cleavage and formation (fig. S3). Detailed material synthesis and characterization methods are provided in the supplementary materials.

Cells were sealed with oxygen and cycled between 2.6 and 3.5 V at an applied current of 0.1 mA cm^{-2} (Fig. 2). The inorganic electrolyte Li-O_2 cell with a Ni-nitrate composite cathode exhibits a very high discharge capacity of 11 mA-hour cm^{-2} (Fig. 2A, red solid curve). This value is more than 20-fold higher than achieved in a cell using an aprotic organic electrolyte and a carbon electrode ($0.5 \text{ mA-hour cm}^{-2}$) (Fig. 2A, black dashed curve) and infinitely higher compared with a Ni electrode in an aprotic cell (Fig. 2A, red dashed line). After fully discharging the cell to 2.6 V, the x-ray diffraction (XRD) pattern of the composite cathode shows two peaks at 34° and 56° assigned to the (111) and (022) reflections of Li_2O [Joint Committee on Powder Diffraction Standards (JCPDS) 01-077-2144] (Fig. 2B). A Raman band definitive of Li_2O at 523 cm^{-1} further supports formation of the oxide (Fig. 2C). The scanning electron microscopy (SEM) image in Fig. 2E reveals that the discharged cathode is covered with large $\sim 5\text{-}\mu\text{m}$ octahedral crystals, a morphology characteristic of the Li_2O antifluorite structure. Because the solubility of Li_2O in molten nitrate is 27 mM at 150°C (fig. S5), we speculate that

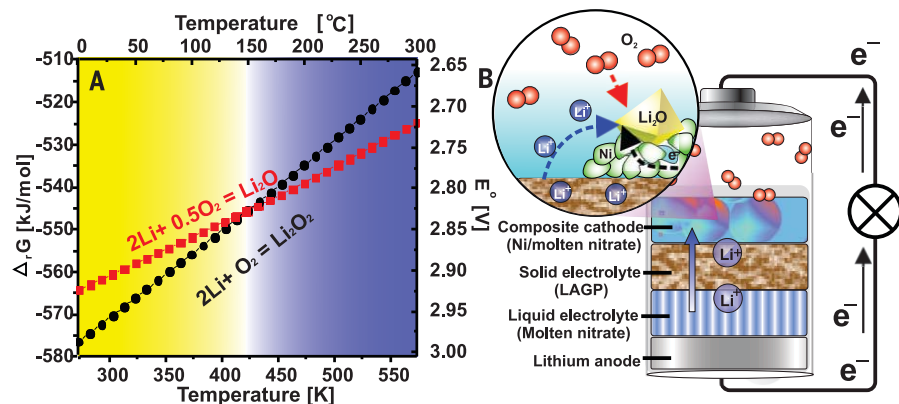
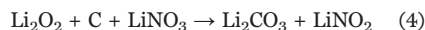


Fig. 1. Thermodynamics and configuration of the Li-O_2 cell. (A) Gibbs reaction energy for formation of Li_2O and Li_2O_2 as a function of temperature. The thermodynamic data were calculated according to the database of HSC chemistry version 5. (B) Configuration of the inorganic electrolyte Li-O_2 cell and schematic illustration of Li_2O formation during discharge.

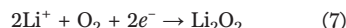
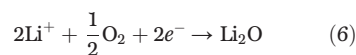
Department of Chemistry and the Waterloo Institute of Nanotechnology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

*Corresponding author. Email: lfnazar@uwaterloo.ca

solution-mediated Li_2O transport—as reported in aprotic $\text{Li}-\text{O}_2$ and $\text{Na}-\text{O}_2$ cells (13, 14)—is responsible for crystal nucleation and growth. A high initial coulombic efficiency (CE) of 96% was achieved after recharging the cell to 3.5 V, accompanied by a very low polarization of 0.2 V. The disappearance of Li_2O in the XRD pattern (Fig. 2B) and Raman spectrum (Fig. 2C) of the recharged cathode indicates Li_2O is fully removed by oxidation. Furthermore, the charged cathode is bare (Fig. 2F, SEM), identical to before discharge (Fig. 2D), indicating excellent electrochemical reversibility. There is a gravimetric energy penalty owing to the higher mass of Ni compared with that of carbon, but this can be improved by optimizing the cathode microstructure and mass. By contrast, the inorganic electrolyte $\text{Li}-\text{O}_2$ cell by using a cathode composed of Super P carbon—run at the same temperature of 150°C —shows about half the discharge capacity ($6.5 \text{ mA}\cdot\text{hour cm}^{-2}$) (Fig. 2A, black solid line). Li_2O_2 is identified as the main discharge product in the carbon electrode cell according to XRD and Raman analysis (Fig. 2, B and C, black curves) and is present as hexagonal-shaped Li_2O_2 crystal agglomerates $\sim 8 \mu\text{m}$ in dimension (Fig. 2G). We assume that the quinone groups on carbon are responsible for peroxide formation, as reported (15). Crystalline Li_2CO_3 and amorphous Li_2O were also identified with XRD and x-ray photoelectron spectroscopy (XPS) analysis, respectively (fig. S4). These by-products are attributed to carbon corrosion by Li_2O_2 as reported in aprotic $\text{Li}-\text{O}_2$ cells (Eq. 3) (16) and to the reaction of Li_2O_2 with nitrite (Eqs. 4 and 5). These electrochemically inert by-products lower the cell rechargeability, evidenced by a low CE of only 75% on the first cycle.



To further examine the electrochemical reversibility, quantitative compositional changes of the cathodic products at the 1st and 10th cycle were analyzed (fig. S6). Cells with either carbon or the Ni-nitrate composite cathode were discharged to 2 $\text{mA}\cdot\text{hour cm}^{-2}$ and recharged to 3.5 V. Li_2O is the product in the composite cathode on discharge ($36.1 \mu\text{mol}$, as quantified by means of acid-base titration) (Fig. 3A and supplementary materials, section S6-3), along with a tiny fraction of Li_2O_2 ($1.1 \mu\text{mol}$). The total amount of Li_2O and Li_2O_2 ($37.2 \mu\text{mol}$) is nearly identical to the theoretical value of $37.3 \mu\text{mol}$, assuming a $2e^-$ transfer per mole of products (Eqs. 6 and 7)



After recharging the cell to 3.5 V, the amount of Li_2O on the cathode is reduced to $2.6 \mu\text{mol}$,

whereas no Li_2O_2 is observed. The residual Li_2O likely arises from its low solubility in the molten nitrate electrolyte and crosses over to the electrically insulating LAGP membrane, rendering it electrochemically inaccessible. There are no notable changes in the product quantity at the 10th cycle, indicating a reversible cathodic reaction. By contrast, the $\text{Li}-\text{O}_2$ cell using a carbon cathode presents poor chemical reversibility (Fig. 3B).

On discharge, in addition to the major product (Li_2O_2 ; $22.1 \mu\text{mol}$), Li_2CO_3 ($4.37 \mu\text{mol}$) and Li_2O ($12.4 \mu\text{mol}$) also form. After recharging the carbon cathode to 3.5 V, Li_2O_2 is fully oxidized, whereas Li_2CO_3 and Li_2O remain. On the basis of mass spectrometry analysis of model cathodes, formation of Li_2CO_3 and Li_2O is likely electrochemically irreversible below 4 V in the absence of an oxygen evolution reaction (OER) catalyst

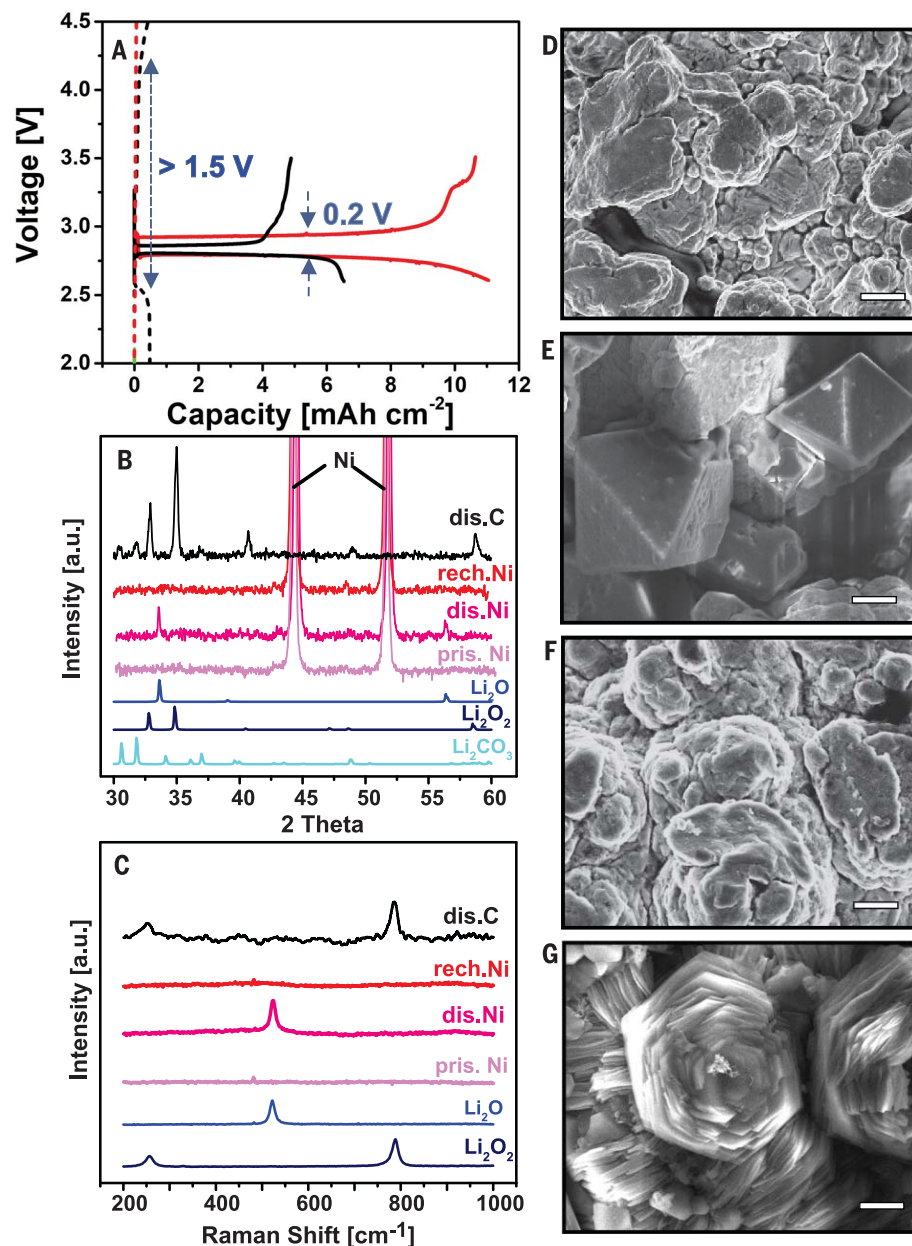
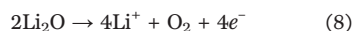


Fig. 2. Characteristics of $\text{Li}-\text{O}_2$ cells using carbon and Ni-nitrate composite cathodes.

(A) First discharge and charge curves of $\text{Li}-\text{O}_2$ cells with a carbon cathode (black) and a Ni-based cathode (red). The cells using aprotic electrolyte [0.5 M lithium bis(trifluoromethanesulfonyl) imide (LiTFSI) in tetraethylene glycol dimethyl ether (TEGDME)] were examined at 25°C (dashed lines), whereas the cells using the molten nitrate electrolyte were measured at 150°C (solid lines). The current density is 0.1 mA cm^{-2} , and cutoff voltages are 2.6 and 3.5 V. (B) XRD patterns, (C) Raman spectra, and (D to G) SEM images of (D) pristine Ni cathode, (E) Ni cathode discharged to 2.6 V then (F) recharged to 3.5 V, and (G) carbon cathode discharged to 2.6 V. Scale bars, 2 μm .

(fig. S7). These by-products accumulate during cycling, leading to cathode passivation (fig. S8). Online electrochemical mass spectrometry (OEMS) monitoring of the gaseous products formed during the 1st and 10th charge showed no other signals aside from O_2 (such as CO_2 , NO , and H_2O) below 3.5 V. The first charge profile of the $Li-O_2$ cell with the Ni-nitrate cathode exhibits two plateaus at 3.0 and 3.3 V (Fig. 3C). The rate of O_2 evolu-

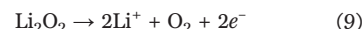
tion on either plateau is exactly equal to the theoretical value based on $4e^-/O_2$, indicating the electrochemical oxidation of Li_2O (Eq. 8)



We speculate that the first plateau at 3.0 V is due to the oxidation of Li_2O crystallites that are

deposited near or on the Ni catalyst at the cathode, whereas the higher plateau at 3.3 V (whose disappearance at the 10th charge is not well understood at present) corresponds to the oxidation of Li_2O deposited on LAGP. In the latter case, the soluble oxide must diffuse back to the Ni particles for OER, creating a kinetic overpotential as previously noted in $Na-O_2$ cells (17). Moreover, rechargeability is improved on cycling as indicated by more oxygen evolution and a more prolonged charge plateau at 3.0 V, with a CE of 100%.

By contrast, the carbon cathode $Li-O_2$ cell exhibits a voltage plateau at ~2.9 V on the 1st charge, where OEMS shows that oxygen is initially evolved at a rate close to theoretical for Li_2O_2 oxidation based on $2e^-/O_2$ (Eq. 9). The voltage is close to expected at 150°C (2.84 V)



Subsequently, the rate of oxygen evolution greatly diminishes and is followed by a false “peak” as a consequence of the voltage cutoff (3.5 V). This behavior is indicative of substantial parasitic reactions that hinder the OER process (18). Passivation of the carbon cathode results in an increase of cell overpotential at the 10th charge (Fig. 3B).

The Ni-nitrate cathode $Li-O_2$ cell was cycled with a limited capacity of $0.5 \text{ mA}\cdot\text{hour cm}^{-2}$ at 0.2 mA cm^{-2} for 150 cycles (Fig. 4A), with low overpotential. At a lower current density of 0.1 mA cm^{-2} (fig. S9), the polarization is further reduced (0.16 V; comparable with Fig. 2A). Shallow cycling was necessary to limit the amount of Li transfer at the negative electrode in order to provide proof of concept of the rechargeability at the cathode. The cell exhibits two discharge plateaus, the first at ~2.9 V corresponding to a very low capacity (~0.1 $\text{mA}\cdot\text{hour cm}^{-2}$), followed by a much longer discharge plateau at ~2.7 V. The Ni-nitrate cathode cell at a much deeper state of discharge of $11 \text{ mA}\cdot\text{hour cm}^{-2}$ exhibits a similar electrochemical profile, although the first short plateau is masked by the large discharge capacity (Fig. 2A). We ascribe the first plateau in Fig. 4A to initial formation of the Li_2O species in solution, which has a different formation energy than that of solid Li_2O . Phase field simulations of $Li-O_2$ electrochemistry report that a higher potential results from nucleation of a supersaturated solution of solvated species before growth of solid products (19). The longer plateau at 2.7 V corresponds to oxygen reduction at the surface of the catalyst to form Li_2O . Upon charge, the cell also exhibits two charge plateaus at ~3.0 and ~3.3 V, which is in agreement with the OEMS measurement (Fig. 3D) and the deep-cycled Ni-nitrate cathode in Fig. 2A. The CE of the cell increases rapidly from 80 to 100% over four cycles and is subsequently stable (Fig. 4B). This performance is superior to that of aprotic organic electrolyte $Li-O_2$ cells and their cousins ($Na-O_2$ and $K-O_2$) (14, 20), where parasitic reactions cause poor cycling performance: A monolayer of carbonate created at the $C-Li_2O_2$ interface causes

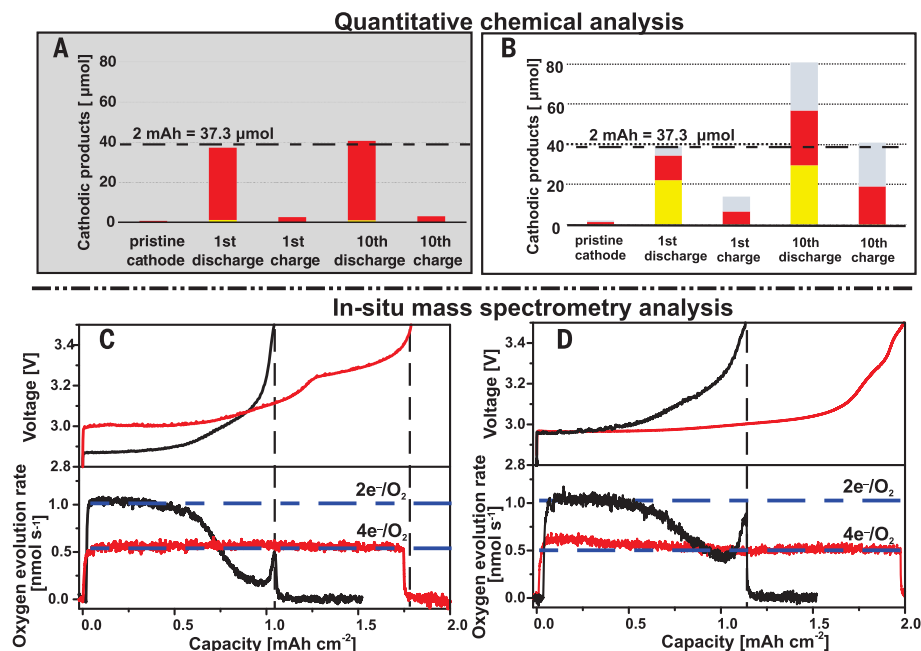


Fig. 3. High reversibility of molten salt electrolyte $Li-O_2$ cells using a Ni-nitrate composite cathode. (A and B) Quantitative analysis of cathodic products. Shown are Li_2O (red), Li_2O_2 (yellow), and Li_2CO_3 (gray), at the 1st and 10th cycle using (A) Ni-nitrate composite cathode and (B) carbon cathode. The discharge limit is $2 \text{ mA}\cdot\text{hour cm}^{-2}$, and the charge limit is 3.5 V. The current density is 0.1 mA cm^{-2} . (C and D) Online mass spectrometry analysis of gaseous oxygen evolution upon (C) 1st charge and (D) 10th charge using a carbon cathode (black) and a Ni-based composite cathode (red). Cells were predischarged to $2 \text{ mA}\cdot\text{hour cm}^{-2}$ at 0.1 mA cm^{-2} . The charging rate was 0.2 mA cm^{-2} in order to enhance the O_2 mass spectrometry signal.

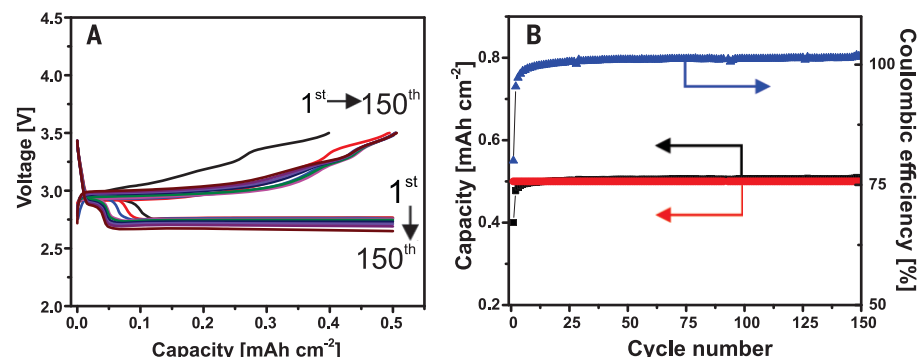


Fig. 4. Cycling performance of a molten salt electrolyte $Li-O_2$ cell with a Ni-nitrate composite cathode. (A) Discharge and charge curves over 150 cycles at a constant current of 0.2 mA cm^{-2} . (B) The corresponding changes of discharge capacity (red), charge capacity (black), and CE (blue) during cycling. The CE ultimately reaches 100.8% because a 3.5 V cutoff was chosen as the charge limit, whereas O_2 evolution stops at a charge voltage of 3.48 V (Fig. 3D); the slight additional charge capacity can be ascribed to a trace of nitrate decomposition.

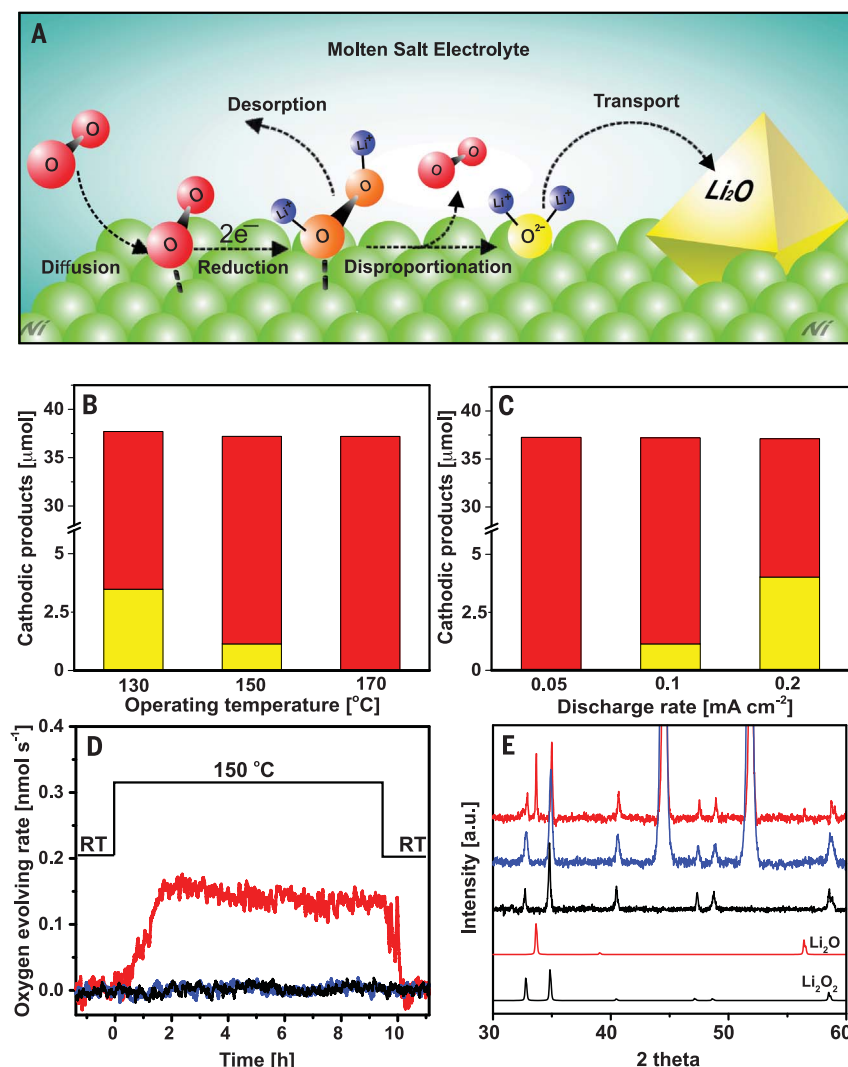


Fig. 5. Mechanistic studies of the oxygen reduction reaction over the Ni-based composite catalyst. (A) Schematic illustration of the pathway of the oxygen reduction reaction. (B and C) Effects of (B) operating temperature and (C) discharge rates on the composition of oxygen reduction products, Li₂O (red) and Li₂O₂ (yellow). A discharge rate of 0.1 mA cm⁻² was applied to study the temperature effects, and 150 °C was chosen to investigate the effects of discharge rate. All cells were discharged to 2 mA-hour cm⁻² before quantitative analysis. (D) Oxygen mass spectrometry signal response on heating a mixture composed of commercial Li₂O₂ powder with Ni-nitrate composite cathode (red), Ni (blue), and molten nitrate (black) at 150 °C, respectively. (E) Their corresponding XRD patterns were obtained after heating the mixtures for 1 week.

an increase of interfacial resistance (16), and decomposition of organic electrolyte is triggered by a superoxide attack that forms carbon-centered radicals (21). By using inorganic electrolytes, electrolyte degradation is avoided. A layer of lithiated nickel (III) oxide and Ni₂O₃ on the surface of the Ni particles (because of the oxidation of Ni by LiNO₃) (fig. S3) provides a protective passivation layer; meanwhile, lithiation improves the electronic conductivity of the oxide layer (22). Hence, the chemically stable inorganic electrolyte and cathode play critical and synergistic roles. Although we used a Li⁺-ion conducting interface between the Li and LAGP in

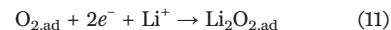
order to prevent direct reduction of the LAGP (23), degradation of the LAGP membrane over 150 cycles—arising from its probable poor stability in molten nitrate and some localized reduction (fig. S10, A and B)—led to increased impedance from the membrane (fig. S10C). This resulted in an increase of the overpotential by 0.15 V over 150 cycles. More effort is needed to address the challenges on the negative electrode side, such as the development of robust polymer-inorganic membranes.

To explain the mechanism, we propose a peroxide-mediated ORR pathway, illustrated in Fig. 5A and outlined below.

Diffusion: Upon discharge, oxygen adsorbs on the surface of the cathode (Eq. 10)



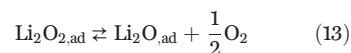
Reduction: Oxygen is electrochemically reduced to form lithium peroxide (Li₂O_{2,ad}), via a two-electron transfer, on the surface of the Ni/Li_xNiO₂ catalyst (Eq. 11)



Desorption: A small amount of Li₂O_{2,ad} slowly desorbs from the catalyst surface, governed by its low solubility and diffusibility in the molten nitrate electrolyte (Eq. 12)



Disproportionation: The major remaining Li₂O₂ is converted to Li₂O by the catalyst through disproportionation (Eq. 13)



Transport: Once formed, Li₂O is soluble in the electrolyte (Eq. 14), and on supersaturation formation of Li₂O nuclei triggers nucleation and growth to result in micrometer-sized Li₂O crystals following (24)



Experimental and computational studies suggest that the ORR pathway via peroxide is operative over a variety of metal catalysts (such as platinum, mercury, and silver) (25, 26). We quantified the amount of Li₂O₂ formed at different discharge conditions after discharge to 2 mA-hour cm⁻². As shown in Fig. 5, B and C, the amount of Li₂O₂ increases from 0 μmol at 0.05 mA cm⁻² to 4 μmol at 0.2 mA cm⁻², whereas it decreases from 3.5 μmol at 135 °C to 0 μmol at 170 °C. According to the proposed ORR pathway, the fast formation of Li₂O₂ at higher discharge rates likely results in some Li₂O₂ remaining owing to relatively slow disproportionation (rate-determining step). However, the elevated temperature accelerates disproportionation, rapidly converting Li₂O₂ to Li₂O. Evidence confirming the catalytic disproportionation of Li₂O₂ is shown in Fig. 5, D and E. When a mixture composed of commercial Li₂O₂ powder and the Ni-nitrate cathode was heated at 150 °C, oxygen evolution was detected with mass spectrometry, accompanied by diffraction peaks of Li₂O in the XRD pattern of the mixture. Neither features are observed in the absence of either component.

Suntivich *et al.* have shown that high ORR activity for transition metal oxide catalysts primarily correlates to σ*-orbital (*e_g*) occupation and the extent of transition-metal-oxygen covalency (27). Optimal activity correlates to an *e_g* occupancy is close to unity. Consistent with this design principle, the surface Li_xNiO₂ species on the Ni particle contains Ni³⁺ with an electron configuration of *t_{2g}⁶e_g¹* (28). Furthermore, the Ni³⁺/Ni²⁺

redox couple promotes charge transfer between surface cations and adsorbates. Both factors give rise to a high ORR activity of Li_xNiO_2 (29). Although the thermodynamic driving force for the disproportionation of Li_2O_2 is very small (only $-0.063 \text{ kJ mol}^{-1}$ at 150°C), the removal of Li_2O from the catalyst surface via solution-mediated transport will shift the equilibrium (Eq. 13) toward the right, exposing the active catalyst surface. The spontaneous disproportionation reaction on oxygen reduction dictates that OER must follow a different pathway, however. Indeed, no Li_2O_2 intermediate is observed when charging a Li-O_2 cell by using a Li_2O -prefilled cathode (fig. S11). We speculate that Li_2O is solubilized and diffuses to the surface of Li_xNiO_2 for electrocatalytic oxidation via a direct $4 e^-$ pathway. Suntivich *et al.* also concluded that the OER activity of metal oxides, similar to ORR, is dependent on the occupancy of 3d electron states with e_g symmetry (30). Thus, although Li_xNiO_2 is an effective ORR catalyst, it also reversibly catalyzes OER (31), leading to a low charge overpotential.

By tuning the operating temperature and using a single bifunctional ORR/OER catalyst, our Li-O_2 battery overcomes the barriers of thermodynamics and kinetics, leading to the electrochemically reversible formation of Li_2O instead of Li_2O_2 . The in situ-generated Li_xNiO_2 electrocatalyst—applicable to other metal oxygen electrochemistries—catalyzes both O–O cleavage to form Li_2O on discharge, and the reverse process that releases oxygen upon charge in a $4e^-/\text{O}_2$ process with excellent CE and low polarization. The latter is aided by electrolyte-solubilized Li_2O transfer. The fact that the Li-O cell is more reversible when Li_2O is the product is a consequence of the less reactive chemical nature of oxide versus superoxide or peroxide. Moreover, the use of chemically stable inorganic electrolytes and a noncarbonaceous cathode circumvents the degradation of organic electrolyte and carbon

corrosion, which form the main failure mechanisms for nonaqueous Li-O_2 cells. The “ Li_2O ” cell presented here is akin to both fuel cells and electrolyzers—which also operate on the basis of a $4 e^-$ electrocatalyzed reaction—in which Li_2O replaces H_2O , and the combination forms a simple reversible energy storage system. Whereas elevated operating temperatures can limit battery applications, commercialized sodium nickel chloride (ZEBRA) cells and Na-S batteries run at much higher temperatures ($\sim 325^\circ\text{C}$), and proton-exchange membrane (PEM) fuel cells are recently trending to temperatures between 120° and 180°C (32). More fundamentally though, our work directly addresses a number of issues associated with Li-O_2 chemistry, showing that it is not intrinsically limited and that four-electron transfer from Li_2O is possible, reversible, and operates with almost theoretical CE.

REFERENCES AND NOTES

1. Z. Peng, S. A. Freunberger, Y. Chen, P. G. Bruce, *Science* **337**, 563–566 (2012).
2. Y. C. Lu *et al.*, *Energy Environ. Sci.* **6**, 750–768 (2013).
3. P. Hartmann *et al.*, *Nat. Mater.* **12**, 228–232 (2013).
4. D. Sharon *et al.*, *Isr. J. Chem.* **55**, 508–520 (2015).
5. D. Aurbach, B. D. McCloskey, L. F. Nazar, P. G. Bruce, *Nat. Energy* **1**, 16128 (2016).
6. J. Lu *et al.*, *Nature* **529**, 377–382 (2016).
7. T. Liu *et al.*, *Science* **350**, 530–533 (2015).
8. C. M. Burke *et al.*, *ACS Energy Lett.* **1**, 747–756 (2016).
9. D. G. Kwabi *et al.*, *J. Phys. Chem. Lett.* **5**, 2850–2856 (2014).
10. R. Black *et al.*, *J. Am. Chem. Soc.* **134**, 2902–2905 (2012).
11. C. L. Bender, P. Hartmann, M. Vračar, P. Adelhelm, J. Janek, *Adv. Energy Mater.* **4**, 1301863 (2014).
12. V. Giordani *et al.*, *J. Am. Chem. Soc.* **138**, 2656–2663 (2016).
13. N. B. Aetukuri *et al.*, *Nat. Chem.* **7**, 50–56 (2015).
14. C. Xia, R. Black, R. Fernandes, B. Adams, L. F. Nazar, *Nat. Chem.* **7**, 496–501 (2015).
15. V. A. Garten, D. E. Weiss, *Rev. Pure Appl. Chem.* **7**, 69–122 (1957).
16. B. D. McCloskey *et al.*, *J. Phys. Chem. Lett.* **3**, 997–1001 (2012).
17. P. Hartmann *et al.*, *J. Phys. Chem. C* **119**, 22778–22786 (2015).
18. B. D. McCloskey *et al.*, *J. Phys. Chem. Lett.* **3**, 3043–3047 (2012).
19. M. J. Welland *et al.*, *J. Chem. Phys.* **143**, 224113 (2015).
20. B. D. Adams *et al.*, *Adv. Energy Mater.* **5**, 1400867 (2015).
21. C. Xia *et al.*, *J. Am. Chem. Soc.* **138**, 11219–11226 (2016).
22. W. Chia-Ching, Y. Cheng-Fu, *Nanoscale Res. Lett.* **8**, 33 (2013).
23. P. Hartmann *et al.*, *J. Phys. Chem. C* **117**, 21064–21074 (2013).
24. B. D. Adams *et al.*, *Energy Environ. Sci.* **6**, 1772–1778 (2013).
25. N. Markovic, *Surf. Sci. Rep.* **45**, 117–229 (2002).
26. C. M. Sánchez-Sánchez, A. J. Bard, *Anal. Chem.* **81**, 8094–8100 (2009).
27. J. Suntivich *et al.*, *Nat. Chem.* **3**, 546–550 (2011).
28. Y. Koyama, T. Mizoguchi, H. Ikeno, I. Tanaka, *J. Phys. Chem. B* **109**, 10749–10755 (2005).
29. S. Srinivasan, *Fuel Cells: From Fundamentals to Applications* (Springer, 2010).
30. J. Suntivich, K. J. May, H. A. Gasteiger, J. B. Goodenough, Y. Shao-Horn, *Science* **334**, 1383–1385 (2011).
31. Z. W. Seh *et al.*, *Science* **355**, eaad4998 (2017).
32. M. E. Scofield, H. Liu, S. S. Wong, *Chem. Soc. Rev.* **44**, 5836–5860 (2015).

ACKNOWLEDGMENTS

The authors thank V. Goodfellow and R. Smith at the University of Waterloo Mass Spectrometry Facility for their scientific input in the gas chromatography–mass spectrometry measurements. We also thank S. H. Vajargah for performing the TEM measurements. **Funding:** Research was supported by the Natural Sciences and Engineering Council of Canada through their Discovery and Canada Research Chair programs (L.F.N.), and a doctoral scholarship to C.Y.K. Partial funding for this work (C.Y.K.) was also provided by the Joint Center for Energy Storage Research, an Energy Innovation Hub funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences. **Author contributions:** C.X. led the design of the study, and all the authors contributed to the implementation and writing of the manuscript; data collection and analysis were conducted by C.X. and C.Y.K. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data are available in the manuscript or in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/777/suppl/DC1
Materials and Methods
Figs. S1 to S11
Table S1
References (33–40)

8 January 2018; accepted 10 July 2018
10.1126/science.aas9343

QUASICRYSTALS

Dirac electrons in a dodecagonal graphene quasicrystal

Sung Joon Ahn^{1*}, Pilkyung Moon^{2,3*}, Tae-Hoon Kim^{4*}, Hyun-Woo Kim¹, Ha-Chul Shin¹, Eun Hye Kim¹, Hyun Woo Cha⁴, Se-Jong Kahng⁵, Philip Kim⁶, Mikito Koshino⁷, Young-Woo Son^{8,†}, Cheol-Woong Yang^{4,†}, Joung Real Ahn^{1,9,†}

Quantum states of quasiparticles in solids are dictated by symmetry. We have experimentally demonstrated quantum states of Dirac electrons in a two-dimensional quasicrystal without translational symmetry. A dodecagonal quasicrystalline order was realized by epitaxial growth of twisted bilayer graphene rotated exactly 30°. We grew the graphene quasicrystal up to a millimeter scale on a silicon carbide surface while maintaining the single rotation angle over an entire sample and successfully isolated the quasicrystal from a substrate, demonstrating its structural and chemical stability under ambient conditions. Multiple Dirac cones replicated with the 12-fold rotational symmetry were observed in angle-resolved photoemission spectra, which revealed anomalous strong interlayer coupling with quasi-periodicity. Our study provides a way to explore physical properties of relativistic fermions with controllable quasicrystalline orders.

Quasicrystals, which can have quasi-periodic orders (such as rotational symmetry) without spatial periodicity, have been used to study quantum states between the limits of periodic order and disorder (1–15). The study of the influence of quasi-periodic order has focused on extended wave functions of ordered states, pseudogaps, and fine structures of density of states (16–28). However, these studies have been limited to nonrelativistic fermions. We have found experimentally that a relativistic Dirac fermion quasicrystal can be realized when the Dirac electrons in a single-layer graphene are incommensurately modulated by another single-layer graphene that is rotated by exactly 30°. Such a twisted bilayer graphene (TBG) can form a 12-fold quasicrystalline order through interactions between the two layers.

This TBG quasicrystal is a direct material realization of the algorithmic quasicrystal construction process first proposed by Stampfli, with two ideal hexagonal grids rotated 30° with respect to each other (29, 30). We observed that the graphene quasicrystal has distinct Dirac cone replicas in the Brillouin zone. Furthermore, it shows characteristic electronic structures that

differ from the periodic TBG crystals with rotation angles near 30°. Angle-resolved photoemission spectroscopy (ARPES) revealed many replicas of Dirac cones with 12-fold rotational symmetry, indicating large enhancement of interlayer coupling with the quasicrystalline order. We show that the graphene quasicrystal can be used to explore quantum states of Dirac electrons in quasi-periodic order, just as a single-layer graphene has served as the quintessential material for studying Dirac fermions in periodic crystals. Furthermore, it has been predicted theoretically that four-dimensional (4D) quantum Hall effects of Dirac fermions can be studied in the 2D quasicrystal (31). Thus, the experimental achievement of the quasicrystal of Dirac fermions in two dimensions can be a starting point to approach higher-dimensional physics of quasiparticles in solids.

We grew graphene quasicrystals on the Si face of the 4H-SiC (0001) surface [see figs. S1 and S2 in supplementary materials (SM)]. The upper- and lower-layer graphene of the TBG had rotational angles of 0° and 30° (hereafter referred to as R0° and R30°), respectively, relative to the orientation of the SiC (0001) surface. Usually, when graphene is thermally grown on a Si face of 4H- or 6H-SiC (0001), the graphene assumes the R30° configuration (32–35). In contrast, by using borazine gas in a vacuum system, a hexagonal boron nitride (h-BN) layer with R0° can be grown epitaxially at 1050°C, where the typical π and σ energy bands of h-BN were observed in the previous ARPES experiments (fig. S1) (36). Thus, we first grew a h-BN layer with R0° by using borazine gas at 1050°C and then heated the system at a higher temperature of 1600°C to replace h-BN with a graphene layer while maintaining the same angle of 0° (fig. S1) (36). A buffer layer exists between graphene with R0° and bulk SiC, where the buffer layer has the same atomic structure as single-layer graphene but is insulating because some carbon atoms bond to underlying

Si atoms (33). This procedure enabled the formation of the upper layer of the graphene quasicrystal (36). Finally, when the sample was heated further at 1600°C, a graphene layer with R30° grew between the first graphene layer with R0° and the SiC (0001) substrate, resulting in TBG with an exact rotational angle of 30° (fig. S2, B, D, and F). In this case, a buffer layer exists between the graphene quasicrystal and bulk SiC (fig. S2, B, D, and F). Another method can be used to grow the graphene quasicrystal. When hydrogen atoms are intercalated between the buffer layer underlying graphene with R0° and bulk SiC, the buffer layer changes into graphene with R30° (fig. S2, C, E, and G) (33). The hydrogen intercalation results in the same graphene quasicrystal on a hydrogen-terminated SiC surface (33). In this case, there is no buffer layer between the graphene quasicrystal and bulk SiC (fig. S2, C, E, and G). The nonexistence of boron and nitrogen atoms in the graphene quasicrystal was proved by x-ray photoemission spectroscopy (fig. S3). The Si 2p and C 1s spectra in fig. S3 suggest that the buffer layer in this system is the same as the buffer layer of epitaxial graphene with R30° on SiC (32, 33). The present epitaxial growth has merit over a manual transfer (37–39) in making the quasicrystal because of the inevitable errors in twisting angles as well as the local distortions during the transfer process.

We determined the crystal orientation of the graphene layers from low-energy electron diffraction (LEED) and transmission electron microscopy (TEM) measurements (Fig. 1 and figs. S4 and S5). The upper and lower layers of the TBG can be clearly distinguished from the LEED pattern (Fig. 1A) because the intensity of the 1×1 LEED pattern of the upper-layer graphene (the red hexagon in Fig. 1A) is higher than that of the lower-layer graphene (the blue hexagon in Fig. 1A). The LEED spots other than the 1×1 LEED patterns come from the quasi-periodic order induced by the interlayer interaction. As shown in Fig. 1B, our observation is fully consistent with the simulated pattern from an atomic structure model of graphene quasicrystal (see also fig. S6). By using the 1-mm by 1-mm electron beam, we also confirmed that the LEED pattern is uniform throughout the entire sample size of up to 4 mm by 7 mm, where the sample size is limited only by our experimental setup and can be scaled up to a wafer scale (fig. S5). Furthermore, the coherent length of the graphene quasicrystal was similar to that of typical epitaxial graphene with R30° grown on SiC (fig. S4).

The graphene quasicrystal can be spatially mapped onto a quasicrystal lattice model constructed by dodecagonal compound tessellations (Fig. 1, D and E, and figs. S7, D and E; S8; and S9) (29, 30). Squares, rhombuses, and equilateral triangles with different orientations can fill the entire space with a 12-fold rotationally symmetric pattern without translational symmetry. Because the Stampfli tiles have a fractal structure with self-similarity, the same pattern emerges at a larger scale with an irrational scaling factor. For the graphene quasicrystal, the Stampfli tiles have

¹Department of Physics and SAINT, Sungkyunkwan University, Suwon, Republic of Korea. ²New York University and NYU-ECNU Institute of Physics at NYU Shanghai, Shanghai, China. ³Department of Physics, New York University, New York, NY, USA. ⁴School of Advanced Materials Science and Engineering, Sungkyunkwan University, Suwon, Republic of Korea. ⁵Department of Physics, Korea University, Seoul, Republic of Korea. ⁶Department of Physics, Harvard University, Cambridge, MA, USA. ⁷Department of Physics, Osaka University, Machikaneyama, Toyonaka, Japan. ⁸Korea Institute for Advanced Study, Seoul, Republic of Korea. ⁹Samsung-SKKU Graphene Center, Sungkyunkwan University, Suwon, Republic of Korea.

*These authors contributed equally to this work.

†Corresponding author. Email: jrahn@skku.edu (J.R.A.); cwyang@skku.edu (C.-W.Y.); hand@kias.re.kr (Y.-W.S.)

the scaling factor of $\sqrt{2 + \sqrt{3}}$ (Fig. 1, D and E, and figs. S7 to S9) (29). As shown in the atomic model (Fig. 1, C and D), the graphene quasicrystal results in Stampfli tiles such as equilateral triangles and rhombuses. In the false-colored TEM image of the graphene quasicrystal transferred from a SiC wafer to a TEM grid (Fig. 1F and fig. S7D), the Stampfli tiles, including squares (blue), rhombuses (red), and equilateral triangles (green) with different orientations, are clearly observed. The LEED pattern and TEM image clearly indicate that the TBG with a rotational angle of 30° has a quasicrystalline order with 12-fold rotational symmetry. The detailed TEM images match the Stampfli tiling very well, with a few minute differences between them (figs. S8 to S12). The slight differences originate mainly from different internal structures of Stampfli tiles due to the different radial distances from the center (figs. S10 to S12). In addition, because our TEM image was taken for our transferred sample on a TEM grid, the height of the freestanding graphene could vary in space slightly, thus defocusing the image locally. The successful TEM experiments show that the graphene quasicrystal can be isolated from a substrate and is chemically and structurally stable at room temperature in air.

The characteristic electronic structures were determined from ARPES measurements. Figure 2A shows the constant energy map of ARPES spectra of the graphene quasicrystal, and Fig. 2C shows the enlarged view in the vicinity of the zero crystal momentum denoted by the Γ point (see also fig. S13). Several replicas of main Dirac cones respecting the 12-fold rotational symmetry were observed throughout the momentum space (Fig. 2), in sharp contrast to the usual TBGs with an arbitrary rotational angle other than 30° (40, 41). Considering the large photon beam size of 3 mm by 3 mm in our ARPES system, the energy-momentum dispersions throughout the entire momentum space ensure the uniform sample quality, as in the LEED experiments. Like the LEED patterns, the intensities of the photoemission spectra of the lower graphene layer are weaker than those of the upper graphene layer (Fig. 2A) because of attenuated intensities for photoemitted electrons from the lower layer.

To understand the origin of the 12-fold Dirac cone replicas, we performed a numerical simulation of the ARPES spectra by using a single-particle tight-binding model that has been successfully used for describing various electronic properties of TBGs of general angles (42, 43) (see SM for the detailed methods). Figure 3A shows the constant energy map at the Fermi energy (i.e., 0.3 eV above the Dirac point) of simulated ARPES spectra. The red (blue) contours indicate that those Dirac cones originate mainly from inter- and intralayer scatterings of the wave functions of the upper (lower) graphene layer. In stacked crystals, a state $|\mathbf{k}\rangle$ in the upper layer is scattered to every $|\mathbf{k}'\rangle$ state in the lower layer, which satisfies the generalized Umklapp scattering condition of

$\mathbf{k} + \mathbf{G} = \mathbf{k}' + \mathbf{G}'$. Here, \mathbf{k} and \mathbf{G} (\mathbf{k}' and \mathbf{G}') are the in-plane wave vector and reciprocal vector of the upper (lower) layer, respectively (44, 45).

The coupling strength between $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$ is proportional to $t(\mathbf{k} + \mathbf{G}) = t(\mathbf{k}' + \mathbf{G}')$, where $t(\mathbf{q})$ is the in-plane Fourier transform of the interlayer transfer integral and where \mathbf{q} ($= \mathbf{k} + \mathbf{G} = \mathbf{k}' + \mathbf{G}'$) is the equivalent wave vector common to both layers (44). Figure 3, B to D, shows some examples of the Umklapp scattering from the Dirac point ($\mathbf{k} = \mathbf{K}$) of the upper layer, which shows that the scattering with different vectors \mathbf{G} is mapped to distinct wave vectors \mathbf{k}' in the lower layer. We can also check this finding given that, though the states

$|\mathbf{k} + \mathbf{G}\rangle$ have the same wave function at the lattice point in the upper layer regardless of \mathbf{G} , the states $|\mathbf{k}' + \mathbf{G}'\rangle$ which satisfy $\mathbf{k}' + \mathbf{G}' = \mathbf{k} + \mathbf{G}$ give distinct wave functions in the lower layer for different \mathbf{G} (see SM for a detailed explanation). Each scattering event makes a replica of the Dirac cone (originally at $\mathbf{k} = \mathbf{K}$) at \mathbf{k}' . Because \mathbf{G} and \mathbf{G}' are neither identical nor commensurate in graphene quasicrystal, there are many distinct Dirac cone replicas in the momentum space. Likewise, the Dirac cone at \mathbf{k}' of the lower layer is replicated at \mathbf{k} of the upper layer, satisfying $\mathbf{k} + \mathbf{G} = \mathbf{k}' + \mathbf{G}'$.

From these calculations, the observed positions of Dirac cones in our ARPES experiment

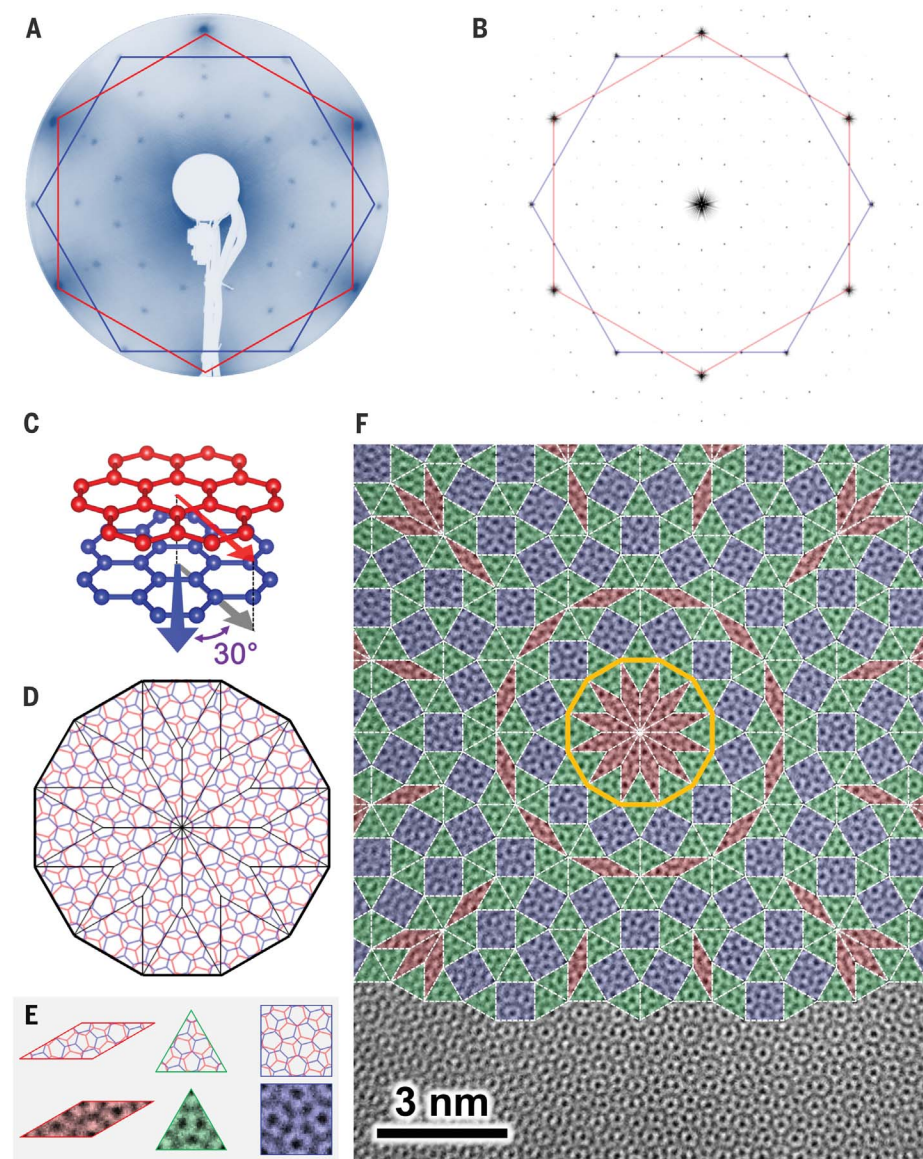


Fig. 1. A LEED pattern and a TEM image of graphene quasicrystal. (A) A LEED pattern of graphene quasicrystal. **(B)** A Fourier-transformed pattern of graphene quasicrystal (see also fig. S6 in SM). **(C and D)** An atomic structure model of TBG with R30°. **(E)** Atomic structures and TEM images of Stampfli tiles [rhombuses (red), equilateral triangles (green), and squares (blue)]. **(F)** A false-colored TEM image of graphene quasicrystal mapped with 12-fold Stampfli-inflation tiling.

are reproduced in our simulations as shown in Fig. 3, E and F. The observed positions, however, cannot be reproduced if the angle deviates from 30° very slightly (fig. S13, D and E). For example, if the twisted angle is 29.9576° , the TBG has a translational symmetry with a periodic supercell that is 1351 times the graphene unit cell, and the computed positions of the Dirac cone replicas in this quasicrystal approximant were not consistent with the experimental ones (fig. S13E).

Close inspection of ARPES intensities of Dirac cone replicas highlights the important role of higher-order Umklapp scatterings in realizing 12-fold Dirac cones in graphene quasicrystals. Within a single Umklapp scattering process, the scattering strength is proportional to the wave vector \mathbf{q} component of the in-plane Fourier transform of the interlayer transfer integral $t(\mathbf{q})$. Figure 3, B to D, shows the examples of the scattering with the shortest three vectors $|\mathbf{q}|$, and Fig. 3E shows the entire map of the Dirac cone replicas up to the shortest eight vectors $|\mathbf{q}|$. The

numbers in the circles indicate the order of the length of $|\mathbf{q}|$ (see SM and fig. S19) (44, 45). For a single Umklapp scattering, the scattering intensity should decrease rapidly with $|\mathbf{q}|$ as shown in Fig. 3G because $t(\mathbf{q})$ decays exponentially with large $|\mathbf{q}|$, as demonstrated in fig. S19C. However, because the number of Umklapp scattering events which give the same $|\mathbf{q}|$ also increases very rapidly with increasing $|\mathbf{q}|$, the intensities of higher-order Dirac cone replicas increase markedly compared with single-scattering intensities, as shown in Fig. 3G (e.g., intensities are enhanced by factors of 1.15×10^5 and 7.40×10^{11} with $|\mathbf{q}| = 6.14$ and 9.01 \AA^{-1} , respectively). Thus, the resulting intensities for larger $|\mathbf{q}|$ become similar (Fig. 3G).

In our experiment, we observed a similar trend in ARPES intensities (Fig. 3G). The first two intensities of Dirac cone replicas decrease exponentially, and the remaining Dirac cones show almost constant intensities. Although a general trend of intensities in our experiment agrees well with theoretical expectations, the

absolute relative intensities do not match the computed ones. The discrepancies between experimental and simulated intensities of the Dirac cone replicas may originate from hidden scattering paths for large $|\mathbf{q}|$ Umklapp processes, such as with the surface reconstruction of the SiC substrate (32, 34, 46–48), or impurities or localized states that may not be properly accounted for in the simulation on the basis of single-particle pictures.

Detailed shapes of Dirac cone replicas in the graphene quasicrystal also show their characteristic features. They were all electron-doped, and the Dirac point energies of replicas on the upper and lower layers were nearly identical (Fig. 4, A and B). The average values of the Dirac points of the upper- and lower-layer graphene of the graphene quasicrystal measured from different samples with different photon energies are approximately 0.29 and 0.30 eV, respectively, where the overall energy resolution was approximately 60 meV (fig. S15). In typical epitaxial graphene on a SiC (0001) substrate, electrons were transferred from the SiC substrate to graphene (32–34, 46). Likewise, we expected a relative shift of the Dirac points between the two layers caused by the effective perpendicular electric field, as shown in recent studies on gated TBGs (49, 50). Unlike gated TBGs (49, 50), however, all Dirac cones in the sample had almost identical Dirac point energies, showing an anomalous doping behavior. Within a single-particle tight-binding approach, the doping dependence shown in Fig. 4 could not be described very well, implying either stronger interlayer coupling or larger interlayer screening. We also could not exclude other dopant sources such as carbon or silicon clusters. Further theoretical and experimental work is needed to resolve the anomalous doping in future.

This anomalous interlayer screening is in accordance with the exponential enhancement of ARPES intensities of high-order Dirac cone replicas (Fig. 3G) because of the higher-order Umklapp scatterings. Regardless of the large variation in ARPES intensities of Dirac cone replicas, their Fermi velocities are almost constant (Fig. 4C; see also fig. S13F) (51, 52). The overall interaction strength between Dirac cones on upper and lower layers can be estimated as $|t(\mathbf{q})/\Delta E|^2$, where ΔE is the energy difference between $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$ states. In TBGs, ΔE for the dominant interaction near the Dirac point is proportional to the rotation angle. Thus, TBGs with small angles exhibit very strong coupling between $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$ because ΔE is very small, causing strong distortion of the Dirac cone structures, such as the additional renormalization of Fermi velocity (42, 53). In the graphene quasicrystal, however, ΔE is very large, so we expect negligible distortion of the Dirac cone structures. As expected, all of the Dirac cones showed a well-defined linear momentum-energy relation (Fig. 4, B and C). Furthermore, the Fermi velocities of the Dirac cones did not differ as much as would be expected from Umklapp scattering with large $|\mathbf{q}|$ values. The energy distribution curves (fig. S14) show that a

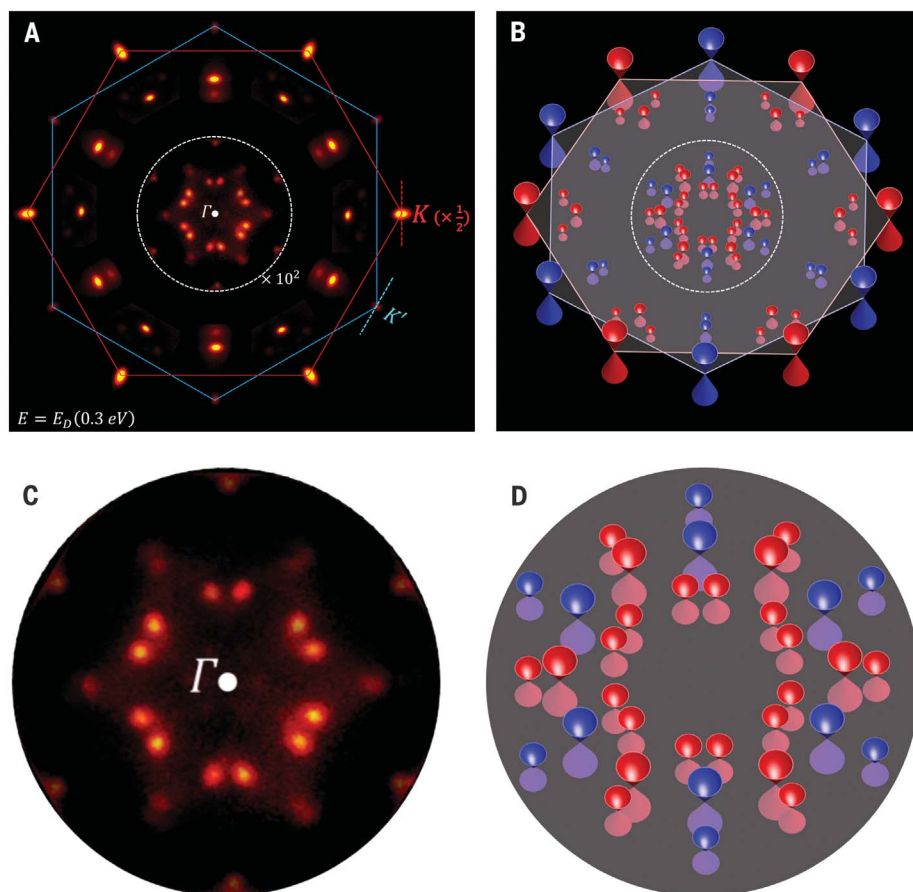


Fig. 2. Constant energy maps of ARPES spectra of graphene quasicrystal. (A) A constant energy map at the Dirac point energy (E_D) located at a binding energy of 0.3 eV, where the K points of the upper- and lower-layer graphene are K and K', respectively. The intensities near the Γ point were magnified by 100, and the intensity of the Dirac cone at the K point was reduced by half. (B) A schematic drawing of Dirac cones shown in (A), where red and blue Dirac cones come from the upper- and lower-layer graphene, respectively, and the gray plane indicates the Dirac point, with the darker and brighter colors indicating binding energies above and below the Dirac point. (C) Enlarged constant energy maps near the Γ point indicated by the circle in (A). (D) A schematic drawing of Dirac cones near the Γ point shown in (C).

Fig. 3. Calculations of the electronic structure of graphene quasicrystal. (A) A constant energy map at the Fermi energy (0.3 eV above the Dirac points) of simulated ARPES spectra of graphene quasicrystal. The color of each Dirac cone indicates whether the Dirac cone originates mainly from the states in the upper (red) or the lower (blue) graphene layer. The intensity of each Dirac cone indicates the scattering intensity plotted on a log scale. The intensities of the Dirac cones near the Γ point were magnified by a factor of 10^9 . Letters A and B designate the original Dirac cones, and letters C to F represent the Dirac cone replicas. (B to D) Umklapp scattering paths from the Dirac point of the upper layer with the shortest three wave vectors $|\mathbf{q}|$ involved in the Umklapp process (see the text). Each panel shows the scattering involving different \mathbf{G} . (E) A schematic drawing of the locations of Dirac cones in calculations, where the Dirac cones were ranked by the length of $|\mathbf{q}|$ (rankings are indicated by the numbers in the circles) and the size of the circle is proportional to the intensity calculated theoretically (see also figs. S19 and S20). (F) A schematic drawing of the locations of Dirac cones in ARPES measurements, where the Dirac cones were also ranked by the length of $|\mathbf{q}|$ as in (E). Both the intensity and the size of the circle are proportional to the intensity measured experimentally. (G) The experimental intensities of Dirac cones plotted as a function of $|\mathbf{q}|$ and the theoretical calculations of the contribution to the ARPES intensity from the single (multiple) scattering plotted as a function of $|\mathbf{q}|$. Black (gray) circles in (G) show the contribution from the single (multiple) scattering (see also SM and fig. S19). Red and blue circles in (G) show experimental intensities of Dirac cones plotted as a function of $|\mathbf{q}|$. I/I_0 is the ratio of the intensity of each Dirac cone replica to that of the main Dirac cone of the upper-layer graphene.

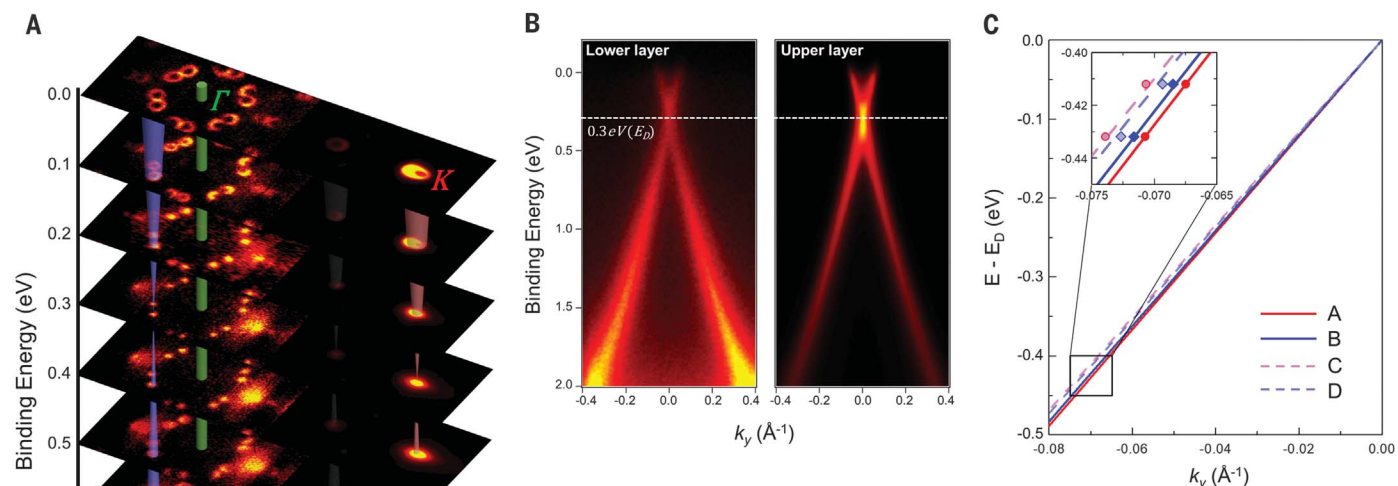
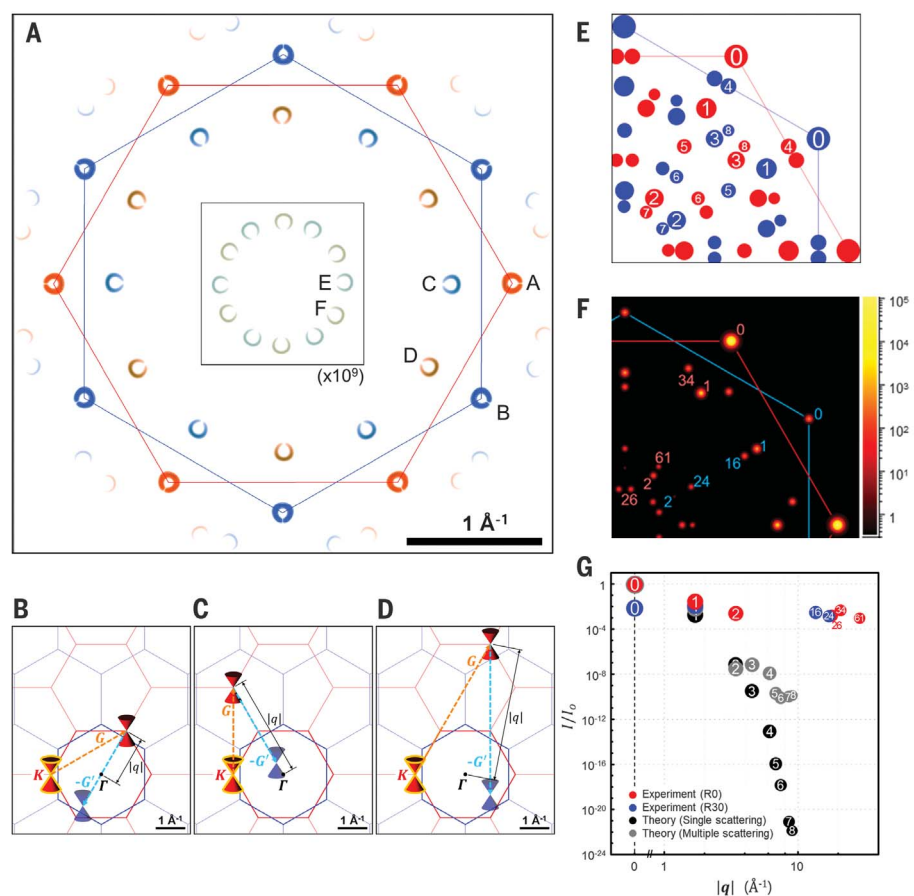


Fig. 4. ARPES spectra and Fermi velocities of graphene quasicrystal. (A) Constant energy maps of ARPES spectra of graphene quasicrystal with different binding energies. (B) Energy-momentum dispersions of Dirac cones at the K points of the upper- and lower-layer graphene. k_y is the electron momentum across the K point, as shown in Fig. 2A. (C) The fitted lines of energy-momentum dispersions of Dirac cones used to extract Fermi velocities (see also fig. S13F for detailed comparison between the experimental dispersions and the fitted lines), where the Fermi velocities of the A, B, C, and D Dirac cones (see Fig. 3A for the notations of the Dirac cones) are 0.93×10^6 , 0.92×10^6 , 0.89×10^6 , and 0.91×10^6 m/s, respectively. In the inset of (C), experimental dispersions are overlapped with the fitted lines.

typical pseudogap feature in epitaxial graphene (46, 54) occurs at the Dirac points for both layers. Having established the quasicrystalline order in our TBG system and shown its consequence in realizing a distinct Dirac electron system, we expect that the TBG rotated by exactly 30° will be an important material for studying various unsolved issues (26–28, 30, 37) in anomalous physical properties of quasicrystals.

REFERENCES AND NOTES

1. D. Shechtman, I. Blech, D. Gratias, J. W. Cahn, *Phys. Rev. Lett.* **53**, 1951–1953 (1984).
2. D. Levine, P. J. Steinhardt, *Phys. Rev. Lett.* **53**, 2477–2480 (1984).
3. L. Bindi, P. J. Steinhardt, N. Yao, P. J. Lu, *Science* **324**, 1306–1309 (2009).
4. E. Abe, Y. Yan, S. J. Pennycook, *Nat. Mater.* **3**, 759–767 (2004).
5. L. Bindi et al., *Sci. Rep.* **5**, 9111 (2015).
6. H. M. Price, O. Zilberberg, T. Ozawa, I. Carusotto, N. Goldman, *Phys. Rev. Lett.* **115**, 195303 (2015).
7. Z. V. Vardeny, A. Nahata, A. Agrawal, *Nat. Photonics* **7**, 177–187 (2013).
8. M. A. Bandres, M. C. Rechtsman, M. Segev, *Phys. Rev. X* **6**, 011016 (2016).
9. M. Verbin, O. Zilberberg, Y. Lahiri, Y. E. Kraus, Y. Silberberg, *Phys. Rev. B* **91**, 064201 (2015).
10. D. Levine, P. J. Steinhardt, *Phys. Rev. B* **34**, 596–616 (1986).
11. F. S. Pierce, Q. Guo, S. J. Poon, *Phys. Rev. Lett.* **73**, 2220–2223 (1994).
12. P. Ebert, M. Feuerbacher, N. Tamura, M. Wollgarten, K. Urban, *Phys. Rev. Lett.* **77**, 3827–3830 (1996).
13. R. Widmer, P. Gröning, M. Feuerbacher, O. Gröning, *Phys. Rev. B* **79**, 104202 (2009).
14. K. Nagao, T. Inuzuka, K. Nishimoto, K. Edagawa, *Phys. Rev. Lett.* **115**, 075501 (2015).
15. S. M. Lubin, W. Zhou, A. J. Hryn, M. D. Huntington, T. W. Odom, *Nano Lett.* **12**, 4948–4952 (2012).
16. T. Odagaki, *Solid State Commun.* **60**, 693–696 (1986).
17. M. Kohmoto, B. Sutherland, *Phys. Rev. B* **34**, 3849–3853 (1986).
18. J. Hafner, M. Krajci, *Phys. Rev. Lett.* **68**, 2321–2324 (1992).
19. E. S. Zijlstra, T. Janssen, *Phys. Rev. B* **61**, 3377–3383 (2000).
20. Z. M. Stadnik, D. Purdie, Y. Baer, T. A. Lograsso, *Phys. Rev. B* **64**, 214202 (2001).
21. E. Rotenberg, W. Theis, K. Horn, *Prog. Surf. Sci.* **75**, 237–253 (2004).
22. R. Mäder, R. Widmer, P. Gröning, W. Steurer, O. Gröning, *Phys. Rev. B* **87**, 075425 (2013).
23. L. C. Collins, T. G. Witte, R. Silverman, D. B. Green, K. K. Gomes, *Nat. Commun.* **8**, 15961 (2017).
24. T. Fujiwara, S. Yamamoto, G. Trambly de Laissardiére, *Phys. Rev. Lett.* **71**, 4166–4169 (1993).
25. S. Roche, D. Mayou, *Phys. Rev. Lett.* **79**, 2518–2521 (1997).
26. S. Roche, T. Fujiwara, *Phys. Rev. B* **58**, 11338–11344 (1998).
27. J. Delahaye, C. Berger, *Phys. Rev. B* **64**, 094203 (2001).
28. L. Levi et al., *Science* **332**, 1541–1544 (2011).
29. P. Stampfli, *Helv. Phys. Acta* **59**, 1260–1263 (1986).
30. E. Koren, U. Duerig, *Phys. Rev. B* **93**, 201404 (2016).
31. Y. E. Kraus, Z. Ringel, O. Zilberberg, *Phys. Rev. Lett.* **111**, 226401 (2013).
32. C. Riedl, C. Coletti, U. Starke, *J. Phys. D Appl. Phys.* **43**, 374009 (2010).
33. C. Riedl, C. Coletti, T. Iwasaki, A. A. Zakharov, U. Starke, *Phys. Rev. Lett.* **103**, 246804 (2009).
34. H. Kageshima, H. Hibino, S. Tanabe, *J. Phys. Condens. Matter* **24**, 314215 (2012).
35. M. Ye et al., *Eur. Phys. J. B* **75**, 31–35 (2010).
36. H.-C. Shin et al., *J. Am. Chem. Soc.* **137**, 6897–6905 (2015).
37. R. W. Havener, H. Zhuang, L. Brown, R. G. Hennig, J. Park, *Nano Lett.* **12**, 3162–3167 (2012).
38. Y. Kim et al., *Phys. Rev. Lett.* **110**, 096602 (2013).
39. Y. Kim et al., *Nano Lett.* **16**, 5053–5059 (2016).
40. T. Ohta et al., *Phys. Rev. Lett.* **109**, 186807 (2012).
41. A. Artaud et al., *Sci. Rep.* **6**, 25670 (2016).
42. P. Moon, M. Koshino, *Phys. Rev. B* **87**, 205404 (2013).
43. M. Van der Donck, C. De Beule, B. Partoens, F. M. Peeters, B. Van Duppen, *2D Mater.* **3**, 035015 (2016).
44. M. Koshino, P. Moon, *J. Phys. Soc. Jpn.* **84**, 121001 (2015).
45. M. Koshino, *New J. Phys.* **17**, 015014 (2015).
46. S. Kim, J. Ihm, H. J. Choi, Y.-W. Son, *Phys. Rev. Lett.* **100**, 176802 (2008).
47. P. Mallet et al., *Phys. Rev. B* **76**, 041403 (2007).
48. F. Varchon et al., *Phys. Rev. Lett.* **99**, 126805 (2007).
49. J. D. Sanchez-Yamagishi et al., *Phys. Rev. Lett.* **108**, 076601 (2012).
50. T.-F. Chung, R. He, T.-L. Wu, Y. P. Chen, *Nano Lett.* **15**, 1203–1210 (2015).
51. M. Sprinkle et al., *Phys. Rev. Lett.* **103**, 226803 (2009).
52. C. Hwang et al., *Sci. Rep.* **2**, 590 (2012).
53. J. M. B. Lopes Dos Santos, N. M. R. Peres, A. H. Castro Neto, *Phys. Rev. Lett.* **99**, 256802 (2007).
54. A. Bostwick, T. Ohta, T. Seyller, K. Horn, E. Rotenberg, *Nat. Phys.* **3**, 36–40 (2006).

ACKNOWLEDGMENTS

We thank S.-H. Kang, H. J. Choi, E.-G. Moon, L. A. Wray, A. Kent, M. Yang, and K. Ihm for fruitful discussions. **Funding:** This work was supported by the National Research Foundation (NRF) of Korea (NRF-2015R1A2A2A01004853 and NRF-2017M2A2A6A01019384). M.K. was supported by JSPS KAKENHI grants JP25107005 and JP15K21722. Y.-W.S. was supported by the NRF of Korea (grant 2017R1A5A1014862, SRC program vdwMRC center). P.M. was supported by the NYU Shanghai (start-up funds), the NYU-ECNU Institute of Physics, the NYU Global Seed Grant for Collaborative Research program, and the NSF of China (grant 11550110177, Research Fund for International Young Scientists program). P.K. was supported by the Global Research Laboratory program (2015K1A1A2033332) through the NRF of Korea. C.-W.Y. was supported by the National Research Council of Science and Technology (NST) (CRC-15-06-KIGAM) and the NRF of Korea (NRF-2011-0030058 and NRF-2015R1D1A1A01059653). **Author contributions:** S.J.A., H.-W.K., H.-C.S., E.H.K., and J.R.A. contributed to the growth of the graphene quasicrystal and ARPES experiments. T.-H.K., H.W.C., and C.-W.Y. contributed to the TEM experiments. P.M., M.K., and Y.-W.S. contributed to the theoretical calculations. P.K. and S.-J.K. contributed to the interpretation of the experiments. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/782/suppl/DC1

Materials and Methods

Figs. S1 to S21

References (55–60)

21 December 2017; accepted 19 June 2018

Published online 28 June 2018

10.1126/science.aar8412

Dirac electrons in a dodecagonal graphene quasicrystal

Sung Joon Ahn, Pilkyung Moon, Tae-Hoon Kim, Hyun-Woo Kim, Ha-Chul Shin, Eun Hye Kim, Hyun Woo Cha, Se-Jong Kahng, Philip Kim, Mikito Koshino, Young-Woo Son, Cheol-Woong Yang and Joung Reul Ahn

Science **361** (6404), 782-786.
DOI: 10.1126/science.aar8412 originally published online June 28, 2018

Dirac fermions in quasicrystalline graphene

Quasicrystal lattices, which can have rotational order but lack translational symmetry, can be used to explore electronic properties of materials between crystals and disordered solids. Ahn *et al.* grew graphene bilayers rotated exactly 30° that have 12-fold rotational order. Electron diffraction and microscopy confirmed the formation of quasicrystals, and angle-resolved photoemission spectroscopy revealed anomalous interlayer electronic coupling that was quasi-periodic. The millimeter-scale layers can potentially be transferred to other substrates.

Science, this issue p. 782

ARTICLE TOOLS

<http://science.sciencemag.org/content/361/6404/782>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2018/06/27/science.aar8412.DC1>

REFERENCES

This article cites 60 articles, 3 of which you can access for free
<http://science.sciencemag.org/content/361/6404/782#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

SURFACE CHEMISTRY

High-affinity adsorption leads to molecularly ordered interfaces on TiO₂ in air and solution

Jan Balajka¹, Melissa A. Hines², William J. I. DeBenedetti², Mojmir Komora^{1,3,4}, Jiri Pavelec¹, Michael Schmid¹, Ulrike Diebold^{1*}

Researchers around the world have observed the formation of molecularly ordered structures of unknown origin on the surface of titanium dioxide (TiO₂) photocatalysts exposed to air and solution. Using a combination of atomic-scale microscopy and spectroscopy, we show that TiO₂ selectively adsorbs atmospheric carboxylic acids that are typically present in parts-per-billion concentrations while effectively repelling other adsorbates, such as alcohols, that are present in much higher concentrations. The high affinity of the surface for carboxylic acids is attributed to their bidentate binding. These self-assembled monolayers have the unusual property of being both hydrophobic and highly water-soluble, which may contribute to the self-cleaning properties of TiO₂. This finding is relevant to TiO₂ photocatalysis, because the self-assembled carboxylate monolayers block the undercoordinated surface cation sites typically implicated in photocatalysis.

Metal oxide photocatalysts formed from earth-abundant metals have attracted a great deal of attention, particularly for environmental remediation and photocatalysis (1). These materials are attractive because they are inexpensive, they are typically very stable under reactive conditions (e.g., in air or solution), and their semiconducting nature enables efficient generation of long-lived photo-carriers that diffuse to the surface and initiate chemical reactions. Although some of these materials have been commercialized, an atomic-scale understanding of their reactivity in ambient and solution environments has proven elusive, in part because of the diversity of possible active sites on metal oxide surfaces and the complex milieu of reactants in ambient environments. Although scientists have made considerable strides in understanding the atomic-scale structure of model catalysts in vacuum (2–4), the study of solution/metal oxide interfaces is just emerging (5–9).

An early indication of complexity at the solution/metal oxide interface was the observation of photoswitching (10). When TiO₂ surfaces in ambient environments are irradiated with ultraviolet (UV) light, they become hydrophilic, but they slowly revert to hydrophobicity in the dark. In contrast, clean TiO₂ is unaffected by UV irradiation in vacuum. The UV-induced hydrophilicity in ambient environments is now understood to be due to facile oxidation of adsorbates, such as hydrocarbons, to produce, for example, hydroxylated species (11). The reverse dark reaction has never been explained. Although there have been suspicions that the hydrophobic dark state results from adventitious contamination, recent reports of the formation of a highly ordered surface structure with (2 × 1) symmetry on TiO₂ rutile (110) in H₂O (12), or after exposure to H₂O (13) or air (14, 15), have called this into question. On the basis of a combination of scanning tunneling microscopy (STM) and x-ray

diffraction (XRD), this structure was recently assigned to a new ordered state of adsorbed H₂O—persistent in vacuum at room temperature—in which one H₂O molecule adsorbs to every other surface Ti atom (12). Others attributed the (2 × 1) structure to an ordering of H₂O at the solid/liquid interface (11), the reaction of CO₂ with H₂O to produce bicarbonate (13), or the formation of hydroperoxyl from O₂ and H₂O (14). We show that none of these explanations are correct.

To clarify the configuration and origin of this persistent ordered structure, we constructed a small side chamber (Fig. 1 and fig. S1), which allows the transfer of an ultrapure H₂O drop to a precleaned TiO₂ surface in vacuum or a controlled gas, followed by characterization in vacuum with no air exposure. This side chamber is connected by an in-vacuum sample transfer system to the surface analysis chamber. The experiments proceeded in four steps (movie S1). First, a small icicle was grown on the cooled finger by vapor transfer from liquid H₂O. The side chamber was then evacuated to ~10^{−7} mbar, and the precleaned sample was transferred from the analysis system to directly beneath the icicle. The side chamber was isolated, and the cold finger was warmed until a drop of liquid H₂O fell onto the sample. A gas was optionally added. After exposure, the H₂O and any additional gases were evacuated with a LN₂-cooled cryopump, and the sample was transferred back for analysis.

Figure 2A shows that exposure of a clean TiO₂(110) surface to ultrapure H₂O in vacuum did not lead to the (2 × 1) structure reported in (12). Instead, the surface was essentially identical to clean, room-temperature TiO₂(110) after exposure to H₂O gas in vacuum (16), albeit with a low density of contaminants (white blobs in

¹Institute of Applied Physics, Technische Universität Wien, Wiedner Hauptstraße 8-10/134, 1040 Vienna, Austria. ²Department of Chemistry, Cornell University, Ithaca, NY 14853, USA. ³Central European Institute of Technology, Purkyňova 123, Brno 612 00, Czech Republic. ⁴Institute of Physical Engineering, Brno University of Technology, Technická 2896/2, Brno 616 69, Czech Republic.

*Corresponding author. Email: diebold@iap.tuwien.ac.at

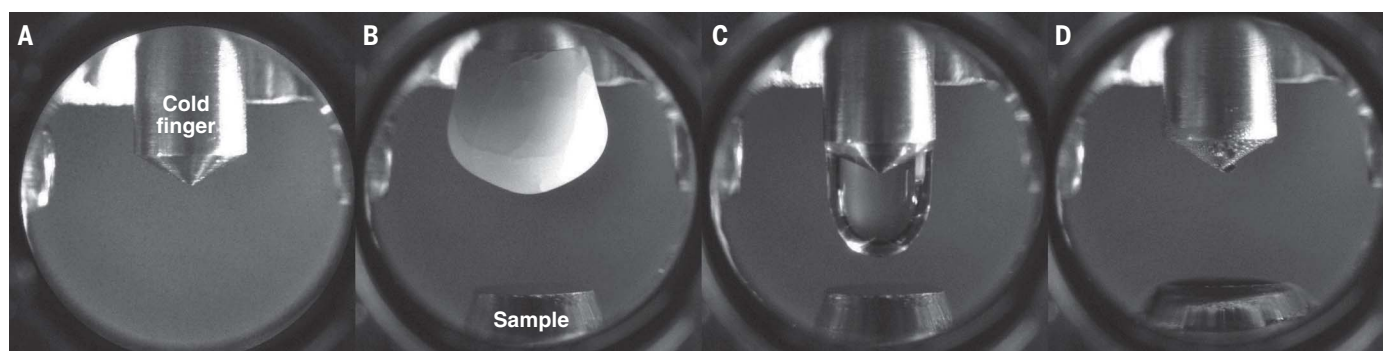


Fig. 1. Dosing liquid H₂O inside a vacuum chamber. (A) The finger is cooled with LN₂. (B) Ice is grown from H₂O vapor, and after evacuation, the sample is introduced. (C) The finger is heated, and liquid H₂O falls onto the sample and (D) remains until evacuation.

Fig. 2A). In STM, the morphology was dominated by rows of undersaturated Ti atoms, which imaged as protrusions with a 3-Å spacing. Both gas and liquid water exposure hydroxylated any pretreatment-induced O vacancies, which are apparent as protrusions between the Ti rows, but resulted in no stable H₂O adsorption.

This assignment was confirmed with x-ray photoemission spectroscopy (XPS) (Fig. 3, A to C). Importantly, a 534-eV O 1s transition was not observed, so no adsorbed H₂O remained in vacuum (5). The only difference between the clean and H₂O-exposed surfaces was minor contamination, as shown by the aliphatic or “adventitious” C transition at 285 eV. Hydroxyls from H₂O dissociation at O vacancies led to the tiny shoulder at 532.5 eV.

In contrast, exposure of a clean TiO₂(110) surface to a H₂O drop in air resulted in a (2 × 1) structure, as shown by STM images (Fig. 2B and fig. S2). Chemical analysis of the air-exposed surface (Fig. 3, D to F, and figs. S4 and S5) revealed carboxylic acid adsorption, as evidenced by the C 1s transition at 289.2 eV and the O 1s transition at 532.5 eV. Similar results were obtained when a nominally dry rutile surface was exposed to air for 2 min (Fig. 3). The similarity is not surprising, given that clean TiO₂(110) surfaces develop a several-monolayer-thick H₂O layer when exposed to air (5). In these experiments, the intensity of the 285-eV transition, which we assigned to aliphatic C in the carboxylic acid and adventitious carbon, varied from run to run.

To confirm that these features are due to carboxylic acids, reference spectra were obtained from a formic acid monolayer deposited in vacuum, as shown in blue in Fig. 3, D to F. This monolayer displayed similar, albeit more intense, transitions in the O 1s and C 1s regions, but no adventitious C. Formic acid deposition self-limits at one monolayer in vacuum (2, 4). To test whether the air exposure is similarly self-limiting, we exposed a clean TiO₂(110) surface to air for 30 min. This exposure led to the development of a complete monolayer of carboxylic acid (figs. S3 and S4), as determined by XPS and the production of a well-ordered monolayer with (2 × 1) symmetry.

Carboxylic acid adsorption explains the (2 × 1) surface structure. Carboxylic acids bind to TiO₂(110) dissociatively in a bridged, bidentate geometry, forming an adsorbed proton (H⁺) on a bridging O and an adsorbed carboxylate (RCOO[−]) bound to two adjacent Ti atoms (2–4), which leads to a (2 × 1) structure.

The previously suggested (13) adsorption of bicarbonate, which has similar binding, can be ruled out from the C 1s spectrum, which responds sensitively to the chemical environment (fig. S6). Carboxylate C is in a +2 oxidation state, whereas bicarbonate C is in a +4 oxidation state. Density functional theory (DFT) calculations predict a 1.32-eV shift in the C 1s core-level energy from this change (Fig. 2B and supplementary materials). Carboxylic acid adsorption is consistent with the 1.95-Å Ti–O distance observed in situ and ex situ XRD analysis of H₂O-exposed TiO₂(110)

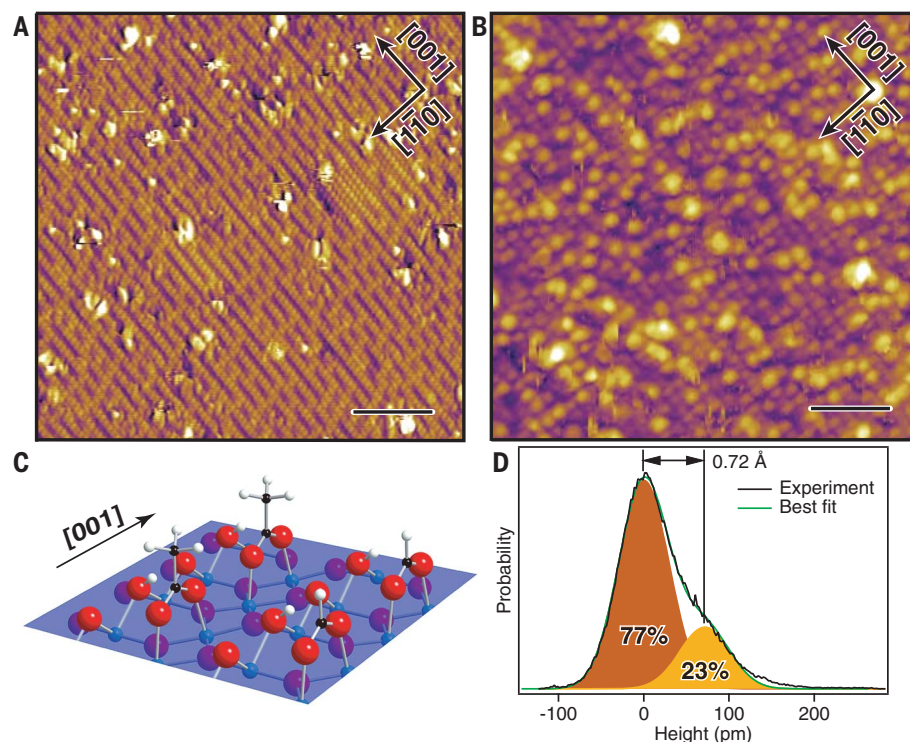


Fig. 2. Effect of air on the TiO₂(110) surface. STM images of TiO₂(110) after contact with an H₂O drop (A) in vacuo (+1.87 V, 112 pA) and (B) in air (+1.42 V, 41 pA). Scale bars, 3 nm. (C) DFT simulation of a formate/acetate monolayer. Ti, O, C, and H atoms are shown in blue, red, black, and white, respectively. (D) STM height distribution of TiO₂(110) after exposure to a H₂O drop in air and fit to equal-width Gaussians (brown and gold).

(12), because these species have a 2.08-Å Ti–O distance in DFT simulations.

STM images of the (2 × 1) surface revealed two species with an apparent height difference of 0.72 Å (Fig. 2C and fig. S2), which is in good agreement with the 0.63-Å height difference predicted by simulated STM (fig. S7) of a formate/acetate monolayer. In contrast, the height difference between two longer carboxylates of the form CH₃(CH₂)_{n+1}COO[−] and CH₃(CH₂)_nCOO[−] (where *n* = 0, 1, ...) would be 0.95 Å. Therefore, Fig. 2 suggests an ~80% formate/20% acetate monolayer.

Polarized infrared spectroscopy of air-exposed surfaces (fig. S8) provided further information. Formate/TiO₂(110) is the only carboxylate terminated by a single C–H bond. This structure leads to a characteristic pair of p-polarized C–H stretch resonances (supplementary materials), which were observed on both air-exposed surfaces and formate monolayers. This suggested that formate, the shortest carboxylate, is the shorter species in the STM images, and acetate is the taller one.

To test whether the adsorbed carboxylates were formed by the reaction of gas-phase molecules with H₂O, we coexposed clean rutile surfaces to liquid H₂O and a variety of atmospheric gases. No reaction was detected with CO₂ (0.4 mbar in air) or CO (10^{−4} mbar in air) (figs. S9 and S10).

To show that the adsorbed carboxylic acids arise from a ubiquitous component of air and

not specific contaminants of our air, we obtained XPS spectra and STM images of TiO₂(110) exposed to air in rural Ithaca, NY, USA, ~7000 km from the measurements taken in urban Vienna, Austria. These data (figs. S2 and S4) display the same chemical signatures, the same height difference, and a similar height distribution.

The formation of two such similar, highly ordered monolayers in ambient environments on different continents—as well as previous observations of (2 × 1) structures in England (12), Italy (11), the United States (13), and Japan (14)—can only be explained by ubiquitous environmental species with high-affinity binding. We address these requirements sequentially.

Formic and acetic acids are the dominant sources of atmospheric acidity in most parts of the world, with longer-chain acids having much lower concentrations (17). Measurements in the United States and Germany, for example, show typical partial pressures of 10^{−6} to 10^{−7} mbar for each species (16). These acids have many origins; however, biogenic sources, particularly the oxidation of isoprene from trees and shrubs, are thought to dominate (16). Both acids are rapidly and efficiently adsorbed by liquid H₂O, leading to calculated concentrations of each acid of ~10 μM in ambient environments (supplementary materials). This facile trapping is also expected in the several-monolayer-thick H₂O layer (5) that forms on air-exposed TiO₂(110).

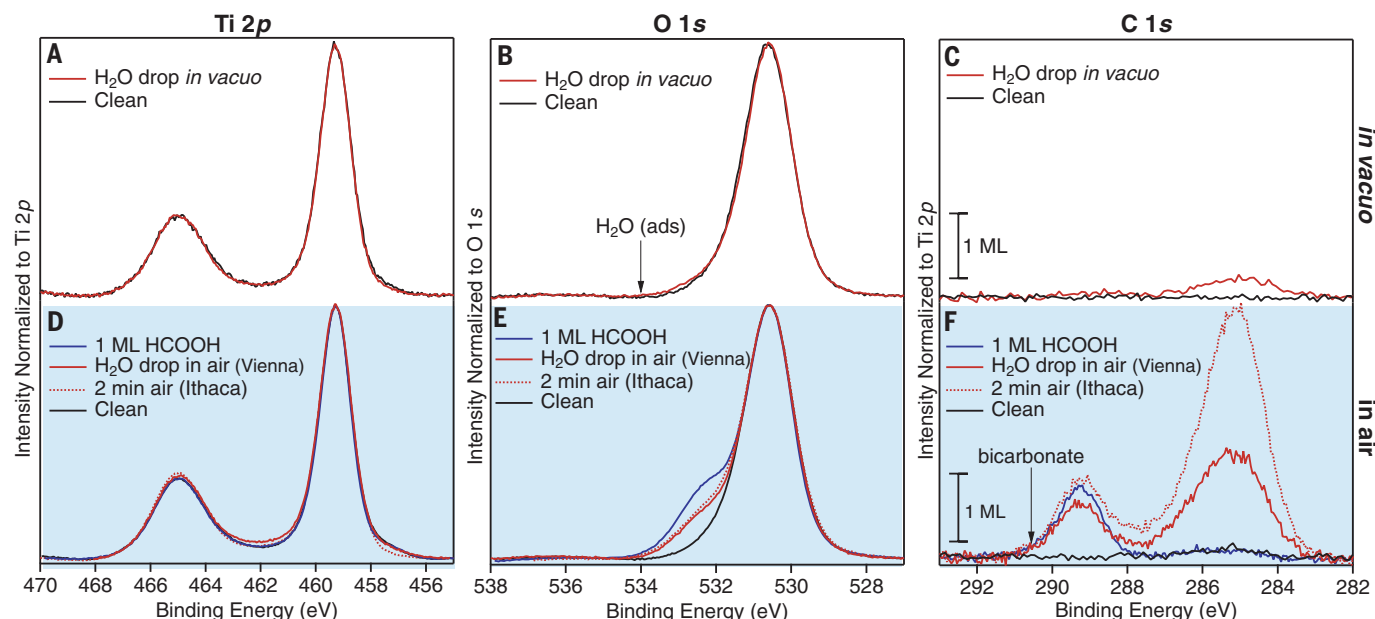


Fig. 3. XPS spectra of $\text{TiO}_2(110)$ after exposure to a H_2O drop. (A to C) In vacuo. (D to F) In air in Vienna, Austria. The lower panels include reference spectra from a 2-min air exposure in Ithaca, NY, and a saturation

HCOOH coverage [\equiv 1 monolayer (ML)]. Energies of the unobserved transitions from $\text{H}_2\text{O}/\text{TiO}_2$ (5) and bicarbonate/ TiO_2 systems are indicated. The “1 ML” calibration bars represent a HCOOH saturation coverage.

The hydrolysis of atmospheric CO_2 and CO leads to much lower concentrations of H_2CO_3 and HCOOH . Thermodynamic calculations (supplementary materials) show that air-equilibrated H_2O contains 10^{-8} M H_2CO_3 and 10^{-7} M HCOOH —orders of magnitude less than from atmospheric formic and acetic acids.

Having established that formic and acetic acids are ubiquitous in air, why do these relatively minor species—as opposed to other small molecules, such as alcohols—form monolayers on $\text{TiO}_2(110)$? The partial pressure of atmospheric methanol (18), for example, is 10^{-2} to 10^{-3} mbar. Although many small molecules adsorb to TiO_2 , species with monodentate binding [e.g., alcohols (19) and amines (20)] typically desorb near room temperature, whereas carboxylic acids, which form two bonds to the substrate, are stable until much higher temperatures (21). Thus, although some alcohols and other monodentate species are present in considerably higher concentrations in air, desorption of these species will be facile.

Our finding that $\text{TiO}_2(110)$ surfaces in ambient environments are terminated by a well-ordered carboxylate monolayer (and not a persistent H_2O monolayer or adventitious molecules) explains a long-standing puzzle: the spontaneous transformation of initially hydrophilic TiO_2 surfaces into hydrophobic surfaces in the dark. We attribute the dark state to the spontaneous formation of mixed formate/acetate monolayers by adsorption from the atmosphere, given that formic acid solutions also produce hydrophobic, self-assembled monolayers on $\text{TiO}_2(110)$ (supplementary materials).

This finding provides further insight into the well-known self-cleaning properties of TiO_2 . Self-assembled carboxylate monolayers have the un-

usual property of being both hydrophobic (from their CH_x -terminated “tails”) and highly water-soluble (from their acidic “heads”). In their hydrophobic state, the carboxylate monolayers will resist contaminant adsorption. However, their high water solubility will cause the surface to become hydrophilic under rinsing, enabling the water sheeting action that is important for self-cleaning.

Perhaps most unexpectedly, this finding suggests that surfaces in ambient environments may be more structured and controlled than previously realized. In this study, a clean metal oxide surface selectively adsorbed species present in parts-per-billion concentrations to form a molecularly ordered interface, while effectively repelling other adsorbates present in much higher concentrations. This finding may have important implications for photocatalysis, given that self-assembled carboxylate monolayers effectively block the undercoordinated cation sites typically implicated in photocatalysis.

REFERENCES AND NOTES

1. A. Fujishima, X. Zhang, D. A. Tryk, *Surf. Sci. Rep.* **63**, 515–582 (2008).
2. U. Diebold, *Surf. Sci. Rep.* **48**, 53–229 (2003).
3. Z. Dohnálek, I. Lyubinetzky, R. Rousseau, *Prog. Surf. Sci.* **85**, 161–205 (2010).
4. A. G. Thomas, K. L. Syres, *Chem. Soc. Rev.* **41**, 4207–4217 (2012).
5. S. Yamamoto et al., *J. Phys. Condens. Matter* **20**, 184025 (2008).
6. M. Salmeron et al., *Faraday Discuss.* **141**, 221–229, discussion 309–346 (2009).
7. P. Fenter, N. C. Sturchio, *Prog. Surf. Sci.* **77**, 171–258 (2004).
8. R. Franking, H. Kim, S. A. Chambers, A. N. Mangham, R. J. Hamers, *Langmuir* **28**, 12085–12093 (2012).
9. M. J. Makowski, R. P. Galhenage, J. Langford, J. C. Hemminger, *J. Phys. Chem. Lett.* **7**, 1732–1735 (2016).
10. R. Wang et al., *Nature* **388**, 431–432 (1997).
11. J. T. Yates Jr., *Surf. Sci.* **603**, 1605–1612 (2009).

12. G. Serrano et al., *Adv. Mater. Interfaces* **2**, 1500246 (2015).
13. H. Hussain et al., *Nat. Mater.* **16**, 461–466 (2017).
14. A. Song, E. S. Skibinski, W. J. I. DeBenedetti, A. G. Ortol-Bloch, M. A. Hines, *J. Phys. Chem. C* **120**, 9326–9333 (2016).
15. A. Sasahara, M. Tomitori, *J. Phys. Chem. C* **120**, 21427–21435 (2016).
16. S. Wendt et al., *Science* **320**, 1755–1759 (2008).
17. P. Khare, N. Kumar, K. M. Kumari, S. S. Srivastava, *Rev. Geophys.* **37**, 227–248 (1999).
18. B. G. Heikes et al., *Global Biogeochem. Cycles* **16**, 1133 (2002).
19. Z. Li, R. S. Smith, B. D. Kay, Z. Dohnálek, *J. Phys. Chem. C* **115**, 22534–22539 (2011).
20. E. Farfan-Arribas, R. J. Madix, *J. Phys. Chem. B* **107**, 3225–3233 (2003).
21. M. A. Henderson, *J. Phys. Chem. B* **101**, 221–229 (1997).

ACKNOWLEDGMENTS

R. Gärtner and H. Schmid are thanked for technical assistance. **Funding:** This work was supported by the Austrian Science Fund FWF (Wittgenstein Prize project Z-250-N27, Doctoral College Solids4Fun project W1243-N16, and project F45), the European Research Council advanced Grant Oxide Surfaces (ERC-2011-ADG-20110209), and the National Science Foundation (CHE-1708025). This work used the National Energy Research Scientific Computing Center (Department of Energy, DE-AC02-05CH11231). **Author contributions:** J.B., J.P., M.K., M.S., and U.D. designed the apparatus and conceived of the experiment. J.B. and M.A.H. performed Vienna experiments, W.J.I.D. performed Ithaca experiments, and M.A.H. performed calculations. M.A.H. wrote the manuscript with input from all. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/786/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (22–29)
Movie S1

23 March 2018; accepted 25 June 2018
10.1126/science.aat6752

GRAPHENE

Interaction-driven quantum Hall wedding cake-like structures in graphene quantum dots

Christopher Gutiérrez^{1,2*}†, Daniel Walkup^{1,2*}, Fereshte Ghahari^{1,2*}, Cyprian Lewandowski^{3*}, Joaquín F. Rodríguez-Nieva⁴, Kenji Watanabe⁵, Takashi Taniguchi⁵, Leonid S. Levitov³, Nikolai B. Zhitenev¹, Joseph A. Stroscio^{1‡}

Quantum-relativistic matter is ubiquitous in nature; however, it is notoriously difficult to probe. The ease with which external electric and magnetic fields can be introduced in graphene opens a door to creating a tabletop prototype of strongly confined relativistic matter. Here, through a detailed spectroscopic mapping, we directly visualize the interplay between spatial and magnetic confinement in a circular graphene resonator as atomic-like shell states condense into Landau levels. We directly observe the development of a “wedding cake”-like structure of concentric regions of compressible-incompressible quantum Hall states, a signature of electron interactions in the system. Solid-state experiments can, therefore, yield insights into the behavior of quantum-relativistic matter under extreme conditions.

Energy quantization owing to quantum confinement takes place when the particle's de Broglie wavelength becomes comparable to the system's length scale. Confinement can arise through spatial constraints imposed by electric fields or through cyclotron motion induced by magnetic fields. Together, confinement and quantization strengthen the effects of electron-electron interactions, providing a setting to probe a range of exotic phenomena in strongly correlated quantum systems. In the solid-state setting, different types of confined strongly correlated states and the transitions between them have been studied in quantum dots (QDs) in the presence of external magnetic fields (1). Evolution

from atomic-like shell structure to magnetic quantization in QDs was first probed by using Coulomb blockade spectroscopy (2–4). For QDs at large magnetic fields—that is, in the quantum Hall regime—it is expected that Coulomb interactions and the redistribution of carriers between Landau levels (LLs) will lead to a characteristic “wedding cake”-like shape in the density of electronic states (Fig. 1E) (4–8). Although similar structures have been observed in ultracold atoms undergoing transition from the superfluid to Mott insulator (9, 10), the wedding cake-like structure has not been mapped spatially in a solid-state system.

Graphene offers an ideal platform for this inquiry, as it hosts a fully exposed two-dimensional

(2D) electron gas amenable to local probes (11–17). Graphene circular p-n junction resonators (18–22) with built-in local potentials (Fig. 1A) are particularly well suited to the present study; they circumvent the problems of edge roughness and edge impurities encountered in lithographically fabricated graphene QDs. Furthermore, they enable fine control of the confining potential as well as QD doping by means of local gate potentials, offering an opportunity to directly visualize the transition of electron states from spatial to magnetic confinement (Fig. 1, B to E).

In the absence of a magnetic field, confinement of graphene carriers in a p-n junction resonator gives rise to a series of quasibound single-particle states. These states result from oblique Klein scattering at the p-n interface (18–22). At the same time, Klein tunneling, although present, is very weak for oblique scattering angles and thus has little impact on confinement (23). Analogous to atomic physics, the many-electron shell-like states are characterized by radial and azimuthal quantum numbers (n, m), forming a ladder of states within the spatially confined potential (Fig. 1B) (20, 22). In weak magnetic fields, energy levels exhibit abnormally large splittings, corresponding to time-reversed $\pm m$ states in the presence of the π Berry phase (22, 24). At higher fields, the system enters

¹Center for Nanoscale Science and Technology, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. ²Maryland NanoCenter, University of Maryland, College Park, MD 20742, USA. ³Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ⁴Department of Physics, Harvard University, Cambridge, MA 02138, USA. ⁵National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan.

*These authors contributed equally to this work. †Present address: Quantum Matter Institute, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada.

‡Corresponding author. Email: joseph.stroscio@nist.gov

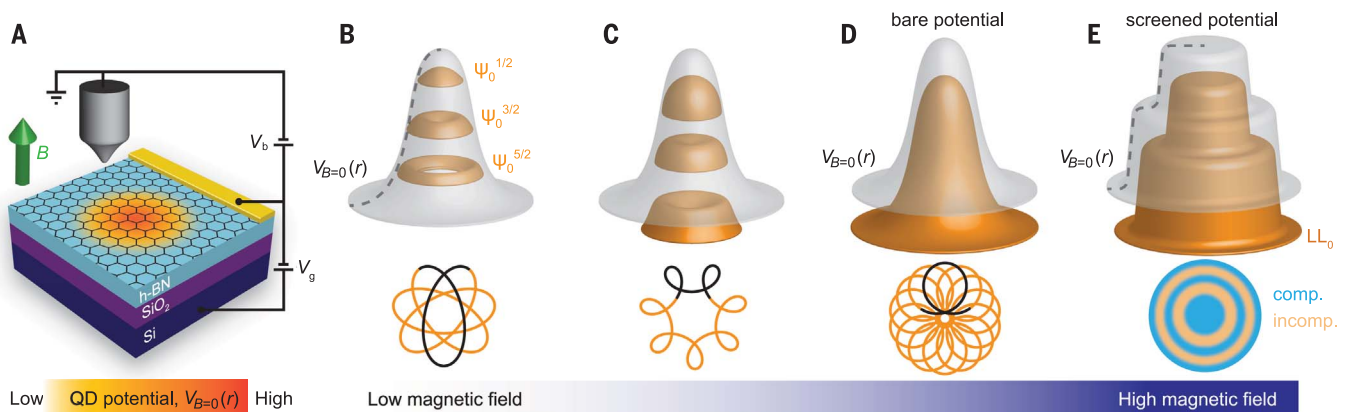


Fig. 1. Schematic evolution of states with magnetic field in a graphene QD. (A) The device geometry for the graphene QD resonator with a p-doped center inside an n-doped background used in the current experiment. V_b , sample bias voltage; V_g , gate voltage; $V_{B=0}(r)$, QD potential at $B = 0$. (B to E) Top panels show schematic of the potential profile (gray surface) and corresponding wave function density (orange surfaces), and bottom panels show semiclassical orbits, as a function of applied magnetic field. Confined states start out as quasibound QD

states [(B) and (C)] and condense into LLs (D) with increasing field. The corresponding screened potential (E) develops a wedding cake-like appearance through electron interactions. Semiclassical orbits start out as expected for a central force potential and then develop into cyclotron motion drifting along equipotential lines forming compressible (blue) and incompressible (yellow) density rings, as shown in (E). ψ_n^m , electron wave functions with n and m radial and azimuthal quantum numbers, respectively; $V_B(r)$, effective potential, Eq. 3.

the quantum Hall regime, with confined states transitioning to highly degenerate LLs (Fig. 1, C and D). A signature of this transition is the formation in the electron density of wedding cake-like structures comprised of a series of compressible and incompressible electron liquid rings (Fig. 1E) (4–8). Extending the single-particle description to include Coulomb interactions is essential in this regime.

Our experiment involves spectroscopic mapping of a graphene QD by tunneling measurements. The QD is formed by ionized impurities in the hexagonal boron nitride (hBN) insulating layer acting as a fixed built-in confining potential [see Fig. 1A and (22, 25) for device fabrication]. The transition from spatial to magnetic confinement occurs when the magnetic length becomes smaller than the confining potential width. By following the bright envelope in the spectral map

in Fig. 2A at zero field, we can model the screened confining potential, as seen by electrons, by $V_{B=0}(r) \approx U_0 \exp\left(-\frac{r^2}{R_0^2}\right) + U_\infty$, with the potential amplitude $U_0 \approx 210$ meV, radius $R_0 \approx 104$ nm, and the background value $U_\infty = -55$ meV, where r is the distance. This defines a characteristic length scale for the confining potential $l_V = \left(\frac{R_0^2 \hbar v_F}{U_0}\right)^{1/3} \approx 32$ nm. Here \hbar is Planck's constant \hbar divided by 2π , e is the elementary charge, and $v_F \approx 10^6$ m/s is the graphene Fermi velocity. Such a potential gives rise to quasibound states with energy splitting $\Delta E \approx \hbar v_F / l_V \approx 20$ meV. Application of a magnetic field B tends to confine electrons in a region of size $l_B = \left(\frac{\hbar}{eB}\right)^{1/2}$ and leads to the characteristic Landau quantization in graphene, $\varepsilon_N = \text{sgn}(N) \hbar v_F \sqrt{2|N|} / l_B$, with the LL number $N = 0, \pm 1, \pm 2, \dots$. Each LL is highly

degenerate and can host $n_{LL} = g/2\pi l_B^2$ carriers per unit area, where $g = 4$ is the valley and spin degeneracy. Therefore, we expect a transition from atomic-like QD states to LL states occurring at $l_B \approx l_V$, which gives a characteristic critical field of $B \approx 0.6$ T.

The transition results in an intricate evolution of QD states from spatial to magnetic confinement with increasing magnetic field, which is displayed in Fig. 2. The measured differential conductance signal, proportional to the local density of states (LDOS), shows the evolution of the QD states in the energy versus radial plane that cuts through the diameter of the QD. The zero-field shell-like QD states are well resolved in Fig. 2A under the bright concave band, which tracks the confining potential. States with $m = \pm 1/2$ have the largest weight in the center at $r = 0$, whereas states with a common radial quantum number n have a

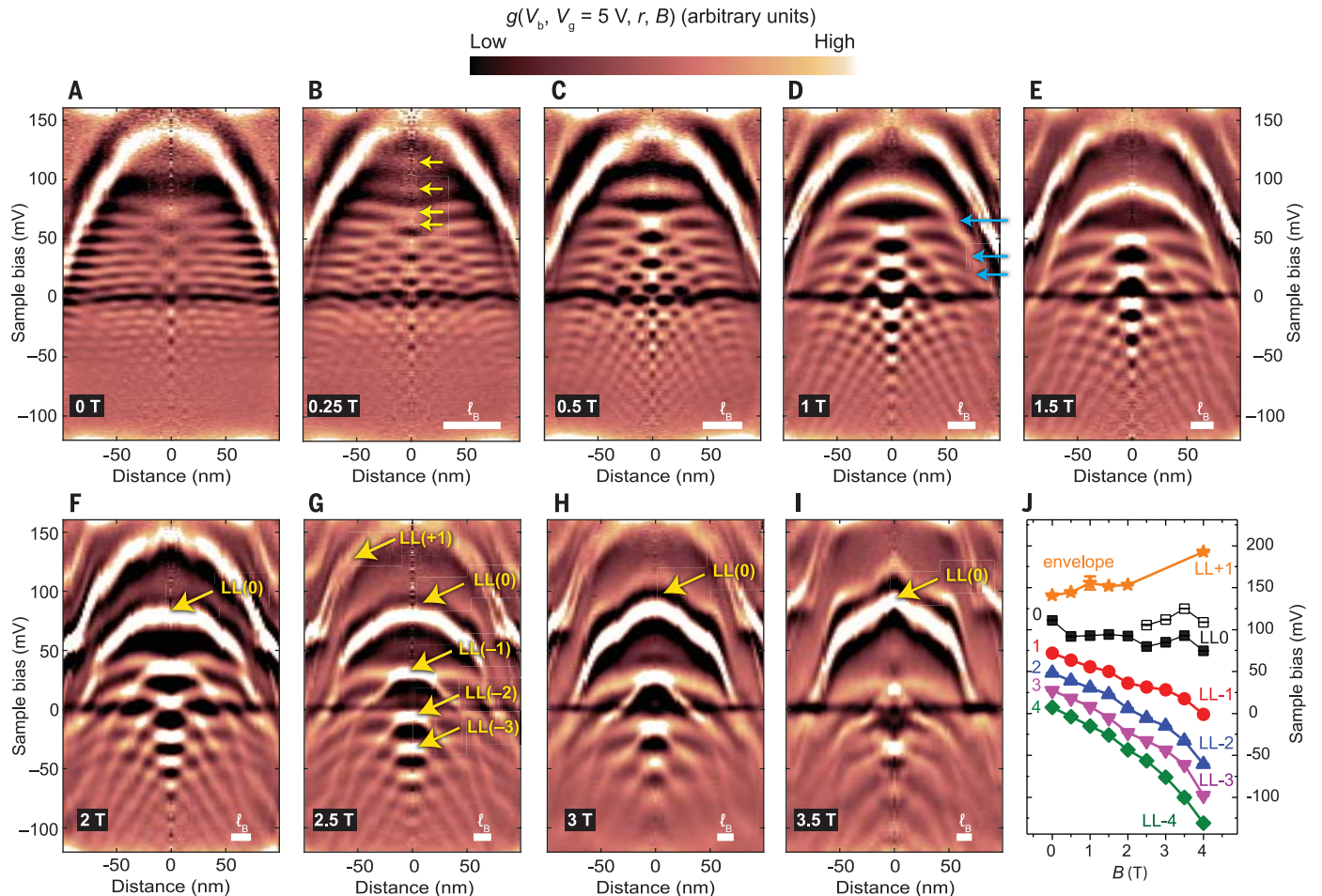


Fig. 2. Visualization of the condensation of states from spatial to magnetic quantization. (A to I) Experimental differential conductance [temperature $T = 4.3$ K], $g(V_b, V_g, r, B)$, maps the local density of states as a function of applied magnetic field, from 0 to 3.5 T, showing manifolds of spatially confined QD states condensing into LLs at higher fields. The magnetic field and corresponding magnetic length l_B is indicated at the bottom of each map. The 2D maps are radially averaged from a 2D grid of spectra. A smooth background was subtracted to remove the graphene dispersive background (25). The yellow arrows in

(B) indicate the splitting of the $m = \pm 1/2$ degeneracy at $r = 0$ owing to the turn on of a π Berry phase. The blue arrows in (D) indicate the shell-like states merging into the $N = 0$ LL edge mode. (J) Energy positions of the $n, m = 1/2$ states (symbols) obtained from the maps in (A) to (I) at $r = 0$ are observed to evolve into separate LLs with increasing applied magnetic field. $LL(0)$ splits into two peaks above 2.5 T, indicated by the open and solid square symbols. The experimental uncertainty, determined from fitting the peak positions in the spectra, represents one standard deviation and is smaller than the symbol size.

large weight in the form of arcs following the concave potential outline. The first critical field is reached by 0.25 T, where the $\pm m$ degeneracy is lifted owing to a π shift in the Berry phase of the QD wave function (22, 24), as seen by the doubling of the antinodes at $r = 0$ (arrows in Fig. 2B). The onset of the transition into the quantum Hall regime can be observed at $B = 0.5$ T (Fig. 2C), in agreement with the estimate above. Here states in the center of the resonator start to flatten out, have increased intensity, and shift lower in energy. Beginning at $B = 1$ T, various interior resonator states (arrows in Fig. 2D) merge into the $N = 0$ LL [LL(0)]. With progressively higher fields, the number of QD resonances decreases as they condense into the flattened central states, forming a series of highly degenerate LLs (Fig. 2, F to I). Beginning at about 2 T (Fig. 2F), LL(0) develops kinks near the QD boundary and an additional concave cusp near the center. Below, we argue that these are related to electron interaction effects. Additionally, LL(0) develops a splitting, which increases with field, whereas LL(-1) continually moves down in energy. In this field range, a decrease in conductance over a small energy range is also observed at the Fermi level, which we attribute to a Coulomb pseudogap (26, 27).

We now discuss the spatial pattern associated with the eigenstate evolution observed in Fig. 2. Experimental spatial maps of the differential con-

ductance, corresponding to resonator LDOS wave function probability amplitudes, are obtained by taking a 2D slice in the x - y plane of the dataset in Fig. 2 at a specific energy (Fig. 3). Only a subset of the data is shown in Fig. 3, corresponding to specific energies of the prominent central states at $r = 0$ in Fig. 2, with increasing magnetic field for each column of maps (a complete view of the dataset can be seen in movie S1). The spatial extent of the $m = \pm 1/2$ states at the selected energies is observed at zero field in the first column of the maps in Fig. 3. As we increase field and progress from left to right in the figure, we observe the formation of rings that contract and, simultaneously, sharpen with increasing field. For higher- n states (progressing down in a column), more rings are seen. Some of these rings originate from the quasibound resonances that have not yet developed into LLs, and some reflect the presence of magnetic confinement. The former can have a relatively narrow spatial profile if they are dispersing up or down in energy. At a still-higher field (Fig. 3D), LL plateaus are formed, as seen in Fig. 2, and show up as bright rings in the spatial maps, as indicated by the arrows pointing to LL(+1) and the valley-split (28) LL(0+), LL(0-) state in Fig. 3D. When LL states cross, or are pinned at the Fermi level, they form compressible (metallic) rings and disks, which start to show Coulomb charging effects (13), as indicated by the fine quartet of rings in the center and out-

side edge of Fig. 3D (see also vertically dispersing lines in Fig. 4E and rings in Fig. 4G).

A notable and unexpected feature observed in Fig. 3 is the appearance of circular nodal patterns in the spatial maps of differential conductance, which are present even at zero field. The origin of these nodal patterns is not clear at present, but they can be attributed either to interactions, as they resemble the shell-like structure predicted for Wigner crystals (7, 29–31), or to deviations from a rotationally symmetric confining potential. Deviations from perfect symmetry will partially lift the m -state degeneracy and give rise to nodal patterns. Moiré superlattice effects can be ruled out as an origin of the potential asymmetry because of the large angular mismatch ($\approx 29^\circ$) between the graphene and hBN insulator for this device, which gives a superlattice period of ≈ 0.5 nm, much smaller than the nodal-separation length scales (22). A nonsymmetric potential can result from the shape of the probe tip, which gets imprinted in the QD potential shape from the electric field generated during the tip voltage pulse.

As a simple theoretical model, we use the edge-state picture of the quantum Hall effect. In a circular geometry, it yields a system of compressible and incompressible rings formed in the electron liquid (Fig. 1D) owing to the interplay between LLs and electron interactions (4). In our measurement, interaction effects are already

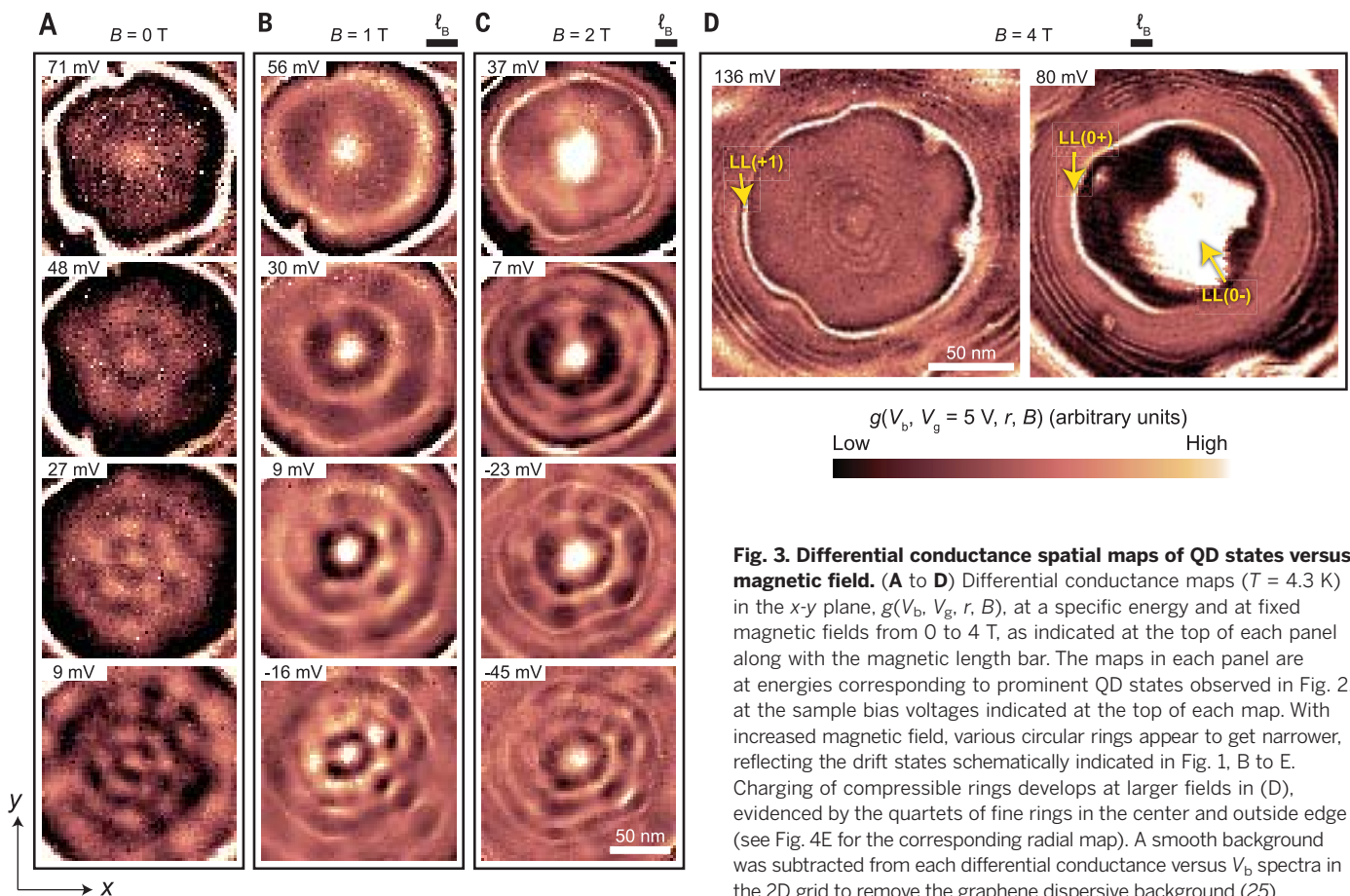


Fig. 3. Differential conductance spatial maps of QD states versus magnetic field. (A to D) Differential conductance maps ($T = 4.3$ K) in the x - y plane, $g(V_b, V_g, r, B)$, at a specific energy and at fixed magnetic fields from 0 to 4 T, as indicated at the top of each panel along with the magnetic length bar. The maps in each panel are at energies corresponding to prominent QD states observed in Fig. 2, at the sample bias voltages indicated at the top of each map. With increased magnetic field, various circular rings appear to get narrower, reflecting the drift states schematically indicated in Fig. 1, B to E. Charging of compressible rings develops at larger fields in (D), evidenced by the quartets of fine rings in the center and outside edge (see Fig. 4E for the corresponding radial map). A smooth background was subtracted from each differential conductance versus V_b spectra in the 2D grid to remove the graphene dispersive background (25).

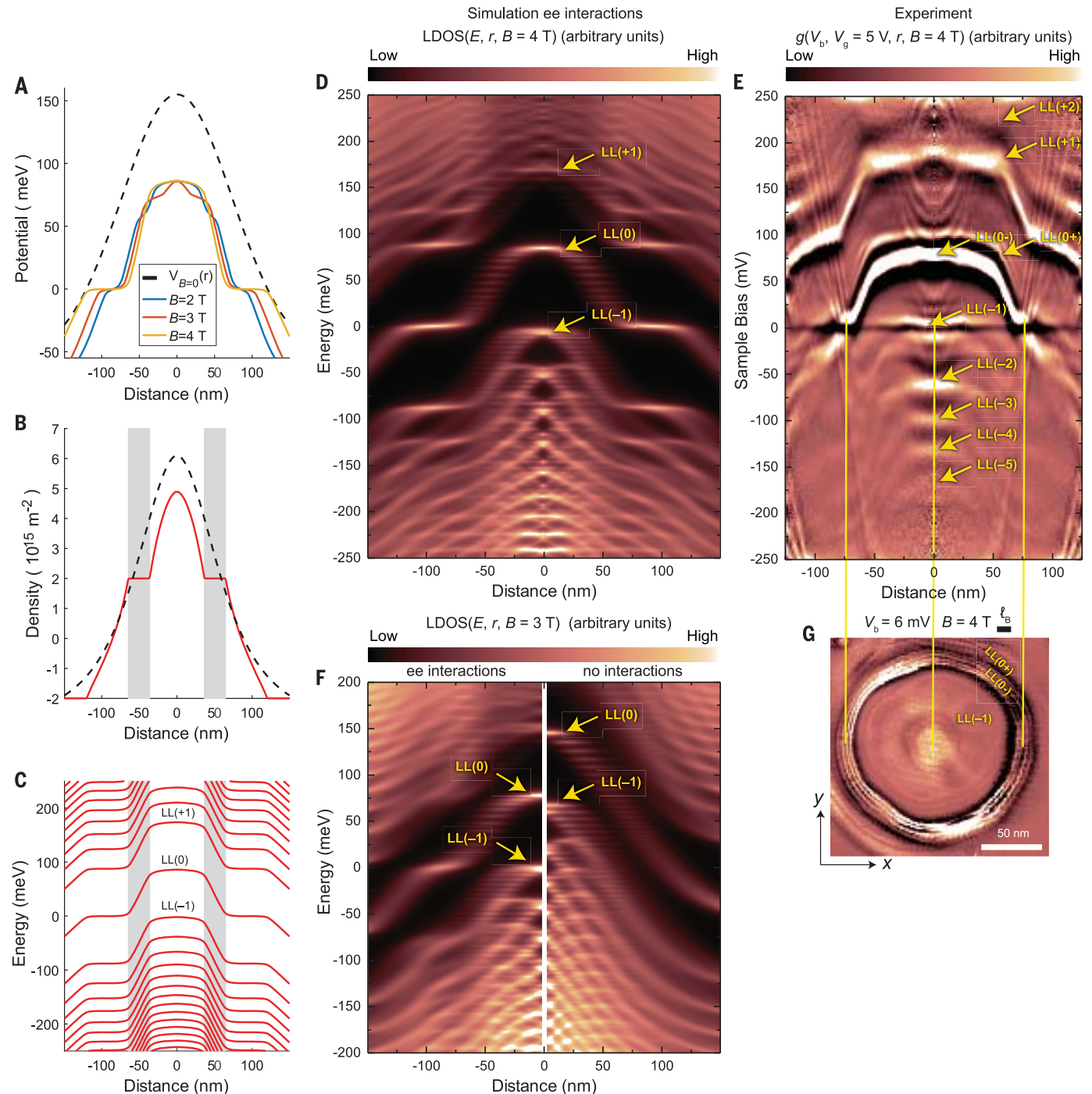


Fig. 4. Electron interactions and the wedding cake-like structure. (A) The effective potential, Eq. 3, for several magnetic field values (solid lines) and the $V_{B=0}(r)$ fit based on Fig. 2A (dashed line). (B) Carrier density and (C) LLs at $B = 4$ T as predicted by the model from Eq. 1. Screening produces compressible regions, where LLs are flat and pinned at the Fermi level, separated by incompressible regions (marked in gray). The size of incompressible regions is estimated in Eq. 4. The dashed line in (B) describes charge density in the compressible limit obtained by excluding the kinetic term $K[n(r)]$ from the functional in Eq. 1. (D) LDOS map calculated with the screened potential from (A). A Fermi velocity

of 1.2×10^6 m/s was used in the calculations to match the LL positions between theory and experiment at $B = 4$ T. (E) Experimental differential conductance map ($T = 4.3$ K), $g(V_b, V_g, r, B)$, as a function of V_b and r at $B = 4$ T, showing the wedding cake-like structure in the LLs in the QD. (F) LDOS simulated by using the potential at $B = 3$ T and $V_{B=0}(r)$ from (A) (see text for discussion). (G) An x-y slice of the $g(V_b, V_g, r, B)$ map in (E) at $V_b = 6$ mV (near the Fermi level), showing the inner compressible disk from LL(-1) and the outer compressible ring from LL(0), as schematically indicated in Fig. 1E. The solid yellow lines show where LL(0) and LL(-1) cut through the Fermi level, creating the compressible rings.

observed at low fields, signaled by the shifting and flattening of the LL states in Fig. 2. A minimal model incorporating interactions at low fields is the energy functional (4)

$$E[n] = \int d^2r \{K[n(r)] + V_{\text{ext}}(r)n(r) + \frac{1}{2} \int d^2r' V_{\text{ee}}(|r - r'|)n(r)n(r')\} \quad (1)$$

where $n(r)$ is charge density at position r . This functional describes the competition between carriers' kinetic energy $K[n(r)]$ and the effective potential $V_B(r)$, which includes the effects of screening. We approximate these quantities as

$$\frac{\delta K}{\delta n(r)} = \epsilon_N, \quad N - \frac{1}{2} < \nu(r) < N + \frac{1}{2} \quad (2)$$

$$V_B(r) = V_{\text{ext}}(r) + \int d^2r' V_{\text{ee}}(|r - r'|)n(r') \quad (3)$$

where $\nu = n/n_{\text{LL}}$ is the filling fraction. These relations are valid in the limit $l_B \ll l_V$. Here $N = 0, \pm 1, \pm 2, \dots$ is the LL number, V_{ext} is the electrostatic potential defining the dot, $V_{\text{ee}}(r) = \frac{e^2}{r}$ is the Coulomb interaction and \tilde{e} is the screened electron charge [see (25) for details].

The calculated effective potential $V_B(r)$ is shown in Fig. 4A for a few magnetic field values. The joint effect of the magnetic field and interactions is the creation of a series of plateaus that form a wedding cake-like pattern of concentric rings within the dot. At the same time, the potential $V_B(r)$ is reduced compared to $V_{B=0}(r)$. Notably, this causes LL(0) to move toward the Fermi level in agreement with the energy dependence of LL(0) in Fig. 2. The extra concave features in the potential in the central region match those in the experimental maps of LL(0) in Fig. 2, F to I. The effect of interactions on the LLs is shown by comparing the LDOS with and without interactions in the left and right panels of Fig. 4F. Before interactions are turned on (right panel), the LLs seen through the LDOS essentially track the potential $V_{B=0}(r)$. After including interactions (left panel), the evolution of the LDOS mimics that of potential $V_B(r)$: LLs shift to lower energy and flatten in the central region, in agreement with the evolution seen in the measurements (Fig. 2).

The incompressible and compressible rings become considerably clearer in higher magnetic field. The experimental spectral map in Fig. 4E shows the LLs becoming flat in the central region of the QD, even though the bare external electrostatic potential is concave (see QD outline in Fig. 2A), and then they progress sharply to a new energy level as new LLs become occupied, forming a wedding cake-like structure. Here LL(N), where $N = -5$ to 2, can be observed as plateaus in the center of the QD (Fig. 4E). Both LL(0) and LL(-1) cross the Fermi level at zero bias, as indicated by the yellow lines, forming a LL(-1) compressible disk in the center and an outer LL(0) compressible ring separated by an

incompressible ring, as shown in the Fermi-level spatial map in the x - y plane (Fig. 4G). We observe Coulomb charging of these LLs as charging lines intersecting the LLs at the Fermi level and progressing upward at sharp angles in Fig. 4E. These lines correspond to a quartet of rings in Figs. 3D and 4G. The charging of the compressible regions occurs in groups of four, reflecting the fourfold (spin and valley) graphene degeneracy (13).

To understand these observations, we use a two-stage approach. We first use the mean field functional Eq. 1 to find LL occupancies and determine the screened potential $V_B(r)$ (25). We then use this potential to calculate the density of microscopic states, which can be directly compared to the measurements. The features seen in the measured LDOS can be understood by comparing them to a simple calculation of the LLs (25), shown in Fig. 4C. The highest LL that is partially filled can be obtained by counting the number of LLs that need to be populated to accommodate the carrier density equal to that in the fully compressible regime (dashed line in Fig. 4B). In the simulated LDOS map in Fig. 4D, we can identify the LL states, which track the screened potential $V_B(r)$ pictured in Fig. 4A. These states exhibit plateaus, as expected from theory (4, 5). This behavior is in good agreement with the experimental results shown in Fig. 4E.

The width of the observed incompressible ring can be estimated from the functional in Eq. 1 following the approach of (5) and yields the strip width (25),

$$l = \left(\frac{4\Delta\epsilon_{\text{LL}}}{\pi^2 \tilde{e}^2 \frac{dn}{dr}} \right)^{\frac{1}{2}} \approx 34 \text{ nm} \quad (4)$$

The estimate in Eq. 4 is slightly greater than the width inferred from our measurement results shown in Fig. 4E. The discrepancy can partly be attributed to the result of (5), derived for LL spacing $\Delta\epsilon_{\text{LL}}$ much smaller than the external potential, being used in the regime when $\Delta\epsilon_{\text{LL}}$ is not small on the $V_B(r)$ scale.

Fingerprints of electron-electron interactions that are as clear and notable as the observed electronic wedding cake-like patterns are relatively rare in solid-state experiments. The measurements reported here suggest, as hinted by the charging lines and nodal patterns in the differential conductance maps, that even more exotic signatures of electronic interactions may be within experimental reach in future scanned-probe measurements at lower temperatures.

REFERENCES AND NOTES

1. S. M. Reimann, M. Manninen, *Rev. Mod. Phys.* **74**, 1283–1342 (2002).
2. R. C. Ashoori, *Nature* **379**, 413–419 (1996).
3. L. Kouwenhoven et al., *Mesoscopic Electron Transp.* **345**, 105–214 (1997).
4. P. L. McEuen et al., *Phys. Rev. B Condens. Matter* **45**, 11419–11422 (1992).

5. D. B. Chklovskii, B. I. Shklovskii, L. I. Glazman, *Phys. Rev. B Condens. Matter* **46**, 4026–4034 (1992).
6. D. B. Chklovskii, K. A. Matveev, B. I. Shklovskii, *Phys. Rev. B Condens. Matter* **47**, 12605–12617 (1993).
7. Y. V. Nazarov, A. V. Khaetskii, *Phys. Rev. B Condens. Matter* **49**, 5077–5080 (1994).
8. M. M. Fogler, E. I. Levin, B. I. Shklovskii, *Phys. Rev. B Condens. Matter* **49**, 13767–13775 (1994).
9. S. Fölling, A. Wiedera, T. Müller, F. Gerbier, I. Bloch, *Phys. Rev. Lett.* **97**, 060403 (2006).
10. N. Gmelke, X. Zhang, C.-L. Hung, C. Chin, *Nature* **460**, 995–998 (2009).
11. J. Martin et al., *Nat. Phys.* **5**, 669–674 (2009).
12. D. L. Miller et al., *Science* **324**, 924–927 (2009).
13. S. Jung et al., *Nat. Phys.* **7**, 245–251 (2011).
14. J. Chae et al., *Phys. Rev. Lett.* **109**, 116802 (2012).
15. M. Zarenia, A. Chaves, G. A. Farias, F. M. Peeters, *Phys. Rev. B* **84**, 245403 (2011).
16. J.-B. Qiao et al., *Phys. Rev. B* **95**, 081409 (2017).
17. Y. Jiang et al., *Nat. Nanotechnol.* **12**, 1045–1049 (2017).
18. Y. Zhao et al., *Science* **348**, 672–675 (2015).
19. N. M. Freitag et al., *Nano Lett.* **16**, 5798–5805 (2016).
20. J. Lee et al., *Nat. Phys.* **12**, 1032–1036 (2016).
21. C. Gutiérrez, L. Brown, C.-J. Kim, J. Park, A. N. Pasupathy, *Nat. Phys.* **12**, 1069–1075 (2016).
22. F. Ghahari et al., *Science* **356**, 845–849 (2017).
23. A. Matulis, F. M. Peeters, *Phys. Rev. B* **77**, 115423 (2008).
24. J. F. Rodriguez-Nieva, L. S. Levitov, *Phys. Rev. B* **94**, 235406 (2016).
25. Additional supplementary text and data are available in the supplementary materials.
26. M. Morgenstern, D. Haude, J. Klijn, R. Wiesendanger, *Phys. Rev. B* **66**, 121102 (2002).
27. S. Becker et al., *Phys. Rev. Lett.* **106**, 156805 (2011).
28. Y. J. Song et al., *Nature* **467**, 185–189 (2010).
29. A. V. Filinov, M. Bonitz, Y. E. Lozovik, *Phys. Rev. Lett.* **86**, 3851–3854 (2001).
30. C.-H. Zhang, Y. Joglekar, *Phys. Rev. B* **75**, 245414 (2007).
31. K. A. Guerrero-Becerra, M. Rontani, *Phys. Rev. B* **90**, 125446 (2014).

ACKNOWLEDGMENTS

We thank S. Blankenship and A. Band for their contributions to this project and M. Stiles for valuable discussions. **Funding:** C.G., D.W., and F.G. acknowledge support from the Cooperative Research Agreement between the University of Maryland and the National Institute of Standards and Technology Center for Nanoscale Science and Technology, grant no. 70NANB10IH193, through the University of Maryland. C.L. and L.S.L. acknowledge support from the STC Center for Integrated Quantum Materials (CIQM) under NSF award 1231319. L.S.L. also acknowledges support from the U.S. Army Research Laboratory and the U.S. Army Research Office through the Institute for Soldier Nanotechnologies, under contract number W911NF-13-D-0001. J.F.R.-N. acknowledges support from the Harvard-MIT Center for Ultracold Atoms, NSF grant no. DMR-1308435. K.W. and T.T. acknowledge support for the growth of hexagonal boron nitride crystals from the Elemental Strategy Initiative conducted by the MEXT, Japan, and the CREST (JPMJCR15F3). **JST. Author contributions:** C.G., D.W., and F.G. performed the experiments. C.G. led the STM-STS data analysis. F.G. designed and fabricated the graphene device. J.A.S. designed and built the apparatus and conceived of the study. C.L., J.F.R.-N., and L.S.L. performed the simulations. K.W. and T.T. grew the hBN crystals used in the graphene device. All authors contributed to writing the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data are available in the manuscript and supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/789/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 and S2
References (32–40)
Movie S1
Data S1

12 October 2017; accepted 15 June 2018
10.1126/science.aar2014

MAGNETISM

Observation of Dicke cooperativity in magnetic interactions

Xinwei Li¹, Motoaki Bamba^{2,3}, Ning Yuan⁴, Qi Zhang⁵, Yage Zhao⁶, Maolin Xiang⁴, Kai Xu⁴, Zuanming Jin⁴, Wei Ren⁴, Guohong Ma⁴, Shixun Cao^{4*}, Dmitry Turchinovich^{7,8}, Junichiro Kono^{1,9,10*}

The interaction of N two-level atoms with a single-mode light field is an extensively studied many-body problem in quantum optics, first analyzed by Dicke in the context of superradiance. A characteristic of such systems is the cooperative enhancement of the coupling strength by a factor of \sqrt{N} . In this study, we extended this cooperatively enhanced coupling to a solid-state system, demonstrating that it also occurs in a magnetic solid in the form of matter-matter interaction. Specifically, the exchange interaction of N paramagnetic erbium(III) (Er^{3+}) spins with an iron(III) (Fe^{3+}) magnon field in erbium orthoferrite (ErFeO_3) exhibits a vacuum Rabi splitting whose magnitude is proportional to \sqrt{N} . Our results provide a route for understanding, controlling, and predicting novel phases of condensed matter using concepts and tools available in quantum optics.

When an ensemble of two-level atoms interacts with a single-mode long-wavelength light field, coherence can develop within the ensemble through photon exchange; the interaction becomes cooperative. This phenomenon, captured by the Dicke superradiance model (1), has profound consequences in cavity quantum electrodynamics (QED) research. In a photonic cavity, the coupling rate Λ between N dipoles and a quantized vacuum photon field is cooperatively enhanced by a factor of \sqrt{N} (2). The ground state of the atom ensemble is predicted to be unstable against a phase transition, known as the superradiant phase transition (SRPT), when Λ reaches a critical value (3). This possibility has stimulated much recent interest in condensed matter cavity QED systems consisting of N dipoles with very large moments (4–7).

The role light plays in the Dicke model can be performed by any fundamental excitation—such as lattice waves (phonons) and spin waves (magnons)—that can be bosonized and quantized in the same way as photons. This concept of Dicke physics without light is crucial for understanding phase transitions in condensed matter. The most relevant example is the cooperative Jahn-Teller (JT) effect, which describes the dynamics of an ensemble of pseudospins with degenerate electronic levels cooperatively coupled

with a phonon mode in the same material. The cooperative coupling leads to a phase transition that distorts the lattice and breaks the degeneracy of the pseudospin energy levels. Theoretically, the transition is believed to be analogous to the SRPT (8); the displacive lattice distortion is comparable to the appearance of a static electromagnetic field in the photon SRPT. Although these theories can explain JT and JT-like transitions phenomenologically, unambiguous evidence for the cooperative coupling of two matter subsystems in one material system is still lacking.

Here we report cooperative exchange coupling of a spin ensemble with a vacuum magnon field within a solid. We used Y^{3+} -doped single-crystal ErFeO_3 samples—namely, $\text{Er}_x\text{Y}_{1-x}\text{FeO}_3$ —and systematically studied the doping, temperature, and magnetic field dependence of their terahertz (THz) absorption spectra. The Er^{3+} electron paramagnetic resonance (EPR) in an external magnetic field strongly coupled with a vacuum magnon mode of the ordered Fe^{3+} spins. The situation is analogous to a standard N -atom cavity QED experiment, in which an ensemble of N two-level atoms couples with a vacuum photon field in an optical cavity. The Fe^{3+} - Er^{3+} coupling rate showed a characteristic scaling behavior with the density of Er^{3+} ions, evidencing Dicke cooperativity. By analyzing this scaling behavior with our micro-

scopic theoretical model, we determined the Fe^{3+} - Er^{3+} exchange coupling constants. These constants are important for understanding the widely discussed 3d-4f magnetic coupling that is responsible for many exotic phenomena in a variety of compounds: examples include novel magnetic phase transitions (9), magnetoelectric effects (10, 11), electromagnons (12), nonlinear spin excitations (13), and heavy fermions (14).

ErFeO_3 crystallizes in an orthorhombic perovskite structure (Fig. 1A) that can be described by the space group $D_{2h}^{16}-Pbnm$. The Fe^{3+} spins order antiferromagnetically below 650 K. Many magnetic phases—such as the Γ_4 , Γ_{24} , Γ_2 , and Γ_{12} phases—can appear as a function of temperature (15). In the temperature range 4.5 K < T < 85 K, the crystal is in the Γ_2 phase, in which the spins in the two Fe^{3+} sublattices, \mathbf{S}_1 and \mathbf{S}_2 , are antiparallel along the c axis but cant toward the a axis by a small angle β , owing to the antisymmetric Dzyaloshinskii-Moriya interaction (Fig. 1, A and B). The sum of the two spins $\mathbf{S}_+ = \mathbf{S}_1 + \mathbf{S}_2$ induces $\mathbf{M}_{\text{Fe}} \parallel \mathbf{a}$, where \mathbf{M}_{Fe} is the macroscopic magnetization vector of the Fe^{3+} subsystem. The quasi-ferromagnetic (qFM) magnon mode of the Fe^{3+} subsystem can be selectively excited by using linearly polarized THz radiation with $\mathbf{H}_{\text{THz}} \perp \mathbf{a}$ (16, 17), where \mathbf{H}_{THz} is the magnetic component of the THz electromagnetic field. Figure 1B shows how the Fe^{3+} spins oscillate in the qFM mode. \mathbf{S}_1 and \mathbf{S}_2 oscillate in phase while the angle between them remains constant, so the model can be reduced to the precession of the combined spin \mathbf{S}_+ about the a axis.

On the other hand, the Er^{3+} ions ($4f^{11}$) occupy low-symmetry sites in the crystal. The crystal field forms Kramers doublets; each doublet

¹Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005, USA. ²Department of Materials Engineering Science, Osaka University, Osaka 560-8531, Japan. ³PRESTO, Japan Science and Technology Agency, Saitama 332-0012, Japan. ⁴Department of Physics, International Center of Quantum and Molecular Structures, and Materials Genome Institute, Shanghai University, Shanghai 200444, China. ⁵Argonne National Laboratories, Lemont, IL 60439, USA. ⁶School of Physics, Peking University, Beijing 100871, China. ⁷Fakultät für Physik, Universität Duisburg-Essen, Lotharstr. 1, Duisburg 47057, Germany. ⁸Max Planck Institute for Polymer Research, Ackermannweg 10, Mainz 55128, Germany. ⁹Department of Materials Science and NanoEngineering, Rice University, Houston, TX 77005, USA. ¹⁰Department of Physics and Astronomy, Rice University, Houston, TX 77005, USA. *Corresponding author. Email: kono@rice.edu (J.K.); sxcao@shu.edu.cn (S.C.)

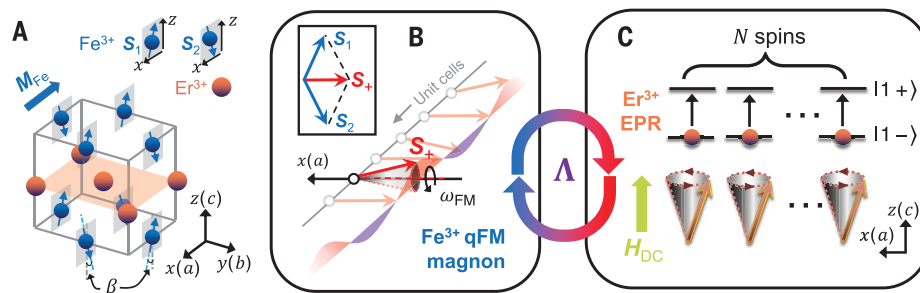


Fig. 1. Cooperatively coupled N -spin-magnon system as an analog of an N -atom cavity QED system. (A) Crystal and magnetic structure of ErFeO_3 in the Γ_2 phase. (B) Oscillations of spins in the qFM magnon mode of the Fe^{3+} subsystem in the Γ_2 phase. (C) The EPR of the Er^{3+} spin ensemble is the transition between the lowest Kramers-doublet states. The two matter subsystems, illustrated in (B) and (C), are resonantly coupled with the coupling constant Λ .

consists of a pair of time-reversed states. As shown in Fig. 1C, when an external static magnetic field (\mathbf{H}_{DC}) is applied, and in the low-temperature limit, the transition between the lowest two time-reversed states ($|1-\rangle$ and $|1+\rangle$) can be interpreted as the EPR of the ion. In a classical picture, the EPR corresponds to the Larmor precession of spins about \mathbf{H}_{DC} . The $N\text{Er}^{3+}$ spins within a characteristic volume interact cooperatively with the Fe^{3+} qFM magnon mode at rate Λ . The rate can be determined by mapping out anticrossings through THz magnetospectroscopy.

Figure 2A shows the THz absorption spectra obtained from a *c*-cut ErFeO_3 single crystal; the measurement configuration was $\mathbf{H}_{\text{THz}} \parallel \mathbf{b}$ and $\mathbf{H}_{\text{DC}} \parallel \mathbf{c}$, and the temperature was 45 K (the crystal was in the Γ_2 phase, both in this experiment and the ones described hereafter). On the basis of a comprehensive series of complementary measurements and theoretical analyses (18), we assigned the constant-frequency line to be the Fe^{3+} qFM magnon mode and the line that increases linearly with \mathbf{H}_{DC} to be the Er^{3+} EPR mode.

We observed a clear temperature-dependent magnon-EPR anticrossing behavior (Fig. 2, B to E). The zero-detuning magnetic field, H_0 , is the magnitude of \mathbf{H}_{DC} at which the uncoupled magnon and EPR mode frequencies are equal; this frequency is the zero-detuning transition frequency, ω_0 . Experimentally, H_0 also coincides with the magnitude of \mathbf{H}_{DC} at which the frequency difference between the hybridized branches reaches a minimum. The frequency splitting between the hybridized branches at H_0 , which we call the vacuum Rabi splitting $\Omega(H_0)$, is equal to twice the magnon-EPR coupling strength Λ . The magnitude of Λ increases with decreasing temperature, indicating that the Fe^{3+} - Er^{3+} coupling becomes stronger.

Here we provide some interpretations to several features in the data. First, there is some fine structure in the middle of the anticrossing region. This can be modeled by considering two species of Er^{3+} ions (due to the \mathbf{H}_{DC} -induced sublattice degeneracy breaking) interacting with the Fe^{3+} magnons; however, the feature can be neglected if we effectively consider the two Er^{3+} species as one (18). Next, the intensity variations of the absorption lines are due to the \mathbf{H}_{DC} -induced mixing of crystal field levels. Finally, the magnon frequency bowing for the 5 K spectra is due to the spin structure instability, as the temperature is close to the critical point where a Γ_2 -to- Γ_{12} spin reorientation transition occurs (18).

To determine how Λ grows with the Er^{3+} spin density, we studied Y^{3+} -doped ErFeO_3 single crystals ($\text{Er}_x\text{Y}_{1-x}\text{FeO}_3$). It has been previously shown that the nonmagnetic Y^{3+} ions simply reduce the density of Er^{3+} spins by a factor of x without changing the crystal or magnetic structure of the system (19). In Fig. 2, panels F to H (and I to K, respectively) show temperature-dependent anticrossing spectra for the $x = 0.75$ ($x = 0.5$) sample. An increasing Λ at lower temperatures is again found. In addition, comparison among Fig. 2, C, F, and I, indicates that, at the same temperature

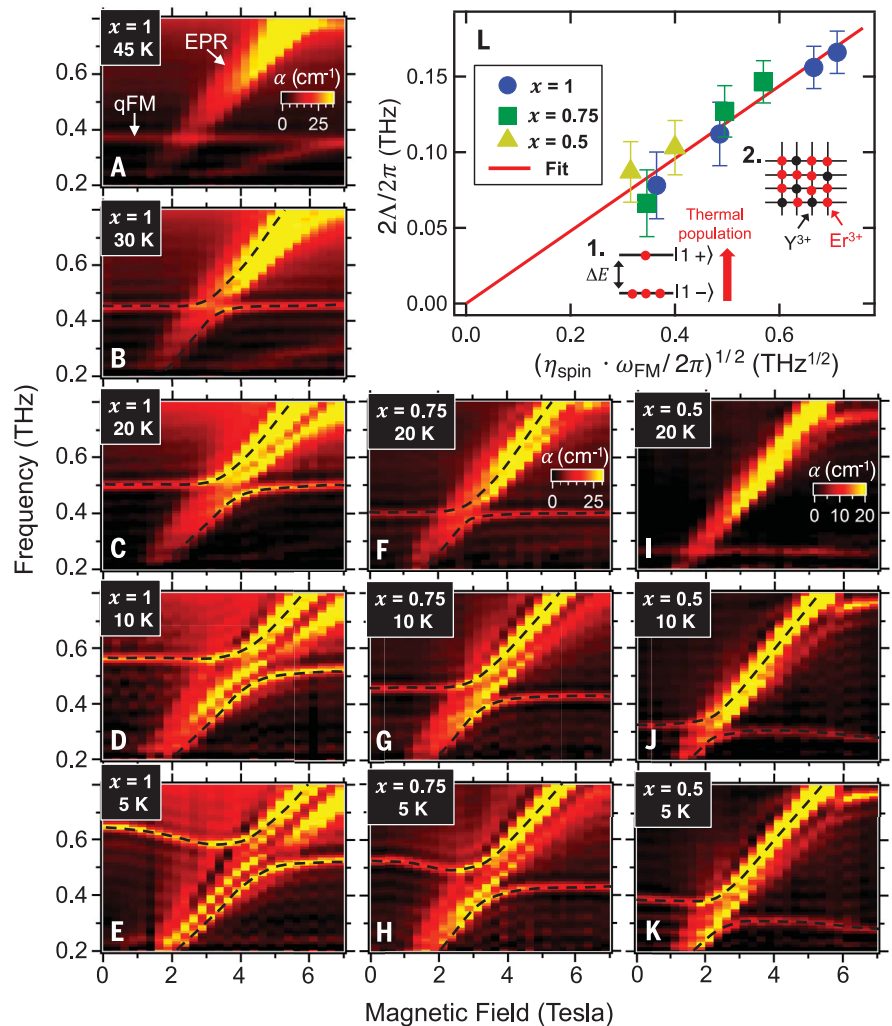


Fig. 2. Experimental evidence for cooperative coupling between paramagnetic Er^{3+} spins and Fe^{3+} vacuum magnons. (A to K) Absorption spectra measured in the $\mathbf{H}_{\text{THz}} \parallel \mathbf{b}$ and $\mathbf{H}_{\text{DC}} \parallel \mathbf{c}$ configurations at various temperatures and Y^{3+} doping levels. Dashed black lines are guides to the eye for identifying the hybridized modes. (L) The coupling rate Λ shows proportionality with $\sqrt{\eta_{\text{spin}} \omega_{\text{FM}}}$. Error bars indicate the uncertainties in determining mode frequencies due to finite mode linewidths. The inset shows two types of mechanisms that determine η_{spin} in the measurements shown in (A) to (K).

(20 K), a larger value of x leads to a larger Λ . The same trend can also be observed from the doping-dependent data at 10 K (Fig. 2, D, G, and J) and 5 K (Fig. 2, E, H, and K). Notably, the system exhibited the largest $\Lambda/\omega_0 = 0.18$ when $x = 0.5$ and $T = 5$ K (Fig. 2K); this puts the system into the ultrastrong coupling regime, defined by $\Lambda/\omega_0 > 0.1$ in the language of cavity QED (20, 21). Most importantly, we found that $\Lambda \propto \sqrt{n_{\text{spin}} \omega_{\text{FM}}}$, where n_{spin} is the net density of EPR-contributing Er^{3+} spins (a function of T , H_0 , and x) and ω_{FM} is the Fe^{3+} qFM magnon frequency. Below, we discuss the physical meaning of this scaling law in terms of Dicke cooperativity.

First, we can understand how n_{spin} varies with T , H_0 , and x , as depicted in the inset to Fig. 2L. At low T , we can approximate the Er^{3+} ions to be two-level systems that obey the Boltzmann statistics (22). Calculating the populations of the spin-up

($|1+\rangle$) and spin-down ($|1-\rangle$) states at T leads to the Curie law, $\langle \mu_{\text{Er}} \rangle / \mu_{\text{sat}} = \tanh(-\Delta E(T, H_0)/2k_B T)$, where $\langle \mu_{\text{Er}} \rangle$ is the average Er^{3+} magnetic moment per spin; μ_{sat} is the Er^{3+} saturation moment per spin, corresponding to $T = 0$; $\Delta E(T, H_0)$ is the T - and H_0 -dependent energy level separation between $|1+\rangle$ and $|1-\rangle$; and k_B is the Boltzmann constant. Hence, $\langle \mu_{\text{Er}} \rangle$ monotonically increases as T decreases. In addition, the doping by nonmagnetic Y^{3+} ions simply reduces the density of Er^{3+} spins by a factor of x through dilution. Combining these effects, we can express the net Er^{3+} spin density as $n_{\text{spin}}(T, H_0, x) = \eta_{\text{spin}} \cdot n_{\text{spin}}^0$, where $\eta_{\text{spin}} \equiv x \cdot \tanh(-\Delta E/2k_B T)$ is the net fraction of EPR-contributing spins and $n_{\text{spin}}^0 \equiv 2/V_0$ is the total Er^{3+} spin density in ErFeO_3 with V_0 being the unit-cell volume.

Second, we explain why the Fe^{3+} qFM magnon frequency ω_{FM} appears in the scaling equation

of Λ . In a semiclassical description of two-level systems interacting with a cavity light field, a photonic mode frequency that is analogous to ω_{FM} here also emerges in the expression of the vacuum Rabi splitting. This is due to the fact that the vacuum Rabi splitting is proportional to the vacuum fluctuation field of a quantized boson mode, and the vacuum field amplitude itself is proportional to the square root of the boson mode frequency. In our case, the Fe^{3+} vacuum magnon field scales with $\sqrt{\omega_{\text{FM}}}$ and affects the hybrid Fe^{3+} - Er^{3+} mode frequency splitting. In our experiments, ω_{FM} remains essentially independent of H_0 , but it is x - and T -dependent because both doping and temperature change the Fe^{3+} *ac* plane anisotropy.

On the basis of the above considerations, we extracted the values of Λ for all spectra in Fig. 2, A to K, where a spectroscopic anticrossing is resolvable and plotted these values against $\sqrt{\eta_{\text{spin}} \omega_{\text{FM}}}$ in Fig. 2L. As shown in the figure, all data points fall onto a line that passes through the origin. This scaling behavior for the vacuum Rabi splitting evidences that the Er^{3+} spins are cooperatively coupled to the Fe^{3+} vacuum magnon field.

To extract more quantitative information from our experimental data, we developed a microscopic theoretical model (18). For the Fe^{3+} subsystem, we followed the treatment of Herrmann (23). The Hamiltonian \hat{H}_{Fe} takes into account the symmetric exchange, antisymmetric exchange, and anisotropy of the spins in the two sublattices, \mathbf{S}_{2l-1} and \mathbf{S}_{2l} , where $l = 1, 2, \dots, N_{\text{UC}}$ is the unit-cell index (with N_{UC} being the total number of unit cells). For the Er^{3+} subsystem, we modeled the EPR associated with the $|1- \rangle \rightarrow |1+ \rangle$ transition as the Larmor precession of an Er^{3+} spin \mathbf{R}_l in the l th unit cell. The EPR Hamiltonian is $\hat{H}_{\text{Er}} = \sum_{l=1}^{N_{\text{UC}}} -\hat{\boldsymbol{\mu}}_l \cdot \boldsymbol{\mu}_0 \mathbf{H}_{\text{DC}}$, where $\hat{\boldsymbol{\mu}}_l$ is the magnetic moment, expressed as the dot product of the anisotropic Landé g -factor and \mathbf{R}_l , and μ_0 is the vacuum permeability.

The Er^{3+} EPR interacts with the Fe^{3+} qFM magnon (represented by the oscillation of $\mathbf{S}_+ = \mathbf{S}_{2l-1} + \mathbf{S}_{2l}$) through both the symmetric and antisymmetric exchange interactions (9). The Fe^{3+} - Er^{3+} coupling Hamiltonian for ErFeO_3 is

$$\hat{H}_{\text{Fe-Er}} = \sum_{l=1}^{N_{\text{UC}}} \left[J_1 \hat{\mathbf{R}}_l \cdot \hat{\mathbf{S}}_{2l-1} + J_2 \hat{\mathbf{R}}_l \cdot \hat{\mathbf{S}}_{2l} + \mathbf{D}_1 \cdot (\hat{\mathbf{R}}_l \times \hat{\mathbf{S}}_{2l-1}) + \mathbf{D}_2 \cdot (\hat{\mathbf{R}}_l \times \hat{\mathbf{S}}_{2l}) \right] \quad (1)$$

where J_1 and J_2 (\mathbf{D}_1 and \mathbf{D}_2) are the symmetric (antisymmetric) exchange constants.

From the total Hamiltonian $\hat{H} = \hat{H}_{\text{Fe}} + \hat{H}_{\text{Er}} + \hat{H}_{\text{Fe-Er}}$, we derived

$$\Lambda = \frac{J}{\hbar} \sqrt{\frac{n_{\text{spin}} V_0 \omega_{\text{FM}}}{4\gamma J_{\text{Fe}}}} = \frac{J}{\hbar \sqrt{2\gamma J_{\text{Fe}}}} \sqrt{\eta_{\text{spin}} \omega_{\text{FM}}} \quad (2)$$

where $J_1 = J_2 = J$ due to symmetry arguments, γ is the gyromagnetic ratio, and J_{Fe} is the isotropic exchange constant between the Fe^{3+} spins; J_{Fe} has been determined through previous experiments (24); and \hbar is Planck's constant h divided by 2π . Equation 2 reveals the characteristic cooperative scaling behavior of Λ , supporting our experimental observation shown in Fig. 2. We determined the exchange constant $J = 2.95$ meV by fitting the slope of the $\Lambda \propto \sqrt{\eta_{\text{spin}} \omega_{\text{FM}}}$ line in Fig. 2L using Eq. 2.

Our observed Dicke cooperativity in ErFeO_3 opens up possibilities to realize a SRPT in thermal equilibrium. A magnon SRPT in a spin-magnon interaction system is much more feasible than a photon SRPT based on an electrical-dipolar-type light-matter interaction system (25). It has been speculated that some magnetic phase transitions in these orthoferrite compounds (such as ErFeO_3) are magnon SRPTs (26). Our work demonstrates the key physics required to rigorously analyze these phenomena. Furthermore, another inspiration from our observation is to control condensed matter phases on the basis of boson-mediated long-range interactions. In JT and JT-like transitions, the critical temperatures can be manipulated by changing the matter-matter interaction strength through tuning the density of the pseudospins. This raises the future possibility that Dicke physics, a well-established concept in quantum optics, can provide a guideline for controlling and engineering novel entangled phase transitions of condensed matter.

REFERENCES AND NOTES

1. R. H. Dicke, *Phys. Rev.* **93**, 99–110 (1954).
2. M. Tavis, F. W. Cummings, *Phys. Rev.* **188**, 692–695 (1969).
3. K. Hepp, E. H. Lieb, *Ann. Phys.* **76**, 360–404 (1973).
4. C. Weisbuch, M. Nishioka, A. Ishikawa, Y. Arakawa, *Phys. Rev. Lett.* **69**, 3314–3317 (1992).
5. Y. Tabuchi *et al.*, *Phys. Rev. Lett.* **113**, 083603 (2014).
6. G. Günter *et al.*, *Nature* **458**, 178–181 (2009).
7. Q. Zhang *et al.*, *Nat. Phys.* **12**, 1005–1011 (2016).
8. D. Porras, P. A. Ivanov, F. Schmidt-Kaler, *Phys. Rev. Lett.* **108**, 235701 (2012).
9. T. Yamaguchi, *J. Phys. Chem. Solids* **35**, 479–500 (1974).
10. T. Lottermoser *et al.*, *Nature* **430**, 541–544 (2004).
11. N. J. Laurita *et al.*, *Phys. Rev. Lett.* **119**, 227601 (2017).
12. L. Chaix *et al.*, *Phys. Rev. Lett.* **112**, 137201 (2014).
13. S. Baierl *et al.*, *Nat. Photonics* **10**, 715–718 (2016).
14. G. R. Stewart, *Rev. Mod. Phys.* **56**, 755–787 (1984).
15. R. L. White, *J. Appl. Phys.* **40**, 1061–1069 (1969).

16. K. Yamaguchi, T. Kurihara, Y. Minami, M. Nakajima, T. Suemoto, *Phys. Rev. Lett.* **110**, 137204 (2013).
17. J. Lu *et al.*, *Phys. Rev. Lett.* **118**, 207204 (2017).
18. See supplementary materials.
19. D. L. Wood, J. P. Remeika, L. M. Holmes, E. M. Gyorgy, *J. Appl. Phys.* **40**, 1245–1246 (1969).
20. Y. Todorov *et al.*, *Phys. Rev. Lett.* **105**, 196402 (2010).
21. P. Forn-Díaz, L. Lamata, E. Rico, J. Kono, E. Solano, arXiv:1804.09275 [quant-ph] (9 May 2018).
22. K. Zhang *et al.*, *Sci. Rep.* **6**, 23648 (2016).
23. G. Herrmann, *J. Phys. Chem. Solids* **24**, 597–606 (1963).
24. N. Koshizuka, K. Hayashi, *J. Phys. Soc. Jpn.* **57**, 4418–4428 (1988).
25. J. M. Knight, Y. Aharonov, G. T. C. Hsieh, *Phys. Rev. A* **17**, 1454–1462 (1978).
26. A. M. Kadomtseva, I. B. Krynetskii, V. M. Matveev, *Sov. Phys. JETP* **52**, 732–737 (1980).

ACKNOWLEDGMENTS

We thank A. H. MacDonald, P. Lapas, and O. Gomonay for useful discussions. **Funding:** This research was primarily supported by the National Science Foundation through the Center for Dynamics and Control of Materials: an NSF MRSEC under Cooperative Agreement DMR-1720595. J.K. acknowledges support from the U.S. Army Research Office (grant W911NF-17-1-0259). M.B. was supported by the JST PRESTO program (grant JPMJPR1767); KAKENHI (grant 26287087); and the ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, government of Japan). D.T. acknowledges financial support from the Deutsche Forschungsgemeinschaft (SFB 1242 “Non-Equilibrium Dynamics of Condensed Matter in the Time Domain,” TP B08), the European Commission (EU Career Integration grant EU CIG 334324 LIGHTER), and the Max Planck Society. S.C. acknowledges support from the National Natural Science Foundation of China (11774217). G.M. acknowledges support from the National Natural Science Foundation of China (11674213 and 61735010). Z.J. acknowledges support from the National Natural Science Foundation of China (11604202), Shanghai Municipal Education Commission (Young Eastern Scholar QD2015020), Science and Technology Commission of Shanghai Municipality (Shanghai Rising-Star Program 18QA1401700), and Shanghai Educational Development Foundation (Chen Guang project 16CG45). W.R. acknowledges support from the State Key Laboratory of Solidification Processing in NWP (SKLSP201703) and the National Natural Science Foundation of China (51672171). **Author contributions:** X.L. performed all measurements, analyzed all experimental data, and prepared the manuscript under the supervision and guidance of J.K. M.B. developed the theoretical model to explain the experimental data. N.Y., K.X., and M.X. grew, cut, and characterized the crystals used in the experiments under the guidance of S.C. Q.Z. and Y.Z. assisted X.L. with measurements and simulations. Z.J. and D.T. characterized the THz response of the samples at low magnetic fields. All authors discussed the results and commented on the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/794/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
Table S1
References (27–33)

7 March 2018; accepted 21 June 2018
10.1126/science.aat5162

ENVIRONMENTAL TOXINS

Diatom ooze—A large marine mercury sink

Sara Zaferani, Marta Pérez-Rodríguez, Harald Biester*

The role of algae for sequestration of atmospheric mercury in the ocean is largely unknown owing to a lack of marine sediment data. We used high-resolution cores from marine Antarctica to estimate Holocene global mercury accumulation in biogenic siliceous sediments (diatom ooze). Diatom ooze exhibits the highest mercury accumulation rates ever reported for the marine environment and provides a large sink of anthropogenic mercury, surpassing existing model estimates by as much as a factor of 7. Anthropogenic pollution of the Southern Ocean began ~150 years ago, and up to 20% of anthropogenic mercury emitted to the atmosphere may have been stored in diatom ooze. These findings reveal the crucial role of diatoms as a fast vector for mercury sequestration and diatom ooze as a large marine mercury sink.

Owing to a scarcity of archive data from the marine environment, the role of primary production and biogenic sediments for natural and anthropogenic mercury (Hg) sequestration in the oceans is poorly understood. In the open ocean, most Hg is received by atmospheric deposition and, to a minor extent, by discharge from rivers, and most atmospheric Hg is thought to be reemitted (1, 2). Available data on global Hg burial in marine sediments are based on model estimates (1, 3). These studies estimate that the amount of Hg buried in open ocean sediments is relatively low (190 to 200 Mg year⁻¹) (2, 4). Several studies indicate that the marine Hg cycle is closely related to biological productivity and that Hg scavenging by organic particles is an important vector for Hg burial in sediments (5–7). However, the contribution of these processes to the accumulation of Hg in marine sediments and the overall marine Hg mass balance in the oceans is not known.

Diatoms are a major group of microalgae. Their remains make up pelagic sediments called diatom ooze, which cover ~11% of the ocean floor (8) and accumulate in largest amounts in the Southern Ocean (9, 10). Until now, no high-resolution (<50 years) Hg record—which allows calculation of global marine Hg sequestration—was available from ocean sediment, and data on Hg in the ocean were limited to water column measurements (7).

Here we show high-resolution Holocene Hg records from diatom ooze sediment cores taken at three basins around Antarctica: Adélie Basin (AB) [Integrated Ocean Drilling Program (IODP) Expedition 318 Site U1357], Prydz Bay (PrB) [Ocean Drilling Program (ODP) Expedition 119 Site 740], and Palmer Basin (PB) (ODP Expedition 178 Site 1098) (Fig. 1). From these cores and modeled marine biogenic silica (BSi) sequestration, we evaluated the role of diatom ooze as

a Holocene Hg sink and estimated the global amount of Hg accumulated in diatom ooze. In addition, we reconstructed the chronology of Hg accumulation and anthropogenic atmospheric Hg pollution during the past 8600 years from annually laminated AB sediments at resolutions of 10, 20, and 200 years.

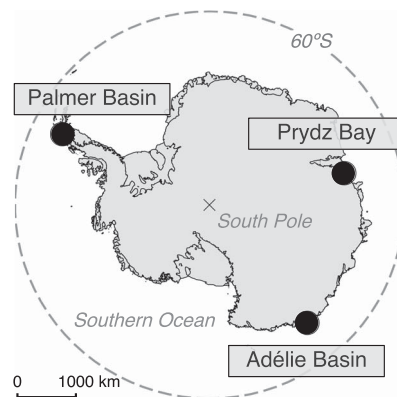


Fig. 1. Map of core sites. Locations of the investigated cores in Antarctica: Prydz Bay core (ODP Expedition 119 Site 740), Palmer Basin core (ODP Expedition 178 Site 1098), and Adélie Basin core (IODP Expedition 318 Site U1357).

Composition of the sediment is similar at each of the three locations (Fig. 2), with high SiO₂ concentrations of 53% (PB), 66% (PrB), and 70% (AB), respectively, identifying the sediments as diatom ooze (11–13). Principal components analyses reveal that Hg generally shows low or no positive covariance with lithogenic elements (e.g., Al, Ti, Zr) (fig. S1). This indicates that binding of atmospheric-derived Hg to microalgae particles is the dominant process of Hg accumulation in these sediments and that terrestrial influences are negligible. The upper core sections contain median Hg concentrations between 55.6 and 70.9 ng g⁻¹ and thereby exceed the median back-

ground (lower core section) concentrations (which are between 32.1 and 36.0 ng g⁻¹) by a factor of 2.2 (Fig. 3). Maximum concentrations reached 73.0 (PB), 84.5 (AB), and 86.6 ng g⁻¹ (PrB). This strong increase in Hg concentrations in the upper core sections likely results from the two- to fivefold increase in global atmospheric Hg loads due to industrial Hg emissions in the past two centuries (14). These concentrations are similar to those found in other marine sediments (15). In the lower core sections, Hg concentrations vary little (Fig. 3) but exhibit dilution effects at higher silica concentrations (see Si-normalized Hg concentrations in fig. S2). Similar Hg concentrations in the three cores taken at locations separated by thousands of kilometers indicate the absence of strong local effects and suggest that Hg concentrations in diatom ooze are related to concentrations of dissolved Hg in seawater, which are controlled by atmospheric Hg deposition and the Henry's law constant, as well as scavenging and mineralization processes of sinking particles.

Owing to the consistency in Hg concentrations, the different sedimentation rates of the three cores control the large differences in Hg accumulation rates (Hg_{AR}). The extremely high sedimentation rates (~2 cm year⁻¹) of AB sediments lead to Hg_{AR} that are the highest ever reported from remote marine areas [median rates of 576 and 1200 μg m⁻² year⁻¹ for the pre-industrial and industrial (after 1850) periods, respectively] (Fig. 3). These values even exceed those reported from estuaries and shelf areas (60 to 700 μg m⁻² year⁻¹) (16), which have hitherto been ranked as the largest marine Hg sinks (17). The lower sedimentation rates of the PrB and PB cores produce lower Hg_{AR} (PrB median: 16.8 and 33.3 μg m⁻² year⁻¹; PB median: 124 and 192 μg m⁻² year⁻¹ for the lower section and upper section of the cores, respectively) than AB cores. However, the PB and PrB rates are still high relative to previously reported values, indicating areas with high accumulation of diatom ooze as large Hg sinks.

To quantify the role of diatom ooze sediments for Hg sequestration in the oceans, we used an estimated total area of pelagic diatom ooze sediments of 23.3 million and 31.1 million km² (10, 18). Owing to the high sedimentation rates of our Antarctic cores, which could lead to large overestimation of global Hg_{AR} in diatom ooze (DOHg_{AR}), we applied two independent approaches for DOHg_{AR} calculation. For calculation one, we used the diatom ooze sedimentation rate suggested for the Southern Ocean (0.75 mm year⁻¹) (19) (dry mass accumulation rate of 0.68 kg m⁻² year⁻¹) in combination with median Hg concentrations (preindustrial and industrial period) derived from our Antarctic cores. For calculation two, we used published estimates of global annual BSi accumulation rate and different diatom preservation rates (30, 60, and 80%) combined with median Si/Hg ratios from our cores for the preindustrial and industrial periods, respectively (Table 1).

Average DOHg_{AR} for the industrial period ranges between 850 and 1166 Mg year⁻¹ (Table 1),

Institut für Geoökologie AG Umweltgeochemie, Technische Universität Braunschweig, Braunschweig, Germany.

*Corresponding author. Email: h.biester@tu-bs.de

which are, respectively, a factor of ~4.5 and 6.1 higher than estimates of existing model approaches (190 Mg year^{-1}) (1, 3) for burial of Hg in deep ocean sediments, indicating that Hg sequestration in biogenic marine sediments appears to have been largely underestimated to date. Preindustrial DOHg_{AR} values range between 497 and 689 Mg year^{-1} , which are a factor of 1.7 lower than those calculated for the industrial period. Although diatom ooze sedimentation and preservation in the Southern Ocean is known to be generally high (>86%) (20), we additionally calculated two conservative approaches using the lowest reported diatom ooze sedimentation rate ($\sim 0.01 \text{ mm year}^{-1}$) (21, 22) and the lowest published BSi preservation rate (3%) (23). These calculations reveal DOHg_{AR} of 18 and 38 Mg year^{-1} , which still account for 9 and 20% of the model estimate (190 Mg year^{-1}), respectively (1, 3). These findings emphasize the importance of Hg sequestration in diatom ooze in particular and the role of marine primary production in general. However, the amount of data used for these calculations is still small and may exhibit large uncertainties.

Streets *et al.* (24) estimated that ~1130 Gg of Hg were released from anthropogenic sources to the environment between 1850 and 2010, and 336 Gg of this amount were emitted directly to the atmosphere. On the basis of the minimum and maximum DOHg_{AR} obtained from calculations one and two (397 to $1322 \text{ Mg year}^{-1}$), the total amount of Hg accumulated in diatom ooze in the past 150 years is 22 to 84 Gg, or about 6.5 to 25% of all Hg emitted to the atmosphere and 2.0 to 7.4% of all Hg released to the environment through anthropogenic emissions in the past 150 years. The remote location and large distance of most diatom ooze sediments from coastlines indicate atmospheric fluxes as the predominant Hg source. Under this assumption, ~20% of the anthropogenic Hg emitted to the atmosphere since 1850 may have been buried in diatom ooze alone. This amount might have been even higher considering that DOHg_{AR} in areas such as AB largely exceeds the rates used in our calculation. The unequal distribution of atmospheric Hg concentrations during the industrial period as well as that of diatom ooze sediments between the hemispheres suggest that Hg sequestration by diatom ooze might be more important in the Southern Hemisphere than in the Northern Hemisphere, which might reduce the amount of anthropogenic Hg accumulated in diatom ooze. It is further unknown to what extent the 1 to 5.5 Gg year^{-1} of global river discharge of Hg, mainly released to the Northern Ocean (25), contributes to Hg burial in diatom ooze. However, recent studies indicate that most of this Hg is buried in estuaries and on the continental shelf (17).

The high Hg accumulation rates in diatom ooze, which largely surpass the reported Hg deposition rates to the oceans (1), might be partly explained by increased Hg fluxes from the atmosphere into the seawater. Similar to CO_2 , such increased atmospheric Hg fluxes

could be caused by a permanent shift of the dissolution equilibrium (based on Henry's law constant) toward the dissolved phase if the dissolved phase is permanently removed through Hg scavenging by a large amount of algae during blooms. We assume that scavenging of water-phase Hg by sinking diatom organic matter, rather than active uptake [see (26) for Hg uptake by algae], is the dominant process of Hg binding in diatom ooze. Furthermore, dissolved Hg in the water column will provide a large Hg pool for scavenging by sinking particles, which could explain the high observed Hg sedimentation rates (see supplementary materials for further discussion). Previous mass-balance models estimated

that 96% ($\sim 1300 \text{ Mg year}^{-1}$) of the atmospheric Hg flux to the ocean ($\sim 1350 \text{ Mg year}^{-1}$) is re-emitted (1). Even the conservative Hg accumulation rates in diatom ooze suggest that these high re-emission fluxes are likely overestimated if used on a global scale. A recent model approach on Hg mass-balance changes during eutrophication in the Baltic Sea estimated a 30 to 40% decrease in Hg evasion during diatom blooms (27), which was attributed to Hg sedimentation by algae organic matter. This is in line with findings of Hg concentrations in settling organic particulate matter in the Atlantic Ocean, which suggest that Hg partitioning to particles is up to 1000 times greater than suggested by models

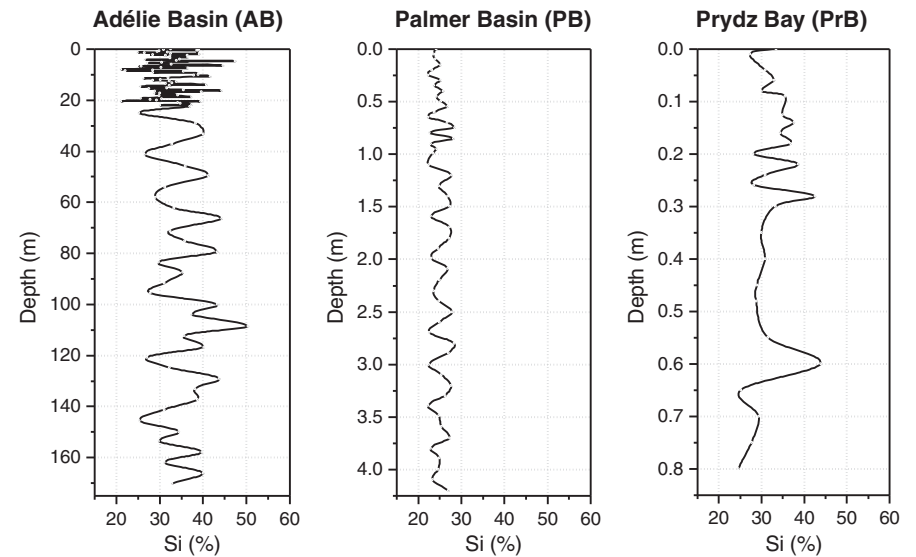


Fig. 2. Silicon concentration in diatom ooze sediments. Downcore records of silicon (Si) concentrations in Adélie Basin (AB), Palmer Basin (PB), and Prydz Bay (PrB) sediments (note the differences in core depth).

Table 1. Global annual mercury accumulation and total accumulated mercury in diatom ooze sediments (DOHg_{AR}). Values for the industrial (after 1850) and preindustrial (before 1850) periods were determined by two approaches. For calculation one (Calc. 1) we used a diatom ooze area of 23.3 million and 31.1 million km^2 (10, 18) and a sedimentation rate of $0.75 \text{ mm year}^{-1}$, as suggested for Antarctic diatom ooze (19), and average Hg concentrations for the preindustrial and the industrial period derived from AB, PrB, and PB cores. Calculation two (Calc. 2) was based on the annual accumulation of biogenic silica (BSi) in the world's oceans (23), assuming BSi preservation of 30, 60, and 80% and average Si/Hg ratios for the preindustrial and industrial periods derived from AB, PrB, and PB cores.				
Calculation	DOHg_{AR} preindustrial period (Mg year^{-1})	DOHg_{AR} industrial period (Mg year^{-1})	Enrichment factor	DO-stored pollution Hg since ~1850 (Gg)
Calc. 1 (23.3 million, 31.1 million km^2)	574, 765	992, 1322	1.7	63, 84
Calc. 2 (30, 60, 80%)	230, 460, 613	397, 758, 1010	1.6	22, 45, 60
Average Calc. 1 and 2 (range based on Calc. 2)	497, 613, 689	850, 1040, 1166	1.7	53, 64, 72

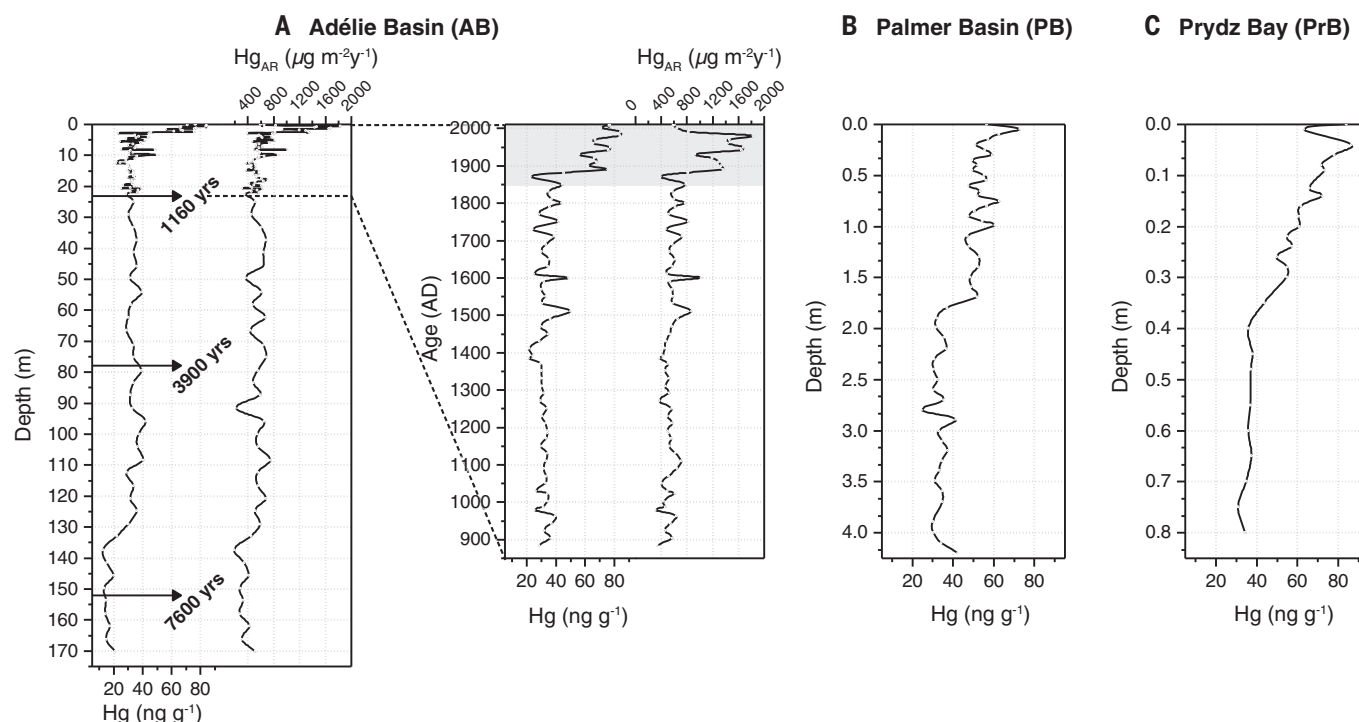


Fig. 3. Mercury concentration and accumulation rate in diatom ooze sediments. (A to C) Downcore records of mercury (Hg) concentrations in Adélie Basin (AB), Palmer Basin (PB), and Prydz Bay (PrB), and mercury accumulation rates (Hg_{AR}) in AB sediments. The shaded area in (A) denotes the industrial period.

An average sedimentation rate of 2 cm year⁻¹ was used for AB chronology reconstruction (13). Median background Hg concentrations equal 32.1, 34.5, and 36.0 ng g⁻¹ with median absolute deviations of 2.74, 2.21, and 1.20 for AB, PB, and PrB, respectively (note the differences in core depth).

resulting in decreasing Hg availability for evasion (28), emphasizing the important role of the marine biological pump for the global Hg cycle.

The undisturbed 170-m-long annually laminated sediments of AB offer an ultrahigh resolution (~2 cm year⁻¹) record of Hg accumulation in the oceans, spanning large parts of the Holocene (Fig. 3) (13). Hg concentrations in AB sediments show a median of 32.1 ng g⁻¹ between ~8600 years ago and ~1850 AD, and Hg accumulation rates largely follow the Hg concentration record, with a median value of 556 μg m⁻² year⁻¹. Since 1850, Hg concentrations and accumulation rates have increased by a factor of ~2.2 to median values of 70.9 ng g⁻¹ and 1296 μg m⁻² year⁻¹, with local peaks around 1890, 1950, and 1990, and highest values of 84.5 ng g⁻¹ and 1830 μg m⁻² year⁻¹ in ~1990. The increase in Hg accumulation was coincident with the onset of the industrial period and intense coal burning, which has been the most important anthropogenic source of Hg emissions to the atmosphere and global Hg dispersion. The local peaks in Hg accumulation during 1900–1920 and 1950–1970 (Fig. 3) correspond to emission peaks calculated on the basis of global Hg production and consumption rates (24, 29). An indication of a global influence of Hg emissions from gold and silver mining in the Americas during 1850 and 1910, as estimated in recent models (29), which nearly reach or even surpass levels of those found at times of maximum Hg emissions during the 20th century,

was not found in our cores. Moreover, no anthropogenic influence before ~1850 that could be linked to Hg emissions from colonial gold mining in South America, as suggested by model estimates, could be detected (16). The amount of Hg emitted to the atmosphere by colonial gold mining was probably too low to be detectable in remote Antarctica. Similarly, the increase in global atmospheric Hg loads, as revealed by our Antarctic marine sediment cores, attributed to all-time anthropogenic activities, is only about a factor of 2.2, which is slightly lower than in most Hg records from terrestrial archives (14, 30–32) but distinctly lower than current model predictions. This might be due to the fact that most anthropogenic Hg has been emitted in the Northern Hemisphere so that a larger Hg portion is sequestered there. However, the Southern Ocean is one of the most productive areas on Earth and, because of the link between primary production and Hg sequestration, is likely a large sink for global atmospheric Hg.

REFERENCES AND NOTES

1. R. P. Mason, G.-R. Sheu, *Global Biogeochem. Cycles* **16**, 1093 (2002).
2. H. M. Amos, D. J. Jacob, D. G. Streets, E. M. Sunderland, *Global Biogeochem. Cycles* **27**, 410–421 (2013).
3. E. M. Sunderland, R. P. Mason, *Global Biogeochem. Cycles* **21**, GB4022 (2007).
4. R. P. Mason, W. Fitzgerald, in *Global and Regional Mercury Cycles: Sources, Fluxes and Mass Balances*, W. Baeyens, R. Ebinghaus, O. Vasiliev, Eds. (NATO Science Partnership Subseries: 2, Springer, 1996), pp. 249–272.

5. S. Strode, L. Jaeglé, S. Emerson, *Global Biogeochem. Cycles* **24**, GB4014 (2010).
6. Y. Zhang, L. Jaeglé, L. Thompson, D. G. Streets, *Global Biogeochem. Cycles* **28**, 1251–1261 (2014).
7. C. H. Lamborg et al., *Nature* **512**, 65–68 (2014).
8. W. H. Berger, in *Chemical Oceanography* (Academic Press, 1976), vol. 5, pp. 265–372.
9. B. Diekmann, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **54**, 2350–2366 (2007).
10. A. Dutkiewicz, R. D. Müller, S. O. O'Callaghan, H. Jónasson, *Geology* **43**, 795–798 (2015).
11. Shipboard Scientific Party, "Palmer Deep (Sites 1098 and 1099)," *Proceedings of the Ocean Drilling Program, Initial Reports*, vol. 178 (ODP, 1999); www.odp.tamu.edu/publications/178_IR/VOLUME/CHAPTERS/CHAP_07.PDF.
12. Shipboard Scientific Party, "Site 740," *Proceedings of the Ocean Drilling Program, Initial Reports*, vol. 119 (ODP, 1989); www.odp.tamu.edu/publications/119_IR/VOLUME/CHAPTERS/ir119_09.pdf.
13. C. Escutia, H. Brinkhuis, A. Klaus, Expedition 318 Scientists, "Site U1357," *Proceedings of the Integrated Ocean Drilling Program*, vol. 318 (IODP, 2011); http://publications.iodp.org/proceedings/318/EXP_REPT/CHAPTERS/318_105.PDF.
14. H. Biester, R. Bindler, A. Martinez-Cortizas, D. R. Engstrom, *Environ. Sci. Technol.* **41**, 4851–4860 (2007).
15. J. Gleason et al., *Geochim. Cosmochim. Acta* **197**, 245–262 (2017).
16. H. M. Amos et al., *Environ. Sci. Technol.* **49**, 4036–4047 (2015).
17. H. M. Amos et al., *Environ. Sci. Technol.* **48**, 9514–9522 (2014).
18. H. U. Sverdrup, M. W. Johnson, R. H. Fleming, *The Oceans: Their Physics, Chemistry, and General Biology* (Prentice-Hall, 1942).
19. D. K. Fütterer, in *Marine Geochemistry*, H. D. Schulz, M. Zabel, Eds. (Springer, ed. 2, 2006).
20. D. J. DeMaster, O. Ragueneau, C. A. Nittrouer, *J. Geophys. Res. Oceans* **101**, 18501–18518 (1996).
21. W. H. Berger, in *The Geology of Continental Margins* (Springer, 1974), pp. 213–241.

22. M. G. Gross, *Oceanography: A View of the Earth* (Prentice-Hall, 1987).
23. P. Tréguer *et al.*, *Science* **268**, 375–379 (1995).
24. D. G. Streets *et al.*, *Environ. Sci. Technol.* **51**, 5969–5977 (2017).
25. J. E. Sonke, L.-E. Heimbürger, *Nat. Geosci.* **5**, 447–448 (2012).
26. S. Le Faucheur, P. G. Campbell, C. Fortin, V. I. Slaveykova, *Environ. Toxicol. Chem.* **33**, 1211–1224 (2014).
27. A. L. Soerensen *et al.*, *Environ. Sci. Technol.* **50**, 11787–11796 (2016).
28. C. H. Lamborg, C. R. Hammerschmidt, K. L. Bowman, *Philos. Trans. R. Soc. A* **374**, 20150297 (2016).
29. H. M. Horowitz, D. J. Jacob, H. M. Amos, D. G. Streets, E. M. Sunderland, *Environ. Sci. Technol.* **48**, 10242–10250 (2014).
30. D. R. Engstrom *et al.*, *Environ. Sci. Technol.* **48**, 6533–6543 (2014).
31. S. A. Beal, B. P. Jackson, M. A. Kelly, J. S. Stroup, J. D. Landis, *Environ. Sci. Technol.* **47**, 12715–12720 (2013).
32. S. Eyrikh *et al.*, *Environ. Sci. Technol.* **51**, 11597–11606 (2017).

ACKNOWLEDGMENTS

We thank P. Rumford (Golf Coast Repository), W. Hale, and H. Kuhlmann (Bremen Core Repository) for providing samples and A. Caele and P. Schmidt for technical assistance. **Funding:** This work was funded by Technische Universität Braunschweig.

Author contributions: S.Z. carried out the laboratory analyses, and M.P.-R. performed the statistical analyses. H.B. conceived of the research. All authors contributed equally to data interpretation

and manuscript writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/797/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S14
References (33–54)

9 February 2018; accepted 7 July 2018
Published online 26 July 2018
10.1126/science.aat2735

VIROLOGY

Serotype-specific immunity explains the incidence of diseases caused by human enteroviruses

Margarita Pons-Salort* and Nicholas C. Grassly

Human enteroviruses are a major cause of neurological and other diseases. More than 100 serotypes are known that exhibit unexplained complex patterns of incidence, from regular cycles to more irregular patterns, and new emergences. Using 15 years of surveillance data from Japan (2000–2014) and a stochastic transmission model with accurate demography, we show that acquired serotype-specific immunity can explain the diverse patterns of 18 of the 20 most common serotypes (including Coxsackieviruses, Echoviruses, and Enterovirus-A71). The remaining two serotypes required a change in viral characteristics, including an increase in pathogenicity for Coxsackievirus-A6, which is consistent with its recent global rise in incidence. On the basis of our findings, we are able to predict outbreaks 2 years ahead of time (2015–2016). These results have implications for the impact of vaccines under development.

More than 100 enterovirus serotypes infect humans and can cause a wide range of diseases, including meningitis, encephalitis, paralysis, myocarditis, haemorrhagic conjunctivitis, upper respiratory infections, herpangina, and hand-foot-and-mouth disease (HFMD) (1). This group contains the Coxsackieviruses A (CV-A) and B (CV-B), the Echoviruses (E), and other serotypes simply called Enterovirus (EV), as well as the three poliovirus serotypes, which are the target of a global eradication program. The epidemiology of enteroviruses is characterized by frequent asymptomatic infections and a peak of disease in summer. Despite the well-defined annual seasonality of enterovirus-associated diseases, individual serotypes show different long-term patterns of circulation, from regular annual or multiannual cycles to more irregular patterns. For example, enterovirus A71 (EV-A71) has a 3-year cycle in Malaysia (2) and Japan but a 1-year cycle in China (3). Some serotypes also sporadically emerge as substantial causes of diseases. For example, CV-A6 has emerged during recent years as the main cause of HFMD worldwide (4), and EV-D68 emerged in 2014 in the United States as an important cause of severe respiratory illnesses (5). The increased incidence of HFMD in Southeast Asia during the past two decades [with >1 million cases reported in China annually (3)], the large proportion of viral meningitis attributed to nonpolio enteroviruses (>75% (6)) and the sporadic emergence of new serotypes attest to their clinical and public health importance.

A fundamental challenge for research on enteroviruses and other endemic and antigenically diverse pathogens (such as rotavirus, influenza,

pneumococcus, and dengue), is to understand what causes the changes in incidence over time of different serotypes. This has direct implications for the prediction of future outbreaks of disease and the impact of serotype-specific vaccines, such as licensed monovalent EV-A71 vaccines (7) and other single- and multivalent vaccines under development (8). There is currently no accepted mechanistic explanation for the com-

plex long-term patterns of circulation of enterovirus serotypes, including why different serotypes predominate in different years. Homotypic (serotype-specific) population immunity, heterotypic (cross-serotype) immunity, and virus evolution (including antigenic drift and recombination) could all contribute.

In most countries, enterovirus surveillance is passive and thus mainly captures the serotypes isolated from severe cases suspected of having an enterovirus infection (such as neonatal infections and clusters of aseptic meningitis, encephalitis, or paralysis). Serotypes isolated from mild diseases are rarely reported. Furthermore, variation in clinical testing and participation of laboratories results in irregular reporting rates over time. In contrast with most countries, Japan has a national disease sentinel-based surveillance system that isolates and identifies enteroviruses from patients attending sentinel sites for four target diseases representative of enterovirus infections: aseptic meningitis (mostly caused by Echoviruses, CV-Bs, and CV-A9), HFMD (mostly caused by EV-A71 and CV-A6, -10, and -16), herpangina (mostly caused by CV-As), and acute haemorrhagic conjunctivitis (mostly caused by a variant of CV-A24 and EV-D70) (9). Enterovirus surveillance in Japan therefore captures the circulation of a broad range of serotypes and does it in a consistent and systematic way. We used time-series analysis and mathematical modeling of the high-quality surveillance data from Japan collected between 2000 and 2016 to identify mechanisms that can explain the transmission dynamics of enterovirus serotypes.

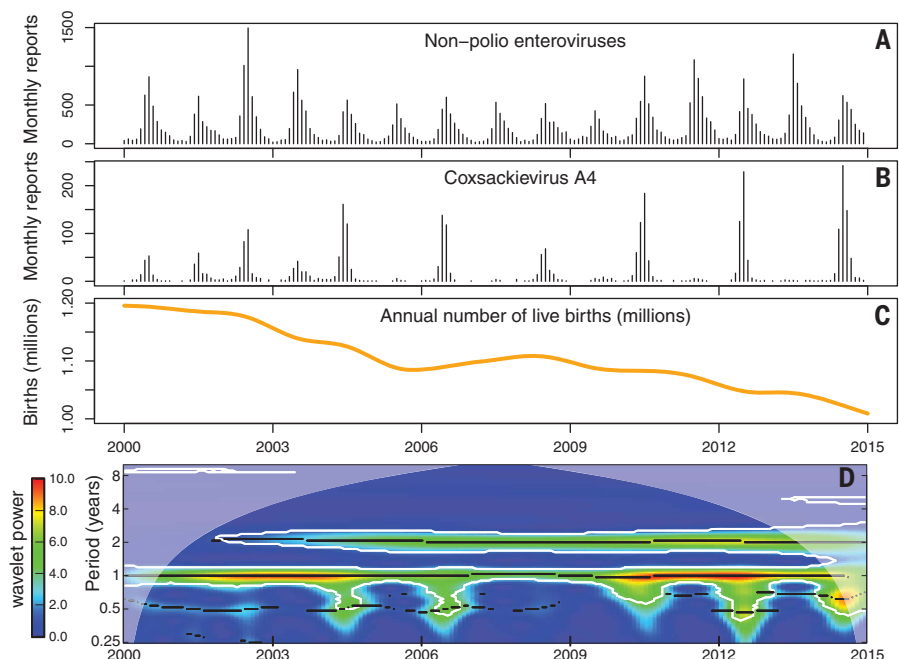


Fig. 1. Nonpolio enterovirus incidence and births in Japan (2000–2014). (A and B) Monthly number of reported enterovirus isolations from January 2000 to December 2014 for (A) nonpolio enteroviruses and (B) CV-A4. (C) Smoothed annual number of live births. (D) Average wavelet power of the square-root-transformed time series for CV-A4 showing the emergence of a biennial pattern of incidence.

Department of Infectious Disease Epidemiology, St Mary's Campus, Imperial College London, London W2 1PG, UK.
*Corresponding author. Email: m.pons-salort@imperial.ac.uk

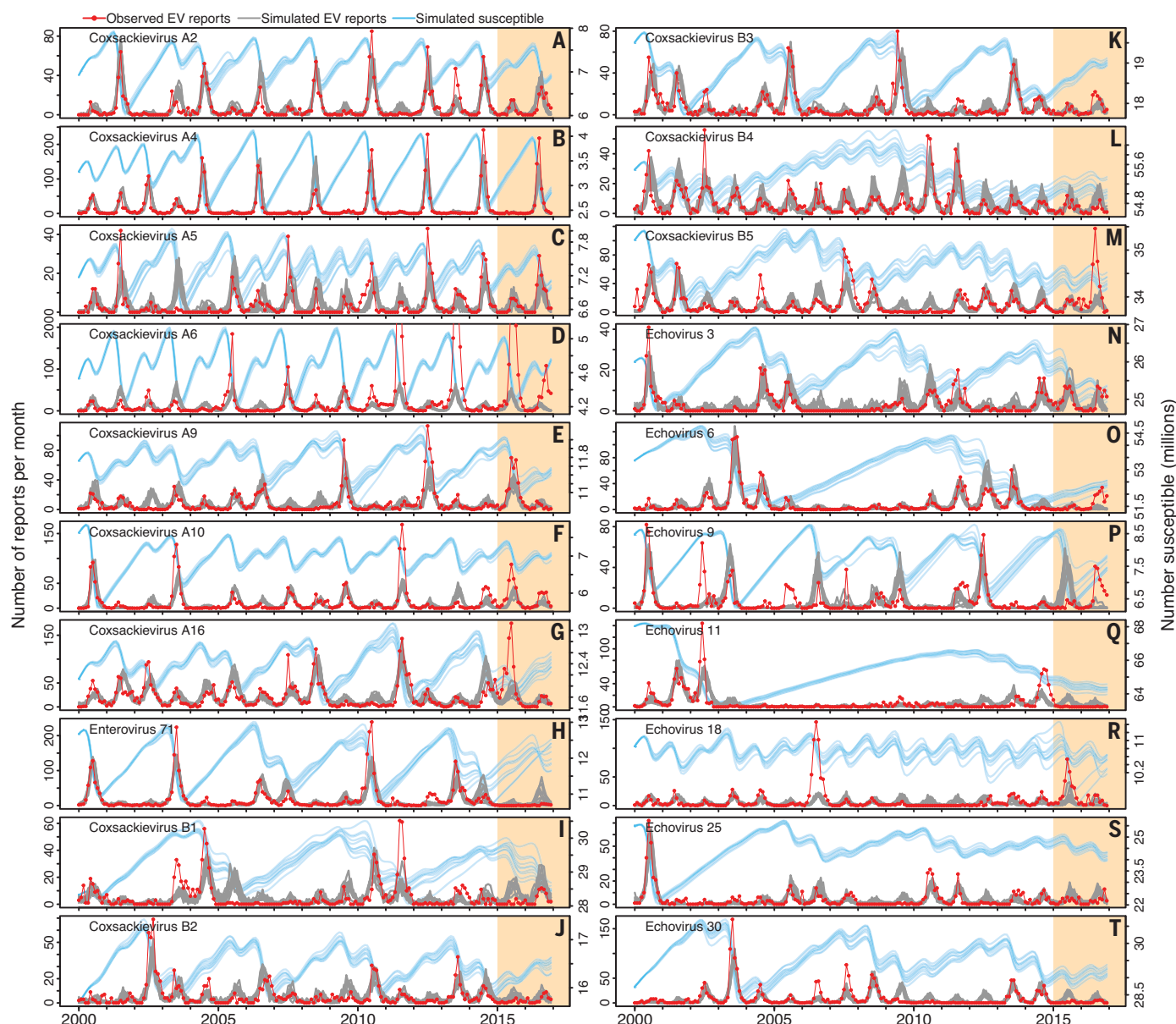


Fig. 2. Model fit to data. (A to T) The reported number of monthly isolations (red line) and 10 out of 100 stochastic simulations of the maximum likelihood model (gray lines) with the closest match to the data are shown. The estimated number of susceptible individuals (blue lines) is shown in the right y axis. Maximum likelihood parameter estimation and selection of simulations was based on data for

January 2000 to December 2014 (white background), and projections continued out-of-sample for January 2015 to December 2016 (orange background). For CV-A6 and E18, parameter estimation was based on data through to December 2010 and December 2005, respectively, because of the poor fit of the model to the entire period. Parameter values are available in table S2.

Nonpolio enteroviruses show strong seasonality in Japan, with aggregate case numbers showing a peak every summer between January 2000 and December 2014 (Fig. 1A). The incidence of serotypes with >500 isolations during that period (total numbers are provided in table S1) exhibited diverse patterns (Fig. 2, red line, and figs. S4 to S6), including serotypes with regular cycles of 1 year (such as CV-B4), 2 years (such as CV-A4), 3 years (such as EV-A71), and 4 years (such as CV-B3), and less regular patterns (such as E25). A wavelet analysis (supplementary text S1.3) identified these periodicities and their

changes over time—that is, regime shifts (figs. S7 to S9). In particular, we observed that CV-A4 switched from a 1-year cycle between 2000 and 2004 to a 2-year cycle from 2004 onward (Fig. 1, B and D). We hypothesized that if CV-A4 behaved as a highly immunizing infection, this temporal pattern could be explained by the general decrease in births that Japan has experienced during the past two decades (Fig. 1C). In this case, only births would replenish the susceptible pool and consequently determine the chances that an outbreak occurs (10). Similar but perhaps less obvious signatures in the pattern of other serotypes also

indicated this mechanism (such as CV-A9 and CV-A16).

To test this hypothesis, we constructed a simple stochastic transmission model that assumed acquired and lifelong homotypic immunity and that included a seasonal transmission rate, asymptomatic infections, underreporting, and realistic birth and death rates (supplementary text S1.4 and S1.5). Using a maximum-likelihood inference framework based on particle filtering (11), we fitted the model to the number of detections reported each month between January 2000 and December 2014 for each of the 20 most prevalent

serotypes with more than one peak of activity. These included seven Coxsackieviruses A, five Coxsackieviruses B, seven Echoviruses, and EV-A71. We then assessed whether the maximum likelihood model was able to predict the 15-year incidence pattern of each serotype when simulating forward from the initial conditions, only varying birth and death rates as observed in Japan. Because demographic stochasticity results in variability between simulated realizations with the same parameter values, we simulated 100 stochastic realizations with the maximum-likelihood estimates (MLEs) of the parameters and show here the 10 that most closely match the observed epidemic trajectories (supplementary text S2.3). Model failure to reproduce the observed dynamics of specific serotypes guided us to generate hypotheses of extra mechanisms that could explain the observed incidence.

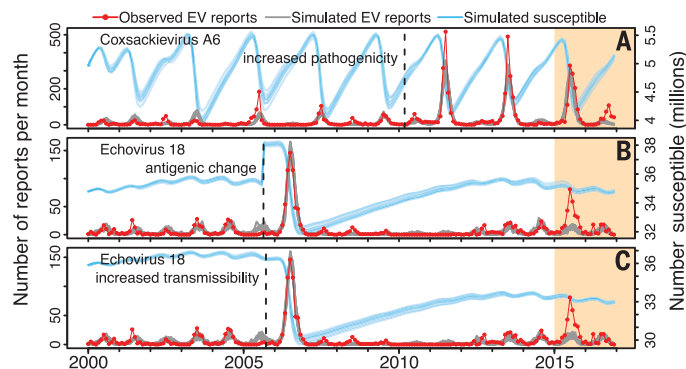
The model fitted the 15 years of incidence data remarkably well for 18 of the 20 serotypes examined—all except CV-A6 and E18 (Fig. 2). These included serotypes with periods of 2 years (such as CV-A2 and CV-A4), 3 years (such as CV-A9 and EV-A71), and even longer (such as E3). The model was also able to capture more irregular dynamics, such as those exhibited by E6 and E25, that showed a few years of activity followed by several consecutive years with barely any cases. Underestimation of the largest peaks of cases for a few serotypes could be due to undocumented enhanced sample collection and surveillance efforts during large outbreaks (such as 2010 for EV-A71). The model also predicted a few small peaks that were not observed in the data (such as 2003 for CV-A5, 2013 for CV-A10, 2005 for CV-B1, 2009–2010 for CV-B5, and 2009–2012 for E30). Possibly, these discrepancies reflect gaps in surveillance resulting from the distribution of sentinel sites and spatial clustering of cases or, alternatively, could correspond to local viral extinctions that are not captured by the model. The number of estimated susceptible individuals (Fig. 2, blue line) shows how births progressively increase population susceptibility after each peak of cases, until a threshold is reached and a new outbreak occurs. With decreasing births, it takes longer for this threshold to be reached, increasing the chances of transition to longer cycles (such as CV-A4 from 2004 onwards, CV-A9 from 2006, and CV-A16 from 2008), which sometimes result in a bigger outbreak.

Although the model did not explain the dynamics of CV-A6 and E18 over the entire 15-year period, it fit the data well until the dynamics suddenly changed (2011 and 2006, respectively), suggesting that mechanisms other than homotypic immunity contributed to these dynamics (Fig. 2, D and R). Virus evolution in antigenic or other regions could perhaps explain these patterns. For example, it has been suggested that diversity in capsid structures of some Echoviruses could have generated different antigenic variants (12), or that recombination in the non-structural proteins could provide a fitness or immunological advantage (13). To test whether a sudden change in transmissibility, pathogenicity,

Fig. 3. Model

extensions. Model fit to data for CV-A6 and E18 by using the best-fitting model extensions. (A) A fivefold increase in pathogenicity from around 2010 for CV-A6. (B) A change in antigenicity from around 2006 for E18. (C) A 9% increase in transmissibility from around 2006 for E18.

Dashed vertical lines indicate the estimated time when the changes occurred. Details are provided in Fig. 2. Parameter values are available in table S4.



or antigenicity alone could explain the pattern of CV-A6 and E18, we fitted three model extensions accounting for those changes (supplementary text S1.6). For CV-A6, we estimated that a fivefold increase in pathogenicity in 2010 provided a better fit to the data than models with a change in transmissibility or antigenicity (Fig. 3A). This is consistent with the recent global emergence of this serotype as the main cause of HFMD with atypical symptoms (4) and a change around 2009 in the disease that it caused in Japan (mostly herpangina before, but HFMD since then) (14). For E18, however, either an antigenic change or an increase in transmissibility at the end of 2005 (but not a change in pathogenicity) provided a better fit to the data (Fig. 3, B and C).

Parameter estimates revealed interesting differences in transmissibility among the serotypes (fig. S55 and table S3). Those associated with HFMD and herpangina (most CV-As and EV-A71) had generally higher estimates of the average basic reproduction number compared with those associated with aseptic meningitis (CV-A9, CV-Bs, and Echoviruses). This may reflect higher transmission efficiency of serotypes mainly affecting children and/or differences in transmission routes (fecal-oral versus respiratory). The estimated reporting probability per infection was between 10^{-5} and 10^{-4} , which is about two orders of magnitude smaller than the probability of paralysis after poliovirus infection (10^{-3} to 10^{-2}) (15). This may reflect differences in pathogenicity and coverage of the sentinel surveillance system.

Our results have implications for outbreak preparedness. By extending model simulations 2 years beyond the data used for parameter estimation (2015–2016), the magnitude of the epidemics was predicted with reasonable accuracy for most of the serotypes (Figs. 2 and 3, orange backgrounds, and fig. S56). The accuracy of these forecasts was similar to that obtained when simulating forward from 2015 only, by using the reconstructed state space variables obtained from the particle-filter (fig. S57). Inaccurate forecasts for a few serotypes could be due to changes in viral properties during the prediction period, local viral extinctions, or inaccurate projections of future number of births. Nevertheless, our simple model could help anticipate years with a large

number of cases of a given serotype. These forecasts could be improved by incorporating data on serotype-specific population immunity from seroprevalence surveys.

Similarly, our model could also be used to predict the impact of vaccines under development. For example, we simulated the effect of vaccinating newborns against EV-A71 and CV-A16 from 2008 onward with a 100% effective vaccine at different coverage levels (fig. S58). For both serotypes, a coverage of 80% resulted in almost no cases from 2 years after the introduction of the vaccine, but lower coverage led to continued incidence, with altered dynamics and occasionally large outbreaks comparable with those observed prevaccination.

Our analysis has some limitations. First, we assumed that acquired homotypic immunity against infection is complete and lifelong. However, it is likely to be incomplete and wane over time for nonpolio enteroviruses, as seen for intestinal immunity to polioviruses (16, 17). Extending our model to include waning immunity (supplementary text S1.7) resulted in a similarly good fit to the data (figs. S62 and S63) and an estimated mean duration of protective immunity that ranged from 8 years to lifelong (table S5). This relatively long-lasting immunity meant that the dynamics were still determined by births, but estimates of the reproduction number were generally lower because infection could be sustained at lower transmission efficiency. Second, we ignored spatial, social, and demographic structure of the transmission dynamics. Although these simplifications may have affected some of our parameter estimates, they do not affect our main result of the importance of homotypic immunity. Third, although stochastic extinction of a serotype and subsequent reintroduction can occur in our model, it is inherently difficult to predict specific periods when infection may be absent. Last, smaller changes in pathogenicity, transmissibility, or antigenicity may have occurred for serotypes other than CV-A6 and E18 but not led to a strong signature in the incidence dynamics.

Although there is evidence that different enterovirus serotypes share some T cell and B cell epitopes (18, 19), cross-reactive antibodies

are not strongly neutralizing, and their potential epidemiological implications remain elusive. Our model can explain the temporal dynamics of all the serotypes we examined without the complexity of adding cross-serotype interactions, thus providing evidence that homotypic immunity is the most important driver of the transmission dynamics of enterovirus serotypes. These findings support current development efforts of serotype-specific mono- and multivalent vaccines because they point against serotype replacement induced by vaccination. However, two of the serotypes we examined showed dynamics that are consistent with a change in viral properties (pathogenicity, antigenicity, or transmissibility) and emergence as an important cause of disease. A better understanding of this process would support the appropriate choice of serotypes in multivalent vaccines that target diseases such as HFMD.

REFERENCES AND NOTES

1. M. A. Pallansch, M. S. Oberste, J. L. Whitton, in *Fields Virology*, D. M. Knipe, P. Howley, Eds. (Lippincott Williams & Wilkins, 2013), vol. 2, chap. 17.
2. Y. Podin *et al.*, *BMC Public Health* **6**, 180 (2006).
3. W. Xing *et al.*, *Lancet Infect. Dis.* **14**, 308–318 (2014).
4. L. Bian *et al.*, *Expert Rev. Anti Infect. Ther.* **13**, 1061–1071 (2015).
5. C. M. Midgley *et al.*, *Morb. Mortal. Wkly. Rep.* **63**, 798–799 (2014).
6. N. G. Martin *et al.*, *Lancet Infect. Dis.* **16**, 1279–1287 (2016).
7. World Health Organization, “China produces world’s first vaccine against virus that causes Hand, Foot and Mouth Disease” (2015); www.wpro.who.int/china/mediacentre/releases/2015/20151208/en.
8. Q. Mao, Y. Wang, L. Bian, M. Xu, Z. Liang, *Emerg. Microbes Infect.* **5**, e75 (2016).
9. National Institute of Infectious Diseases (NIID) of Japan, Enterovirus surveillance in Japan, 1982–1999, *IASR* **21**, 212–213 (2000); <http://idsc.nih.gov/ja/iasr/21/248/tpc248.html>.
10. R. M. Anderson, B. T. Grenfell, R. M. May, *J. Hyg. (Lond.)* **93**, 587–608 (1984).
11. E. L. Ionides, D. Nguyen, Y. Atchadé, S. Stoev, A. A. King, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 719–724 (2015).
12. M. S. Oberste, W. A. Nix, D. R. Kilpatrick, M. R. Flemister, M. A. Pallansch, *Virus Res.* **91**, 241–248 (2003).
13. E. C. McWilliam Leitch *et al.*, *J. Virol.* **84**, 9292–9300 (2010).
14. National Institute of Infectious Diseases (NIID) of Japan, Hand, foot, and mouth disease and herpangina, 2007 to September 2017 (week 38), Japan, *IASR* **38** (2017); www.niid.go.jp/niid/en/iasr-vol38-e/865-iasr/7617-452te.html.
15. K. Penttinen, R. Patiala, *Ann. Med. Exp. Biol. Fenn.* **39**, 195–202 (1961).
16. J. John *et al.*, *Lancet* **384**, 1505–1512 (2014).
17. J. John *et al.*, *J. Infect. Dis.* **215**, 529–536 (2017).
18. T. Hovi, M. Roivainen, *J. Clin. Microbiol.* **31**, 1083–1087 (1993).
19. K. Katrak, B. P. Mahon, P. D. Minor, K. H. Mills, *J. Gen. Virol.* **72**, 1093–1098 (1991).

ACKNOWLEDGMENTS

We thank H. Shimizu for providing insightful comments on the enterovirus surveillance system in Japan; B. T. Grenfell, S. Takahashi, and E. P. K. Parker for providing valuable comments to the manuscript; and N. Imai for her help with the Japanese language. **Funding:** M.P.-S. is funded by the Wellcome Trust (grant 106073/Z/14/Z). **Author contributions:** M.P.-S. designed research, performed the analyses, and wrote the paper. N.C.G. supervised the work and contributed to the writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All the data used in the manuscript are publicly available in the website of the National Institute of Infectious Diseases (NIID) of Japan: www.niid.go.jp.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/800/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S63
Tables S1 to S5
References (20–31)
26 March 2018; accepted 3 July 2018
10.1126/science.aat6777

OCEAN OXYGEN

Large-scale ocean deoxygenation during the Paleocene-Eocene Thermal Maximum

WeiQi Yao^{1*}, Adina Paytan², Ulrich G. Wortmann¹

The consequences of global warming for fisheries are not well understood, but the geological record demonstrates that carbon cycle perturbations are frequently associated with ocean deoxygenation. Of particular interest is the Paleocene-Eocene Thermal Maximum (PETM), where the carbon dioxide input into the atmosphere was similar to the IPCC RCP8.5 emission scenario. Here we present sulfur-isotope data that record a positive 1 per mil excursion during the PETM. Modeling suggests that large parts of the ocean must have become sulfidic. The toxicity of hydrogen sulfide will render two of the largest and least explored ecosystems on Earth, the mesopelagic and bathypelagic zones, uninhabitable by multicellular organisms. This will affect many marine species whose ecozones stretch into the deep ocean.

The geological record contains many examples in which the Earth system was out of equilibrium and large parts of the ocean were inhospitable to life. However, only few of these events can provide insight into the effects of modern fossil fuel burning. This is because either the boundary conditions are substantially different (e.g., the plate-tectonic configuration) or the rate of change is not comparable. The short-lived Paleocene-Eocene Thermal Maximum (PETM) event [~ 55 million years ago] (1) is a notable exception. Current data suggest that in the PETM, the atmosphere had to accommodate about 2500 to 4500 Gt of carbon released within 4000 years (2). This is an increase of the same order of magnitude as the IPCC RCP8.5 emission scenario, which projects a cumulative anthropogenic CO₂ release of 2000 GtC by 2100 (3). Although the carbon dioxide

release rate during the PETM was about a factor of 10 slower, it is our best analog for studying nonlinear feedbacks and consequences of the anthropogenic carbon cycle perturbation.

The geochemical cycles of carbon and sulfur are linked through microbial sulfate reduction (MSR), where the electron transfer from sulfate to sulfide provides the energy to respire organic matter (OM) back to CO₂. Combined, these cycles constitute the dominant control on atmospheric oxygen (4, 5). Owing to their drastically different residence times (0.1 versus 10 million years) (4), they are rarely considered together. Our data suggest, however, that MSR can alter the redox state of the marine sulfur reservoir on time scales that are comparable to that of the carbon cycle. This has three important implications: (i) Unlike oxic respiration, MSR also produces H₂S, which is toxic to most life forms even at low concentra-

tions; (ii) if we accept the premise that the PETM is a model for the present-day oceans, the time scales of the observed changes in the redox state of marine sulfur suggest that similar processes could affect the oceans in the near future; (iii) the development of oxygen-free waters creates a sizable but intermittent reservoir in the global sulfur cycle, with fluxes exceeding traditional weathering and burial flux estimates.

The mass and sulfur-isotope ratio of seawater sulfate are controlled by sulfur delivered to the oceans via weathering and volcanic degassing and sulfur removed from the oceans via burial of sulfur-bearing minerals. In marine sediments, sulfur is present in its oxidized form as sulfate-bearing evaporite salts and in its reduced state as metal sulfides. Sulfate-bearing evaporites precipitate with little or no fractionation relative to seawater, and thus have a limited effect on the S-isotope ratio of seawater sulfate (5, 6). By contrast, MSR has a preference for ³²S, and thus exerts a substantial influence on the seawater ³⁴S/³²S ratio (7, 8). In this study, we use pristine authigenic marine barite crystals to trace the evolution of the marine S-isotope ratio across the PETM. We report our results in the traditional delta notation ($\delta^{34}\text{S}$), which expresses the isotopic ratio of ³⁴S/³²S as a difference relative to a reference standard:

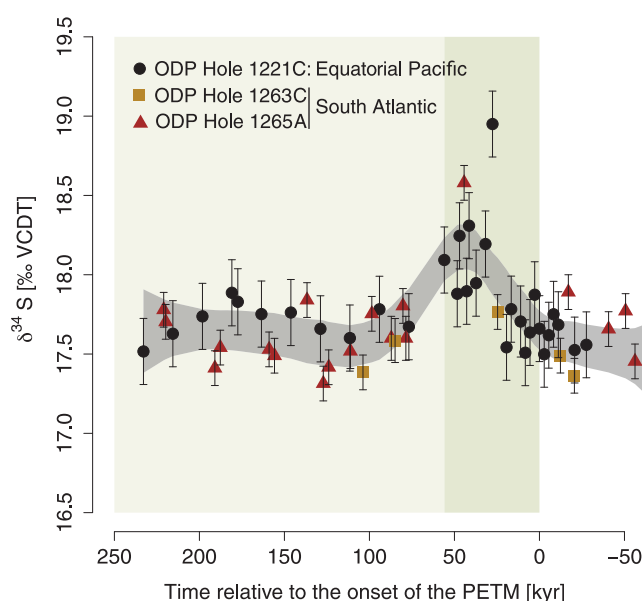
$$\delta^{34}\text{S} = \left(\frac{\left(\frac{^{34}\text{S}}{^{32}\text{S}} \right)_{\text{Sample}}}{\left(\frac{^{34}\text{S}}{^{32}\text{S}} \right)_{\text{Standard}}} - 1 \right) \times 1000 [\text{‰}]$$

where X^{S} denotes the concentration of the respective sulfur isotope. All results are reported relative to the Vienna Canyon Diabolo Troilite (VCDT).

We observe a 1‰ shift in $\delta^{34}\text{S}$ during the PETM (Fig. 1), which implies the removal of 8×10^{16} mol of reduced sulfur within 50 thousand years from the oceans (assuming a marine sulfate concentration of 5 mM) (9). This requires more than double the pre-PETM steady-state pyrite burial flux. More important, the rapid decay of the S-isotope excursion occurs within ~ 100 thousand years. Owing to the large size of the marine sulfate reservoir and the small input and output fluxes to the ocean, the marine S-isotope ratio is expected to retain the peak S-isotope values of the excursion for millions of years before returning to pre-excursion conditions. The observed rapid decay of the S-isotope excursion requires the input of 8×10^{16} mol of isotopically light sulfur back into the ocean before the end of the PETM.

Changes in pyrite burial are traditionally invoked to explain changes in seawater $\delta^{34}\text{S}$. It is assumed that most of the pyrite burial happens on continental shelves (10, 11). As such, pyrite burial rates correlate with sea-level variations, i.e., during high stands, shelf areas expand and pyrite burial increases, whereas during low stands, shelf areas contract and previously buried pyrite is reoxidized (11). Sea-level estimates suggest a rise

Fig. 1. $\delta^{34}\text{S}$ data of authigenic marine barite crystals across the PETM. The gray envelope denotes the 95% confidence interval of the LOESS regression (see supplementary materials). Shaded boxes indicate the extents of the main and total PETM interval (1). Error bars are 1σ .



¹Department of Earth Sciences, University of Toronto, Toronto, Ontario, M5S 3B1, Canada. ²Institute of Marine Science, University of California–Santa Cruz, Santa Cruz, CA 95064, USA.

*Corresponding author. Email: weiqi.yao@mail.utoronto.ca

on the order of 25 m during the PETM (12). Using previously published parametrizations (11), this increase results in the additional burial of 4.45×10^{10} mol pyrite per year, about 4% of that required to explain the observed isotope shift. The idea that the shelf acts as a transient pyrite reservoir has been expressed previously (11, 13), but the observed isotope signal requires an increase in the global shelf area equivalent to 13 to 32 times the size of the contiguous United States. Clearly, these numbers depend on a variety of assumptions, chiefly, the OM and pyrite burial rates on the shelves; however, they do demonstrate the magnitude of required flux changes.

Previous authors proposed that the importance of continental shelves for pyrite burial may have been overstated (11, 14) and that the long-term burial of pyrite is controlled by the intensity of anaerobic methane oxidation (AMO) on continental slopes (14). This idea is particularly intriguing because it links pyrite burial to subsurface methane flux, which in turn has been suggested as a possible explanation for the negative carbon-isotope excursion in the PETM (14, 15). In this scenario, the rapid release of isotopically light methane would not only drive the observed C-cycle excursion but also increase pyrite burial rates.

Linking pyrite burial flux to methane flux is elegant and provides a plausible explanation for the rise in the marine $\delta^{34}\text{S}$ ratio during the PETM. Furthermore, once the methane has been vented, gas hydrate reservoirs will have to recharge and thus AMO fluxes and AMO-associated pyrite burial rates will cease (14). However, even if we assume the extreme case that all pyrite burial is AMO-controlled, it would take at least 0.2 million years to return the marine $\delta^{34}\text{S}$ ratio back to its pre-excursion value (see supplementary materials). That means it is not sufficient just to inhibit all pyrite burial; rather, we need to inject large quantities of isotopically light S into the oceans to explain the observed S-isotope signal. The speed and magnitude of the required ^{32}S flux require a transient reservoir of ^{32}S -enriched sulfur that is sufficiently large to hold 8×10^{16} mol of isotopically depleted S and can be charged and released in less than 100 thousand years.

Here we propose that the marine oxygen minimum zone (OMZ) constitutes such a reservoir. With the exception of a few upwelling zones, oxygen concentrations in the OMZ are usually high enough to render MSR energetically unfavorable. However, the marine oxygen pool is small, and changes in circulation or export production can alter the oxygen content of the OMZ rapidly. Once oxygen concentrations in the OMZ drop below $4 \mu\text{M}$ (16), MSR becomes the dominant OM remineralization pathway. In the modern ocean, MSR is limited by OM supply and the diffusive supply of sulfate into the sediment column. Shifting the location of sulfate reduction from the sediment into the intermediate water column would not only increase the availability of sulfate, but drastically increase the quality and quantity of OM. This will accelerate the globally integrated sulfate reduction flux and transfer ^{32}S into the

OMZ, effectively separating the ocean into two sulfur reservoirs.

To estimate the required changes to the OMZ, we use S-isotope fractionation factors (α) between 43 and 70‰ (7, 8), and an average sulfide concentration up to 0.5 mM, well below the 1 mM threshold where chemocline upwelling and the subsequent release of H_2S into the atmosphere become a possibility (17). The value of 0.5 mM is high compared to sulfide concentrations observed in anoxic upwelling zones in the modern ocean (16), but well within the limits of marine sulfide concentrations observed within more restricted settings (e.g., 6.14 mM in Framvaren Fjord) (18). Modeling the data using the statistically significant maximum of 18.35‰ (Fig. 2) requires sulfidic water volumes between 1.6×10^{17} to $2.6 \times$

10^{17} m^3 . The highest measured $\delta^{34}\text{S}$ value (18.95‰) can be explained with a sulfidic water volume of no more than $2.7 \times 10^{17} \text{ m}^3$ for α of 70‰. To put these estimates into perspective: In the modern ocean, the OMZ sensu stricto ($\text{O}_2 < 20 \mu\text{M}$) encompasses about 1% of the global ocean volume ($1.38 \times 10^{18} \text{ m}^3$) (19), whereas the above numbers suggest an order-of-magnitude increase to 10 to 20% of the ocean volume. Indeed, several studies suggest that low O_2 or anoxic waters were present during the PETM at intermediate depths in many ocean basins (20–23).

Modeling predictions exploring the effects of anthropogenic climate change suggest that measurable oxygen loss from the subarctic North Pacific will occur by 2030 to 2040 (24) and that the total volume of suboxic ocean water will

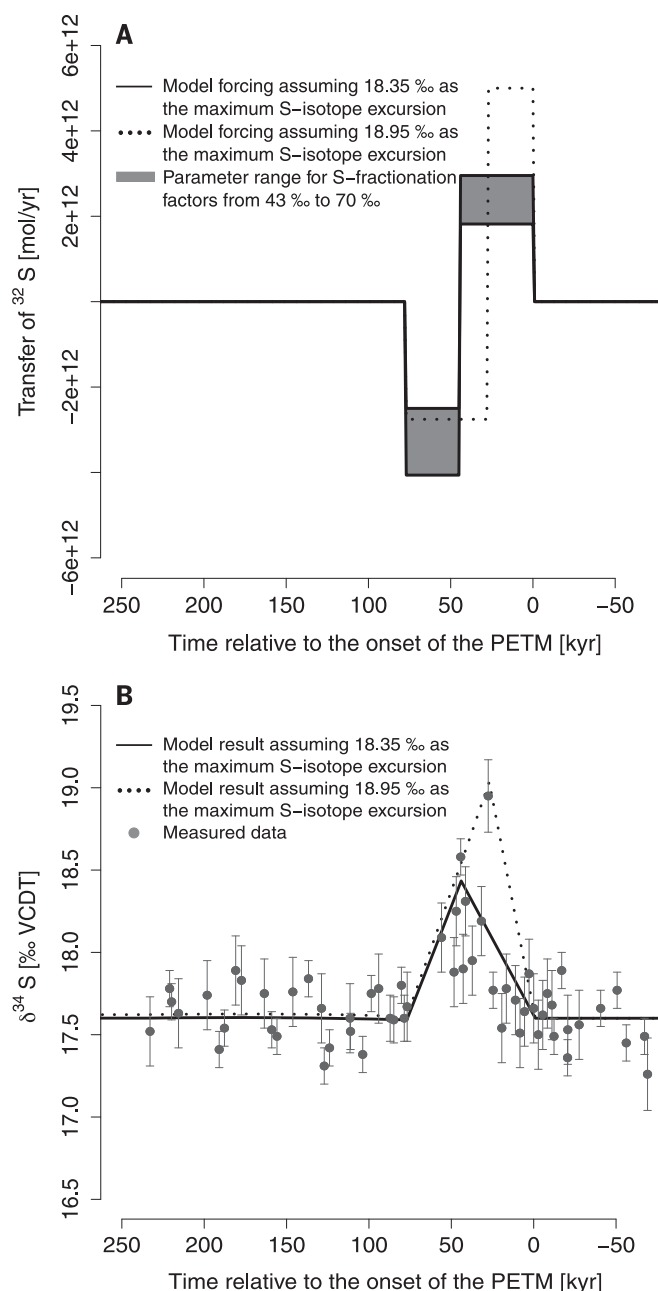


Fig. 2. Model forcing and the resulting S-isotope ratios versus measured data. (A) Transfer of ^{32}S into and out of the OMZ. (B) Resulting changes in the marine sulfate $\delta^{34}\text{S}$ across the PETM. See supplementary materials for full model results. Note that in order to model the peak at 18.95‰, the timing of the flux changes has to be adjusted.

expand by 50% by 2100 (25). Once local oxygen concentrations drop below 4 μM (16), sulfate reduction will commence, resulting in the production of H_2S , which is toxic at levels as low as 4 $\mu\text{g/liter}$ (26). This will (A) create an ocean internal reservoir of reduced sulfur; (B) create an ocean that is no longer well mixed with respect to sulfate; and (C) compress the ecozones of fish species that venture in the mesopelagic and bathypelagic zones (27) and change their ecosystem structure, which could jeopardize 10 to 50% of worldwide pelagic predator diversity (27) with unknown consequences for global fish stocks. Our findings suggest that (A) and (B) have been well expressed during the PETM, and given the similarities between the PETM and the IPCC RCP8.5 emission scenario, (C) will be a possibility in the not-too-distant future.

REFERENCES AND NOTES

1. F. Nunes, R. Norris, in *Proceedings of the Ocean Drilling Program, Scientific Results*, P. Wilson, M. Lyle, J. Firth, Eds. (College Station, TX, 2005), vol. 199, pp. 1–12.
2. R. Zeebe, A. Ridgwell, J. Zachos, *Nat. Geosci.* **9**, 325–329 (2016).
3. "IPCC, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker et al., Eds. (Cambridge Univ. Press, 2013).
4. J. C. Walker, *Mar. Geol.* **70**, 159–174 (1986).
5. U. G. Wortmann, B. M. Chernyavsky, *Nature* **446**, 654–656 (2007).
6. G. Claypool, W. Holser, I. Kaplan, H. Sakai, I. Zak, *Chem. Geol.* **28**, 199–260 (1980).
7. I. Kaplan, K. Emery, S. Rittenberg, *Geochim. Cosmochim. Acta* **27**, 297–331 (1963).
8. U. Wortmann, S. Bernasconi, M. Böttcher, *Geology* **29**, 647–650 (2001).
9. U. G. Wortmann, A. Paytan, *Science* **337**, 334–336 (2012).
10. A. Turchyn, D. Schrag, *Earth Planet. Sci. Lett.* **241**, 763–779 (2006).
11. S. Markovic, A. Paytan, U. Wortmann, *Biogeosciences* **12**, 3043–3060 (2015).
12. K. G. Miller et al., *Science* **310**, 1293–1298 (2005).
13. J. Higgins, D. Schrag, *Earth Planet. Sci. Lett.* **245**, 523–537 (2006).
14. G. Dickens, *Clim. Past* **7**, 831–846 (2011).
15. G. Dickens, J. O'Neil, D. Rea, R. Owen, *Paleoceanography* **10**, 965–971 (1995).
16. V. Brüchert et al., *Geochim. Cosmochim. Acta* **67**, 4505–4518 (2003).
17. L. Kump, A. Pavlov, M. Arthur, *Geology* **33**, 397–400 (2005).
18. W. Landing, S. Westerlund, *Mar. Chem.* **23**, 329–343 (1988).
19. A. Paulmier, D. Ruiz-Pino, *Prog. Oceanogr.* **80**, 113–128 (2009).
20. M. Nicolo, G. Dickens, C. Hollis, *Paleoceanography* **25**, PA4210 (2010).
21. A. Dickson, A. Cohen, A. Coe, *Geology* **40**, 639–642 (2012).
22. A. Winguth, E. Thomas, C. Winguth, *Geology* **40**, 263–266 (2012).
23. X. Zhou et al., *Paleoceanography* **31**, 1532–1546 (2016).
24. M. C. Long, C. Deutsch, T. Ito, *Global Biogeochem. Cycles* **30**, 381–397 (2016).
25. A. Oschlies, K. Schulz, U. Riebesell, A. Schmittner, *Global Biogeochem. Cycles* **22**, GB4008 (2008).
26. L. Smith, D. Oseid, I. Adelman, S. Broderius, Effect of hydrogen sulfide on fish and invertebrates. Part I – Acute and chronic toxicity studies. *Ecol. Res. Tech. Rep. Ser. No.3* (U.S. Environmental Protection Agency, 1976).
27. L. Stramma et al., *Nat. Clim. Chang.* **2**, 33–37 (2012).

ACKNOWLEDGMENTS

We thank G.R. Dickens and S. Markovic for discussions and comments on an early draft of this paper; H. Li for support with the isotope analysis; and the three anonymous reviewers for helping to improve this manuscript. **Funding:** This research was supported by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada (NSERC) to U.G.W. and a National Science Foundation (NSF) CAREER grant OCE-0449732 to A.P. **Author contributions:** All authors contributed to the ideas expressed in this manuscript. W.Y. and U.G.W. developed the project. A.P. provided barite samples. W.Y. performed isotope measurements, data analysis, and computations and did the art work. W.Y., U.G.W., and A.P. were responsible for data interpretation and manuscript writing. **Competing interests:** The authors declare no competing interest. **Data and materials availability:** All data are available in supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/804/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 and S2
References (28–58)

27 December 2017; accepted 7 July 2018
Published online 19 July 2018
10.1126/science.aar8658

An intrinsic S/G₂ checkpoint enforced by ATR

Joshua C. Saldivar¹, Stephan Hamperl^{1*}, Michael J. Bocek¹, Mingyu Chung¹, Thomas E. Bass², Fernanda Cisneros-Soberanis^{3,4}, Kumiko Samejima³, Linfeng Xie⁵, James R. Paulson⁵, William C. Earnshaw³, David Cortez², Tobias Meyer¹, Karlene A. Cimprich^{1†}

The cell cycle is strictly ordered to ensure faithful genome duplication and chromosome segregation. Control mechanisms establish this order by dictating when a cell transitions from one phase to the next. Much is known about the control of the G₁/S, G₂/M, and metaphase/anaphase transitions, but thus far, no control mechanism has been identified for the S/G₂ transition. Here we show that cells transactivate the mitotic gene network as they exit the S phase through a CDK1 (cyclin-dependent kinase 1)-directed FOXM1 phosphorylation switch. During normal DNA replication, the checkpoint kinase ATR (ataxia-telangiectasia and Rad3-related) is activated by ETAA1 to block this switch until the S phase ends. ATR inhibition prematurely activates FOXM1, deregulating the S/G₂ transition and leading to early mitosis, underreplicated DNA, and DNA damage. Thus, ATR couples DNA replication with mitosis and preserves genome integrity by enforcing an S/G₂ checkpoint.

To avoid loss of genetic information, cells must replicate their DNA before mitosis (1). Recent data show that replication delays the onset of mitosis (2), yet how cells sense the end of the S phase and coordinate these processes remains a fundamental biological question. A checkpoint monitoring the end of DNA replication has been proposed (3), but effectors of such an S/G₂ checkpoint have not been identified. DNA damage can activate the checkpoint kinase ATR (ataxia-telangiectasia and Rad3-related), arresting S or G₂ cells by inactivating the CDC25-cyclin-dependent kinase (CDK) pathway (4). Because ATR ensures completion of replication before mitosis during normal proliferation (5, 6), we hypothesized that it may regulate the S/G₂ transition.

To test this hypothesis, we used quantitative image-based cytometry (QIBC) (fig. S1) (7) combined with pulse-chase assays (fig. S2) and measured S-to-M and G₂-to-M progression. S-phase cells treated with ATR inhibitors (ATRi) underwent accelerated mitotic entry (Fig. 1, A and B). Unexpectedly, ATR inhibition in G₂ did not accelerate mitotic entry rates, whereas WEE1 inhibition did (Fig. 1C). This extends previous findings (5, 6) suggesting that ATR acts in the

S phase and not in the G₂ phase of normal cell cycles to delay mitosis.

These observations are consistent with ATR inhibition shortening the S phase. Nevertheless, a pulse-chase-pulse assay (fig. S3) and live-cell imaging with an EYFP-PCNA (enhanced yellow fluorescent protein–proliferating cell nuclear antigen) biosensor (8) (fig. S4) revealed that S-phase shortening alone cannot explain the combined shortening of S and G₂ after ATR inhibition (fig. S3, B to D). We hypothesized that ATR controls subsequent G₂ duration by delaying accumulation of promitotic factors (9). Inhibiting ATR led to premature cyclin B protein accumulation in the S phase (Fig. 1, D to F, and fig. S5). Moreover, ATR inhibition prematurely increased cyclin B1 mRNA in early-S-phase synchronized cells, whereas in mock-treated synchronized cells, cyclin B mRNA increased concomitantly with the S/G₂ transition (Fig. 1G and fig. S6). To determine whether ATR similarly controls the transcription of additional genes, we performed RNA sequencing (RNA-seq). Unsupervised clustering analysis of sequenced reads revealed a group of genes prematurely upregulated with ATRi in the S phase (Fig. 1H, group 5; fig. S6E; and table S1). Consistent with their normal G₂ expression, Gene Ontology (GO)-term analysis of group 5 showed enrichment for promitotic factors (Fig. 1, I and J). These findings suggest a G₂/M gene network poised for transcription in the S phase but repressed by ATR until G₂.

Next, we analyzed publicly available ChIP (chromatin immunoprecipitation) sequencing, also known as ChIP-seq, data for transcription factor enrichment at the promoters of group 5 genes. B-MYB and FOXM1, transactivators of a mitotic transcription program (10, 11), were highly enriched at these sites (fig. S7A). We asked whether knockdown of either would prevent premature

cyclin B accumulation. B-MYB knockdown caused a G₁ arrest, precluding further study (fig. S7, B to D). FOXM1 knockdown, which did not block S-phase entry, prevented premature cyclin B accumulation in ATR-inhibited cells (fig. S7, B and E), confirming that FOXM1 drives premature cyclin B expression.

To determine whether ATR regulates B-MYB and FOXM1, we monitored their phosphorylation, a requirement for their activation (11, 12). Whereas both proteins normally exhibit a profound phosphorylation shift at the S/G₂ transition, ATR inhibition in the early S phase triggered immediate hyperphosphorylation (Fig. 2, A and B, and fig. S7F). In asynchronous cells, single-cell QIBC analysis revealed that FOXM1 threonine-600 (T600) phosphorylation (fig. S8A), a CDK-dependent activating modification (12), rises in the early S phase, then again in G₂. ATR inhibition caused rapid and premature FOXM1 phosphorylation throughout the S phase, raising pFOXM1 to G₂ levels (Fig. 2, C to F, and fig. S8, B to D). This premature phosphorylation occurred even after blocking replication with thymidine (fig. S9A). Moreover, mRNA FISH (fluorescence in situ hybridization) analysis of the FOXM1 target PLK1 revealed that ATRi induced its premature expression with or without thymidine (fig. S9, B and C). Thus, ATR suppresses FOXM1 phosphorylation and downstream transcription in S-phase cells, irrespective of S-phase progression.

The marked difference in S- and G₂-phase pFOXM1 levels suggests that a specific regulatory event defines the S/G₂ transition. Thus, we measured pFOXM1 levels during the late S phase, the S/G₂ transition (fig. S10), and the early G₂ phase. pFOXM1 levels rose abruptly in the S/G₂ population (Fig. 2, G and H), a finding confirmed using EYFP-PCNA and pFOXM1 imaging (fig. S11, A and B). Given that this population makes up less than 4% of the late S and early G₂ cells, phosphorylation is switchlike, in that it happens rapidly at the S/G₂ transition.

Paradoxically, cyclin A-CDK phosphorylates FOXM1 (12), but total cyclin A-CDK activity gradually increases throughout the S phase (13), which is inconsistent with switchlike behavior. To test whether a specific cyclin A-CDK complex mediates this behavior, we measured pFOXM1 levels after either CDK2 or CDK1 inhibition. CDK2 inhibitors had little effect on FOXM1 phosphorylation (Fig. 3, A and B, and fig. S12A), although they decreased EdU (5-ethynyl-2'-deoxyuridine) incorporation (fig. S12B). In contrast, CDK1 inhibition prevented FOXM1 phosphorylation at the S/G₂ transition and after ATR inhibition (Fig. 3, A and C, and fig. S12). Furthermore, analog-sensitive CDK1 inhibition (14) reduced FOXM1 phosphorylation, and CDK2 inhibition had no additional effect (fig. S13). These results show that CDK1 phosphorylates FOXM1 at the S/G₂ transition.

Given that ATR suppresses FOXM1 phosphorylation until G₂, we hypothesized that ATR activity may abruptly decline as the S phase ends, relieving an intrinsic checkpoint. Consistent

¹Department of Chemical and Systems Biology, Stanford University School of Medicine, 318 Campus Drive, Stanford, CA 94305-5441, USA. ²Department of Biochemistry, Vanderbilt University School of Medicine, 2215 Garland Avenue, Nashville, TN 37232, USA. ³Wellcome Centre for Cell Biology, University of Edinburgh, King's Buildings, Max Born Crescent, Edinburgh EH9 3BF, Scotland, UK. ⁴Unidad de Investigación Biomédica en Cáncer, Instituto de Investigaciones Biomédicas–Universidad Nacional Autónoma de México; Instituto Nacional de Cancerología, México City 14080, México. ⁵Department of Chemistry, University of Wisconsin–Oshkosh, 800 Algoma Boulevard, Oshkosh, WI 54901, USA.

*Present address: Institute of Epigenetics and Stem Cells, Helmholtz Zentrum München, Marchioninistraße 25, 81377 München, Germany. †Corresponding author. Email: cimprich@stanford.edu

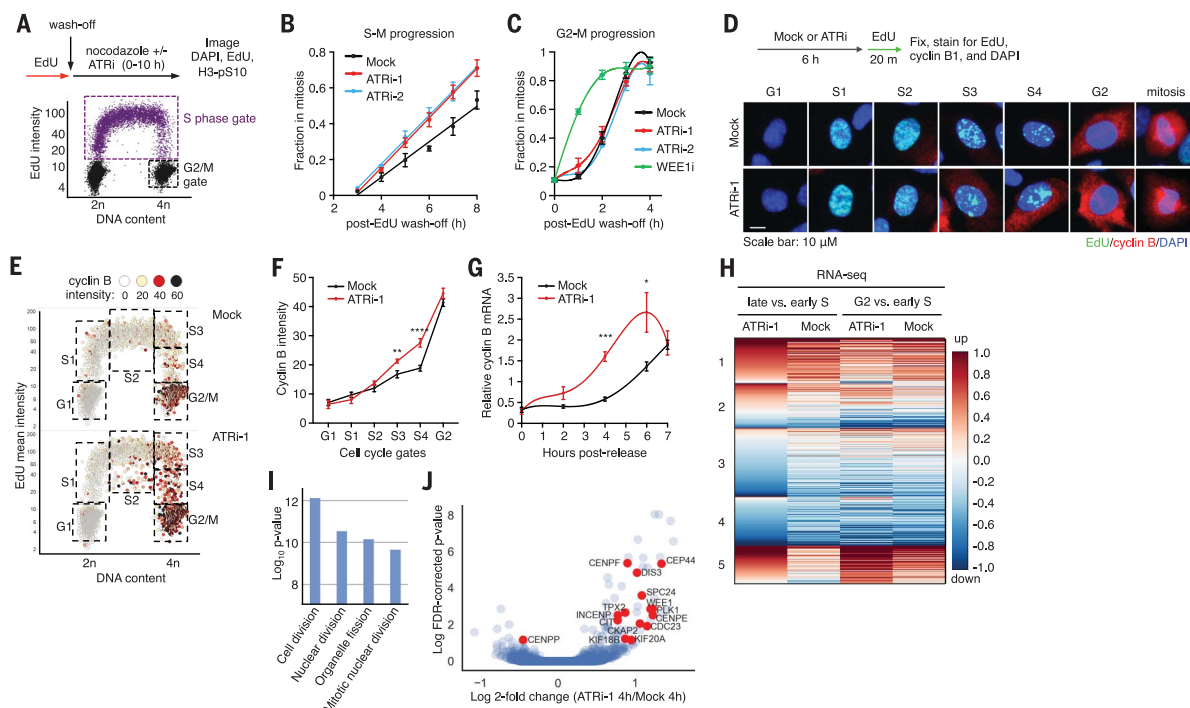


Fig. 1. ATR represses a mitotic gene network during the S phase.

(A) S-to-M or G₂-to-M progression assay showing the cellular distribution at time = 0 and the gating scheme. Details are shown in fig. S2. 2n and 4n refer to the number of sets of chromosomes. h, hours; DAPI, 4',6-diamidino-2-phenylindole. (B and C) Fraction of S-gated (B) or G₂/M-gated (C) hTERT RPE-1 cells in mitosis as a function of time. Error bars, SEM of $n = 3$. (D) Experimental design and representative images of hTERT RPE-1 cells in G₁, early to late S phase (S₁ to S₄), G₂, or mitosis. m, minutes. (E) QIBC plots of DNA content versus EdU mean intensity, with mean cyclin B cytoplasmic intensity shown by a color scale. Boxes indicate gated populations used for analyses. (F) Mean cyclin B cytoplasmic intensities

of gated populations from (E). Error bars, SEM of $n = 4$. $^{**}P < 0.01$, $^{****}P < 0.0001$. (G) RT-qPCR (quantitative reverse transcription polymerase chain reaction) analysis of cyclin B1 mRNA from synchronized hTERT RPE-1 cells described in fig. S6. Error bars, SEM of $n = 4$. $^{*}P < 0.05$, $^{***}P < 0.001$. (H) Unsupervised clustering analysis of mock or ATRi and early S, late S, or G₂ synchronized cells (selected comparisons shown). Heatmap indicates the fold change of normalized RNA-seq reads between indicated samples ($n = 3$). (I) Top 4 GO terms for group 5 genes from (H). (J) Volcano plot of log₁₀ false discovery rate-corrected P values and log₂ fold change in gene expression for ATRi versus mock treatment of cells in the late S phase.

with this hypothesis, ATR-dependent H2AX phosphorylation (γ H2AX) increased as cells entered the S phase, peaked in the mid-S phase, and rapidly decreased as cells completed DNA replication (Fig. 3, D and E, and fig. S14, A to D). Thus, ATR is active throughout the normal S phase, but its activity drops at the S/G₂ transition, allowing rapid FOXM1 phosphorylation.

In mammalian cells, either a RAD9A-TOPBP1 or ETAA1-dependent pathway activates ATR (4). ETAA1 knockdown reduced ATR activity in an unperturbed S phase, whereas RAD9A knockdown did not, even though it reduced γ H2AX after replication stress (Fig. 4A and fig. S14E). Moreover, deleting the ETAA1 ATR-activation domain greatly reduced ATR activity in an unperturbed S phase, but not in response to replication stress, whereas auxin-mediated degradation of a TOPBP1-mAID fusion had the opposite effect (Fig. 4, B and C, and fig. S15). Thus, ETAA1 activates ATR during an unperturbed S phase, whereas TOPBP1-RAD9A activates ATR to enforce the replication stress response, providing a rationale for the existence of multiple ATR activators.

ATR enforces its checkpoint functions in part by activating CHK1 (checkpoint kinase 1), an

effector kinase that inhibits CDKs (4). CHK1 inhibition also triggered premature FOXM1 phosphorylation and cyclin B accumulation (fig. S16, A to C). Thus, the repressive activity of an ETAA1-ATR-CHK1 pathway on the cyclin A-CDK1-FOXM1 axis controls the S/G₂ transition (fig. S16D).

FOXM1 overexpression is sufficient to accelerate mitotic entry (10); therefore, we asked whether ATR slows the S-to-M progression by controlling FOXM1. Consistent with this hypothesis, reducing FOXM1 to levels that permit mitotic entry, while suppressing premature cyclin B expression in ATR-inhibited cells (fig. S7E), prevented ATRi-induced premature mitosis (Fig. 4, D and E). Furthermore, partial FOXM1 knockdown reduced ultrafine anaphase bridges (UFBs) and 53BP1 bodies after ATR inhibition (Fig. 4, F and G, and fig. S17). Because these both indicate a failure to complete DNA replication before mitosis (5, 15, 16), we conclude that proper control of the S/G₂ phosphorylation switch promotes the completion of DNA replication and prevents genome instability.

Our data unveil an ATR-controlled S/G₂ phosphorylation switch that initiates the G₂/M tran-

scription program. By repressing CDK1, ATR blocks this switch until the S phase ends to properly time G₂-specific events. Although the molecular signal activating ATR in an unperturbed S phase is unclear, ETAA1's role suggests that single-stranded DNA (ssDNA) generated by ongoing replication, rather than unreplicated DNA, is a critical component. Transiently formed RPA-coated ssDNA could serve as a platform for colocalizing ETAA1 and ATR, sustaining ATR activity throughout the S phase. The decline in ATR activity that occurs once replication is complete then signals the S/G₂ transition, executed via the CDK1-pFOXM1 switch (fig. S18).

Because ATR ensures that G₂ depends on S-phase completion, we refer to this as an intrinsic S/G₂ checkpoint and propose that this checkpoint monitors replication completion (3). This checkpoint prevents a cellular identity crisis in which the S and G₂ phases overlap, which would cause underreplication, early mitosis, and subsequent DNA damage. Given the frequent overexpression of FOXM1 in cancer (17), deregulation of this fundamental cell cycle transition could be a common event contributing to cancer genome instability.

Fig. 2. ATR controls an S/G₂ FOXM1 phosphorylation switch. (A) Western blots of hTERT RPE-1 cells synchronized as in fig S6A. Async, asynchronous. (B) Quantification of FOXM1 band intensities from (A) plotted as red lines (right y axis). Black lines (left y axis) represent the fraction of S-phase synchronized cells, calculated as described in fig. S6. B to D. (C) Experimental design and representative images of S-phase hTERT RPE-1 cells. Scale bar, 15 μ m. (D) QIBC plots of DNA content versus EdU mean intensity, with mean pFOXM1 T600 intensity shown by a color scale. Boxes indicate gated populations used for analyses. (E) Median pFOXM1 intensities from (D). Error bars, SEM of $n = 4$. (F) pFOXM1 T600 mean intensity in cells treated as described in (C). (G) QIBC plot of EdU-labeled cells with 4n DNA content. Boxes indicate gated populations used for analysis in (H). Numbers indicate cells in each gated population. The S/G₂ population was determined as described in fig. S10. (H) pFOXM1 T600 mean intensity in indicated populations. In (F) and (H), whiskers indicate the 10th and 90th percentiles, boxes span the 25th to 75th percentiles, and lines inside boxes represent medians.

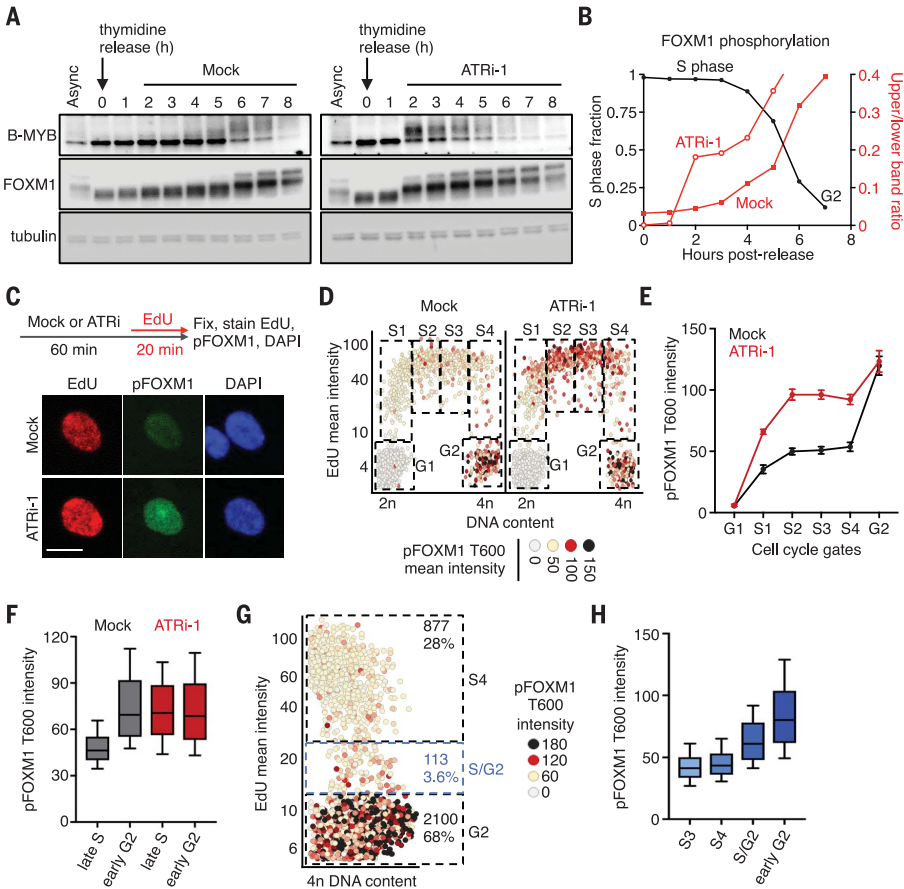
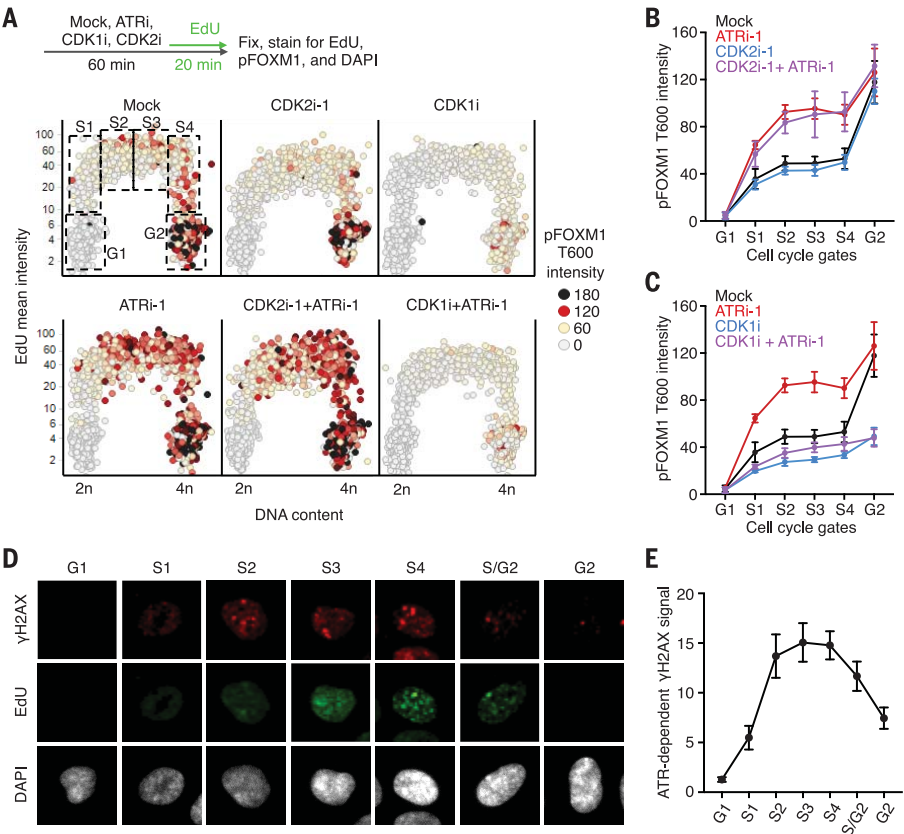


Fig. 3. ATR is active until G₂, preventing CDK1-dependent FOXM1 phosphorylation. (A) Experimental design and QIBC plots of DNA content versus EdU mean intensity, with mean pFOXM1 T600 intensity shown by a color scale, in hTERT RPE-1 cells. Boxes indicate gated populations used for analyses. (B and C) Median pFOXM1 intensities of populations shown in (A). Error bars, SEM of $n = 3$. (D) Representative images of indicated cell stage in hTERT RPE-1 cells. (E) ATR activity during an unperturbed cell cycle as described in fig. S14. Error bars, SEM of $n = 3$.



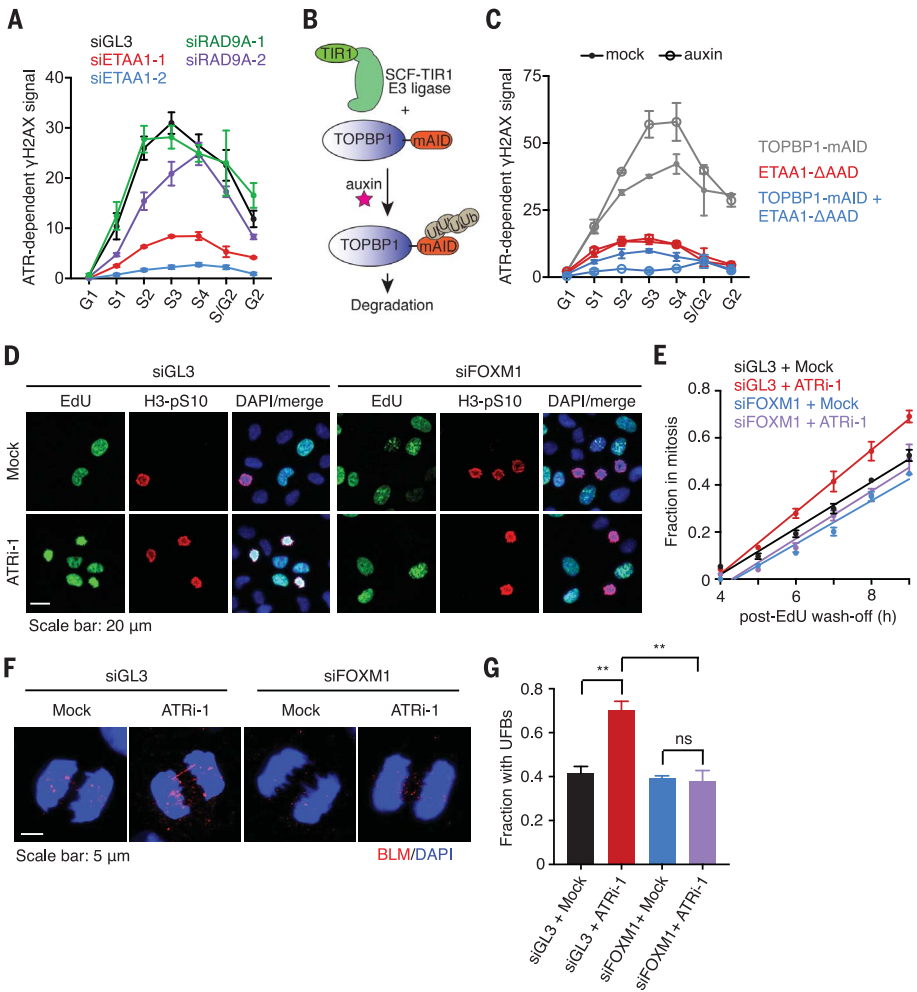


Fig. 4. The ETAA1-ATR pathway couples the S phase and mitosis. (A) ATR activity in small interfering RNA (siRNA)-transfected hTERT RPE-1 cells during an unperturbed cell cycle as described in fig. S14. Error bars, SEM of $n = 3$. (B) Illustration of the mAID degron system fused to TOPBP1. Ub, ubiquitin. (C) ATR activity in HCT116 cell lines mock-treated or auxin-treated (2 hours) as described in fig. S15A. Error bars, SEM of $n = 3$. AAD, ATR-activation domain. (D and E) S-to-M progression assay (see Fig. 1, A and B) in hTERT RPE-1 cells 40 hours after siRNA transfection. Representative images are 6 hours after EdU wash-off. (E) Fraction of S phase-gated cells in mitosis as a function of time after EdU wash-off. Error bars, SEM of $n = 3$. (F) Representative images of anaphase cells mock- or ATRi-treated for 8 hours, 40 hours after siRNA transfection. (G) Fraction of cells with UFBs. Error bars, SEM of $n = 3$. $**P < 0.01$; ns, not significant.

REFERENCES AND NOTES

1. D. O. Morgan, *The Cell Cycle: Principles of Control* (Primers in Biology, New Science Press in association with Oxford Univ. Press, 2007).
2. B. Lemmens *et al.*, *Mol. Cell* **71**, 117–128.e3 (2018).
3. S. J. Elledge, *Science* **274**, 1664–1672 (1996).
4. J. C. Saldivar, D. Cortez, K. A. Cimprich, *Nat. Rev. Mol. Cell Biol.* **18**, 622–636 (2017).
5. J. K. Eykelboom *et al.*, *Cell Rep.* **5**, 1095–1107 (2013).
6. S. Ruiz *et al.*, *Mol. Cell* **62**, 307–313 (2016).
7. L. I. Toledo *et al.*, *Cell* **155**, 1088–1103 (2013).
8. A. T. Hahn, J. T. Jones, T. Meyer, *Cell Cycle* **8**, 1044–1052 (2009).
9. K. Akopyan *et al.*, *Mol. Cell* **53**, 843–853 (2014).
10. J. Laoukili *et al.*, *Nat. Cell Biol.* **7**, 126–136 (2005).
11. S. Sadasivam, S. Duan, J. A. DeCaprio, *Genes Dev.* **26**, 474–489 (2012).
12. J. Laoukili *et al.*, *Mol. Cell Biol.* **28**, 3076–3087 (2008).
13. S. L. Spencer *et al.*, *Cell* **155**, 369–383 (2013).
14. H. Hochegger *et al.*, *J. Cell Biol.* **178**, 257–268 (2007).
15. K. L. Chan, T. Palmai-Pallag, S. Ying, I. D. Hickson, *Nat. Cell Biol.* **11**, 753–760 (2009).
16. C. Lukas *et al.*, *Nat. Cell Biol.* **13**, 243–253 (2011).
17. J. Laoukili, M. Stahl, R. H. Medema, *Biochim. Biophys. Acta* **1775**, 92–102 (2007).

ACKNOWLEDGMENTS

We thank F. Ochs, J. Ferrell, and the Cimprich laboratory for comments on the manuscript. **Funding:** This work was supported by grants from the NIH to K.A.C. (ES016486), D.C. (CA102729), and T.M. (GM127026); the Wellcome Fund (107022 and 203149) to W.C.E.; the American Cancer Society (PF-15-165-01-DMC) and the Burroughs Wellcome Fund Postdoctoral Enrichment Program to J.C.S.; the German Research Foundation DFG (HA 6996/1-1) to S.H.; and by a Mexican Government CONACYT fellowship to F.C.-S. **Author contributions:** J.C.S., W.C.E., D.C., K.S., and K.A.C. designed experiments. J.C.S., S.H., F.C.-S., K.S., J.R.P., L.X., T.E.B., and M.C. performed experiments. M.J.B. and M.C. performed the bioinformatics analyses. J.C.S., M.J.B., M.C., T.M., and K.A.C. analyzed the data. J.C.S. and K.A.C. wrote the manuscript. **Competing interests:** None declared. **Data and materials availability:** RNA-seq data are available in the GEO repository (accession number GSE116131).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/806/suppl/DC1
Materials and Methods
Figs. S1 to S18
Table S1
References (18–22)

13 September 2017; resubmitted 7 February 2018
Accepted 13 July 2018
10.1126/science.aap9346

GENETICS

Biallelic *RIPK1* mutations in humans cause severe immunodeficiency, arthritis, and intestinal inflammation

Delphine Cuchet-Lourenço^{1*}, Davide Eletto^{1*}, Changxin Wu^{1*†}, Vincent Plagnol², Olivier Papapietro¹, James Curtis¹, Lourdes Ceron-Gutierrez³, Chris M. Bacon^{4,5}, Scott Hackett⁶, Badr Alsaleem⁷, Mailis Maes¹, Miguel Gaspar¹, Ali Alisaac^{1,8}, Emma Goss¹, Eman AlIdrissi⁹, Daniela Siegmund¹⁰, Harald Wajant¹⁰, Dinakantha Kumararatne³, Mofareh S. AlZahrani⁹, Peter D. Arkwright¹¹, Mario Abinun¹², Rainer Doffinger³, Sergey Nejentsev^{1†}

RIPK1 (receptor-interacting serine/threonine kinase 1) is a master regulator of signaling pathways leading to inflammation and cell death and is of medical interest as a drug target. We report four patients from three unrelated families with complete *RIPK1* deficiency caused by rare homozygous mutations. The patients suffered from recurrent infections, early-onset inflammatory bowel disease, and progressive polyarthritis. They had immunodeficiency with lymphopenia and altered production of various cytokines revealed by whole-blood assays. In vitro, *RIPK1*-deficient cells showed impaired mitogen-activated protein kinase activation and cytokine secretion and were prone to necroptosis. Hematopoietic stem cell transplantation reversed cytokine production defects and resolved clinical symptoms in one patient. Thus, *RIPK1* plays a critical role in the human immune system.

Primary immunodeficiencies (PIDs) are a heterogeneous group of disorders characterized by increased susceptibility to infection. In many cases, PIDs are monogenic disorders that follow Mendelian inheritance; mutations in more than 300 genes have been shown to cause PIDs (1). However, in many PID patients, causative mutations remain unknown. Identification of such mutations not only facilitates diagnosis of PIDs, but also can provide fundamental knowledge about the roles of the affected proteins in the human immune system.

Here, we used exome sequencing to identify causative mutations in a heterogeneous cohort of PID patients with unknown genetic etiology (2). We excluded known polymorphisms (2) and studied rare variants. We noticed that four patients (P1 to P4) from three unrelated consanguineous families (Fig. 1A) had homozygous loss-of-function mutations in the same gene, *RIPK1* (ENST00000380409). In patients P1 and P2 from family A, the mutation was a 4-nucleotide frameshift deletion in *RIPK1* exon 6 that led to a premature stop codon (Fig. 1B). In patient P3 from family B, we found a 21-nucleotide deletion that removed one nucleotide in *RIPK1* exon 4 and 20 nucleotides in the following intron (Fig. 1B). This deletion activated an alternative splice site

in intron 4, so that the *RIPK1* transcript lacked the last nucleotide of exon 4 and had an insertion of 48 nucleotides from the following intron (fig. S1). Patient P4 from family C had a homozygous 2064-nucleotide deletion that completely removed *RIPK1* exon 4 (Fig. 1B and fig. S2). Parents of all patients were heterozygous carriers of these *RIPK1* mutations (Fig. 1A and fig. S3). Given that homozygous loss-of-function mutations in the *RIPK1* gene had never been reported in humans [e.g., absent from more than 120,000 subjects in the gnomAD database (3)], such mutations in our patients are likely to be pathogenic. The *RIPK1* gene encodes a 671-amino acid serine/threonine kinase (Fig. 1C). All three mutations mapped to the N-terminal kinase domain and produced premature stop codons. The *RIPK1* protein was absent in cells of patients P2, P3, and P4 (Fig. 1D). Therefore, all three homozygous mutations led to complete *RIPK1* deficiency. The four patients had lymphopenia, suffered from recurrent viral, bacterial, and fungal infections, exhibited early-onset inflammatory bowel disease (IBD) involving the upper and lower gastrointestinal tract, and developed arthritis (fig. S4 and tables S1 and S2) (2). Therefore, these clinical features characterize *RIPK1* deficiency in humans.

RIPK1 is a widely expressed cytosolic protein kinase that controls multiple signaling pathways leading to inflammation and apoptotic or necroptotic cell death (4, 5). *RIPK1* is present in protein complexes that mediate signal transduction from cell surface receptors, including tumor necrosis factor receptor 1 (TNFR1), Toll-like receptor 3 (TLR3), and TLR4 (4, 5). Stimulation of these receptors activates the canonical nuclear factor (NF)- κ B pathway and the mitogen-activated protein kinases (MAPKs). This leads to phosphorylation of NF- κ B and AP-1 transcription factors, which induces expression of pro-inflammatory and prosurvival genes (4, 6). To assess the functioning of these signaling pathways, we stimulated patients' skin fibroblasts with tumor necrosis factor- α (TNF- α) and polyinosinic-polycytidylic acid [poly(I:C)] in vitro. We found that phosphorylation of MAPK p38 and the AP-1 subunit cJun was markedly reduced (Fig. 2A and fig. S5), whereas phosphorylation of MAPK p42/44 (ERK2/1) and NF- κ B p65 was partially reduced (fig. S6). TNF- α -induced secretion of the cytokines interleukin-6 (IL-6) and RANTES by patients' fibroblasts was also diminished (fig. S7). Thus, similar to studies in *RIPK1*-deficient mice (7–11), proinflammatory signaling downstream of TNFR1 and TLR3 is impaired in the patients. We next investigated whether *RIPK1* deficiency affected cell viability. After stimulation with TNF- α or poly(I:C), we found significantly fewer viable fibroblasts of patients than those of healthy controls ($P < 0.001$; Fig. 2B). When we transduced patients' fibroblasts and expressed wild-type *RIPK1*, this viability defect was reversed (fig. S8) and activation of the MAPK and NF- κ B pathways was also rescued (fig. S9). To investigate the mechanism of cell death, we studied patients' fibroblasts 24 hours after poly(I:C) stimulation and found increased phosphorylation of *RIPK3* and *MLKL* (mixed lineage kinase domain-like pseudokinase), proteins that mediate necroptosis (12, 13) (Fig. 2C). We detected no cleavage of caspase-8 and minimal cleavage of caspase-3, indicating that apoptosis was not a major death mechanism of these cells (fig. S10). Consistent with these results, the *MLKL* inhibitor necrosulfonamide and, to a lesser extent, the *RIPK3* inhibitor GSK'872 rescued patients' cells from poly(I:C)-induced death, whereas the pan-caspase inhibitor zVAD-fmk had no effect (Fig. 2D). As expected, the *RIPK1* inhibitors Nec-1s and GSK2982772 had no effect on these *RIPK1*-deficient cells (Fig. 2D).

To investigate the molecular basis of the immune dysfunction responsible for the disease, we studied cytokine release in whole-blood assays. After stimulation with phytohemagglutinin

¹Department of Medicine, University of Cambridge, Cambridge CB2 0QQ, UK. ²University College London Genetics Institute, University College London, London, UK. ³Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK. ⁴Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, UK. ⁵Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. ⁶Paediatric Immunology Department, Birmingham Heartland Hospital, Birmingham, UK. ⁷Children's Hospital, King Fahad Medical City, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Kingdom of Saudi Arabia. ⁸Faculty of Applied Medical Sciences, Albaha University, Albaha, Kingdom of Saudi Arabia. ⁹Children's Hospital, University of King Saud for Health Sciences, King Fahad Medical City, Riyadh, Kingdom of Saudi Arabia. ¹⁰Division of Molecular Internal Medicine, Department of Internal Medicine II, University Hospital Würzburg, Würzburg, Germany. ¹¹University of Manchester, Royal Manchester Children's Hospital, Manchester, UK. ¹²Primary Immunodeficiency Group, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK.

*These authors contributed equally to this work. †Present address: Institute of Biomedical Sciences (IBMS), Shanxi University, Taiyuan, China.

†Corresponding author. Email: sn262@cam.ac.uk

(PHA), patients' blood cells produced markedly increased amounts of IL-1 β (Fig. 3A). This enhanced IL-1 β response to PHA normalized in patient P2 after hematopoietic stem cell transplantation (HSCT) (Fig. 3A). Also, after PHA stimulation, patients' blood produced reduced amounts of

IL-17 and interferon- γ (IFN- γ), whereas the production of TNF- α , IL-6, and IL-10 was similar to that of controls (fig. S11); this finding indicates dysregulated but not generally suppressed responses of stimulated T cells. After stimulation with lipopolysaccharide (LPS), whole blood of

healthy controls produced large amounts of IL-6 and IL-10 and showed strongly up-regulated production of TNF- α and IL-12 upon costimulation with LPS and IFN- γ (Fig. 3B). These responses were markedly reduced in the RIPK1-deficient patients but normalized in patient P2 after HSCT

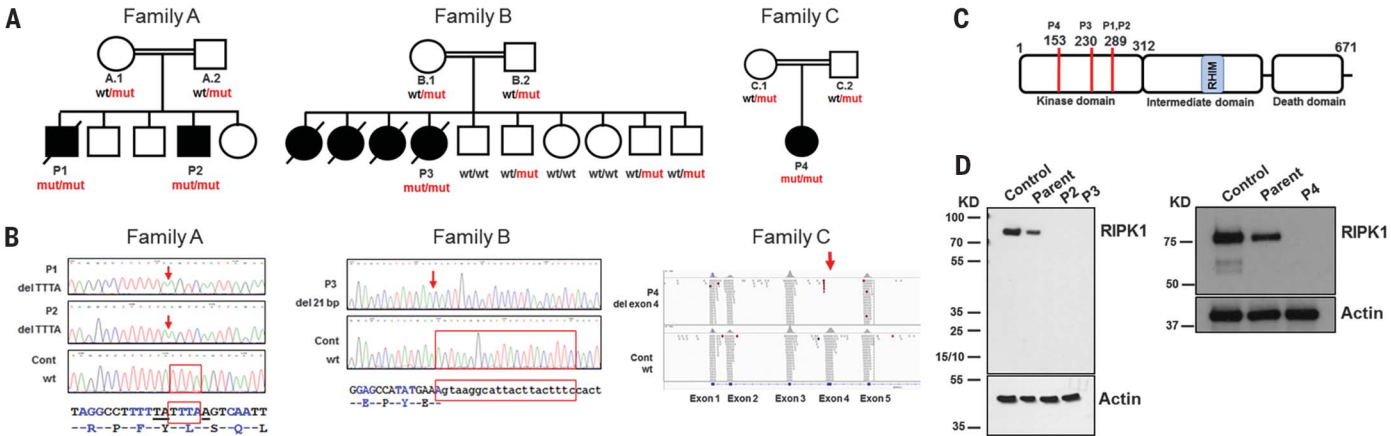


Fig. 1. Mutations of the *RIPK1* gene cause complete RIPK1 deficiency. (A) Three families with the *RIPK1* gene mutations (wt, wild-type allele; mut, mutant allele). Open circles and squares, unaffected; solid circles and squares, affected. (B) Patients' mutations. Locations of deletions are shown by red arrows; deleted nucleotides are shown by red frames. Right: Deletion of exon 4 in P4. (C) Domains of the

RIPK1 protein. RHIM denotes receptor-interacting protein (RIP) homotypic interaction motif. Codons affected by mutations are shown by red lines; codon numbers and corresponding patients are indicated above. (D) Western blot assays for detection of RIPK1 protein in fibroblasts (left) and T cell blasts (right); actin was used as a loading control.

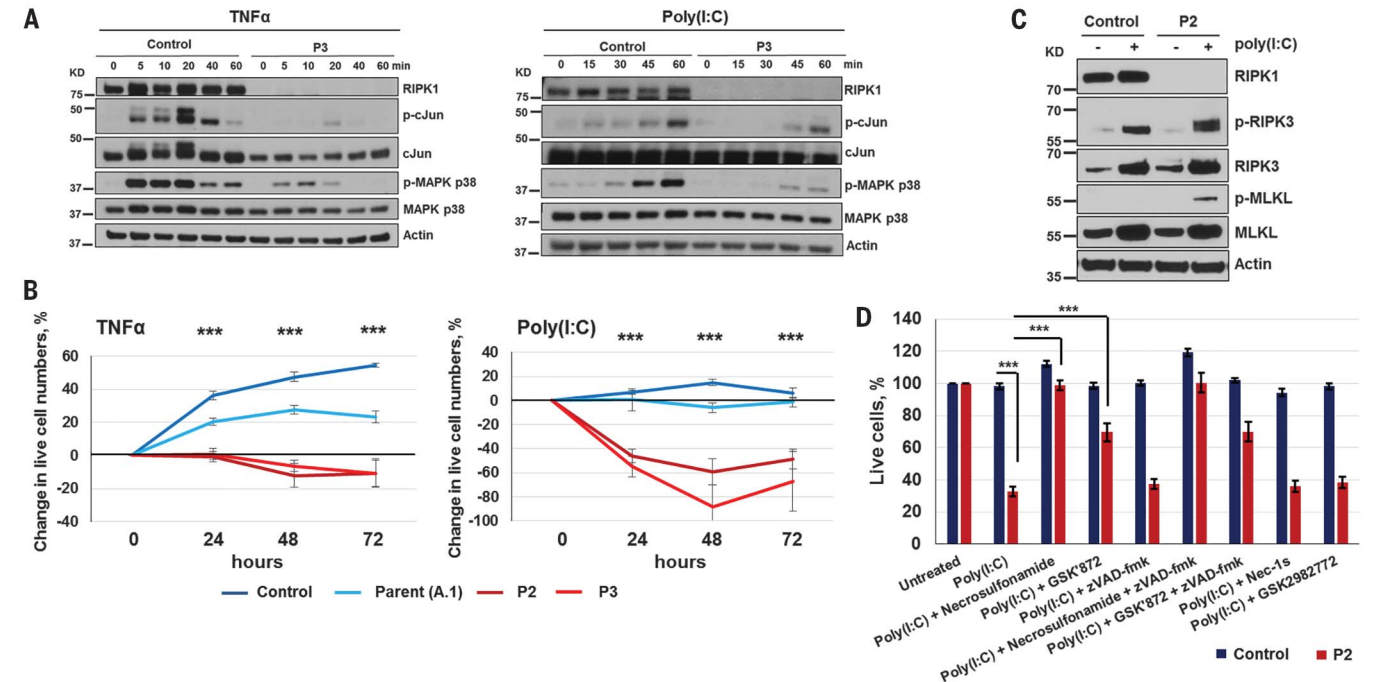


Fig. 2. RIPK1-deficient fibroblasts show impaired MAPK signaling and necroptosis. (A) Primary fibroblasts were stimulated with TNF- α (50 ng/ml) or poly(I:C) (100 μ g/ml) and protein extracts were subjected to immunoblotting ($N = 2$). (B) Viability of primary fibroblasts after stimulation with TNF- α (100 ng/ml) or poly(I:C) (20 μ g/ml). Experiments were repeated at least three times. Differences between stimulated and unstimulated cells are shown at each time point relative to the 0 time point. P values were calculated using two-tailed unpaired t tests, comparing combined data

from healthy subjects (parent and control) versus patients (P2 and P3); graphs show mean values \pm SEM. (C) Fibroblasts were stimulated with poly(I:C) (20 μ g/ml) for 24 hours and protein extracts were subjected to immunoblotting ($N = 1$). (D) Fibroblasts of P2 and controls (unrelated and parent A.1) were stimulated with poly(I:C) (20 μ g/ml) for 24 hours in the presence of the indicated compounds, and cell viability was measured ($N \geq 6$); graphs show mean values \pm SEM. P values were calculated using two-tailed unpaired t tests. *** $P < 0.001$.

(Fig. 3B). At the same time, we found normal production of IL-10, IL-6, and TNF- α in patients' whole-blood assays after stimulation with the TLR1/2 and TLR2/6 ligands Pam3CSK4 and Pam2CSK4 (fig. S12), consistent with the RIPK1-independent signaling downstream of these receptors (4). The production of IL-1 β in whole blood of the patients was within normal range after LPS stimulation and was slightly reduced

after LPS and IFN- γ costimulation (fig. S13). Similar to whole-blood assays, primary monocytes isolated from the blood of patient P4 and stimulated with LPS showed reduced production of IL-6, TNF- α , and IL-12 (Fig. 3C). However, the production of IL-1 β was increased (Fig. 3C).

We next analyzed in more detail the response of RIPK1-deficient human immune cells to LPS stimulation. We were unable to conduct further

studies of primary blood cells from the patients because P1 and P3 had died, P2 had undergone HSCT, and P4 was not available. We therefore used CRISPR-Cas9 technology to knock out RIPK1 in the human monocyte-like THP-1 cell line (2). After LPS stimulation, THP-1^{RIPK1-/-} cells secreted reduced amounts of IL-6 and IL-10 and released increased amounts of IL-1 β (Fig. 3D), resembling the cytokine response of patients' cells (Fig. 3C). The impaired cytokine production was preceded by the reduced phosphorylation of MAPK p38 (Fig. 3E), mirroring the defective MAPK p38 phosphorylation in the patients' primary fibroblasts (Fig. 2A). Activation of NF- κ B p65 and other branches of the MAPK pathway was not affected in the THP-1^{RIPK1-/-} cells (fig. S14). Differential activation of the MAPK and NF- κ B pathways after LPS stimulation was reported previously (14), and our results indicate that RIPK1 deficiency in THP-1 cells preferentially alters LPS-induced MAPK p38 activation.

After LPS stimulation, THP-1^{RIPK1-/-} cells showed increased phosphorylation of RIPK3 and MLKL and enhanced cell death (Fig. 3, F and G). No cleavage of caspase-8 or caspase-3 was found (fig. S15). Necrosulfonamide and GSK872, but not zVAD-fmk, reduced death of THP-1^{RIPK1-/-} cells (Fig. 3G), again pointing at necroptosis as the main death mechanism of these cells. Necroptosis of THP-1^{RIPK1-/-} cells was accompanied by the release of cleaved caspase-1 that we found in the supernatant together with IL-1 β (Fig. 3H), suggesting concurrent activation of inflammasome. Next, we compared mechanisms that mediate IL-1 β release from LPS-stimulated THP-1^{RIPK1-/-} cells with those that mediate IL-1 β release during pyroptosis of wild-type THP-1 cells stimulated by LPS and nigericin. We found that different mechanisms are at play. After treatment with necrosulfonamide, IL-1 β release during pyroptosis was only slightly reduced, whereas it was completely prevented in THP-1^{RIPK1-/-} cells; these findings suggested that IL-1 β release is secondary to necroptosis in these cells (fig. S16). Of note, treatment of THP-1^{RIPK1-/-} cells with necrosulfonamide did not restore reduced IL-6 secretion (fig. S17), indicating that it is reduced not because of enhanced necroptosis but is an earlier phenomenon. Taken together, these data show that human RIPK1-deficient cells have impaired pro-inflammatory signaling leading to dysregulated cytokine secretion and are prone to necroptosis, which, in myeloid cells, is accompanied by IL-1 β release.

RIPK1 has been extensively studied in mouse models (7–9, 15–18) and its inhibitors are considered for the treatment of acute and chronic organ injury, including stroke, myocardial infarction, and renal ischemia-reperfusion injury (19, 20), but the phenotype associated with RIPK1 deficiency in humans was unknown. While the *Ripk1* knockout mice displayed systemic inflammation and cell death in multiple tissues and died during the postnatal period (7–9), the clinical presentation of our RIPK1-deficient patients was less severe. Nonetheless, they suffered from immunodeficiency, gut inflammation, and

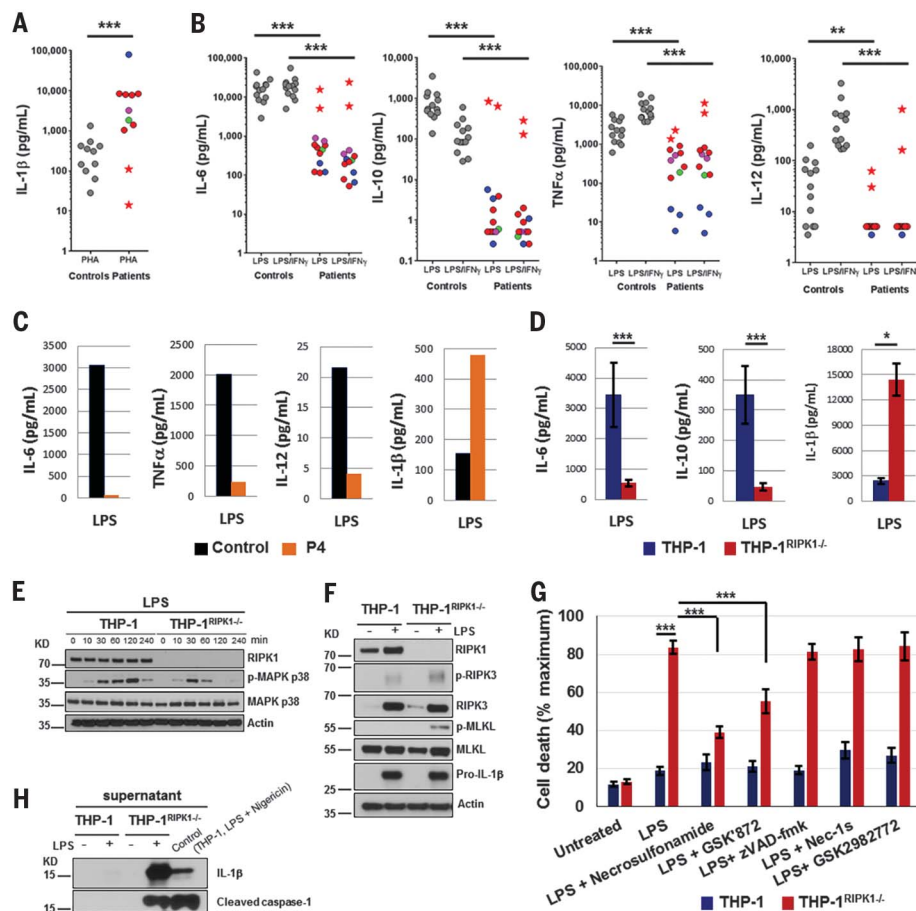


Fig. 3. RIPK1-deficient immune cells show dysregulated cytokine production, impaired MAPK signaling, and necroptosis. (A and B) Cytokines were measured in whole blood after 24 hours of stimulation with PHA (10 μ g/ml) (A) or LPS (1 μ g/ml) with or without IFN- γ (20,000 IU/ml) (B). Data were corrected for lymphocyte counts to account for lymphopenia. Controls are shown as gray circles, patients as colored circles (magenta, P1; blue, P3; green, P4; red, P2 before HSCT); red star denotes P2 after HSCT. *P* values were calculated using two-tailed Mann-Whitney tests, excluding the data of P2 after HSCT. (C) CD14⁺ monocytes were purified from PBMC of patient P4 (age 3 years) and a healthy adult (travel control), stimulated overnight with LPS (5 μ g/ml), and then cytokines were measured in supernatants. *N* = 1, two technical replicates; graphs show mean values. (D) THP-1 cells (2 wild-type and 7 THP-1^{RIPK1-/-} clones) were treated with phorbol 12-myristate 13-acetate (PMA; 50 ng/ml) for 3 days, rested for 1 day, and stimulated overnight with LPS (5 μ g/ml). Cytokines were measured in supernatants. The data were corrected for the estimated number of live cells and are means combining data for different clones \pm SEM. *P* values were calculated using two-tailed unpaired *t* tests. (E) THP-1 cells were treated with PMA as in (D) and stimulated with LPS (1 μ g/ml), and the extracted proteins were analyzed by immunoblotting (*N* = 2). (F) THP-1 cells were stimulated with LPS (5 μ g/ml) for 48 hours, and the extracted proteins were analyzed by immunoblotting (*N* = 2). (G) THP-1 cells were stimulated with LPS (5 μ g/ml) for 48 hours in the presence of indicated compounds, and cell death was measured using lactate dehydrogenase (LDH) release assay (*N* \geq 4). Data are means \pm SEM. *P* values were calculated using two-tailed unpaired *t* tests. (H) Supernatants of THP-1 cells studied in (F) were analyzed by immunoblotting (*N* = 1). **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

progressive polyarthritis. During infection, activation of the pattern-recognition receptors TLR3 and TLR4 and stimulation of TNFR1 by secreted TNF- α induces proinflammatory effects. Consistent with the established role of RIPK1 in the signal-transducing protein complexes assembled downstream of these receptors (4–7, 10, 11), we found impaired proinflammatory signaling in RIPK1-deficient cells and reduced production of multiple cytokines. These defects, as well as lymphopenia, likely explain susceptibility to infections in the patients.

Our data show that LPS stimulation of RIPK1-deficient monocytes resulted in increased necroptosis and IL-1 β release, similar to a previous observation in the RIPK1-deficient monocytes transdifferentiated from immortalized human B cells (21). Furthermore, we found high IL-1 β levels after PHA stimulation of the patients' whole blood. This may suggest that IL-1 β production by T cells, which was reported recently (22, 23), is also increased in the context of RIPK1 deficiency. Alternatively, dysregulated secretion of T cell factors after PHA stimulation may have augmented the IL-1 β production by patients' monocytes in whole blood. IL-1 β is a proinflammatory cytokine involved in the pathogenesis of arthritis and IBD (24, 25). Another cytokine dysregulated in our patients, IL-10, is essential for balancing immune response in the gut, and impaired signaling in the IL-10 pathway has been associated with IBD (26, 27). Therefore, it is likely that low secretion of IL-10 and increased IL-1 β production contributed to the pathogenesis of arthritis and IBD in our RIPK1-deficient patients. Accordingly, treatment with IL-1 inhibitors may be considered in RIPK1 deficiency, although none of our patients received such therapy.

Whereas patients P1, P2, and P3 developed severe IBD in the first months of life, P4 had no IBD signs up until the age of 4 years. Therefore, genetically determined RIPK1 deficiency was not the sole cause of IBD in these patients. Rather, it led to the dysregulated cytokine production, which set the immune system in a predisposition mode, while additional factors (e.g., distinct microbiomes) likely affected progression to IBD in these patients. This scenario resembles

adult-onset IBD, where immune predisposition is determined by multiple common genetic polymorphisms, whereas progression to clinical disease is driven by environmental factors (28).

The intestinal epithelium provides a physical barrier and participates in maintaining immune homeostasis in the gut. Mice genetically deficient in *Ripk1* in the intestinal epithelial cells (IECs) developed severe lethal intestinal pathology due to caspase-8-mediated apoptosis of IECs (9, 29). Histological examination of gastrointestinal biopsies from P1 and P2 showed only occasional cells with apoptotic morphology and cells positive for cleaved caspase-3 (fig. S18). Such cells were also present in biopsies from children with idiopathic IBD and histologically normal biopsies (fig. S18). Thus, in contrast to the mouse model, no extensive IEC apoptosis was found, which suggests that it is not a characteristic feature of RIPK1 deficiency in humans. Likewise, skin disorders were not typical in our patients, indicating that human RIPK1 has no special protective role in keratinocytes, in contrast to mice with epidermis-specific RIPK1 knockout that developed severe skin inflammation (9, 30). Given that HSCT in patient P2 resolved IBD and arthritis and reduced the frequency of infections (2), it is likely that dysfunction of the immune system, rather than dysfunction of other cell types, was critical for disease development.

Our findings indicate that RIPK1 has a more narrow function in humans than in mice, with the effects of RIPK1 deficiency being largely confined to the immune system. Accordingly, HSCT performed at a young age can be an effective treatment in RIPK1-deficient patients.

REFERENCES AND NOTES

1. A. Bousfiha et al., *J. Clin. Immunol.* **38**, 129–143 (2018).
2. See supplementary materials.
3. M. Lek et al., *Nature* **536**, 285–291 (2016).
4. D. Ofengeim, J. Yuan, *Nat. Rev. Mol. Cell Biol.* **14**, 727–736 (2013).
5. M. Pasparakis, P. Vandenabeele, *Nature* **517**, 311–320 (2015).
6. M. Najjar et al., *Immunity* **45**, 46–59 (2016).
7. M. A. Kelliher et al., *Immunity* **8**, 297–303 (1998).
8. J. A. Rickard et al., *Cell* **157**, 1175–1188 (2014).
9. M. Dannappel et al., *Nature* **513**, 90–94 (2014).
10. T. H. Lee et al., *Mol. Cell. Biol.* **23**, 8377–8385 (2003).
11. E. Meylan et al., *Nat. Immunol.* **5**, 503–507 (2004).
12. D. W. Zhang et al., *Science* **325**, 332–336 (2009).
13. L. Sun et al., *Cell* **148**, 213–227 (2012).
14. R. A. Gottschalk et al., *Cell Syst.* **2**, 378–390 (2016).
15. K. Newton et al., *Science* **343**, 1357–1360 (2014).
16. A. Polykratis et al., *J. Immunol.* **193**, 1539–1543 (2014).
17. W. J. Kaiser et al., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7753–7758 (2014).
18. J. E. Roderick et al., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14436–14441 (2014).
19. A. Degterev et al., *Nat. Chem. Biol.* **1**, 112–119 (2005).
20. H. Zhao et al., *Cell Death Dis.* **6**, e1975 (2015).
21. M. M. Gaidt et al., *Immunity* **44**, 833–846 (2016).
22. B. N. Martin et al., *Nat. Immunol.* **17**, 583–592 (2016).
23. G. Arbore et al., *Science* **352**, aad1210 (2016).
24. G. Schett, J. M. Dayer, B. Manger, *Nat. Rev. Rheumatol.* **12**, 14–24 (2016).
25. M. F. Neurath, *Nat. Rev. Immunol.* **14**, 329–342 (2014).
26. R. Kühn, J. Löhler, D. Rennick, K. Rajewsky, W. Müller, *Cell* **75**, 263–274 (1993).
27. E. O. Glocker et al., *N. Engl. J. Med.* **361**, 2033–2045 (2009).
28. B. Khor, A. Gardet, R. J. Xavier, *Nature* **474**, 307–317 (2011).
29. N. Takahashi et al., *Nature* **513**, 95–99 (2014).
30. J. Lin et al., *Nature* **540**, 124–128 (2016).

ACKNOWLEDGMENTS

We thank S. Prothero, who participated in the patients' treatment. **Funding:** Supported by the Wellcome Trust (095198/Z/10/Z), MRC (MR/M012328/1), and ERC Starting grant (260477) (S.N.); the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre (S.N., D.K., and R.D.); Deutsche Forschungsgemeinschaft (WA 1025/31-1) (H.W.); and the MRC/EPSRC Newcastle Molecular Pathology Node. **Author contributions:** D.C.-L., D.E., C.W., O.P., M.M., M.G., A.A., E.G., D.S., and H.W. performed cell experiments and analyzed the data; V.P. analyzed exome data; J.C. performed sequencing; M.A., P.D.A., M.S.A.I.Z., D.K., E.A.I., B.A., and S.H. looked after the patients and collected patients' data; C.M.B. analyzed biopsies; R.D. and L.C.-G. performed cytokine analyses; and S.N. planned the experiments, analyzed the data, and wrote the first draft of the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** Cells are available upon signing a material transfer agreement. Sequence data for the *RIPK1* gene region are available in Sequence Read Archive under accession number SRP136541. Whole exome data are available upon signing a data transfer agreement.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6404/810/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S18
Tables S1 and S2
References (31–35)

24 October 2017; accepted 28 June 2018
Published online 19 July 2018
10.1126/science.aar2641

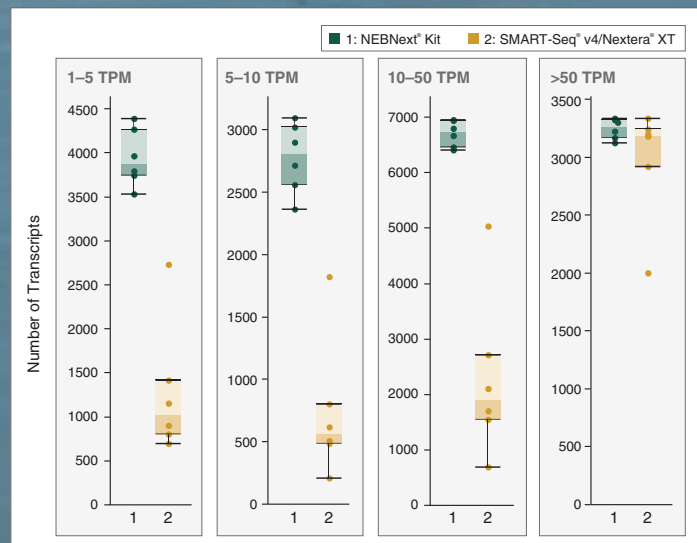
How low can you go?

NEBNext[®] Single Cell/ Low Input Library Prep Kit

With this unique streamlined method, high-quality, full-length transcript sequencing libraries are made directly from single cells, or from as little as 2 pg – 200 ng of total RNA.

- Detect more transcripts, including low-abundance transcripts
- Obtain uniform, full-length transcript coverage, regardless of input amount or sample type
- Save time with a streamlined workflow, minimal handling steps and hands-on time

Superior transcript detection with the NEBNext Single Cell/Low Input RNA Library Prep Kit



Sequencing libraries were generated from Jurkat single cells (6 replicates) using the NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq[®] v4 Ultra[®] Low Input RNA Kit for Sequencing plus the Nextera[®] XT DNA Library Prep Kit. Libraries were sequenced on an Illumina[®] NextSeq[®] 500. Each dot represents the number of transcripts identified at the given Transcripts Per Kilobase Million (TPM) range, and each box represents the median, first and third quartiles per replicate and method. Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 transcripts. Increased identification of low abundance transcripts is observed with the NEBNext libraries.

Visit **NEBNext.com**
to request your sample today.



2019 **MRS**® SPRING MEETING & EXHIBIT

April 22–26, 2019 | Phoenix, Arizona

CALL FOR PAPERS

Abstract Submission Opens
September 28, 2018

Abstract Submission Closes
October 31, 2018

Spring Meeting registrations include MRS Membership July 1, 2019 – June 30, 2020

BROADER IMPACT

- BI01 High Impact Practice—Increasing Ethnic and Gender Diversification in Engineering Education

CHARACTERIZATION, PROCESSING AND THEORY

- CP01 Advances in *In Situ* Experimentation Techniques Enabling Novel and Extreme Materials/Nanocomposite Design
- CP02 Design and *In Situ* TEM Characterization of Self-Assembling Colloidal Nanosystems
- CP03 Advances in *In Situ* Techniques for Diagnostics and Synthetic Design of Energy Materials
- CP04 Interfacial Science and Engineering—Mechanics, Thermodynamics, Kinetics and Chemistry
- CP05 Materials Evolution in Dry Friction—Microstructural, Chemical and Environmental Effects
- CP06 Smart Materials for Multifunctional Devices and Interfaces
- CP07 From Mechanical Metamaterials to Programmable Materials
- CP08 Additive Manufacturing of Metals
- CP09 Mathematical Aspects of Materials Science—Modeling, Analysis and Computations

ELECTRONICS AND PHOTONICS

Soft Organic and Biomolecular Electronics

- EP01 Liquid Crystalline Properties, Self-Assembly and Molecular Order in Organic Semiconductors
- EP02 Photonic Materials and Devices for Biointerfaces
- EP03 Materials Strategies and Device Fabrication for Biofriendly Electronics
- EP04 Soft and Stretchable Electronics—From Fundamentals to Applications
- EP05 Engineered Functional Multicellular Circuits, Devices and Systems
- EP06 Organic Electronics—Materials and Devices

Semiconductor Devices, Interconnects, Plasmonic and Thermoelectric Materials

- EP07 Next-Generation Interconnects—Materials, Processes and Integration
- EP08 Phase-Change Materials for Memories, Photonics, Neuromorphic and Emerging Application
- EP09 Devices and Materials to Extend the CMOS Roadmap for Logic and Memory Applications
- EP10 Heterovalent Integration of Semiconductors and Applications to Optical Devices
- EP11 Hybrid Materials and Devices for Enhanced Light-Matter Interactions
- EP12 Emerging Materials for Plasmonics, Metamaterials and Metasurfaces
- EP13 Thermoelectrics—Materials, Methods and Devices

ENERGY AND SUSTAINABILITY

Energy Storage

- ES01 Organic Materials in Electrochemical Energy Storage
- ES02 Next-Generation Intercalation Batteries
- ES03 Electrochemical Energy Materials Under Extreme Conditions
- ES04 Solid-State Electrochemical Energy Storage

Catalysis, Alternative Energy and Fuels

- ES05 Cooperative Catalysis for Energy and Environmental Applications
- ES06 Atomic-Level Understanding of Materials in Fuel Cells and Electrolyzers
- ES07 New Carbon for Energy—Materials, Chemistry and Applications
- ES08 Materials Challenges in Surfaces and Coatings for Solar Thermal Technologies
- ES10 Rational Designed Hierarchical Nanostructures for Photocatalytic System
- ES11 Advanced Low Temperature Water-Splitting for Renewable Hydrogen Production via Electrochemical and Photoelectrochemical Processes
- ES12 Redox-Active Oxides for Creating Renewable and Sustainable Energy Carriers

Water-Energy Materials and Sustainability

- ES09 Advanced Materials for the Water-Energy Nexus
- ES13 Materials Selection and Design—A Tool to Enable Sustainable Materials Development and a Reduced Materials Footprint
- ES14 Materials Circular Economy for Urban Sustainability

Photovoltaics and Energy Harvesting

- ES15 Fundamental Understanding of the Multifaceted Optoelectronic Properties of Halide Perovskites
- ES16 Perovskite Photovoltaics and Optoelectronics
- ES17 Perovskite-Based Light-Emission and Frontier Phenomena—Single Crystals, Thin Films and Nanocrystals
- ES18 Frontiers in Organic Photovoltaics
- ES19 Excitonic Materials and Quantum Dots for Energy Conversion
- ES20 Thin-Film Chalcogenide Semiconductor Photovoltaics
- ES21 Nanogenerators and Piezotronics

QUANTUM AND NANOMATERIALS

- QN01 2D Layered Materials Beyond Graphene—Theory, Discovery and Design
- QN02 Defects, Electronic and Magnetic Properties in Advanced 2D Materials Beyond Graphene
- QN03 2D Materials—Tunable Physical Properties, Heterostructures and Device Applications
- QN04 Nanoscale Heat Transport—Fundamentals
- QN05 Emerging Thermal Materials—From Nanoscale to Multiscale Thermal Transport, Energy Conversion, Storage and Thermal Management
- QN06 Emerging Materials for Quantum Information
- QN07 Emergent Phenomena in Oxide Quantum Materials
- QN08 Colloidal Nanoparticles—From Synthesis to Applications

SOFT MATERIALS AND BIOMATERIALS

- SM01 Materials for Biological and Medical Applications
- SM02 Progress in Supramolecular Nanotheranostics
- SM03 Growing Next-Generation Materials with Synthetic Biology
- SM04 Translational Materials in Medicine—Prosthetics, Sensors and Smart Scaffolds
- SM05 Supramolecular Biomaterials for Regenerative Medicine and Drug Delivery
- SM06 Nano- and Microgels
- SM07 Bioinspired Materials—From Basic Discovery to Biomimicry

www.mrs.org/spring2019

Meeting Chairs

Yuping Bao The University of Alabama
Bruce Dunn University of California, Los Angeles
Subodh Mhaisalkar Nanyang Technological University
Ruth Schwaiger Karlsruhe Institute of Technology—
 Institute for Applied Materials
Subhash L. Shinde University of Notre Dame

Don't Miss These Future MRS Meetings!

2019 MRS Fall Meeting & Exhibit
 December 1–6, 2019, Boston, Massachusetts

2020 MRS Spring Meeting & Exhibit
 April 13–17, 2020, Phoenix, Arizona



MATERIALS RESEARCH SOCIETY®
Advancing materials. Improving the quality of life.

506 Keystone Drive • Warrendale, PA 15086-7573
 Tel 724.779.3003 • Fax 724.779.8313 • info@mrs.org • www.mrs.org

DOES YOUR LAB COMBINE COMPUTATIONAL AND EXPERIMENTAL STRATEGIES TO INVESTIGATE SIGNALING NETWORKS?

Michal Hershfinkel, Elias Aizenman, Glen Andrews, and Israel Sekler *Sci. Signal.* 3, mr2 (6 July 2010) Image: Omer Barad, Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.

ScienceSignaling | AAAS
CELL SIGNALING IN PHYSIOLOGY AND DISEASE

Find out more about the scope of the journal and submit your research today. ScienceSignaling.org

AAAS *Travels*



Christmas in Guatemala December 18-27, 2018

Visit the Highlands of Guatemala during the festive season. Explore the historic cities of Chichicastenango and Antigua, and village life around the jewel of the Americas, Lake Atitlan, which is surrounded by spectacular volcanoes. Explore the finest Mayan ruins in Central America—Tikal, Quirigua and Copan with fantastic archaeological sites, spider monkeys, macaws and toucans! \$3,995 per person + air.

For a detailed brochure, call (800) 252-4910
All prices are per person twin share + air



BETCHART EXPEDITIONS Inc.
17050 Montebello Rd, Cupertino, CA 95014
Email: Info@betchartexpeditions.com
www.betchartexpeditions.com

FORCEFORSCIENCE.ORG

STAND TOGETHER Be a Force for Science

GET THE FACTS

Understand the science behind the issues that matter.

FOLLOW AAAS ADVOCACY

Champion public discussion and evidence-based policy.

TAKE ACTION

Learn ways you can become an advocate and stand up for science.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE



**Lundbeck Foundation
calls for nominations for**

THE BRAIN PRIZE 2019

Nominations must be received by 1 September 2018

The prize of € 1 million is awarded to researchers for outstanding contributions to neuroscience – from basic to clinical.

The Brain Prize is open to researchers around the world.

Nominations will be reviewed by the Selection Committee:

Anders Björklund, Sweden, Chairman
Joseph Coyle, USA
Geoffrey Donnan, Australia
Catherine Dulac, USA
Story Landis, USA
Richard Morris, United Kingdom, Vice-Chairman
Mu-ming Poo, China
Philip Scheltens, The Netherlands
Irene Tracey, United Kingdom

Lundbeck Foundation encourages diversity in nominees.

For the nomination form and details of the nomination procedure, please visit www.thebrainprize.org.



Lundbeck Foundation



Seeing is believing: R&D applications of computer vision

Once a technology that only existed in the realm of science fiction, artificial intelligence (AI) has moved out of research labs and into everyday products. Powering much of the progress in AI is computer vision, a technology concerned with building software that can gather and process visual information. Recognizing individuals in photos, reading X-rays, and making robotic systems on factory floors smarter are just a few of the applications of this promising new field.

Most people take vision for granted, not thinking twice about the massive amount of computation our brains carry out just to pick up a fork or catch a ball. It is only in recent years that computers have become quick enough, powerful enough, and small enough to enable practical applications for computer vision. The most cutting-edge computer vision uses deep learning, an area of AI inspired by the human brain. Deep learning algorithms use artificial neural networks (ANNs)—layers of connected nodes that analyze and pass information along to each other, in a way similar to the process by which neurons communicate.

Show a neural network a selfie, and one layer of neurons will recognize broad strokes like the outline of a face. Another will focus on the space between features on the face, like the distance from the eyes to the mouth. More still will look at the shape of an ear. From there, the algorithm can figure out if it's a photo of a person or even who that person is.

"In most computer vision tasks, neural networks generate the best algorithms, hands down," says Jiaya Jia, Distinguished Scientist in the YouTu Lab at Tencent, a global leader in Internet-related services and products, entertainment, and AI located in Shenzhen, China.

The engineers behind computer vision try to mimic what ANNs do within the human visual system. But machines have an advantage, in that they don't have to rely on the limited light spectrum that people use, and can employ sensors to fill out how they see the world.

"Computer vision works better than human vision in many tasks, such as facial recognition and image classification;

while in others requiring reasoning, there is still a long way to go," says Jia. "It is easy for humans to know how objects relate to each other. We can make a story from just one picture—computers are still far from this level of understanding and imagination."

As this technology is developed, new realms of discovery are possible. Computer vision and AI are still in their early days, and there is much more to come. Advances in computer vision could supercharge AI and give us all the things science fiction stories promise—self-driving cars, robotic butlers, and even long-distance space travel.

Learning by doing: How AI works

Computer vision, like most AI systems, needs a trove of data from which to learn. Researchers go through and carefully label this data with features they want the AI to interpret. For computer vision tasks, researchers collect hundreds—or more likely thousands—of images to be analyzed. That labeled data becomes the set of examples on which the AI is trained to categorize or find patterns. To test its learning efficacy, researchers challenge the AI to correctly categorize new, unlabeled images.

Besides the human labor that goes into collecting, annotating, and preparing massive datasets, a significant hurdle is the large amount of computing power required to run the training algorithm. Cheap, online services let researchers train their algorithms in the cloud instead of investing thousands of dollars on powerful computers. However, it can still take hours or days to get results.



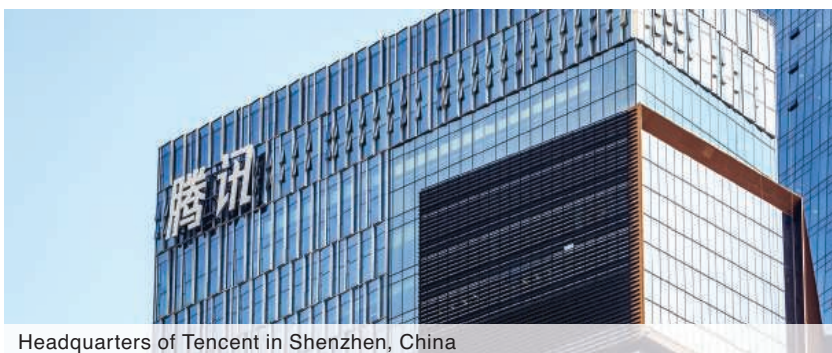
Smile for the camera: Image and video recognition

One of the most important advancements in ANNs came from ImageNet, a database of more than 14 million labeled images collected and released in 2009. The ImageNet Challenge, a contest asking participants to create an algorithm that could categorize photos as well as humans, didn't produce winners until 2012, when a team using a deep learning algorithm achieved significantly better results than any previous attempts.

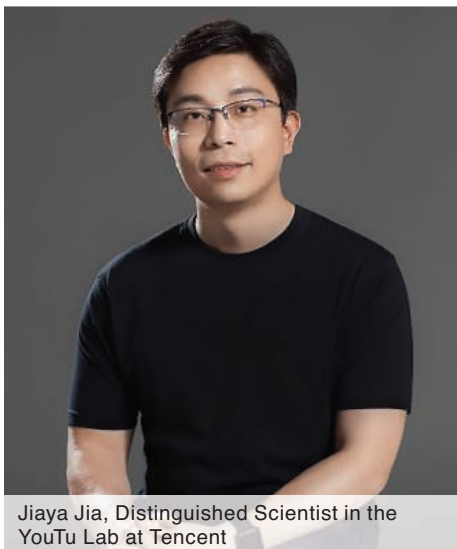
Today, some of the most common ways people interact with computer vision—including automated tagging of images and face recognition in photographs—originated with the technology that bested ImageNet. These applications can be helpful with certain tasks, like doing visual searches when online shopping and automatically tagging social media photos.

In addition to image recognition, photo editing has also gotten a boost from this new technology. There are now algorithms for image segmentation, a component of computer vision that helps break down different parts of a picture; for example, by differentiating the background from people posing in the foreground. Users can edit photos rapidly, making them look like they've been professionally retouched.

Visual recognition capabilities are also being applied to video footage. Computer vision algorithms can assess a camera's video feed and flag critical parts so that a human doesn't have to comb through hours of footage. Understanding the emotions of individuals in footage is still a work in progress but has a few early adopters. Some Chinese schools—like Hangzhou Number 11 High School in the country's eastern Zhejiang Province—are experimenting with cameras called “smart eyes” that track student engagement in class.



Headquarters of Tencent in Shenzhen, China



Jiaya Jia, Distinguished Scientist in the YouTu Lab at Tencent

20/20 vision for robots

Going from a 2D world of images to 3D data presents a different set of challenges as well as a host of new opportunities. One area showing promise involves endowing robots with computer vision.

Robots began hitting manufacturing floors in the 1960s. These devices could lift heavy weights, do repetitive tasks, and carry out precise measuring for hours at a time, easily besting humans.

“It's always been about building things that can

do something humans couldn't,” says Danica Kragic, a robotics professor at the KTH Royal Institute of Technology in Stockholm, Sweden.

With 40% of our brain dedicated to processing visual information, it's important to understand to what extent visual information is needed if we are creating machines that mimic us and share our world, says Kragic. “We humans use visual feedback in just about everything we do,” she explains.

Machines able to process visual information could do more sophisticated work in factories and might even make their way into our homes. Some skills, like picking up squishy items that change shape with pressure, are still out of reach for robots. That's because humans gain more than just visual information from sight; we also pick up cues about the physical property of an object and the physics needed to interact with it. Machines need to be able to gather that type of information to move effortlessly through the physical world as humans do.

“Among the five senses, vision is one of the most important, since it gives us the ability to comprehend a complicated world,” says Jia. “Likewise, computer vision commits to the task of letting computers see and directly interact with the environment as we do.”

Empowering robots with sensors that could better map out the world is the next iteration of the technology that could have them doing things they are unable to do today. *cont.>>*



Making drones autonomous means they will be able to reduce costs, bringing the technology into the hands of more people and companies around the world.

Coming to a road near you: Self-driving cars

A holistic understanding of the world is also crucial for autonomous vehicles (AVs), one of the best-funded and most talked about areas in AI development. Besides cameras, most driverless cars use lidar, radar, GPS, and perception algorithms to navigate.

"A lot of the algorithms we use draw from computer vision, but now it's more than camera data," says Raquel Urtasun, an associate professor at the University of Toronto and head of Uber's Advanced Technologies Group, also in Toronto, Canada. "We want to give our cars more than just our eyes."

Companies like Uber hope to get AVs on the road and carrying passengers by 2020. These cars will only drive preselected routes or require people in the driver's seat ready to take over in case any issues come up. The ultimate hope is for this technology to become truly autonomous so that riders can do other activities besides watching the road.

To get there, advances in both hardware and software are needed, Urtasun says. On the hardware side, lidar can cost tens of thousands of dollars, making it too expensive for wide-scale deployment. For AV software development, engineers need to find a way to make AI better able to "generalize," or to differentiate between different objects. For example, if a human driver sees something unexpected in the road like a downed power line, they understand that they need to go around it. But if a self-driving car approaches something it hadn't experienced in training, it might not react safely because it can't tell it apart from other objects.

Urtasun says she is hopeful her work to improve sensors and training algorithms won't be in vain, even if AVs are not quite ready for prime time. Fortunately, she says, "You can tackle many other problems with the technology." Improved lidar could make mapping and land surveys more accurate, and even nonautonomous cars equipped with sensors could help reduce traffic.

Special delivery: Drones

Cars aren't the only thing that researchers want to be able to move autonomously: Drones are also being trained to fly on their own. Some of the same kinds of problems that slow AV developers are faced by drone researchers. It's difficult and expensive to get quality training data; different flying styles create new scenarios that drones need to be trained on; and

regulations make it hard to test in some areas. Even after a drone is trained, flying it can still be very tricky.

"Anyone who has tried to steer a drone knows it's not so easy," says Tinne Tuytelaars, a professor at KU Leuven, a research university in Belgium. However, unlike errors with self-driving cars, mistakes with drones are less costly. "If a drone crashes," Tuytelaars shrugs, "so be it."

Drones are already being put to work in projects such as disaster relief and pipeline inspection. They could one day make deliveries and provide rides to passengers. Already, companies like Amazon and Boeing are testing drones that could eventually be able to drop packages off like mail carriers do today.

In some circumstances, multiple drones could occupy a single airspace and could better coordinate flight with each other than human pilots do. Making drones autonomous means they will be able to reduce costs, bringing the technology into the hands of more people and companies around the world.

Dr. Robot

After vehicular transportation, the medical field could see the most significant impact from computer vision. Already, AI algorithms can spot complications in medical images, such as bone fractures and pneumonia, more accurately than a radiologist can.

"The explosion of big data, specifically in the medical field, means we have so much to work with," says Beatriz Remeseiro, an assistant professor at the Department of Computer Science at the University of Oviedo in Spain. "We're using the data to solve more complicated problems than ever before."

Last year, Google announced that it had developed image recognition algorithms to detect signs of diabetic retinopathy, an eye condition that causes blindness if left untreated. The algorithms worked as well as human experts at spotting small aneurysms in photographs of patient's retinas, which indicate the early stages of the disease.

In 2017, Tencent also developed a system that can screen gastroscopy pictures for esophageal cancer, with a 90% accuracy rate in diagnosing preliminary cases of the disease. The software, called the AI Medical Innovation System (AIMIS), is already in use at more than 100 hospitals across China and could one day be applied to help diagnose diabetic retinopathy, lung nodules, cervical cancer, and breast cancer.

Other AI-powered tools from other companies, such as InferVision, in Beijing, China, are being used to spot strokes faster, giving patients a better chance of surviving. The U.S. Food and Drug Administration recently announced it was going to streamline its process to help AI products get approval more quickly.

Instead of replacing doctors, however, these tools from other companies are meant to work as advisers rather than practitioners. But computer vision could speed up this work and allow regions with fewer doctors to receive more care. These innovations are also being used to minimize invasive or dangerous procedures. For example, computerized tomography (CT) scans hold more information than X-rays but expose a patient to more radiation. Using an X-ray, AI can now approximate what a CT scan would show a doctor.

"Medical imaging is an area where computer vision could really have an impact, by providing additional information," says Ramin Zabih, a professor of computer science at Cornell Tech and a research scientist at Google Research, in New York City, New York. "And the history of medicine has shown that when you can provide more data to practitioners, it can mean better outcomes for patients."

Near and far: Edge devices and spacecraft

As impressive as all this work seems, the future could be even brighter for computer vision. One of the biggest changes that could soon impact the industry is the edge device, defined as any hardware that "controls data flow at the boundary between two networks." Most AI processing needs to be done on big, remote cloud servers, because running these algorithms is so computationally intensive. Edge devices, on the other hand, are built with enough processing power to work locally. And with companies like Nvidia and Facebook building chips specifically to run AI, edge devices are becoming more commonplace. This allows for faster, more secure data processing, and opens the possibility of more custom training of AI with a user's own data and increased customization of results for individuals.

"It's going to drive a greater amount of innovation," says Remi El-Ouazzane, former CEO of Movidius—a California-based company that designs specialized low-power processor chips for computer vision—and current chief operating officer of the AI products group at Intel, another California tech company that produces semiconductor chips and microprocessors.

El-Ouazzane says billions of devices could be infused with AI and work on the edge, from smart home devices and surveillance cameras to autonomous cars. This could create technology capable of finding missing people by scanning images of crowds, for instance, or able to alert a parent when their child hasn't brushed their teeth before bed.

"It's not an 'if,' it's a 'when,'" he says.

Looking beyond the microworld of edge devices, astronomers—who work with massive datasets gathered from the expanses of space—are especially interested in computer vision. For example, a contest on Kaggle, an online platform for predictive modeling and analytics competitions, had researchers look at astronomical images to discover more about the dark matter that dominates our universe; they relied on deep learning and computer vision techniques.

Moreover, an entire research incubator is dedicated to using AI to further our understanding of space.

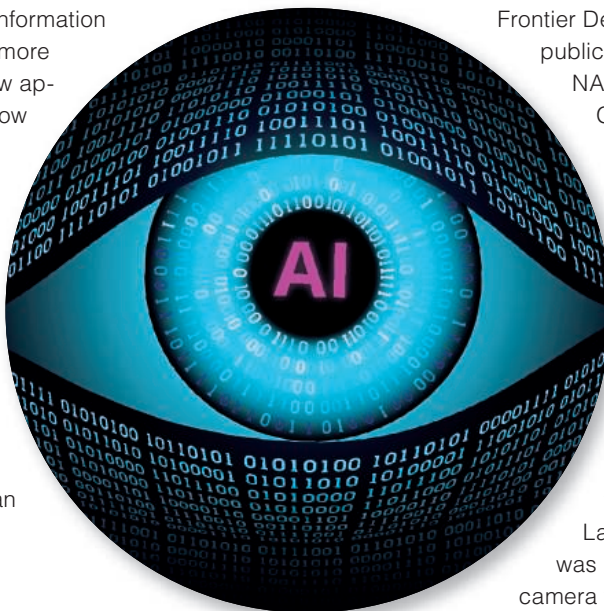
Frontier Development Lab (FDL) is a public-private partnership between NASA and companies like Intel AI, Google Cloud, Lockheed, and IBM. FDL brings astronomers and computer scientists together for eight weeks in Silicon Valley, California, to tackle problems such as understanding solar flares, mapping the moon, and finding asteroids. According to James Parr, one of the founders of FDL, the program couldn't succeed without computer vision. In fact, NASA's Jet Propulsion Laboratory, in Pasadena, California, was essential in the invention of the camera technology that powers most of the computer vision software currently in operation.

"There's a symbiosis between computer vision and the space program," Parr says. "It's a story the space industry doesn't talk about enough."

Space exploration will be impacted similarly, with AI being essential for travel to Mars and beyond. The lag in communication between space travelers and earth command centers means systems will have to be capable of autonomous decision-making, with visual data driving many of those decisions.

"As we expand outwards, we will need robots and autonomous systems that will prepare and assist astronauts, help build structures, and locate and extract resources," Parr says. "It's an exciting time for discovery and exploration."

This exhilarating time of discovery extends to finding out how computer vision will change life on earth. There will likely be new technological hurdles to overcome as AI expands into even more areas. But when we look back on the history of AI, giving computers the gift of sight may prove to be the most important advancement of all. Machines that can see will surely take us into a brighter future.



Tencent YouTu Lab



Exceptional scientists wanted

Present your work to the world

Are you a representative of the upcoming generation of thought leaders in your field? Together we look forward to your application for the new Sartorius & Science Prize for Regenerative Medicine & Cell Therapy.

Apply now!

www.passionforscience.com/prize



The Sartorius & Science
Prize for Regenerative
Medicine & Cell Therapy

Awarded by



sartorius

Science



Anaerobic Chambers

Coy Laboratory Products' Anaerobic Chambers provide a strict anaerobic atmosphere of 0–5 parts per million, using a palladium catalyst and a nonflammable hydrogen gas mix of 5%. The heavy-duty vacuum airlock allows sample

transfer without changes to the internal atmosphere. The airlock is also programmable, enabling researchers to tailor an interchange sequence ideal for their lab experiments. The unit's flexible nature allows the user to reach deeper and higher inside it, and its padded base and larger glove ports offer superior ergonomics compared to those of a rigid glove box. The vinyl expands and contracts to compensate for changes in internal pressure, without expelling the expensive gas mix to room atmosphere, saving you up to 60% annually on your operational gas cost. Quick Change Cuffs are available on any new or existing vinyl anaerobic chamber—change your vinyl anaerobic chamber gloves in less than 30 seconds!

Coy Laboratory Products

For info: 734-475-2200
coylab.com

Customized Resins

Production of agarose resins requires precise control of many parameters, and the product range is based on the most common research and production requirements. However, we recognize that there are many specialized applications that are not currently served, and many separations that could be improved by a more appropriate bead, so we now offer customized resins to extend the range. Parameters that can be specified include size of beads, degree of crosslinking, concentration of agarose, type of activation, spacer arm length, and density of active groups. Agarose Bead Technologies Customized Resins are now available to cover all different types of orders, enabling you to choose specifications for your resin at an affordable price. In order to produce exactly what you require, we will work closely with you and share technical details as appropriate.

Agarose Bead Technologies

For info: 305-262-0177
www.abtbeads.com

ELISA Kits

Bioaim Scientific is an ELISA company dedicated to developing high-quality immunoassay kits and antibodies for both industrial and academic scientific research. We have more than 800 ELISA kits, covering human, mouse, rat, bovine, canine, equine, and porcine species. All kits are quality guaranteed and risk-free. We offer sandwich-based EasyTest ELISA kits (which require only one wash, decreasing variations and significantly saving time, as compared to standard ELISA), sandwich-based normal ELISA kits, and competitive ELISA (enzyme immunoassay) kits. In addition to providing testing services for all assay kits, Bioaim also offers a full complement of biomarker discovery services, diagnostics kit development, and assay development services.

Bioaim Scientific

For info: 416-286-6868
bioaimscientific.com

Cell-Free Protein Expression

myTXTL is a fast, easy-to-use solution for protein expression *in vitro*. Gene transcription (TX) and translation (TL) is executed in a single reaction tube by a highly efficient cell-free system that utilizes the endogenous TXTL machinery from *Escherichia coli*. This all-in-one solution offers convenient, one-step gene expression from a simple nucleotide template for various applications in synthetic biology and biomanufacturing. The system employs endogenous core RNA polymerase and primary sigma factor 70 (σ^{70}) present in the *E. coli* cytoplasm. The technology has been well characterized and proven useful for applications including high-yield protein synthesis, prototyping of biomolecular networks, bacteriophage production, and high-throughput protein expression analysis. myTXTL kits contain ready-to-use Master Mix comprised of *E. coli* cell extract, energy buffer, and amino acids mix optimized for coupled TXTL.

Arbor Biosciences

For info: 734-998-0751
www.arborbiosci.com

N-Glycan Assay

The N-Glycan Assay allows for in-depth profiling of immunoglobulin G glycosylation. This physicochemical assay, developed by Sartorius Stedim Biotech's subsidiary, BioOutsource, and offered in combination with its unique Antibody Dependent Cell Cytotoxicity (ADCC) assay platform, combines data on molecular structure with a key therapeutic mechanism of action. This combination will generate comprehensive, highly accurate biosimilar comparability data, driving shorter drug-development timelines. BioOutsource's new assay enzymatically removes glycans from the antibody, functionally derivatizes them, and separates the glycans by UHPLC. The glycans are then identified by online electrospray ionization mass spectrometry, enabling confident assignment and precise quantification of their different structures. The N-Glycan Assay is one of a range of orthogonal methods, including analysis by CD16a (FcγRIIIa) and surface plasmon resonance binding, as well as ADCC functional assays that BioOutsource offers to provide a detailed testing package for generating evidence suitable for regulatory submissions.

Sartorius Stedim BioOutsource

For info: +44-(0)-141-946-4222
www.biooutsource.com

Kinase Inhibition Assay

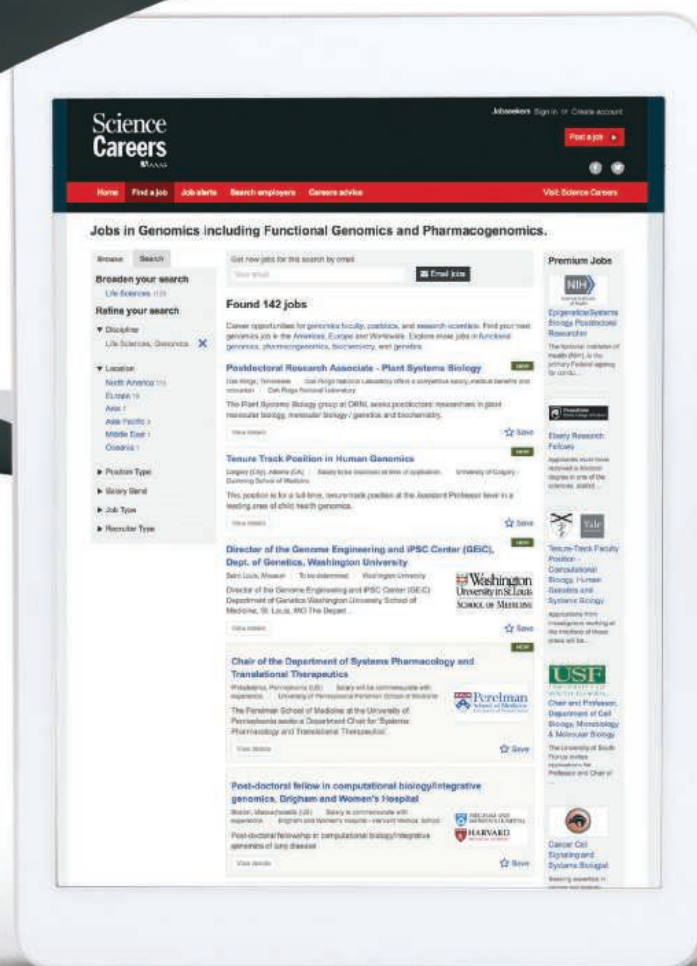
BÜHLMANN's Kinase Inhibition Assay offers a quick, precise method to evaluate inhibitors of immune-cell signaling pathways, using inhibition of basophil degranulation as a surrogate readout cell-based system by means of flow cytometry. PI3K, BTK, and SYK signaling pathways involved in basophil activation have been implicated in allergic responses, autoimmune disorders, and cancer. Inhibitors of these pathways are a large focus for pharmaceutical companies, with several small-molecule inhibitors currently at different stages of development. Alterations in the expression of cell-surface basophil activation markers such as CD63 can be measured by flow cytometry. These surrogate markers allow for high-throughput assessment of compound efficacy, resulting in faster failure of ineffective compounds.

BÜHLMANN Diagnostics

For info: 844-300-9799
kinaseinhibitionassays.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.



Step up your job search with *Science Careers*

- Access thousands of job postings
- Sign up for job alerts
- Explore career development tools and resources



Search jobs on **ScienceCareers.org** today

ARTIFICIAL INTELLIGENCE ■ AUTONOMOUS VEHICLES ■ BIOMATERIALS IN ROBOTICS ■ HUMANOIDS ■ LAND & UNDERSEA ROBOTS ■ MEDICAL & SURGICAL ROBOTS
MICRO/NANO ROBOTS ■ ROBOT ENGINEERING ■ ROBOTS IN EDUCATION ■ SPACE ROBOTS ■ THEORETICAL ADVANCES WITH POSSIBLE APPLICATIONS

Transforming the **Future** of **Robotics**



As a multidisciplinary online-only journal, *Science Robotics* publishes original, peer-reviewed, research articles that advance the field of robotics. The journal provides a central forum for communication of new ideas, general principles, and original developments in research and applications of robotics for all environments.

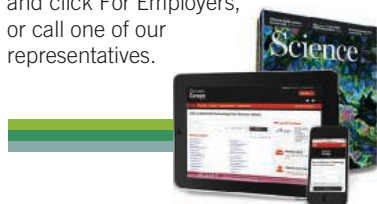
Learn more at: [ScienceRobotics.org](https://www.sciencerobotics.org)

**Science
Robotics**
AAAS

Science Careers

SCIENCE CAREERS ADVERTISING

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.



AMERICAS

+1 202 326-6577
+1 202 326-6578
advertise@sciencecareers.org

EUROPE, INDIA, AUSTRALIA, NEW ZEALAND, REST OF WORLD

+44 (0) 1223 326527
advertise@sciencecareers.org

CHINA, KOREA, SINGAPORE, TAIWAN, THAILAND

+86 131 4114 0012
advertise@sciencecareers.org

JAPAN

+81 3-6459-4174
advertise@sciencecareers.org

CUSTOMER SERVICE

AMERICAS

+1 202 326-6577

REST OF WORLD

+44 (0) 1223 326528

advertise@sciencecareers.org

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. *Science* reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. *Science* encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

ScienceCareers

FROM THE JOURNAL SCIENCE

ScienceCareers.org

myIDP: A career plan customized for you, by you.



There's only one *Science*.



Recommended by leading professional societies and the NIH

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.
- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- A certificate of completion for users that finish myIDP and more.

Start planning today!

myIDP.sciencecareers.org

— **Science Careers** In partnership with: —



POSITIONS OPEN

Assistant Professor, Mechanical Engineering

The Department of Mechanical Engineering at University of Connecticut at Storrs is seeking applicants to conduct research at the interface of biomedicine, materials, and nano/micro technology in an NIH-funded project to create a biodegradable, highly-efficient piezoelectric Poly-L-Lactid acid simulator, and integrate it with a biological chondrocyte-seeded cartilage graft, forming a bionic cartilage tissue. Teach undergrad/grad courses in this line of expertise. Experience developing nanoscale biointegrated materials, devices, and platform technology for applications in vaccine/drug delivery and medical implants, and creating nanoscale electromechanical materials, devices, and their interfaces with biological cells and tissues for harvesting, sensing, and engineering cellular mechanics, required. Salary range: \$80-\$100k. Apply to: **Attn Dr. Horea Ilies, Department of Mechanical Engineering, University of Connecticut, 191 Auditorium Rd. U-3139, Storrs, CT, 06269.**

Advance your career with expert advice from *Science Careers*.



Download Free Career Advice Booklets!

ScienceCareers.org/booklets



Science Careers

FROM THE JOURNAL SCIENCE

Search more jobs online

Access hundreds of job postings on ScienceCareers.org.

Expand your search today.



Career Feature:

Artificial Intelligence

Issue date: November 30

Book ad by November 15

Ads accepted until November 21 if space allows



129,562

subscribers in print
every week

503,472

monthly unique browsers
on ScienceCareers.org

56 %

of our weekly readers
are Ph.D.s

To book your ad:
advertise@sciencecareers.org

The Americas

+ 202 326 6577

Europe

+44 (0) 1223 326527

Japan

+81 3 6459 4174

China/Korea/Singapore/ Taiwan

+86 131 4114 0012

Produced by the Science/AAAS
Custom Publishing Office.

SCIENCECAREERS.ORG

Artificial Intelligence (AI) is impacting science in new and exciting ways as scientists are using it to better understand society to find solutions to problems across diverse disciplines. This feature will give an overview of AI, and explore the hotspots/centers of excellence and applications for AI. Typical career paths for those working in AI will be explored as well as the opportunities that exist for careers in AI.

Your organization can brand itself as a leader in AI by raising your visibility alongside relevant content while attracting potential candidates. Contact us for further details.

What makes Science the best choice for recruiting?

- Read and respected by 400,000 readers around the globe
- Your ad dollars support AAAS and its programs, which strengthens the global scientific community.

Why choose this AI Feature for your advertisement?

- Relevant ads lead off the career section with a special "AI" banner.

Expand your exposure by posting your print ad online:

- Link on the job board homepage directly to AI jobs
- Dedicated landing page for AI positions.



ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

FOR RECRUITMENT IN SCIENCE, THERE'S ONLY ONE SCIENCE.



GOVERNMENT OF INDIA

Ministry of Human Resource Development (Department of Higher Education) Technical Section - I

Appointment of Director IIT Kharagpur, IIT Guwahati and IIT Bombay

Applications are invited for appointment to the post of Director of Indian Institute of Technology (IIT) at Kharagpur, Guwahati and Bombay. The Director of an IIT is the academic and administrative head of the Institution. He/she is expected to have a minimum of 5 years' administrative experience and leadership qualities to head an Institute of National importance. The candidate/person should be a Ph.D. with first class or equivalent at the preceding degree, preferably in a branch of Engineering. In exceptional cases candidates with Science, Mathematics or Management degrees may be considered. He/she should have an outstanding academic record throughout and a minimum of 10 years teaching experience as a Professor in a reputed Engineering or Technology Institute or University and should have guided Ph.D students. The applicant should preferably be less than 60 years of age on the last date of receipt of the applications. The post carries a fixed pay of Rs. 2,25,000/- (Revised) per month, with allowances as per rules.

2. Interested individuals may apply giving their detailed resume in the prescribed format clearly bringing out research, teaching, industry-academia collaborations and administrative achievements, alongwith a two-page justification in support of their candidature, a two-page vision statement for the institution and contact details of at least two distinguished individuals well acquainted with their work. The application typed in the prescribed format along with enclosures may be sent by Registered/ Speed Post to **The Under Secretary (TS.1), Department of Higher Education, Ministry of Human Resource Development, Room No. 428 "C" Wing, Shastri Bhawan, New Delhi -110 001** so as to reach the Ministry on or before **15th September, 2018**. The detailed advertisement and the format of application is available on the website (www.mhrd.gov.in)

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

Follow us for jobs,
career advice & more!



@ScienceCareers



/ScienceCareers



Science Careers

ScienceCareers.org

Yale University
School of Medicine

**FACULTY POSITION AT THE ASSISTANT
PROFESSOR LEVEL**

**DEPARTMENT OF CELLULAR AND
MOLECULAR PHYSIOLOGY**

The Department of Cellular and Molecular Physiology is conducting a search for new faculty members at the assistant professor level.

The search seeks candidates whose research connects the properties of molecules to the properties of physiological systems.

Excellent opportunities are available for collaborative research, as well as for graduate and medical student teaching. Candidates must hold a Ph.D., M.D., or equivalent degree. Applicants should include a curriculum vitae, a statement of research interests and goals, and should arrange to have three letters of reference sent. Applicants should apply at the following website:

<http://apply.interfolio.com/53471>

Application Deadline: **October 19, 2018**

Yale University is an Affirmative Action/Equal Opportunity Employer and welcomes applications from women, persons with disabilities, covered veterans, and members of minority groups.

By Jon Hellin

Inspiration from the outdoors

It was an extraordinarily busy time of year, with deadlines fast approaching for reporting my progress to donors and a journal article requiring major revisions. I felt some pressure to head straight home after the conference I had attended at the last-minute request of my major research funder. But I had traveled halfway around the world to be there, in Morocco. I did not want to miss the opportunity to explore. So I headed to the mountains for 2 days of therapeutic hiking. As I marveled at the sunrise and the view from the icy summit of Mount Toubkal, the highest peak in North Africa, my addled mind calmed. Re-energized, I knew that I could cope with the work that waited for me upon my return. When I arrived back in the office a few days later, I was rejuvenated, and—yes—I successfully submitted the revised journal article and met the donor reporting requirements.

It was almost 20 years ago that I learned how important it is for me to carve out time to escape to the mountains. I had finished my Ph.D., studying soil and water conservation in Honduras, but I wasn't sure what career direction I wanted to take. So, I decided to act on a dream I had been nursing for a few years, since my mom showed me a magazine article about a potter who had cycled from Los Angeles, California, to the southern tip of Tierra del Fuego in Argentina. I figured that I could combine cycling with research for a book on Andean agriculture while I worked out my next career step.

My partner and I spent 12 months cycling from Tierra del Fuego to Ecuador, battling ferocious winds in Patagonia, gasping for air as we inched up mountains in Peru, and rejoicing as the rain became deliciously warmer during the descent from Quito to the Ecuadorian Amazon. Along the way, I learned a fundamental lesson: Being outdoors inspires me, and the beauty of my surroundings combined with physical exertion helps me think broadly and distill my ideas.

During the cycling trip, I found the clarity I needed to figure out my next career move. Talking with farmers we met along the way, from shepherds in Patagonia to potato producers in Ecuador, I learned that their biggest challenge was selling their produce. I decided to broaden my focus from land management to improving farmers' access to markets. My new interest led to working for a nongovernmental organization and eventually conducting agricultural research at international nonprofit organizations, which is what I do now. And my partner and I



“The key is to find your source of inspiration and nurture it.”

hiking. Some things, however, have not changed. I still find that the ideas for many research articles have come to me while surrounded by snow-capped peaks, or even on urban streets on early morning runs. When I'm home, I make sure to head to the hills, at least for a few hours every weekend. I always travel with my running shoes, and occasionally with my hiking boots. And when I find myself straying from this routine and forgetting the nourishment I get from being active in the outdoors, as I nearly did during that conference trip a few years ago, I remind myself of the words of Scottish-American naturalist John Muir: “Wildness is a necessity.” ■

Jon Hellin is a researcher at the International Rice Research Institute in Los Baños, the Philippines. Send your career story to SciCareerEditor@aaas.org.