

How Musk's monster rocket could transform space science p. 702

Vocal simplification led to speech complexity pp. 706 & 760

Protecting biodiversity with tools from the insurance sector p. 714

Science

\$15
12 AUGUST 2022
science.org

 AAAS



GUIDANCE SYSTEM

Death's-head moths correct course based on an internal "compass" p. 764

CALL FOR PAPERS



Plant Phenomics

 OPEN ACCESS

Plant Phenomics is a Science Partner Journal published in affiliation with the State Key Laboratory of Crop Genetics & Germplasm Enhancement, Nanjing Agricultural University (NAU) and distributed by the American Association for the Advancement of Science (AAAS). *Plant Phenomics* publishes novel research that advances both in field and indoor plant phenotyping, with focus on data acquisition systems, data management, data interpretation into structural or functional traits, integration into process based or machine learning based models, and connects phenomics to applications and other research domains.

Submit your research to *Plant Phenomics* today!

Learn more: spj.sciencemag.org/plantphenomics

The Science Partner Journals (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the non-profit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, Open-Access publications produced in collaboration with international research institutions, foundations, funders, and societies. Through these collaborations, AAAS furthers its mission to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility, and publishing expertise that AAAS is uniquely positioned to offer as the world's largest general science membership society.

Learn more at spj.sciencemag.org



@SPJournals



@SPJournals

ARTICLE PROCESSING CHARGES WAIVED UNTIL 2022



Chong Tang, deputy director of research and development, BGI Tech.

Innovative data visualization for gene-related research

A biomedical engineering Ph.D. student at the University of California, Davis, Yongheng Wang had a problem: He wanted to identify the top 10 targets from thousands of genes affected by a drug he developed. The process to mine and extract information would normally take a huge investment of time and effort, switching back and forth between different programming languages and software packages. Wang was delighted to learn of a solution: an innovative data visualization system called "Dr. Tom" from BGI Genomics, a leading global provider of genomic sequencing and proteomics services, headquartered in Shenzhen, China.

Developed by a team of expert scientists and bioinformaticians at BGI Genomics, Dr. Tom is a powerful tool for analysis, visualization, and interpretation of many types of data, including RNA sequencing (RNA-Seq), long noncoding RNA-Seq, microRNA-Seq, whole-genome bisulfite sequencing, single-cell RNA-Seq, and proteomics data.

Having an easy learning curve, Dr. Tom is also versatile, and most importantly, allows researchers without backgrounds in bioinformatics to quickly, efficiently, and reliably create meaningful visualizations. "The intuitive interface allows us to select genes of interest and rank differentially expressed genes based on their fold changes," says Wang. "Dr. Tom covers all the practical functionalities. I think a high school student could do it."

Designed by scientists for scientists

Dr. Tom is a project fueled by the passion of Chong Tang, deputy director of research and development at BGI Tech. A biochemist and molecular biologist with expertise in software engineering, Tang's extensive experience helped him uncover the specific pain point experienced by his research colleagues around the world: Given the tsunami of 'omics data, there is a critical need for a flexible, easy, and turnkey data analysis program. "I designed it in such a way that scientists will like it," he says.

Dr. Tom is designed to be flexible and simple for users with limited computing knowledge, notes Margot Maurer, research associate in the Wolfson Centre for Age-Related Diseases at King's College London. Maurer uses Dr. Tom to analyze the transcriptomics differences in sensory ganglia to help her investigate RNA and age-

related diseases. "I don't need to go through all the calculations," she explains. "Dr. Tom is the bridge. With just a couple of clicks, I can see the graphs, which are easy to read." No programming skills are needed for analyzing these genetic sequences, she adds. "I really like this very visual Venn diagram comparing the different gene expressions in samples."

This visualization feature is designed to be interactive so it is easier to bounce ideas around the team, Tang says. He spent 5 years designing, testing, and fine-tuning Dr. Tom to serve his community, and now Dr. Tom's web-based interface is inspiring innovative research and reducing researchers' headaches.

"The BGI culture is a culture of scientists and for scientists," Tang says. The organization considers it so vital to advance scientific discovery that it is gifting this software to the world.

Making a mark

Relying on cutting-edge sequencing and bioinformatics technology, BGI Genomics has provided scientists and researchers in 100 countries and regions with integrated solutions across a broad range of applications spanning basic life sciences research, clinical research in human health, and agriculture and biodiversity preservation and sustainability.

Since its launch in 2018, Dr. Tom has enchanted more than 15,000 users from over 20 countries and regions. "More than 60 published papers in the field of disease treatment, developmental regulation, immunity, and environmental adaptation have shown Dr. Tom to be a valuable and important tool in addition to any institution's own internal data curation and analysis efforts," says Tang.

BGI Genomics has partnered with academic institutions to offer workshops on leveraging this system effectively, and will be expanding these opportunities to more countries and regions in the future.

Sponsored by

BGI

QUALITY CONTENT FOR THE GLOBAL SCIENTIFIC COMMUNITY

Multiple ways to stay informed on issues related to your research



Posters



Podcasts



Sponsored Collection Booklets



Advertorials



Webinars



Science
AAAS



Scan the code and start exploring the latest advances in science and technology innovation!

[Science.org/custom-publishing](https://www.science.org/custom-publishing)

CONTENTS



706
& 760

12 AUGUST 2022 • VOLUME 377 • ISSUE 6607

NEWS

IN BRIEF

694 News at a glance

IN DEPTH

696 **Scientists scramble to set up monkeypox vaccine trials**

Logistical and ethical challenges are complicating the design of efficacy studies *By K. Kupferschmidt*

697 **Harassment researchers decry proposed reporting rule**

U.S. Title IX law update requiring mandatory reporting of sexual misconduct would cause harm, they say *By K. Langin*

698 **Star's midlife crisis illuminates our Sun's history—and future**

Long magnetic lull mimics Maunder Minimum, when sunspots largely disappeared 400 years ago *By Z. Savitsky*

699 **Star marine ecologist guilty of misconduct, university says**

University of Delaware finding vindicates whistleblowers *By M. Enserink*

700 **Webb reveals early universe's galactic bounty**

Star formation after the big bang appears much faster than models had forecast *By D. Clery*

FEATURES

702 **Prime mover**

Starship will be the biggest rocket ever. Are space scientists ready to take advantage of it? *By S. Scoles*

INSIGHTS

PERSPECTIVES

706 **When less is more in the evolution of language**

Did loss of vocal fold membranes typical of nonhuman primates enable human speech?

By H. Gouzoules

REPORT p. 760



707 **Ultra-sounding out a technique that sticks**

Ultrasound can be used to promote the physical interlocking of adhesives and tissues

By J. Es Sayed and M. Kamperman

REPORT p. 751

709 **Perfluorocubane—a tiny electron guzzler**

Perfluorination gives cubane the capacity to host an extra electron in its inner structure

By M. P. Krafft and J. G. Riess

REPORT p. 756

710 **A soft active matter that can climb walls**

Mechanical activity of an active fluid can be used to control its dynamics at the boundaries *By J. Palacci*

REPORT p. 768

711 **Synthetic gene circuits take root**

Complex spatial patterns of gene expression are engineered in plants to modulate root morphology

By S. Alamos and P. M. Shih

RESEARCH ARTICLE p. 747

712 **Activating translation with phase separation**

Ribonucleoprotein granules allow activation of translation to complete mouse spermatogenesis

By A. Ramat and M. Simonelig

RESEARCH ARTICLE p. 727

PHOTOS: (TOP TO BOTTOM) ZSSD/MINDEN PICTURES; REUTERS/EDUARDO MUNOZ

Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: science.org/journal/sciadv

Science
Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

POLICY FORUM

714 Insurance-sector tools to combat biodiversity loss

Risk transfer can facilitate nature-positive investments *By C. Kousky*

BOOKS ET AL.

717 Our rhythmic existence

A neuroscientist interrogates sleep and circadian clocks *By M. Torres and A. Kramer*

718 Longevity lessons

Animal aging could hold clues to healthier human life spans
By C. Brenner

LETTERS

720 Protect Brazil's overlooked Pampa biome

By J. H. Ellwanger et al.

720 Brazil's Pantanal threatened by livestock

By L. Ferrante and P. M. Fearnside

721 Tech companies liable for illegal wildlife trade

By T. Q. Morcatty et al.

RESEARCH

IN BRIEF

722 From *Science* and other journals

REVIEW

725 Seismology

Deep-learning seismology
S. M. Mousavi and G. C. Beroza
REVIEW SUMMARY; FOR FULL TEXT:
DOI.ORG/10.1126/SCIENCE.ABM4470

RESEARCH ARTICLES

726 Microbiology

Prokaryotic innate immunity through pattern recognition of conserved viral proteins *L. A. Gao et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABM4096

727 Molecular biology

LLPS of FXR1 drives spermiogenesis by activating translation of stored mRNAs *J.-Y. Kang et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABJ6647
PERSPECTIVE p. 712

**Coronavirus**

728 Broadly neutralizing antibodies target the coronavirus fusion peptide
C. Dacon et al.

735 ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies *J. S. Low et al.*

742 Physical chemistry

Proton-coupled energy transfer in molecular triads
B. Pettersson Rimgard et al.

747 Plant science

Synthetic genetic circuits as a means of reprogramming plant roots
J. A. N. Brophy et al.
PERSPECTIVE p. 711

REPORTS

751 Biomaterials

Controlled tough bioadhesion mediated by ultrasound *Z. Ma et al.*
PERSPECTIVE p. 707

756 Organic chemistry

Electron in a cube: Synthesis and characterization of perfluorocubane as an electron acceptor *M. Sugiyama et al.*
PERSPECTIVE p. 709

760 Speech evolution

Evolutionary loss of complexity in human vocal anatomy as an adaptation for speech
T. Nishimura et al.
PERSPECTIVE p. 706

764 Insect migration

Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth *M. H. M. Menz et al.*

768 Fluid dynamics

Dynamics of active liquid interfaces
R. Adkins et al.
PERSPECTIVE p. 710

773 Evolution

The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization *A. J. Barley et al.*

777 Nanophotonics

Cavity-mediated electron-photon pairs
A. Feist et al.

DEPARTMENTS

693 Editorial

Save the Supreme Court and democracy
By M. Sen

782 Working Life

Finding a new calling *By O. Lanes*

ON THE COVER

A death's-head hawkmoth (*Acherontia atropos*), fitted with a miniaturized radio transmitter, is released at dusk in Konstanz, Germany, before a night of tracking. These enigmatic large moths migrate in the autumn, moving through Europe toward the Mediterranean and possibly as far as Africa. Individual tracking has revealed that these moths are able to correct for wind drift to maintain their course during migration. See page 764.
Photo: Christian Ziegler



Science Staff 692
Science Careers 781

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2022 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): \$165 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$2212; Foreign postage extra: Air assist delivery: \$98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #125488122. Publications Mail Agreement Number 1069624. Printed in the U.S.A.

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$15 each plus shipping and handling available from backissues.science.org; bulk rate on request. Authorization to reproduce material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

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

Executive Editor Valda Vinson

Editor, Research Jake S. Yeston **Editor, Insights** Lisa D. Chong **Managing Editor** Lauren Kmec

DEPUTY EDITORS Stella M. Hurlty (UK), Phillip D. Szurmi, Sacha Vignieri **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Groholski, Pamela J. Hines, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Mattia Maroko, Yevgeniya Nusinovich, Ian S. Osborne (UK), L. Bryan Ray, Seth Thomas Scanlon (UK), H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Yuen Yiu **ASSOCIATE EDITORS** Michael A. Funk, Bianca Lopez, Corinne Simonti, Yury V. Suleymanov, Ekeoma Uzogara **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Chris Filiatreau, Harry Jach **SR. CONTENT PRODUCTION EDITOR** Amelia Beyna **CONTENT PRODUCTION EDITORS** Jeffrey E. Cook, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SR. EDITORIAL MANAGERS** Carolyn Kyle, Beverly Shields **SR. PROGRAM ASSOCIATE** Maryrose McMadrid **EDITORIAL ASSOCIATE** Joi S. Granger **SR. EDITORIAL COORDINATORS** Aneera Dobbins, Jeffrey Hearn, Lisa Johnson, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alice Whaley (UK), Anita Wynn **EDITORIAL COORDINATORS** Maura Byrne, Alexander Kief, Ronnel Navas, Isabel Schnaidt, Qiyam Stewart, Brian White **RESEARCH & DATA ANALYST** Jessica L. Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldo (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travers **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Shradha Chakradhar, Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Kelly Servick **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Zack Savitsky, Viviana Flores, Katherine Irving **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Tokyo), Nicholas Pain (Careers), Charles Piller, Gabriel Popkin, Michael Price, Joshua Sokol, Richard Stone, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Morgan Everett, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Chris Bickel **MULTIMEDIA MANAGING PRODUCER** Kevin McLean **WEB STRATEGY MANAGER** Kara Estelle **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR SCIENTIFIC ILLUSTRATOR** Valerie Altounian **SCIENTIFIC ILLUSTRATORS** Kellie Holoski, Ashley Mastin **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITORS** Charles Borst, Kaitlyn Dolan **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski **INTERN** Khafia Choudhary

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark **FINANCIAL ANALYST** Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **SENIOR PRODUCTION MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **CONTENT OPERATIONS MANAGER** Rebecca Doshi **SENIOR CONTENT & PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wolk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhery

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

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst **PROJECT MANAGER** Melissa Collins

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

MARKETING MANAGER Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellani, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Amranta Reyes

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Rie Rambelli (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

MAIN HEADQUARTERS

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

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

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

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

EDITORIAL

science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING

advertising.science.org/
products-services

science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.science.org/
science-careers

advertise@sciencecareers.org

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

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/benefits

AAAS BOARD OF DIRECTORS
CHAIR Susan G. Amara
PRESIDENT Gilda A. Barabino
PRESIDENT-ELECT Keith Yamamoto
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Ann Bostrom
Janine Austin Clayton
Kaye Husbands Fealing
Maria M. Klawe
Jane Maienschein
Robert B. Millard
Babak Parviz
William D. Provine
Juan S. Ramirez Lugo
Susan M. Rosenberg

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

Erin Adams, *U. of Chicago*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Fdn.*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
James Analytis, *UC Berkeley*
Trevor Archer, *NIEHS, NIH*
Paola Arlotta, *Harvard U.*
Delia Baldassarri, *NYU*
Nenad Ban, *ETH Zurich*
Nandita Basu, *U. of Waterloo*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *UT Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Kiros T. Berhane, *Columbia U.*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med.*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Malcolm Brenner, *Baylor Coll. of Med.*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA (\$)*
Christiane Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
György Buzsáki, *NYU School of Med.*
Mariana Byndloss, *Vanderbilt U. Med. Ctr.*
Annmarie Carlton, *UC Irvine*
Simon Cauchemez, *Inst. Pasteur*
Ling-Ling Chen, *SIBCB, CAS*
Ib Chorkendorff, *Denmark TU*
Karlene Cimprich, *Stanford U.*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Virginia Cornish, *Columbia U.*
Carolyn Coyne, *Duke U.*
Roberta Cozzani, *VU Amsterdam*
Christina Curtis, *Stanford U.*
Ismail Dabo, *Penn State U.*
Jeff L. Dangl, *UNC*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Frans de Waal, *Emory U.*
Claude Desplan, *NYU*
MIRA Diaz, *U. Nacional de Córdoba*
Ana Magurran, *U. of St. Andrews*
Ari Pekka Mäöhönen, *U. of Helsinki*
Ulrike Diebold, *TU Wien*
Stefanie Dimmeler, *Goethe-U. Frankfurt*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Ruth Drdla-Schutting, *Med. U. Vienna*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Nader Egheta, *UPenn*
Karen Ersche, *U. of Cambridge*
Beate Escher, *UFZ & U. of Tübingen*
Ray Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzi, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *UCLA*
Ramon Gonzalez, *U. of South Florida*
Sandra González-Bailón, *UPenn*
Nicolas Gruber, *ETH Zurich*
Hua Guo, *U. of New Mexico*
Tekjip Ha, *Johns Hopkins U.*
Daniel Haber, *Mass. General Hos.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zurich*
Louise Harra, *U. Coll. London*
Carl-Philipp Heisenberg, *IST Austria*
Janet G. Hering, *Eawag*
Christoph Hess, *U. of Basel & U. of Cambridge*
Heather Hickman, *NIAID, NIH*
Hans Hilgenkamp, *U. of Twente*
Janneke Hille Ris Lambers, *ETH Zurich*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Gwyneth Ingram, *ENS Lyon*
Darrell Irvine, *MIT*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS & U. of Arizona*
Erich Jarvis, *Rockefeller U.*
Peter Jonas, *IST Austria*
Johanna Joyce, *U. de Lausanne*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *UC Berkeley*
Kisuk Kang, *Seoul Natl. U.*
Sabine Kastner, *Princeton U.*
V. Narry Kim, *Seoul Natl. U.*
Robert Kingston, *Harvard Med.*
Nancy Knowlton, *Smithsonian Institution*
Etienne Kochelin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Paul Kubes, *U. of Calgary*
Chris Kuzawa, *Northwestern U.*
Laura Lackner, *Northwestern U.*
Gabriel Lander, *Scripps Res. (\$)*
Mitchell A. Lazar, *UPenn*
Hedwig Lee, *Duke U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *Wash. U. in St. Louis*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Ana Magurran, *U. of St. Andrews*
Ari Pekka Mäöhönen, *U. of Helsinki*
Asifa Majid, *U. of Oxford*
Oscar Marin, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth*
Rodrigo Medellín, *U. Nacional Autónoma de México*
C. Jessica Metcalf, *Princeton U.*
Baoxia Mi, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Alison Motsinger-Reif, *NIEHS, NIH (\$)*
Suresh Naidu, *Columbia U.*
Beate Escher, *UFZ & U. of Tübingen*
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Helga Noheda, *U. of Groningen*
Beatriz Nowotny, *Vienna Sci. & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst. (\$)*
Daniel Pauly, *U. of British Columbia*
Ana Pêgo, *U. do Porto*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Inst. Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Lei Stanley Qi, *Stanford U.*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial Coll. London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *UT Austin*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Erica Ollmann Saphire, *La Jolla Inst.*
Jochim Saur, *U. zu Köln*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UCSF*
Robert Siliciano, *JHU School of Med.*
Lucia Sivilotti, *U. Coll. London*
Richard Smith, *UNC (\$)*
John Speakman, *U. of Aberdeen*
Tara Spres-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Northwestern U.*
Ira Tabas, *Columbia U.*
Eriko Takano, *U. of Manchester*
A. Alec Talin, *Sandia Natl. Labs*
Patrick Tan, *Duke-UNC Med. School*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titicomi, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Maria-Elena Torres Padilla, *Helmholtz Zentrum München*
Kimiari Toussaint, *Brown U.*
Barbara Treutlein, *ETH Zurich*
Jason Tylianakis, *U. of Canterbury*
Wim van der Putten, *Netherlands Inst. of Ecology*
Matthew Vander Heiden, *MIT*
Ivo Vankelecom, *KU Leuven*
Judith Varner, *UC San Diego*
Henrique Veiga-Fernandes, *Champalimaud Fdn.*
Reinhilde Veugelers, *KU Leuven*
Beit Vogelstein, *Johns Hopkins U.*
Julia Von Blume, *Yale School of Med.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis (\$)*
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Chris Wikle, *U. of Missouri (\$)*
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res. (\$)*
Hao Wu, *Harvard U.*
Li Wu, *Tsinghua U.*
Wei Xie, *Tsinghua U.*
Benjamin Youngblood, *St. Jude*
Yu Xie, *Princeton U.*
Jan Zaenen, *Leiden U.*
Kenneth Zaret, *UPenn School of Med.*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

Save the Supreme Court and democracy

The US Supreme Court has been busy. It recently overturned a nearly 50-year-old precedent protecting abortion rights, upheld the right to carry guns outside the home, and hamstrung the Environmental Protection Agency's ability to regulate emissions—all while signaling an aversion to contemporary empirical evidence and instead favoring “history and tradition.” Although the majority of Americans disagree with many of these decisions, the court has only just begun to reshape the country. When it resumes in October, the court will be poised to outlaw affirmative action, undercut federal regulations regarding clean water, and possibly allow state legislatures to restrict voting rights without oversight by state courts. What explains the court's shift to an ideological extreme, and what can be done about it?

The court's eye-popping move to the conservative right is confirmed by research that compares its decisions to public opinion. Decades-long surveys reveal that the court's rulings were in step with the opinions of most Americans through 2020. However, around 2021, its views on important issues (including abortion and religion) swerved to the right of those held by a majority of the public—and are now more aligned with the views of Republican voters. This sharp turn coincided with the 2020 death of Ruth Bader Ginsburg, a liberal-leaning justice.

The shift can be best understood by looking at the identity of the court's ideological “median voter.” In considering the nine Supreme Court justices from most liberal to most conservative, the justice who sits in the median fifth position is the one whose vote will create a five-person majority in many cases. This gives the median enormous power over decisions. In 2018, the median justice was Anthony Kennedy, a moderate conservative but also an occasional cross-over voter on issues such as abortion and capital punishment. Upon retiring that year, he was replaced by staunch conservative Brett Kavanaugh, shifting the median rightward to Chief Justice John Roberts. As the median, Roberts would have kept the court more or less in line with the American ideological mainstream. However, when Ginsburg was replaced by conservative Amy Coney Barrett, the median moved further to the right—somewhere around Kavanaugh or Neil Gorsuch. Although

Roberts could be considered center-right, new medians Kavanaugh and Gorsuch are firm conservatives. By some estimates, Gorsuch is more conservative than around 85% of Americans.

Such a shift matters because public perception of the court's ideology is critical. If people think the court is ideologically opposed to them, they will be more likely to think that it is acting purely politically, and they will be more likely to support proposals that curb the court's powers. Indeed, a method of estimating judicial ideology based on justices' voting—known as Martin-Quinn scores—suggests that the justices have separated into two ideological blocs defined by political party, implying not just a strongly consolidated conservative majority but also less ideological “cross-over” voting. If the majority of Americans continue to strongly disagree with the court's decisions, then the public may well turn against it, which may not only lead to greater political conflict but also undermine the rule of law.

Reforming the court to prevent extreme ideological movements may be difficult, but not impossible. For example, 18-year term limits for justices would regularize appointments—eliminating gamesmanship around

vacancies and reducing incentives for justices to strategically time retirements. This would help prevent extreme partisan imbalance and thus keep the court closer to the ideological mainstream. Term limits enjoy wide bipartisan support and would put the US in line with other democratic peer nations, all of which have term or age limits for their high courts. Additional promising proposals by scholars to help reduce ideological imbalance include reconfiguring how the US selects justices and expanding the size of the court. Others—such as stripping the court's jurisdiction—would address the argument that the court wields too much power.

As gridlock and polarization continue to undercut the efficacy of elected branches of government, the Supreme Court's salience in matters of public importance will only rise. And the court's conservative supermajority (6-to-3)—one closely aligned with Republican Party policy interests—has many years left. The discussion of court reform will not go away soon.

—Maya Sen



Maya Sen is a professor of public policy at the John F. Kennedy School of Government, Harvard University, Cambridge, MA, USA. msen@hks.harvard.edu

“What explains the court's shift to an ideological extreme, and what can be done about it?”

“ We [worry it] will hinder Lilly’s and Indiana’s ability to attract diverse scientific talent from around the world. ”

Pharmaceutical giant **Eli Lilly and Co.** on the state’s passage of what the corporation called “one of the most restrictive anti-abortion laws in the United States.”



A South Korean lunar probe was launched on 4 August from Cape Canaveral in Florida.

IN BRIEF

Edited by **Shraddha Chakradhar** and **Jeffrey Brainard**

PLANETARY SCIENCE

South Korea sends orbiter to Moon to search for ice

South Korea’s first Moon probe was lofted into space from Cape Canaveral in Florida on 4 August by a SpaceX rocket. The \$200 million Korea Pathfinder Lunar Orbiter, also called Danuri—“enjoy the Moon” in Korean—will study the Moon from a polar orbit for at least a year. One of the probe’s five instruments will capture polarized light to measure the grain sizes of lunar dust, an indicator of

“weathering” by the solar wind and hence of the age of features such as lava flows and impact craters. In another first, a highly sensitive camera on Danuri will peek into the depths of the Moon’s permanently shadowed craters to inventory the water ice known to lurk there. Other instruments on the satellite will find suitable locations for a lander, planned for the early 2030s, as South Korea’s next step into deep-space exploration.

U.S. boosts monkeypox response

PUBLIC HEALTH | President Joe Biden’s administration last week designated the monkeypox outbreak a national public health emergency, allowing U.S. health officials easier access to funds and procedural flexibility as they respond to rising cases (more than 8900 as of 8 August). Earlier in the week, the White House appointed Robert Fenton, a senior official at the Federal Emergency Management Agency, as national monkeypox response

coordinator. Demetre Daskalakis, a physician who directs the Division of HIV Prevention at the U.S. Centers for Disease Control and Prevention, will serve as deputy coordinator. Daskalakis has experience working with the LGBTQ community; 97.5% of monkeypox cases with available data on sexual behavior have been in men who have sex with men, according to a 3 August report from the World Health Organization. As *Science* went to press, the United States had the world’s largest number of confirmed monkeypox cases.

Few get prompt hep C treatment

INFECTIOUS DISEASE | Less than one-third of people infected with the hepatitis C virus in the United States who have medical insurance receive lifesaving treatment for the disease within 1 year of their diagnosis, the U.S. Centers for Disease Control and Prevention (CDC) reported this week. The liver-damaging virus caused 14,000 deaths in the United States in 2019. But an 8- to 12-week course of pills that were hailed as a medical breakthrough when

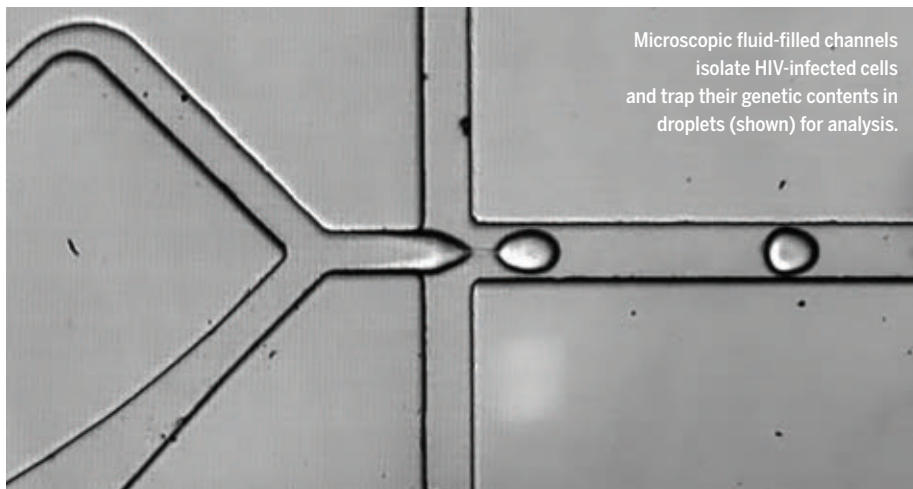
they came to market nearly a decade ago can cure most infections. CDC analyzed the treatment records of 47,687 insured people from January 2019 through October 2020 and blamed the low uptake on eligibility restrictions and what are known as pre-authorization requirements that insurance plans often impose. The treatment rate was lowest, 23%, among people on Medicaid, the federal program that covers low-income Americans.

Small reactor design wins OK

NUCLEAR POWER | NuScale Power, a startup based in Portland, Oregon, came closer to bringing a small, modular nuclear reactor to market when the U.S. Nuclear Regulatory Commission last week certified its design. NuScale says it would be safer, more versatile, and more economical than larger conventional reactors. The company has a deal with an electricity supplier, the Utah Associated Municipal Power Systems (UAMPS), to build a power plant comprising six of the factory-built reactors at Idaho National Laboratory. NuScale isn't done with NRC reviews, however. The regulator certified the new reactor to produce 50 megawatts of electrical power—about 5% of the capacity of a typical power reactor—but the UAMPS plan now calls for running each of the six at a more cost-efficient 77 megawatts. That 54% “uprate” in power will require further NRC review before the Idaho plant can be licensed for operations. Other countries are also developing small reactor designs.

Primate center director stays on

SCIENTIFIC INTEGRITY | Critics have panned a decision by the Southwest National Primate Research Center (SNPRC) to retain its director with full duties despite a finding by the U.S. Office of Research Integrity (ORI) that he engaged in research misconduct. As Retraction Watch first reported, ORI found that Deepak Kaushal fabricated data on testing tuberculosis drugs in monkeys in a study published in 2020 and in two grant applications. He and colleagues produced the data while he was working at the Tulane National Primate Research Center; he became SNPRC's director in 2019. Kaushal, who did not respond to several interview requests from *Science*, has not escaped all consequences: His settlement agreement with ORI requires a committee of senior faculty members to supervise his research for 1 year. SNPRC, based in San Antonio, said the misconduct finding does not affect his leadership of the center, however. The Texas Biomedical



Microscopic fluid-filled channels isolate HIV-infected cells and trap their genetic contents in droplets (shown) for analysis.

BIOMEDICINE

Probing HIV's cellular fortresses opens cure avenues

A research team has developed a way to assess the gene activity of single cells that harbor latent HIV genes—a technique that could aid the search for a cure. People living with HIV who take existing antiretrovirals invariably retain infected cells that dodge the drugs and natural immune responses. Even though scientists could identify these rare reservoir cells, technical constraints prevented them from evaluating the cells' gene activity. The new method, revealed at the 24th International AIDS Conference last week, hinges on “microfluidic” devices (above) that allow investigators to retrieve genetic material from the infected cells for sequencing. The team found that the reservoir cells had unique patterns of gene activity, turning on genes that protect them from immune attack and self-destruction. Targeting these genes could, in theory, reduce, if not eliminate, the HIV reservoirs.

Research Institute, which oversees SNPRC, noted that Kaushal's team retracted the original study and published a corrected version, whose conclusions did not change. But some outside scientists say the misconduct finding could feed public doubts about animal research generally and should prevent Kaushal from leading SNPRC. The center received \$10.3 million in federal funding in 2021 and houses about 2500 marmosets, baboons, and macaques.

Long Covid gets U.S. office

COVID-19 | President Joe Biden's administration last week published a plan for research on Long Covid, the long-lasting sequelae of acute SARS-CoV-2 infection, which the U.S. Centers for Disease Control and Prevention recently estimated affects up to one in five U.S. adults previously infected with the pandemic coronavirus. The blueprint creates an Office of Long COVID Research and Practice at the Department of Health and Human Services. It will be headed by Rachel Levine, the department's assistant secretary for health, who has been overseeing the administration's Long Covid response.

The new research will explore diagnostics and treatments and build on existing studies by the National Institutes of Health and the Department of Veterans Affairs to seek genes associated with Long Covid.

Radar mapper declared dead

REMOTE SENSING | In a blow to scientists who monitor tiny movements of Earth's surface caused by earthquakes, volcanoes, and glaciers, the European Space Agency said last week that Sentinel-1B, a radar mapping satellite, cannot be revived after a power supply malfunction caused it to go dark in December 2021. Investigators say a defective capacitor probably caused the fault. The loss of the satellite, launched in 2016, leaves the agency with a twin, Sentinel-1A, which was lofted into orbit two years earlier and is operating past its designed lifetime of 7 years. The duo had surveyed the globe, measuring surface changes as small as a few millimeters every 6 days as it orbited the planet. Until the agency launches Sentinel-1C in 2023, the intervals between observations will drop to 12 days, hampering data collection and mapping efforts.

Empty monkeypox vaccine vials and syringes at a vaccination point in Valhalla, New York, on 28 July.

not the same thing as just brushing up against somebody,” says Anne Rimoin, an epidemiologist at the University of California, Los Angeles.

So far, there’s scant data on how well the vaccine is working in the current outbreak. Among 276 individuals who received a shot at a Paris hospital as post-exposure prophylaxis (PEP) after reporting a high-risk contact, 12 developed a monkeypox infection, French scientists reported in a recent preprint. But there was no control group, making it impossible to tell how many people would have developed monkeypox if no one had been vaccinated. And people eager to be vaccinated may have lied about having had a high-risk contact. “That makes results from these studies on PEP really hard to evaluate,” says immunologist Leif Erik Sander of the Charité clinic in Berlin, who’s setting up a vaccine study in Germany.

A randomized trial—in which one group receives the vaccine and the other does not—would avoid such problems. Without a randomized study, “you can end up in this evidence limbo and find that if you had just done the trial, you would have been in a much better situation,” says virologist Natalie Dean of the University of Florida.

Giving a control group a placebo instead of a presumably effective vaccine is ethically dicey, many researchers say. But University of Oxford epidemiologist Richard Peto sees another way. Because demand for the vaccine is so much higher than supply, “Why not randomize the order in which people in the highest risk group are called in?” Peto asks. So far, however, no one seems to have taken up that idea.

Sander considered a randomized design but decided against it. “There was a lot of pushback,” he says. Instead, he has started a so-called cohort study in which he hopes to enroll 5000 vaccinated and 10,000 unvaccinated people at risk of monkeypox and follow them for 12 months. (Over time, some of the unvaccinated people will receive the shot so the groups may become more similar in size.) So far, about 800 people have been enrolled.

The groups may differ in ways other than their vaccine status—people with lots of sexual contacts may try harder to get vaccinated, for example—but there is still an element of randomization, Sander says: Many doctors are using lottery-type procedures to decide who gets the vaccine first.

A cohort study in France is taking another approach. There, MSM already enrolled in a study of sexually transmit-

PHOTO: REUTERS/EDUARDO MUNOZ

IN DEPTH

INFECTIOUS DISEASES

Scientists scramble to set up monkeypox vaccine trials

Logistical and ethical challenges are complicating the design of efficacy studies

By Kai Kupferschmidt

When monkeypox suddenly started spreading globally in May, the world was fortunate in one respect: a vaccine was available. MVA, originally developed by Bavarian Nordic as a smallpox vaccine, was already licensed for monkeypox in Canada and the United States. EU regulators have since followed suit. By now clinics have delivered thousands of doses to people in high-risk groups.

There’s little doubt the vaccine can help, but that’s about all that’s certain. Exactly how well MVA protects against monkeypox and for how long is not known. Nor is it clear how much protection is lost by giving just a single dose rather than the recommended two doses, as some countries are doing to stretch supply, or how much protection a vaccine given after exposure can offer.

But the ethical and logistical complexities of the monkeypox crisis, which is overwhelmingly affecting men who have sex with men (MSM), are making these questions hard to answer. Placebo-controlled clinical trials are fraught because MVA is

already licensed and people are clamoring to get it. And vaccine clinics are often set up at short notice as doses become available, making it harder to organize a trial and enroll subjects. Researchers are responding with a plethora of inventive trial designs.

The first evidence that smallpox vaccines also protect against monkeypox came from a study in the 1980s in the Democratic Republic of the Congo (then called Zaire), where the virus occasionally jumps from animals to people, who then infect others in their household. A study among patients’ contacts suggested smallpox vaccination was 86% effective at preventing monkeypox as well. But the study looked at a small number of cases, the virus was genetically quite different than the one now spreading, and the smallpox vaccine was an older one with more side effects; MVA was developed as a safer alternative.

MVA was licensed for monkeypox based on data from animal experiments and the immune response it triggers in humans. But its efficacy has barely been tested in people, and not at all for preventing sexual transmission, which results in “very significant mucosal exposure, which is

ted diseases—and deemed at high risk of monkeypox—will get MVA in the next 2 months. Jade Ghosn of Bichat Hospital, who runs the study, hopes to have all participants vaccinated by the end of September and plans to compare infection rates before and after vaccination.

Another option is a “test-negative” design, in which researchers look at people who seek testing for monkeypox and compare the percentages of people who were vaccinated among those who test positive and negative. This is “probably the strongest nonrandomized approach to measuring vaccine efficacy,” says Michael Marks, an epidemiologist at the London School of Hygiene & Tropical Medicine who is planning a vaccine trial soon with colleagues in Spain.

The test-negative setup requires good linkage between vaccination and testing data. “If we can solve that issue we may use such a design in our study,” Marks says. The Canadian province of Ontario is going ahead with a similar design, says Jeff Kwong of the University of Toronto. The drawback is that testing and vaccination data alone can’t answer many other questions, such as how immunity develops over time or whether disease severity is different among the vaccinated and the unvaccinated; that requires additional studies.

The U.S. National Institute of Allergy and Infectious Diseases (NIAID) does plan a randomized trial, aimed at finding out if the vaccine supply can be stretched by giving people much smaller doses. Participants will either get two full doses or two one-fifth doses 4 weeks apart; a third arm may be added to test one-tenth of the normal dose, says NIAID’s John Beigel, who is involved in designing the study. The lower doses will be injected into the skin, which is known to cause a more vigorous immune response than the standard subcutaneous shot. But the study, expected to start in September, will only test whether fractional doses trigger a similar reaction as the full dose; it won’t measure vaccine efficacy directly.

One strategy not tested in the trial, even though it is being used, is giving just one full dose. Available data suggest that regimen is inferior to two full doses, Beigel says: “We don’t think it’s scientifically supported.”

With so many unanswered questions it’s hard to provide good vaccine information to those at risk, says Will Nutland, a U.K. community organizer who runs an organization for MSM sexual health. That should not deter people from seeking the shots, he says: “I think most people understand ... that it is better to receive some level of protection than no protection at all.” ■

WORKFORCE

Harassment researchers decry proposed reporting rule

U.S. Title IX law update requiring mandatory reporting of sexual misconduct would cause harm, they say

By **Katie Langin**

When Allison Cipriano read the U.S. Department of Education’s recently proposed amendments to Title IX, the federal statute prohibiting sexual discrimination at educational institutions that receive federal funding, she was “quickly disappointed.” The 700-page document includes many rule changes she wanted to see, including protections for sexual minorities. But Cipriano, a Ph.D. student in psychology at the University of Nebraska, Lincoln, who studies sexual misconduct investigations in academia, and others are dismayed by one set of changes. They would compel most university employees to be “mandatory reporters”—required to notify their institution’s Title IX office of any alleged sexual misconduct involving students they become aware of, regardless of whether the student wants them to.

Researchers who study such policies have spoken out against them, saying they re-traumatize targets of sexual misconduct and reduce their ability to seek the support they need. Mandatory reporting, which proponents argue helps universities unearth and address misconduct, is already common at U.S. universities; a 2018 study found 88% of institutions required most or all employees to serve as mandatory reporters. But the government’s proposed rule changes—open for comment until 12 September—would make the requirement universal, blocking efforts to try other approaches.

“I was expecting all the changes to be good, so I’m caught off guard ... by what I perceive to be a potential—if it doesn’t get fixed—disaster,” says Jennifer Freyd, a trauma researcher and professor emerita at the University of Oregon. She fears the new mandatory reporting rules will be particularly harmful for graduate students, who are often highly dependent on faculty members and are more vulnerable than undergraduates to “career-killing” retaliation. “If ... a grad student can speak privately with a supportive faculty member in their department about how to navigate the bind they are in, it can save a career,” she says.

Past regulations have required universities to designate some employees as mandatory reporters but didn’t define who those employees should be. The new regulations, in contrast, would require any university employee who “has the authority to institute corrective measures” or the “responsibility for administrative leadership, teaching, or advising” to serve as a mandatory reporter. That definition encompasses most faculty members, says Lilia Cortina, a professor at the University of Michigan, Ann Arbor, who studies gender and co-authored the National Academies of Sciences, Engineering, and Medicine’s 2018 sexual harassment report.

“In the report, we talked about how problematic that is,” Cortina says of broad mandatory reporting policies. “[There’s]



Faculty members could be required to report information shared with them in confidence.

research showing that when you take control away from victims, that’s actually associated with an increase in psychological distress.” The requirements “essentially amount to nonconsensual reporting [of] their traumatic or humiliating or otherwise extremely distressing experience.”

In a statement to *Science*, a Department of Education spokesperson wrote that the updates were developed “with the aim of ensuring full protection under Title IX for students, employees, and others and to end all forms of sex discrimination.” They declined to comment about why universal mandatory reporting is necessary, but the proposed regulations note that students “may be less capable of self-advocacy” and

may share Title IX violations with those who fall under the proposed mandatory reporter definition “with the expectation that doing so would obligate the recipient to act.”

But targets of sexual harassment may go to trusted faculty members for reasons other than to trigger a formal investigation, as Cipriano found when she interviewed graduate students about their experiences with mandatory reporting. For instance, one student opened up to her adviser because she wanted to let him know she was seeking mental health services, Cipriano and her colleagues report in a preprint posted last month. “I had to do it on work time, so I felt like he deserved to know why I wasn’t at my desk,” the student said.

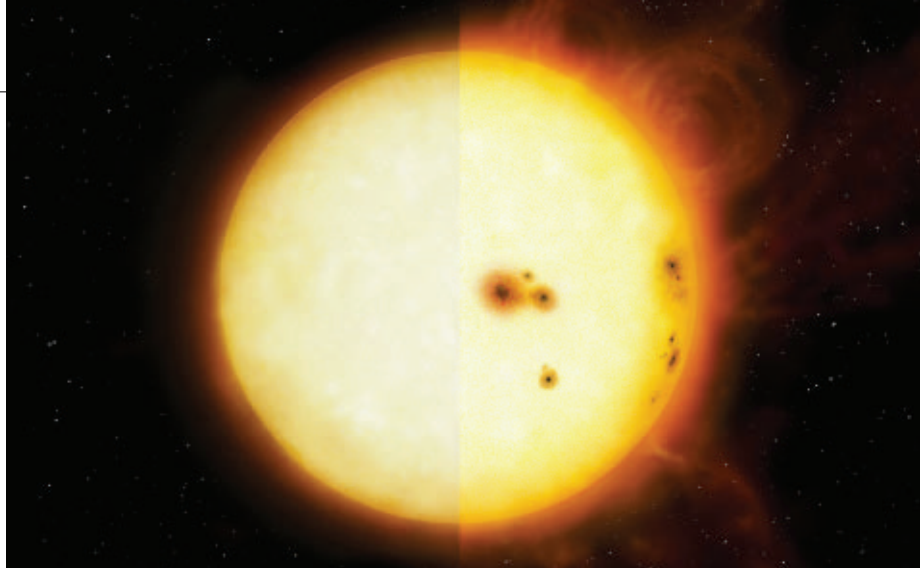
Most of the students in Cipriano’s study weren’t aware their institution even had a mandatory reporting policy. “I was under the impression that talking to my adviser was a safe communication,” one said. Another told Cipriano she freaked out after learning a report would be filed. “When you’re in the midst of trauma ... you’re not going to be thinking about the ins and outs of university policy,” Cipriano says. “Students are not going to see their professors as a de facto arm of the Title IX office.”

At a listening session the Department of Education held last year to help guide the development of the new regulations, Cipriano shared research she and others have done on mandatory reporting—which made it all the more frustrating when she read the resulting amendments. “It really does feel like a bit of a slap in the face,” she says.

Others, including Brett Sokolow, a lawyer who serves as president of the Association of Title IX Administrators, see the issue differently. “What most research on mandated reporting misses is that mandated reporting largely does not result in ‘forced prosecutions’ but instead results in outreach, sharing of resources, and discussion of options. That’s not really taking agency away from survivors,” he says.

The proposed regulations will block efforts to test other approaches. The University of Oregon, for instance, has a “mandatory supporter” policy, which requires most employees—including most faculty members—to tell students how they could report sexual misconduct and then let the student decide how to move forward. (Those in higher administrative positions, such as deans and department heads, are mandatory reporters.)

“We need some room for experimentation,” says Freyd, who helped develop the University of Oregon policy. “I don’t think we’re at a point yet where we have enough information for the federal government to dictate to institutions everything about how they should handle Title IX issues.” ■



Sunspots wax and wane with the 11-year solar cycle, as shown in an illustration of minimum (left) and maximum activity, but they mostly vanished for decades soon after telescopes emerged.

ASTROPHYSICS

Star’s midlife crisis illuminates our Sun’s history—and future

Long magnetic lull mimics Maunder Minimum, when sunspots largely disappeared 400 years ago

By **Zack Savitsky**

Soon after European astronomers developed the first telescopes at the start of the 17th century, they observed dark spots speckling the Sun’s surface. They also handed their modern successors a mystery. From about 1645 to 1715, the spots, now known to be indicators of solar activity, all but disappeared. Gathering sunspot counts and other historical observations, astronomer John Eddy concluded nearly 50 years ago that the Sun had essentially taken a 70-year nap, which he called the Maunder Minimum after an astronomer couple who had previously studied it (*Science*, 18 June 1976, p. 1189).

Now, it appears the Sun is not the only star that takes long naps. By building a decades-long record of observations of a few dozen stars at specific wavelengths that trace stellar activity, a team of astronomers has identified another star going through its own Maunder Minimum period. “I am more convinced this is a Maunder Minimum star than anything else I’ve seen,” says Jennifer van Saders, an astronomer at the University of Hawaii, Manoa, who was not involved in the discovery.

The finding, reported in a preprint last month on arXiv, could help explain what triggered the Sun’s strange behavior 400 years ago and suggests more such episodes are likely. “This is the way to study the past and

future of the Sun,” van Saders says. She adds the discovery supports a theory she and colleagues have advanced: that such events are an occasional symptom of a critical transition in the magnetic field of Sun-like stars about halfway through their lifetime—a midlife crisis of sorts. Some astronomers speculate that the Sun’s transition helped favor the emergence of life on Earth, and that searching for stars in a similar stage could help identify other solar systems conducive to complex life.

Scientists have known for decades that our Sun’s activity surges and ebbs on a roughly 11-year cycle, which corresponds to how often its magnetic poles flip their orientation. During a solar maximum, sunspots proliferate, marking weak points in the magnetic field, where plasma from the Sun’s atmosphere can lash out in violent loops. Astronomers have spotted young Sun-like stars with similar cycles, and older ones that have totally stable activity. But no one had spotted a cycling star suddenly turning flat.

In 2018, as part of undergraduate research at Pennsylvania State University, University Park, Anna Baum set out to combine observations of the telltale wavelengths from 59 stars taken by the Mount Wilson Observatory and the W. M. Keck Observatory to produce a 50-year chronology of star evolution. During a 7-year gap in data while Keck was upgrading a detector, one star appeared to show a drastic shift. Its activity went from cycling over a

17-year period to being virtually flat, and it's stayed that way for the past 18 years.

Baum thought at first she'd made an error; perhaps the observatories were even looking at two different stars. But earlier this year, her colleagues came across additional observations that filled in the data gap, capturing the star's emissions as it switched from active to quiet. The recovered data set "hit the jackpot," says Jacob Luhn, an astronomer at the University of California, Irvine, and lead author on the preprint.

The discovery reinforces one popular theory about why these extended quiescent periods happen. Stars spin more slowly with age because their solar winds act as "magnetic brakes," like a child sticking out their arms while revolving in a chair. In 2016, van Saders and her colleague Travis Metcalfe of the White Dwarf Research Corporation noticed that at some point, stars stop hitting the brakes and their velocity stabilizes—a shift, they proposed, that stems from a change in the stars' magnetic field. Then, last year, Dibyendu Nandi and colleagues at the Center of Excellence in Space Sciences India pinned down the idea with computer simulations that linked the stabilizing of the spin rate to a weakening magnetic field. During this transition, as the star heads toward a "lazy" state in which its activity is flat rather than cycling, random perturbations in its magnetic field can result in temporary cycle shutoffs like the Maunder Minimum, Nandi says.

The theory predicts that this transition state will emerge in middle-aged stars—just like our Sun and this newly identified napping star. "Everything about this discovery has actually corroborated what we've been talking about for the last 5 years," Metcalfe says. "We definitely knew about stars that were not cycling, but we didn't know how they got there—this is like the missing link in that evolutionary picture."

Our Sun's magnetic transition probably began around the same time life on Earth first crawled out of the sea, and that may be no coincidence, Metcalfe suggests. The incoming particles and radiation from active stars damage DNA and promote mutations, speeding evolution. They "may be part of the necessary ingredients to get life started," he says. But at some point, energetic space weather poses a threat to complex life—"like a giant cosmic reset button that's always going off," he adds.

Stars undergoing a transition from cycling to stable could provide the ideal balance of spark and protection to nurture life. "If we're looking for technological civilizations," Metcalfe says, "maybe the best place to look is around stars that are in the second half of [their] lifetimes"—in other words, just entering a midlife crisis. ■

RESEARCH ETHICS

Star marine ecologist guilty of misconduct, university says

University of Delaware finding vindicates whistleblowers

By **Martin Enserink**

A long-running controversy in marine biology took a new twist last week when the University of Delaware (UD) found one of its star scientists guilty of research misconduct. The university accepted an investigative panel's conclusion that marine ecologist Danielle Dixon committed fabrication and falsification in work on fish behavior and coral reefs. UD is seeking the retraction of three of Dixon's papers and "has notified the appropriate federal agencies," a spokesperson says.

Among the papers is a study about coral reef recovery that Dixon published in *Science* in 2014 and for which the journal issued an Editorial Expression of Concern in February. *Science*—whose News and Editorial teams operate independently of each other—retracted that paper on 9 August.

The investigative panel's draft report, which *Science's* News team has seen in heavily redacted form, paints a damning picture of Dixon's scientific work. "The Committee was repeatedly struck by a serial pattern of sloppiness, poor recordkeeping, copying and pasting within spreadsheets, errors within many papers under investigation, and deviation from established animal ethics protocols," wrote the panel of three UD researchers.

Dixon did not respond to requests for comment. She "adamantly denies any and all allegations of wrongdoing, and will vigorously appeal any finding of research misconduct," her lawyer, Kristina Larsen, wrote in an email to *Science*. Dixon and two co-authors on the 2014 paper did not agree to the retrac-

tion. UD says it will not comment on Dixon's future there.

Dixon is known as a highly successful scientist and fundraiser. She obtained her Ph.D. at James Cook University (JCU), Townsville, in Australia in 2012; worked as a postdoc and assistant professor at the Georgia Institute of Technology for 4 years; and in 2015 started her own group at UD's marine biology lab in Lewes, a small town on the Atlantic Coast.

Together with one of her Ph.D. supervisors, JCU marine biologist Philip Munday, Dixon pioneered research into the effects on fish of rising carbon dioxide levels in the atmosphere, which cause the oceans to acidify. In a series of studies published since 2009 they showed acidification can disorient fish, lead them to swim toward chemical cues emitted by predators, and affect their hearing and vision. Dixon's later work focused on coral reef ecology.

But a group of other marine biologists had long questioned the very big effect sizes and unusually small variances in many of Dixon and Munday's results. In a 2020 *Nature* paper, they reported being unable to reproduce several of the claims. Later, four of them, along with former members of Dixon's UD lab, decided to ask for a misconduct investigation (*Science*, 7 May 2021, p. 560).

In the draft report, UD's investigative committee says Dixon simply did not have enough time to collect the vast amount of data described in the *Science* paper, seaweed-covered reefs in Fiji fail to attract juvenile individuals of 15 fish and three coral species. The draft report also confirms claims



Danielle Dixon studied coral reefs in Fiji for her 2014 study in *Science*, which has now been retracted.

by the whistleblowers and independent statistical experts that a large Excel file of the study's raw data was riddled with inexplicably duplicated columns.

One of the paper's two co-authors, Georgia Tech marine ecologist Mark Hay, was Dixon's supervisor at the time. He says he wasn't in Fiji during the study. But Dixon "was among the most hard-working and persistent colleagues I've worked with," and the amount of data she reported "did not seem unreasonable given Dr. Dixon's time in Fiji and strong work ethic," he says. Hay says he has not seen the panel's report, but "if UD found fault with her studies, I'm surprised." Dixon "seemed honest and genuine." Hay says he was interviewed by a separate investigative committee at Georgia Tech but does not know the status or outcome of that investigation. A Georgia Tech spokesperson did not respond to a request for comment.

Dixon and Hay published another paper in *Science* in 2012 showing corals can "recruit" plant-eating fish to trim toxic seaweeds. "We have been discussing the 2012 paper but were not given any information about it by the university," says Holden Thorp, *Science's* editor-in-chief, "so we will begin our usual process to determine whether retraction or correction is warranted."

The draft report also found misconduct in a 2016 paper on whether anemone fishes can sniff out the condition of potential host anemones, published in the *Proceedings of the Royal Society B* by Dixon and marine biologist Anna Scott of Southern Cross University in Australia. Again, the timeline was implausible, the committee concluded. Collecting the data would have taken 22 working days of 12 hours, it wrote, "working continuously without any breaks." Yet the paper said the study was done in 13 days, between 12 and 24 November 2014.

Scott and Dixon posted a correction to the *Proceedings B* paper in early July, stating the studies actually took place between 5 October and 7 November 2014, adding 20 days to the timeline. They also uploaded the raw data for the study, which had been missing. But the whistleblowers say that file also has duplications and other problems.

Scott tells *Science* she signed off on the correction. She did not respond to questions about UD's misconduct findings or the alleged problems in the data set. Spencer Barrett, editor-in-chief of *Proceedings B*, says UD has asked for a retraction of the paper after the correction ran, and that an investigative panel at the journal will now reconsider the issue. The journal is

still investigating another paper by Dixon, Barrett says.

The UD committee noted that a 2014 paper in *Nature Climate Change* co-authored by Dixon and Munday also "had serious issues regarding the datasheet files," which it called "signatures of fabrication and falsification of data." But it's unclear from the redacted draft report whether the committee thinks that paper should be retracted. On another paper co-authored by Munday, in *Animal Behaviour* in 2012, the panel notes "extreme effect sizes and low to zero variance" in the data, but says, because no notebooks or files exist from the study, "there is insufficient evidence to support a finding of research misconduct." The same is true for a 2009 *Ecology Letters* paper that Munday co-authored. Munday, in an email to *Science*, says he is "shocked to hear the outcome of the investigation" but declined to elaborate.

The outcome is vindication for the group of whistleblowers, whose work had sharply divided marine ecologists, with some scientists accusing them of acting out of professional envy or to advance their own careers.

UD "did a decent investigation. I think it's one of the first universities that we've seen actually do that," says one of the whistleblowers, Fredrik Jutfelt of the Norwegian University of Science and Technology. But he is disappointed that the committee appears to only have looked at seven of the 20 papers that the team flagged as suspicious. The group had

also hoped UD would release the committee's final report and detail any sanctions against Dixon.

One of the whistleblowers, former Ph.D. student Paul Leingang, worked at Dixon's lab but started to secretly monitor her work after he became increasingly suspicious. Leingang, who left UD in February 2021 after formally accusing Dixon of misconduct—and has since left science altogether—provided the investigation with detailed notes, chat conversations, and other evidence to show she did not spend enough time on her studies to collect the data she was jotting down in her lab notebooks.

The investigative panel found Leingang's account convincing and singled him out for praise. "It is very difficult for a young scholar seeking a Ph.D. to challenge their advisor on ethical grounds," the draft report says. "The Committee believes it took great bravery for him to come forward so explicitly." ■

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

ASTRONOMY

Webb reveals early universe's galactic bounty

Star formation after the big bang appears much faster than models had forecast

By Daniel Clery

The James Webb Space Telescope has only been watching the sky for a few weeks, and it has already delivered a startling finding: tens, hundreds, maybe even 1000 times more bright galaxies in the early universe than astronomers anticipated.

"No one was expecting anything like this," says Michael Boylan-Kolchin of the University of Texas, Austin. "Galaxies are exploding out of the woodwork," says Rachel Somerville of the Flatiron Institute.

Galaxy formation models may now need a revision, as current ones hold that gas clouds should be far slower to coalesce into stars and galaxies than is suggested by Webb's galaxy-rich images of the early universe, less than 500 million years after the big bang. "This is way outside the box of what models were predicting," says Garth Illingworth of the University of California (UC), Santa Cruz.

Webb, a NASA-led orbiting observatory with contributions from the European and Canadian space agencies, began observing in late June from its vantage point 1.5 million kilometers from Earth. Much of its time so far has been devoted to projects meant to show off its capabilities, such as the Cosmic Evolution Early Release Science (CEERS) Survey. Webb is designed to delve deeper into cosmic history than its predecessor, the Hubble Space Telescope. Its 6.5-meter mirror—with six times the area of Hubble's—can catch more light from distant sources, and unlike Hubble it operates at infrared wavelengths, making Webb more sensitive to those far-away sources, whose light is stretched to longer, redder wavelengths by cosmic expansion.

Within days after Webb began observations, it spotted a candidate galaxy that appears to have been shining brightly when the universe was just 230 million years old, 1.7% of its current age, which would make it the most distant ever seen. Surveys since then have shown that object is just one of a stunning profusion of early galaxies, each small

The University of Delaware "did a decent investigation."

Fredrik Jutfelt,
Norwegian University
of Science and Technology



Members of the Cosmic Evolution Early Release Science Survey team analyze a recent galaxy-strewn James Webb Space Telescope image.

by today's standards, but more luminous than astronomers had expected.

Some researchers caution that the abundance, based on images of a small patch of sky, may be an illusion. Boylan-Kolchin wonders whether Webb just got "extra lucky" and stared into a huge clump of galaxies, denser than the rest of the early universe. That question will be resolved when CEERS broadens its scope later this year and results come in from other wide-ranging surveys.

It is also possible that astronomers are misidentifying galaxies from slightly more recent times as very early ones. Spectra are the gold standard for gauging a galaxy's age because they allow the reddening of its light to be measured precisely.

But gathering spectra from many galaxies takes time. Instead, Webb surveys so far have estimated galaxy ages from the color they appear in images—a relatively crude method. Webb's near-infrared camera filters their light into a few wide wavelength bins, giving astronomers a rough measurement of color; redder equals more distant. But dust surrounding a galaxy can fool observers, as it can absorb starlight and re-emit it at longer wavelengths, making the galaxy look redder.

Webb's early science teams have already identified a few such masquerading galaxies, as they report in several recent preprints. But if the profusion of early galaxies is real, astronomers may have to fundamentally rethink galaxy formation or the reigning cosmology.

Viewing nearby galaxies, researchers

have concluded that heat within gas clouds slows how quickly gravity would otherwise condense the matter into stars—making star formation take about 100 times longer than if gravity alone was in charge. As the first stars in a protogalaxy begin to shine, they inject more heat into the gas, pumping the brakes on further star formation. And the first stars are short-lived giants; when they explode as supernovae they heat up gas clouds even more or blast them out of a forming galaxy completely.

Studies with Hubble have shown that the rate of star formation has been relatively constant as far back as about 600 million years after the big bang, says Charlotte

Mason of the Niels Bohr Institute. But the Webb results imply that at earlier times its pace was much more rapid—as fast, Somerville suggests, as if gas clouds were collapsing freely, without any braking from heat or supernovae.

Indeed, Tommaso Treu of UC Los Angeles, who leads another Webb survey called GLASS, says his team is seeing these early galaxies "form stars like crazy." They

look, he adds, "like giant balls of star formation and nothing else."

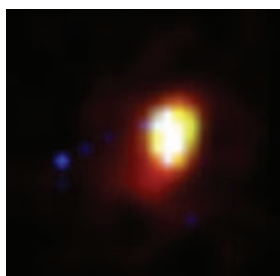
Theorists don't know if the higher density of matter and higher temperatures of the early universe might have sped star formation. Another theory is that the first stars could have formed faster because they took shape from just the primordial matter leftover from the big bang—hydrogen and helium—without the heavier elements forged by later generations of stars.

Or something may be wrong in the current understanding of how the universe evolves. The prevailing theory of cosmology, known as lambda-CDM (referring to cold dark matter), describes how, soon after the big bang, the unseen dark matter that makes up most of the stuff of the universe clumped together under its own gravity into "halos." These halos then drew in normal matter and created the conditions for it to condense into galaxies. Lambda-CDM predicts the number and size of halos that should exist in the early universe, and hence the number of galaxies. "There's not much wiggle room," Boylan-Kolchin says.

Somerville says it may be possible to tweak lambda-CDM to create something closer to what Webb is seeing. Or, she says, cosmo-logists may be forced to reassess the first moments of the big bang itself: the era of inflation, a period of rapid growth when quantum fluctuations grew into areas of higher or lower matter density—the seeds of later halos. "If inflation is wrong that could be very fundamental," she says. "But I wouldn't bet on it being that."

Having revealed the early galaxies problem, Webb may provide the data needed to answer it. So far Webb is only seeing young, hot, bright stars in the newfound early galaxies. Follow-up observations of these galaxies at longer wavelengths with Webb's midinfrared instrument or ground-based radio telescopes sensitive to submillimeter waves could reveal the gas clouds actively building stars. Those observations might help astronomers confirm that early galaxies were unusually prodigious star factories—and hold clues to how they did so.

"In 6 months we'll have a much better picture of all this," Boylan-Kolchin says. "It's a very exciting time." ■



The James Webb Space Telescope captured this galaxy at a record distance.



PRIME MOVER

Starship will be the biggest rocket ever. Are space scientists ready to take advantage of it?

By Sarah Scoles

NASA's Lunar Crater Observation and Sensing Satellite mission was brutish and short. It began on 9 October 2009, when the hull of a spent Centaur rocket stage smashed into Cabeus crater, near the south pole of the Moon, with the force of about 2 tons of TNT. And it ended minutes later, when a trailing spacecraft flew through and analyzed the lofted plume of debris before it, too, crashed. About 6% of the plume was water, presumably from ice trapped in the shadowed depths of the crater, where the temperature never rises above -173°C . The Moon, it turned out, wasn't as bone dry as the Apollo astronauts believed. "That was our first ground truth that there is water ice," says Jennifer Heldmann, a planetary scientist at NASA's Ames Research Center who worked on the mission.

Today, Heldmann wants to send another rocket to probe lunar ice—but not on a one-way trip. She has her eye on Starship, a behemoth under development by private rocket company SpaceX that would be the largest flying object the world has ever seen. With Starship, Heldmann could send 100 tons to the Moon, more than twice the lunar payload of the Saturn V, the workhorse of the Apollo missions. She dreams of delivering robotic excavators and drills and retrieving ice in freezers onboard Starship, which could return to Earth with tens of tons of cargo. By analyzing characteristics

SpaceX expects to launch the 120-meter-tall Starship (left) on its first orbital test flight in the coming months. In 2021, workers stacked a Starship upper stage on a Super Heavy booster (right).



such as the ice's isotopic composition and its depth, she could learn about its origin: how much of it came from a bombardment of comets and asteroids billions of years ago versus slow, steady implantation by the solar wind. She could also find out where the ice is abundant and pure enough to support human outposts. "It's high-priority science, and it's also critical for exploration," Heldmann says.

When SpaceX CEO Elon Musk talks up Starship, it's mostly about human exploration: Set up bases on Mars and make humans a multiplanetary species! Save civilization from extinction! But Heldmann and many others believe the heavy lifter could also radically change the way space scientists work. They could fly bigger and heavier instruments more often—and much more cheaply, if SpaceX's projections of cargo launch costs as low as \$10 per kilogram are to be believed. On Mars, they could deploy rovers not as one-offs, but in herds. Space telescopes could grow, and fleets of satellites in low-Earth orbit could become commonplace. Astronomy, planetary science, and Earth observation could all boldly go, better than they ever have before.

Of course, Starship isn't real yet. All eyes will be on a first orbital launch test, expected sometime in the coming months. Even if it is a success, no one knows whether SpaceX will be able to achieve its vision of launching the rockets daily and reusing them many times. Also unsettled is whether a market will materialize for a rocket that could put so much into orbit. But scientists need to prepare, Heldmann says. "We on the science side need to be ready to take advantage of those capabilities when they come online."

So do NASA centers such as the Jet Propulsion Laboratory, which designs and builds many space science missions, says Casey Handmer, a former JPL software engineer. In a series of provocative blog posts with titles like "Starship is still not understood," he has argued that Starship will upset the traditional way of doing space science—spending billions of dollars to make one-of-a-kind instruments that work perfectly. If the NASA centers don't find ways to take risks and make more stuff more cheaply, he says, they will find themselves displaced by companies willing to do so. "The writing is on the wall," Handmer says. "And all the NASA centers should be thinking really carefully."

ON A BALMY NIGHT in February, Musk strode onto a stage in Boca Chica, Texas, home of SpaceX's Starbase launch site, for a public update on the status of Starship. Towering behind him, bathed in lights, was the latest prototype, about 120 meters tall: the Starship vehicle, which carries people or payloads, resting on top of a Super Heavy booster. The prototype wasn't flight ready, nor had the Federal Aviation Administration (FAA) given SpaceX permission to launch it from Starbase—but it was still a spectacular backdrop, packed with coiled purpose. After welcoming the crowd of faithful rocket geeks, Musk launched into an impromptu lecture on the philosophy propelling him and his company beyond Earth. "Why build a giant, reusable rocket? Why make life multiplanetary? I think this is an incredibly important thing for the future of life itself."

SpaceX's workhorse rocket, the 70-meter-tall Falcon 9, has already shaken up the aerospace business. With that rocket, SpaceX

pioneered reusability, employing retrorockets and steerable fins to guide the first stage to a landing after it re-enters the atmosphere. Today, SpaceX routinely slaps on a fresh coat of paint and launches it again; in June, the company flew one of these "flight tested" stages a record 13th time. Another record is on the horizon: The company is on track to launch more than 50 Falcon 9 and Falcon Heavy rockets this year, or about one per week on average. The dependable reuse and rapid launch cadence are two of the reasons why SpaceX can charge \$67 million for a Falcon 9 launch, much less than its competitors. But Musk wasn't satisfied.

In 2016, at an International Astronautical Congress in Mexico, Musk sketched out plans for a rocket to colonize Mars, one he would soon be calling BFR (Big Falcon Rocket, in family-friendly terms, but you get the joke). The concept evolved into Starship, but the focus remained on affordability and reusability—making launches as dull and routine as FedEx cargo flights. The body of the rocket is stainless steel, heavier than the aluminum alloys of most rockets, but cheaper and more easily manufactured. The 33 Raptor engines crammed into the backend of Super Heavy burn methane rather than the traditional kerosene-based rocket fuels, not only because it is cheaper, but also because it could be harvested on Mars by combining carbon dioxide and water. The booster is designed to return to the launchpad after its 6-minute ride; the company believes it can be refueled and ready to relaunch in an hour. Starship is also reusable. The goal is to be able to launch each vehicle three times a day.

Once in orbit, a loaded Starship could be gassed up by a "tanker" version of the vehicle—

enabling it to take its 100 tons of payload on to the Moon or Mars. At the February event, Musk explained how a single Starship, launching three times per week, would loft more than 15,000 tons to orbit in a year—about as much as all the cargo that has been lifted in the entire history of spaceflight. Musk has claimed the price of each launch might eventually be as low as \$1 million, or \$10 per kilogram to low-Earth orbit. The only rocket close to Starship in its capabilities is NASA's Space Launch System, set to fly for the first time this month. Earlier this year, the agency's auditor found each launch would cost about \$4 billion, or nearly \$60,000 per kilogram.

Pierre Lionnet, a space economist at Euro-space, an industry trade group, is skeptical SpaceX can achieve such a low price point. It may not correctly account for the costs of developing and building the rocket, for example. "When I look at Starship, I'm looking at what seems to be a very expensive device." To achieve profitability with such high capital costs, SpaceX will have to attain its ambitious launch rates, which means it will need paying customers to soak up all that cargo capacity. SpaceX hopes to develop new markets in space mining, tourism, or other activities not yet dreamed of, but Lionnet is not so sure the heavy lifter will whet that appetite all by itself. "If you're vegetarian, and I'm offering you a burger, I can offer it at the cheapest possible price, and you don't eat it."

The debate will soon graduate beyond the theoretical. In May 2021, after several spectacularly explosive failures, a Starship upper stage flew 10 kilometers up into the atmosphere. After landing, it briefly caught fire, but the company deemed the suborbital flight a success. Since then, SpaceX has built out Starbase, constructing a launch tower that can catch returning boosters with two robotic arms the company calls "chopsticks." It has refined its rocket assembly line, which can now build four Raptor engines per week. And in June, FAA gave SpaceX approval to launch from Starbase, provided it takes steps to minimize the impact on the environment.

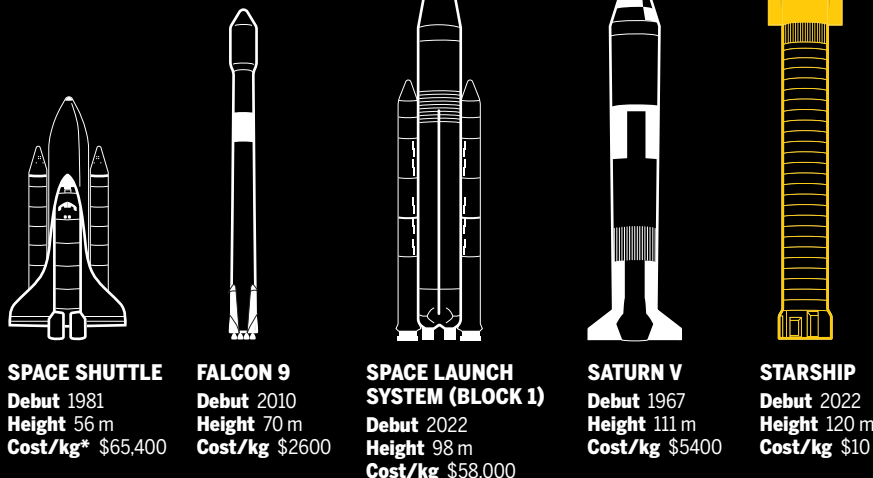
At the February event, Musk said he was confident Starship would make its first orbital attempt this year. For Musk, the sci-fi dreams are tantalizingly within reach. "Let's make this real!" he exhorted the crowd, pumping his fists.

SCIENCE HAS MOSTLY been an afterthought for Musk. But Heldmann has been surprised that, for many planetary scientists, Starship has also been an afterthought.

In 2020, she and a team of researchers and industry insiders submitted a white paper touting the benefits of Starship to the "decadal survey" in planetary science, an influential community exercise that helps

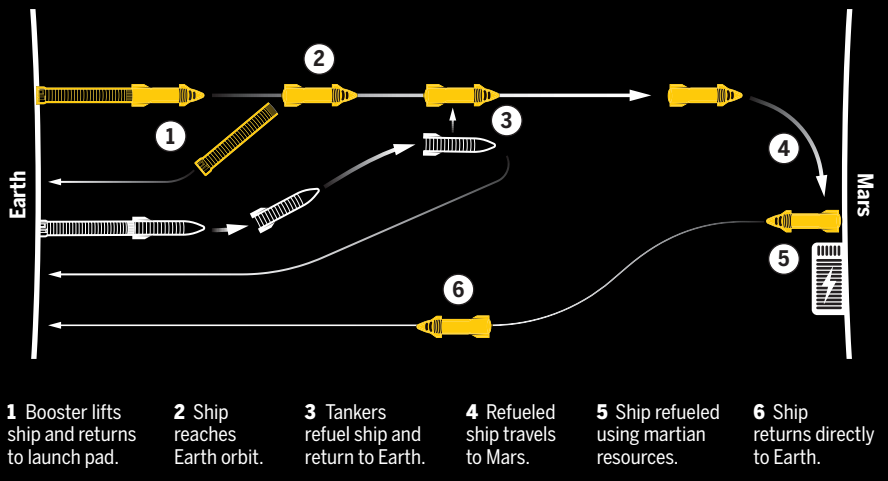
Heavy lifters

Starship would be the largest rocket the world has ever seen. It could also offer the cheapest launches. To achieve that, SpaceX will need to launch fleets of Starships daily—and it will need customers for the immense cargo capacity.



Long hauler

By refueling in orbit, each Starship can carry its full 100-ton payload to Mars. On Mars, water, carbon dioxide, and electricity can be used to make methane fuel.



NASA and Congress set long-term priorities. "It's a good time to try and get this idea in the consciousness of others," she says. Heldmann and her colleagues suggested NASA create a dedicated funding line for missions relying on Starship.

The survey embraced the ideas. In its April report, the survey committee explicitly mentioned Starship and cited ideas in Heldmann's paper. The committee recommended a funding line relevant to Starship's specs and said NASA should plan to capitalize on the rocket's potential. "Both cargo and crew flights to Mars offer significant potential

science opportunities," the committee said.

The benefits wouldn't be limited to the Moon and Mars, points out Daniel Baker, director of the Laboratory for Atmospheric and Space Physics at the University of Colorado, Boulder. For instance, Mercury, with its weak gravity, has been a tough target because of the extra fuel required to slow a spacecraft enough to get into orbit; the Sun's heat is another problem. But Starship wouldn't mind the big gas tank or the sunshade needed to keep the spacecraft from melting. Baker also envisions faster missions to the outer planets that don't require time-consuming

gravitational assists from other planets. Even farther afield, Interstellar Probe, a proposed mission to follow in the footsteps of NASA's famed Voyager mission (*Science*, 29 July, p. 460), could carry more capable instruments aboard Starship—and get a faster ride to interstellar space.

Some astronomers also have Starship in their eyes. “There’s no way to talk about it without resorting to clichés, but ‘best rocket engine ever,’ probably, by most metrics,” says David Rubin, a cosmologist at the University of Hawaii, Manoa. He wonders how much simpler the \$10 billion James Webb Space Telescope (JWST) might have been if its 6.5-meter-wide segmented mirror hadn’t had to fold up to fit on its rocket. Engineers could have built a monolithic mirror and launched it as is within the 9-meter-wide Starship fairing, which encloses a volume about half as big as a hot air balloon.

Rubin also dreams of using Starship to construct a giant telescope—say 30 meters—in space. Limbed robots could precisely lay down mirror segments on a scaffolding, forming a giant mirror that could pick out the universe’s first galaxies and look for signs of life in the atmospheres of Earth-like exoplanets. “The science gains scale really quickly as you build larger and larger telescopes,” Rubin says.

Not all astronomers are fans of SpaceX, which has already launched nearly 3000 of its Starlink internet satellites into low orbits, where sunlight glinting off them leaves streaks on the cameras of ground-based telescopes. The problems could multiply with Starship, which could launch hundreds of Starlinks at a time, enabling the company to build its planned constellation of up to 42,000 satellites even faster. SpaceX now equips the satellites with “sunshades” to reduce the reflective glare, but astronomers are still worried. “Making access to space and Earth orbit easier has a lot of benefits,” says Meredith Rawls, an astronomer at the University of Washington, Seattle, and a member of an International Astronomical Union center set up to mitigate satellite interference. “But we need to make sure that we’re doing it in a mindful way and not just having it be a Wild West disaster.”

As another reality check for the dreamers, Lionnet points out that discounted rides will only reduce the cost of missions by so much. For major scientific projects, Lionnet says, launch costs are usually between just 5% and 10% of the total price tag. For JWST, the fraction was even smaller. The typical cost for a ride on an Ariane 5 rocket, JWST’s launcher, is about \$175 million, just 2% of the mission’s total price tag. “A complex telescope will still be a complex telescope,” Lionnet says.

But cheaper launches could allow the probes themselves to be cheaper, with less need for space-rated parts that save on weight or bulk. With Starship, planetary scientists wanting to outfit a rover with a spectrometer could just buy one on the internet. Astronomers could use a glass mirror instead of a featherweight beryllium one, like JWST’s. And, Rubin says, “You should just be able to shield your way into radiation hardness,” rather than soldering circuits out of specialized materials.

Earth-observation researchers have already been down that road. For many years, remote sensing satellites were big and pricey—little different from JWST, says Aravind Ravichandran, founder of the space-industry consultancy TerraWatch Space. But about a decade ago came the “small satel-

“If thousands and thousands of tons are going to orbit, someone’s going to figure out how to put a telescope up there without NASA.”

David Rubin, University of Hawaii, Manoa

lite” revolution: Researchers shrank and standardized equipment and took advantage of rideshares on relatively cheap Falcon 9 launches and other small rockets.

Suddenly, university students were sending shoebox-size CubeSats to space. Using cheap cameras and consumer electronics, the company Planet built up a fleet of about 200 small satellites that gather daily images of all of Earth’s land. Ravichandran sees Starship making it easier to assemble ever bigger fleets—enough eyes on the planet to revisit a given spot multiple times an hour, rather than every few hours, days, or weeks. “Why can’t you do every 5 minutes? Every 10 minutes?” Ravichandran asks. Imagine, he says, what that kind of revisit rate might do for, say, tracking wildfires or floods.

HANDMER, WHO NOW WORKS as a clean-energy entrepreneur, wants astronomers and planetary scientists to adopt this sort of bold thinking. Instead of a 30-meter telescope, why not a 1000-meter one? Why not mass-produce probes that could survey dozens of asteroids? Why not fly by all the outer planets in the next decade? Or land on most planets annually?

In Handmer’s view, the problem is partly cultural: NASA engineers try to get everything right on the first try, at all costs—the vastly expensive, long-delayed JWST being a prime example. “It’s kind of like a medieval cathedral,” he says, of such flagship missions. To exploit Starship’s immense capacity, Handmer estimates NASA will need to make 100 times as much stuff for a fraction of the

usual cost. It will need to be a fast-fashion factory, not a boutique. But having worked at JPL, Handmer isn’t necessarily hopeful that will happen. “It was just not set up to mass-produce anything,” he says.

Rubin says NASA centers could get left behind by nimble companies, or privately funded scientists. “If thousands and thousands of tons are going to orbit, someone’s going to figure out how to put a telescope up there without NASA,” he says. Astrobotic, which calls itself a “lunar logistics company,” has sprung up with this sort of business model. It plans to send landers and rovers to the Moon, carrying instruments for paying customers.

Robert Manning, JPL’s chief engineer, doesn’t think the facility is quite so resistant to change. But he also questions Handmer’s vision of cheap, mass-produced probes. The equipment used at the scientific frontier is rarely standard. Every mission, with its fresh targets and questions, requires innovation. And it’s hard to take risks as a public agency, Manning says.

“We have an obligation to make sure that we are not wasteful of taxpayers’ dollars,” he says. “We can’t throw things to space frivolously and say, ‘Well, if it doesn’t work, let’s build another one.’”

In 1992, then-NASA Administrator Daniel Goldin pushed the agency to pursue a “faster, better, cheaper” approach—a mantra that was discarded later in the decade after high-profile losses of the Mars Climate Orbiter and Mars Polar Lander. If a faster, better, cheaper culture is to return, Manning says both NASA and Congress will have to bless the risk-taking—and stand up for it when things go wrong. “It’s going to be difficult to politically communicate that that’s OK for us to try it out,” Manning says.

But assuming all those issues worked themselves out, he acknowledges that frequent, low-cost, high-mass-and-volume launches could provide “an incredible opportunity for us to change how we get things done and be willing to take more risks.” JPL has already been thinking about it, he says—for example, considering how to incorporate standardized, lower cost components in NASA’s custom deep-space missions.

There’s a major asterisk on the rocket revolution that Starship heralds. “We haven’t been able to act on it yet,” Manning says, “because it’s not true yet.” When Starship makes its first orbital launch attempt, many researchers will be watching, waiting to see whether that giant silvery rocket is a vision of space science’s future, or a mirage. ■

Sarah Scoles is a journalist in Denver.

INSIGHTS

PERSPECTIVES



EVOLUTION

When less is more in the evolution of language

Did loss of vocal fold membranes typical of nonhuman primates enable human speech?

By **Harold Gouzoules**

Most studies of the biological evolution of language address human features, both cognitive and neurological, that appear absent, or less well developed, in nonhuman species. A picture emerges of what has been added to the broader vocal and communicative abilities of nonhuman primates to produce language. For example, insights into the linguistic capabilities and limitations of the great apes came from projects in which primate subjects were tutored to communicate with humans using either a form of sign language, physical tokens, or lexigrams. Extant apes appear to have most of the cognitive skills for language but lack shared intentionality and a capacity for speech (1). On

page 760 of this issue, Nishimura *et al.* (2) provide evidence that loss of vocal fold membranes that are universally present in extant nonhuman primates resulted in a stable vocal source in humans that was critical to speech evolution.

Speech and language are critically related, but not synonymous. Speech is the audible sound-based manner of language expression and humans, alone among the primates, can produce it. Sound production mechanisms in human and nonhuman primates are similar, with air from the lungs driving oscillations of vocal folds in the larynx. The acoustical energy generated then passes through the pharyngeal, oral, and nasal cavities, and emerges in a form that is governed by the filtering of specific frequencies by the vocal tract. This filtering mechanism is important to speech, generating the formant frequency patterns that convey the sounds of lan-

guage. However, this aspect of vocalization production is not distinct to humans, and all mammals appear to perceive formants in conspecific calls (3).

Identifying the adaptations that enabled speech has been challenging, and many unanswered questions remain. Notably, the core findings of Nishimura *et al.* are not even hinted at in earlier studies of the evolution of speech (3, 4). Indeed, most recent research attention has focused away from the peripheral structural dimensions of speech production to central neural control mechanisms (3, 4). Peripheral features that had received earlier interest included hypothesized speech-related specializations of the human vocal tract, in particular the descent of the tongue root into the pharynx. For years this feature was considered a distinctly human trait associated with speech evolution because it was thought to provide a greatly expanded phonetic range relative

Department of Psychology, Emory University, Atlanta, GA, USA. Email: psyhg@emory.edu

Nonhuman primates, such as this chimpanzee (*Pan troglodytes*), possess vocal fold membranes that were lost in the evolution of human speech.

to other primates (5). But it turns out that the vocal range of nonhuman primates had been previously underestimated, and the descended larynx is not distinctly human (3). Thus, the descended larynx is clearly not sufficient for speech evolution, but it might have been a necessary step.

The study of Nishimura *et al.* refocuses attention on the peripheral adaptations for speech. Using magnetic resonance imaging (MRI) and computed tomography (CT), the authors examined larynges from 29 genera and 44 species of primates and found that all the nonhuman taxa possessed a vocal membrane that is absent in humans. They also documented membrane activity during vocalizations in chimpanzees, squirrel monkeys, and rhesus macaques and used mathematical modeling to compare the acoustic effects of nonhuman vocal membrane vibration to that of vocal fold vibration alone, which is characteristic of humans. From this evidence, they propose that selection for the elimination of the laryngeal air sacs and vocal fold membranes (thin upward extensions of the vocal folds, or “vocal cords,” that are universally found in the larynx of nonhuman primates) was critical to speech evolution because this loss resulted in a stable vocal source.

Although the vocal repertoires of many nonhuman primate species include some clear, harmonically structured calls, their vocalizations frequently bifurcate to subharmonics or chaos at higher intensities. The vocal membrane that has been lost in humans is implicated in generating these nonlinear phenomena. The loss of the vocal membrane in humans therefore reduces the risk of chaotic irregularities and noise contaminating the stable vocal fold oscillations that are critical to human speech. According to the authors, this stability, combined with fine neural control involving extensive dorsal intracortical connections between motor and auditory cortices, yields the quality of vocal signal needed to rapidly encode complex linguistic information.

What makes this study especially noteworthy is that the authors conclude that the human larynx differs from that of extant nonhuman primates in that it has become simpler or less complex with the loss of the membranes. Evolutionary biology textbooks emphasize that natural selection does not inevitably produce greater complexity. For example, the Mexican cavefish (*Astyanax mexicanus*) no longer has eyes, and there is genetic, developmental, and physiological evidence for adaptive eye regression in

this species and other cave animals (6). In essence, the cave dwellers represent a “use it or lose it” evolutionary scenario. But what did humans lose with the loss of their vocal membranes (and air sacs)?

In nonhuman species, including primates, nonlinear vocal phenomena (NLP) are prominent features of communicative calls. NLP vocalizations sound harsh and have been hypothesized to lower perceived voice pitch, exaggerating the body size of the vocalizer to the listener (7). This suggests that selection can favor anatomical structures and physiological mechanisms that promote vocalization instability and support complex subharmonics and chaos. Although the descended larynx in humans likely offers some compensation with respect to lowering pitch, the loss of vocal membranes and the concomitant abatement of NLP would seem to reduce this dimension of human nonverbal vocal communication. Despite the absence of vocal fold membranes, humans retain the ability to produce vocalizations (often intentionally and voluntarily) with nonlinear and chaotic features, such as screams. Indeed, human screams are evolutionarily conserved acoustically and resemble those of other mammalian species (8, 9). Humans perceive and respond to NLP in vocalizations in ways that are consistent with animal communication studies. NLP lowers perceived voice pitch, increases acoustic roughness, and causes vocalizers to sound larger, more formidable, and more aggressive (10). With the loss of vocal fold membranes, humans have apparently reduced the production of NLP in speech, but still retain the ability to generate harsh sounds in nonverbal vocalizations such as screams and roars, a complexity of vocal control that has yet to be adequately investigated. If there has been a loss of a tendency or capacity to exaggerate size vocally in humans, there are other means to offset them: Humans can, of course, speak softly and carry a big stick. ■

REFERENCES AND NOTES

1. M. Tomasello, “What did we learn from the ape language studies?” in *Bonobos: Unique in Mind, Brain, and Behavior*, B. Hare, S. Yamamoto, Eds. (Oxford Univ. Press, 2017).
2. T. Nishimura *et al.*, *Science* **377**, 760 (2022).
3. W. T. Fitch, *Annu. Rev. Linguist.* **4**, 255 (2018).
4. E. D. Jarvis, *Science* **366**, 50 (2019).
5. P. H. Lieberman, E. S. Crelin, D. H. Klatt, *Am. Anthropol.* **74**, 287 (1972).
6. J. Krishnan, N. Rohner, *Philos. Trans. R. Soc. B* **372**, 20150487 (2017).
7. A. Anikin, K. Pisanski, M. Massenet, D. Reby, *Proc. Biol. Sci.* **288**, 20210872 (2021).
8. J. W. M. Engelberg, J. W. Schwartz, H. Gouzoules, *PeerJ* **9**, e10990 (2021).
9. S. Frühholz, J. Dietziker, M. Staib, W. Trost, *PLOS Biol.* **19**, e3000751 (2021).
10. A. Anikin, D. Reby, *Ethol. Ecol. Evol.* **34**, 303 (2022).

MEDICAL DEVICES

Ultra-sounding out a technique that sticks

Ultrasound can be used to promote the physical interlocking of adhesives and tissues

By **Julien Es Sayed** and **Marleen Kamperman**

Advances in robotics, surgical instrumentation, and real-time imaging inside the body have made it possible to switch various surgeries to minimally invasive ones. However, for many surgical procedures that require reconnecting tissue or attaching prosthetic materials, a minimally invasive way to perform them does not yet exist. Stapling and suturing are harsh techniques that can cause local tissue damage from deep piercing. In some cases, their removal after wound healing can also inflict damage. Tissue adhesives could potentially reduce the harm of surgical procedures, decrease operative times, and improve the surgical outcomes. Although

“Tissue adhesives could potentially reduce the harm of surgical procedures, decrease operative times, and improve the surgical outcomes.”

promising approaches have been developed, tissue adhesives for wound closure are now mainly used in combination with sutures and staples. On page 751 of this issue, Ma *et al.* (1) introduce an ultrasound-mediated approach to enhance the adhesive performance of soft hydrogels on various tissues.

Researchers have long been searching for adhesives that can withstand the environment of the human body (2). It is difficult to “glue” things together when they are wet and slippery. Water acts as a boundary layer between adhesive and tissue, preventing bonds

Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherland. Email: marleen.kamperman@rug.nl

10.1126/science.add6331

from forming. Mucous membranes on the surfaces of internal organs, such as the gastrointestinal tract and lungs, also act as a barrier for adhesives (3). Some tissues also change shape constantly, such as the beating heart or the breathing diaphragm, making adhesive-tissue bonding difficult to attain. One approach to achieve adhesion is to chemically modify the adhesive with functional groups such as amines, carboxylic acids, and thiols. Peptide coupling reactions can then create a strong bond between the modified adhesive and tissue proteins (4). Challenges in developing such an adhesive include timing the chemical reaction in the time window of the surgical procedure; tuning the properties of the adhesive to the chemistry of the specific tissue; preventing inflammation triggered by unreacted residues, degradation, or reaction by-products; and ensuring that the generated adhesive joint can withstand stretching (5). Ma *et al.* address the latter challenge by increasing the toughness of the adhesive joint.

Toughness of an adhesive joint is characterized by the energy needed to break it. The more an adhesive joint can stretch and absorb energy, the harder it is to break. Stretching reduces the force per unit area and consequently decreases the likelihood that the adhesive joint will fail (6). An effective way to increase toughness is by interlocking polymer chains into the substrate (7). This design principle takes advantage of the roughness as well as the relative permeability of biological tissues and hydrogels in an approach called topological adhesion (8, 9). This technique relies on first diffusing polymer chains into the surface of the tissue substrate and then inducing the formation of strong interchain interactions between the polymer chains and the substrate by manipulating the environmental conditions. The average size and concentration of the polymer chains, the affinity for the solvent as well as for the substrate, and the interactions between polymer chains are all important parameters for achieving effective interlocking. For example, a pH-sensitive natural polymer such as chitosan can diffuse into a biological tissue at low pH and form strong hydrogen bonds at

higher pH (9). In this approach, water is no longer considered detrimental to adhesion but instead helps facilitate polymer diffusion into the tissue. Still, the relatively slow diffusion of polymer chains inside the tissue limits the speed of the adhesion process.

Building on these considerations, Ma *et al.* advance the topological adhesion concept by using ultrasound to propel polymer chains deeper and faster into the biological tissue. These sound waves prepare biological tissue to interact with a hydrogel that is placed on top (see the figure). Before establishing contact with the hydrogel, a natural polymer solution (e.g., chitosan, cellulose nanocrystals, or gelatin) is spread on the tissue, and the contact area is then treated with

ultrasound. The ultrasound creates cavitation microbubbles that propel the primer molecules into the tissue, resulting in strong mechanical interlocking. In addition to this entanglement, the primer polymer chains also form either electrostatic, hydrophobic, or hydrogen bonds with the surrounding tissue. This combination of topological and physical affinity results in strong adhesion between hydrogel and tissue without the need for a chemical reaction.

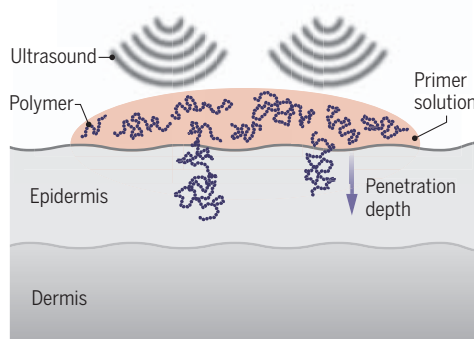
The concept of using ultrasound to enhance the permeability of tissues has been successfully applied for drug delivery systems (10), which may make the translational process to the clinic faster and easier. In the work of Ma *et al.*, and that of many other research groups, it becomes clear that large variations between tissues lead to substantial differences in the effectiveness of adhesion strategies. Future research should focus on tissue-specific adhesive development. Better insight is needed into the mechanical characteristics and surface chemistry of different tissues, including damaged and diseased tissues, because this is where adhesives will be used most often. In addition, specific adhesives will need to match tissues in such a way as to actively promote repair by means of cell or drug delivery. In this way, tissue-specific bioactive adhesives can help the healing process by stimulating the patient's body to use its own systems to recreate cells and rebuild tissues and organs. ■

Ultrasound-enhanced topological adhesion to tissues

Tough adhesion between tissues and hydrogels has been achieved by stitching them together with anchored dangling polymer chains. Ma *et al.* show that ultrasound treatment enhances the interlocking of the chains in the tissue.

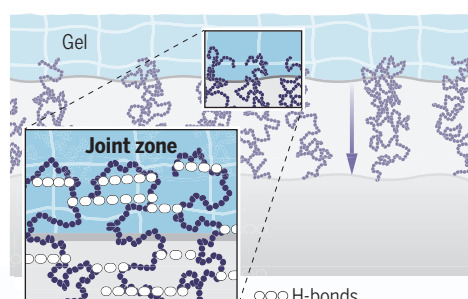
Chain penetration or interlocking

Primer polymer chains are propelled into the tissue by ultrasound.



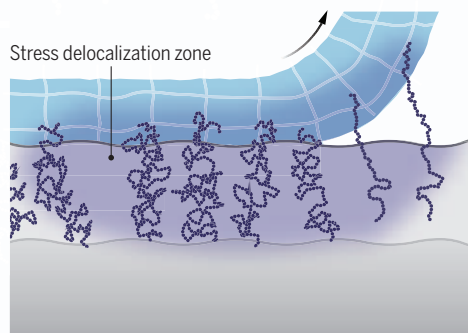
Hydrogel bonding

A hydrogel is applied to the tissue. The polymer chains interlock into the hydrogel, and hydrogen bonds form between the polymer chains.



Debonding mechanism

As the hydrogel is removed, the interlocked polymer chains anchored in the tissue are able to stretch and deform, which drastically increases the toughness of the adhesive joint by spreading the stress over a larger volume.



REFERENCES AND NOTES

1. Z. Ma *et al.*, *Science* **377**, 751 (2022).
2. Y.J. Z. Young, P.B. Medawar, *Lancet* **236**, 126 (1940).
3. C. Karamaoun, B. Sobac, B. Mauroy, A. Van Muylem, B. Haut, *PLOS ONE* **13**, e0199319 (2018).
4. C. Cui, W. Liu, *Prog. Polym. Sci.* **116**, 101388 (2021).
5. P.A. Leggat, D.R. Smith, U. Kedjarune, *ANZ J. Surg.* **77**, 209 (2007).
6. C. Creton, M. Ciccotti, *Rep. Prog. Phys.* **79**, 046601 (2016).
7. L. Leger, C. Creton, *Philos. Trans. R. Soc. London Ser. A* **366**, 1425 (2008).
8. J. Li *et al.*, *Science* **357**, 378 (2017).
9. J. Yang, R. Bai, Z. Suo, *Adv. Mater.* **30**, e1800671 (2018).
10. M. R. Prausnitz, R. Langer, *Nat. Biotechnol.* **26**, 1261 (2008).

ACKNOWLEDGMENTS

Support for this work was provided by the European Research Council (European Union's Horizon 2020 research and innovation program, consolidator grant agreement no. 864982).

10.1126/science.abq7021

Perfluorocubane—a tiny electron guzzler

Perfluorination gives cubane the capacity to host an extra electron in its inner structure

By Marie Pierre Krafft¹ and Jean G. Riess²

High symmetry arouses a sense of perfection and perennity. The highly symmetrical cubane molecule (C_8H_8), despite its improbable bonding angles, has been synthesized using a complex, multistep approach (1). Unexpectedly, the structurally strained and thermodynamically unstable molecule proved to be inert. In the molecular world, symmetry impedes reactivity because no weak spot is available for intrusion. Achieving perfluorination—that is, replacing all hydrogens of cubane with much larger fluorine atoms—was a challenge, with the reward of creating a perfect nest for hosting an additional electron, thus generating a radical anion. On page 756 of this issue, Sugiyama *et al.* (2) report the synthesis of perfluorocubane, C_8F_8 , contributing to the understanding of electron-accepting perfluorinated polyhedral structures.

Since its inception (1), cubane has inspired many investigations, derivatizations, functionalizations, and suggestions for potential uses, particularly in medicine and materials science (3–9). Perfluorination drastically changes the structure and electron charge distribution of cubane. Compared with hydrogen, fluorine—with its nine protons and nine electrons—is much larger and has a higher electron density, ionization potential, electron affinity, and electronegativity as well as lower polarizability. Perfluorocarbons are among the most thermally stable and chemically inert materials ever synthesized. Perfluorination generates new properties or enhances known ones to levels often unattainable otherwise (10). Specific features include combined hydrophobicity and lipophobicity, combined oxygen-carrying capacity and biocompatibility, extreme surface activity, and low-temperature resilience (10). Cautious use is needed, however, because perfluorochemicals can raise environmental problems, particularly as a result of persistence and accumulation in living organisms (11). Regarding perfluorocubane, the predicted property of netting an electron in a carbon polyhedron (12) offered a tantalizing additional challenge.

To synthesize C_8F_8 , Sugiyama *et al.* selected direct fluorination—i.e., replacement of the hydrogens with fluorines using fluorine molecules (F_2). The reaction of F_2 with organic material is highly exothermic and provokes profuse carbon-carbon (C-C) bond breaking. Taming such explosive reactions is not straightforward. Direct fluorination was originally done by diluting F_2 in an inert gas, typically nitrogen, at low temperature in a chlorofluorocarbon solvent (13). The technique was later improved by attaching a fluorinated ligand to the molecule that is to be perfluorinated (14). Heptafluorocubane (C_8HF_7) was thus obtained, and the remaining hydrogen was replaced by a fluorine using standard monofluorination reactions to produce C_8F_8 .

Notably, the experiments of Sugiyama *et al.* showed little C-C bond length change upon perfluorination. This implies that the C-C bond shortening expected from highly electronegative fluorine substitution is coun-

“Perfluorination causes sweeping steric and electronic alterations, entailing profound changes in properties and reactivity...”

tered by the increased spatial needs of the fluorine atoms. Their data also revealed a rare directional noncovalent carbon bond between a fluorine atom and the nearest perfluorocubane molecule. Most notably, their data confirmed the prediction that the overlap (so-called sigma stellation) of all eight antibonding orbitals of the C-F bonds form an extended star-shaped molecular orbital. The result of this sigma stellation is a pronounced electron affinity and the possibility to host an electron inside the eight-carbon cubic frame, generating the radical anion $C_8F_8^{\cdot-}$ (12). Distinctively, the spin density of the hosted electron is chiefly located inside the cubane frame, in contrast to the common electron acceptors that store extra electrons on their surface. The radical anion is, however, unstable.

Perfluorination causes sweeping steric and electronic alterations, entailing profound changes in properties and reactivity, making forecasts hazardous. Only further

experimentation and modeling studies will tell what properties and reactivity perfluorocubane has. Future studies may investigate the synthesis of other functional fluorinated cubane derivatives, devising means of introducing such building blocks into molecules and polymers and assessing their chemistry elicited by the extra electron hosted inside the structure. The rare carbon-bonding donor capacity could be explored using cyanide and similarly elongated anions. The use of functional perfluorocubane building blocks as linkers, spacers, or appendages in polymers may also be rewarding.

The inspiring asset of perfluorocubane and derivatives lies in their distinctive electron hosting capacity and redox activity. Perfluorocubane-derived compounds may offer an opportunity to experimentally determine what differentiates an internally netted electron from an externally located one. New means of grafting a small, space-filling electron-sequestering building block into molecules would be valuable for developing redox-active, functional electron-carrying and -transferring materials, which could find applications in organic conducting and semiconducting materials, magnetic frameworks, or photothermal materials. Currently, the availability and cost of functional perfluorocubane-derived reagents on an appropriate scale may impede extensive investigations. The findings of Sugiyama *et al.* also warrant the investigation of the perfluorination of further polyhedral molecular structures—especially those susceptible to achieving sigma stellation—as well as exploration of their electron chemistry. ■

REFERENCES AND NOTES

1. P. E. Eaton, T. W. Cole, *J. Am. Chem. Soc.* **86**, 962 (1964).
2. M. Sugiyama *et al.*, *Science* **377**, 756 (2022).
3. B. A. Chalmers *et al.*, *Angew. Chem. Int. Ed.* **55**, 3580 (2016).
4. J. Wlochal *et al.*, *Org. Lett.* **16**, 4094 (2014).
5. K. F. Biegasiwicz, J. R. Griffiths, G. P. Savage, J. Tsanaktsidis, R. Priefer, *Chem. Rev.* **115**, 6719 (2015).
6. D. E. Collin *et al.*, *Chem. Eur. J.* **26**, 374 (2020).
7. T. A. Reekie *et al.*, *J. Med. Chem.* **62**, 1078 (2019).
8. S. D. Houston *et al.*, *Org. Biomol. Chem.* **17**, 6790 (2019).
9. H.-T. Huang *et al.*, *J. Am. Chem. Soc.* **142**, 17944 (2020).
10. M. P. Krafft, J. G. Riess, *Chemosphere* **129**, 4 (2015).
11. M. P. Krafft, J. G. Riess, *Curr. Opin. Colloid Interface Sci.* **20**, 192 (2015).
12. K. K. Irikura, *J. Phys. Chem. A* **112**, 983 (2008).
13. R. J. Lagow, in *Fluorine Chemistry at the Millennium: Fascinated by Fluorine*, R. E. Banks, Ed. (Elsevier, 2000), pp. 283–296.
14. T. Okazoe, *J. Fluor. Chem.* **174**, 120 (2015).

¹University of Strasbourg, Institut Charles Sadron (CNRS), 67034 Strasbourg, France. ²Harangoutte Institute, 68160 Ste-Croix-aux-Mines, France. Email: krafft@unistra.fr

ACTIVE MATTER

A soft active matter that can climb walls

Mechanical activity of an active fluid can be used to control its dynamics at the boundaries

By Jérémie Palacci

If you mix fruit syrups with alcohol to make a schnapps, the two liquids will remain perfectly blended forever. But if you mix oil with vinegar to make a vinaigrette, the oil and vinegar will soon separate back into their previous selves. Such liquid-liquid phase separation is a thermodynamically driven phenomenon and plays an important role in many biological processes (1). Although energy injection at the macroscale can reverse the phase separation—a strong shake is the normal response to a separated vinaigrette—little is known about the effect of energy added at the microscopic level on phase separation. This fundamental question has deep ramifications, notably in biology, because active processes also make the interior of a living cell different from a dead one. On page 768 of this issue, Adkins *et al.* (2) examine how mechanical activity at the microscopic scale affects liquid-liquid phase separation and allows liquids to climb surfaces.

To investigate the phenomenon of liquid-liquid phase separation, the authors used a mixture of the polymer poly(ethylene glycol) (PEG) with dextran, a polysaccharide derived from glucose. Because both PEG and dextran are biocompatible, they can be combined with functional protein nanomotors—kinesins. PEG-dextran can phase separate into PEG-rich drops in a dextran-rich phase (3). In addition, PEG and dextran exhibit ultralow surface tension, in which their phase boundaries are a thousand times flimsier than in an oil-water mixture and deform with minimal mechanical stress.

microtubules are added to a PEG-dextran fluid undergoing phase separation, the kinesin motors and microtubules exclusively reside in the dextran phase, whereas the PEG phase remains passive.

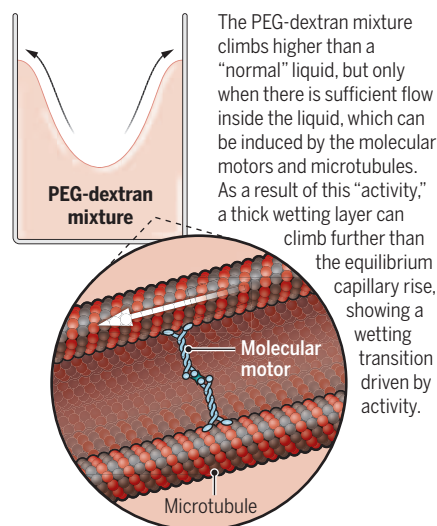
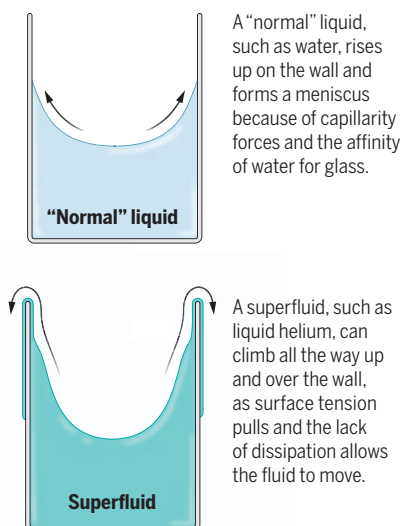
The authors observed the behavior of the passive drops of PEG immersed in the active fluid powered by kinesins motors. At moderate activity of the fluid, the growth of the PEG droplets during phase separation appeared faster than in a passive system. The activity in the dextran phase enhanced the movement of the PEG droplets and favored the droplets to collide and merge with each other. In effect, the system behaved as a hot thermal bath. Unexpectedly, at high activity, the PEG droplets stopped growing, and the phase separation arrested. At this point, the fluid mixture reached a steady state, in which larger droplets shed off smaller droplets at the same rate that smaller droplets merged to form large droplets. The authors posit that deformation of the boundary between the two phases is caused by active stresses of the fluid competing with the surface tension between the two phases of PEG and dextran. This competition can be observed as interfacial waves travel along the PEG-dextran boundary, and the interface shows giant fluctuations in shape. This model is consistent with simulations based on a hydrodynamic model, which agree with the observations and furthermore allow the estimation of active stresses, a fundamental but elusive parameter of the mechanics of active systems.

On the basis of their results, the authors then provide an example of how active stresses of liquids may be leveraged to create a “super” fluid, in which internal forces allow the active fluids to climb against gravity—almost like a superfluid but with broadly different physical reasons (see the figure). Such a liquid can be used as an experimental platform for active, shape-shifting soft materials.

Another notable feature of this study is the connection established between thermodynamics and active matter. Active matter is a research area that focuses on the emerging and dynamic properties of a large number of active agents, such as a flock of starlings or a colony of bacteria. These systems can be found across all length and time scales in nature. The field has expe-

A “super” fluid, but not a superfluid

When experimenting with a polymer poly(ethylene glycol) (PEG)–dextran mixture with molecular motors and microtubules, Adkins *et al.* discovered a nonequilibrium wetting transition that allows the active fluid to climb walls.



Adkins *et al.* investigated how energy injected at the microscopic scale affects the phase separation between PEG and dextran. To do so, they made the fluid “active”—that is, one that contains individual constituents that exert forces—by adding kinesin molecular motors and microtubules to the fluid, a combination that has been done before (4). Powered by the kinesin molecular motors, the microtubules continuously slide past each other, leading to chaotic flows inside the fluid. The strength of those flows characterizes the activity of the fluid. When the kinesin and

rienced a surge in the past decade, with important experimental progress both on the synthetic (5–8) and biology side (9, 10) and conceptual breakthroughs highlighting flocking behavior (8, 11), nonequilibrium phase separation (12), and the role of topology and topological defects in controlling the dynamics of active systems (13–15).

The work by Adkins *et al.* opens a new horizon for active matter, in which an active phase can exert forces and generate flows to control the behavior of a passive system. Active matter constitutes a potent toolbox to devise new phases of matter, leveraging the interplay of an active phase onto a “passive one.” The study of these hybrid materials addresses fundamental questions, such as the role of activity to control the phase diagram or to promote phases of matter that have no conventional equivalent. The effect of the activity on the thermodynamics of liquid-liquid phase separation also remains an open question.

“This work... will be inspirational for others to investigate the rich interplay between active and passive phases....”

Active stresses can couple to passive constituents and control their organization. Its relevance in biology as well as the design of adaptable soft materials and machines remains to be explored. A considerable challenge is to establish experimental platforms to devise and study those phases of matter. This work is an important step in that direction. It will be inspirational for others to investigate the rich interplay between active and passive phases—an uncharted territory. ■

REFERENCES AND NOTES

1. A. A. Hyman, C. A. Weber, F. Jülicher, *Annu. Rev. Cell Dev. Biol.* **30**, 39 (2014).
2. R. Adkins *et al.*, *Science* **377**, 768 (2022).
3. Y. Liu, R. Lipowsky, R. Dimova, *Langmuir* **28**, 3831 (2012).
4. T. Sanchez *et al.*, *Nature* **491**, 431 (2012).
5. C. Bechinger *et al.*, *Rev. Mod. Phys.* **88**, 045006 (2016).
6. J. F. Boudet *et al.*, *Sci. Robot.* **6**, eabd0272 (2021).
7. D. Needleman, Z. Dogic, *Nat. Rev. Mater.* **2**, 17048 (2017).
8. A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, D. Bartolo, *Nature* **503**, 95 (2013).
9. V. Schaller, C. Weber, C. Semmrich, E. Frey, A. R. Bausch, *Nature* **467**, 73 (2010).
10. H. H. Wensink *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14308 (2012).
11. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, *Phys. Rev. Lett.* **75**, 1226 (1995).
12. M. E. Cates, J. Tailleur, *Annu. Rev. Condens. Matter Phys.* **6**, 219 (2015).
13. G. Duclos *et al.*, *Science* **367**, 1120 (2020).
14. V. Soni *et al.*, *Nat. Phys.* **15**, 1188 (2019).
15. S. Shankar *et al.*, *at. Rev. Phys.* **4**, 380 (2022).

10.1126/science.adc9202

SYNTHETIC BIOLOGY

Synthetic gene circuits take root

Complex spatial patterns of gene expression are engineered in plants to modulate root morphology

By **Simon Alamos**^{1,2,3} and **Patrick M. Shih**^{1,2,3,4}

A long-standing aim of synthetic biology has been to engineer genetic circuits that are able to confer prescribed spatiotemporal patterns of gene expression. This feat would not only unlock the next tier in the technology tree of translational biology but would also constitute a powerful demonstration of a predictive and quantitative basic understanding of genetic regulation in higher eukaryotes. On page 747 of this issue, Brophy *et al.* (1) report a strategy to engineer spatial transcriptional patterns in the root of the model plant *Arabidopsis thaliana* to alter its morphology. This effort constitutes a milestone in the genetic engineering of a whole, fully developed multicellular organism and points to the challenges ahead.

Early studies of the cis and trans regulation of plant and animal genes showed that spatiotemporal patterns of transcription follow a combinatorial logic (2, 3). Whether a gene is expressed in a cell type depends on the combination of DNA binding sites in its regulatory region and whether these sites are occupied by activator or repressor transcription factors (TFs) present in these cells. These regulatory combinations can be described as boolean logical operations. For example, a gene may be transcribed only if its activator is present while its repressor is absent, a logic operation known as a “NIMPLY gate.” It also became clear that TF proteins themselves are modular, with DNA binding and regulatory activity being physically and functionally separated (4). Together, these findings inspired what might be considered the earliest synthetic biology tools (5). It became possible to fuse heterologous DNA binding domains to transcription activation domains and express these synthetic TFs in specific cell types. Genes under the control of synthetic promoters that bear binding sites to these synthetic TFs could thus be expressed in specific cells and/or under defined conditions (5).

Although these tools proved invaluable for basic biology, the spatiotemporal control of gene expression in whole organisms remained modest. Plant genetic engineering has barely explored the combinatorial logic of cis and trans elements that underpin plant developmental complexity. Synthetic efforts in the *Drosophila melanogaster* embryo—a prime model for animal genetic circuit interrogation—have so far been limited to simple activation, a long way from the choreography of repressors and activators that establish the fly body plan (6). Thus, the repertoire of transcriptional logical operations engineered in complex eukaryotes is still limited, in contrast to microbes and cultured cells (7). Indeed, there persists a humbling gap between our understanding of the endogenous genetic circuitry that controls gene expression patterns in animals and plants and our ability (or lack thereof) to engineer these patterns (8).

These limitations stem in part from a dearth of “genetic parts” that can perform logical operations in whole organisms. This is not entirely unexpected because engineering synthetic TFs and the synthetic promoters responsive to them can take multiple iterations of testing and optimization, whereas the turnaround time to generate a transgenic plant or animal is several months. To close this gap, Brophy *et al.* turned to a model system that is somewhat intermediate between cell culture and a transgenic plant. It is possible to transiently express multiple transgenes in the leaves of the tobacco relative *Nicotiana benthamiana* and measure the circuit performance using green fluorescent protein (GFP) fluorescence in just 2 days (see the figure). Using this platform, they developed a library of synthetic TFs based on bacterial DNA binding domains and eukaryotic activation and repression domains. Combined with synthetic promoters that bear bacterial DNA binding sites, these form an impressive collection of parts that can perform all major logical operations.

Given the completeness of this toolbox, it should, in principle, be possible to use cell type-specific promoters to drive synthetic TFs and create spatially embedded logical operations in a cell type-specific manner. Of all multicellular model systems, the *A. thaliana* root has emerged as one of the premier systems for this type of experiment

¹Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. ²Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA, USA. ³Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁴Innovative Genomics Institute, University of California, Berkeley, CA, USA. Email: pmshih@berkeley.edu

owing to its relative simplicity and accessibility for imaging. Moreover, cell type-specific messenger RNA (mRNA) patterns in the root have been thoroughly characterized and have resulted in a collection of tissue-specific promoters. In contrast to the findings of Brophy *et al.*, which were obtained using transient expression in *N. benthamiana*, synthetic circuits designed to perform spatial logical operations in stably transformed *A. thaliana* tended to deviate from their expected behavior. For example, it was found that the activity of some synthetic TFs varied greatly across cell types depending on the promoter under which they were expressed. The authors carefully documented these negative findings and proposed concrete strategies to overcome them. The field would benefit greatly from such transparency. Chief among the debugging strategies was simply swapping synthetic cis and trans elements, which high-

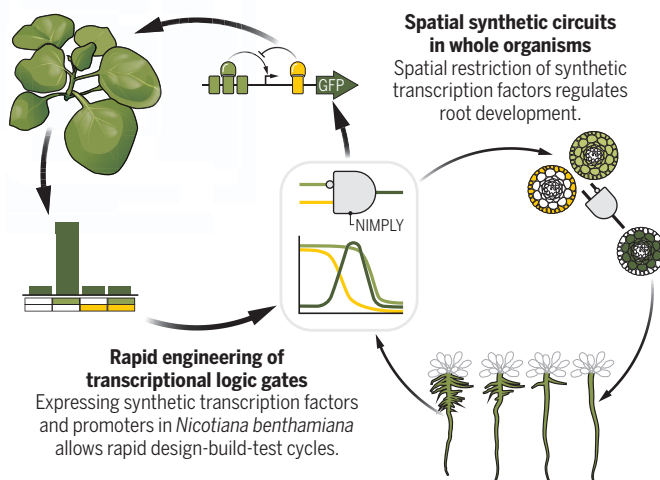
behaviors that are only beginning to be appreciated, much less understood. It is likely that a deeper basic understanding of multicellularity and development will be necessary to leverage synthetic biology for translational applications in plants and animals.

As the term suggests, translational research depends on prior basic biology. Translating implies a sequential hierarchy: discover first, engineer second. Synthetic biology flips this script, arguing that this sequential hierarchy can be productively inverted (9). This “build to understand” ethos could be the path forward. Synthetic biology approaches tailored to embrace the complexity of developing organisms are well-positioned to provide sophisticated and quantitative tools to dissect the inner workings of genetic networks specifically in an organismal context. Discovering the intricacies of genetic networks in whole organisms would in turn enable more advanced synthetic bi-

Synthetic transcriptional logic gates

Logic gates such as NIMPLY can be implemented at the transcriptional level to engineer biological computation. Brophy *et al.* developed a library of different synthetic transcription factors and promoters. These tools were optimized and then used to drive root development in *Arabidopsis thaliana*.

GFP, green fluorescent protein



lights the importance of a diverse toolbox. These results underscore that those genetic tools that perform well at a lower level of biological complexity are not guaranteed to work in a more complex context. This is a cautionary tale for those in the broader synthetic biology community who are seeking to engineer multicellular organisms.

Models such as mammalian cell culture or transient expression in *N. benthamiana* allow for a speed and throughput that is unthinkable in whole organisms. These simpler systems have been pivotal in the development of molecular toolboxes for bioengineers such as the one introduced by Brophy *et al.* However, implementing these tools in a transgenic animal or plant line is far from trivial. Why this should be the case is an interesting question that the field will have to grapple with in the future. Cells in intact tissues give rise to emergent collective

biology applications, creating a virtuous cycle. Moving forward, the *A. thaliana* root could serve as an ideal model to better understand the principles of natural and engineered gene circuits in an organismal context. The lessons learned from Brophy *et al.* are key to the future success of implementing and translating combinatorial circuits not just into plants but more broadly into other complex biological systems. ■

REFERENCES AND NOTES

1. J. A. N. Brophy *et al.*, *Science* **377**, 747 (2022).
2. D. Stanojević *et al.*, *Nature* **341**, 331 (1989).
3. P. N. Benfey, N.-H. Chua, *Science* **250**, 959 (1990).
4. L. Keegan *et al.*, *Science* **231**, 699 (1986).
5. J. Haseloff, *Methods Cell Biol.* **58**, 139 (1999).
6. J. Crocker *et al.*, *Cell Rep.* **18**, 287 (2017).
7. S. Ausländer, M. Fussenegger, *Cold Spring Harb. Perspect. Biol.* **8**, a023895 (2016).
8. B. J. Vincent *et al.*, *Integr. Biol.* **8**, 475 (2016).
9. M. Elowitz, W. A. Lim, *Nature* **468**, 889 (2010).

MOLECULAR BIOLOGY

Activating translation with phase separation

Ribonucleoprotein granules allow activation of translation to complete mouse spermatogenesis

By Anne Ramat and Martine Simonelig

Phase separation allows cellular organization through the coordination of biochemical reactions into membraneless biomolecular condensates (1). Studies have focused on the biophysical properties and dynamics of biomolecular condensates. Now, a key issue to address is the biological functions of these condensates: Can phase separation induce specific functions and, if so, how (2, 3)? In the cytoplasm, messenger RNAs (mRNAs) and RNA binding proteins assemble into biomolecular condensates called ribonucleoprotein (RNP) granules that are intimately linked to RNA biology. RNP granules have been associated with mRNA storage and translational repression. On page 727 of this issue, Kang *et al.* (4) report translational activation as a previously unknown function of biomolecular condensates, showing that phase separation of the RNA binding protein FXR1 (fragile X mental retardation syndrome-related protein 1) is required for completion of spermatogenesis in mice.

Since the pioneering study that applied a concept well known in polymer chemistry to understand the dynamics of *Caenorhabditis elegans* germline P granules (5), liquid-liquid phase separation has become a flourishing topic in cell biology. Phase separation occurs when molecules locally reach a concentration threshold at which they demix from the surrounding milieu, leading to the formation of biomolecular condensates. Demixing depends on a multitude of multivalent interactions within and between nucleic acids and proteins that contain intrinsically dis-

Institute of Human Genetics, Université Montpellier, CNRS, Montpellier, France. Email: martine.simonelig@igh.cnrs.fr

10.1126/science.add6805

ordered regions, low-complexity domains, or multiple interaction domains (6).

Cytoplasmic RNP granules have been linked to mRNA storage and translational repression. This is the case for stress granules that form upon environmental stress and store mRNAs in a translationally repressed state until stress is alleviated (7). Germ granules also store maternal mRNAs in a repressed state for prolonged periods of time (8). Germ granules store mRNAs that drive germ cell development. Therefore, an outstanding question is how translation of these mRNAs is activated. A recent study in *C. elegans* showed that translation takes place after the exit of mRNAs from germ granules (9). However, other studies hinted at a link between germ granules and translation activation. In both mice and *Drosophila melanogaster*, the PIWI proteins (core components of germ granules) were shown

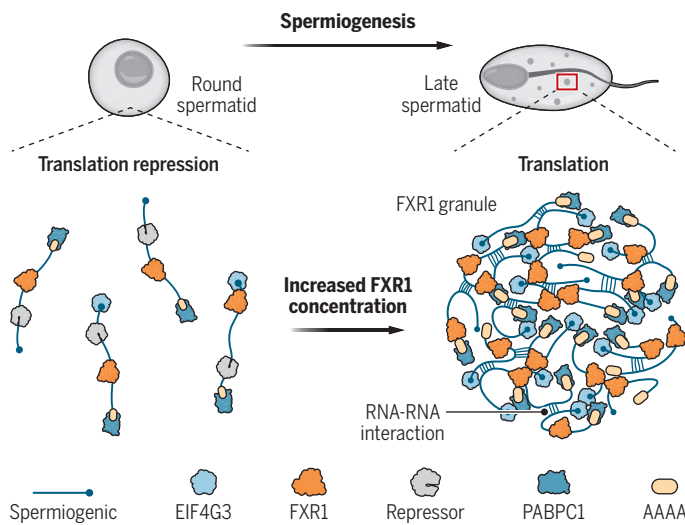
to activate translation by recruiting translation initiation factors to germ granules (10, 11). Moreover, using a technique that records translation in living cells, it was shown that mRNAs can be actively translated in stress granules (12). Thus, translation can occur in RNP granules, but our understanding of this process is limited.

Spermatogenesis in mammals is divided into several steps, the last one being spermiogenesis. During spermiogenesis, haploid spermatids are transformed into spermatozoa through a series of morphologic changes. Transcription gradually stops during this process; therefore, spermiogenic mRNAs that encode factors involved in spermatid differentiation are produced earlier and stored in a translationally repressed state until their timely translation. The molecular mechanisms underlying this switch from translational repression to activation have remained a mystery.

Analyzing the proteome associated with spermiogenic mRNAs that undergo active translation, Kang *et al.* identified FXR1. Phase separation of the related protein, fragile X mental retardation protein (FMRP), has been observed using in vitro systems and linked to translational repression (13, 14). Kang *et al.* found that FXR1 binds actively translated spermiogenic mRNAs, and *Fxr1* germ cell-specific deletion revealed that FXR1-mediated translational activation is important for late spermatid differentiation and mouse fertility. Prompted by the

A molecular function of phase separation

Cytoplasmic ribonucleoprotein (RNP) granules are classically associated with translational repression. But, in mouse spermatids, phase separation of FXR1 (fragile X mental retardation syndrome–related protein 1) is required for translational activation of spermiogenic messenger RNAs (mRNAs). FXR1 associates with the translation factors EIF4G3 (eukaryotic translation factor 4γ3) and PABPC1 [poly(A) binding protein C1] and recruits them to granules in late spermatids.



presence of intrinsically disordered regions in FXR1, as in FMRP, the authors found that FXR1 levels increase in late spermatids when spermiogenic mRNAs are translated, correlating with the formation of foci, which the authors established were phase-separated condensates (see the figure).

Kang *et al.* managed to identify a point mutation in one of the KH RNA binding domains of FXR1 that prevents FXR1 phase separation while preserving its RNA binding capacities. Using this mutant in cultured cells and in mouse testes, they

“...understanding how granule formation contributes to biological functions in vivo remains a major challenge.”

showed that translational activation depended on the capacity of FXR1 to form RNP granules. Moreover, they restored the ability of mutated FXR1 to form granules by inserting an additional intrinsically disordered region from a heterologous protein, thus increasing its potential for multiple interactions. This additional region rescued both the capacity of FXR1 to form granules and its function in activating translation of spermiogenic mRNAs, leading to spermatid differentiation into functional spermatozoa. Furthermore, in-

roducing the FXR1 point mutation in mouse germ cells led to male sterility, phenocopying *Fxr1* germ cell deletion. These results clearly establish the critical role of FXR1 phase separation for translational activation and spermatogenesis in mice.

Although RNP granule dynamics during developmental processes have been described (15), understanding how granule formation contributes to biological functions in vivo remains a major challenge. This study represents an important advance in addressing this issue. However, many aspects remain unresolved. It is unknown how FXR1 phase separation activates translation and where in the granule translation takes place. The length of mRNAs, rather than the presence of motifs, seems to determine mRNA inclusion in FXR1 granules as well as other RNP granules. This raises the question of how speci-

ficity is achieved. Additionally, phase separation of FXR family members induces translational repression or activation. It is unclear whether the same protein orchestrates these opposite functions and through which molecular determinants. Many RNP granules are not homogeneous, but rather are made of immiscible phases. Understanding how these different phases participate in different functions of RNP granules represents another important challenge. ■

REFERENCES AND NOTES

1. S. Boeynaems *et al.*, *Trends Cell Biol.* **28**, 420 (2018).
2. S. Alberti, A. Gladfelter, T. Mittag, *Cell* **176**, 419 (2019).
3. A. S. Lyon, W. B. Peeples, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **22**, 215 (2021).
4. J.-Y. Kang *et al.*, *Science* **377**, 727 (2022).
5. C. P. Brangwynne *et al.*, *Science* **324**, 1729 (2009).
6. W. Borchers, A. Bremer, M. B. Borgia, T. Mittag, *Curr. Opin. Struct. Biol.* **67**, 41 (2021).
7. C. L. Riggs, N. Kedersha, P. Ivanov, P. Anderson, *J. Cell Sci.* **133**, jcs242487 (2020).
8. A. E. Dodson, S. Kennedy, *Dev. Cell* **55**, 4 (2020).
9. C.-Y. S. Lee *et al.*, *eLife* **9**, e52896 (2020).
10. P. Dai *et al.*, *Cell* **179**, 1566 (2019).
11. A. Ramat *et al.*, *Cell Res.* **30**, 421 (2020).
12. D. Mateju *et al.*, *Cell* **183**, 1801 (2020).
13. T. H. Kim *et al.*, *Science* **365**, 825 (2019).
14. B. Tsang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4218 (2019).
15. A. Hubstenberger, S. L. Noble, C. Cameron, T. C. Evans, *Dev. Cell* **27**, 161 (2013).

ACKNOWLEDGMENTS

The Simonelig laboratory is supported by grants from Agence Nationale de la Recherche (ANR-17-CE12-0011-01, ANR-19-CE12-0031, and ANR-21-CE12-0035-01), Fondation pour la Recherche Médicale, Fondation pour la Recherche sur le Cancer, and the Labex EpiGenMed.

10.1126/science.add6323

POLICY FORUM

ENVIRONMENTAL ECONOMICS

Insurance-sector tools to combat biodiversity loss

Risk transfer can facilitate nature-positive investments

By Carolyn Kousky

We are in the middle of a global mass extinction event (1). In addition to heightened risks related to food security, disease, natural disasters, air and water quality, and nonmaterial aspects of life (1), loss of biodiversity also threatens the livelihoods of billions of people and creates risks to the economy. The World Economic Forum estimated that over half of global gross domestic product (\$44 trillion) is highly or moderately dependent on nature (2). Some of these risks are then transmitted to financial and insurance markets (3). Addressing all these risks has been articulated in the “nature-positive” goal of halting and reversing nature loss globally by 2030. Achieving this goal will require profound changes in the relationship between nature, economies, and social and governing institutions. As leading risk managers, the insurance sector, which is not typically engaged widely on nature issues, has multiple tools to help mainstream the transition.

Engagement of the insurance sector would generate widespread global benefits, but also lower the risks insurers face from biodiversity loss, including risks of higher claims payments, lower investment returns, higher operating costs, reputation impacts, and/or increased regulatory and transition risks (4). Below, I describe five primary pathways that insurers can pursue to support nature-positive goals: creating new risk-transfer approaches, incorporating impacts on nature into underwriting, pricing for the risk reduction benefits provided by natural systems, applying a nature-positive lens to investments, and insuring nature itself (5). Despite myriad reports and papers suggesting possible approaches for insurers to have an impact in this area, the number of pilots or programs that have actually been implemented is much smaller and is the focus of this discussion. Not discussed here, given the many other treatments of corporate environmental stew-

ardship—but equally important—is insurers improving the environmental practices of their own operations.

NEW RISK-TRANSFER APPROACHES

Conservation and restoration projects need a variety of insurance coverages, including some familiar ones, such as liability and worker’s compensation, as well as new products. Many conservation groups, however, have found that insurers tend to be unfamiliar with their needs and specific risk profile. This has made it difficult for them to find affordable insurance or, sometimes, any coverage at all. In response, new organizations and products are emerging that unite expertise in conservation with expertise in risk transfer.

One example is Conservation United, a brokerage firm dedicated to finding tailored insurance solutions for conservation groups. Founded by professionals from both conservation and the insurance industry, the brokerage offers standard policies, including workers compensation, liability, auto, accident, umbrella policies, and also pollution insurance, but tailored to the actual risk profile of conservation groups. By educating insurers about the specific risks, as well as the risk reduction measures that conservation groups can adopt, Conservation United has often found more affordable policies for these groups.

Conservation efforts may also require new products to meet specific needs, such as insurance and performance bonds to assure the achievement of ecological performance standards for wetland mitigation banks. In the US, when impacts to wetlands trigger Clean Water Act regulations, they must be mitigated by restoring other wetlands and preserving them in perpetuity. To receive the necessary permits, insurance, performance bonds, or other financial assurance must be provided to guarantee sufficient funds for completion of the wetland mitigation activity. Conservation United cofounder, Chris Baker, has developed such a product where failure to achieve a defined ecological performance standard for a wetland triggers

a claim, with the payout used to bring the wetland into compliance.

Sometimes, the risk transfer solution for conservation may lie outside the traditional insurance market. Consider land trusts in the US, non-governmental conservation groups that protect land using legal tools such as conservation easements. They may have to defend conservation lands against legal challenges, which can be expensive and for which they need insurance coverage. Several years ago, land trusts were struggling to find this coverage in the private market. Insurers were reluctant to take on this risk because the potential market was small and there was not much historical data on the size of the risk (and if insurers did take it on, they would charge a high premium). In response, the Land Trust Alliance created TerraFirma Risk Retention Group LLC. It has an unusual organizational structure, being both a nonprofit organization and a captive, which is an insurance provider that is owned by the members to whom it provides coverage. Captives are a useful solution for risks that are difficult or too pricey to find coverage for in the market and have been used by many sectors. Being owned by members, they may also be better able to support risk reduction and risk education. This structure allows TerraFirma to provide the insurance more efficiently than the private sector, because they do not take a profit and can reinvest premiums not needed to pay claims. As of the time of writing in 2022, TerraFirma offered conservation defense liability insurance policies to 545 land trusts covering 10 million acres of conservation land.

Additional new risk-transfer products are also emerging to support nature-positive activities. One example is human-wildlife conflict insurance. A review in 2017 found that since 1980, programs in 50 different countries have been established to make payments to farmers or ranchers if wildlife from conservation areas destroy crops or livestock (6). The aim is to reduce retaliatory killings of wildlife. There has been increasing interest in harnessing insurance as a solution to human-wildlife conflict as it might speed payouts and improve operations. An ongoing challenge, however, is designing a product that beneficiaries are willing to pay for and can afford. To solve this, some approaches have channeled funding from government or wildlife-related tourism toward premiums (7). Indeed, public-sector support for insurance is often necessary for products aimed at lower-income populations.

INCORPORATE NATURE INTO UNDERWRITING

Underwriting is the process of evaluating risks and deciding whether to insure them.

Environmental Defense Fund, Philadelphia, PA, USA.
Email: ckousky@edf.org

Typically, underwriting decisions are driven by risk assessments. Explicitly accounting for risks stemming from degradation of nature might highlight cases where firms engaged in harmful environmental practices also have higher risks, such as liability risk, supply chain disruption risk, or reputational risk (8), making the firms poorer risks to assume. Some limited steps have been taken in this direction, such as several insurers pulling back insurance for vessels engaged in illegal fishing (9). Greater research on when, and the extent to which, nature-positive activities lower risks could guide the industry further.

Insurers could also restrict underwriting for certain entities to support broader nature-positive goals, even if insuring those entities would not be a higher risk for the insurer. Recently, advocates have pushed this position, arguing that insurers contribute to solving the climate crisis by refusing to insure carbon-intensive firms or facilities. In response, some insurers are pledging to align their underwriting with a net-zero emissions goal by 2050 through the United Nations-convened Net-Zero Insurance Alliance, which claims that such pledges currently represent about 11% of global premium written. A similar approach could be extended to risks associated with actions that harm biodiversity and species' populations.

Beyond refusing to insure firms that contribute to degradation of nature, insurers could demand substantial modifications to their operations to lessen their environmental impact. For example, AXA has pledged to stop underwriting firms that contribute to biodiversity loss in key sectors, such as production of palm oil, soy, cattle, and timber, but also to work with others in these supply chains to adopt more sustainable approaches (10).

PRICING

Some natural systems reduce disaster risks to property. Wetlands can store floodwaters, ecological forestry practices can reduce wildfire risk, mangroves can buffer storm surge, green infrastructure can absorb storm water, and vegetation can stabilize slopes. Because the price of insurance reflects the underlying risk, when insured property is protected by these ecosystems, it should lower the cost of disaster insurance. Although there is hope that such reductions in insurance prices could incentivize greater investments in conservation and restoration, it is not so easy.

First, a lower insurance price only incentivizes investments in risk reduction for those who are actually insured—and there are many entities not insured against disasters.

Second, to lower prices due to protection from natural systems, insurance firms need to be able to quantify how much the conservation will affect the likelihood or magnitude of claims they would have to pay on insured property. Modeling the benefits of nature at this scale is relatively new and typically highly uncertain. Still, progress is being made. A recent analysis by Willis Towers Watson (WTW) and The Nature Conservancy (TNC), for example, used catastrophe models (the industry standard for simulation-based modeling of the probability and impact of disasters) to estimate premium reductions of 10 to 40% when protected by landscape-scale



A first-of-its-kind insurance policy will support repair of damages inflicted by Hurricane Delta in October 2020 on portions of the Mesoamerican Reef, shown here near Isla Mujeres, Mexico, in September 2018.

ecological forestry approaches (11). AXA XL and the Ocean Risk and Resilience Action Alliance are developing a coastal risk index to quantify the impact of mangroves on flood risk; one use case they envision is helping insurers price the value of natural systems.

The final challenge is that natural systems tend to provide protection to many properties, and the premium reductions on just one are not enough to offset the costs of expanded conservation. Indeed, the WTW-TNC report on ecological forestry found that premium reductions would be sufficient to finance the debt service on a bond to undertake the restoration work—but only when accounting for the many property owners that would benefit. Even for large insured parties, like a timber company or electric utility, the economics did not work on their own.

That is a problem that plagues all investments in conservation and restoration. Natural systems are public goods, and everyone has an incentive to free ride and let others pay to provide the public good. So, how can we develop an institutional structure that can overcome free-riding and harness the reduced insurance premiums from dozens to thousands of individual policyholders and use that as a financial flow to pay for ecological investments?

There are some innovative options to harness premium reductions as a financing flow, none of which are yet in operation. One is to consider insurance at the level of the community instead of the individual. Such community-based insurance could provide financial protection for all residents against disasters (helping close the global disaster insurance gap), and the insurance savings from community-level investments in nature-based protection could financially justify greater investments in nature (12). A study examining a levee setback project (which provides ecological benefits by expanding the natural floodplain) found that such an investment could be financed by premium reductions if the community was purchasing a community-wide insurance policy (13). Community insurance, though, would need to be paid for through taxes or general revenue. No community has yet been willing to tax households or use other funds to purchase this type of community insurance.

Without a single policy, it is trickier to use insurance savings to pay for conservation or restoration. One option is for a local government or special-purpose district to enter an agreement with insurers about the premium reductions they would offer to all residents if a conservation or restoration project was undertaken. The government could then finance that investment with a bond and assess fees on residents to pay the debt service, knowing that the residents would be financially whole given their lower insurance costs. This faces political and institutional challenges more than financial ones.

INVESTING FOR NATURE

Insurers are large holders of capital, with trillions of dollars under management globally. Much of this must be held in lower-risk assets to ensure future claims-paying ability. But insurers still have some flexibility in how they allocate their investments. And like other firms, many insurance and

reinsurance companies are applying an environmental, social, and governance (ESG) lens to their investment portfolio. This can involve refusing to invest in certain companies, such as those with substantial carbon emissions; proactively investing in firms that are helping solve environmental problems; and actively engaging with companies on their environmental practices. Although this type of investing has advanced around climate, nature-positive investing is still in its infancy. Identification of risks related to biodiversity loss and environmental degradation, as well as measurement and disclosure, is still limited, but growing. The new Taskforce on Nature-related Financial Disclosures is developing a framework to mirror climate disclosures. Smaller groups guided toward certain issues, such as the Plastic Solutions Investor Alliance, are also forming and create opportunities for insurer engagement. Only an estimated 14% of global investments in conservation and restoration come from the private sector; this must at least triple by 2030 to meet global climate and biodiversity targets (14). Insurers, with their large holdings of capital, could help lead and develop best practices in this area. For example, AXA has created an impact fund that supports efforts like sustainable agroforestry and cocoa production.

Some insurers are finding that companies with negative environmental impacts create other risks that make them poor financial investments, such as liability risks, risk of losing market share, or risks of losing their license to operate. This can strengthen the rationale for focusing on investments that are nature-positive. Insurers may also find additional motivation for impact investments that support conservation or restoration of natural systems that lower disaster risks, such as wetlands or green infrastructure that lowers flood risk. When insurers are providing property coverage for disasters, they could then directly benefit from such investments in natural systems. Recognizing this, a group of Canadian insurers have recently joined Ducks Unlimited and provided funding to a new partnership, Nature Force, that will identify investments in natural systems to manage flood risk in urban-adjacent areas of British Columbia, Ontario, and Quebec.

INSURING NATURE

Natural systems can be damaged from extreme events and need to be repaired. Consider coral reefs. Very severe hurricanes can break off coral and make it hard for coral to regrow. If teams of divers enter the water quickly after a storm, however, they can reattach coral and help repair the reef. This has been done, with the help of insurance, to protect a portion of the Mesoamerican Reef.

Multiple groups in Quintana Roo, Mexico, that care about the health of the reef formed the Coastal Zone Management Trust. In 2019 (renewed in 2020 and 2021), the Trust achieved a first: It purchased an insurance policy on the reef itself to secure funding immediately after a storm to pay for divers to reattach coral. This was a parametric policy, meaning that payment was not based on damages but on observable measures of the hazard itself. Parametric insurance can be useful, as in this case, when it would be difficult for the insurance company to estimate a dollar value of the actual damage. In the fall of 2020, when Hurricane Delta hit, the policy paid around \$800,000 to the Trust for restoration work on the reef (see photo).

Although the pilot is now expanding the concept to other sections of the Mesoamerican Reef, the approach has not yet been widely replicated in other systems, partly because there are limited cases where this is a useful model and where the institutional challenges in implementation, such as convincing many beneficiaries to jointly pay for insurance, have been solved (15). For many natural systems, after a natural disaster, it is best to leave the ecosystem alone to heal itself. So, the first criteria for replicability of this concept is identifying where restoration intervention is needed in response to some type of disaster. Second, insurance is not always cost-effective. There are many times where it makes more financial sense to self-insure—that is, set money aside in a savings account to use immediately when needed, rather than pay annual premiums for an insurance policy. Third, there must be an institution to finance the insurance that spans the users and can overcome free-riding. Finally, many serious threats facing ecosystems now are uninsurable because threats are chronic and/or there is not a clear restoration or repair activity to be financed. Coral reefs are threatened more by warming and acidifying ocean waters from climate change than by hurricanes, for example, but insurance cannot address those threats.

Although those criteria narrow the possible applications, this is still an innovative approach that could prove useful in the right circumstances—for example, insuring culturally important natural assets or the ecosystems in popular tourist locations to help quickly reestablish cultural and recreational values. Another possible application is to insure green infrastructure for its performance, or against possible setbacks in initial construction.

MOVING BEYOND IDEAS

Global coordination on the insurance industry's role in broader environmental topics is emerging, such as through the

United Nations Environment Programme Finance Initiative's Principles for Sustainable Insurance. There are many theoretical discussions on possible roles for insurers to contribute to building a nature-positive future, but far fewer implemented pilots or projects. Moving beyond ideas will require better education about risk transfer among those in the conservation field, as well as more time spent understanding the needs of conservation and restoration practitioners on the part of insurers. It will require insurance regulators to support nature-positive innovations and practices and the public sector to partner and potentially cost share on some efforts. It will also require insurance-sector firms to make the decision to proactively lead and make mission-driven decisions with respect to underwriting and investments. Finally, it will require deeper research on the risks associated with environmental activities, an exploration of current barriers to implementation, and a sharing of best practices to move from theory to practice. ■

REFERENCES AND NOTES

1. IPBES, "Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services" (IPBES secretariat, 2019).
2. World Economic Forum, "Nature risk rising: Why the crisis engulfing nature matters for business and the economy" (2019).
3. U.N. D. P. Sustainable Insurance Forum, "Nature-related risks in the global insurance sector" (Sustainable Insurance Forum, United Nations Development Programme, 2021).
4. J. Chandellier, M. Malacain, *Biodiversity and Re/insurance: An Ecosystem at Risk* (Muséum National d'Histoire Naturelle, Paris, 2021).
5. C. Kousky, *Understanding Disaster Insurance: New Tools for a more Resilient Future* (Island Press, 2022).
6. J. Ravenelle, P. J. Nyhus, *Conserv. Biol.* **31**, 1247 (2017).
7. O. Wilson-Holt, P. Steele, "Human-wildlife conflict and insurance: Can insurance reduce the costs of living with wildlife?" in IIED Discussion Paper. IIED, London, 2019.
8. L. Cai, J. Cui, H. Jo, *J. Bus. Ethics* **139**, 563 (2016).
9. Oceana and UNEP PSI, "Risk assessment and control of IUU fishing for the marine insurance industry" (UNEP-Finance Initiative, Principles for Sustainable Insurance, and Oceana, 2018).
10. AXA, "AXA Group Ecosystem conversion & Deforestation policy and Natural World Heritage Sites policy" (2021).
11. N. Martinez *et al.*, "Wildfire resilience insurance: Quantifying the risk reduction of ecological forestry with insurance" (Willis Towers Watson and The Nature Conservancy, 2021).
12. A. Bernhardt *et al.*, "Community-based catastrophe insurance: A model for closing the disaster protection gap" (Marsh McLennan Advantage and the Wharton Risk Center at the University of Pennsylvania, 2021).
13. Munich Re and The Nature Conservancy, "Nature's remedy: Improving flood resilience through community insurance and nature-based mitigation" (Munich Re and The Nature Conservancy, 2021).
14. UNEP, "State of Finance for Nature: Tripling Investments in Nature-Based Solutions by 2030" (United Nations Environment Programme, The World Economic Forum, and The Economics of Land Degradation, Nairobi, 2021).
15. C. Kousky, S. Light, *Duke Law J.* **69**, 323 (2019).

ACKNOWLEDGMENTS

I thank the National Science Foundation (award 1939913) for support of this work.



PHYSIOLOGY

Our rhythmic existence

A neuroscientist interrogates sleep and circadian clocks

By **Manon Torres** and **Achim Kramer**

Did you wake up with an alarm clock this morning or work during the night last night? If the answer to either question is “yes,” then you have been doing yourself a disservice by ignoring your internal clock. Current science indicates that our bodies’ natural rhythms are essential for health, well-being, and productivity. So why do so many of us ignore or attempt to overcome them? In his new book, *Life Time*, neuroscientist Russell Foster takes the reader on a vivid and knowledgeable quest to understand sleep and biological rhythms that is part scientific treatise, part self-help guide, and part plea for societal and personal lifestyle changes that better prioritize these biological functions.

Humans evolved in a rhythmic environment resulting from the rotation of Earth on its axis. We have therefore, like most species, developed internal or circadian clocks (from the Latin *circa diem*, meaning

“about a day”) to anticipate the daily recurring changes and challenges in our environment. These clocks are molecular oscillators present in every cell of our bodies and orchestrated by a master clock in the suprachiasmatic nucleus of the hypothalamus, a structure located in the brain.

The science of circadian clocks has exploded in recent decades, impressively demonstrating that they are genetically encoded, varying from person to person, and that the interplay of the clock with an organism’s behavior, environment, and age profoundly influences every aspect of the organism’s physiology. Foster’s own research was instrumental in demonstrating that light synchronizes our internal clocks with our environment, by means of special blue light-sensitive receptors in the retina.

Circadian clocks not only govern our sleep but also exert an extremely sophisticated temporal control over nearly every physiological and cognitive function. Foster explains these complex relationships in simple but always scientifically correct language, highlighting the ubiquity of circadian regulation and the enormous impact it

Night work fails to account for our bodies’ natural rhythms and can lead to sleep and health problems.

has on metabolism, as well as on immunity and cognitive, cardiovascular, and kidney function (among other things). He spares the reader the molecular details, paraphrasing instead by saying that a clock anticipates events to ensure “the correct materials in the right place, in the right amount, at the right time of day.”

But what happens when this system gets out of balance, for example, when one works nights, frequently travels between time zones, or gets up earlier than one’s biological clock dictates? The most direct consequence is often sleep problems, but given the strong linkage between biological rhythms and physiological function, such disruptions can also have longer-term health consequences. Here, Foster describes how circadian disruptions can contribute to or exacerbate common diseases such as schizophrenia, diabetes, obesity, multiple sclerosis, migraine, infectious diseases, and cancer.

Many conditions themselves perturb sleep quality and circadian rhythms, and a number of disorders are characterized by time-of-day-dependent symptoms. Foster notes that better-timed medications could materially improve patient outcomes in the latter category. Together, however, such conditions illustrate how sleep and circadian disruption is both a potential risk factor for, and a possible symptom of, disease.

It is a distinct merit of *Life Time* that the biology of circadian clocks and that of sleep are treated as they are: intrinsically intertwined fields of science. Nevertheless, we need more research to disentangle the two phenomena, where possible, to determine which health problems are caused primarily by poor or insufficient sleep and which are caused by disruption of the temporal physiological structure. Foster elides these knowledge gaps by simply subsuming all forms of sleep and rhythm problems under the term “sleep and circadian rhythm disruption.”

Using relatable examples and tools to determine our own biological rhythms, Foster highlights aspects of our daily routines that we can target to improve our sleep quality and overall well-being.

But he does not leave it at recommendations for individuals. He advocates forcefully for societal change, especially for better education on sleep and circadian rhythms, better physician training, the abolition of daylight saving time, later school start times for teenagers, and more. ■



Life Time
Russell Foster
Yale University Press,
2022. 480 pp.

10.1126/science.add6488

The reviewers are at the Division of Chronobiology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, 10117 Berlin, Germany. Email: manon.torres@charite.de

AGING

Longevity lessons

Animal aging could hold clues to healthier human life spans

By Charles Brenner

Readers of several recently published books about longevity have been told that human aging is a solved problem. One recent book argues, for example, that dominantly acting longevity genes, identified in yeast, are conserved in animals including humans and can be activated by natural products such that we do not have to age (*1*). Given that the most powerful animal “longevity genes” are actually recessive alleles of genes required for growth and fertility, the certainty with which such claims have been made is astonishing.

It was therefore refreshing to read *Methuselah's Zoo*, the latest book from biologist Steven Austad, who expresses with clarity what we know and what we don't know about aging. Austad does not believe that aging is a solved problem, citing uncertainty about phenomena as fundamental as the timing of menopause with respect to women's life span. However, he believes that by dissecting the mechanisms by which particular animals age well, we will be able to develop medicines that promote longevity in the future.

Indeed, Austad is so confident of this assertion that in 2001, he bet gerontologist Jay Olshansky that the first person who will live to 150 had already been born. Their heirs will settle this wager no later than 2150.

Definitions matter in aging, so I appreciated that Austad establishes some boundaries early on. While he tips his hat to trees that are thousands of years old, he reasons that the ability of plants to expand clonally puts them in a different category than nearly every animal.

A former professional lion trainer and someone who spent decades studying animals in wild habitats, Austad introduces the longevity quotient (LQ) as a general system of classification that indicates whether a species' life span is above or below a trend line that adjusts for its size. (Big animals generally live longer than little animals, so a

species' absolute life span does not capture whether it really is adept at aging.)

Scoring longevity on a size-normalized scale thus allows Austad to characterize mice as poor agers and to identify tuataras, Greenland sharks, rougheye rockfish, and olms (a type of salamander) as extraordinarily good agers. Although their estimated LQ is not as high as these four species, naked mole rats appear to age better than humans and, indeed, they are the only animal model organism to have a higher LQ than humans. (Note that the longest-lived



Wisdom, the oldest known wild bird, is shown just shy of 68 years old.

whales, which outlive humans by a century, have a lower LQ than humans because their size goes in the denominator.)

Having established these terms, Austad argues that it is unfortunate that we depend on the laboratory mouse as much as we do to learn how to extend a healthy life span. He quips that humans have much to teach mice about longevity, rather than vice versa, and suggests broadening geroscience to animals with higher LQs in order to identify ways that organisms are able to remain cancer-free, physically active, and mentally astute into later decades. Cellular or organ-

**Methuselah's Zoo:
What Nature Can Teach
Us About Living Longer,
Healthier Lives**

Steven N. Austad
MIT Press, 2022. 320 pp.



oid models of tissue resiliency and repair could potentially make up for our inability to cultivate the longest-lived animals as laboratory aging models, he argues.

I agree that we will find mechanisms of resiliency in animals of high LQ, but I would make a distinction between extended life span and extended health span. In large part, the selective pressures that formed animal gene sets are bottlenecks that optimized successful reproduction, which integrates many complex functions, including the ability to acquire food, avoid predation, attract a mate, and—in some cases—care for offspring until they can accomplish these functions themselves. These abilities require thousands of gene functions in the nervous, musculoskeletal, respiratory, and circulatory systems, and more.

Because animals that succeed in reproducing multiple times pass on the most genes, there is an indirect selection for longevity. However, clever as evolution is, it doesn't produce something for which there isn't a selection, namely postreproductive vigor, unless, for example, grandmothering sufficiently improves the reproductive success of an individual's offspring, as has been suggested in human longevity research (*2*).

If I were refereeing a debate between Austad and Olshansky, I'd imagine Olshansky would argue that any potential longevity medicine is more likely to help more people get to 100-plus years—that is, closer to a genetically encoded longevity maximum—than to extend the observed maximum (*3*). In this case, the mouse being a poor ager might make it a good model for health-span research, with end

points such as improved wound healing and infection resistance in 20-month-old mice rather than extending the average life span of a mouse from 2 years to 2.5. If, however, we someday derive longevity medicine from the cave-dwelling olm, Steven Austad will deserve an assist. ■

REFERENCES AND NOTES

1. D. A. Sinclair, M. D. LaPlante, *Lifespan: Why We Age—and Why We Don't Have To* (Atria Books, 2019).
2. K. Hawkes, J. E. Coxworth, *Evol. Anthropol.* **22**, 294 (2013).
3. S. J. Olshansky, *JAMA* **320**, 1323 (2018).

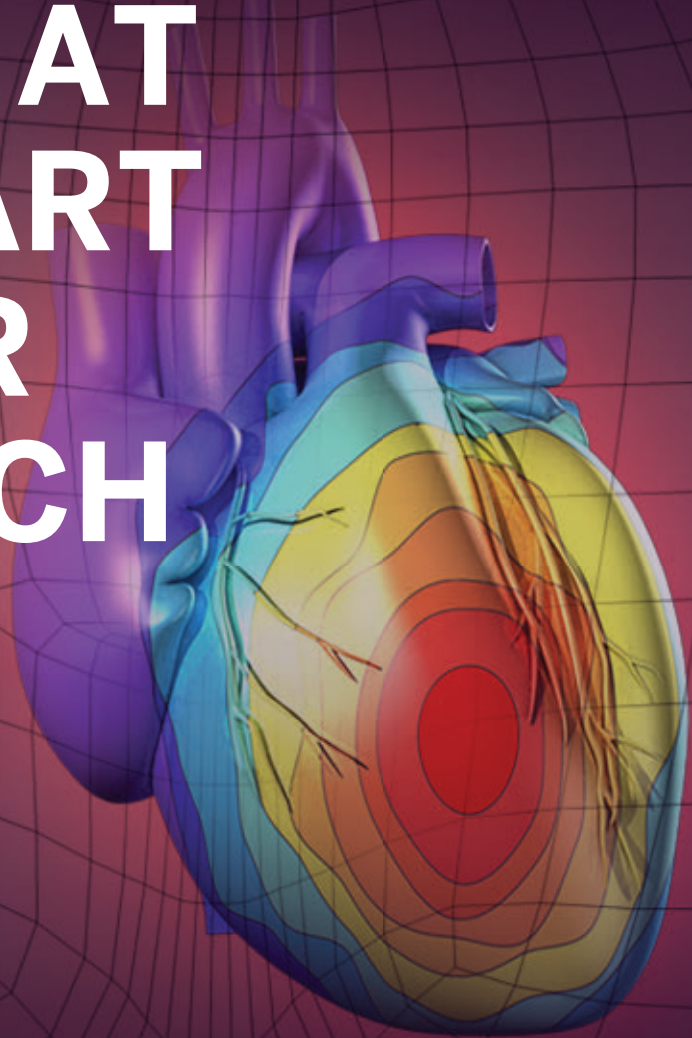
10.1126/science.add9130

The reviewer is at the Department of Diabetes and Cancer Metabolism, City of Hope Beckman Research Institute, Duarte, CA 91010, USA. Email: cbrenner@coh.org


science.org/journal/stm

PUT HUMAN HEALTH AT THE HEART OF YOUR RESEARCH

Submit your research:
cts.ScienceMag.org



Science
Translational
Medicine


 Twitter: @ScienceTM

 Facebook: @ScienceTranslationalMedicine



LETTERS

Edited by Jennifer Sills

Protect Brazil's overlooked Pampa biome

Brazil's Pampa biome is characterized by grassland vegetation and covers half of Rio Grande do Sul, the southernmost state in Brazil (1), bordering Argentina and Uruguay. The Pampa is the biome at the most risk in Brazil according to the Conservation Risk Index (1), and recent data point to an alarming 92% increase in the loss of Brazilian Pampa vegetation in 2021 compared to data from 2020 (2). Brazil must increase conservation efforts in this overlooked biome.

Although the Pampa does not support striking arboreal vegetation as seen in Brazil's forests, the region contains substantial biodiversity and performs critical ecological functions. Just 1 m² in the Pampa can harbor more than 50 plant species (3). This biome is home to 476 bird species (1) and serves as a breeding and resting place for migratory birds (4). The Pampa contributes to CO₂ capturing, climate and water cycle regulation, pollination, and food production, among other benefits to humans and animal populations (1, 4, 5).

Maintaining the Pampa's biodiversity also contributes to human health. High

biodiversity reduces, or "dilutes," infectious zoonotic disease risk (6, 7). Access to green spaces promotes mental health (8). Exposure to the ecosystem's antigens and microorganisms is important for immune development and healthy microbiota, respectively, and the vegetation helps remove air pollutants and contain noise pollution (8).

The Pampa's ecosystems are neglected in Brazil's environmental protection policies. Only about 1% of the biome is under protection, the smallest protected area of any biome in Brazil (5). Approximately half of natural Pampa vegetation has been lost, mostly to be replaced by pasture and cash crops (5, 9). Pesticides used on the crops have contaminated nearby rivers and soil, threatening humans and native species (10, 11). Invasive species, mining projects, and urban expansion (9) put the Pampa's biodiversity under increasing risk.

Recently, 19 Brazilian entities published 10 guidelines to guarantee the Pampa's protection, including planning the occupation and use of the biome's land, establishing sustainable agricultural and livestock production chains in the Pampa, and regulation and protection of traditional territories (12). Brazil should work to implement these suggestions. In addition, more research and conservation efforts should focus on the Pampa's

Brazil's vulnerable Pampa biome hosts biodiversity and provides ecosystem services.

animal biodiversity, including increased efforts to prevent poaching. Also, invasive species in the region must be better controlled. Recognizing and protecting South American grasslands diversity is fundamental to safeguarding environmental and human health.

Joel Henrique Ellwanger*, Marina Ziliotto, José Artur Bogo Chies

Department of Genetics, Institute of Biosciences, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, 91501-970, Brazil.

*Corresponding author.

Email: joel.ellwanger@gmail.com

REFERENCES AND NOTES

1. G. E. Overbeck *et al.*, *Divers. Distrib.* **21**, 1455 (2015).
2. "RAD 2021: Relatório anual do desmatamento do Brasil," Map Biomas Alerta (2022); <http://alerta.mapbiomas.org/relatorio> [in Portuguese].
3. L. da Silva Menezes *et al.*, *Braz. J. Bot.* **41**, 817 (2018).
4. A. D. Krob *et al.*, *Biol. Divers.* **1**, 132 (2021).
5. J. P. Metzger *et al.*, *Perspect. Ecol. Conserv.* **17**, 91 (2019).
6. K. A. Schmidt, R. S. Ostfeld, *Ecology* **82**, 609 (2001).
7. F. Keesing *et al.*, *Nature* **468**, 647 (2010).
8. R. Aerts, O. Honnay, A. Van Nieuwenhuysse, *Br. Med. Bull.* **127**, 5 (2018).
9. S. Ribeiro, L. F. B. Moreira, G. E. Overbeck, L. Maltchik, *J. Land Use Sci.* **16**, 260 (2021).
10. C. Gonçalves *et al.*, *Ecotoxicol. Environ. Saf.* **205**, 111314 (2020).
11. J. A. M. de Castro Lima *et al.*, *Environ. Sci. Pollut. Res. Int.* **27**, 10581 (2020).
12. "Coalizão pelo Pampa," *O Eco* (2022); <https://oeco.org.br/wp-content/uploads/2022/07/Carta-Aberta-Coalizao-Pampa.pdf> [in Portuguese].

10.1126/science.ade1838

Brazil's Pantanal threatened by livestock

Brazil's Pantanal region is a UN Educational, Scientific, and Cultural Organization World Heritage Site renowned for its spectacular wildlife (1). The vast wetland faces threats ranging from record-breaking fire outbreaks (2) to waterway plans that would dredge its main river and reduce water levels in the ecosystem (1). Cattle ranching is an ever-increasing threat to the Pantanal (3, 4), and a new bill has just been approved by the state of Mato Grosso that will open the protected portions of ranches within the Pantanal to use for livestock. If the policies in this bill are implemented, the changes will be devastating to the Pantanal's biodiversity and ecosystems.

On 29 June, the Mato Grosso Legislative Assembly approved Bill (PL) No. 561/2022 (5), which would allow cattle in the "permanent preservation areas" (APPs) and "legal reserves" (RLs) required by

Brazil's forest code (Law 12.727/2012), as well as in all other areas in the Pantanal designated for permanent conservation. Previously, cattle were only allowed in seasonally flooded grasslands in this part of Mato Grosso (6). Opening protected land for cattle is incompatible with the goals of protected status. Cattle ranching is extremely harmful to protected lands, with the capacity to change species composition and richness (7).

Fallacious arguments have been used to expand livestock in the Pantanal, such as President Jair Bolsonaro's statements that cattle help to prevent fire. Such claims have no scientific basis, in contrast to the well-documented damage caused by cattle. The presence of cattle in protected areas kills trees, fragments habitats, and causes erosion that results in siltation of water bodies (7). Degradation of the vegetation in these areas makes them more susceptible to fires (8). These fires are expected to increase further with global warming and could take the Pantanal beyond its threshold for tolerating degradation, resulting in collapse of this important region's ecosystem services (2).

The Pantanal is also home to Indigenous peoples who would be affected by the newly approved law (6). These people have a right to free and informed consultation in advance about activities that may affect their territory and way of life, as established by International Labour Organization Convention 169. Although Brazil signed and ratified this convention, violations are frequent (9).

This law sacrifices Brazil's people and ecosystems in the name of profits. The country has idle areas with the capacity to expand its livestock production without using any additional protected areas with native vegetation (10). Moreover, the expansion of ranching allowed by this bill would increase exports rather than meat available for domestic consumption (11).

A complaint has been filed with the Federal Public Ministry recommending the immediate suspension of PL No. 561/2022 (12). Brazilian courts must base their decision on technical and scientific information and immediately suspend this new law. In addition, countries that import commodities from Brazil need to assess the impacts of this bill and review their imports of Brazilian beef.

Lucas Ferrante^{1*} and Philip M. Fearnside²

¹Ecology Graduate Program, National Institute for Research in the Amazon (INPA), 69060-001, Manaus, AM, Brazil. ²National Institute for Research in the Amazon (INPA), 69067-375 Manaus, AM, Brazil.

*Corresponding author.

Email: lucasferrante@hotmail.com

REFERENCES AND NOTES

1. M. G. Coelho-Junior et al., *Nat. Ecol. Evol.* **6**, 484 (2022).
2. A. F. S. Ribeiro et al., *Environ. Res. Lett.* **17**, 065015 (2022).
3. M. Dick et al., *J. Clean. Prod.* **311**, 127750 (2021).
4. A. Guerra et al., *Land Use Pol.* **91**, 104388 (2020).
5. F. Garcia, "Deputados aprovam projeto que altera política de proteção da bacia pantaneira." Assembleia Legislativa do Mato Grosso (2022); <https://diariodecaceres.com.br/noticia/26755/Deputados-aprovam-projeto-que-altera-politica-de-protecao-da-bacia-pantaneira> [in Portuguese].
6. M. Esquer, "Deputados aprovam projeto que libera uso de áreas protegidas no Pantanal mato-grossense," *O Eco* (2022); <https://oeco.org.br/noticias/deputados-aprovam-projeto-que-libera-uso-de-areas-protegidas-no-pantanal-mato-grossense> [in Portuguese].
7. L. Ferrante et al., *J. Biogeogr.* **44**, 1911 (2017).
8. L. Ferrante, "Passando a boiada no Pantanal, o boi não é "bombeiro" é ameaça ao bioma," *O Eco* (2020); <https://oeco.org.br/analises/passando-a-boiada-no-pantanal-o-boi-nao-e-bombeiro-e-ameaca-ao-bioma/> [in Portuguese].
9. "The Brazilian Government systematically violates Convention 169, as denounced by organizations to the ILO," *Terra de Direitos* (2021); <https://terradedireitos.org.br/en/news/news/the-brazilian-government-systematically-violates-convention-169-as-denounced-by-organizations-to-the-ilo/23625>.
10. B. B. N. Strassburg et al., *Glob. Environ. Change* **28**, 84 (2014).
11. S. Hecht et al., in *Amazon Assessment Report 2021: Science Panel for the Amazon (SPA)* (United Nations Sustainable Development Solutions Network, New York, 2021), chap. 14.
12. M. Esquer, "Cientista denuncia projeto que flexibiliza Lei do Pantanal," *O Eco* (2022); <https://oeco.org.br/noticias/cientista-denuncia-projeto-que-flexibiliza-lei-do-pantanal/> [in Portuguese].

10.1126/science.ade0656

Tech companies liable for illegal wildlife trade

In a move heralded by conservationists and tech policy reformists alike, Brazil's environmental regulatory agency has penalized Meta, the parent company of Facebook and WhatsApp, for failing to remove thousands of online posts that had been flagged for selling illegal wildlife (1). The ~US\$2 million fine will not materially affect the world's largest social media company, which in 2021 earned more than US\$117 billion in global revenue (2), but it was nonetheless an important salvo in the fight to hold Meta responsible for vast amounts of illegal wildlife trading occurring on its platforms (3). To the best of our knowledge, the July ruling marked the first time any tech company has been sanctioned for hosting wildlife crime. The decision should set a precedent for holding tech companies accountable.

The fine was imposed after Brazilian conservationists provided authorities with evidence that illegal wildlife traders operated groups on Facebook and WhatsApp and that Meta had failed to remove the

content after being made aware of it, claiming it did not violate its community standards (1). Meta's response was typical. Although most tech companies officially forbid the trading of endangered species on their platforms (4), most parts of the world provide tech platforms with immunity for user-generated content, even when the activity is locally illegal. As a result, many tech companies do not enforce their own rules.

Legal reform is on the horizon. The Digital Services Act agreed upon by the European Union in April (5) and the UK's Online Safety Bill currently under debate (6) will impose a duty of care on tech platforms to restrict and remove organized crime activity on their systems. We support these reforms and urge other countries to adopt similar laws. Cybercrime legislation should explicitly address illegal wildlife trading, which is now largely dependent upon social media (3, 7). Failure to comply should result in substantial fines, as modeled by Brazil. Only successive financial sanctions and public pressure will incline noncompliant companies to invest in effectively combating wildlife crime.

Thais Q. Morcatty^{1,2*}, Gretchen Peters³, K. A. I. Nekaris¹, Pedro Cardoso⁴, Caroline S. Fukushima⁴, Hani R. El Bizri^{1,2,5}, Vincent Nijman¹

¹Oxford Wildlife Trade Research Group, Oxford Brookes University, Oxford, UK. ²Instituto de Desenvolvimento Sustentável Mamirauá, Tefé, Amazonas, Brazil. ³Alliance to Counter Crime Online, Washington DC, USA. ⁴Laboratory for Integrative Biodiversity Research, Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland. ⁵School of Science, Engineering, and Environment, University of Salford, Salford, UK. *Corresponding author. Email: tqeiroz-morcatty@brookes.ac.uk

REFERENCES AND NOTES

1. "Facebook fined US\$2 million for facilitating the trafficking of wild animals in Brazil," *Renctas* (2022); <https://renctas.org.br/facebook-fined-u2-million-for-facilitating-the-trafficking-of-wild-animals-in-brazil/>.
2. M. Iqbal, "Facebook revenue and usage statistics," *Business of Apps* (2022); www.businessofapps.com/data/facebook-statistics/.
3. K. A. Paul, M. K. Miles, D. Huffer, "Two clicks away: Wildlife sales on Facebook," *The Alliance to Counter Crime Online* (2020); www.counteringcrime.org/wildlife-sales-on-facebook.
4. "Facebook Community Standards: Restricted goods and services," *Meta* (2022); <https://transparency.fb.com/en-gb/policies/community-standards/regulated-goods/>.
5. European Commission, "Digital Services Act: Commission welcomes political agreement on rules ensuring a safe and accountable online environment" (2022); https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545.
6. UK Parliament, "Online Safety Bill" (2022); www.gov.uk/government/publications/online-safety-bill-supporting-documents/online-safety-bill-factsheet.
7. A. Lavorgna, *Crime Sci.* **3**, 1 (2014).

10.1126/science.ade0843

RESEARCH

IN SCIENCE JOURNALS

Edited by Michael Funk



EVOLUTION

Divergent species create distinct hybrids

Hybridization between species is common in many taxa and can lead to a variety of outcomes. Hybrids can backcross with parent lineages, increasing genetic diversity, or they can develop into new species. In squamate lizards, hybridization can create polyploid species with unisexual reproduction (parthenogenesis). Barley *et al.* used North American whiptail lizards, which include 15 recognized unisexual species, as a model system to examine the possible outcomes of hybridization. Phylogenetic analyses showed that divergence time between parents predicts the effects of their hybridization. Hybrid speciation and unisexuality were more common with more time since the parent species diverged from a common ancestor. —BEL *Science*, abn1593, this issue p. 773

A study of many species of North American whiptail lizards, such as the individual pictured above, identifies outcomes of hybridization.

CORONAVIRUS

Blocking a range of coronaviruses

Seven coronaviruses cause human disease, and three have caused serious outbreaks in the past 20 years. The potential for future coronavirus outbreaks in humans and the ongoing resistance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants to existing antibodies make it important to identify cross-reactive antibodies that can be the basis of therapeutics and guide vaccine design. Low *et al.* and Dacon *et al.* isolated antibodies from convalescent individuals that show broad neutralizing activity against a range of coronaviruses, including Omicron variants of SARS-CoV-2. The antibodies target a conserved region of the viral spike protein known as the fusion peptide and may act by preventing the cell fusion that is required for infection of new host cells. —VV

Science, abq2679, abq3773, this issue p. 735, p. 728

PHYSICAL CHEMISTRY

Concerted proton and energy transfer

Proton-coupled electron transfer (PCET) reactions have generated considerable interest because of their essential role in various energy conversion processes in biology. Petterson Rimgard *et al.* report another type of mechanism denoted as proton-coupled energy transfer (PCEnT), in which proton transfer is coupled to electronic excitation energy transfer. PCEnT was experimentally detected when analyzing the excited state behavior for a series of anthracene-phenol-pyridine trimers in low-temperature conditions, where PCET is thermodynamically hindered. Theoretical calculations showed that the observed PCEnT is a non-adiabatic singlet-singlet energy transfer coupled to proton tunneling, which, unlike PCET, occurs with no charge transfer between donor and acceptor. PCEnT is potentially important

PHOTO: ADAM G. CLAUSE

for light-activated chemistry, photonic materials, and photobiology, but it has yet to be identified for natural systems. —YS
Science, abq5173, this issue p. 742

NANOPHOTONICS

Generating electron-photon pairs

The interaction of electron beams with cavities and resonant structures represents a universal scheme for generating electromagnetic radiation. Feist *et al.* fabricated structures with phase-matched interactions between free electrons and the vacuum cavity field of a photonic chip-based microresonator. As the electrons passed near the resonator, coupling between them and the vacuum field resulted in the spontaneous generation of photons within the cavity. Because the electron-photon pairs are correlated, they should be a useful source for the development of free-electron quantum optics providing enhanced imaging and sensing capabilities. —ISO
Science, abo5037, this issue p. 777

FLUID DYNAMICS

When active and passive fluids interact

Incompatible liquids such as oil and water will phase separate with low interfacial tension. Adkins *et al.* investigated the dynamics of a one-dimensional interface separating an active nematic phase with a passive isotropic phase (see the Perspective by Palacci). They found a rich behavior of fluctuating interfaces in which the phase-separating fluids could form active emulsions that did not coarsen and in which droplets formed spontaneously. Macroscopic interfaces can also display propagating waves with a characteristic wave number and speed. Furthermore, the activity of one of the fluids, in which the addition of energy drove the ordering of that fluid, was able to modify the wetting transitions. The authors also observed active wetting of a solid surface whereby active extensile stresses parallel

to the surface drove the fluid to climb a solid wall against gravity. —MSL
Science, abo5423, this issue p. 768;
see also adc9202, p. 710

GPCR SIGNALING

Sniffing out odorant receptors

Odorant receptors (ORs) are G protein-coupled receptors (GPCRs) found in the cilia of olfactory sensory neurons (OSNs) in the nasal epithelium. Because only one OR is usually present in an OSN, there are too few ORs to enable identification of the specific odor that activates an OR in vivo. Omura *et al.* developed a platform to profile human ORs ex vivo. Cilia were isolated from OSNs of mice expressing a specific human OR and were found to contain all of the components of the OR signaling pathway. The cilia were screened for their responses to potential ligands in both liquid and vapor phases. —JFF
Sci. Signal. **15**, eabm6112 (2022).

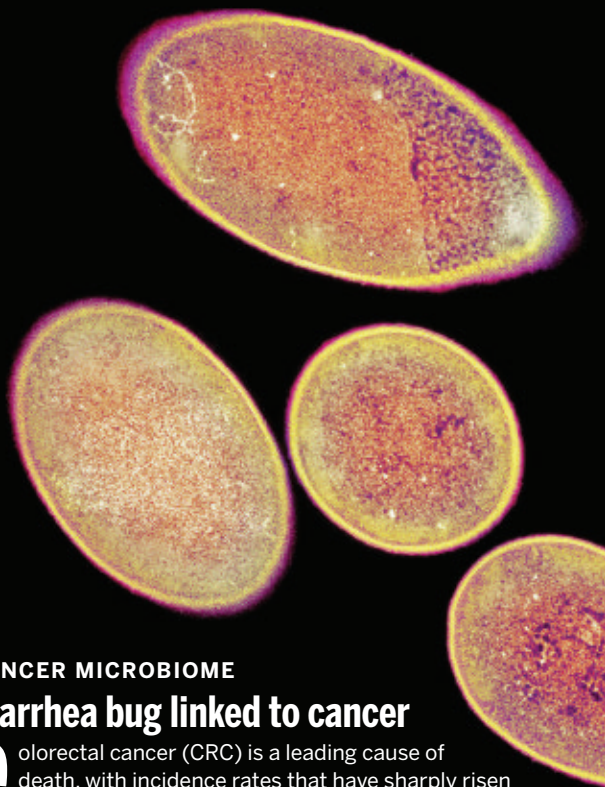
ANTHROPOLOGY

Dual protection

Most of the world's primate species are threatened with extinction due to climate change and reduction of their natural habitats caused by forest clearance for agriculture, infrastructure development, and industrial-scale resource extraction. These factors also pose significant threats to the lifeways and cultures of the Indigenous peoples who live in these areas. Estrada *et al.* conducted an extensive review of the literature of the global distribution of primate species that included a spatial analysis of their current habitats and the degree of threat they face. The authors found that as lands managed by Indigenous peoples increase, threats of extinction to primate species decrease. They propose that the national governments of these areas promote policies that protect both primates and Indigenous lifeways. —MSA
Sci. Adv. **10**, 1126/
sciadv.abn2927 (2022).

IN OTHER JOURNALS

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



CANCER MICROBIOME

Diarrhea bug linked to cancer

Colorectal cancer (CRC) is a leading cause of death, with incidence rates that have sharply risen in recent years. Drewes *et al.* found that a bacterium called *Clostridioides difficile*, which is commonly associated with infectious diarrhea in humans, may also be involved in the development of CRC. Bacterial isolates were obtained from human CRC specimens and tested to see if they could promote tumors when inoculated into mice. The toxigenic version of *C. difficile* that causes human diarrhea did trigger tumor formation. When the researchers inactivated the TcdB toxin made by the bacteria, they observed an impaired ability to form tumors, suggesting that TcdB exposure is key to the formation of CRC in the model. —PNK
Cancer Discov. **10**, 1158/2159-8290.CD-21-1273 (2022).

A diarrhea-causing bacterium, *Clostridioides difficile*, shown here in a colored transmission electron micrograph, is also associated with colorectal cancer.

GENETICS

Admixture reveals epistatic effects

Variants found in genetic studies of disease replicate poorly across populations. Explanations range from local genetic context to epistatic interactions with other genes or the environment. Patel *et al.* tested whether epistasis influences the effects of variation in regions of European ancestry that are shared among

Americans of different ethnicity. If local genetic context is the key factor determining a variant's effect size, then it should act similarly in European Americans and African Americans with European ancestry at or around that variant. However, for both low-density lipoprotein cholesterol and gene expression in an atherosclerosis study, these variants were found statistically to have dissimilar effects between the groups. These results indicate

CONSERVATION

Sequencing “Benny”

Monk seals were once widespread, but now only the Mediterranean and Hawaiian species remain. Both species are endangered, with fewer than 1500 individuals each, and both have experienced multiple population bottlenecks because of hunting and other anthropogenic causes. Many studies have used genetic markers to reveal that genetic diversity is low in these species. Mohr *et al.* went further by creating a chromosome-length genomic assembly of the Hawaiian monk seal from a well-known male, “Benny,” which confirmed the very low levels of variation (including at major histocompatibility loci) and showed that population sizes have been small for thousands of years. Purging of deleterious alleles may have occurred in this species, which could protect them against some of the impacts of inbreeding, but it also makes them vulnerable to disease. —SNV *Genes* **13**, 1270 (2022).



Genetic analysis of a Hawaiian monk seal (*Neomonachus schauinslandi*) individual showed that the population has been small for thousands of years.

that gene-by-gene or gene-by-environment interactions modify the effect size of variants associated with complex human traits. —CNS

Am. J. Hum. Genet. **109**, 1286 (2022).

OPTICS

Limits on space compaction

Optical devices such as telescopes, microscopes, and cameras contain three main components: lenses, detectors, and space. Although metalenses and plasmonic sensors have shrunk the physical size of lenses and detectors, the space

between them has remained largely the same. However, the development of “spaceplates,” a slab of lower-index material placed between the lens and detector, can reduce this separation, effectively compressing space. Shastri *et al.* theoretically studied the bounds and limitations of space compression, finding that there are trade-offs to be made in the bandwidth, the numerical aperture, and the extent of compression. The findings should help guide development of space-compressing spaceplates and the further miniaturization of optical devices. —ISO

Optica **9**, 738 (2022).

SUPERCONDUCTIVITY

Charge order in a cuprate look-alike

Infinite-layer nickelates, which were recently found to be superconducting, have striking structural similarities with the still-enigmatic cuprate superconductors. However, how far the similarities go in terms of the phases that they exhibit is still an open question. Rossi *et al.* used resonant x-ray scattering in the nickelate family $\text{La}_{1-x}\text{Sr}_x\text{NiO}_2$ to detect an incommensurate charge order that is known to be ubiquitous in the cuprates. The ordering was strongest in the parent compound, and its doping

dependence suggests that it is caused by electronic correlations. —JS

Nat. Phys. **10**, 1038/s41567-022-01660-6 (2022).

OCEAN MIXING

All mixed up

Mesoscale ocean eddies have significant impacts on general circulation and mixing. Although their behavior and effects are clearly observable, how to model them more accurately is a continuing challenge for physical oceanographers. Radko and Kamenkovich incorporated an old idea put forth by James Clerk Maxwell about the kinetics of molecular dispersion into numerical simulations of mesoscale ocean eddies and show that this results in clear improvements in eddy representation. This approach thus offers a basis for the development of better mixing parameterizations in models of physical ocean processes. —HJS

Geophys. Res. Lett. **10**, 1029/2022GL099670 (2022).

IMMUNOMETABOLISM

B cells and obesity? You (T-)bet!

Previous work has implicated B cells in obesity-associated chronic inflammation and metabolic syndrome. Hägglöf *et al.* report that B cells expressing the transcription factor T-bet expand in the adipose tissue of obese mice and humans. Expansion of T-bet⁺ B cells is supported by the production of inflammatory cytokines, including interferon gamma and interleukin-21, by an innate-like population of T cells called invariant natural killer T cells. Once activated, T-bet⁺ B cells secrete the chemokine CXCL10 and IgG2c antibodies, which contribute to macrophage recruitment and the aggravation of metabolic syndrome. T-bet⁺ B cells could thus serve as a target for treating the symptoms of obesity. —STS

Cell Metab. **10**, 1016/j.cmet.2022.07.002 (2022).

ALSO IN SCIENCE JOURNALS

Edited by Michael Funk

MOLECULAR BIOLOGY

Driving translation in spermatids

Translation control becomes a key mechanism for gene regulation during postmeiotic male germ cell differentiation because of the uncoupling between transcription and translation. Kang *et al.* identified a member of the fragile X-related (FXR) protein family, FXR1, as a translation activator for stored mRNAs to instruct spermiogenesis in mice (see the Perspective by Ramat and Simonelig). Mechanistically, highly elevated FXR1 in late spermatids forms condensates through phase separation to organize the stored mRNAs into FXR1 granules and then recruits the translational machinery to activate translation. Either germline-specific depletion of Fxr1 or modeling a phase separation-deficient mutation in Fxr1 similarly causes spermatogenic failure and male infertility in mice. —DJ

Science, abj6647, this issue p. 727;
see also add6323, p. 712

PLANT SCIENCE

Synthetic gene circuits

Synthetic engineering offers the opportunity to build new functional circuits into existing developmental programs. Brophy *et al.* have now designed a collection of synthetic regulatory elements that they can use to control gene expression (see the Perspective by Alamos and Shih). A design-build-test strategy optimized function. With these synthetic regulatory elements, the authors redesigned root development in the model plant *Arabidopsis*, quantitatively controlling lateral root density. The approach developed here may be useful for generating precise changes in a variety of phenotypes and organisms. —PJH

Science, abo4326, this issue p. 747;
see also add6805, p. 711

BIOMATERIALS

A sound way to make it stick

Tissue adhesives play a role in temporary or permanent tissue repair, wound management, and the attachment of wearable electronics. However, it can be challenging to tailor the adhesive strength to ensure reversibility when desired and to maintain permeability. Ma *et al.* designed hydrogels made of polyacrylamide or poly(N-isopropylacrylamide) combined with alginate that are primed using a solution containing nanoparticles of chitosan, gelatin, or cellulose nanocrystals (see the Perspective by Es Sayed and Kamperman). The application of ultrasound causes cavitation that pushes the primer molecules into the tissue. The mechanical interlocking of the anchors eventually results in strong adhesion between hydrogel and tissue without the need for chemical bonding. Tests on porcine or rat skin showed enhanced adhesion energy and interfacial fatigue resistance with on-demand detachment. —MSL

Science, abn8699, this issue p. 751;
see also abq7021, p. 707

ORGANIC CHEMISTRY

A fluorinated cube

The synthesis of a cube-shaped hydrocarbon nearly 60 years ago was a major event in chemistry, both because of the compound's beautiful, high symmetry and its distinctly unnatural bonding geometry. Sugiyama *et al.* have now synthesized and structurally characterized a cubane derivative in which the hydrogen atoms at each vertex are replaced by fluorines (see the Perspective by Krafft and Riess). Consistent with theoretical predictions, low-temperature electron spin resonance spectroscopy suggests that the molecule centrally internalizes an electron upon reduction. —JSY

Science, abq0516, this issue p. 756;
see also adc9195, p. 709

SPEECH EVOLUTION

Complexity from simplification

Human speech and language are highly complex, consisting of a large number of sounds. The human phonal apparatus, the larynx, has acquired the capability to create a wider array of sounds, even though previous work has revealed many similarities between our larynx and those in other primates. Looking across a large number of primates, Nishimura *et al.* used a combination of anatomical, phonological, and modeling approaches to characterize sound production in the larynx (see the Perspective by Gouzoules). They found that instead of the human larynx having increased complexity, it has actually simplified relative to other primates, allowing for clearer sound production with less aural chaos. —SNV

Science, abm1574, this issue, p. 760;
see also add6331, p. 706

SEISMOLOGY

Large-scale learning

The large amount and availability of datasets in seismology create a great opportunity to apply machine learning and artificial intelligence to data processing. Mousavi and Beroza provide a comprehensive review of the deep-learning techniques being applied to seismic datasets, covering approaches, limitations, and opportunities. The trends in data processing and analysis can be instructive for geoscience and other research areas more broadly. —BG

Science, abm4470, this issue p. 725

MICROBIOLOGY

STAND against viral invaders

The innate immune systems of animals, plants, and fungi universally use nucleotide binding oligomerization domain-like receptors (NLRs) of the STAND

superfamily to detect molecular patterns common to pathogens. Gao *et al.* show that NLR-based immune pattern recognition is also prevalent in bacteria and archaea, something that was not known before. In particular, the authors characterized four families of NLR-like genes, finding that they are specific sensors for two highly conserved bacteriophage proteins. Upon binding to the target, these NLRs activate diverse effector domains, including nucleases, to prevent phage propagation. These findings demonstrate that pattern recognition of pathogen-specific proteins is a common mechanism of immunity across all domains of life. —DJ

Science, abm4096, this issue p. 726

INSECT MIGRATION

Staying on course

We still know little about how many migrating species navigate across vast distances. This is especially true for invertebrates, which are challenging to monitor. However, new technologies leading to extremely light, animal-mounted tags are opening up new research avenues in this area. Menz *et al.* used such tags to track the flight of death's-head hawkmoths that migrate between Europe and sub-Saharan Africa. They found that the moths were able to correct to their specific course even in the face of disruptive winds and high mountains. This work suggests that the moths are not merely passively moving in the right direction, but instead actively navigating based on an internal map or compass. —SNV

Science, abn1663, this issue p. 764

TUMOR IMMUNOLOGY

Do not feel excluded

Chimeric antigen receptor (CAR)-T cells have revolutionized the treatment of blood cancers, but the selection of

target antigens is hampered by the lack of understanding about how CARs are activated. Xiao *et al.* tested the size-exclusion hypothesis, whereby the bulky tyrosine phosphatase CD45, which is present in the immune synapse, inhibits CAR-T cell activation. They found that lengthening CARs decreased CD45 exclusion and CAR-T cell activation, whereas using larger CD45s in CAR-T cells or shortening target antigens led to increased CD45 exclusion and CAR-T cell activation. Lengthening CARs abrogated the *in vivo* efficacy of CAR-T cells in a mouse tumor model. Thus, manipulating the ability of CAR-T cells to better exclude CD45 from the immune synapse could improve CAR-T cell therapy. —DAE

Sci. Immunol. **7**, eabl3995 (2022).

PROGERIA

Understanding premature aging syndromes

Heterochromatin suppresses repetitive and transposable elements, including long interspersed nuclear element (LINE) retrotransposons, but is lost during aging, leading to derepression of these LINE elements. However, the consequences of this derepression are not fully understood. Della Valle *et al.* studied LINE-1 in typical and atypical progeroid syndromes. Nuclear LINE-1 RNA expression occurred early in progeroid cells, resulting in suppression of the histone-lysine *N*-methyltransferase SUV39H1, leading to heterochromatin loss and senescent phenotypes. Using antisense oligonucleotides to deplete LINE-1 RNA restored heterochromatin epigenetic marks and decreased the expression of senescence-associated genes in human cells and in a mouse model of Hutchinson-Gilford Progeria syndrome, resulting in increased life span. These findings suggest LINE-1 as a therapeutic target for premature aging syndromes. —MN

Sci. Transl. Med. **14**, eabl6057 (2022).

REVIEW SUMMARY

SEISMOLOGY

Deep-learning seismology

S. Mostafa Mousavi* and Gregory C. Beroza

BACKGROUND: Seismology is the study of seismic waves to understand their origin—most obviously, sudden fault slip in earthquakes, but also explosions, volcanic eruptions, glaciers, landslides, ocean waves, vehicular traffic, aircraft, trains, wind, air guns, and thunderstorms, for example. Seismology uses those same waves to infer the structure and properties of planetary interiors. Because sources can generate waves at any time, seismic ground motion is recorded continuously, at typical sampling rates of 100 points per second, for three components of motion, and on arrays that can include thousands of sensors. Although seismology is clearly a data-rich science, it often is a data-driven science as well, with new phenomena and unexpected behavior discovered with regularity. And for at least some tasks, the careful and painstaking work of seismic analysts over decades and around the world has also made seismology a data label-rich science. This facet makes it fertile ground for deep learning, which has entered almost every subfield of seismology and outperforms classical approaches, often dramatically, for many seismological tasks.

ADVANCES: Seismic wave identification and onset-time, first-break determination for seismic P and S waves within continuous seismic

data are foundational to seismology and are particularly well suited to deep learning because of the availability of massive, labeled datasets. It has received particularly close attention, and that has led, for example, to the development of deep learning-based earthquake catalogs that can feature more than an order of magnitude more events than are present in conventional catalogs. Deep learning has shown the ability to outperform classical approaches for other important seismological tasks as well, including the discrimination of earthquakes from explosions and other sources, separation of seismic signals from background noise, seismic image processing and interpretation, and Earth model inversion.

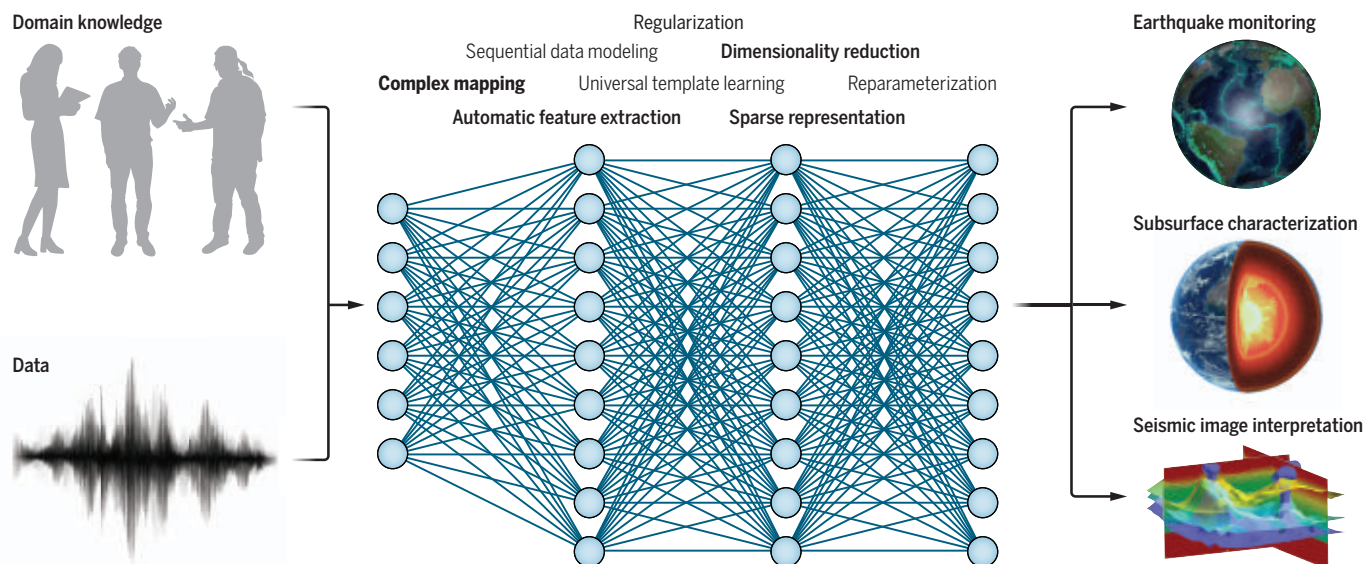
OUTLOOK: The development of increasingly cost-effective sensors and emerging ground-motion sensing technologies, such as fiber optic cable and accelerometers in smart devices, portend a continuing acceleration of seismological data volumes, so that deep learning is likely to become essential to seismology's future. Deep learning's nonlinear mapping ability, sequential data modeling, automatic feature extraction, dimensionality reduction, and reparameterization are all advantageous for processing high-dimensional seismic data, particularly because those data are noisy and,

from the point of view of mathematical inference, incomplete. Deep learning for scientific discovery and direct extraction of insight into seismological processes is clearly just getting started.

Aspects of seismology pose interesting additional challenges for deep learning. Many of the most important problems in earthquake seismology—such as earthquake forecasting, ground motion prediction, and rapid earthquake alerting—concern large and damaging earthquakes that are (fortunately) rare. That rarity poses a fundamental challenge for the data-hungry methods of deep learning: How can we train reliable models, and how do we validate them well enough to rely on them when data are scarce and opportunities to test models are infrequent? Further, how can we operationalize deep-learning techniques in such a situation, when the mechanisms by which they make predictions from data may not be easily explained, and the consequences of incorrect models are high? Incorporating domain knowledge through physics-based and explainable deep learning and setting up standard benchmarking and evaluation protocols will help ensure progress, as is the nascent emergence of a seismological data science ecosystem. More generally, a combination of data science literacy for geoscientists as well as recruiting data science expertise will help to ensure that deep-learning seismology reaches its full potential. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: mousavim@google.com
Cite this article as S. M. Moustavi, G. C. Beroza, *Science* 377, eabm4470 (2022). DOI: 10.1126/science.abm4470

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abm4470>



Deep-learning processing of seismic data and incorporation of domain knowledge can lead to new capabilities and new insights across seismology.

REVIEW

SEISMOLOGY

Deep-learning seismology

S. Mostafa Mousavi^{1,2*} and Gregory C. Beroza¹

Seismic waves from earthquakes and other sources are used to infer the structure and properties of Earth's interior. The availability of large-scale seismic datasets and the suitability of deep-learning techniques for seismic data processing have pushed deep learning to the forefront of fundamental, long-standing research investigations in seismology. However, some aspects of applying deep learning to seismology are likely to prove instructive for the geosciences, and perhaps other research areas more broadly. Deep learning is a powerful approach, but there are subtleties and nuances in its application. We present a systematic overview of trends, challenges, and opportunities in applications of deep-learning methods in seismology.

Seismology is a data-rich and data-driven science in which the application of deep-learning techniques is growing rapidly (Fig. 1). Deep learning (1) uses deep neural networks (DNNs) to learn complex relationships, find multiscale structure and patterns, and address ill-posed inverse problems by developing high-level representations of data using stacked layers of neurons and multiple nonlinear transformations. The abilities of DNN variants—such as convolutional neural networks (CNNs) to extract multiscale features from data and recurrent neural networks (RNNs) to model temporal relations in sequential data—make them powerful tools for seismology.

Deep-learning approaches have entered almost every subfield of seismology, for which they have shown the ability to outperform classical approaches, often dramatically, for seismological tasks such as denoising, earthquake detection, phase picking, seismic image processing and interpretation, and inverse and forward modeling. Some properties of DNNs—such as their universal approximation capability, automatic feature extraction, and dimensionality reduction—have been shown to be particularly advantageous in processing high-dimensional seismic recordings, which often are noisy and incomplete (Fig. 2). Seismological deep learning can process massive amounts of multifidelity seismic observations with unprecedented spatiotemporal coverage and lead to new insights and discoveries (2). Deep learning may be particularly effective for seismological problems for which the underlying physical processes are incompletely understood but for which the data are abundant and of high quality.

We present a systematic overview of deep-learning methods in seismology by sampling the current state of the art and through a meta-analysis of 637 journal papers published between January 1988 and January 2022 (3).

We conclude with some speculation on future trends and with recommendations intended to accelerate progress.

Data processing automation

Seismic data are recorded (often irregularly or heterogeneously) as time series of ground motion by sensors that are deployed either sparsely (in a network) or densely (in an array) and that register acceleration, velocity, or displacement as output. Typically, these quantities are recorded in three perpendicular directions, so that we work with three-component vector ground motion. Important extensions to conventional seismic recording include rotational motion (4) and tensorial strain (5).

In both passive and active-source seismology, usually a sequence of processing and analyzing steps exists that needs to be performed. Despite decades of earlier efforts to

develop algorithms to automate such data-processing and -analyzing tasks, many cases remain for which manual processing by skilled analysts is the most reliable option (such as phase picking or seismic interpretation). The accelerating expansion of seismic data volumes poses new challenges and brings with it the need to develop a new generation of robust processing tools by using data-driven approaches such as deep learning.

Seismic event monitoring

The classification capability of DNNs makes them ideal for processing tasks such as event discrimination (6–8), event detection (9, 10), phase picking (11, 12), first-polarity determination (13, 14), and phase association (15, 16) that together form a seismicity monitoring workflow.

Seismic event discrimination

Various studies have shown outstanding performance of deep learning for different event discrimination tasks, such as differentiating between human-made explosions and natural earthquakes (6, 7), volcano-tectonic earthquakes and volcanic tremor (8), mining-related events (17), tectonic tremor and local earthquakes (18), and teleseismic and local earthquakes (19). The main challenge in developing discriminator models is the scarcity of large-scale labeled datasets. Efforts to address this include semisupervised learning (20), unsupervised learning (19), and transfer learning (8).

Earthquake detection and phase picking

Identification of earthquake signals among nonstationary background noise and a variety

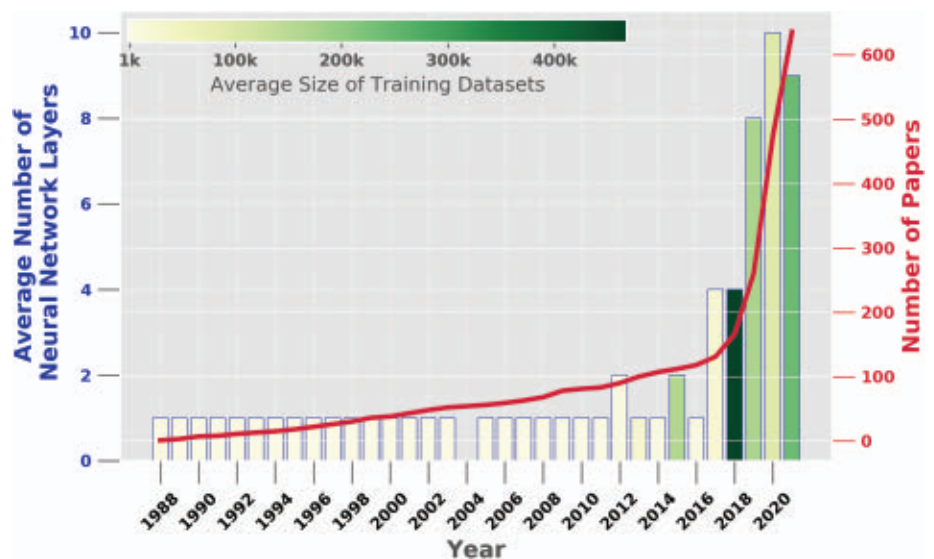


Fig. 1. A surge of interest in deep-learning applications in seismology. Cumulative number of published papers in seismology between January 1988 and January 2022 that used neural networks (red curve). Bars show the average number of hidden layers of neural networks—the depth of networks. The color scale depicts the average size of training data used to train the networks.

¹Department of Geophysics, Stanford University, Stanford, CA 94305, USA. ²Google, Mountain View, CA 94043, USA. *Corresponding author. Email: mousavim@google.com

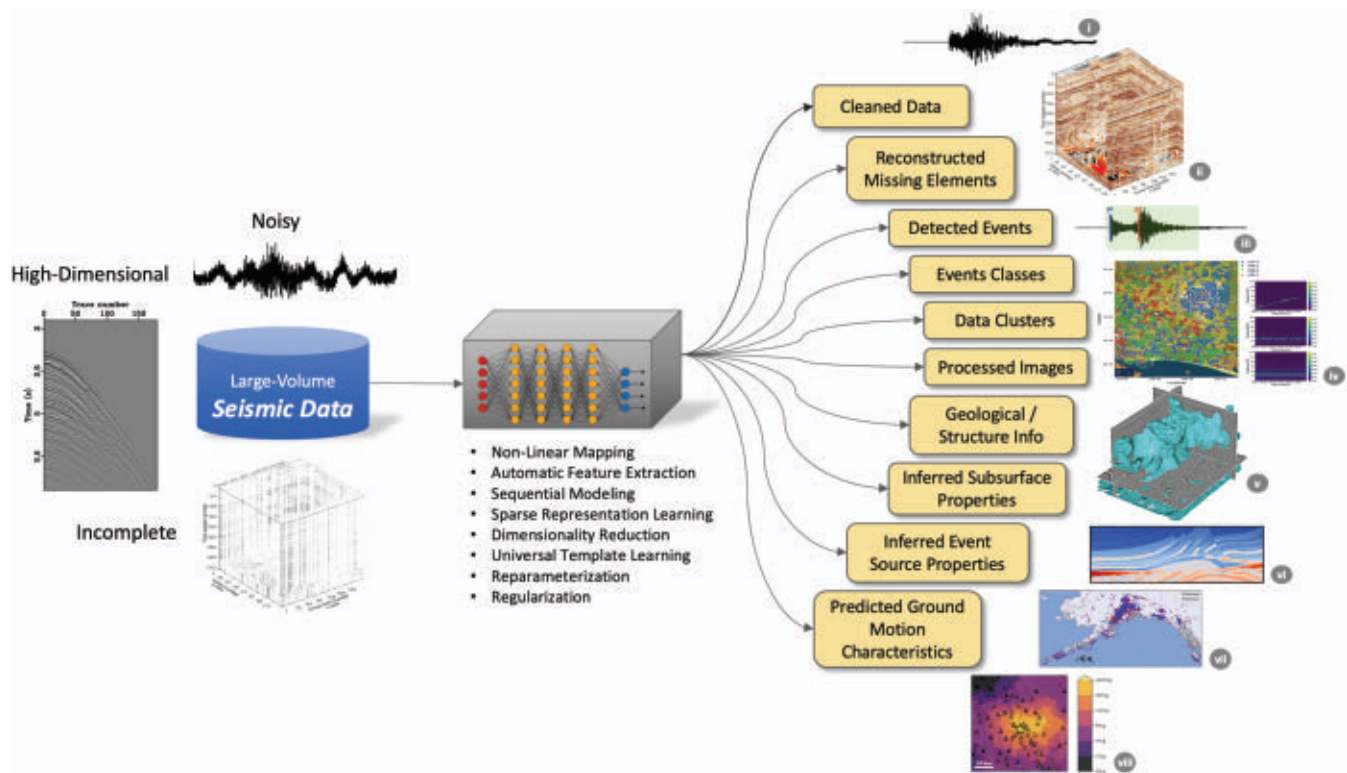


Fig. 2. The properties of DNNs for the processing of seismic data. Various properties of DNNs such as their universal approximation capability, automatic feature extraction, and dimensionality reduction are advantageous for processing large-volume, high-dimension seismic recordings, which often are noisy and incomplete. These attributes are used to build powerful models for various seismological tasks such as (i) seismic data denoising, (ii) interpolation of missing traces, (iii) earthquake signal detection and phase picking [reprinted with permission from (43)], (iv) exploratory data analysis [reprinted with permission from (123), copyright Seismological Society of America], (v) salt body detection, (vi) seismic velocity model building, (vii) earthquake location estimation, or (viii) ground motion prediction.

of nonearthquake signals, from both natural and anthropogenic sources, as well as precise measurement of P wave or S wave arrival times, is a challenging task—we differentiate between discrimination and the detection task in which the goal is to identify only one signal of interest among a variety of recorded events. The huge archives of seismic data and the availability of hand-picked labels from earthquake catalogs made these tasks to be among the earliest and the most popular applications of data-hungry deep learning in seismology. The potential of deep learning for these tasks was evident from the earliest deep learning-based studies (9–11, 21). Since then, more than 80 models have been developed to explore the performance gain obtained through more complex deep-learning architectures, such as attentive CNN-RNN (12, 22) and Capsule Networks (23). Researchers have also explored the advantages and disadvantages of performing these tasks in the time domain (10, 11), the time-frequency domain (21), and a fusion of time and time-frequency (24); of using multistation data (25, 26) and multisensor data (GPS and seismic); of using two observation levels (seismic station level and network level) (27); and of performing the two tasks sequentially (28, 29) or simultaneously (12), on data recorded with

low-cost microelectromechanical systems (MEMS) accelerometers (30), distributed acoustic sensing (DAS) (31), or ocean-bottom seismometers (OBSs), and geophones (25).

Deep-learning detectors and phase-pickers can be thought of as universal templates of earthquake signals represented by a set of kernels of different sizes at different layers of a neural network and the connection weights between them. The outstanding performance of these models comes from their ability to learn a deep and compact representation of diverse templates; however, because of the lack of standardized benchmarking, the optimal deep-learning approach for these tasks is unclear. Nevertheless, common observations and conclusions are that (i) deep-learning models systematically outperform classical characteristic function-based models; (ii) deep-learning model performance approaches that of template-matching methods, but with much lower computational cost and no need for template events; (iii) deep-learning models can pick phase-arrival times with a similar picking error to that of skilled analysts while picking far more phases in much less time; (iv) the power of artificial intelligence (AI)-based pickers is most evident for the challenging task of picking S waves that emerge from P coda; (v)

deep-learning models can generalize well to other regions or data types beyond their training data; and (vi) compared with the classical methods, deep-learning models perform better in conditions with lower signal-to-noise ratio (SNR). As a result, deep-learning models provide more complete earthquake catalogs and much-higher-resolution maps of seismicity by detecting more small events.

Seismic image processing

Reflected and refracted seismic waves recorded by an array of seismic sensors are used to image subsurface geology and structure. Deep learning has proven an effective tool in the processing steps used to improve the quality of seismic images and to transform them into an interpretable image of the subsurface by removing data acquisition artifacts and wave propagation effects to highlight events that more accurately portray the true geology and structure.

Seismic denoising

Recorded seismic data are always contaminated by unwanted background noise or interference from other signals. This unwanted energy is often nonstationary with incoherent (random or Gaussian) and/or coherent components

that can mask the signal of interest and affect subsequent inversion or interpretation. Seismic denoising (32) and closely related tasks such as multiple removal (33) and debending (34) are a focus area for the applications of deep learning in seismology. These studies have shown the promising performance of deep learning for suppressing a wide variety of noise in seismic data, including the random noise (35), coherent low-frequency noise (36), swell noise (37), ground roll (38), and coherent and incoherent noise (32) in conventional and unconventional seismic data such as DAS (39) and in one-dimensional (1D) (32), 2D (40), and 3D (41) recordings. The effectiveness of deep-learning models in suppressing various types and levels of seismic noise with low computational cost is remarkable.

In deep-learning denoising, the statistics of signal and noise are directly learned from training data. As a result, the learned models achieve superior performance and are more flexible than conventional methods that rely on prior assumptions about the statistics of the noise or signal. The out-of-distribution generalization of their superior performance, however, depends highly on how well their training data represent the distributions of field data. A challenge in building training data for deep-learning denoisers is to obtain noise-free signals and signal-free noise from field recordings. An effective approach to meet this challenge is to create semisynthetic data in which clean synthetic signals are combined with manually picked noise from the field data (36, 39). Performing unsupervised denoising by using either autoencoders (35, 40) or generative adversarial networks (GANs) (36, 38) are other common approaches. The idea behind autoencoder denoisers is that the most representative features of the data that are automatically extracted by the autoencoder have a more coherent structure and represent signals of interest. Hence, autoencoders are usually only effective for random noise suppression. Generative models, on the other hand, learn the domain mapping from noisy data domain into the signal domain and could work for incoherent noise as well, but require examples of clean and noisy data.

Seismic image acquisition and processing

Tasks under this heading include identification of shots and receivers with position errors (42), interpolation for reconstructing traces missing during acquisition (43), extrapolation to low-frequency components on the basis of high-frequency energy (44) or vice versa (45), identification of the first arriving seismic energy in time-space domain (46), automation of velocity picking (47), extraction and classification of dispersion curves (48), and seismic migration. Although velocity picking and migration can be treated as an inversion pro-

cess, we discuss them here under seismic image processing. Tasks in which the main goal is the inversion for some physical properties are discussed in a subsequent section (49, 50). The results of these studies suggest impressive capabilities of DNNs in performing subtle and complex processing tasks with unprecedented computational efficiency. For example, deep-learning seismic interpolation methods can successfully reconstruct missing regularly or irregularly spaced data with flexibility similar to that of traditional methods based on wave equation and signal processing, but without their high computational costs or limiting assumptions such as sparsity or low rank. The efficiency gain is evident from applications for seismic migration. Using a CNN, Liu *et al.* (49) accelerated prestack time migration by more than 100 times, and Kaur *et al.* (50) improved both computational efficiency and image resolution in least-squares migration by estimating the inverse Hessian operator using a cyclic GAN and eliminating the need for iteration.

Seismic image interpretation

Seismic image interpretation aims to extract useful information regarding subsurface geology (such as rock type) or structure (such as fault, horizon, and salt body) from seismic images that have limited resolution and are contaminated with noise. Many deep-learning techniques have been developed to automate and assist seismic-image interpretations that are subjective and often require substantial time and domain knowledge from the interpreter.

The feature-learning and dimensionality-reduction abilities of CNNs are well suited for clustering of seismic response in terms of geological properties or facies (51, 52) and outperform machine-learning approaches for seismic facies mapping based on hand-engineered features. Deep-learning models, however, have the potential to go beyond simple pattern recognition to provide accurate end-to-end classification results (53–55) while quantifying the uncertainties (56).

Besides geological interpretation, the power of deep learning for image segmentation and object detection makes it a popular choice for structural-interpretation tasks such as automatic identification of geological channels (57), bright spots (58), collapsed karst features (59), scatterers (60), bottom-simulating reflections (61), salt bodies (62), horizons (63, 64), and faults (65–67) in seismic images.

Seismic image interpretation is a type of image understanding. Hence, applications of popular neural network architectures developed for computer vision are a natural choice. The remarkable performance of these methods for seismic image interpretation—despite the scarcity of large, complete, and precisely labeled training datasets—is attributable to a combination of the extensive use of data aug-

mentation and the generation of sophisticated synthetic data.

Forward problems

The impressive predictive power of DNNs makes them popular tools for forward modeling problems in seismology. We outline three categories—seismic wave simulation, ground motion characterization, and earthquake forecasting—as instructive examples of forward-modeling tasks in seismology for which DNNs have been used.

Seismic wave simulation

To date, applications of deep learning for seismic wave simulation have been limited. Nevertheless, deep learning appears to provide an effective alternative for standard numerical methods and can address issues such as discretization errors and high computational complexity. The use of physics-informed neural networks (PINNs) (68–71) and GANs (72–74) are two current trends. In PINNs, a DNN is trained to learn the solution of the wave equation for a medium by implicitly defining the boundary conditions and the wave equation in the loss function used in training of the neural network. These deep-learning models can learn to solve the wave equation in 2D or 3D media, even for complex faulted structure or topography, and generalize well beyond the time stamps of their training dataset. They are much more efficient in computing arbitrary space-time points in the wavefield than are traditional numerical simulations and reduce computation time by at least an order of magnitude.

GAN simulators use the universal function approximation ability of DNNs to learn the probability distributions of attributes of training data by optimizing a generator network. This generator model can be used as reparameterization of such distributions to generate new samples drawn from the learned distributions during inference. GANs have been used to generate synthetic earthquake and non-earthquake seismograms as a data augmentation tool for training deep-learning earthquake detectors (72); to generate broadband seismic signals by blending the low-frequency output of numerical physics-based simulations with sparsely sampled broadband observations (73); and to generate 3C strong motion time series for different magnitude, distance, and site conditions (74). They provide an efficient framework for generating large-scale synthetic training data to improve the performance of deep-learning classifiers and detectors.

Ground motion attribute modeling

Another category of forward models provides estimates of distinct attributes of ground motion caused by a seismic source such as peak ground acceleration (PGA) (75, 76) or

the travel time of a seismic wave between two points in a medium (77).

Neural networks have been used to build nonparametric ground-motion models (GMMs) based on a fixed set of input parameters by using field (78) or synthetic (79) data. This group of AI-based GMMs (which is more than 15 years old) has similar input-output and performance to those of traditional methods; however, another group of deep learning-based GMMs introduces a more progressive generation of GMMs (75, 76, 80) under a new paradigm. These models can provide a rapid estimation of ground shaking directly from observed waveforms without an interpretation through earthquake source parameters or site conditions. A network-level pattern of ground-motion propagation contains enough information for a DNN to nowcast PGA at stations that have not been shaken by the earthquake waves yet, on the basis of those stations that have been shaken (76, 80). This group of end-to-end approaches seems particularly relevant for earthquake early warning (EEW).

Deep learning-based Eikonal solvers (77) have much lower space and time complexity compared with those of the traditional finite-difference approaches (especially for complex structures) because they eliminate the need for travel-time lookup tables and greatly benefit from the parallel computing capability of deep-learning platforms.

Earthquake forecasting

Solving the earthquake forecasting puzzle by using neural networks is an appealing topic (87). Early attempts relied on hand-engineered features, but the emphasis has shifted to neural network-based forecasting approaches that more directly incorporate the complete spatiotemporal structure of seismic catalogs (82, 83). This exploits the automatic feature extraction ability of some DNNs along with their potential for learning the dynamics of data to take a fresh look at an old problem. Most of these studies are in an experimental phase and focus on comparing the relative performance of (usually not fine tuned) different deep-learning and/or classical machine-learning predictors. It remains to be established whether deep-learning earthquake forecasting models offer new insights into earthquake predictability and whether they perform better than traditional empirical approaches (84).

The rapid progress in deep learning-based earthquake monitoring has resulted in deeper earthquake catalogs with unprecedented spatial resolution and magnitude of completeness that should improve both traditional statistical and physics-based forecasting methods (85) as well as AI-based approaches. Here too, the predictive value of newly uncovered small earthquakes in deep learning-based catalogs needs to be clarified (86). A key challenge is to

effectively explore these complex and denser catalogs and extract new relationships. This is a task that deep learning can help with as well. For example, Zhu *et al.* (87) present a reinforcement-learning approach to model complex spatiotemporal relationships in earthquake catalogs that are not possible to model by using classical approaches such as the epidemic-type aftershock sequences (ETAS) method.

Deep learning offers new opportunities in earthquake forecasting by blending different potential precursory data and known earthquake physics with seismicity data and underlying physics that remain unknown. The growing number of forecasting competitions such as in (88), richterx (89), or AETA (90) is important for engaging broader contributions beyond the earthquake seismology community. More effort in setting up the technical rigor and benchmarking practices is needed to direct this new wave of enthusiasm into constructive contributions toward earthquake forecasting.

Inverse problems

Solving often nonlinear and ill-posed inverse problems forms another popular category of applications of deep learning in seismology. Deep learning for estimating physical parameters from observed seismic data can be divided into two main categories: (i) characterizing Earth's subsurface properties and (ii) characterizing the earthquake source.

Subsurface characterization

Various characteristics of DNNs have proven effective for solving inverse problems in seismology. Their ability to learn highly nonlinear relationships is the main idea behind fully data-driven end-to-end approaches for inferring a wide variety of subsurface properties such as seismic velocity (91–93), impedance (94, 95), petrophysical attributes (96, 97), dynamic reservoir properties (98, 99), groundwater distribution (100), and anelastic attenuation (101) directly from either pre-stack (92) or post-stack (102) seismic data.

Deep-learning applications for seismic inversion are not limited to end-to-end mappings. In conventional seismic inversion, deep learning can be used to learn an optimization algorithm (103), to introduce data-driven elements into an optimization procedure (104), or to replace computationally expensive steps (105). The representation-learning and dimensionality-reduction capabilities of DNNs are valuable attributes for seismic inversion. In full-waveform inversion (FWI), these characteristics are used to transform the large inverse problems by either skeletonizing the data (106) or reparametrizing the Earth model (107, 108) as the connection weights and/or learnable kernels in a DNN. Although reducing

the dimension of observed and synthetic seismic data by use of DNNs can reduce the nonlinearity of the inverse problem and the risk of converging to a local minimum (106), the provided sparse representation of an Earth model in a lower-dimensional latent space reduces the computational cost and effectively regularizes the problem (103, 109). Representing Earth models with sparse multiscale feature maps through a DNN acts like a regularization technique by constraining the updates to the effective dominant features that represent the velocity model (107, 109). In addition, the differentiable representation of the Earth model facilitates the inclusion of physics-based constraints (95, 110). The flexibility of the open-source deep-learning platforms (such as Tensorflow or PyTorch) and their appealing computational capabilities (such as automatic differentiation and GPU parallel computing) motivate reformulating the entire FWI into a DNN with forward modeling guided by physics laws through neuron computations (111). A more advanced example of this is the development of specialized deep-learning architectures that are capable of learning the internal mechanism of general wave-based inversions (112).

These studies demonstrate the exceptional performance of deep-learning techniques in providing fast and reliable solutions to large-scale seismic inverse problems. Moreover, the processing time for both training and inference of deep-learning models can be an order of magnitude less than optimization time of classical seismic inversion approaches. Deep-learning inversion methods also tend to be less sensitive to noise (113) and missing low-frequency components (110) compared with those of classic FWI. They either are not sensitive to the choice of initial velocity models or do not require such initialization (91, 93). Last, the modularity of deep learning enables the processing of multimodal input data such as well-log and seismic or gravity and electromagnetic data, which facilitates joint inversion (114).

Earthquake source parameterization

Deep learning has great potential to provide near-real-time earthquake source parameters, such as hypocenters, magnitudes, and source mechanisms. Deep learning is effective for end-to-end learning of earthquake source properties based on waveform patterns propagating across a network of seismic stations (115–117). It can also be used to go beyond conventional multistation approaches by characterizing the source parameters from single-station data (118, 119).

Although multistation approaches are aligned with conventional methods, to date they suffer from smaller training datasets (usually a few tens of thousands of events). Single-station

approaches are in an experimental phase that has explored the feasibility of maximizing the information that can be extracted from smaller units of observation (short windows of waveforms recorded by a single instrument) by using the feature learning and high non-linearity of DNNs. The latter approach has the potential to enable monitoring of smaller events by using limited observations (or sparse network coverage) and benefits from much larger training datasets (hundreds of thousands of single-station observations) available for training deeper DNNs.

Overall, deep-learning applications for source characterization are in an earlier stage than those for subsurface inversion or automatic data processing. They encounter the challenges of the unavailability of genuine ground truth and the scarcity of large events (earthquakes of 7.0 magnitude or greater) in the training set. Incorporating the aleatory and epistemic uncertainties into the modeling procedure and physical constraints through physics-guided (120) or physics-informed (121) DNNs could be a promising direction to overcome these limitations.

Exploratory data analyses

Most applications of deep learning in seismology have been related to automation and modeling tasks; however, deep learning can be used in discovery mode to find patterns and insight in seismic data as well. Seydoux *et al.* (2) developed an unsupervised framework that combines a deep scattering network and a Gaussian mixture model to identify patterns in continuous seismic recordings through segment clustering. They used this to reveal the repeating precursory seismicity preceding a landslide in Greenland. This method has also been used for clustering of daily quasi-periodicity and numerous transient microevents on continuous seismic data recorded on Mars (122). A similar approach has been used to cluster urban seismic noise in Long Beach to identify their sources (123). In this approach, features that represent distinct clusters most effectively are extracted from input data by using a deep CNN through simultaneous optimization of clustering and feature-learning tasks. Unsupervised deep learning holds great potential for uncovering new insights in continuous seismic data. The intelligent feature extraction and dimensional reduction mechanism of this new group of DNNs (19, 124) render them well suited for exploratory analyses of high-dimensional seismic data such as DAS.

Beyond conventional seismology

Seismology has many unconventional applications. The sensitivity of seismic sensing to the Doppler effect makes it an effective approach for the monitoring of moving objects

such as humans, vehicles, and aircraft (125). Deep learning has also been used to detect human footsteps (126, 127), discriminate between different classes of targets (such as humans versus vehicles) (128, 129) and different types of vehicles (such as light vehicles, heavy vehicles, wheeled vehicles, and tracked vehicles) (130), or discriminate between different wildlife species and their behavior (131).

These unconventional applications of seismic sensing bring challenges and motivate technical approaches that could be of interest to the larger seismological community. For example, Bin *et al.* (128) showed that through on-site feature learning and classification by using CNNs to compress data and only transmitting the classification result to the server, it is possible to obtain similar results to those of state-of-the-art cloud-based deep-learning methods at only 1/10 of the computational time. Similar approaches could help reduce current challenges in seismic monitoring by use of fiber optic cables. Compressed sensing of seismic data (132), fusion of multisensor data (129), and federated learning (133) are other areas in which deep-learning techniques could assist in processing large data volumes.

The Nobel prize-winning detections of gravitational waves and subsequent groundbreaking discoveries made by using the advanced Laser Interferometer Gravitational-wave Observatory (LIGO) were made possible through a series of multidisciplinary research efforts. The advances in deep-learning seismology discussed above have the potential to contribute to gravity-wave measurements in multiple ways. Despite sophisticated seismic isolation systems, ground-based gravitational wave interferometers remain susceptible to ground shaking from large teleseismic earthquakes, which are one of the primary sources of interruption of the proper functioning of the interferometers and substantially reduce their duty cycle (134). Machine learning for ground motion characterization and EEW could help reduce that impact. Deep learning can also be used to predict on the basis of archival seismic data the ground-motion level and state of the gravitational wave detector during the event of an earthquake (135).

The template-matching algorithms used for the detection of gravitational waves can be complex and computationally expensive. Similar to earthquake signal detection and discrimination, deep learning has been shown to be an effective alternative for efficient detection and classification of gravitational-wave signals from different cosmic sources (136). DNN models can detect and estimate the true parameters of real events observed by LIGO with comparable sensitivity and lower error than that of template-matching, while being several orders of magnitude faster and more resilient to transient noise artifacts (136). These gravitational waves are extremely weak signals

embedded in highly non-Gaussian and nonstationary noise owing to instrumental artifacts and environmental contamination. Denoising by use of deep-learning approaches and with the supervision of accompanied recorded seismograms is another approach for the detection of new events below the noise floor (137).

Discussion Meta-analysis

To date, automation of seismic data processing and solving of inverse problems form the majority of deep-learning applications in seismology (Fig. 3). Although deep learning can be used to recognize patterns and extract insights from high-dimensional seismic data, their application for exploratory analysis of seismic data and scientific findings are largely unexplored (Fig. 4A). This might be due to an atmosphere of skepticism regarding the black box nature of deep-learning models. In most seismic data-processing tasks, however, the cost of a wrong answer is low relative to the value of a correct answer. For example, the risk of picking wrong or inaccurate phase arrivals in earthquake monitoring is low because wrong or inaccurate arrivals will usually be eliminated or refined in subsequent processing. Thus, having a deep-learning picker that can do this task with high efficiency but some brittleness, which is defined as the tendency to get fooled (138), and imperfect precision does not undermine the utility of the method. Even in applications such as EEW, for which the cost of a wrong answer may be high, deep learning may offer the best available approach (76). Last, although deep learning cannot be used to establish causation or understand processes directly from data, it can be used to inspire and direct researchers toward scientific discoveries.

The scarcity of large-scale labeled field data for model building is more evident in active seismology and for applications such as seismic image processing, interpretation, and subsurface characterization (Fig. 3). This scarcity is due to both the limited number of large-scale and publicly available datasets and the subjectivity of labeling for these tasks. Earthquake signal processing tasks—such as seismic phase picking, phase association, and earthquake detection—benefit from larger and publicly available training datasets.

Supervised learning has dominated deep-learning approaches used in seismology to date. However, important applications for alternative unsupervised and semisupervised approaches are being increasingly found. Unsupervised learning has focused on seismic facies analyses, seismic denoising, event discrimination, and exploratory data analyses. As the number of developed models for each seismological task grows, techniques such as ensemble learning are likely to find more applications.

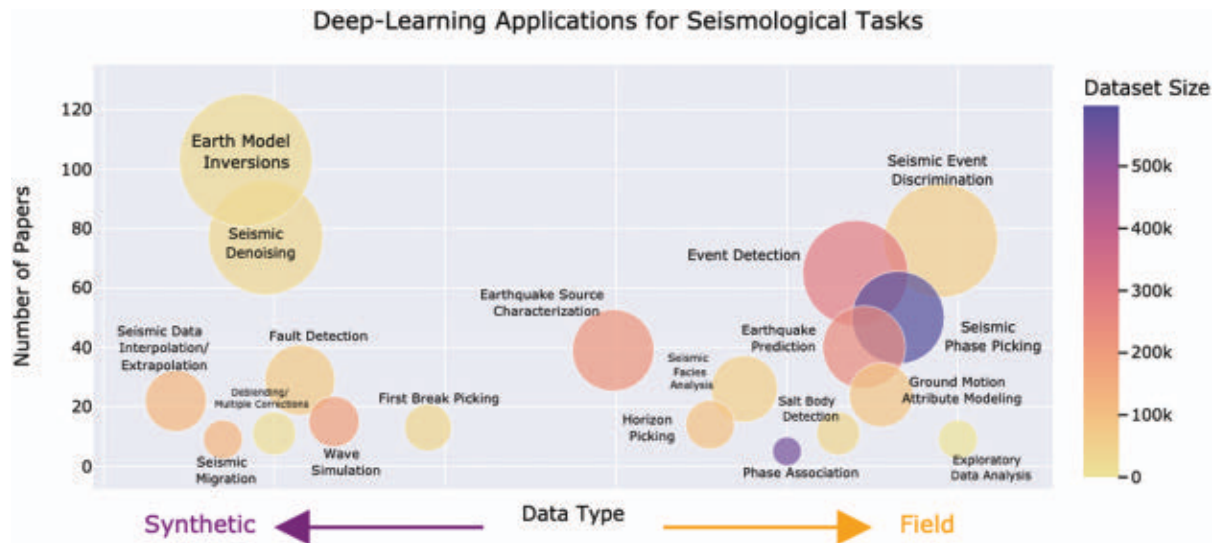


Fig. 3. Deep-learning applications for seismological tasks. The size of each circle is scaled to the number of published papers for each application and color-coded according to the average size of training data. Horizontal axes represent the average percentage of observational and recorded data versus synthetic training data.

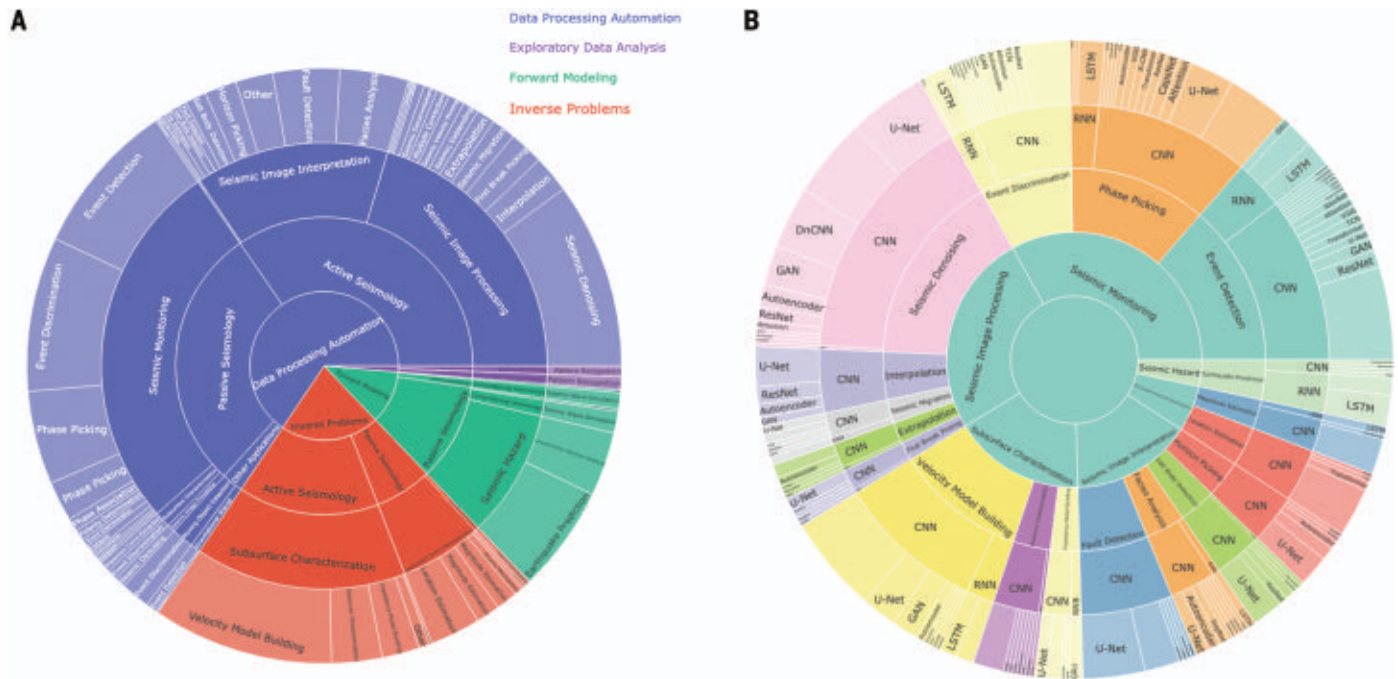


Fig. 4. Hierarchical distribution of studies in our database. (A and B) Studies are distributed according to (A) the seismological applications and (B) used DNNs.

Fine tuning of trained models, especially those trained with synthetic data, is popular in seismology, and we expect that meta-learning will be a growth area in seismology.

CNNs dominate seismological models across many sectors (Fig. 4B) because of their effectiveness for automatic feature extraction and sparse representation learning of seismic data. Some characteristics of seismic data such as their limited bandwidth, sequential nature, or data structure (for example, 3D and 4D imaging in active seismology) have motivated

developments of specialized and hybrid DNN architectures composed of a mixture of neural-layer types. RNNs are popular in tasks in which temporal relations in data play a crucial role in the modeling. Network architectures with contracting encoders and expanding decoders such as U-Net and Autoencoders are highly suitable for modeling seismic data across many applications. GANs are another popular approach for seismic data processing. Among RNNs, long short-term memory (LSTM) units are the most popular network type. The se-

quential nature of seismic data makes RNNs appealing for many seismological tasks. More recently, neural networks with attention mechanisms, dilated CNNs, and physics-informed DNNs are gaining momentum.

Explainable deep learning

Despite the widespread application of deep learning in seismology and its breakthrough performance in multiple areas, the black-box nature remains a source of skepticism. Accepting indecipherable models on their own terms

can contribute to scientific goals of the community (139), but presenting the properties of these models in understandable terms to humans (interpretability) and revealing the underlying reasons for their output decision (explainability) (140) would allay such skepticism and lead to wider adoption and new insights.

Currently, most deep-learning approaches in seismology concern building a model based on known input-output pairs to predict the outputs corresponding to previously unseen inputs without interpretability or explainability. In most of these cases, it is not possible to explain why specific outcomes were obtained from an algorithmic point of view nor what else these models can tell us.

Visualizing the learned kernels at different layers of a CNN is the main approach used in some of the seismological studies (93, 107, 141, 142) to understand something about the representations a deep learning model has learned and how it drives the output; however, the extracted multiscale feature representations from seismic data are too complex for direct visualization to provide interpretable and understandable insights.

Although it may not be possible to elucidate how a deep-learning model works, gaining useful information about learned models is possible by using interpretation tools such as backward propagation techniques, saliency maps, or heatmaps to reveal relevant patterns in the input based on feature importance or relevance scores (143, 144). These interpretable representations of the input can be used along with domain knowledge to explain the output.

Mousavi *et al.* (12), for example, used attention mechanisms in their detection and picker DNN. They used a hierarchy of attention modules to gain insight into the interaction of task-specific decoder branches and where and on what the DNN focused. Doing so gave a degree of interpretability to their results. Attention was used in (54) to build an interpretable deep-learning model for seismic facies analysis. They used 3D spatial-spectral attention maps to reveal the relations between the seismic spectral response (input) and the geological formations (output). Liao *et al.* (22) used a backward propagation technique to compare different deep-learning models for phase-picking on the basis of the learned representations from data.

In addition to explaining the outcomes, designing an interpretable deep-learning model is possible in which the model itself and its components became explainable by using domain knowledge (111, 145). In this approach, the network architecture is designed to provide insight by promoting or restricting specific attributes of the data being modeled (146, 147). Zhu *et al.* (87), for example, used a mixture of Gaussian diffusion

kernels to build an interpretable model of anisotropic spatial dependency in seismicity patterns. The parameters of these kernels are reparameterized by DNNs and learned from training data by using an imitation learning approach. Such an interpretable model can help uncover patterns [and hopefully causal relationships (148)] in data.

Incorporating geophysics into DNNs

Deep-learning approaches that incorporate physical laws have gained momentum in the machine learning community (149) and a growing number of implementations in seismology. The main idea is to integrate data and mathematical physics (domain knowledge) models, even if only partially understood. The key objectives are to provide interpretability and explainability for a deep-learning method, to regularize them in the presence of missing or noisy data, and to enhance their out-of-distribution generalization by restricting the solution space to physically plausible solutions.

Domain knowledge can be incorporated into the training data, the hypothesis space, and the training procedure of a deep-learning model (Fig. 5). Among these, incorporating the domain knowledge into training data through data generation and augmentation by using numerical simulations is the most commonly used approach in seismology for active seismic data processing (47, 150), wave simulation (68), and velocity estimation (97). These models are usually fine tuned through some examples of field data.

The hypothesis space can be constrained with domain knowledge through design of specialized neural network architectures based on the characteristics of the problem. Such

physics-inspired neural network architectures have been developed for seismic FWI (111, 145). Sun *et al.* (111) used a recurrent architecture to model time-dependent dynamics of seismic wave propagation. To enforce physical rules for wave propagation in the model, they set up LSTM-like cells—including some physical operators, which take the wavefield at a past instant as input—estimated the shot record at the current instant as output, and saved the modeled wavefield in the memory cell for calculating the next time step. In these models, the trainable parameters of the DNN can be viewed as a reparameterization of an Earth model. The forward propagation of information through this representation resembles the forward propagation of a seismic wavefield through a heterogeneous medium. Thus, training such a network and updating its weights by using DNN back-propagation and based on the misfit between the simulated and observed seismic data amounts to back-propagation of the residual wavefield and under certain assumptions is equivalent to the optimization process in conventional gradient-based FWI. This can result in a more interpretable model in which the latent representation is physically meaningful.

Physical constraints can be introduced into the training process either by incorporating them into the loss function (implicit approach), by explicit inclusion of a numerical modeling into the training loop, or both. The first approach is used more recently in seismology and primarily for forward problems such as simulation of the pressure wavefield (68) or estimation of first arrival times (77). In these studies, a form of the wave equation (68) or Eikonal equation (77) is used to penalize the

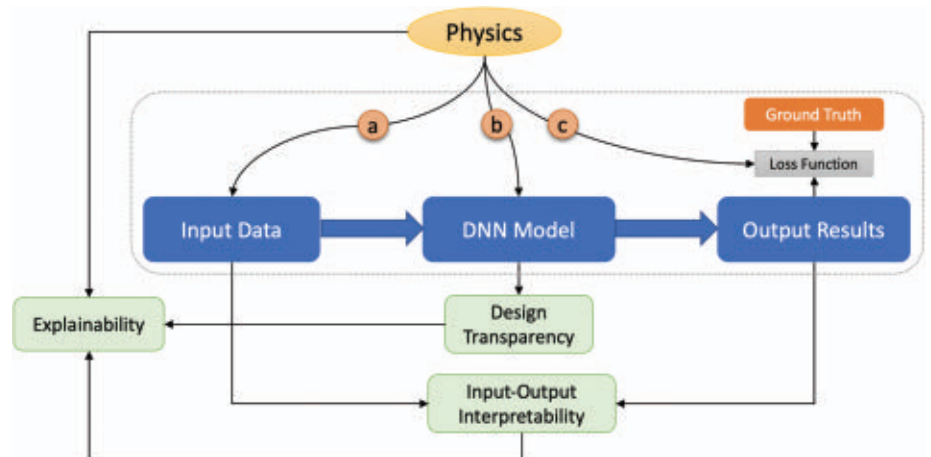


Fig. 5. Explainable DNN. The domain knowledge, physics, can be incorporated into the training data through (A) data simulation, (B) the hypothesis space by designing specialized neural network architectures, and (C) the training procedure of a deep-learning model by enforcing consistency with governing equations. The architecture design based on domain knowledge provides some transparency and interpretability of the model. The interpretable representation of the input data along with some domain knowledge can be used to explain the output.

loss functions used for training the DNNs. Thus, during the training the model is simultaneously optimized to fit the observed data while respecting physical constraints. The examples of the second approach are mostly in seismic FWI in which a physics-based forward modeling of synthetic seismograms is performed as a part of the training loop to embed some physics-based constraints in the model (95, 107, 108, 110). In this approach, the mismatch between observed and numerically simulated seismic data is used to compute the gradient of the misfit function with respect to the DNN weights and update the initial Earth model represented by a DNN.

Physics-guided (or physics-informed) deep learning is effective for solving ill-posed inverse problems as well as for forward problems in seismology. They hold a great potential for wider applications in seismology beyond examples mentioned above. Developing hybrid approaches by combining multiple physical constraints (for example, simultaneously in architecture design and loss function), using physics-based GANs to enforce realism for data generation (151), constraining a model statistically (151), quantifying uncertainties (152), and learning physical operators (153) are other directions of interest.

Uncertainty quantification

As development of deep-learning models for some seismological tasks progresses toward routine operations, quantification of corresponding aleatory and epistemic uncertainties becomes essential to assess the reliability of output predictions. Bayesian DNNs can play a key role in this area by providing probabilistic treatment of neural networks (154). In these models, a prior probability distribution is placed over DNN parameters, and a posterior of these parameters is learned during the training. This strategy can be seen as learning by an infinite number of neural networks given the data in which the inference occurs by aggregating the results through ensembling. The high time complexity of Bayesian DNNs is the main challenge in their practical application; however, advances in scalable and easy-to-implement adaption of Bayesian DNNs are set to change that (155–158).

Monte Carlo dropout sampling (158) is a popular technique for uncertainty estimation because of its flexibility and ease of implementation (57, 94, 118, 159). Mousavi and Beroza (118), for example, developed a Bayesian CNN for earthquake location by using a customized loss function to estimate the aleatory uncertainties directly from data during the training process. The epistemic (model) uncertainties were estimated through dropout during inference. In this approach, the loss function acts like an intelligent loss function and reduces the effect of uncertain data, whether because

of noise or unreliable labels; however, the effectiveness of this approach in estimating model uncertainties is limited.

Recent computational advances (160, 161) for building deep probabilistic models provide alternative options for a more robust estimation of model uncertainties. Feng *et al.* (56) developed a Bayesian deep-learning framework for seismic facies analysis in which the Bayesian variational inference for approximating the posterior distribution of DNN parameters is formulated as an optimization process. In this approach, multiple random samples are drawn from the estimated posterior distribution to quantify the model uncertainties by using the prediction entropy.

Besides uncertainty quantification, the Bayesian DNN approach has some implicit regularization effects that can suppress overfitting, especially when there is insufficient data. We expect that there will be an increased effort to develop probabilistic deep-learning models, especially for tasks such as seismic image interpretation, in which the cost of a wrong decision is high. Developing physics-informed Bayesian models would be one approach in this direction.

Training data limitations

The existence of large-scale realistic training datasets with diverse characteristics and precise labels is crucial for the success of deep learning. Uncertainties and nonuniqueness of manual labels owing to limited resolution, presence of noise, different levels of expertise, cognitive biases, and inherent ambiguity of tasks is a limiting factor in the development of training data in seismology. Two analysts may disagree about the existence or location of a particular fault, or a deep horizon in a seismic image, or arrival times in an earthquake signal when labeling them, which biases training data. A challenging task in building training datasets is quality control of the labels.

In addition to labeling uncertainties, some tasks such as seismic velocity inversion, image processing, and interpretation face the challenge of a scarcity of public large-scale labeled data from field observations. One approach to address this limitation is to generate physically realistic synthetic data that represent a variety of complex geological settings (162). For seismic velocity inversion, Wang *et al.* (163) used smoothed versions of natural images as seismic velocity models to produce synthetic seismic traces to use as training data. Feng *et al.* (164) developed a real-time style transform method to create a large set of physically realistic subsurface velocity maps from natural images. They used these data in their multi-scale deep learning-based FWI. Araya-Polo *et al.* (165) used GANs to generate a large number of geologically feasible models to augment their original training data set.

Another approach uses inductive transfer learning. This was used to fine-tune pre-trained CNN models on handwritten images (8) or natural images (166) to classify volcano-seismic events according to their time-frequency representations. Developing semisupervised (20, 95, 167) and unsupervised (19, 35, 110) approaches for reducing a need for labels is a growing trend. Other options are adding constraints to the training procedure (167, 168) and reducing the degrees of freedom (the number of trainable parameters) in supervised approaches by incorporating physics-based constraints.

Out-of-distribution generalization

To evaluate the quality of the obtained model in most seismological applications, authors usually perform a blind test on a subset of the data not used in training and/or samples of data from other datasets (or regions). A blind-test dataset that is representative of the full diversity of the real world is needed to evaluate the generalization of a model robustly. Statistical significance of such tests is usually not presented, and it is unclear to what extent a trained network can generalize.

For applications such as earthquake signal detection and phase picking, in which the model generalizations have been studied extensively (12, 22, 23, 25, 29, 169), neural network architecture, neural network type, and the size of a training set have been shown to be not the only defining factors for good out-of-distribution generalization (12) and that hyperparameter tuning (29), training procedure, and data augmentation (170) also play important roles. These effects have seen less attention in seismic image denoising, processing, and interpretation, in which model building is often done by using the synthetic data and data augmentation is limited (in part because some common image augmentation techniques, such as rotation, result in physically implausible models).

Generalization of a deep-learning model always has a limit because each model may have blind spots, and out-of-distribution cases can always arise. This makes it difficult to have a distinct and robust measure of the generalizability. It remains unclear how much improvement might be obtained by using physics-guided DNNs or Bayesian frameworks. More systematic studies are required to understand and characterize this important property of deep-learning models.

Benchmarking pitfalls and concerning trends

Despite the recent increase of applications of deep learning in seismology, not all the reported improvements lead to real-world performance increases. Often it is not clear what worked and why or which deep-learning approaches actually lead to better performance.

The choice of weak or often not properly fine-tuned baselines, different datasets, data pre-processing, evaluation protocols, performance metrics, the lack of result reproducibility, ablation studies, error analysis, and robustness checks are all factors that make it difficult to assess progress. To overcome this concerning trend and avoid the “phantom progress” phenomena observed in other fields (171), encouragement and ultimately enforcement of standard benchmarking need to be a necessary foundation of empirical machine learning research.

The implementation of benchmarking is currently diverse and inconsistent across deep-learning seismology. Applications such as subsurface characterization have well-established benchmarks, such as the Marmousi (172) or BP2004 (173) models that are widely used; but even in this case, performance evaluation can be qualitative. Applications such as event discrimination and seismic denoising have seen a relatively better adaptation of reporting metrics (for example, precision and recall, SNR, or mean-squared error) but suffer from the lack of canonical benchmarks. Because this is still a young field, there is no standard benchmark or evaluation setup for most of the deep-learning applications in seismology. As a result, researchers choose their own test data and evaluation criteria to showcase their method and (perhaps unknowingly) tend to find good performance for their models. This lack of community-driven sets of benchmarks and evaluation guidelines hinders progress in the field.

Given the indispensable role of benchmarking for measuring progress, the community involved in deep-learning applications needs to set up more unified, standard, and rigorous evaluations that will ultimately lead to the development of insight and understanding because the community also are those who will endorse and enforce these practices. This requires first determining the guidelines for the reproducibility of the results, the choice of the baselines, building of benchmark datasets, and defining of task-specific metrics and standards for reporting the results. When creating a benchmark dataset for machine learning applications, documentation of its composition, collection process, labeling uncertainties, and applicable tasks, among other criteria, are all important. Only a few such benchmarks currently exist in seismology (174–177) specifically designed for machine learning. Evaluating a model on multiple benchmarks clearly would provide a more robust performance measurement. In this spirit, setting up guidelines on how to use benchmarks is also important. Clear reporting of the training data and its size, model size, time and memory complexities, inference latency, suitable metrics indicating model

variance, its generalization to the out-of-distribution data, amount of hyperparameter tuning, and a discussion of statistical significance of the results are all important factors that have received too little attention in the seismological community. Setting up clear guidelines for reviewers is the next step that could help promote and enforce greater rigor in the field and render more accurate and fair judgments when comparing models.

Conclusions

Deep learning has entered almost every sub-field of seismology and has the potential to dominate algorithm development. Yet as a general-purpose tool for scientific discovery, deep learning is in its infancy. Investing more effort on physics-based and explainable deep learning will help the field to mature. Research is rapidly moving from proof-of-concept efforts to real-world applications, and estimation of uncertainty will become essential. Given common pitfalls and concerning trends, setting up standard benchmarking and evaluation protocols is critical. An understanding of the limitations and breaking points of these models will be extremely helpful. Despite these challenges, deep-learning approaches will clearly contribute substantively to seismology by optimizing results, providing value, and sparking inspiration. Given the incredible ability of DNNs for modeling and processing large-volume, high-dimensional seismic data and their remarkable performance in varied seismological tasks, deep learning will soon become the norm in multiple seismic data analyses, notwithstanding issues of interpretability and brittleness. Deep learning will not replace expert seismological analysts, but those analysts who use it will outperform those who do not for many types of problems.

REFERENCES AND NOTES

1. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015). doi: 10.1038/nature14539; PMID: 26017442
2. L. Seydoux et al., Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. *Nat. Commun.* **11**, 3972 (2020). doi: 10.1038/s41467-020-17841-x; PMID: 32769972
3. S. M. Mousavi, G. Beroza, A dataset of published journal papers using neural networks for seismological tasks. *Zenodo* (2022); doi: 10.5281/zenodo.6386952
4. W. K. Lee, M. Celebi, M. Todorovska, H. Igel, Introduction to the special issue on rotational seismology and engineering applications. *Bull. Seismol. Soc. Am.* **99** (2B), 945–957 (2009). doi: 10.1785/0120080344
5. E. R. Martin, N. Lindsey, J. Ajo-Franklin, B. Biondi, in *Distributed Acoustic Sensing in Geophysics: Methods and Applications* (American Geophysical Union, 2021), pp. 111–129.
6. L. Linville, K. Pankow, T. Draelos, Deep learning models augment analyst decisions for event discrimination. *Geophys. Res. Lett.* **46**, 3643–3651 (2019). doi: 10.1029/2018GL081119
7. B. Ku, G. Kim, J.-K. Ahn, J. Lee, H. Ko, *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5 (2020). doi: 10.1109/LGRS.2022.3143118
8. M. Titos, A. Bueno, L. Garcia, C. Benitez, J. C. Segura, Classification of isolated volcano-seismic events based on inductive transfer learning. *IEEE Geosci. Remote Sens. Lett.* **17**, 869–873 (2020). doi: 10.1109/LGRS.2019.2931063

9. T. Perol, M. Gharbi, M. Denolle, Convolutional neural network for earthquake detection and location. *Sci. Adv.* **4**, e1700578 (2018). doi: 10.1126/sciadv.1700578; PMID: 29487899
10. Z. E. Ross, M.-A. Meier, E. Hauksson, T. H. Heaton, Generalized seismic phase detection with deep learning. *Bull. Seismol. Soc. Am.* **108**, 2894–2901 (2018). doi: 10.1785/0120180080
11. W. Zhu, G. C. Beroza, Passive seismic imaging of subwavelength natural fractures: Theory and 2-D synthetic and ultrasonic data tests. *Geophys. J. Int.* **216**, 1831 (2019). doi: 10.1093/gji/ggy529
12. S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, G. C. Beroza, Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* **11**, 3952 (2020). doi: 10.1038/s41467-020-17591-w; PMID: 32770023
13. Z. E. Ross, M.-A. Meier, E. Hauksson, P Wave arrival picking and first-motion polarity determination with deep learning. *J. Geophys. Res. Solid Earth* **123**, 5120–5129 (2018). doi: 10.1029/2017JB015251
14. X. Tian et al., Comparison of single-trace and multiple-trace polarity determination for surface microseismic data using deep learning. *Seismol. Res. Lett.* **91**, 1794–1803 (2020). doi: 10.1785/0220190353
15. Z. E. Ross, Y. Yue, M.-A. Meier, E. Hauksson, T. H. Heaton, PhaseLink: A deep learning approach to seismic phase association. *J. Geophys. Res. Solid Earth* **124**, 856–869 (2019). doi: 10.1029/2018JB016674
16. I. W. McBrearty, A. A. Delorey, P. A. Johnson, Pairwise association of seismic arrivals with convolutional neural networks. *Seismol. Res. Lett.* **90** (2A), 503–509 (2019). doi: 10.1785/0220180326
17. P. Peng, Z. He, L. Wang, Y. Jiang, Microseismic records classification using capsule network with limited training samples in underground mining. *Sci. Rep.* **10**, 13925 (2020). doi: 10.1038/s41598-020-70916-z; PMID: 32811883
18. M. Nakano, D. Sugiyama, T. Hori, T. Kuwatani, S. Tsuboi, Discrimination of seismic signals from earthquakes and tectonic tremor by applying a convolutional neural network to running spectral images. *Seismol. Res. Lett.* **90**, 530–538 (2019). doi: 10.1785/0220180279
19. S. M. Mousavi, W. Zhu, W. Ellsworth, G. Beroza, Unsupervised clustering of seismic signals using deep convolutional autoencoders. *IEEE Geosci. Remote Sens. Lett.* **16**, 1693–1697 (2019). doi: 10.1109/LGRS.2019.2909218
20. L. Linville, D. Anderson, J. Michalenko, J. Galasso, T. Draelos, *Seismol. Soc. Am.* **92**, 388 (2021).
21. S. M. Mousavi, W. Zhu, Y. Sheng, G. C. Beroza, *Sci. Rep.* **9**, 1 (2019). doi: 10.1038/s41598-018-37186-2; PMID: 30626917
22. W.-Y. Liao, E.-J. Lee, D. Mu, P. Chen, R.-J. Rau, *Seismol. Res. Lett.* **10.1785/02202010274** (2021). doi: 10.1785/02202010274
23. O. M. Saad, Y. Chen, *IEEE Trans. Geosci. Remote Sens.* **69**, 709–726 (2021).
24. G. Kim, B. Ku, H. Ko, Multifeature fusion-based earthquake event classification using transfer learning. *IEEE Geosci. Remote Sens. Lett.* **18**, 974–978 (2020). doi: 10.1109/LGRS.2020.2993302
25. S. Yang, J. Hu, H. Zhang, G. Liu, Simultaneous earthquake detection on multiple stations via a convolutional neural network. *Seismol. Res. Lett.* **92**, 246–260 (2021). doi: 10.1785/0220200137
26. K. Yano et al., *J. Geophys. Res.* **126**, e2020JB020269 (2021).
27. K. Fauvel et al., *Proc. AAAI Conf. AI* **34**, 403–411 (2020).
28. Y. Zhou, H. Yue, Q. Kong, S. Zhou, Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks. *Seismol. Res. Lett.* **90**, 1079–1087 (2019). doi: 10.1785/0220180319
29. H. Soto, B. Schurr, *Geophys. J. Int.* **227**, 1268 (2021).
30. X. Huang, J. Lee, Y.-W. Kwon, C.-H. Lee, in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2020), pp. 3261–3271.
31. A. L. Stork et al., Application of machine learning to microseismic event detection in distributed acoustic sensing data. *Geophysics* **85**, KS149–KS160 (2020). doi: 10.1190/geo2019-0774.1
32. W. Zhu, S. M. Mousavi, G. C. Beroza, Seismic signal denoising and decomposition using deep neural networks. *IEEE Trans. Geosci. Remote Sens.* **57**, 9476–9488 (2019). doi: 10.1109/TGRS.2019.2926772
33. Z. Li, N. Sun, H. Gao, N. Qin, Z. Li, Adaptive subtraction based on U-Net for removing seismic multiples. *IEEE Trans. Geosci. Remote Sens.* **59**, 9796–9812 (2021). doi: 10.1109/TGRS.2021.3051303

34. B. Wang, J. Li, J. Luo, Y. Wang, J. Geng, Intelligent debinding of seismic data based on U-Net and transfer learning. *IEEE Trans. Geosci. Remote Sens.* **59**, 8885–8894 (2021). doi: [10.1109/TGRS.2020.3048746](https://doi.org/10.1109/TGRS.2020.3048746)
35. M. Zhang, Y. Liu, M. Bai, Y. Chen, Seismic noise attenuation using unsupervised sparse feature learning. *IEEE Trans. Geosci. Remote Sens.* **57**, 9709–9723 (2019). doi: [10.1109/TGRS.2019.2928715](https://doi.org/10.1109/TGRS.2019.2928715)
36. Y. Li, H. Wang, X. Dong, *IEEE Geosci. Remote Sens. Lett.* **18**, 2016–2020 (2020). doi: [10.1109/LGRS.2020.3011130](https://doi.org/10.1109/LGRS.2020.3011130)
37. J. You, Y. Xue, J. Cao, C. Li, Attenuation of seismic swell noise using convolutional neural networks in frequency domain and transfer learning. *Interpretation (Tulsa)* **8**, T941–T952 (2020). doi: [10.1190/INT-2019-0303.1](https://doi.org/10.1190/INT-2019-0303.1)
38. H. Kaur, S. Fomel, N. Pham, Seismic ground-roll noise attenuation using deep learning. *Geophys. Prospect.* **68**, 2064–2077 (2020). doi: [10.1111/1365-2478.12985](https://doi.org/10.1111/1365-2478.12985)
39. Q. Feng, Y. Li, *IEEE Trans. Geosci. Remote Sens.* **60**, 1–11 (2021). doi: [10.1109/TGRS.2021.3071189](https://doi.org/10.1109/TGRS.2021.3071189)
40. O. M. Saad, Y. Chen, Deep denoising autoencoder for seismic random noise attenuation. *Geophysics* **85**, V367 (2020). doi: [10.1190/geo2019-0468.1](https://doi.org/10.1190/geo2019-0468.1)
41. E. Wang, J. Nealon, Applying machine learning to 3D seismic image denoising and enhancement. *Interpretation 7*, SE131 (2019). doi: [10.1190/INT-2018-0224.1](https://doi.org/10.1190/INT-2018-0224.1)
42. W. Jiang, J. Zhang, L. Bell, 3D seismic geometry quality control and corrections by applying machine learning. *Geophysics* **84**, P87 (2019). doi: [10.1190/geo2018-0617.1](https://doi.org/10.1190/geo2018-0617.1)
43. X. Chai, G. Tang, S. Wang, K. Lin, R. Peng, Deep learning for irregularly and regularly missing 3-D data reconstruction. *IEEE Trans. Geosci. Remote Sens.* **59**, 6244–6265 (2020). doi: [10.1109/TGRS.2020.3016343](https://doi.org/10.1109/TGRS.2020.3016343)
44. O. Ovcharenko, V. Kazei, M. Kalita, D. Peter, T. Alkhalifah, Deep learning for low-frequency extrapolation from multioffset seismic data. *Geophysics* **84**, R989–R1001 (2019). doi: [10.1190/geo2018-0884.1](https://doi.org/10.1190/geo2018-0884.1)
45. Y. Li, J. Song, W. Lu, P. Monkam, Y. Ao, Multitask learning for super-resolution of seismic velocity model. *IEEE Trans. Geosci. Remote Sens.* **59**, 8022–8033 (2020). doi: [10.1109/TGRS.2020.3034502](https://doi.org/10.1109/TGRS.2020.3034502)
46. S. Yuan, J. Liu, S. Wang, T. Wang, P. Shi, Seismic waveform classification and first-break picking using convolution neural networks. *IEEE Geosci. Remote Sens. Lett.* **15**, 272–276 (2018). doi: [10.1109/LGRS.2017.2785834](https://doi.org/10.1109/LGRS.2017.2785834)
47. M. J. Park, M. D. Sacchi, Automatic velocity analysis using convolutional neural network and transfer learning. *Geophysics* **85**, V33–V43 (2020). doi: [10.1190/geo2018-0870.1](https://doi.org/10.1190/geo2018-0870.1)
48. S. Dong, Z. Li, X. Chen, L. Fu, *Bull. Seismol. Soc. Am.* **111**, 3420–3431 (2021).
49. W. Liu, Q. Cheng, L. Liu, Y. Wang, J. Zhang, Accelerating high-resolution seismic imaging by using deep learning. *Appl. Sci.* **10**, 2502 (2020). doi: [10.3390/app10072502](https://doi.org/10.3390/app10072502)
50. H. Kaur, N. Pham, S. Fomel, Improving the resolution of migrated images by approximating the inverse Hessian using deep learning. *Geophysics* **85**, WA173–WA183 (2020). doi: [10.1190/geo2019-0315.1](https://doi.org/10.1190/geo2019-0315.1)
51. F. Qian *et al.*, Unsupervised seismic facies analysis via deep convolutional autoencoders. *Geophysics* **83**, A39–A43 (2018). doi: [10.1190/geo2017-0524.1](https://doi.org/10.1190/geo2017-0524.1)
52. Y. Duan, X. Zheng, L. Hu, L. Sun, Seismic facies analysis based on deep convolutional embedded clustering. *Geophysics* **84**, IM87–IM97 (2019). doi: [10.1190/geo2018-0789.1](https://doi.org/10.1190/geo2018-0789.1)
53. D. Grana, L. Azevedo, M. Liu, A comparison of deep machine learning and Monte Carlo methods for facies classification from seismic data. *Geophysics* **85**, WA41–WA52 (2020). doi: [10.1190/geo2019-0405.1](https://doi.org/10.1190/geo2019-0405.1)
54. F. Li, H. Zhou, Z. Wang, X. Wu, *IEEE Trans. Geosci. Remote Sens.* **59**, 1733–1744 (2020). doi: [10.1109/TGRS.2020.2999365](https://doi.org/10.1109/TGRS.2020.2999365)
55. M. Liu, M. Jervis, W. Li, P. Nivlet, Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks. *Geophysics* **85**, O47–O58 (2020). doi: [10.1190/geo2019-0627.1](https://doi.org/10.1190/geo2019-0627.1)
56. R. Feng, N. Balling, D. Grana, J. S. Dramsch, T. M. Hansen, Bayesian convolutional neural networks for seismic facies classification. *IEEE Trans. Geosci. Remote Sens.* **59**, 8933–8940 (2021). doi: [10.1109/TGRS.2020.3049012](https://doi.org/10.1109/TGRS.2020.3049012)
57. N. Pham, S. Fomel, D. Dunlap, Automatic channel detection using deep learning. *Interpretation (Tulsa)* **7**, SE43–SE50 (2019). doi: [10.1190/INT-2018-0202.1](https://doi.org/10.1190/INT-2018-0202.1)
58. J. El Zini, Y. Rizk, M. Awad, A deep transfer learning framework for seismic data analysis: A case study on bright spot detection. *IEEE Trans. Geosci. Remote Sens.* **58**, 3202–3212 (2019). doi: [10.1109/TGRS.2019.2950888](https://doi.org/10.1109/TGRS.2019.2950888)
59. X. Wu, S. Yan, J. Qi, H. Zeng, *J. Geophys. Res. Solid Earth* **125**, e2020 (2020).
60. V. Tschannen, N. Ettrich, M. Delescluse, J. Keuper, Detection of point scatterers using diffraction imaging and deep learning. *Geophys. Prospect.* **68**, 830–844 (2020). doi: [10.1111/1365-2478.12889](https://doi.org/10.1111/1365-2478.12889)
61. Z. Geng, Y. Wang, *Nat. Commun.* **11**, 1 (2020). doi: [10.1038/s41467-019-13993-7](https://doi.org/10.1038/s41467-019-13993-7); pmid: 31911652
62. Y. Shi, X. Wu, S. Fomel, SaltSeg: Automatic 3D salt segmentation using a deep convolutional neural network. *Interpretation 7*, SE113–SE122 (2019). doi: [10.1190/INT-2018-0235.1](https://doi.org/10.1190/INT-2018-0235.1)
63. B. Peters, J. Granek, E. Haber, Multiresolution neural networks for tracking seismic horizons from few training images. *Interpretation 7*, SE201–SE213 (2019). doi: [10.1190/INT-2018-0225.1](https://doi.org/10.1190/INT-2018-0225.1)
64. H. Di, Z. Li, H. Maniar, A. Abubakar, Seismic stratigraphy interpretation by deep convolutional neural networks: A semisupervised workflow. *Geophysics* **85**, WA77–WA86 (2020). doi: [10.1190/geo2019-0433.1](https://doi.org/10.1190/geo2019-0433.1)
65. X. Wu, L. Liang, Y. Shi, S. Fomel, FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics* **84**, IM35–IM45 (2019). doi: [10.1190/geo2018-0646.1](https://doi.org/10.1190/geo2018-0646.1)
66. K. Gao, L. Huang, Y. Zheng, *IEEE Trans. Geosci. Remote Sens.* **60**, 1–15 (2021). doi: [10.1109/TGRS.2021.3073840](https://doi.org/10.1109/TGRS.2021.3073840)
67. Z. Wang, B. Li, N. Liu, B. Wu, X. Zhu, *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5 (2020). doi: [10.1109/LGRS.2020.3034960](https://doi.org/10.1109/LGRS.2020.3034960)
68. B. Moseley, A. Markham, T. Nissen-Meyer, *arXiv:2006.11894* [physics.comp-ph] (2020).
69. B. Moseley, A. Markham, T. Nissen-Meyer, *arXiv:2107.07871* [physics.comp-ph] (2021).
70. C. Song, T. Alkhalifah, U. b. Waheed, *arXiv:2008.01865* [physics.comp-ph] (2020).
71. C. Song, T. Alkhalifah, U. B. Waheed, A versatile framework to solve the Helmholtz equation using physics-informed neural networks. *Geophys. J. Int.* **228**, 1750–1762 (2022). doi: [10.1093/gji/ggab434](https://doi.org/10.1093/gji/ggab434)
72. T. Wang, D. Trugman, Y. Lin, *J. Geophys. Res. Solid Earth* **126**, e2020 (2021).
73. F. Gatti, D. Clouteau, Towards blending physics-based numerical simulations and seismic databases using generative adversarial network. *Comput. Methods Appl. Mech. Eng.* **372**, 113421 (2020). doi: [10.1016/j.cma.2020.113421](https://doi.org/10.1016/j.cma.2020.113421)
74. M. A. Florez *et al.*, *arXiv:2011.09038* [physics.geo-ph] (2020).
75. T.-Y. Hsu, C.-W. Huang, *Front. Earth Sci.* **9**, 247 (2021).
76. J. Münchmeyer, D. Bindi, U. Leser, F. Tilmann, The transformer earthquake alerting model: A new versatile approach to earthquake early warning. *Geophysics. J. Int.* **225**, 646–656 (2021). doi: [10.1093/gji/ggab609](https://doi.org/10.1093/gji/ggab609)
77. J. D. Smith, K. Azizzadenesheli, Z. E. Ross, EikoNet: Solving the Eikonal equation with deep neural networks. *IEEE Trans. Geosci. Remote Sens.* **59**, 10685–10696 (2020). doi: [10.1109/TGRS.2020.3039165](https://doi.org/10.1109/TGRS.2020.3039165)
78. F. Khosravik, P. Clayton, Z. Nagy, Artificial neural network-based framework for developing ground-motion models for natural and induced earthquakes in Oklahoma, Kansas, and Texas. *Seismol. Res. Lett.* **90** (2A), 604–613 (2019). doi: [10.1785/0220180218](https://doi.org/10.1785/0220180218)
79. K. B. Withers, M. P. Moschetti, E. M. Thompson, A machine learning approach to developing ground motion models from simulated ground motions. *Geophys. Res. Lett.* **47**, e2019GL086690 (2020). doi: [10.1029/2019GL086690](https://doi.org/10.1029/2019GL086690)
80. D. Jozinovic, A. Lomax, I. Stajduhar, A. Michelini, Transfer learning: Improving neural network based prediction of earthquake ground shaking for an area with insufficient training data. *Geophys. J. Int.* **229**, 704–718 (2020). doi: [10.1093/gji/ggab488](https://doi.org/10.1093/gji/ggab488)
81. P. M. R. DeVries, F. Viégas, M. Wattenberg, B. J. Meade, Deep learning of aftershock patterns following large earthquakes. *Nature* **560**, 632–634 (2018). doi: [10.1038/s41586-018-0438-y](https://doi.org/10.1038/s41586-018-0438-y); pmid: 30158606
82. B. Feng, G. C. Fox, *arXiv:2012.14336* [physics.geo-ph] (2020).
83. Q. Wang, Y. Guo, L. Yu, P. Li, Earthquake prediction based on spatio-temporal data mining: An LSTM network approach. *IEEE Trans. Emerg. Top. Comput.* **8**, 148–158 (2017). doi: [10.1109/TETC.2017.2699169](https://doi.org/10.1109/TETC.2017.2699169)
84. A. Mignan, M. Broccardo, Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations. *Seismol. Res. Lett.* **91**, 2330–2342 (2020). doi: [10.1785/0220200021](https://doi.org/10.1785/0220200021)
85. M. Segou, The physics of earthquake forecasting. *Seismol. Res. Lett.* **91**, 1936–1939 (2020). doi: [10.1785/0220200127](https://doi.org/10.1785/0220200127)
86. G. C. Beroza, M. Segou, S. Mostafa Mousavi, Machine learning and earthquake forecasting-next steps. *Nat. Commun.* **12**, 4761 (2021). doi: [10.1038/s41467-021-24952-6](https://doi.org/10.1038/s41467-021-24952-6); pmid: 34362887
87. S. Zhu, S. Li, Z. Peng, Y. Xie, *arXiv:1906.05467* [cs.LG] (2019).
88. P. A. Johnson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011362118 (2021).
89. Richtex; <https://www.richtex.com> (accessed 30 October 2020).
90. Aeta; <https://aeta.io> (accessed 30 October 2020).
91. M. Araya-Polo, J. Jennings, A. Adler, T. Dahlke, Deep-learning tomography. *Leading Edge (Tulsa Okla.)* **37**, 58–66 (2018). doi: [10.1190/le37010058.1](https://doi.org/10.1190/le37010058.1)
92. Y. Zheng, Q. Zhang, A. Yusifov, Y. Shi, Applications of supervised deep learning for seismic interpretation and inversion. *Leading Edge* **38**, 526–533 (2019). doi: [10.1190/le38070526.1](https://doi.org/10.1190/le38070526.1)
93. S. Li *et al.*, Deep-learning inversion of seismic data. *IEEE Trans. Geosci. Remote Sens.* **58**, 2135–2149 (2019). doi: [10.1109/TGRS.2019.2953473](https://doi.org/10.1109/TGRS.2019.2953473)
94. V. Das, A. Pollack, U. Wollner, T. Mukerji, Convolutional neural network for seismic impedance inversion. *Geophysics* **84**, R869–R880 (2019). doi: [10.1190/geo2018-0838.1](https://doi.org/10.1190/geo2018-0838.1)
95. M. Alfarraj, G. AlRegib, Semisupervised sequence modeling for elastic impedance inversion. *Interpretation 7*, SE237–SE249 (2019). doi: [10.1190/INT-2018-0250.1](https://doi.org/10.1190/INT-2018-0250.1)
96. V. Das, T. Mukerji, Petrophysical properties prediction from prestack seismic data using convolutional neural networks. *Geophysics* **85**, N41–N55 (2020). doi: [10.1190/geo2019-0650.1](https://doi.org/10.1190/geo2019-0650.1)
97. W. Weinzierl, B. Wiese, Deep learning a poroelastic rock-physics model for pressure and saturation discrimination. *Geophysics* **86**, MR53–MR66 (2021). doi: [10.1190/geo2020-0049.1](https://doi.org/10.1190/geo2020-0049.1)
98. Z. Zhong, A. Y. Sun, X. Wu, *J. Geophys. Res. Solid Earth* **125**, e2019 (2020).
99. D. Li, S. Peng, Y. Guo, Y. Lu, X. Cui, CO₂ storage monitoring based on time-lapse seismic data via deep learning. *Int. J. Greenh. Gas Control* **108**, 103336 (2021). doi: [10.1016/j.jggc.2021.103336](https://doi.org/10.1016/j.jggc.2021.103336)
100. T. Lähivaara, A. *et al.*, Estimation of groundwater storage from seismic data using deep learning. *Geophys. Prospect.* **67**, 2115–2126 (2019). doi: [10.1111/1365-2478.12831](https://doi.org/10.1111/1365-2478.12831)
101. H. Zhang, J. Han, Z. Li, H. Zhang, *IEEE Geosci. Remote Sens. Lett.* **10**, 1109/LGRS.2020.3048171 (2021).
102. W. Zhang, J. Gao, *IEEE Trans. Geosci. Remote Sens.* **60**, 1–18 (2021). doi: [10.1109/TGRS.2021.3062688](https://doi.org/10.1109/TGRS.2021.3062688)
103. B. Sun, T. Alkhalifah, ML-descent: An optimization algorithm for full-waveform inversion using machine learning. *Geophysics* **85**, R477–R492 (2020). doi: [10.1190/geo2019-0641.1](https://doi.org/10.1190/geo2019-0641.1)
104. Z. Gao *et al.*, *IEEE Geosci. Remote Sens. Lett.* **31**, 3449 (2020).
105. M. Aleari, Combining discrete cosine transform and convolutional neural networks to speed up the Hamiltonian Monte Carlo inversion of pre-stack seismic data. *Geophys. Prospect.* **68**, 2738–2761 (2020). doi: [10.1111/1365-2478.13025](https://doi.org/10.1111/1365-2478.13025)
106. Y. Chen, G. T. Schuster, Seismic inversion by Newtonian machine learning. *Geophysics* **85**, WA185–WA200 (2020). doi: [10.1190/geo2019-0434.1](https://doi.org/10.1190/geo2019-0434.1)
107. Y. Wu, G. A. McMechan, Parametric convolutional neural network-domain full-waveform inversion. *Geophysics* **84**, R881–R896 (2019). doi: [10.1190/geo2018-0224.1](https://doi.org/10.1190/geo2018-0224.1)
108. L. Mosser, O. Dubrule, M. J. Blunt, Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. *Math. Geosci.* **52**, 53–79 (2020). doi: [10.1007/s11004-019-09832-6](https://doi.org/10.1007/s11004-019-09832-6)
109. Z. Gao *et al.*, *IEEE Trans. Geosci. Remote Sens.* **59**, 2269–2280 (2020). doi: [10.1109/TGRS.2020.3000684](https://doi.org/10.1109/TGRS.2020.3000684)
110. R. Biswas, M. K. Sen, V. Das, T. Mukerji, Prestack and poststack inversion using a physics-guided convolutional neural network. *Interpretation 7*, SE161–SE174 (2019). doi: [10.1190/INT-2018-0236.1](https://doi.org/10.1190/INT-2018-0236.1)
111. J. Sun, Z. Niu, K. A. Innanen, J. Li, D. O. Trad, A theory-guided deep-learning formulation and optimization of seismic waveform inversion. *Geophysics* **85**, R87 (2020). doi: [10.1190/geo2019-0138.1](https://doi.org/10.1190/geo2019-0138.1)

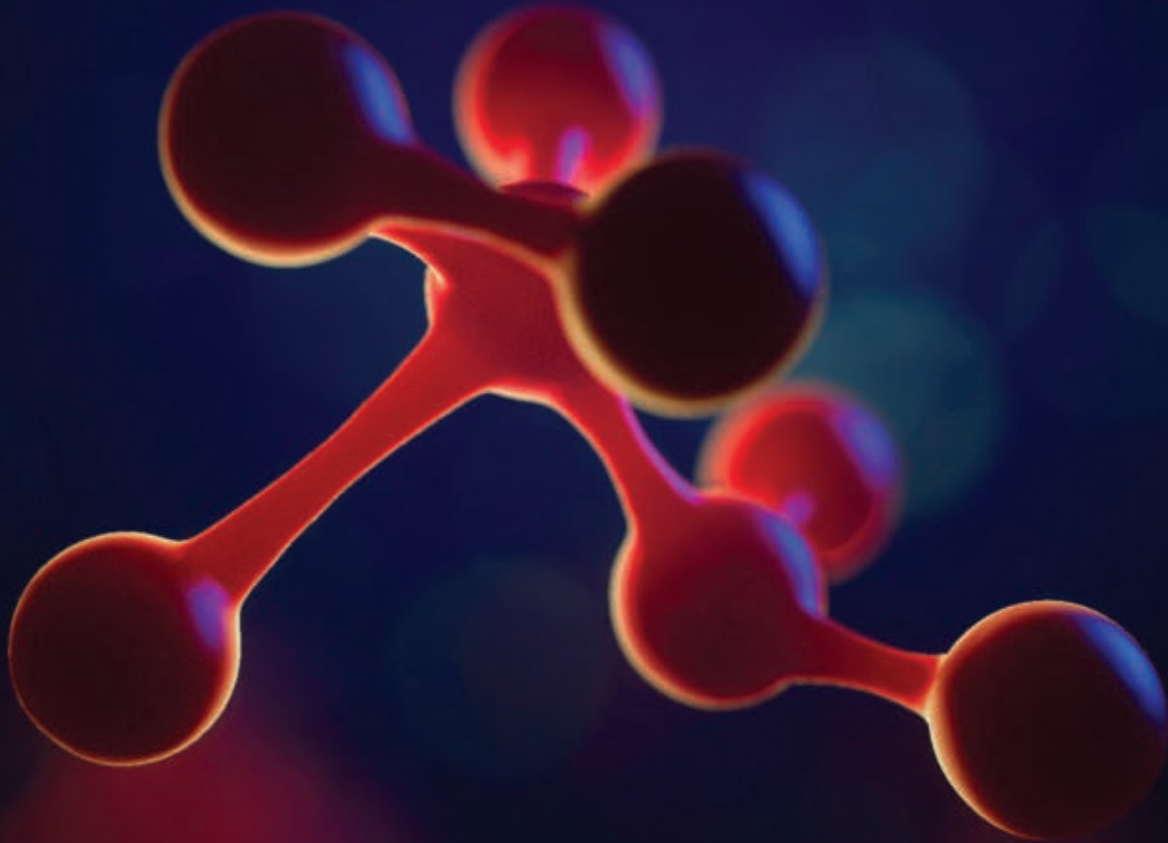
112. K. Kothari, M. de Hoop, I. Dokmanic, *arXiv:2006.05854* [eess.IV] (2020).
113. S. Gupta, K. Kothari, M. V. de Hoop, I. Dokmanic, *arXiv:1805.11718* [cs.CV] (2018).
114. A. Adler, M. Araya-Polo, T. Poggio, Deep learning for seismic inverse problems: Toward the acceleration of geophysical analysis workflows. *IEEE Signal Process. Mag.* **38**, 89–119 (2021). doi: [10.1109/MSP.2020.3037429](https://doi.org/10.1109/MSP.2020.3037429)
115. M. P. van den Ende, J.-P. Ampuero, *Geophys. Res. Lett.* **47**, e2020 (2020).
116. J. Münchmeyer, D. Bindi, U. Leser, F. Tilmann, Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. *Geophys. J. Int.* **226**, 1086–1104 (2021). doi: [10.1093/gji/ggab139](https://doi.org/10.1093/gji/ggab139)
117. A. Steinberg, H. Vasyura-Bathke, P. Gaebler, M. Ohrnberger, L. Ceranna, *JGR Solid Earth* **126**, e2021JB022685 (2021).
118. S. Mousavi, G. C. M. Beroza, Bayesian-deep-learning estimation of earthquake location from single-station observations. *IEEE Geosci. Remote Sens. Lett.* **58**, 8211–8224 (2020).
119. S. M. Mousavi, G. C. Beroza, A machine-learning approach for earthquake magnitude estimation. *Geophys. Res. Lett.* **47**, 2019 (2020). doi: [10.1029/2019GL085976](https://doi.org/10.1029/2019GL085976)
120. H. Zhang, K. A. Innanen, D. W. Eaton, Inversion for shear-tensile focal mechanisms using an unsupervised physics-guided neural network. *Seismol. Res. Lett.* **92**, 2282–2294 (2021). doi: [10.1785/0220200420](https://doi.org/10.1785/0220200420)
121. J. D. Smith, Z. E. Ross, K. Aizzadenesheli, J. B. Muir, HypoSVI: Hypocentre inversion with Stein variational inference and physics informed neural networks. *Geophys. J. Int.* **228**, 698–710 (2021). doi: [10.1093/gji/ggab309](https://doi.org/10.1093/gji/ggab309)
122. S. Barkaoui et al., Anatomy of continuous Mars SEIS and pressure data from unsupervised learning. *Bull. Seismol. Soc. Am.* **111**, 2964–2981 (2021). doi: [10.1785/0120210095](https://doi.org/10.1785/0120210095)
123. D. Snover, C. W. Johnson, M. J. Bianco, P. Gerstoft, *Seismolog. Soc. Am.* **92**, 1011 (2021).
124. E. Min et al., A survey of clustering with deep learning: From the perspective of network architecture. *IEEE Access* **6**, 39501–39514 (2018). doi: [10.1109/ACCESS.2018.2855437](https://doi.org/10.1109/ACCESS.2018.2855437)
125. K. Bina et al., *Measurement* **181**, 109584 (2021). doi: [10.1016/j.measurement.2021.109584](https://doi.org/10.1016/j.measurement.2021.109584)
126. S. Jakkampudi et al., Footstep detection in urban seismic data with a convolutional neural network. *Leading Edge* **39**, 654–660 (2020). doi: [10.1190/tle390906541](https://doi.org/10.1190/tle390906541)
127. T. Xu, N. Wang, X. Xu, Seismic target recognition based on parallel recurrent neural network for unattended ground sensor systems. *IEEE Access* **7**, 137823–137834 (2019). doi: [10.1109/ACCESS.2019.2934893](https://doi.org/10.1109/ACCESS.2019.2934893)
128. K. Bin, J. Lin, X. Tong, *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5 (2021). doi: [10.1109/LGRS.2021.3055795](https://doi.org/10.1109/LGRS.2021.3055795)
129. Y. Wang, X. Cheng, X. Li, B. Li, X. Yuan, *IEEE Sens. J.* **10.1109/JSEN.2021.3067648** (2021).
130. G. Jin, B. Ye, Y. Wu, F. Qu, Vehicle classification based on seismic signatures using convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **16**, 628–632 (2019). doi: [10.1109/LGRS.2018.2879687](https://doi.org/10.1109/LGRS.2018.2879687)
131. A. Szenicer et al., Seismic savanna: Machine learning for classifying wildlife and behaviours using ground-based vibration field recordings. *Remote Sens. Ecol. Conserv.* **8**, 236–250 (2021). doi: [10.1002/rse2.242](https://doi.org/10.1002/rse2.242)
132. X. R. Li, N. Mitsakos, P. Lu, Y. Xiao, X. Zhao, Seismic compressive sensing by generative inpainting network: Toward an optimized acquisition survey. *Leading Edge (Tulsa Okla.)* **38**, 923–933 (2019). doi: [10.1190/tle38120923.1](https://doi.org/10.1190/tle38120923.1)
133. R. Tehseen, M. S. Farooq, A. Abid, A framework for the prediction of earthquake using federated learning. *PeerJ Comput. Sci.* **7**, e540 (2021). doi: [10.7717/peerj-cs.540](https://doi.org/10.7717/peerj-cs.540); PMID: 34141879
134. S. Biscans et al., Control strategy to limit duty cycle impact of earthquakes on the LIGO gravitational-wave detectors. *Class. Quantum Gravity* **35**, 055004 (2018). doi: [10.1088/1361-6382/aaa4aa](https://doi.org/10.1088/1361-6382/aaa4aa)
135. N. Mukund et al., Ground motion prediction at gravitational wave observatories using archival seismic data. *Class. Quantum Gravity* **36**, 085005 (2019). doi: [10.1088/1361-6382/ab0d2c](https://doi.org/10.1088/1361-6382/ab0d2c)
136. D. George, E. A. Huerta, Deep learning for real-time gravitational wave detection and parameter estimation: Results with advanced LIGO data. *Phys. Lett. B* **778**, 64–70 (2018). doi: [10.1016/j.physletb.2017.12.053](https://doi.org/10.1016/j.physletb.2017.12.053)
137. R. Ormiston, T. Nguyen, M. Coughlin, R. X. Adhikari, E. Katsavounidis, Noise reduction in gravitational-wave data via deep learning. *Phys. Rev. Res.* **2**, 033066 (2020). doi: [10.1103/PhysRevResearch.2.033066](https://doi.org/10.1103/PhysRevResearch.2.033066)
138. G. Marcus, *arXiv:1801.00631* [cs.AI] (2018).
139. E. A. Holm, In defense of the black box. *Science* **364**, 26–27 (2019). doi: [10.1126/science.aax0162](https://doi.org/10.1126/science.aax0162); PMID: 30948538
140. R. Roscher, B. Bohn, M. F. Duarte, J. Garcke, Explainable machine learning for scientific insights and discoveries. *IEEE Access* **8**, 42200–42216 (2020). doi: [10.1109/ACCESS.2020.2976199](https://doi.org/10.1109/ACCESS.2020.2976199)
141. A. Cunha, A. Pochet, H. Lopes, M. Gattass, Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data. *Comput. Geosci.* **135**, 104344 (2020). doi: [10.1016/j.cageo.2019.104344](https://doi.org/10.1016/j.cageo.2019.104344)
142. J. Zheng, S. Shen, T. Jiang, W. Zhu, Deep neural networks design and analysis for automatic phase pickers from three-component microseismic recordings. *Geophys. J. Int.* **220**, 323–334 (2020). doi: [10.1093/gji/ggz441](https://doi.org/10.1093/gji/ggz441)
143. F. M. Hohman, M. Kahng, R. Pienta, D. H. Chau, Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE Trans. Vis. Comput. Graph.* **25**, 2674–2693 (2018). doi: [10.1109/TVCG.2018.2843369](https://doi.org/10.1109/TVCG.2018.2843369); PMID: 29933551
144. G. Montavon, W. Samek, K.-R. Müller, Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **73**, 1–15 (2018). doi: [10.1016/j.dsp.2017.10.011](https://doi.org/10.1016/j.dsp.2017.10.011)
145. Y. Ren, X. Xu, S. Yang, L. Nie, Y. Chen, A physics-based neural-network way to perform seismic full waveform inversion. *IEEE Access* **8**, 112266–112277 (2020). doi: [10.1109/ACCESS.2020.2997921](https://doi.org/10.1109/ACCESS.2020.2997921)
146. B. Lusch, J. N. Kutz, S. L. Brunton, Deep learning for universal linear embeddings of nonlinear dynamics. *Nat. Commun.* **9**, 4950 (2018). doi: [10.1038/s41467-018-07210-0](https://doi.org/10.1038/s41467-018-07210-0); PMID: 30470743
147. J. Ma et al., Using deep learning to model the hierarchical structure and function of a cell. *Nat. Methods* **15**, 290–298 (2018). doi: [10.1038/nmeth.4627](https://doi.org/10.1038/nmeth.4627); PMID: 29505029
148. B. Scholkopf, *arXiv:1911.10500* [cs.LG] (2019).
149. G. E. Karniadakis et al., Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021). doi: [10.1038/s42254-021-00314-5](https://doi.org/10.1038/s42254-021-00314-5)
150. Y. Wang, B. Wang, N. Tu, J. Geng, Seismic trace interpolation for irregularly spatially sampled data using convolutional autoencoder. *Geophysics* **85**, V119–V130 (2020). doi: [10.1190/geo2018-0699.1](https://doi.org/10.1190/geo2018-0699.1)
151. J.-L. Wu et al., Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems. *J. Comput. Phys.* **406**, 109209 (2020). doi: [10.1016/j.jcp.2019.109209](https://doi.org/10.1016/j.jcp.2019.109209)
152. L. Yang, X. Meng, G. E. Karniadakis, B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *J. Comput. Phys.* **425**, 109913 (2021). doi: [10.1016/j.jcp.2020.109913](https://doi.org/10.1016/j.jcp.2020.109913)
153. Z. Li et al., *arXiv:2010.08895* [cs.LG] (2020).
154. Z. Ghahramani, Probabilistic machine learning and artificial intelligence. *Nature* **521**, 452–459 (2015). doi: [10.1038/nature14541](https://doi.org/10.1038/nature14541); PMID: 26017444
155. C. Blundell, J. Cornebise, K. Kavukcuoglu, D. Wierstra, in *International Conference on Machine Learning* (PMLR, 2015), pp. 1613–1622.
156. W. J. Maddox, P. Izmailov, T. Garipov, D. P. Vetrov, A. G. Wilson, *Adv. Neural Inf. Process. Syst.* **32**, 13153 (2019).
157. M. Khan et al., in *International Conference on Machine Learning* (PMLR, 2018), pp. 2611–2620.
158. Y. Gal, Z. Ghahramani, in *International Conference on Machine Learning* (PMLR, 2016), pp. 1050–1059.
159. M. van den Ende, J.-P. Ampuero, Automated seismic source characterisation using deep graph neural networks. *Neural Netw.* **10.31223/OSF.IO/NBMTZ** (2020). doi: [10.31223/OSF.IO/NBMTZ](https://doi.org/10.31223/OSF.IO/NBMTZ)
160. O. Duerr, B. Sick, E. Murina, *Probabilistic Deep Learning: With Python, Keras and TensorFlow Probability* (Manning Publications, 2020).
161. J. V. Dillon et al., *arXiv:1711.10604* (2017).
162. B. Liu et al., Deep-learning seismic full-waveform inversion for realistic structural models. *Geophysics* **86**, R31–R44 (2021). doi: [10.1190/geo2019-0435.1](https://doi.org/10.1190/geo2019-0435.1)
163. W. Wang, J. Ma, Velocity model building in a crosswell acquisition geometry with image-trained artificial neural networks. *Geophysics* **85**, U31–U46 (2020). doi: [10.1190/geo2018-0591.1](https://doi.org/10.1190/geo2018-0591.1)
164. S. Feng, Y. Lin, B. Wohlberg, *arXiv:2103.04007* [physics.geo-ph] (2021).
165. M. Araya-Polo, S. Farris, M. Florez, Deep learning-driven velocity model building workflow. *Leading Edge* **38**, 872a1 (2019). doi: [10.1190/tle38110872a1](https://doi.org/10.1190/tle38110872a1)
166. F. Lara, R. Lara-Cueva, J. C. Larco, E. V. Carrera, R. Leon, A deep learning approach for automatic recognition of seismo-volcanic events at the Cotopaxi volcano. *J. Volcanol. Geotherm. Res.* **409**, 107142 (2021). doi: [10.1016/j.jvolgeores.2020.107142](https://doi.org/10.1016/j.jvolgeores.2020.107142)
167. Y. Wang, Q. Ge, W. Lu, X. Yan, Well-Logging Constrained Seismic Inversion Based on Closed-Loop Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* **58**, 5564–5574 (2020). doi: [10.1109/TGRS.2020.2967344](https://doi.org/10.1109/TGRS.2020.2967344)
168. J. Zhang et al., Robust deep learning seismic inversion with a priori initial model constraint. *Geophys. J. Int.* **225**, 2001–2019 (2021). doi: [10.1093/gji/ggab074](https://doi.org/10.1093/gji/ggab074)
169. J. Münchmeyer et al., *J. Geophys. Res. Solid Earth* **127**, e2021JB023499 (2022).
170. W. Zhu, S. M. Mousavi, G. C. Beroza, Seismic signal augmentation to improve generalization of deep neural networks. *Adv. Geophys.* **61**, 151–177 (2020). doi: [10.1016/bs.agph.2020.07.003](https://doi.org/10.1016/bs.agph.2020.07.003)
171. M. F. Dacrema, P. Cremonesi, D. Jannach, in *Proceedings of the 13th ACM Conference on Recommender Systems* (2019), pp. 101–109.
172. G. S. Martin, R. Wiley, K. J. Marfurt, Marmousi2: An elastic upgrade for Marmousi. *Leading Edge* **25**, 156–166 (2006). doi: [10.1190/1.2172306](https://doi.org/10.1190/1.2172306)
173. F. Billette, S. Brandsberg-Dahl, in *67th EAGE Conference and Exhibition* (European Association of Geoscientists and Engineers, 2005), pp. cp-1.
174. S. M. Mousavi, Y. Sheng, W. Zhu, G. C. Beroza, STanford Earthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access* **7**, 179464–179476 (2019). doi: [10.1109/ACCESS.2019.2947848](https://doi.org/10.1109/ACCESS.2019.2947848)
175. Y. Alaudah, P. Michałowicz, M. Alfarraj, G. AlRegib, A machine-learning benchmark for facies classification. *Interpretation* **7**, SE175–SE187 (2019). doi: [10.1190/INT-2018-0249.1](https://doi.org/10.1190/INT-2018-0249.1)
176. A. Michellini et al., INSTANCE – the Italian seismic dataset for machine learning. *Earth Syst. Sci. Data* **13**, 5509–5544 (2021). doi: [10.5194/essd-13-5509-2021](https://doi.org/10.5194/essd-13-5509-2021)
177. J. Woolam et al., *arXiv:2111.00786* [physics.geo-ph] (2021).

ACKNOWLEDGMENTS

We thank the reviewers for insightful comments and suggestions. We also thank J. Münchmeyer for providing a figure presenting an example of TEAM model performance for EEW. **Funding:** G.C.B. was supported by the Air Force Research Laboratory under contract FA9453-19-C-0073. **Authors contributions:** S.M.M. compiled the dataset, designed the structure of the study, conducted the analyses, and wrote the initial draft. G.C.B. led the project. S.M.M. and G.C.B. wrote the final manuscript. Both authors discussed extensively the results. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Our database, supporting materials, and a link to an updating glossary of seismological tasks and machine learning terms are provided here: https://smousavi05.github.io/dl_seismology. This GitHub page hosts our full database, which can be used, updated, or edited by the researchers in the community. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/content/page/science-licenses-journal-article-reuse> <https://www.science.org/about/science-licenses-journal-article-reuse>

Submitted 17 September 2021; accepted 1 July 2022
10.1126/science.abm4470

Science
JOURNALS AAAS



Publish your research in the *Science* family of journals

The *Science* family of journals (*Science*, *Science Advances*, *Science Immunology*, *Science Robotics*, *Science Signaling*, and *Science Translational Medicine*) are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **[Science.org/journals](https://www.science.org/journals)**

RESEARCH ARTICLE SUMMARY

MICROBIOLOGY

Prokaryotic innate immunity through pattern recognition of conserved viral proteins

Linyi Alex Gao*†, Max E. Wilkinson†, Jonathan Strecker†, Kira S. Makarova, Rhiannon K. Macrae, Eugene V. Koonin, Feng Zhang*

INTRODUCTION: Many organisms have evolved specialized immune pattern-recognition receptors, including nucleotide-binding oligomerization domain-like receptors (NLRs) of the STAND superfamily that are ubiquitous in plants, animals, and fungi. NLRs oligomerize upon recognition of pathogen-associated molecular patterns, leading to the activation of an effector domain that mediates an inflammatory or cell death response. Although the roles of NLRs in eukaryotic immunity are well established, it is unknown whether prokaryotes use similar defense mechanisms.

RATIONALE: We previously identified a set of bacterial and archaeal STAND nucleoside triphosphatases (NTPases), dubbed Avs (antiviral STAND), that protect bacteria from tailed phages through an unknown mechanism. Like eukaryotic NLRs, Avs proteins have a characteristic tripartite domain architecture consisting of a central NTPase, an extended C-terminal sensor, and an N-terminal effector. Here, we investigate the defense mechanism of these Avs proteins.

RESULTS: Using genetic screens in *Escherichia coli*, we characterized four Avs families (Avs1 to

Avs4) and found that they detect hallmark viral proteins that are expressed during infection. In particular, Avs1 to Avs3 recognize the large terminase subunit, and Avs4 recognizes the portal. These two proteins together make up the conserved DNA packaging machinery of tailed phages. Coexpression of an Avs protein with its cognate target in *E. coli* resulted in cell death.

We assessed the specificity of Avs target recognition with a panel of terminases and portals from 24 tailed phages, spanning nine major families. Notably, a single Avs protein was capable of recognizing a large variety of targets (terminase or portal), with less than 5% sequence identity in some cases.

We next reconstituted Avs activity in vitro, focusing on representatives from *Salmonella enterica* (SeAvs3) and *E. coli* (EcAvs4), both of which contain N-terminal PD-DExK nuclease effectors. In the presence of their cognate target, SeAvs3 and EcAvs4 mediated degradation of double-stranded DNA. Nuclease activity required the presence of Mg²⁺ and adenosine triphosphate (ATP); however, the hydrolysis of ATP was not strictly required. Single-stranded DNA and RNA substrates were not cleaved.

We determined the cryo-electron microscopy structures of the SeAvs3-terminase and EcAvs4-

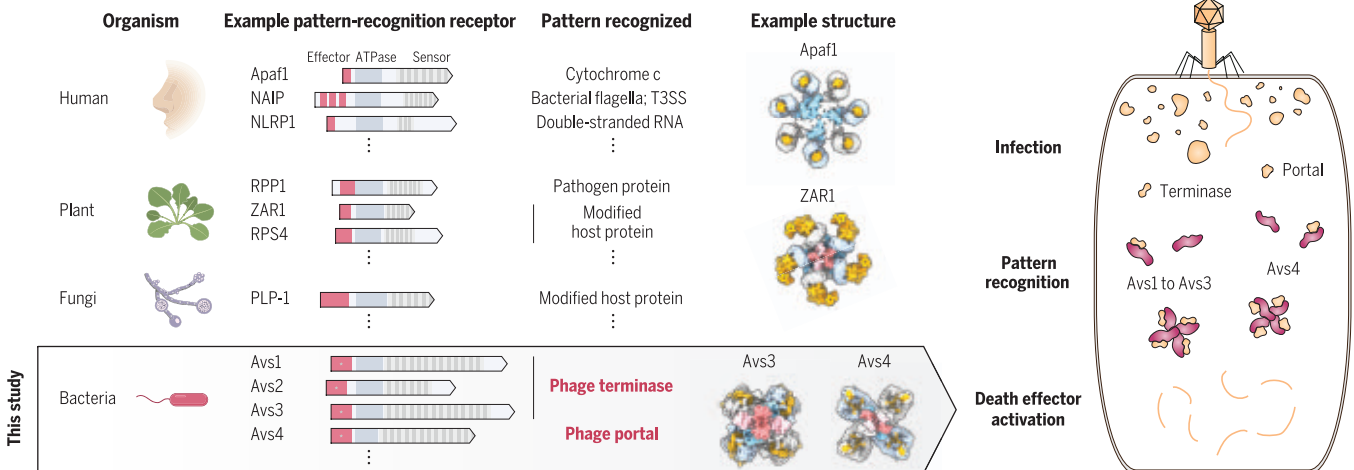
portal complexes, revealing that both form tetramers in which the C-terminal sensor domain of each Avs subunit binds to a single target protein. Binding is mediated by shape complementarity across an extended interface, consistent with fold recognition. Additionally, SeAvs3 directly recognizes terminase active-site residues and its ATP ligand. Tetramerization of both SeAvs3 and EcAvs4 is mediated by their STAND ATPase domains and allows the N-terminal nucleases to adopt active dimeric configurations.

Bioinformatic analysis of Avs proteins across prokaryotic lineages revealed at least 18 distinct types of N-terminal effectors that are modularly swapped between Avs homologs, as well as widespread distribution of *avs* genes across phyla with extensive horizontal gene transfer. Finally, we also discovered phage-encoded Avs inhibitors, highlighting an extensive arms race between prokaryotes and viruses.

CONCLUSION: NLR-like defense proteins in bacteria and archaea recognize the conserved folds of hallmark viral proteins and assemble into tetramers that activate diverse N-terminal effectors. The mechanism of these proteins highlights the similarity between the defense strategies of prokaryotes and eukaryotes and extends the paradigm of pattern recognition of pathogen-specific proteins across all three domains of life. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: zhang@broadinstitute.org (F.Z.); algao@stanford.edu (L.A.G.)
†These authors contributed equally to this work.
Cite this article as L. A. Gao et al., *Science* 377, eabm4096 (2022). DOI: 10.1126/science.abm4096

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abm4096>



Mechanism and structures of NLR-like defense proteins in prokaryotes. (Left) Comparison of the domain architectures of 11 representative NLR-like pattern-recognition receptors across four kingdoms of life. Selected structures of activated complexes are shown as examples. T3SS, type 3 secretion system. (Right) Defense mechanism of Avs proteins in bacteria and archaea (this study). Target binding triggers the formation of Avs tetramers, which activates an N-terminal effector that disrupts the viral life cycle.

RESEARCH ARTICLE

MICROBIOLOGY

Prokaryotic innate immunity through pattern recognition of conserved viral proteins

Linyi Alex Gao^{1,2,3,4,5,6*}†, Max E. Wilkinson^{1,2,3,4,5}†, Jonathan Strecker^{1,2,3,4,5}†, Kira S. Makarova⁷,
Rhiannon K. Macrae^{1,2,3,4,5}, Eugene V. Koonin⁷, Feng Zhang^{1,2,3,4,5*}

Many organisms have evolved specialized immune pattern-recognition receptors, including nucleotide-binding oligomerization domain-like receptors (NLRs) of the STAND superfamily that are ubiquitous in plants, animals, and fungi. Although the roles of NLRs in eukaryotic immunity are well established, it is unknown whether prokaryotes use similar defense mechanisms. Here, we show that antiviral STAND (Avs) homologs in bacteria and archaea detect hallmark viral proteins, triggering Avs tetramerization and the activation of diverse N-terminal effector domains, including DNA endonucleases, to abrogate infection. Cryo-electron microscopy reveals that Avs sensor domains recognize conserved folds, active-site residues, and enzyme ligands, allowing a single Avs receptor to detect a wide variety of viruses. These findings extend the paradigm of pattern recognition of pathogen-specific proteins across all three domains of life.

Bacteria and archaea have evolved numerous defense mechanisms against viral infections that involve a wide range of strategies and enzymatic activities (1–5). Defense systems are activated by viral nucleic acids, in the case of restriction-modification and CRISPR-Cas systems, or by different types of infection-induced cellular stress, including DNA double-strand breaks (6), inhibition of host transcription (7), cytosolic nucleotide depletion (8), and the disruption of translation elongation factor EF-Tu (9) or RecBCD repair nuclease (10). Alternatively, some systems constitutively synthesize small molecules that interfere with phage replication (11, 12). However, for numerous defense systems, the mechanisms of activation remain uncharacterized, and it appears likely that distinct modes of activation exist within the diverse repertoire of recently discovered systems (1, 3, 4, 13).

STAND nucleoside triphosphatases (NTPases), which include nucleotide-binding oligomerization domain-like receptors (NLRs) such as animal inflammasomes and plant resistosomes, are among the key players in immunity, cell signaling, and particularly programmed cell

death in eukaryotes (14–21). STAND NTPases have a conserved tripartite domain architecture, consisting of a central NTPase domain, a C-terminal sensor that contains superstructure-forming repeats, and, in many cases, an N-terminal effector that mediates inflammation or cell death. In animal and plant innate immunity, STAND proteins function by recognizing diverse pathogen-associated molecular patterns (PAMPs), including peptidoglycan fragments from the bacterial cell wall (18), double-stranded RNA (21), bacterial flagellin and type 3 secretion systems (16, 17), and endogenous host proteins that have been modified by pathogens (19, 20). In all of these cases, recognition of the PAMP leads to oligomerization of the STAND NTPase and activation or recruitment of effector proteins.

Bacteria and archaea, especially those with complex signaling systems, also encode a diverse repertoire of STAND NTPases that are predicted to be involved in signal transduction and possibly in programmed cell death (14, 15). However, the functions of these proteins are largely unknown, with the exception of several that have been characterized as transcriptional regulators (22–24). We recently identified a group of prokaryotic STAND NTPases, dubbed Avs (antiviral STAND) (4), that are often encoded next to restriction-modification and other defense systems (fig. S1) and protect bacteria from tailed phages (fig. S2). Here, we investigate the mechanism of Avs proteins.

Avs systems are activated by two conserved phage proteins

Although the domain architectures of Avs proteins resemble those of eukaryotic NLRs (Fig. 1, A and B), it is unclear whether they

function through similar molecular mechanisms. We identified four distinct families of Avs proteins (Avs1 to Avs4) (fig. S3), each of which contains highly divergent tetratricopeptide repeat (TPR)-like sensor domains, and selected two representatives for further characterization: SeAvs3 from *Salmonella enterica* NCTC13175 and EcAvs4 from *Escherichia coli* NCTC11132, both of which provide robust protection against the T7-like coliphage PhiV-1 (fig. S2). We first asked how phage infection leads to Avs activation and whether a specific phage-encoded trigger exists for these defense systems. We cloned fragments covering the whole PhiV-1 phage genome into expression plasmids and transformed the resulting fragment library into *E. coli* containing either Avs proteins or empty-vector controls (Fig. 1C and data S1). We hypothesized that coexpression of Avs proteins with their putative triggers might lead to cell death and depletion of the respective phage genes from the pool, and we performed deep sequencing to detect enrichment or depletion of phage genes. Four phage genes were generally toxic to all cells; however, two genes were depleted only in the presence of Avs proteins, namely the large terminase subunit (gp19) when coexpressed with SeAvs3 and the portal protein (gp8) when expressed with EcAvs4 (Fig. 1D and fig. S4). By Southern blot, we observed that Avs3- and Avs4-mediated depletion of phage DNA during infection was abolished in gp19 and gp8 knockout phage strains, respectively (Fig. 1E and fig. S5), indicating that gp19 and gp8 are both necessary and sufficient for Avs activation.

To validate these findings, we transformed plasmids expressing gp19 or gp8 into *E. coli* harboring SeAvs3 or EcAvs4 and measured cell viability. Consistent with our previous results, we observed cell death after coexpression of SeAvs3 and gp19, as well as coexpression of EcAvs4 and gp8, but not with the reciprocal pairs (Fig. 1E). This toxicity depended on the predicted nuclease activity of both SeAvs3 and EcAvs4 and, importantly, was not due to any intrinsic features of the natural phage gene sequence, because recoded gene sequences also led to cell death (Fig. 1F). Furthermore, the enzymatic activity of the phage terminase, which contains adenosine triphosphatase (ATPase) and nuclease domains unrelated to those of Avs proteins, was not required for SeAvs3-mediated toxicity (Fig. 1F).

Avs proteins recognize a diverse range of terminase and portal proteins

To investigate the specificity of Avs activation, we cloned the portal and large terminase subunit genes from 24 tailed phages, spanning nine major phage families, and coexpressed these genes in *E. coli* with 15 Avs systems spanning all four Avs families (data S2 and S3).

¹Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ²Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. ³McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ⁴Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ⁵Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ⁶Society of Fellows, Harvard University, Cambridge, MA 02138, USA. ⁷National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA.
*Corresponding author. Email: zhang@broadinstitute.org (F.Z.); algao@stanford.edu (L.A.G.)
†These authors contributed equally to this work.

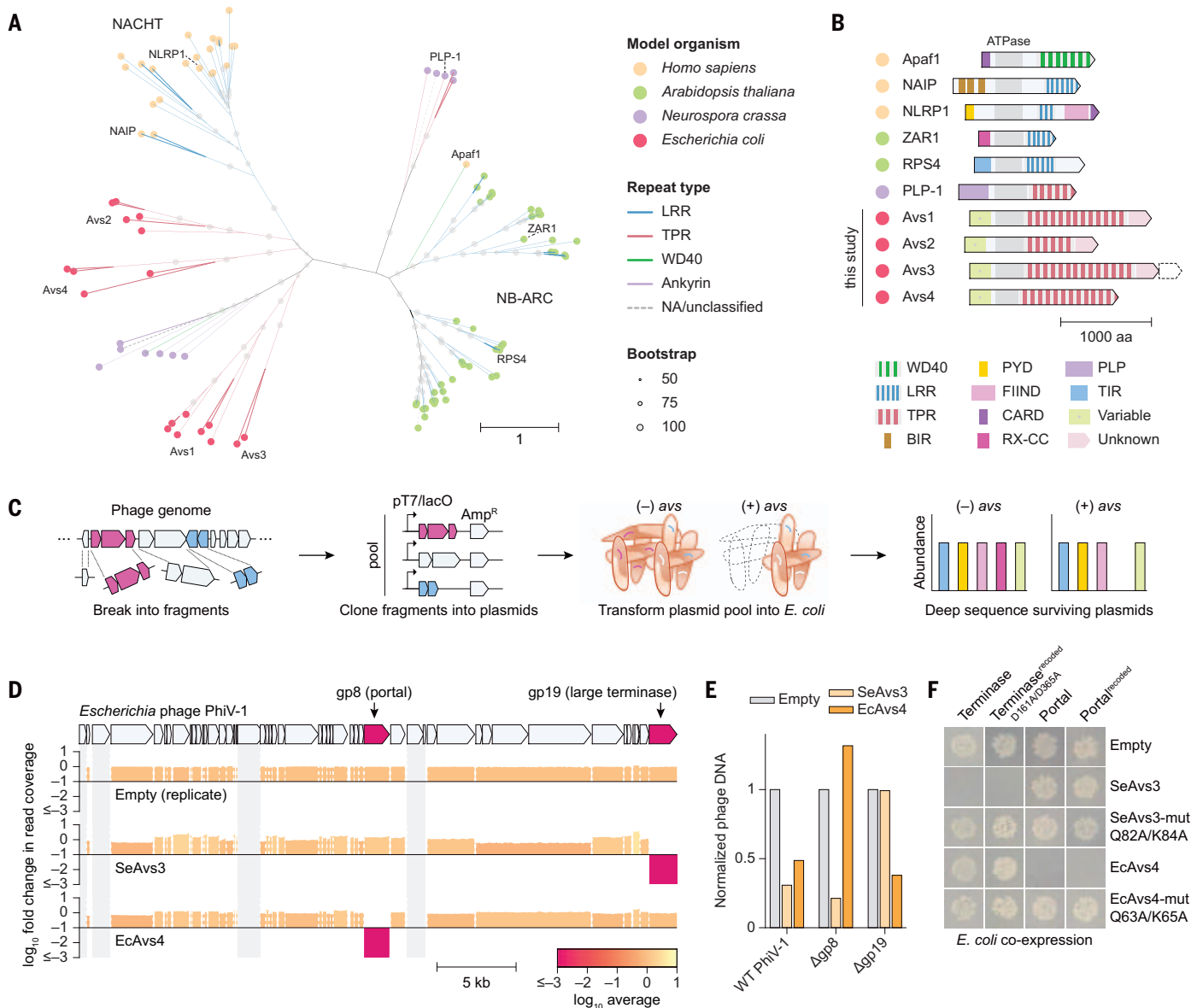


Fig. 1. Prokaryotic STAND NTPases recognize phage terminase and portal

proteins. (A) Maximum likelihood tree of the ATPase domain of selected NLR-like STAND NTPases in four model organisms across kingdoms of life. (B) Domain architectures of representative NLR-like genes in (A). Ankyrin, ankyrin repeat; BIR, baculoviral inhibitor of apoptosis repeat; CARD, caspase activation and recruitment domain; FIIND, function to find domain; LRR, leucine-rich repeat; PLP, patatin-like phospholipase; PYD, pyrin domain; RX-CC, potato virus X

resistance protein coiled-coil domain; TPR, tetratricopeptide repeat; WD40, WD40 repeat. (C) Schematic of the genetic screening approach used to identify phage-encoded activators of Avs proteins that induce cell death. Amp^R, ampicillin resistance gene. (D) Genetic screen results for phage-encoded activators. (E) Quantification of the phage DNA band intensity in a Southern blot of DNA isolated from phage-infected *E. coli*. WT, wild type. (F) Photographs of *E. coli* cotransformation assays with *avs* genes and phage activators identified in (D).

We quantified cellular toxicity and depletion of specific Avs-phage protein pairs for all 720 combinations by deep sequencing (Fig. 2A and data S2). These experiments revealed precise target specificity: Avs1 to Avs3 recognized only large terminase subunits, whereas Avs4 recognized only the portal protein (Fig. 2B and fig. S6). To assess the robustness of the assay, we repeated these experiments by varying the Avs promoter or the amount of terminase and portal induction and obtained similar results (fig. S7). Surprisingly, Avs1 and Avs2

also recognized terminases despite the lack of substantial sequence similarity among the C-terminal TPRs of Avs1 to Avs3, although we detected a structurally similar β sheet-rich domain at the end of the TPR arrays in all three proteins (fig. S8). These findings demonstrate conservation of target recognition across Avs families and suggest that the portal and large terminase subunit are key PAMPs that are recognized by prokaryotes. Moreover, Avs systems recognize PAMPs from diverse phages; for example, SeAvs3 and EcAvs2

were strongly activated by 20 of 24 and 19 of 24 tested terminases, respectively, and EcAvs4 was strongly activated by 15 of 24 tested portals (>100-fold depletion) (Fig. 1F). Because the portals and terminases from different phage families have limited sequence similarity, with less than 5% pairwise sequence identity in some cases (Fig. 2C), but share the same core fold (figs. S9 to S11), this broad range of activity implies that Avs proteins are triggered by conserved structural features rather than by specific peptide sequences. Consistent with

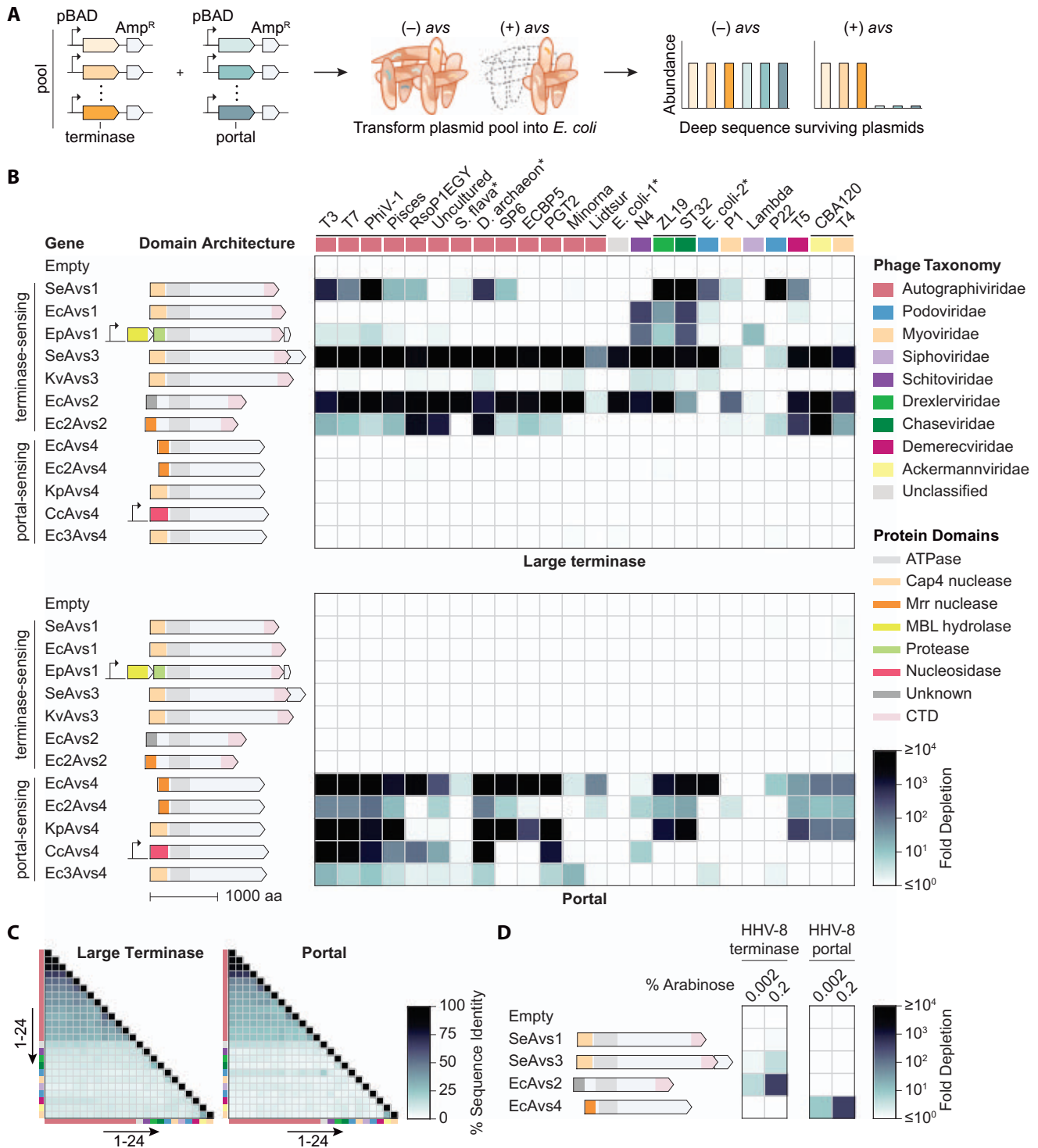


Fig. 2. Avs proteins are pattern-recognition receptors for the terminase and portal of diverse tailed phages. (A) Schematic of the plasmid depletion assay. (B) Heatmaps of plasmid depletion for the terminase and portal proteins of representative phages spanning nine major tailed phage families. The native avs promoter was retained for all homologs except for those outside of the Enterobacteriaceae family (EpAvs1 and CcAvs4). Terminases and portals were induced with 0.002% arabinose. Horizontal black lines at the top of the panel

indicate groups of terminase proteins with at least 20% pairwise sequence identity. Asterisks indicate prophages. aa, amino acids; CTD, C-terminal domain; *D. archaeon*, *Desulfurococcales archaeon ex4484_217_2*; *E. coli-1*, *E. coli* NCTC9020; *E. coli-2*, *E. coli* M885; *S. flava*, *Sphingopyxis flava* R11H. (C) Pairwise amino-acid sequence identity between the core folds of the terminases and portals in (B), excluding nonconserved regions. (D) Activity of four Avs proteins against the human herpesvirus 8 (HHV-8) terminase and portal.

this possibility, EcAvs2 and EcAvs4 displayed weak but clear recognition of the terminase and portal, respectively, of human herpesvirus 8 (Fig. 2D), which is a highly diverged evolutionary derivative of tailed phages (25) and does not infect prokaryotes.

SeAvs3 and EcAvs4 are phage-activated DNA endonucleases

SeAvs3 and EcAvs4 contain predicted N-terminal PD-DEXK-family nuclease domains (Fig. 3A), which we hypothesized degrade phage and cellular DNA upon target recognition. The nuclease domain of SeAvs3 is most similar to Cap4 effector nucleases of cyclic oligonucleotide-based defense systems (26, 27),

whereas EcAvs4 has an Mrr-like restriction endonuclease (28). Both Avs proteins contain conserved D-QxK or E-QxK catalytic motifs (Fig. 3B), and, in addition to the STAND NTPase, the SeAvs3 system contains a small open reading frame (ORF), the deletion of which reduced antiphage activity in *E. coli* (fig. S12A).

To biochemically reconstitute Avs activity in vitro, we purified recombinant SeAvs3, the protein encoded by the small ORF, EcAvs4, and the PhiV-1 terminase (gp19) and portal (gp8) proteins (fig. S13A). We incubated SeAvs3 and the small ORF product with linear double-stranded DNA (dsDNA) and observed progressive degradation of the substrate in the presence of gp19 but not gp8 (Fig. 3, C and

D). This nuclease activity was dependent on the catalytic residues of SeAvs3 but did not require the small ORF product (Fig. 3C). We further investigated the substrate specificity of Avs systems and found that the nuclease activity was specific for dsDNA, whereas single-stranded DNA and RNA were not cleaved (fig. S13B). Moreover, SeAvs3 cleaved both linear and circular dsDNA, including *E. coli* genomic DNA (fig. S13C), indicative of endonuclease activity with no specificity for phage DNA, which is consistent with an abortive infection defense mechanism.

We next investigated cofactor requirements of SeAvs3 and found that in vitro activity depends on both Mg^{2+} and adenosine triphosphate

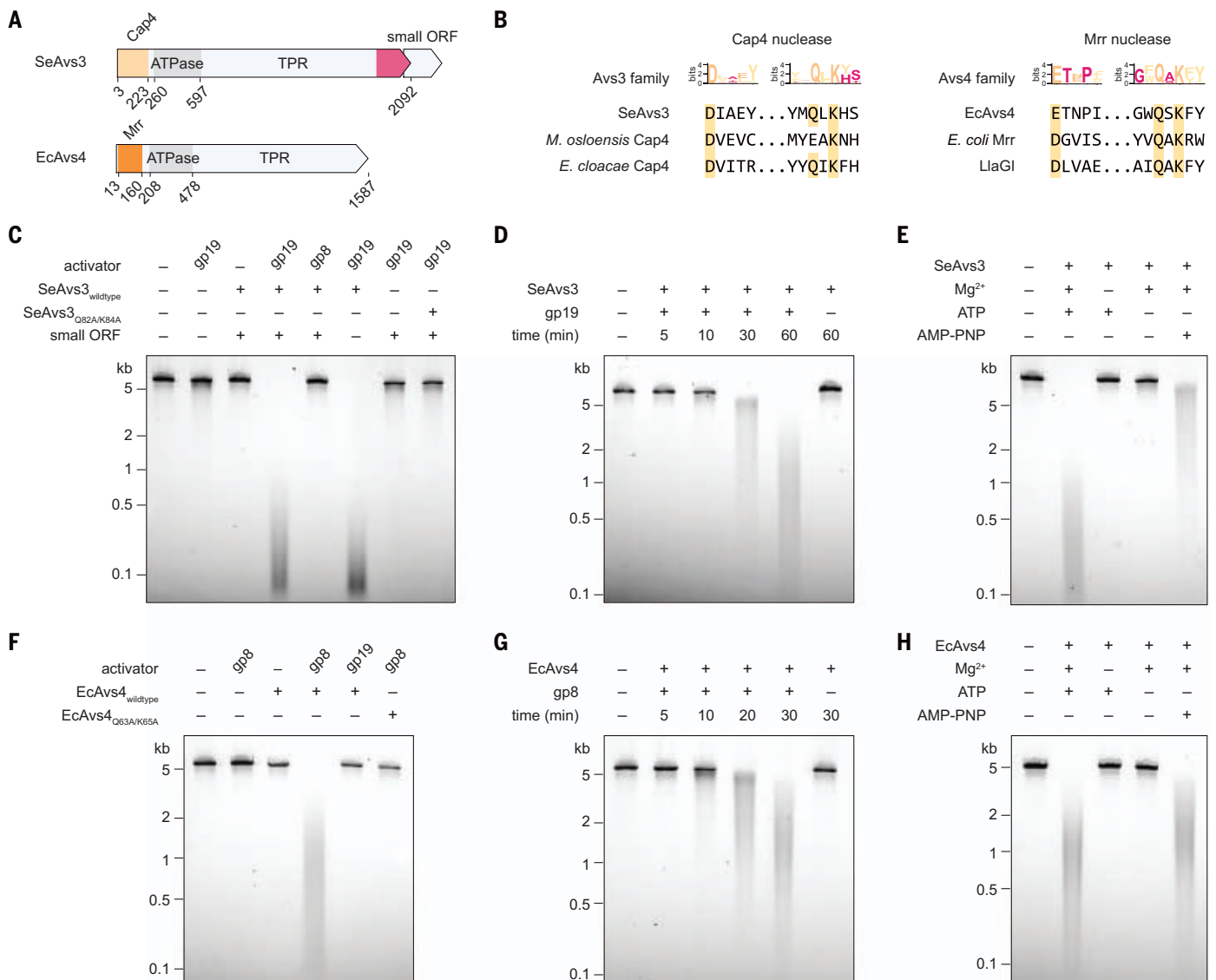


Fig. 3. SeAvs3 and EcAvs4 are phage-activated DNA endonucleases.

(A) Domain architecture of SeAvs3 and EcAvs4. (B) Alignment of Avs D-QxK nuclease motifs with characterized Cap4 and Mrr representatives. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro;

Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (C to E) Agarose gel analysis of SeAvs3 nuclease activity in vitro with a linear dsDNA substrate [(C) and (D)] and cofactor requirements (E). (F to H) Agarose gel analysis of EcAvs4 nuclease activity in vitro with a linear dsDNA substrate [(F) and (G)] and cofactor requirements (H).

(ATP); however, ATP hydrolysis is not strictly required because nuclease activity was observed at a reduced level in the presence of the nonhydrolyzable ATP analog adenylyl-imidodiphosphate (AMP-PNP) (Fig. 3E). We also found that the nuclease activity of EcAvs4 is activated by gp8, but not by gp19, and was abolished in an EcAvs4 Q63A/K65A (Gln⁶³→Ala/Lys⁶⁵→Ala) nuclease mutant (Fig. 3F). Similar to SeAvs3, nuclease activity of EcAvs4 required the presence, but not the hydrolysis, of ATP (Fig. 3H), consistent with phage plaque assays of SeAvs3 and EcAvs4 ATPase active-site mutants (fig. S12B). Together, these experiments indicate that SeAvs3 and EcAvs4 are promiscuous DNA endonucleases that are activated by distinct, highly conserved phage proteins in an ATP-dependent manner.

Structural basis of Avs binding and target recognition

To investigate how Avs systems recognize and bind their cognate phage proteins, we solved cryo-electron microscopy (cryo-EM) structures of the SeAvs3-terminase and EcAvs4-portal complexes in the presence of ATP and Mg²⁺ (figs. S14 to S17 and table S1). A reconstruction at 3.4-Å resolution revealed that SeAvs3 forms a tetramer, with each C-terminal TPR domain gripping the ATPase and nuclease domains of the gp19 terminase (Fig. 4, A and B). These TPR lobes are flexible and required symmetry expansion to improve their local resolution to 3.4 Å (Materials and methods and figs. S15 to S17). For EcAvs4 bound to the PhiV-1 gp8 portal (Fig. 4, C and D), image processing revealed equal abundances of a tetrameric complex and an octameric complex corresponding to tetramer head-to-head dimerization (fig. S14). At lower protein concentrations, however, we observed only the tetramer, indicating that it is most likely the functional complex (fig. S14). Negative-stain and cryo-EM imaging of SeAvs3 and EcAvs4 in the absence of phage proteins revealed only smaller monomeric particles (fig. S14), indicating that phage protein binding is required for the assembly of SeAvs3 and EcAvs4 into tetramers.

Tetramerization of both SeAvs3 and EcAvs4 is mediated through their STAND ATPase domains, which interact in a manner distinct from each other and from other characterized STAND ATPase oligomers like the Roq1 resistosome tetramer (29) or the Apaf1 apoptosome heptamer (30) (fig. S18). The SeAvs3 ATPase domain forms a C4-symmetric tetramer by interactions between the nucleotide-binding (NBD) and winged-helix (WHD) subdomains and the NBD subdomain of the adjacent protomer, whereas the EcAvs4 ATPase domain forms a C2-symmetric dimer of dimers with a tighter interface (1232 Å² of buried surface area, compared to 436 Å² for SeAvs3),

with adjacent WHDs and NBDs both interacting (fig. S18). The smaller interface in SeAvs3 is compensated for by additional contacts between its C-terminal TPR domains (Fig. 4, A and B). SeAvs3 and EcAvs4 both maintain ATP in their active sites with an adjacent magnesium ion coordinated by the canonical Walker A and B motifs (Fig. 4, E and F). Notably, in both cases, tetramerization of the STAND ATPase domains brings adjacent N-terminal nuclease domains close together, forming two nuclease dimers with overall C2 symmetry (Fig. 4, G and L).

SeAvs3 and EcAvs4 contain nuclease effectors of the PD-DExK superfamily. Conventional PD-DExK nucleases (e.g., restriction endonucleases) use a pair of acidic residues to coordinate at least one metal ion and a conserved lysine residue to bind the scissile phosphate and stabilize the transition state for nucleolytic cleavage (Fig. 4J) (31, 32). In the SeAvs3 Cap4 tetramer, this arrangement of residues is found in all four protomers; however, in the two “outward-facing” protomers, an extended β strand makes a steric block for DNA binding and/or metal coordination (Fig. 4I). In the two “inward-facing” protomers, a homotypic interface between the N-terminal α helices prevents the formation of this β strand, enabling the catalytic site to adopt the active configuration (Fig. 4H). Based on the crystal structure of the Hind III restriction endonuclease (33), the inward-facing protomers can be predicted to form a cavity for DNA binding, with each protomer likely cleaving opposite strands of the DNA (Fig. 4K). The EcAvs4 Mrr tetramer shows a similar principle, whereby the two inward-facing protomers contain active sites that resemble canonical PD-DExK nucleases, but in the outward-facing protomers, Glu⁴⁹, which is part of the conserved trio of active-site residues, is displaced (Fig. 4, M and N). Glu⁴⁹ is found in the loop that spans residues 33 to 52, and interactions between this loop on an “inward” protomer and an adjacent “outward” protomer likely stabilize its position in the inward protomer. Like the SeAvs3 Cap4 tetramer, these two inward protomers form a cavity that accommodates DNA in a manner similar to Hind III (Fig. 4K).

SeAvs3 and EcAvs4 both contain extensive TPR domains for binding their cognate phage proteins, which we confirmed using a bacterial two-hybrid system and protein copurification (figs. S19 and S20). The SeAvs3 TPR domain forms a left-hand-like structure capped by a β sheet-rich C-terminal domain (Fig. 5A). This domain has two cavities in which the terminase ATPase and nuclease domains are nestled. Consistent with the ability of SeAvs3 to bind terminases with little sequence similarity (Fig. 2C), there are few specific residue-residue pair contacts between SeAvs3 and the PhiV-1 terminase. Instead, binding is determined by shape and charge complementarity

between the two proteins, burying more than 3700 Å² of solvent-accessible surface area. This complementarity is maintained across a diverse range of experimental structures and AlphaFold models of phage terminases (Fig. 5B). Additionally, SeAvs3 directly recognizes residues within the two terminase active sites. In particular, Asp¹⁷¹⁰ in SeAvs3 forms a salt bridge with the highly conserved Arg⁶¹ within the Walker A motif of the terminase ATPase (Fig. 5, C and E). An arginine in this position is found in most terminase ATPases that activate SeAvs3 but not in nonactivating terminases (Fig. 5E and fig. S21). These observations suggest that Arg⁶¹ within the Walker A motif is a determinant of recognition specificity, and indeed, mutation of the cognate arginine in the T4 terminase ATPase domain substantially reduced SeAvs3 activation by the ATPase domain (Fig. 5F). Notably, an arginine is not typically found in this position in endogenous cellular ATPases (34), suggesting a possible mechanism for avoiding off-target activation. Furthermore, Arg¹¹⁹⁶ and Lys¹¹⁹⁸ in SeAvs3 form salt bridges to the four conserved aspartates that make up the active site of the terminase nuclease (Fig. 5, D and E), and mutation of Asp³⁶⁵ in the PhiV-1 terminase nuclease notably reduced SeAvs3 activation by the nuclease domain (Fig. 5F). Thus, SeAvs3 directly reads the active-site residues of both domains of the terminase. Furthermore, the ATP ligand bound by the terminase is detected by interactions between the gamma phosphate and His¹⁷⁷⁰ and Tyr¹⁷¹⁴ of SeAvs3 (Fig. 5C). Targeting this ligand presumably helps avoid phage escape mutations, because ATP binding is required for the function of the terminase.

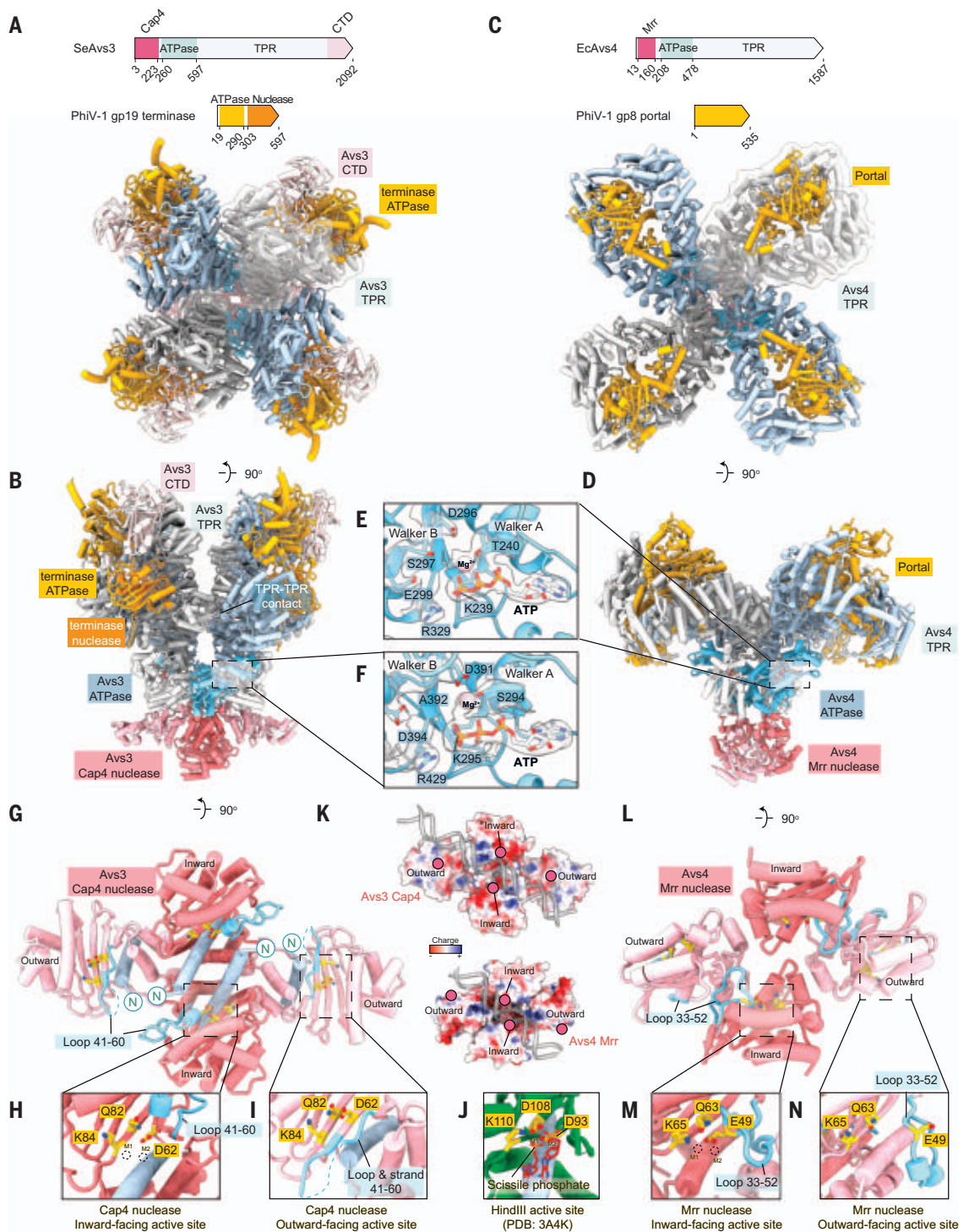
Because SeAvs3 detects both domains of the terminase, we hypothesized that there might be some functional redundancy in these interactions. Indeed, SeAvs3 was activated by the nuclease domain alone from some phages, including T7, but was also activated by the ATPase domain alone from T4 and ZL19, a T1 family phage (fig. S22). Likewise, SeAvs1 was activated by the nuclease domain from T7, but in the case of ST32, both the nuclease and ATPase domains were required. These results suggest that Avs1 and Avs3 recognize both the nuclease and ATPase domains but differ in the extent of activation by either domain, depending on the terminase. By contrast, deletion of the nuclease domain had no impact on Avs2 activity for any of the five tested terminases, suggesting that Avs2 recognizes the ATPase domain only (fig. S22). This pattern of recognition is consistent with the larger size of Avs1 and Avs3 compared with Avs2.

The TPR domain of EcAvs4 also binds the PhiV-1 portal with a large interface, burying 5800 Å² of solvent-accessible surface area, that includes notably few residue-residue contacts (Fig. 5G). The portal protein is recognized

Fig. 4. Cryo-EM structures of SeAvs3 and EcAvs4 in complex with their cognate triggers.

(A and B) Structure of the SeAvs3-terminase complex. (C and D) Structure of the EcAvs4-portal complex. (E and F) ATP molecule in the STAND ATPase active site of EcAvs4 and SeAvs3. The cryo-EM density is shown as a transparent surface. (G) SeAvs3

Cap4-like nuclease effector domain. (H and I) Active sites for the inward- and outward-facing protomers of the SeAvs3 Cap4-like nuclease. (J) Equivalent view of the active site of Hind III bound to target DNA with two divalent metal ions [Protein Data Bank (PDB) ID 3A4K]. (K) Electrostatic surface potential for the SeAvs3 Cap4-like nuclease and the EcAvs4 Mrr-like nuclease. Active sites are indicated by purple circles. Ideal B-form DNA is modeled on both surfaces based on the crystal structure of Hind III bound to its target (PDB ID 3A4K). (L) EcAvs4 Mrr-like nuclease effector domain. (M and N) Active sites for the inward- and outward-facing protomers of the EcAvs4 Mrr-like nuclease.



through its stem, clip, and part of its wing domain. In an assembled dodecameric portal complex, these regions are found toward the interior and are therefore more constrained in their fold requirements (Fig. 5I). Consistent with this observation, we performed random mutagenesis by polymerase chain reaction (PCR) to screen for portal mutations that abrogate Avs4 activation and found that all 29 identified mutations were nonconservative

and located in the core wing or stem regions (fig. S23), possibly disrupting the core portal fold. The clip domain, which contains a conserved antiparallel β sheet with an intervening α helix, is recognized by β -sheet augmentation with a hairpin of EcAvs4 (Fig. 5H), a mode of fold recognition that does not depend on the amino acid sequence of the target. Because portal proteins are not enzymes, there are no active-site residues to target as in the

SeAvs3-terminase complex. Finally, portal oligomerization is not compatible with the Avs4-bound state (Fig. 5H), suggesting that Avs4 recognizes portal monomers before they assemble into the procapsid.

Avs proteins are widespread and possess diverse, modular effector domains

To assess the diversity of *avs* genes in prokaryotes, we collected all intact homologs

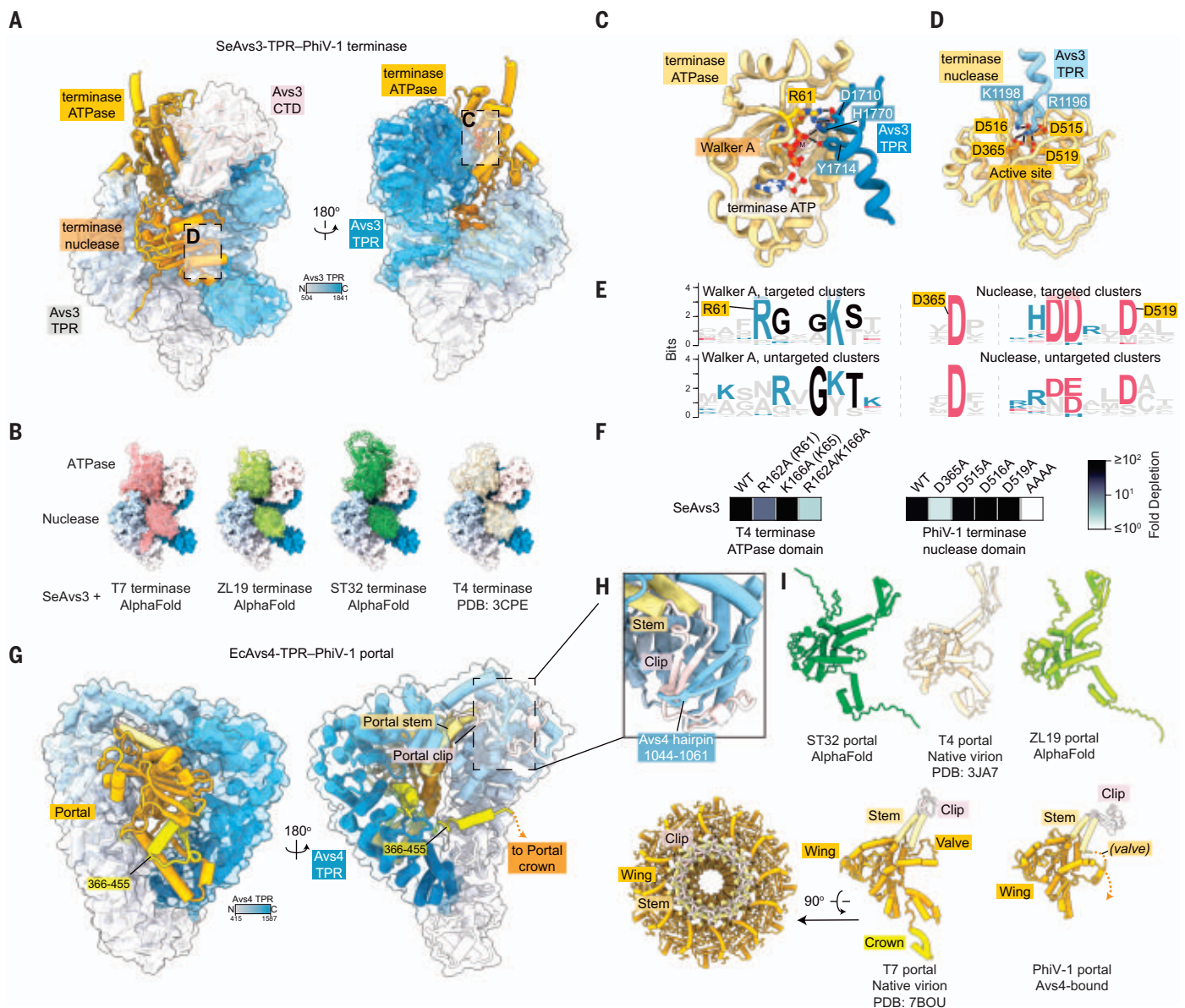


Fig. 5. Structural basis for viral-fold recognition by SeAvs3 and EcAvs4.

(A) The interface between SeAvs3 and the PhiV-1 terminase. An SeAvs3 surface view is shown in transparency. SeAvs3 is colored from the N to C terminus according to the key. (B) AlphaFold or crystal structures of different terminases modeled into SeAvs3. The ATPase and nuclease domains were individually aligned to the PhiV-1 terminase domains. (C and D) Recognition of the PhiV-1 terminase ATPase and nuclease active sites by the SeAvs3 TPR domain. (E) Sequence logos for terminase ATPase Walker A motifs and terminase nuclease active sites. A total of 11,000 terminase sequences were clustered at 30% sequence identity, and motifs were extracted from clusters containing

terminases targeted or not targeted by SeAvs3 according to Fig. 2B (see also fig. S21). (F) Plasmid depletion assay for SeAvs3 coexpressed in *E. coli* with a terminase ATPase or nuclease domain harboring active-site mutations. (G) The interface between EcAvs4 and the PhiV-1 portal. An EcAvs4 surface view is shown in transparency. EcAvs4 is colored from the N to C terminus according to the key. (H) β -sheet augmentation between EcAvs4 and the portal clip domain. (I) Comparison of the EcAvs4-bound state of the PhiV-1 portal, the cryo-EM structure of the highly homologous T7 portal in its native virion, and AlphaFold models of diverse portals. A top view of the assembled dodecamer of the T7 portal is also shown.

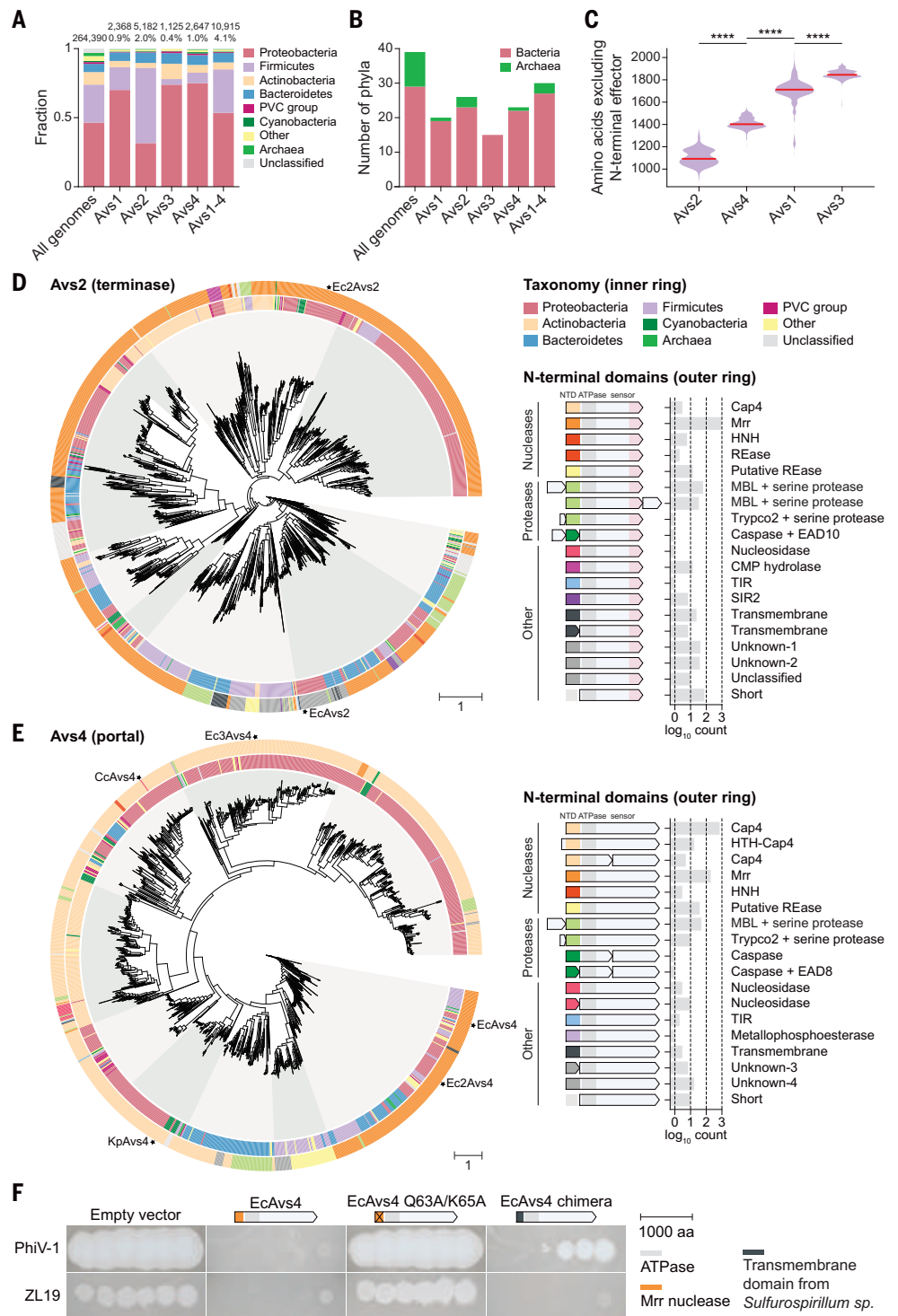
from each of the four families present in the National Center for Biotechnology Information (NCBI) nonredundant protein sequence database (data S4 to S8). The *avs* genes were identified in ~4 to 5% of sequenced prokaryotic genomes and are broadly distributed (Fig. 6A), with at least one *avs* gene detected in 27 of 29 and 3 of 10 well-represented bacterial and

archaeal phyla, respectively (Fig. 6B and fig. S24A). Each Avs family has a characteristic protein size (Fig. 6C). We next constructed phylogenetic trees of each of the four families (Fig. 6, D and E; and fig. S24, B and C) and found that these trees did not follow bacterial and archaeal phylogenies, suggesting extensive horizontal gene transfer, particularly for *avs2* and *avs4*, in

agreement with previous analyses of STAND NTPases (14, 15). Furthermore, we detected at least 18 distinct types of N-terminal effector domains present in Avs proteins, including non-nuclease domains such as proteases, nucleosidases, sirtuins (SIR2), Toll/interleukin-1 receptor homology (TIR) domains, cytidine monophosphate (CMP) hydrolases, transmembrane

Fig. 6. Taxonomic distribution and domain architectures of Avs families.

(A) Distribution of *avs* genes across phyla. The values above the bars indicate the number and percentage of genomes containing each gene. PVC, Planctomycetota, Verrucomicrobiota, and Chlamydiota. **(B)** Number of bacterial and archaeal phyla (minimum 100 sequenced isolates) with at least one detected instance of an *avs* gene. **(C)** Kernel density plots of the length distribution of Avs proteins, excluding the N-terminal domain. The red lines indicate medians. **** $p < 0.0001$ (Mann-Whitney). **(D and E)** Maximum likelihood tree of representatives of the ATPase + C-terminal domain of Avs2 ($n = 1255$) (D) and Avs4 ($n = 1089$) (E) clustered at 95% sequence identity. See fig. S24 for the trees for Avs1 and Avs3. Stars on the outer ring indicate homologs investigated experimentally in this study. HTH, helix-turn-helix; MBL, metallo- β -lactamase; REase, restriction endonuclease. **(F)** Phage plaque assays showing antiphage defense activity of a chimeric Avs4 with transmembrane N-terminal helices from *Sulfurospirillum* sp. replacing the Mrr-like nuclease domain of EcAvs4. The X indicates a nuclease domain mutation.



helices, and domains with unknown functions (data S4 to S8). Some less common variants are predicted to participate in intracellular signaling networks through effector-associated domains (EADs) that may recruit a caspase-like protease by protein-protein interaction (35, 36) (fig. S25), reminiscent of animal NLRs.

The apparent frequent exchange of N-terminal domains in the evolution of the Avs

families emphasizes the modular organization characteristic of STAND NTPases (15) and implies that closely related ATPase and TPR domains can activate a wide range of effector functions beyond DNA cleavage. To test this hypothesis, we chose an Avs4 homolog from *Sulfurospirillum* sp. that contains ATPase and TPR domains that are highly similar to those of EcAvs4, with 44% overall amino acid iden-

tity, but encompasses an N-terminal region with predicted transmembrane helices instead of a nuclease (fig. S26A). We generated a chimeric Avs4 protein by transplanting the transmembrane domain to EcAvs4 and found that the chimera conferred protection against T7, PhiV-1, and ZL19 (Fig. 6F and fig. S26B) while retaining the ability to recognize the portal proteins from diverse phages (fig. S26C).

Phages inhibit Avs defense through diverse antidefense proteins

Bacterial and archaeal viruses have evolved diverse mechanisms to counteract defense systems (37), including numerous anti-restriction and anti-CRISPR proteins (38, 39). We hypothesized that Avs inhibitors might exist among phage early genes, which are expressed before the portal and terminase genes during

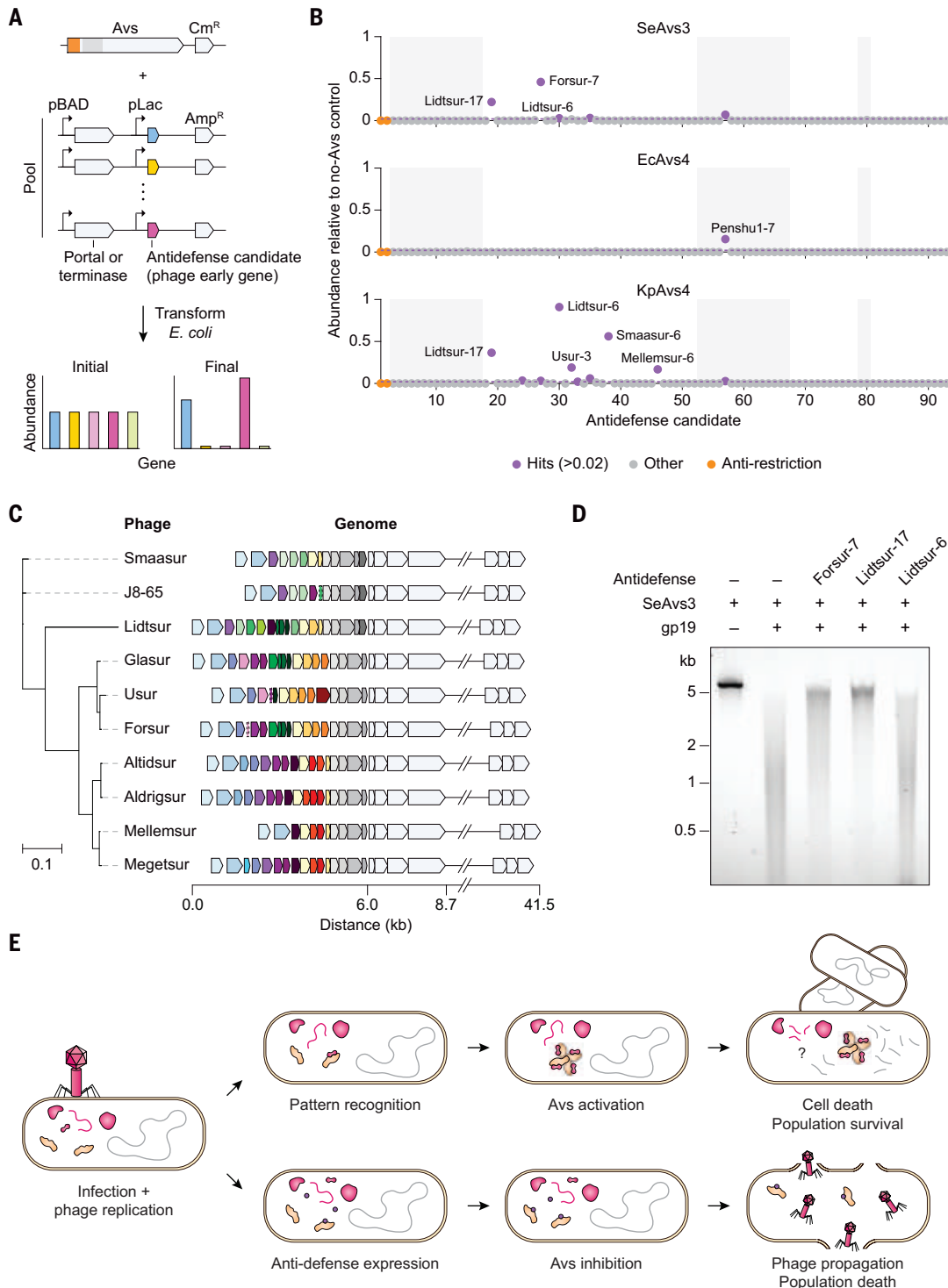
the phage life cycle. Focusing on the Autographiviridae family of T7-like coliphages, which have readily identifiable early genes, as well as portals and terminases that strongly activate Avs proteins, we identified a set of 122 representative early genes that typically encode small proteins (median length 77 amino acids), typical of antidefense genes (data S9). We performed a genetic screen for suppressors

of Avs toxicity by coexpressing these genes with SeAvs3, EcAvs4, or KpAvs4 and their cognate trigger (Fig. 7A). We identified several early genes that rescued cell growth (Fig. 7B), most of which originate from a hypervariable region within a group of closely related phages isolated from wastewater (Fig. 7C) (40, 41).

To validate these observations, we produced three of the Avs inhibitors by cell-free

Fig. 7. Phage-encoded genes inhibit Avs activity.

(A) Schematic of a pooled screen in *E. coli* for phage early genes that rescue Avs-mediated toxicity. Cm^R, chloramphenicol resistance gene. (B) Deep-sequencing readout of antidefense candidate genes coexpressed with SeAvs3, EcAvs4, or KpAvs4. (C) A hypervariable early-gene locus within a closely related set of Autographiviridae phages contains abundant antidefense genes. The tree was constructed from a concatenated alignment of conserved proteins present in all 10 phages. Colors represent groups of proteins clustered at 40% sequence identity and 70% coverage. (D) Agarose gel analysis showing in vitro reconstitution of anti-SeAvs3 activity by three antidefense candidates. (E) Schematic of the mechanism of Avs proteins as antiphage pattern-recognition receptors.



translation and observed inhibition of SeAvs3 nuclease activity *in vitro* by Lidtsur-17 and, to a lesser degree, by Forsur-7 (Fig. 7D). Lidtsur-17, Forsur-7, and Lidtsur-6 were also active in phage plaque assays and restored phage propagation on Avs-containing *E. coli* (fig. S27). Surprisingly, these inhibitors were active against different Avs families, including the chimeric Avs4, where the effector nuclease of EcAvs4 was replaced with a transmembrane domain. Furthermore, the lack of detectable sequence similarity between these inhibitors suggests distinct modes of action, which resembles the case of the highly diverse anti-CRISPRs (38, 39). Further studies will be required to elucidate how these phage proteins block Avs activity.

Discussion

Here, we characterize four families of prokaryotic STAND NTPases and demonstrate that they are pattern-recognition receptors against two phage proteins: the large terminase subunit and the portal. These proteins, along with the major capsid protein, are the signature proteins of the virus realm Duplodnaviria, which unites tailed phages and tailed archaeal viruses with herpesviruses (25). Members of this realm, particularly tailed phages, are the most abundant among known viruses (42, 43). The portal protein nucleates virion assembly, occupying the distinctive pentameric vertex of capsids and providing the attachment site for the phage tail, and serves as the channel for genome entry into and exit from the capsid (44). The terminase is the motor that packages the phage genome into the capsid at high density and pressure, using the energy of ATP hydrolysis, and cleaves DNA concatemers into genome-size units (45). The universal, complex molecular functions of these proteins engender strong selective constraints and hence evolutionary conservation. It is therefore logical that these particular proteins were selected as the targets for pattern recognition during the coevolution of prokaryotes with viruses. The three groups (Avs1 to Avs3) that recognize terminases do not form a clade in the phylogeny of the STAND domain (fig. S3), suggesting that defense based on terminase recognition evolved independently on multiple occasions.

The *in vitro* reconstitution and structural analysis of two Avs systems described here provide insight into their mechanism of defense, including promiscuous DNA endonuclease activity (Fig. 7E). Although many Avs proteins contain predicted nucleases, we identified diverse N-terminal effector domains throughout Avs families, indicating distinct mechanisms of defense that remain to be characterized. The demonstration that at least some of these effectors can be swapped without compromising Avs defense function highlights the modular functionality of these proteins, which appears important for the diversification of defense

mechanisms. The effectors of SeAvs3 and EcAvs4 are both activated by tetramerization, suggesting that the diverse Avs effectors are unified by the requirement for oligomerization for activity, a common mechanism for signal transduction by STAND NTPases (46). Indeed, oligomerization is also involved in the activation of Cap4 nucleases in cyclic oligonucleotide-based antiphage signaling systems (26).

Notably, Avs proteins recognize conserved structural features of their cognate targets across an extreme variety of amino acid sequences, including those originating from both tailed phages and archaeal viruses, as well as eukaryotic herpesviruses, which are only distantly related and do not infect prokaryotes. Structural analysis of Avs3 revealed that it directly detects the active-site residues and ATP ligand of the terminase, thereby targeting the moieties that are the most difficult for phages to mutate without abrogating function.

Notable similarities, but also several differences, exist between eukaryotic NLRs and prokaryotic Avs proteins. Both are intracellular receptors of the STAND superfamily that detect PAMPs through C-terminal repetitive structures. Both exhibit triggered oligomerization, but with distinct interfaces between the central ATPase domains (fig. S18). Similar to plant NLRs like RPP1 and the Roq1 resistosome, both Avs3 and Avs4 form tetramers, with the effector domains activated by forming a twofold symmetric dimer of dimers (29, 47). In the absence of their ligands, animal and plant NLRs have autoinhibited states that prevent oligomerization and effector activation (19). Whether such states exist for Avs proteins will require further investigation. In contrast to Avs proteins, NLRs often contain N-terminal adaptor domains that recruit other proteins to form macromolecular complexes known as inflammasomes or resistosomes (19). In these complexes, the effector (e.g., caspase-1) is a separate protein rather than a domain of the NLR. By contrast, Avs effectors are usually the N-terminal domain of the STAND NTPase. This simpler organization might be advantageous because counteracting phage replication requires a rapid, direct cellular response. This contrast parallels the distinction between the mechanisms of prokaryotic and eukaryotic STING proteins, whereby bacterial STING homologs directly activate TIR domain effectors rather than regulate transcription, as mammalian STINGs do (19, 48).

Bacteria and archaea encode numerous diverse STAND NTPases beyond the four families characterized in this study (15). Although some are not involved in defense, such as the transcriptional regulators MalT (22), AfsR (23), and GutR (24), several are confirmed defense genes or are predicted to have a defense function based on their enrichment in genome regions adjacent to known defense systems (4, 49). We

investigated several of these other defense-associated systems (data S2) but observed no detectable toxicity when coexpressed with any of the 48 tested terminases or portals, suggesting that they are triggered by other pathogen-related patterns that remain to be identified. Further investigation will reveal whether these triggers are proteins and whether they are phage-specific or endogenous to the host. For instance, most characterized plant and fungal NLRs sense the state of host pathways rather than pathogen-specific proteins (19, 20), and it remains a possibility that other groups of prokaryotic STAND NTPases function similarly.

Given the extensive sequence divergence among the STAND NTPases, it is unclear whether Avs proteins are direct evolutionary ancestors of eukaryotic NLRs, although this remains a possibility. Alternatively, or additionally, the characteristic tripartite domain architectures of diverse STAND NTPases could have evolved convergently, suggesting that this modular organization is a facile way to create allosterically activated enzymes that respond to PAMPs and could inspire the design of engineered molecular sensors. Overall, the results of this work advance our understanding of host-virus interactions in diverse microbes and extend the paradigm of pattern recognition of pathogen-specific proteins to all three domains of life.

Materials and methods

Phylogenetic analysis of STAND NTPases

For STAND phylogenetic analysis (fig. S3), PSI-BLAST searches (50) against the database of complete bacterial and archaeal genomes (extracted from GenBank, March 2019) were performed for three iterations using ATPase domains of seven Avs1 to Avs3 homologs (WP_126523998.1, WP_115407481.1, WP_084007836.1, WP_060615938.1, WP_139964370.1, WP_063118745.1, and WP_001017806.1) investigated experimentally. The 2000 best hits from each run were taken and combined with 949 Avs4 homologs found in the NCBI nonredundant protein sequence database (nr) in 2021. A nonredundant set of 4843 sequences was used for phylogenetic reconstruction using a hybrid UPGMA/FastTree approach as follows. At the first step, sequence clusters were obtained using MMseqs2 (51) with a sequence similarity threshold of 0.5, and the sequences within each cluster were aligned using MUSCLE (52). At the second step, cluster-to-cluster similarity scores were obtained using HHSEARCH (53) (including trivial clusters consisting of a single sequence each) and normalized by the minimum of the self-scores. Relative similarity scores (s) were converted to distances (d) defined as $d = -\ln s$, and a UPGMA (unweighted pair group method with arithmetic mean) dendrogram was constructed from the distance matrix (54). At the third step, sequence-based

trees were constructed from the cluster alignments using FastTree (55) (WAG evolutionary model, gamma-distributed site rates) and rooted by midpoint; these trees were grafted onto the tips of the profile similarity-based UPGMA dendrograms. FastTree was also used to calculate support values. Only the second step of the above procedure was applied to reconstruct the UPGMA dendrogram using multiple alignments of selected well-supported branches identified by the first procedure.

Construction of *avs* phylogenetic trees

Homologs of each of the four clades of *avs* genes were identified using PSI-BLAST searches against the NCBI nr database in June 2021 using position-specific scoring matrices for each clade derived from manually curated multiple sequence alignments (MSAs) of conserved regions. After a round of curation to remove false positives hits and partial proteins, referencing the corresponding genome assemblies to correct misannotated start codons, a list of 1584, 2342, 1018, and 1813 nonredundant full-length proteins were obtained for Avs1, Avs2, Avs3, and Avs4, respectively. To reduce sampling bias, sequences were then clustered at 95% sequence identity (minimum 80% coverage) using MMseqs2 with parameters—min-seq-id 0.95 -c 0.8—cov-mode 1. One representative from each cluster was selected for subsequent analyses, resulting in 843, 1255, 630, and 1089 sequences for Avs1, Avs2, Avs3, and Avs4, respectively.

MSAs of each Avs clade, excluding the variable N-terminal domains, were generated using MAFFT v7.450 (56) with global pairwise alignment (parameters—maxiterate 1000—globalpair). Alignments were trimmed using trimAl 1.2 with a gap threshold of 0.25 (-gt 0.25). Phylogenetic trees were built from the trimmed MSAs using IQ-TREE 1.6.12 (57) with the LG+G4 model and 2000 ultrafast bootstrap replicates (parameters—nstop 500 -bb 2000 -m LG+G4). To categorize the N-terminal domains, the N-terminal sequences were clustered using MMseqs2 with parameters—min-seq-id 0.4 -c 0.8, and a representative sequence from each cluster was analyzed using HHpred (58). Phyla classification was determined from the NCBI taxonomy database, and trees were rooted by midpoint and visualized using iTOL (59).

The phylogenetic tree comparing the ATPase domains of NLR-like genes across model organisms (Fig. 1A) was constructed in a similar manner, incorporating the set of 23 human NLRs and the best characterized NLRs from *Arabidopsis thaliana* (60) and *Neurospora crassa* (20, 61).

Taxonomic distribution of *avs* genes

To determine the taxonomic distribution of *avs* genes, genome assemblies containing one or more full-length Avs homologs were

identified through the NCBI Identical Protein Groups (IPG) database (data S8). Redundant assemblies were removed on the basis of their nine-digit accession numbers. To determine the percentage of genomes containing *avs* genes, the list of all available prokaryotic assemblies was downloaded from https://ftp.ncbi.nlm.nih.gov/genomes/GENOME_REPORTS/prokaryotes.txt.

Construction of terminase and portal alignments

Structures of all tested terminase and portal proteins were predicted using AlphaFold2 (62), and structures were aligned and visualized using PyMOL 2.3.4. Representatives of the predicted structures were used as input for MSA construction using PROMALS3D (63). Before computing pairwise sequence identities, MSAs were trimmed to retain only the regions corresponding to the core terminase or portal fold.

Cloning

Genes were chemically synthesized or amplified with Q5 [New England Biolabs (NEB)] or Phusion Flash (Thermo Scientific) polymerase. Plasmids were assembled using Gibson assembly. Plasmid sequences were verified by Tn5 tagmentation and high-throughput sequencing, as previously described (4, 64).

Bacterial strains

E. coli NovaBlue and NovaBlue(DE3) were obtained from Millipore Sigma. *E. coli* K-12 (ATCC 25404) and strain C (ATCC 13706) were obtained from the American Type Culture Collection. All genetic assays were performed with *E. coli* NovaBlue(DE3) unless indicated otherwise.

Competent cell production

E. coli strains were cultured in ZymoBroth with 25 µg/ml chloramphenicol and made competent using Mix & Go buffers (Zymo) according to the manufacturer's recommended protocol.

PhiV-1 fragment screen

DNA fragments consisting of intact ORFs were amplified from phage PhiV-1 and cloned into expression plasmids after a LacI-repressed T7 promoter. Plasmids were pooled with an mNeonGreen-expressing control plasmid and transformed into *E. coli* NovaBlue(DE3) containing either SeAvs3, EcAvs4, or a pACYC184 empty vector. An additional sample consisting of the plasmid pool transformed into empty vector—containing *E. coli* NovaBlue, which does not express from T7 promoters, was also included to assess the basal toxicity of the phage genes.

After 1 hour of outgrowth in SOC (super optimal broth with catabolite repression) medium at 37°C, cells were plated on LB agar

plates containing 25 µg/ml chloramphenicol and 100 µg/ml ampicillin in the absence of isopropyl-β-D-thiogalactopyranoside (IPTG). Plates were incubated for an additional 12 hours at 37°C, after which surviving plasmids were isolated by miniprep (Qiagen). A total of 200 ng of plasmid for each condition was tagged with Tn5 to yield an average fragment size of ~500 base pairs. After addition of 0.5 volumes of 0.1% SDS and column purification, tagged fragments were amplified over eight cycles by Q5 DNA polymerase (NEB) with sample-specific i5 and i7 index primers. Amplicons were gel extracted and sequenced on a NextSeq (Illumina) using 150 cycles for the forward read. Reads were mapped to reference sequences using Geneious Prime. The read coverage of each sample was then normalized to the read coverage of the mNeonGreen control within the same sample. Finally, for each sample, the read coverage per base was divided by the corresponding read coverage per base for the empty vector NovaBlue(DE3) control (Fig. 1D) or by that of the empty-vector NovaBlue control (fig. S4).

Terminase and portal depletion screens

Terminase and portal genes were amplified directly from phage samples or chemically synthesized (Twist Bioscience) with codon optimization for *E. coli*. Genes were expressed under the control of a pBAD promoter. Plasmids were pooled with an mCherry-expressing control plasmid and transformed into *E. coli* NovaBlue(DE3) containing an Avs system or a pACYC184 empty vector. After 1 hour of outgrowth in SOC at 37°C, cells were plated on LB agar plates containing 25 µg/ml chloramphenicol and 100 µg/ml ampicillin with 0.002% arabinose or, in some cases, with 0.2% arabinose, as detailed in the figures. After an additional 12 hours at 37°C, plasmids were isolated and sequenced, and depletion values (Fig. 2, B and D; and figs. S7 and S26C) were computed as described for the PhiV-1 depletion screen.

Portal and terminase mutant screens

Two synonymous versions of the T7 portal gene were randomly mutagenized by PCR using KAPA HiFi HotStart ReadyMix DNA polymerase (Roche) and cloned by Gibson assembly into a plasmid backbone containing a LacI-repressed T7 promoter. Plasmids were column purified and transformed into *E. coli* NovaBlue(DE3) containing EcAvs4, KpAvs4, or CcAvs4. Cells were plated on LB agar plates containing 25 µg/ml chloramphenicol and 100 µg/ml ampicillin in the absence of IPTG. After overnight growth, surviving colonies were sampled at random, cultured, and sequenced, and those containing single amino acid substitutions in the portal were retained for subsequent analysis. To reduce the number of stop codons and frameshift mutants sampled, a fluorescent protein (mNeonGreen) was included in

the plasmid backbone immediately after the portal ORF with a single-nucleotide overlap, such that mNeonGreen was translated only if the portal ORF remained intact (65) (fig. S23, A and B). Both portal and mNeonGreen were translated as separate polypeptides.

To quantitatively assess the effect of each mutant on Avs4-mediated toxicity (fig. S23C), mutant plasmids were pooled and retransformed into Avs4-containing *E. coli* as described for the PhiV-1 depletion screen. Fold depletion was also quantified as described, with the exception that only reads containing 20-mer sequences specific to one mutant (i.e., mapping to the mutation site) were counted in the analysis. A similar procedure was followed to quantify the effect of truncation of the terminase or portal (figs. S22 and S23E), as well as terminase-domain mutations (fig. S5F).

Antidefense screen

Putative early genes from Autographiviridae coliphages were tabulated and clustered at 50% sequence identity and 50% coverage using MMseqs2 (-min-seq-id 0.5 -c 0.5), resulting in 120 clusters. One representative was selected from each cluster, along with two additional sequences, for a total of 122 initial candidates (data S9). Genes were synthesized by Twist Bioscience and cloned by Gibson assembly into expression vectors containing either the portal or terminase from phage PhiV-1 driven by a pBAD promoter. Antidefense candidates were expressed under the control of a *lac* promoter. Plasmids were pooled and transformed into *E. coli* containing SeAvs3, EcAvs4, KpAvs4, or an empty vector. Cells were grown at 37°C for 16 hours on LB agar plates containing 25 µg/ml chloramphenicol and 100 µg/ml ampicillin with no added arabinose. After plasmid isolation, antidefense candidates were amplified over two rounds of PCR to attach eight-nucleotide i7 and i5 index barcodes and sequenced with a 600-cycle MiSeq kit to ensure maximal coverage of each ORF. Reads containing mutations were discarded in the subsequent analysis.

Phage plaque assays

E. coli host strains were grown to saturation at 37°C in LB or Terrific Broth. To 10 ml of top agar (10 g/liter tryptone, 5 g/liter yeast extract, 10 g/liter NaCl, 7 g/liter agar), chloramphenicol was added to a final concentration 25 µg/ml and, if needed, ampicillin to a final concentration of 100 µg/ml. Ten-fold dilutions of phage in phosphate-buffered saline were spotted on the plates. After overnight incubation at 37°C, plates were photographed in a dark room with a white backlight.

Construction of mutant phages

PhiV-1 gp8 and gp19 knockout phages were constructed as previously described (66, 67)

using plasmid donors with homology arms to gp8 or gp19 in a *trxA*⁻ *E. coli* strain [JW5856 from the Keio collection (68)]. The *trxA* gene was inserted via the donor into the PhiV-1 genome as a selection marker.

Protein purification

Avs genes were cloned into pCDF-Duet expression plasmids containing a C-terminal 6xHis tag. PhiV-1 gp8 and gp19 genes were cloned into TwinStrep-SUMO expression plasmids. Proteins were expressed in *E. coli* BL21(DE3) cells (NEB no. C2527H). Cells were grown in Terrific Broth to mid-log phase, and the temperature was lowered to 18°C. Expression was induced at OD₆₀₀ 0.6 with 0.25 mM IPTG for 16 to 20 hours before harvesting and freezing cells at -80°C. Cell paste was resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5% glycerol) supplemented with EDTA-free cOmplete protease inhibitor (Roche). Cells were lysed using a LM20 microfluidizer device (Microfluidics), and cleared lysate was bound to either Strep-Tactin Superflow Plus (Qiagen) or Ni-NTA Superflow resin (Qiagen). For TwinStrep-SUMO phage proteins, the resin was washed with lysis buffer, and proteins were eluted with lysis buffer supplemented with 5 mM desthiobiotin. The TwinStrep-SUMO tag was removed by overnight digest at 4°C with homemade SUMO protease Ulp1 at a 1:100 weight ratio of protease to target. Cleaved proteins were run on a Superose 6 Increase column (GE Healthcare Life Sciences) with a final storage buffer of 25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% glycerol, and 1 mM dithiothreitol (DTT).

Avs proteins containing 6xHis tags were bound to Ni-NTA resin in the presence of 25 mM imidazole, washed with lysis buffer containing 50 mM imidazole, and eluted with lysis buffer containing 300 mM imidazole. SeAvs3 was diluted to a final concentration of 100 mM NaCl and purified using a Resource Q column on an AKTA pure 25 L (GE Healthcare Life Sciences) with a 100 mM-1M NaCl gradient. EcAvs4 was further purified by diluting to a final concentration of 100 mM NaCl and absorbing contaminants by flowing the protein over a Resource Q and Heparin HP column. SeAvs3 and EcAvs4 were concentrated and loaded onto a Superose 6 Increase column with a final storage buffer of 25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% glycerol, and 1 mM DTT. SeAvs3 for cryo-EM analysis was purified in the same buffer without glycerol and only 300 mM NaCl, then concentrated to 1.4 mg/ml in a 500-µl 100,000 molecular weight cutoff (MWCO) Amicon spin concentrator.

Avs complex purification

Avs-TwinStrep constructs were cotransformed with plasmids expressing either gp8 or gp19 into electrocompetent *E. coli* BL21(DE3) (Sigma

Aldrich CMC0016) and grown and induced as before. Avs pull-downs using Strep-Tactin Superflow Plus resin were run on SDS-PAGE (SDS-polyacrylamide gel electrophoresis) gels, and gp8 and gp19 bands were excised and confirmed by mass spectrometry (Taplin Biological Mass Spectrometry Facility, Harvard Medical School). For tandem affinity purification, plasmids containing SeAvs3-6xHis and gp19-StrepTag were cotransformed into electrocompetent *E. coli* BL21(DE3) and grown and induced as before. An SeAvs3-gp19 complex was purified using Ni-NTA followed by Strep-Tactin Superflow Plus resin. The final elution was run on a Superose 6 Increase column and yielded a peak elution at 13 ml containing a 1:1 ratio of SeAvs3 and gp19, as determined by SDS-PAGE band intensity analysis. A standard curve was generated using the Bio-Rad Gel Filtration Standard (1511901), and the gel-phase distribution coefficient (K_{ave}) was calculated as (elution volume - void column)/(column volume - void volume).

SeAvs3-terminase complex formation for cryo-EM

A total of 20 µg of SeAvs3 was mixed with 8.3 µg of PhiV-1 gp19 terminase in a total volume of 24 µl in the presence of 17 mM Tris-HCl pH 7.5, 280 mM NaCl, 0.8 mM DTT, 2% glycerol, 5 mM MgCl₂, and 1 mM ATP. The reaction was incubated at 37°C for 30 min and placed on ice for ~1 hour before cryo-EM grid preparation. Cryo-EM grids were prepared on a Thermo Scientific Vitrobot Mark IV at 4°C and 100% humidity. A total of 3 µl of reaction was applied to a freshly glow-discharged (12 s at 15 mA) Cu 300 R2/2 holey carbon grid with a 2-nm layer of amorphous carbon (Quantifoil). After 30 s, the grid was manually blotted with Whatman Grade 1 filter paper and plunged into liquid ethane.

EcAvs4-portal complex formation for cryo-EM

PhiV-1 gp8 was cloned into an MBP-bdSUMO expression plasmid, and EcAvs4 was cloned into a pCDF-Duet plasmid with an internal TwinStrep tag added between residues 114 and 115. The EcAvs4 Mrr-like nuclease active site was mutated (Q63A/K65A) to allow co-expression with the portal. These two plasmids were cotransformed into *E. coli* BL21(DE3). A total of 6 liters of culture was grown in Terrific Broth to mid-log phase, and the temperature was lowered to 18°C. IPTG (0.25 mM) was added to induce expression, and growth was continued overnight. Cell paste was resuspended in lysis buffer (50 mM Tris-HCl pH 7.4, 250 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol, 2 mM MgCl₂, 0.1 mM ATP) supplemented with EDTA-free cOmplete protease inhibitor (Roche). Cells were lysed using a LM20 microfluidizer (Microfluidics) and cleared lysate was bound to Amylose Resin

High Flow (NEB). After extensive washing with lysis buffer, the resin was eluted overnight at 4°C by addition of 10 µg of homemade bdSENPI protease. Eluted protein was incubated with Strep-Tactin Superflow Plus resin, washed with lysis buffer, then eluted with lysis buffer supplemented with 5 mM desthiobiotin. The eluate was concentrated in a 6-ml Vivaspin spin concentrator (30,000 MWCO) and run on a Superose 6 Increase column using 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 2 mM MgCl₂, and 0.1 mM ATP. Peak fractions containing EcAvs4 and gp8 were concentrated to 1.7 mg/ml using a 0.5 ml Amicon spin concentrator (100,000 MWCO) then immediately used for cryo-EM grid preparation. Cryo-EM grids were prepared on an Thermo Scientific Vitrobot Mark IV at 4°C and 100% humidity. A total of 3 µl of sample was applied to a freshly glow-discharged (60 s at 15 mA) Cu 300 R1.2/1.3 holey carbon grid (Quantifoil). The grid was blotted for 4 s with blot force +5 and drain time 1 s, then plunged into liquid ethane.

Cryo-EM data collection

All data were collected using the Thermo Scientific Titan Krios G3i cryo TEM at MIT.nano using a K3 direct detector (Gatan) operated in super-resolution mode with twofold binning and an energy filter with slit width of 20 eV.

For SeAvs3-gp19, 15,422 movies were collected at 105,000× magnification giving a real pixel size of 0.8697 Å, with defocus ranging from 1 to 3.5 µm with an exposure time of 1.15 s, fractionated into 30 frames and a flux of 19.7 e⁻/pix/s giving a total fluence per micrograph of 30 e⁻/Å². For EcAvs4-gp8, 22,902 movies were collected at 130,000× magnification giving a real pixel size of 0.6788 Å, with defocus ranging from 1 µm to 2.5 µm with an exposure time of 0.6 s, fractionated into 24 frames and a flux of 23.6 e⁻/pix/s giving a total fluence per micrograph of 30.8 e⁻/Å².

Cryo-EM data processing

All cryo-EM data were processed using RELION-4.0 (69). Movies were corrected for motion using the RELION implementation of MotionCor2, with 5-by-5 patches and dose-weighting. Contrast transfer function (CTF) parameters were estimated using CTFIND-4.1. All reported resolutions use the gold-standard Fourier shell correlation with a cutoff of 0.143.

For the SeAvs3-gp19 dataset, particle picking was first carried out on 800 micrographs using the Topaz general model (70). A good subset of these particles, as determined by three-dimensional (3D) classification, was used to train Topaz, and this trained model was used to pick 128,500 particles from the entire dataset. Extracted particles, downsampled fourfold, were subjected to 3D classification without imposing symmetry using a reference

derived from a preliminary dataset. A total of 44,489 particles, corresponding to 34.5% of picked particles, showed sharp features and apparent C4 symmetry and were reextracted without binning and refined with C4 symmetry imposed. After refining per-particle defocus and global magnification, beamtilt, trefoil, and performing Bayesian polishing, a reconstruction was yielded at 3.8-Å resolution with clear density for the SeAvs3 ATPase domain but blurred density for both the N-terminal nuclease and C-terminal TPR+terminase domains.

To improve density for the N-terminal nuclease domains, 3D classification without alignment was performed while imposing C2 symmetry. This revealed two equal populations of particles each with clear density for the nuclease domains, related by a 90° rotation in the *z* axis. In the refinement STAR file, the parameter *rlnAngleRot* was therefore incremented by 90° for one of these populations, before focused refinement starting at 1.8° local angular searches with a soft mask around the nuclease domains. This produced a reconstruction at 3.4 Å, measured using the same soft mask.

To improve density for the C-terminal TPR+terminase domains, C4 symmetry expansion was performed on the C4 refinement data.star file, followed by particle subtraction with recentering using a mask around one of the four asymmetric units. This generated four subparticles for each original particle. Refinement starting at 1.8° local angular searches with a soft mask, followed by CTF refinement and another round of refinement, produced a reconstruction at 3.4-Å resolution.

For the EcAvs4-gp8 dataset, 1825 particles from 80 micrographs were manually picked and used to train Topaz. The trained Topaz model then picked 444,626 particles from the entire dataset, which were extracted with fourfold binning and subjected with 3D classification without imposing symmetry using the octameric (pseudo-D4) EcAvs4-gp8 reference derived from a preliminary dataset. A total of 133,133 particles (29.9%) showing the same pseudo-D4 symmetry were reextracted at 1.034 Å/pix and refined with D4 symmetry imposed. After Bayesian polishing, this yielded a 3.7-Å resolution reconstruction. D1 symmetry expansion followed by particle subtraction was then used to convert these particles to 266,266 subparticles that correspond to the tetrameric complex. Like SeAvs3-gp19, these also had blurry density for the N-terminal nuclease and C-terminal TPR+terminase domains but additionally had poor density for the ATPase domains, suggesting a C2 reconstruction might be suitable for the whole tetramer.

To improve overall density, a 3D classification without alignment was first performed with C2 symmetry imposed. This produced

two equally occupied classes, collectively representing a 169,977-particle subset (63.8%), that appeared identical but for a 90° rotation, but less clearly distinguished than the same analysis on SeAvs3-gp19. Therefore, they were refined together with local 1.8° angular searches and C2 symmetry but with “Relax symmetry: C4” to account for the pseudo-C4 symmetry. This produced a consensus C2 refinement but still with relatively blurred densities for the nuclease and C-terminal TPR domains. The nuclease domain density was improved by focused refinement with a soft mask, followed by refining anisotropic magnification, per-(sub)particle defocus, and beamtilt, trefoil and fourth-order aberrations, and second refinement, yielding a 2.9-Å resolution reconstruction. The C-terminal TPR domains were improved by C4 symmetry expansion, followed by C1 focused refinement with a soft mask and CTF refinement, but still had unclear density at the periphery at the site of an important EcAvs4-portal contact. Therefore, a final 3D classification was performed with a soft mask just around this contact and a regularization parameter (T) of 20. A total of 500,066 selected subparticles (73%) were then focus-refined with the same mask to yield a reconstruction at 3.0-Å resolution with better density for this region.

Model building

Initial models for SeAvs3, EcAvs4, PhiV-1 gp8, and PhiV-1 gp19 were generated using AlphaFold and fit into the cryo-EM maps using ISOLDE (71) with adaptive distance restraints, followed by manual rebuilding in Coot (72) and further refinement in ISOLDE. Coordinates were refined in real space using PHENIX (73), performing one macrocycle of global minimization and atomic displacement parameter (ADP) refinement and skipping local grid searches.

In vitro cleavage reactions

Purified Avs proteins were incubated with nucleic acid substrates in reaction buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, 5% glycerol). Typical reactions contained ~100 ng of DNA substrate, 100 ng Avs protein, and 100 ng gp8 or gp19 in a 10-µl reaction volume. MgCl₂ was added at 5 mM where indicated, and ATP and AMP-PNP at 1 mM. Reactions were carried out at 37°C for the indicated time, and products were purified using a QIAquick PCR Purification column (Qiagen) before agarose gel analysis with a 1% E-Gel EX (Thermo Fisher Scientific).

Bacterial two-hybrid assays

Expression plasmids were cloned by fusing either the T18 or T25 fragments of CyaA from *Bordetella pertussis* to nuclease-deficient Avs proteins, as well as the PhiV-1 gp8 portal and

gp19 terminase. *E. coli* BTH101 cells (F^- , *cya-99*, *araDI39*, *galE15*, *galK16*, *rpsL1* (Str^r), *hsdR2*, *mcrA1*, *mcrB1*) were cotransformed with pairs of T18 and T25 containing plasmids. Overnight cultures were diluted 1:20 and plated on indicator plates containing 50 mg/ml ampicillin, 25 mg/ml kanamycin, 500 μ g/ml ammonium iron(III) citrate, 300 μ g/ml S-gal, and 0.5 mM IPTG. Cells were grown at 30°C overnight before imaging.

Southern blot analysis

E. coli K-12 (ATCC 25404) cultures were grown to mid-log phase (OD₆₀₀ 0.5), and for each sample, 6 ml of culture was infected with wild-type or mutant PhiV-1 at a multiplicity of infection of 1. After 20 min at 37°C, before cell lysis, infected cells were pelleted and resuspended in 200 μ l of media. After further incubation at 37°C, for a total of 90 min, samples were frozen in liquid nitrogen. DNA was extracted from 200- μ l cultures by adding 200 μ l of lysis buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl, 1% SDS, 2% Triton X-100), 100 μ l of glass beads, and 200 μ l of phenol-chloroform (1:1) followed by brief vortexing. Samples were centrifuged at 4°C, and DNA from the upper layer was extracted with chloroform and precipitated with the addition of 1 ml of ice-cold 100% ethanol and centrifugation at 4°C. DNA pellets were resuspended in 200 μ l of Tris-EDTA (TE) with 300 μ g RNase A (Sigma-Aldrich) and incubated at 37°C for 1 hour. DNA was once again precipitated with the addition of 1 ml of ice-cold 100% ethanol and 20 μ l of 4M ammonium acetate, centrifuged, dried, and resuspended in TE.

DNA was digested with Eco47 III and run on a 1% agarose gel in 1 \times Tris-borate-EDTA (TBE) at 100 V. The gel was denatured with 0.5 M NaOH and 1.5 M NaCl for 30 min and neutralized with 1.5 M NaCl and 0.5 M Tris-HCl pH 7.5 for 30 min. DNA was transferred to a Hybond N+ membrane (GE Healthcare Life Sciences) using overnight capillary flow and 10X SSC buffer (1.5 M NaCl, 150 mM sodium citrate, pH 7). Membranes were UV-crosslinked (Stratalinker 1800, Agilent) and blocked at 61°C with Church hybridization buffer (250 mM Na₂HPO₄ pH 7.2, 1 mM EDTA, 7% SDS). Radiolabeled probes complementary to the gp13 gene were generated from purified PCR products using the Prime-It Random labeling kit (Agilent) and [α -³²P]-dCTP. Membranes were probed overnight, washed three times with 61°C Church hybridization buffer, and exposed overnight with x-ray film (GE Healthcare Life Science) before developing. Quantification of phage DNA bands was performed in Fiji with background signal subtracted.

REFERENCES AND NOTES

1. K. S. Makarova, Y. I. Wolf, S. Snir, E. V. Koonin, Defense islands in bacterial and archaeal genomes and prediction of novel

- defense systems. *J. Bacteriol.* **193**, 6039–6056 (2011). doi: 10.1128/JB.05535-11; pmid: 21908672
2. E. V. Koonin, K. S. Makarova, Y. I. Wolf, Evolutionary genomics of defense systems in archaea and bacteria. *Annu. Rev. Microbiol.* **71**, 233–261 (2017). doi: 10.1146/annurev-micro-090816-093830; pmid: 28657885
3. S. Doron *et al.*, Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018). doi: 10.1126/science.aar4120; pmid: 29371424
4. L. Gao *et al.*, Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020). doi: 10.1126/science.aba0372; pmid: 32855333
5. H. G. Hampton, B. N. J. Watson, P. C. Fineran, The arms race between bacteria and their phage foes. *Nature* **577**, 327–336 (2020). doi: 10.1038/s41586-019-1894-8; pmid: 31942051
6. D. Klaiman, E. Steinfelds-Kohn, G. Kaufmann, A DNA break inducer activates the anticonductase RloC and the adaptive immunity in *Acinetobacter baylyi* ADP1. *Nucleic Acids Res.* **42**, 328–339 (2014). doi: 10.1093/nar/gkt851; pmid: 24062157
7. C. K. Guegler, M. T. Laub, Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. *Mol. Cell* **81**, 2361–2373.e9 (2021). doi: 10.1016/j.molcel.2021.03.027; pmid: 33838104
8. R. Cheng *et al.*, A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res.* **49**, 5216–5229 (2021). doi: 10.1093/nar/gkab277; pmid: 33885789
9. R. Bingham, S. I. Ekunwe, S. Falk, L. Snyder, C. Kleanthous, The major head protein of bacteriophage T4 binds specifically to elongation factor Tu. *J. Biol. Chem.* **275**, 23219–23226 (2000). doi: 10.1074/jbc.M002546200; pmid: 10801848
10. A. Millman *et al.*, Bacterial retrons function in anti-phage defense. *Cell* **183**, 1551–1561.e12 (2020). doi: 10.1016/j.cell.2020.09.065; pmid: 33157039
11. S. Kronheim *et al.*, A chemical defence against phage infection. *Nature* **564**, 283–286 (2018). doi: 10.1038/s41586-018-0767-x; pmid: 30518855
12. A. Bernheim *et al.*, Prokaryotic vipers produce diverse antiviral molecules. *Nature* **589**, 120–124 (2021). doi: 10.1038/s41586-020-2762-2; pmid: 32937646
13. K. S. Makarova, Y. I. Wolf, E. V. Koonin, Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* **41**, 4360–4377 (2013). doi: 10.1093/nar/gkt157; pmid: 23470997
14. E. V. Koonin, L. Aravind, Origin and evolution of eukaryotic apoptosis: The bacterial connection. *Cell Death Differ.* **9**, 394–404 (2002). doi: 10.1038/sj.cdd.4400991; pmid: 11965492
15. D. D. Leipe, E. V. Koonin, L. Aravind, STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: Multiple, complex domain architectures, unusual phylogenetic patterns, and evolution by horizontal gene transfer. *J. Mol. Biol.* **343**, 1–28 (2004). doi: 10.1016/j.jmb.2004.08.023; pmid: 15381417
16. Y. Zhao *et al.*, The NLRCA4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* **477**, 596–600 (2011). doi: 10.1038/nature10510; pmid: 21918512
17. E. M. Kofoed, R. E. Vance, Innate immune recognition of bacterial ligands by NALPs determines inflammasome specificity. *Nature* **477**, 592–595 (2011). doi: 10.1038/nature10394; pmid: 21874021
18. R. Caruso, N. Warner, N. Inohara, G. Núñez, NOD1 and NOD2: Signaling, host defense, and inflammatory disease. *Immunity* **41**, 898–908 (2014). doi: 10.1016/j.immuni.2014.12.010; pmid: 25526305
19. J. D. G. Jones, R. E. Vance, J. L. Dangl, Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, aaf6395 (2016). doi: 10.1126/science.aaf6395; pmid: 27934708
20. J. Heller, C. Clavé, P. Gladieux, S. J. Saupé, N. L. Glass, NLR surveillance of essential SEC-9 SNARE proteins induces programmed cell death upon allorecognition in filamentous fungi. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2292–E2301 (2018). doi: 10.1073/pnas.1719705115; pmid: 29463729
21. S. Bauenfried, M. J. Scherr, A. Pichlmair, K. E. Duderstadt, V. Hornung, Human NLRP1 is a sensor for double-stranded RNA. *Science* **371**, eabd0811 (2021). doi: 10.1126/science.abd0811; pmid: 33243852
22. O. Danot, A complex signaling module governs the activity of MalT, the prototype of an emerging transactivator family. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 435–440 (2001). doi: 10.1073/pnas.98.2.435; pmid: 11209048
23. S. Horinouchi *et al.*, Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in

- Streptomyces coelicolor* A3(2). *Gene* **95**, 49–56 (1990). doi: 10.1016/0378-1119(90)90412-K; pmid: 2253887
24. R. Ye, S. N. Rehemtulla, S. L. Wong, Glucanidation induction in *Bacillus subtilis* is mediated by a regulatory factor, GutR. *J. Bacteriol.* **176**, 3321–3327 (1994). doi: 10.1128/jb.176.11.3321-3327.1994; pmid: 8195087
25. E. V. Koonin *et al.*, Global organization and proposed megatranscriptome of the virus world. *Microbiol. Mol. Biol. Rev.* **84**, e00061-19 (2020). doi: 10.1128/MMBR.00061-19; pmid: 32132243
26. B. Lowey *et al.*, CBASS immunity uses CARF-related effectors to sense 3'-5'- and 2'-5'-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* **182**, 38–49.e17 (2020). doi: 10.1016/j.cell.2020.05.019; pmid: 32544385
27. K. S. Makarova *et al.*, Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antiviral defense. *Nucleic Acids Res.* **48**, 8828–8847 (2020). doi: 10.1093/nar/gkaa635; pmid: 32735657
28. J. Heitman, P. Model, Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.* **169**, 3243–3250 (1987). doi: 10.1128/jb.169.7.3243-3250.1987; pmid: 3036779
29. R. Martin *et al.*, Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector PopQ. *Science* **370**, eabd9993 (2020). doi: 10.1126/science.abd9993; pmid: 33273074
30. M. Zhou *et al.*, Atomic structure of the apoptosome: Mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes Dev.* **29**, 2349–2361 (2015). doi: 10.1101/gad.272728.115; pmid: 26543156
31. K. Steczkiewicz, A. Muszewska, L. Krniewski, L. Rychlewski, K. Ginalski, Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res.* **40**, 7016–7045 (2012). doi: 10.1093/nar/gks382; pmid: 22638584
32. A. Pingoud, M. Fuxreiter, V. Pingoud, W. Wende, Type II restriction endonucleases: Structure and mechanism. *Cell. Mol. Life Sci.* **62**, 685–707 (2005). doi: 10.1007/s00018-004-4513-1; pmid: 15770420
33. N. Watanabe, Y. Takasaki, C. Sato, S. Ando, I. Tanaka, Structures of restriction endonuclease HindIII in complex with its cognate DNA and divalent cations. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1326–1333 (2009). doi: 10.1107/S0907444909041134; pmid: 19966419
34. A. M. Burroughs, L. M. Iyer, L. Aravind, Comparative genomics and evolutionary trajectories of viral ATP dependent DNA-packaging systems. *Genome Dyn.* **3**, 48–65 (2007). doi: 10.1159/000107603; pmid: 18753784
35. G. Kaur, L. M. Iyer, A. M. Burroughs, L. Aravind, Bacterial death and TRADD-N domains help define novel apoptosis and immunity mechanisms shared by prokaryotes and metazoans. *eLife* **10**, e70394 (2021). doi: 10.7554/eLife.70394; pmid: 34061031
36. G. Kaur, A. M. Burroughs, L. M. Iyer, L. Aravind, Highly regulated, diversifying NTP-dependent biological conflict systems with implications for the emergence of multicellularity. *eLife* **9**, e52696 (2020). doi: 10.7554/eLife.52696; pmid: 32101166
37. J. E. Samson, A. H. Magadán, M. Sabri, S. Moineau, Revenge of the phages: Defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013). doi: 10.1038/nrmicro3096; pmid: 23979432
38. A. Pawluk, A. R. Davidson, K. L. Maxwell, Anti-CRISPR: Discovery, mechanism and function. *Nat. Rev. Microbiol.* **16**, 12–17 (2018). doi: 10.1038/nrmicro.2017.120; pmid: 29062071
39. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013). doi: 10.1038/nature11723; pmid: 23242138
40. N. S. Olsen, L. Forero-Junco, W. Kot, L. H. Hansen, Exploring the remarkable diversity of culturable *Escherichia coli* phages in the Danish wastewater environment. *Viruses* **12**, 986 (2020). doi: 10.3390/v12090986; pmid: 32899836
41. M. Schmerer, I. J. Molinuevo, J. J. Bull, Synergy as a rationale for phage therapy using phage cocktails. *PeerJ* **2**, e590 (2014). doi: 10.7717/peerj.590; pmid: 25279269
42. G. F. Hatfull, R. W. Hendrix, Bacteriophages and their genomes. *Curr. Opin. Virol.* **1**, 298–303 (2011). doi: 10.1016/j.coviro.2011.06.009; pmid: 22034588
43. H.-W. Ackermann, 5500 Phages examined in the electron microscope. *Arch. Virol.* **152**, 227–243 (2007). doi: 10.1007/s00705-006-0849-1; pmid: 17051420
44. C. L. Dedeo, G. Cingolani, C. M. Teschke, Portal protein: The orchestrator of capsid assembly for the dsDNA tailed

- bacteriophages and herpesviruses. *Annu. Rev. Virol.* **6**, 141–160 (2019). doi: [10.1146/annurev-virology-092818-015819](https://doi.org/10.1146/annurev-virology-092818-015819); pmid: 31337287
45. S. R. Casjens, The DNA-packaging nanomotor of tailed bacteriophages. *Nat. Rev. Microbiol.* **9**, 647–657 (2011). doi: [10.1038/nrmicro2632](https://doi.org/10.1038/nrmicro2632); pmid: 21836625
46. O. Danot, E. Marquet, D. Vidal-Ingigliardi, E. Richet, Wheel of life, wheel of death: A mechanistic insight into signaling by STAND proteins. *Structure* **17**, 172–182 (2009). doi: [10.1016/j.str.2009.01.001](https://doi.org/10.1016/j.str.2009.01.001); pmid: 19217388
47. S. Ma *et al.*, Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science* **370**, ea3069 (2020). doi: [10.1126/science.abe3069](https://doi.org/10.1126/science.abe3069); pmid: 33273071
48. B. R. Morehouse *et al.*, STING cyclic dinucleotide sensing originated in bacteria. *Nature* **586**, 429–433 (2020). doi: [10.1038/s41586-020-2719-5](https://doi.org/10.1038/s41586-020-2719-5); pmid: 32877915
49. S. A. Shmakov, K. S. Makarova, Y. I. Wolf, K. V. Severinov, E. V. Koonin, Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5307–E5316 (2018). doi: [10.1073/pnas.1803440115](https://doi.org/10.1073/pnas.1803440115); pmid: 29784811
50. S. F. Altschul *et al.*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997). doi: [10.1093/nar/25.17.3389](https://doi.org/10.1093/nar/25.17.3389); pmid: 9254694
51. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017). doi: [10.1038/nbt.3988](https://doi.org/10.1038/nbt.3988); pmid: 29035372
52. R. C. Edgar, MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004). doi: [10.1186/1471-2105-5-113](https://doi.org/10.1186/1471-2105-5-113); pmid: 15318951
53. M. Steinegger *et al.*, HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **20**, 473 (2019). doi: [10.1186/1471-2105-5-113](https://doi.org/10.1186/1471-2105-5-113); pmid: 15318951
54. K. S. Makarova *et al.*, Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020). doi: [10.1038/s41579-019-0299-x](https://doi.org/10.1038/s41579-019-0299-x); pmid: 31857715
55. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010). doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490); pmid: 20224823
56. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002). doi: [10.1093/nar/gkf436](https://doi.org/10.1093/nar/gkf436); pmid: 12136088
57. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015). doi: [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300); pmid: 25371430
58. L. Zimmermann *et al.*, a completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018). doi: [10.1016/j.jmb.2017.12.007](https://doi.org/10.1016/j.jmb.2017.12.007); pmid: 29258817
59. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021). doi: [10.1093/nar/gkab301](https://doi.org/10.1093/nar/gkab301); pmid: 33885785
60. A.-L. Van de Weyer *et al.*, A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell* **178**, 1260–1272.e14 (2019). doi: [10.1016/j.cell.2019.07.038](https://doi.org/10.1016/j.cell.2019.07.038); pmid: 31442410
61. W. Dyrka *et al.*, Diversity and variability of NOD-like receptors in fungi. *Genome Biol. Evol.* **6**, 3137–3158 (2014). doi: [10.1093/gbe/evu251](https://doi.org/10.1093/gbe/evu251); pmid: 25398782
62. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). doi: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2); pmid: 34265844
63. J. Pei, B. H. Kim, N. V. Grishin, PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300 (2008). doi: [10.1093/nar/gkn072](https://doi.org/10.1093/nar/gkn072); pmid: 18287115
64. S. Picelli *et al.*, Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014). doi: [10.1101/gr.177881.114](https://doi.org/10.1101/gr.177881.114); pmid: 25079858
65. M. Huber *et al.*, Translational coupling via termination-reinitiation in archaea and bacteria. *Nat. Commun.* **10**, 4006 (2019). doi: [10.1038/s41467-019-11999-9](https://doi.org/10.1038/s41467-019-11999-9); pmid: 31488843
66. U. Qimron, B. Marintcheva, S. Tabor, C. C. Richardson, Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19039–19044 (2006). doi: [10.1073/pnas.0609428103](https://doi.org/10.1073/pnas.0609428103); pmid: 17135349
67. A. M. Grigonyte *et al.*, Comparison of CRISPR and marker-based methods for the engineering of phage T7. *Viruses* **12**, 193 (2020). doi: [10.3390/v12020193](https://doi.org/10.3390/v12020193); pmid: 32050613
68. T. Baba *et al.*, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, 0008 (2006). pmid: 16738554
69. D. Kimanius, L. Dong, G. Sharov, T. Nakane, S. H. W. Scheres, New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021). doi: [10.1042/BCJ20210708](https://doi.org/10.1042/BCJ20210708); pmid: 34783343
70. T. Bepler *et al.*, Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019). doi: [10.1038/s41592-019-0575-8](https://doi.org/10.1038/s41592-019-0575-8); pmid: 31591578
71. T. I. Croll, ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018). doi: [10.1107/S2059798318002425](https://doi.org/10.1107/S2059798318002425); pmid: 29872003
72. A. Casañal, B. Lohkamp, P. Emsley, Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* **29**, 1069–1078 (2020). doi: [10.1002/pro.3791](https://doi.org/10.1002/pro.3791); pmid: 31730249
73. D. Liebschner *et al.*, Macromolecular structure determination using x-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019). doi: [10.1107/S2059798319011471](https://doi.org/10.1107/S2059798319011471); pmid: 31588918

ACKNOWLEDGMENTS

We thank D. Li, H. Altae-Tran, and G. Faure for valuable discussions; R. Tomaino for assistance with mass spectrometry; E. Brignole and C. Borsa for the smooth running of the MIT.nano cryo-EM facility; and the entire Zhang lab for support and advice. **Funding:** L.A.G. is supported by the Society of Fellows of Harvard University. K.S.M. and E.V.K. are supported by intramural funds of the US Department of Health and Human Services (to the National Library of Medicine). F.Z. is supported by NIH grants 1DP1-HL141201 and 2R01HG009761-05; the Howard Hughes Medical Institute; Open Philanthropy; the Edward Mallinckrodt, Jr. Foundation; the Poitras Center for Psychiatric Disorders Research at MIT; the Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT; the Yang-Tan Molecular Therapeutics Center at McGovern; and the Phillips family and J. and P. Poitras. **Author contributions:** L.A.G. performed genetic assays and computational analysis. M.E.W. performed cryo-EM and determined the structures. J.S. performed in vitro biochemistry and bacterial two-hybrid assays. L.A.G. and J.S. carried out phage plaque assays. L.A.G. and K.S.M. performed the phylogenetic analysis. L.A.G., M.E.W., J.S., K.S.M., E.V.K., and F.Z. analyzed data. F.Z. supervised the research with support from R.K.M. L.A.G., M.E.W., and J.S. wrote the manuscript with input from all authors. **Competing interests:** F.Z. is a scientific advisor and cofounder of Editas Medicine, Beam Therapeutics, Pairwise Plants, Arbor Biotechnologies, and Sherlock Biosciences. The Broad Institute has filed a patent application that includes work described in this manuscript. **Data and materials availability:** Expression plasmids for all Avs systems and inhibitors are available from Addgene (#188786-188861). The genome sequence of phage ZL19 has been deposited in GenBank under accession number OM258170.1. The cryo-EM maps have been deposited in the Electron Microscopy Data Bank with the following codes: EMD-27421 (SeAvs3-gp19, C2 refinement), EMD-27424 (SeAvs3-gp19, C4-expanded C1 refinement), EMD-27425 (EcAvs4-gp8, C2 refinement), EMD-27422 (EcAvs4-gp8, C2 refinement focused on Mrr), and EMD-27426 (EcAvs4-gp8, C4-expanded C1 refinement). The coordinates for the composite atomic models have been deposited in the Protein Data Bank under accession codes 8DGC (SeAvs3-gp19) and 8DGF (EcAvs4-gp8). **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm4096

Figs. S1 to S27

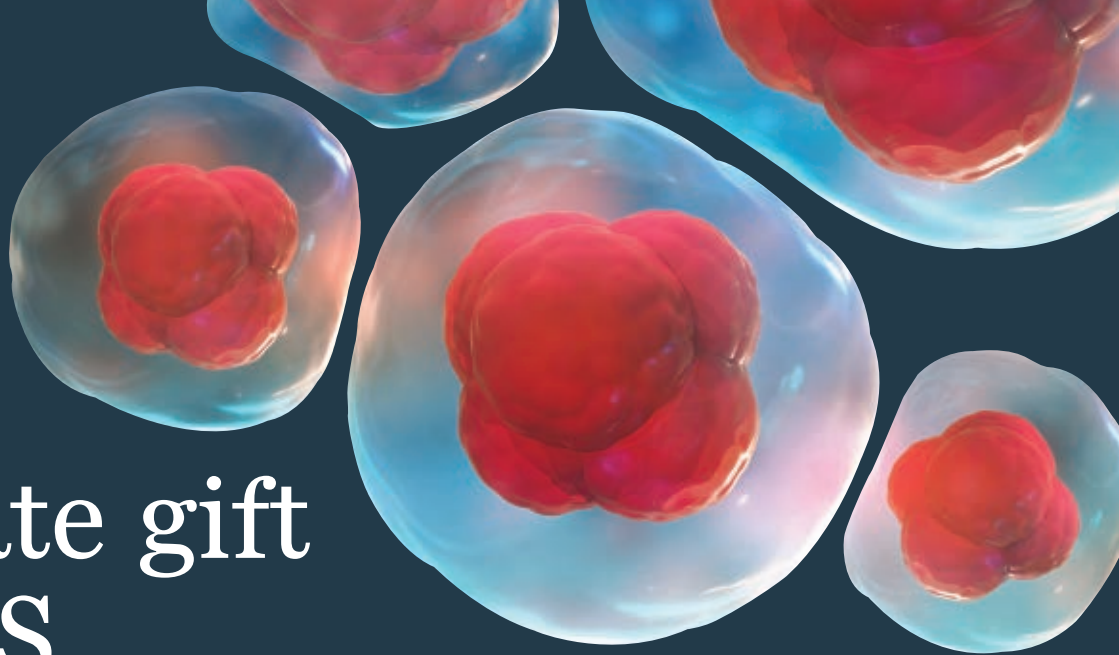
Table S1

References (74–82)

MDAR Reproducibility Checklist

Data S1 to S19

Submitted 15 September 2021; accepted 7 July 2022
10.1126/science.abm4096



An estate gift to AAAS

Going all the way back to 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

Today, we are dedicated to advocating for science and scientific evidence to be fully and positively integrated into public policy and for the community to speak with one voice to advance science and engineering in the United States and around the world.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you will become a member of our 1848 Society and will help fuel our work on behalf of science and society – including publishing the world’s most promising, innovative research in the *Science* family of journals and engaging in the issues that matter locally, nationally and around the world.

“As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all.”

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science

Federal Tax ID Number: 53-0196568

Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at www.aaas.org/1848Society.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____



RETURN THIS FORM TO:

AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLE SUMMARY

MOLECULAR BIOLOGY

LLPS of FXR1 drives spermiogenesis by activating translation of stored mRNAs

Jun-Yan Kang[†], Ze Wen[†], Duo Pan[†], Yuhang Zhang[†], Qing Li[†], Ai Zhong[†], Xinghai Yu[†], Yi-Chen Wu, Yu Chen, Xiangzheng Zhang, Peng-Cheng Kou, Junlan Geng, Ying-Yi Wang, Min-Min Hua, Ruiting Zong, Biao Li, Hui-Juan Shi, Dangsheng Li, Xiang-Dong Fu, Jinsong Li, David L. Nelson, Xuejiang Guo, Yu Zhou, Lan-Tao Gou, Ying Huang^{*}, Mo-Fang Liu^{*}

INTRODUCTION: In mammals, spermiogenesis (postmeiotic male germ cell differentiation) is a highly orchestrated developmental process controlled by a group of genes collectively referred to as spermiogenic genes. Because nuclear condensation during spermiogenesis gradually halts transcription, spermiogenic genes are transcribed in advance during the earlier stages of male germ development and stored as translationally inert messenger ribonucleoproteins (mRNPs) in developing spermatids until they are needed for translation. Such inert mRNPs are usually organized into mRNP granules called germ granules, which serve as storage facilities for nontranslating mRNAs in various types of germ cells. However, little is known about how those mRNAs stored in inert mRNPs are activated during late spermiogenesis.

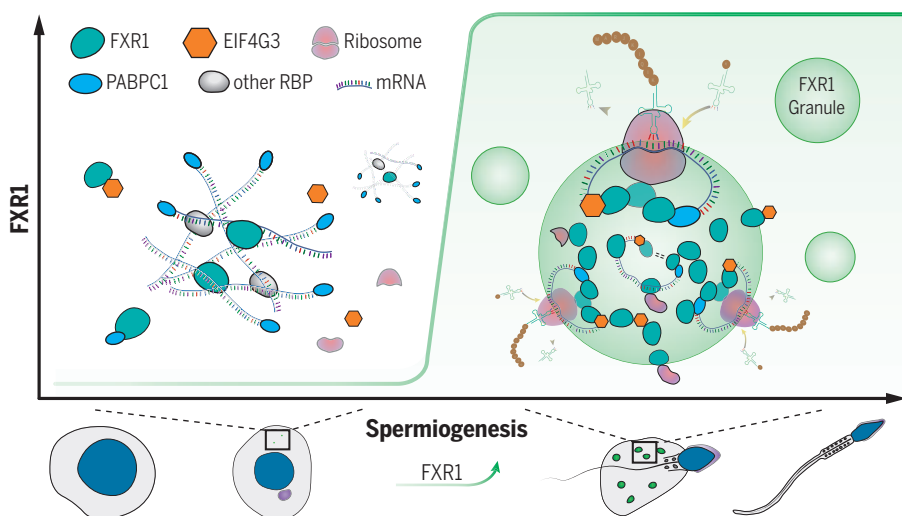
RATIONALE: To understand how translationally inert mRNAs are activated during spermiogenesis,

we screened potential translational regulators by proteomic analysis of polysomes from mouse testes. FXR1, a member of the fragile X-related (FXR) protein family, stood out from the screen as a translational regulator in late spermatids. By performing eCLIP and polysome profiling, in combination with generating a germline-specific *Fxr1* knock-out (*Fxr1*^{cko}) mouse model, we investigated whether FXR1 is required for translation activation in late spermatids. To decipher the mechanism underlying FXR1-mediated translation regulation, we identified the potential cofactor(s) of FXR1 in mouse testes using immunoprecipitation coupled with mass spectrometry. We observed the formation of FXR1 granules through liquid-liquid phase separation (LLPS), which recruits translation factors in late spermatids, and used the TRICK (translating RNA imaging by coat protein knock-off) reporter system to determine whether FXR1 LLPS is required for target translation in cultured

cells. To further investigate whether FXR1 LLPS is critical for target translation in mouse spermatids, we ectopically expressed wild-type FXR1, LLPS-deficient FXR1^{L351P} mutants, or LLPS-restored FXR1^{L351P}-IDR^{FUS} mutants in *Fxr1*^{cko} testes using lentiviral testis transduction. Finally, by generating germline-specific *Fxr1*^{L351P} knock-in mice, we determined whether FXR1 LLPS is indispensable to translation activation in late spermatids, spermiogenesis, and male fertility in mice.

RESULTS: We found that FXR1 was much more enriched in polysomes from 35-day postpartum (dpp) testes relative to 25-dpp testes, suggesting a role for FXR1 in translation activation in late spermatids. We identified a group of 770 mRNAs as being likely direct FXR1-activated targets, and demonstrated that germline-specific *Fxr1* deletion in mice markedly reduced target translation in late spermatids. Consistent with FXR1 functioning in translation activation in late spermatids, *Fxr1*^{cko} male mice were infertile and displayed spermatogenic failure at late spermiogenesis. Interestingly, we observed a pronounced up-regulation of FXR1 and the formation of abundant, distinct condensates in late spermatids, suggesting concentration-dependent LLPS. Mechanistic studies revealed that FXR1 undergoes LLPS to form condensates that assemble target mRNAs as mRNP granules and then recruit translational machinery to activate the stored mRNAs. Consistently, ectopic expression of wild-type FXR1 or FXR1^{L351P}-IDR^{FUS}, but not FXR1^{L351P}, activated target translation in cultured cells and successfully rescued target translation in late spermatids and spermiogenesis in *Fxr1*^{cko} mice. Furthermore, *Fxr1*^{L351P} knock-in mutant mice highly phenocopy *Fxr1*^{cko} mice, directly supporting the indispensability of FXR1 LLPS to target translation in late spermatids, spermiogenesis, and male fertility in mice.

CONCLUSION: Our findings demonstrate that FXR1 is an essential translation activator that instructs spermiogenesis in mice and unveil a key contribution of FXR1 LLPS to the translation activation of stored mRNAs in mouse spermatid and male fertility in mice. In addition, our study pinpoints the importance of LLPS in a developmental process in vivo. ■



FXR1-containing granules mediate translation activation in late spermatids. During late spermiogenesis, elevated FXR1 undergoes LLPS to assemble target mRNAs as FXR1 mRNP granules that recruit translational machinery by interacting with the eukaryotic translation initiation factor 4 gamma 3 (EIF4G3) to activate the stored mRNAs in late spermatids. These phase-separated FXR1 granules drive a large translation program to instruct spermatid development and sperm production in mice.

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

*Corresponding author. Email: mliu@sibcb.ac.cn (M.-F.L.); huangying@xinhua.com.cn (Y.H.)

[†]These authors contributed equally to this work.

Cite this article as J.-Y. Kang *et al.*, *Science* 377, eabj6647 (2022). DOI: 10.1126/science.abj6647

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abj6647>

RESEARCH ARTICLE

MOLECULAR BIOLOGY

LLPS of FXR1 drives spermiogenesis by activating translation of stored mRNAs

Jun-Yan Kang^{1†}, Ze Wen^{1†}, Duo Pan^{1†}, Yuhan Zhang^{2†}, Qing Li^{1†}, Ai Zhong^{1†}, Xinghai Yu^{3,4†}, Yi-Chen Wu¹, Yu Chen⁵, Xiangzheng Zhang⁵, Peng-Cheng Kou¹, Junlan Geng¹, Ying-Yi Wang¹, Min-Min Hua^{1,6}, Ruiting Zong⁷, Biao Li⁸, Hui-Juan Shi⁶, Dangsheng Li¹, Xiang-Dong Fu⁹, Jinsong Li^{1,10,11}, David L. Nelson⁷, Xuejiang Guo⁵, Yu Zhou^{3,4}, Lan-Tao Gou¹, Ying Huang^{2*}, Mo-Fang Liu^{1,10,11*}

Postmeiotic spermatids use a unique strategy to coordinate gene expression with morphological transformation, in which transcription and translation take place at separate developmental stages, but how mRNAs stored as translationally inert messenger ribonucleoproteins in developing spermatids become activated remains largely unknown. Here, we report that the RNA binding protein FXR1, a member of the fragile X-related (FXR) family, is highly expressed in late spermatids and undergoes liquid-liquid phase separation (LLPS) to merge messenger ribonucleoprotein granules with the translation machinery to convert stored mRNAs into a translationally activated state. Germline-specific *Fxr1* ablation in mice impaired the translation of target mRNAs and caused defective spermatid development and male infertility, and a phase separation-deficient FXR1^{L351P} mutation in *Fxr1* knock-in mice produced the same developmental defect. These findings uncover a mechanism for translational reprogramming with LLPS as a key driver in spermiogenesis.

Spermiogenesis (postmeiotic male germ cell development) in mammals is a highly orchestrated process during which haploid spermatids undergo sequential programmed transitions, including nuclear condensation and elongation, formation of flagella and acrosome, and elimination of cytoplasm, eventually transforming into spermatozoa (1). Such a substantial differ-

entiation process is strictly controlled by a group of spatially and temporally regulated genes collectively referred to as spermiogenic genes. Because nuclear condensation during spermatid elongation gradually halts transcription, spermiogenic genes are transcribed early and then stored as translationally repressed messenger ribonucleoproteins (mRNPs) until they are needed for translation (2, 3). Although this uncoupling between transcription and translation provides an elegant mechanism for the temporal control of gene expression during spermiogenesis, how the stored mRNAs are converted into a translationally activated state in developing spermatids has remained largely unclear.

Recent studies suggest that biomolecular condensates are widely present as non-membrane-bound subcellular compartments in eukaryotic cells, which likely enrich proteins and/or nucleic acids to achieve spatiotemporal control in specific cellular processes (4, 5). Although many proposed condensates remained to be fully characterized, one well-documented example is RNP granules, in which RNAs bound by RNA binding proteins (RBPs) are thought to self-organize into membraneless organelles (6). This process, known as liquid-liquid phase separation (LLPS), has been attributed to intrinsically disordered regions (IDRs) of RBPs (7–9), and increasing evidence suggests that multivalent interactions between RNA molecules also contribute to RNP assembly (10). In developing germ cells, RNP granules are better known as germ granules. They are prominently present in the cytoplasm and store nontranslat-

ing mRNAs (11). For example, the chromatoid body (CB), the largest RNP granule in mammalian male germ cells, stores >800 mRNAs in postmeiotic spermatids in mice (12). However, as with other inert mRNPs, little is known about how these mRNAs stored in germ granules are reactivated during germ cell development.

FXR1 belongs to the fragile X-related (FXR) family consisting of FMR1/FMRP, FXR1, and FXR2, which are canonical RBPs containing two KH domains and an RGG box (13). FXR family proteins have been primarily documented as translation suppressors that transport target mRNAs to different intracellular compartments (14). In particular, FMR1/FMRP is also known to repress local translation at synapses, and defects in this critical translational control step are responsible for the pathophysiology of fragile X syndrome (15). FXR1 has also been found in translationally inert mRNP granules (16), where it may be responsible for destabilizing certain target mRNAs in adult neural stem cells (17). Additionally, FMR1/FMRP has been reported to enhance the translation of mRNAs stored in *Drosophila* oocytes (18). Here, we found that the FXR1 protein undergoes LLPS to merge mRNP granules with translational machinery to activate the translation of the stored mRNAs during spermiogenesis.

Results

FXR1 is a potential translational regulator during late spermiogenesis

To understand how translationally inert mRNAs become activated during late spermiogenesis, we performed quantitative proteomic analysis to identify potential translational regulators in polysomes, which are active in protein synthesis (19), in mouse late spermatids (LS), which include elongating and elongated spermatids. Because polysomes in purified male germ cells collapsed, we prepared polysome fractions from mouse testes at two developmental time points: 25 days postpartum (dpp), which contain round spermatids (RS) but lack LS, and 35 dpp, which contain both RS and LS (Fig. 1A). Using proteomic analysis, we identified 50 proteins in testicular polysomes annotated to translation regulation by Gene Ontology (GO) analysis, among which 12 proteins were significantly elevated in polysomes from 35- relative to 25-dpp testes (Fig. 1B and data S1). TARBP2, a translation activator for *Protamine* mRNA during late spermiogenesis (20), was more enriched in the 35-dpp testicular polysomes (Fig. 1B and data S1), thus validating our screening strategy.

We focused on FXR1 given the previously documented roles of its paralog, FMR1/FMRP, in regulating translation (14, 18). The enrichment of FXR1 protein in 35- over 25-dpp testes (Fig. 1B) led us to hypothesize that FXR1 is involved in translation regulation in LS. Sucrose

¹State Key Laboratory of Molecular Biology, State Key Laboratory of Cell Biology, Shanghai Key Laboratory of Molecular Andrology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China. ²Shanghai Key Laboratory of Biliary Tract Disease Research, Shanghai Research Center of Biliary Tract Disease, Department of General Surgery, Xinhua Hospital, affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai, China. ³State Key Laboratory of Virology, Modern Virology Research Center, College of Life Sciences, Wuhan University, Wuhan, China. ⁴Frontier Science Center for Immunology and Metabolism, Wuhan University, Wuhan, China. ⁵State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing, Jiangsu, China. ⁶NHC Key Lab of Reproduction Regulation (Shanghai Institute of Planned Parenthood Research), Pharmacy School, Fudan University, Shanghai, China. ⁷Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. ⁸Department of Nuclear Medicine, Rui Jin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China. ⁹Department of Cellular and Molecular Medicine, University of California, San Diego, CA, USA. ¹⁰Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. ¹¹School of Life Science and Technology, Shanghai Tech University, Shanghai, China.

*Corresponding author. Email: mflu@sicb.ac.cn (M.-F.L.); huangying@xinhuaamed.com.cn (Y.H.)

†These authors contributed equally to this work.

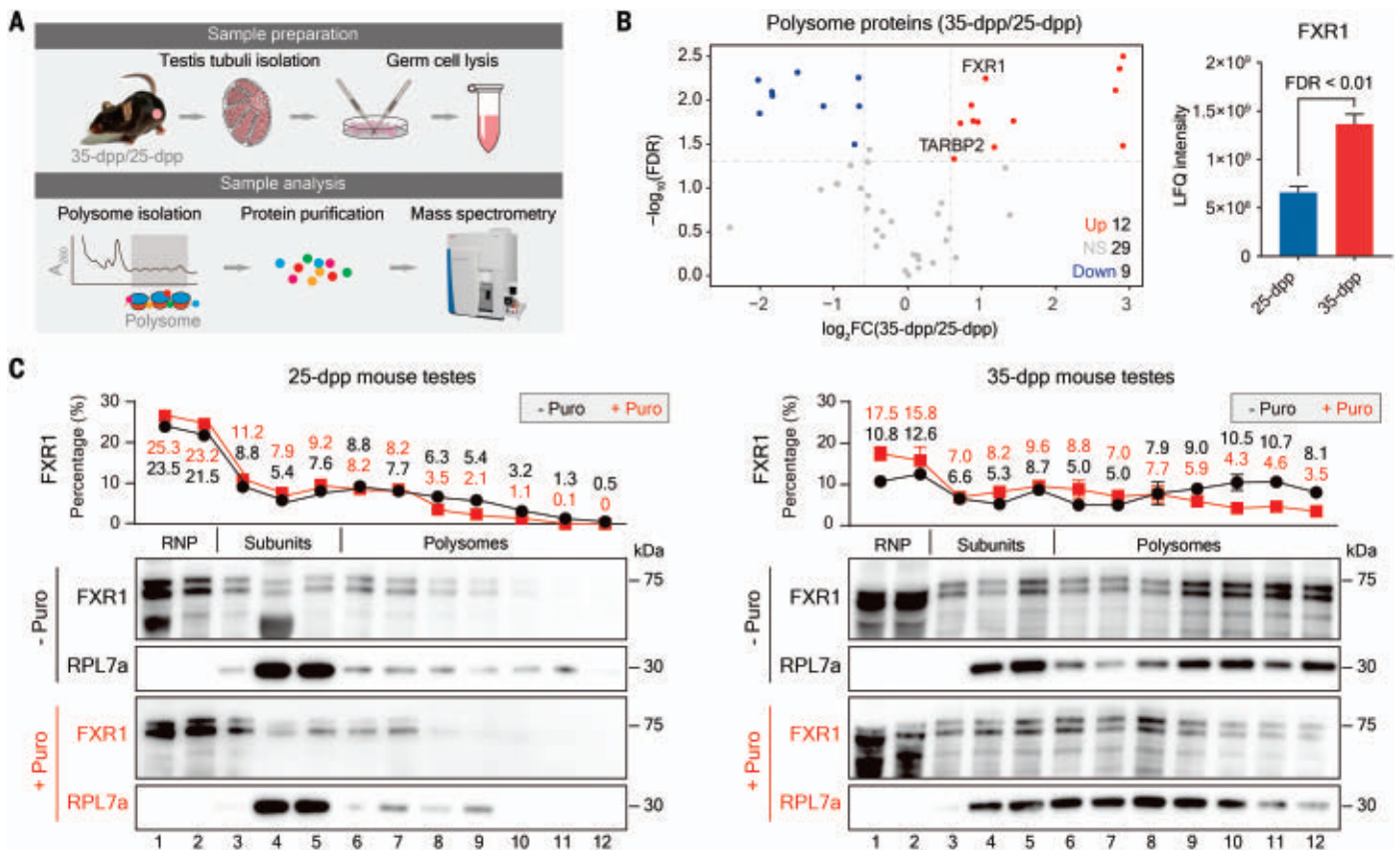


Fig. 1. Screening potential translational regulators in mouse testes.

(A) Schematic illustrating the experimental design for identifying translational regulators in mouse testes by polysome profiling coupled with mass spectrometry. (B) Identification of unique translational regulators in LS. Left, scatter plot showing the levels of all detected translational regulators in 35-dpp testicular polysomes relative to a 25-dpp control with two replicates, with the up-regulated (Up), down-regulated (Down), and unchanged proteins (NS)

indicated in red, blue, and gray dots, respectively. Horizontal dashed line indicates the FDR cutoff at 0.05; two vertical dashed lines indicate the fold-change cutoff at 1.5; right, label-free quantitation (LQF) intensity of FXR1 in 35- and 25-dpp testicular polysomes. (C) Quantification of FXR1 in sucrose gradient fractions from 25- or 35-dpp mouse testes with or without puromycin (Puro) treatment, with RPL7a serving as a reference ($n = 3$; data are shown as mean \pm SEM).

gradient analysis of polysome profiles showed that FXR1 level was elevated in the polysome fractions from 35- relative to 25-dpp testes, and puromycin treatment reduced FXR1 protein present in the heavy polysome fractions from 35-dpp testes (Fig. 1C and fig. S1A). These results indicate that FXR1 is much more associated with polysomes in adult mouse testes. *Fxr1* was much more expressed in testes compared with other tissues from adult mice (fig. S1B), and, moreover, *Fxr1* expression in testes increased as the mice aged (fig. S1C) and *Fxr1* was expressed more highly in LS than in spermatocytes (SC) and RS (fig. S1D). Together, these results suggest FXR1 as a potential translational regulator in LS.

FXR1 binds translationally active mRNAs in mouse LS

To determine whether FXR1 regulates translation during late spermiogenesis, we identified FXR1 target mRNAs using eCLIP-seq (21) and found 4272 FXR1-bound mRNAs in 35-dpp testes (data S2). Like FMR1/FMRP (22), FXR1

binds predominately within coding regions of target mRNAs (~70%; fig. S2, A and B), and we did not detect consensus sequence motifs in FXR1-binding sites. GO analysis of FXR1-bound mRNAs revealed a particular enrichment in spermatid differentiation and development, including flagellum formation, chromatin remodeling, and sperm motility (fig. S2C).

Many FXR1 targets have established roles in postmeiotic phase of male germ cell development, including *Gpd2* (23), *Akap3* (24), *Hkl1* (25), *Prkar2a* (26), *Stat4* (27), *Cap2* (28), and *Cytl1* (29). RNA coimmunoprecipitation in combination with quantitative polymerase chain reaction (RIP-qPCR) assays showed significant enrichment of these seven mRNAs in anti-FXR1 immunoprecipitates in adult mouse testes (Fig. 2A, left). These seven mRNAs were significantly associated with FXR1 protein in polysome fractions from adult testes (Fig. 2A, right), suggesting that FXR1 is involved in their translation. RS and LS of adult mouse testes contained similar levels of each mRNA, but LS contained higher levels of proteins

(Fig. 2B), suggesting that these candidate FXR1 targets are transcribed in earlier stages of male germ cells but are translated in LS. Immunostaining of testis sections confirmed that GPD2, AKAP3, HK1, and STAT4 proteins were mainly translated in LS (Fig. 2C). These results together suggest an involvement of FXR1 in translation activation during late spermiogenesis.

FXR1 depletion impairs translation of a subset of genes in adult mouse testes

To study the function of FXR1 in translation activation in LS, we generated germline-specific *Fxr1* knockout mice (referred to as *Fxr1*^{cko}; fig. S3, A and B). GPD2, AKAP3, HK1, and STAT4 protein levels were reduced in *Fxr1*^{cko} LS (Fig. 2D, left), but their mRNA levels were not changed (Fig. 2D, right). Immunostaining on testis sections confirmed their significant reduction in *Fxr1*^{cko} LS (Fig. 2E). Consistently, sucrose gradient analysis showed that all tested target mRNAs were reduced in the polysome fractions, with a concomitant increase in lighter RNP and 40S and 80S fractions from *Fxr1*^{cko}

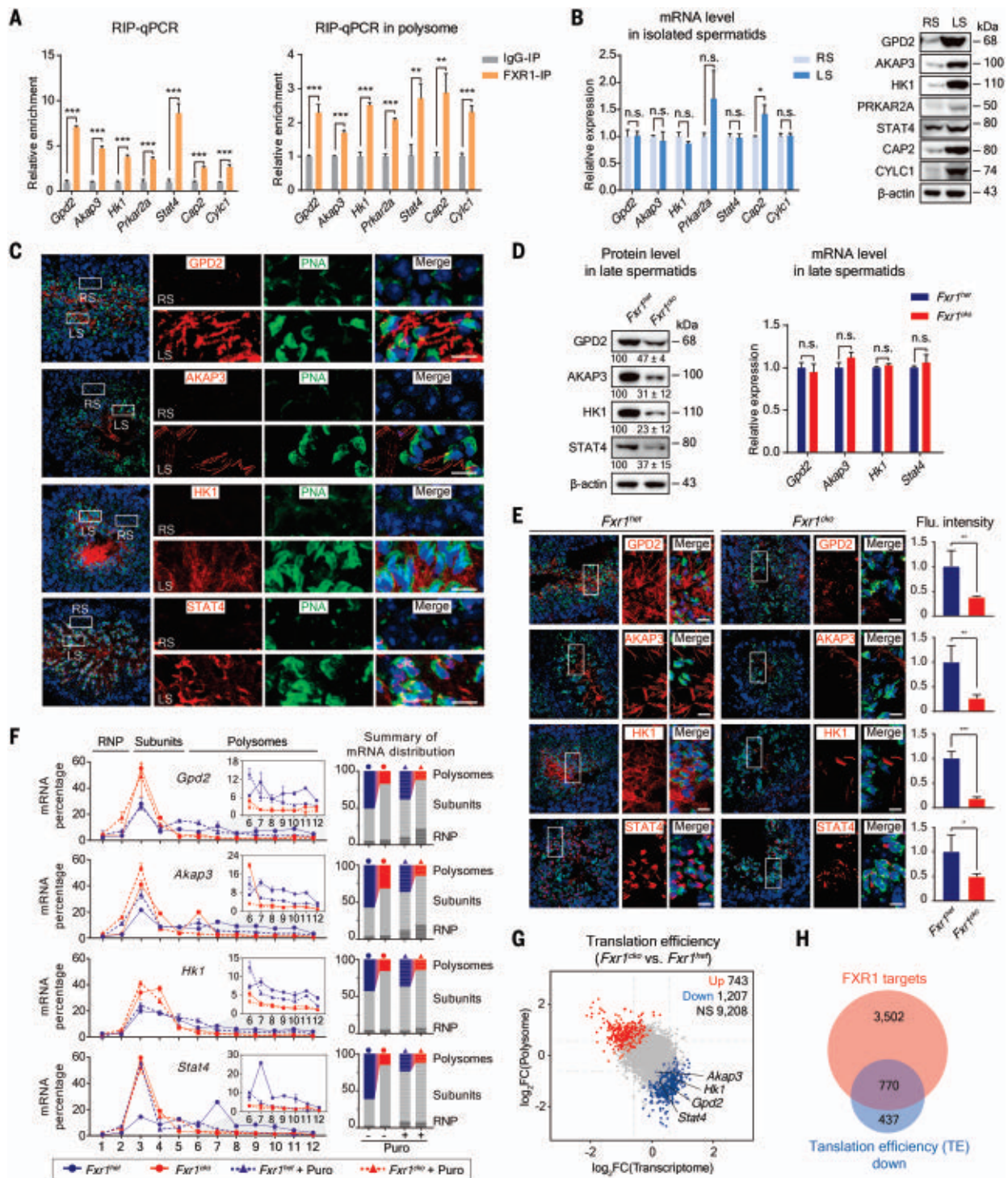


Fig. 2. FXR1 is required for the translation activation of a large group of genes in mouse testes. (A) RIP combined with RT-qPCR analyses of FXR1 binding to seven potential targets in adult mouse testes (left) or polysomes from adult mouse testes (right). (B) RT-qPCR (left) and Western blotting analyses (right) of FXR1 target expression in enriched RS and LS. (C and E) Immunostaining of FXR1 target expression (red) in testis sections from adult wild-type (C) and *Fxr1^{cko}* or *Fxr1^{het}* control mice (E), with an enlargement of white framed region on the right. Nuclei and acrosomes are counterstained with DAPI (blue) and PNA (green), respectively. Scale bar, 10 μ m. (D) Western blotting (left) and RT-qPCR analyses (right) of FXR1 target expression in *Fxr1^{cko}* and control LS. (F) Sucrose gradient combined with RT-qPCR analyses of FXR1 target mRNAs from adult

Fxr1^{cko} (red) or control (blue) testes with (dashed line) or without (solid line) puromycin treatment, with the polysome region shown as an inset. (G) Scatter plot showing mRNA-level changes (x-axis) against polysome-profiling changes (y-axis) between adult *Fxr1^{cko}* and control testes, with translation-deficient genes in *Fxr1^{cko}* testes shown in blue. Arrows indicate the four validated FXR1-activated targets. (H) Venn diagram showing the comparison of 4272 FXR1-bound targets (red) in adult mouse testes with 1207 translational efficiency–down-regulated (TE down) genes in *Fxr1^{cko}* testes (blue). Results shown in (A) to (F) and (G) and (H) are representative of three and two independent experiments, respectively. Significant differences by Student's *t* test are marked with asterisks: **P* < 0.05, ***P* < 0.01, ****P* < 0.001, n.s., not significant.

testes (Fig. 2F). As expected, puromycin treatment led to the reduction of all tested mRNAs in the polysome fractions (Fig. 2F and fig. S3C). By contrast, *Fxr1* deletion did not change the polysome-to-monosome ratio in adult testes (fig. S3C), indicating that FXR1 is not essential for global mRNA translation in mouse testes.

We assessed how many genes might be subjected to activation by FXR1 in mouse testes. Because of the poor quality of ribosome profiling (ribo-seq) libraries generated in adult mouse testes, we performed polysome profiling in parallel with RNA sequencing (RNA-seq) and transcriptome sequencing (fig. S3, D and E) and identified 1207 mRNAs that showed significantly reduced translational efficiency (polysome-associated mRNA over total) in adult *Fxr1*^{cko} testes relative to controls (Fig. 2G and data S3). These results suggest that FXR1 is involved in regulating the translation of a large group of mRNAs in mouse testes. GO analysis showed that many affected genes were involved in spermatid differentiation and development (fig. S3F). All four validated FXR1 targets, *Gpd2*, *Akap3*, *Hkl1*, and *Stat4*, showed significantly reduced translational efficiency in *Fxr1*-null testes (Fig. 2G and data S3). Cross analysis revealed that 770 of those 1207 mRNAs with reduced translational efficiency in *Fxr1*-null testes were directly

bound by FXR1 (Fig. 2H), suggesting that they are direct FXR1 targets. More than 85% of FXR1-activated target mRNAs were not enriched in the CB granule (fig. S3G), implying that FXR1 likely organizes a distinct RNP platform to regulate spermiogenic mRNA translation. Moreover, FXR1 target mRNAs are significantly longer, especially their coding regions (fig. S3H), implying that FXR1 preferentially promotes the synthesis of large proteins. These results indicate that FXR1 is involved in translation activation of a subset of mRNAs in mouse testes.

Fxr1 is essential for spermiogenesis and male fertility in mice

Despite the fact that *Fxr1*^{cko} male mice had normal size testes and epididymides (fig. S4A), they were sterile (Fig. 3A), whereas control *Fxr1*^{het} males were fertile. *Fxr1*^{cko} testes had normal numbers of RS and elongating spermatids at steps 14 to 16 (stages II to VIII) in the seminiferous tubules (Fig. 3, B and C), concomitant with apoptotic signals detected in LS (Fig. 3D). Consistently, transmission electron microscopy (TEM) of seminiferous epithelium revealed extensive defects in *Fxr1*^{cko} LS (fig. S4B). Moreover, we observed few spermatozoa in *Fxr1*^{cko} epididymides (Fig. 3E and fig. S4C).

These results demonstrated the functional importance of FXR1 in spermatid development and male fertility in mice.

Fxr1 recruits translation factors to FXR1 granules in LS

To understand the mechanism underlying FXR1-mediated translation activation, we next performed immunoprecipitation (IP) coupled with mass spectrometry to identify the potential cofactor(s) of FXR1 in mouse testes. A total of 125 proteins were significantly enriched in anti-FXR1 IP pellets (fig. S5A and data S4), with a GO term enrichment in positive regulation of translation (fig. S5B). Among them was the eukaryotic translation initiation factor 4 gamma 3 (EIF4G3) (fig. S5C), a scaffold protein that organizes the assembly of the EIF4F complex onto mRNAs for the initiation of cap-dependent translation (30, 31). The mouse genome encodes three EIF4G members, but only EIF4G3 has been shown essential for male fertility (32). We confirmed the interaction between FXR1 and EIF4G3 by co-IP in both mouse testes and cotransfected 293T cells (Fig. 4, A and B) and found their interaction to be resistant to RNase A treatment, suggesting a direct interaction between them. RIP-qPCR showed that all tested FXR1-activated targets were substantially reduced in anti-EIF4G3

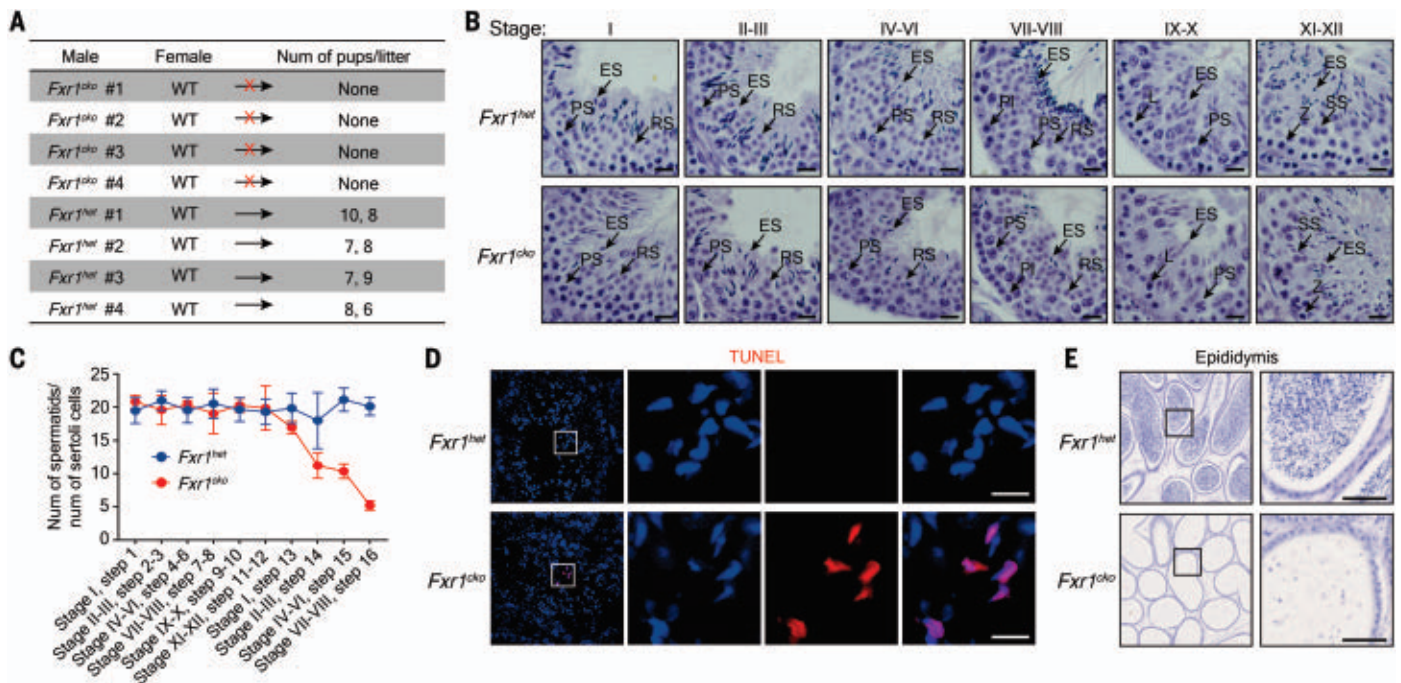


Fig. 3. Germline-specific depletion of FXR1 causes male infertility in mice.

(A) All tested *Fxr1*^{cko} males were infertile. (B) H&E staining of paraffin-embedded testis sections from adult control *Fxr1*^{het} and *Fxr1*^{cko} mice. The stages of seminiferous epithelium cycles were determined by the morphology of spermatocytes and spermatids. PI, preleptotene; L, leptotene; Z, zygotene; PS, pachytene spermatocyte; SS, secondary spermatocyte; ES, elongating spermatids. Scale bar, 20 μ m. (C) Comparison of the ratios of postmeiotic spermatids versus Sertoli cells in

tubule cross sections of specific stages of seminiferous epithelial cycles and corresponding spermatid development steps from adult control and *Fxr1*^{cko} mice. (D) Terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick end labeling (TUNEL) assay of testis sections from adult control and *Fxr1*^{cko} mice. Scale bar, 10 μ m. (E) H&E staining of paraffin-embedded cauda epididymis sections from adult control and *Fxr1*^{cko} mice. Scale bar, 100 μ m. Results of three independent experiments are shown.

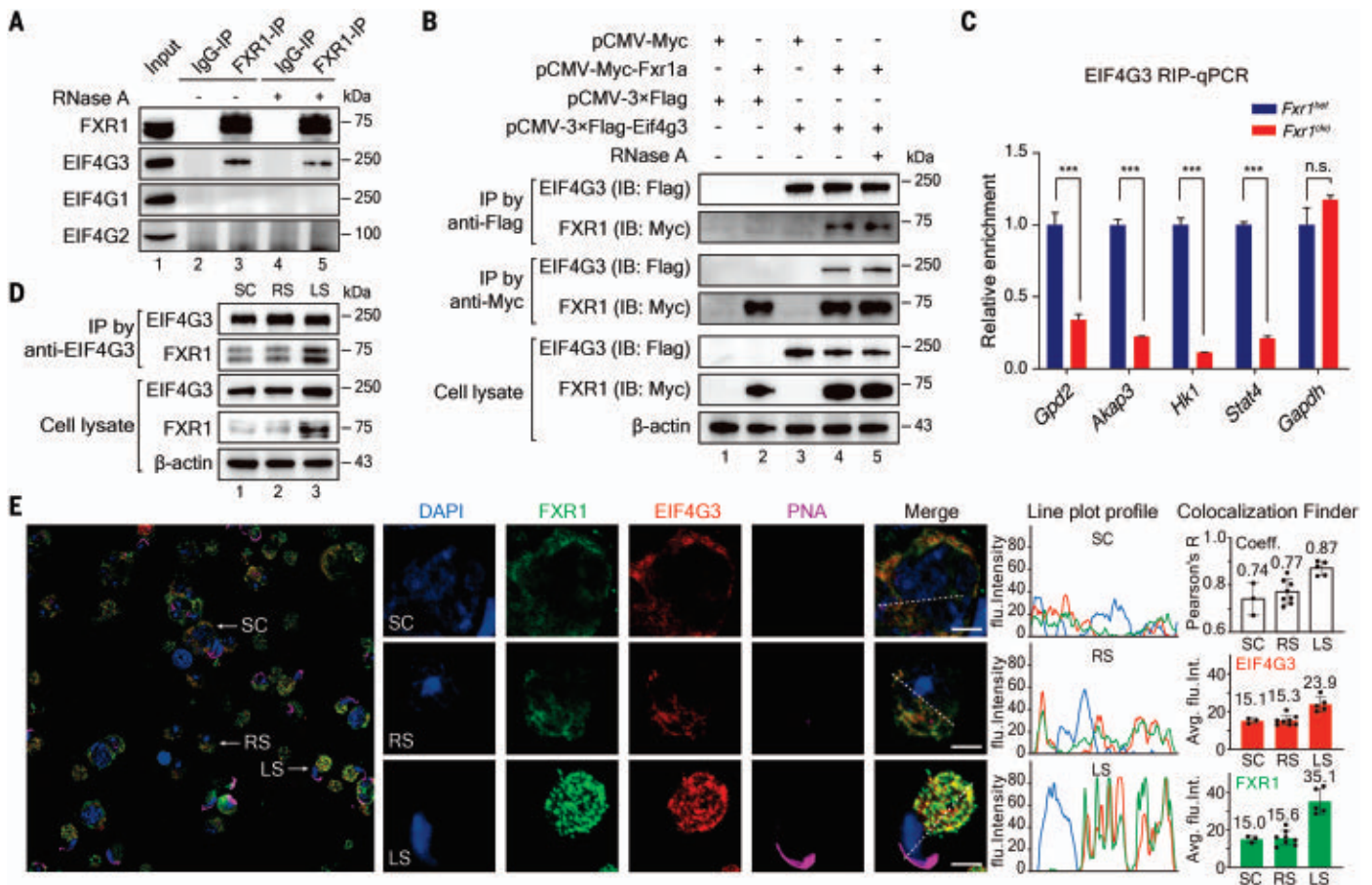


Fig. 4. FXR1 interacts with EIF4G3 in mouse testes. (A, B, and D) Co-IP assay of FXR1-EIF4G3 interaction in mouse testes (A), Myc-FXR1a⁻ and Flag-EIF4G3-cotransfected 293T cells (B), or enriched SC, RS, and LS (D). (C) RIP combined with RT-qPCR analyses of EIF4G3 binding to the four FXR1 targets in adult *Fxr1^{cko}* testes relative to controls. (E) Immunostaining of FXR1 (green) and EIF4G3 (red) in mouse male germ cells isolated from adult testes, with acrosomes counterstained with PNA (purple) and nuclei counterstained

with DAPI (blue). Middle, representative enlarged staining images of SC, RS, and LS. Right, line plot of fluorescence intensity and Colocalization_Finder (<http://punias.free.fr/ImageJ/colocalization-finder.html>) quantifying EIF4G3-FXR1 colocalization and relative protein abundance in SC ($n = 3$), RS ($n = 8$), and LS ($n = 5$). Scale bar, 5 μ m. Results of three independent experiments are shown. Significant differences by Student's *t* test are marked with asterisks: *** $P < 0.001$, n.s., not significant.

immunoprecipitates from *Fxr1^{cko}* testes compared with controls (Fig. 4C), indicating that FXR1 is required for EIF4G3 to efficiently bind to target mRNAs. Despite little interaction between FXR1 and EIF4G1 or EIF4G2 in mouse testes (Fig. 4A and data S4), both EIF4G1 and EIF4G2 interacted with FXR1 under overexpression conditions in cotransfected 293T cells (fig. S5D). The specificity of the FXR1-EIF4G3 interaction in mouse testes can likely be attributed to their similar expression pattern during spermatogenesis (fig. S5E). These findings point to a role of EIF4G3 in FXR1-mediated translation activation in LS.

Given that FXR1 targets are translationally activated primarily in LS (Fig. 2), we next investigated whether the FXR1-EIF4G3 interaction is differentially regulated during spermatogenesis in mice. The FXR1-EIF4G3 interaction was detected in SC and RS but became more evident in LS because of their

elevated expression there (Fig. 4D). FXR1 and EIF4G3 were largely present throughout the cytoplasm of SC and RS but preferentially formed abundant, distinct condensates that became more colocalized in LS (Fig. 4E). Given the markedly elevated FXR1 and EIF4G3 proteins in LS (Fig. 4, D and E), the formation of condensates may be driven by a concentration-dependent process characteristic of LLPS.

In addition to EIF4G3, several other translation factors, including poly(A) tail binding protein C1 (PABPC1), EIF4A3, and ribosome components, were enriched in FXR1 complexes from adult mouse testes (fig. S5C and data S4). PABPC1 has been reported to be actively involved in translation during spermatogenesis (33), likely facilitating translation initiation through association with eIF4G (34). Indeed, PABPC1 colocalized with FXR1 in SC and RS and was recruited to FXR1 condensates in LS (fig. S5F). Moreover, the large ribosomal subunit protein RPL7a was in close proximity to

FXR1-containing granules in LS (fig. S5G). These results suggest that FXR1 orchestrates translation activation of its target mRNAs by recruiting multiple components of the translation machinery.

FXR1 undergoes phase separation to form mRNA granules in LS

The observation of FXR1-containing condensates in LS suggests its distinct functions in early versus later stages of spermiogenesis (Fig. 4E and fig. S5, F and G). FMR1/FMRP, another FXR family protein, can form biomolecular condensates through LLPS (35). We investigated whether FXR1 also undergoes phase separation to form condensates in mouse LS. *Fxr1* encodes seven isoforms, FXR1a to FXR1g, through alternative splicing (36), and all of these contain IDRs at the C terminus (fig. S6A). We found all seven *Fxr1* transcript isoforms in mouse testes (data S5). FXR1a (a shorter form) and FXR1d (a longer form)

proteins appeared to be relatively more abundant in LS (fig. S6C). Purified FXR1 isoforms formed spherical droplets in vitro (Fig. 5A and fig. S6, D and E). After photobleaching, FXR1a droplets recovered most fluorescence signals within ~180 s in fluorescence recovery after photobleaching (FRAP) analysis in vitro (Fig. 5B). Moreover, all FXR1 isoforms formed round condensates that exhibited fusion behavior in C2C12 cells (Fig. 5C and fig. S6F), whereas FXR1a condensates recovered most fluorescence signals within ~120 s after photobleaching (Fig. 5D). Given that the diameters of FXR1 droplets/granules (1.0 to 5.0 μm) were

much larger than the resolution of confocal microscopy (~120 nm), the captured FXR1 droplets and/or granules may represent their real structures or features (37). These results demonstrated that FXR1 proteins could undergo LLPS and that FXR1 condensates display liquid-like features in vitro and in cellulo. In addition, we estimated the cytoplasmic concentration of FXR1 proteins to be ~4 μM in RS and ~23 μM in LS (fig. S7), corresponding to abundant, distinct FXR1 condensates in LS but not in RS (Fig. 4E and fig. S5, F and G).

To determine whether FXR1 condensates function to assemble target mRNAs as mRNP

granules, we mixed Alexa Fluor 488-labeled-FXR1a protein with Cy3-labeled-*Gpd2* mRNA. FXR1 droplets enriched target mRNA *Gpd2* in vitro, and the addition of *Gpd2* promoted FXR1 LLPS in a dosage-dependent manner (Fig. 5E). The FXR1 target mRNAs *Gpd2* and *Akap3* were dispersed in the cytoplasm and were significantly colocalized with FXR1 but not with DDX4 (a CB marker) in RS, and then became accumulated into FXR1 condensates in LS (fig. S8A). Conversely, the MIWI target *Tbpl1* was highly colocalized with DDX4 but not with FXR1. Moreover, biotin-labeled *Gpd2* or *Akap3* antisense pull-down showed that

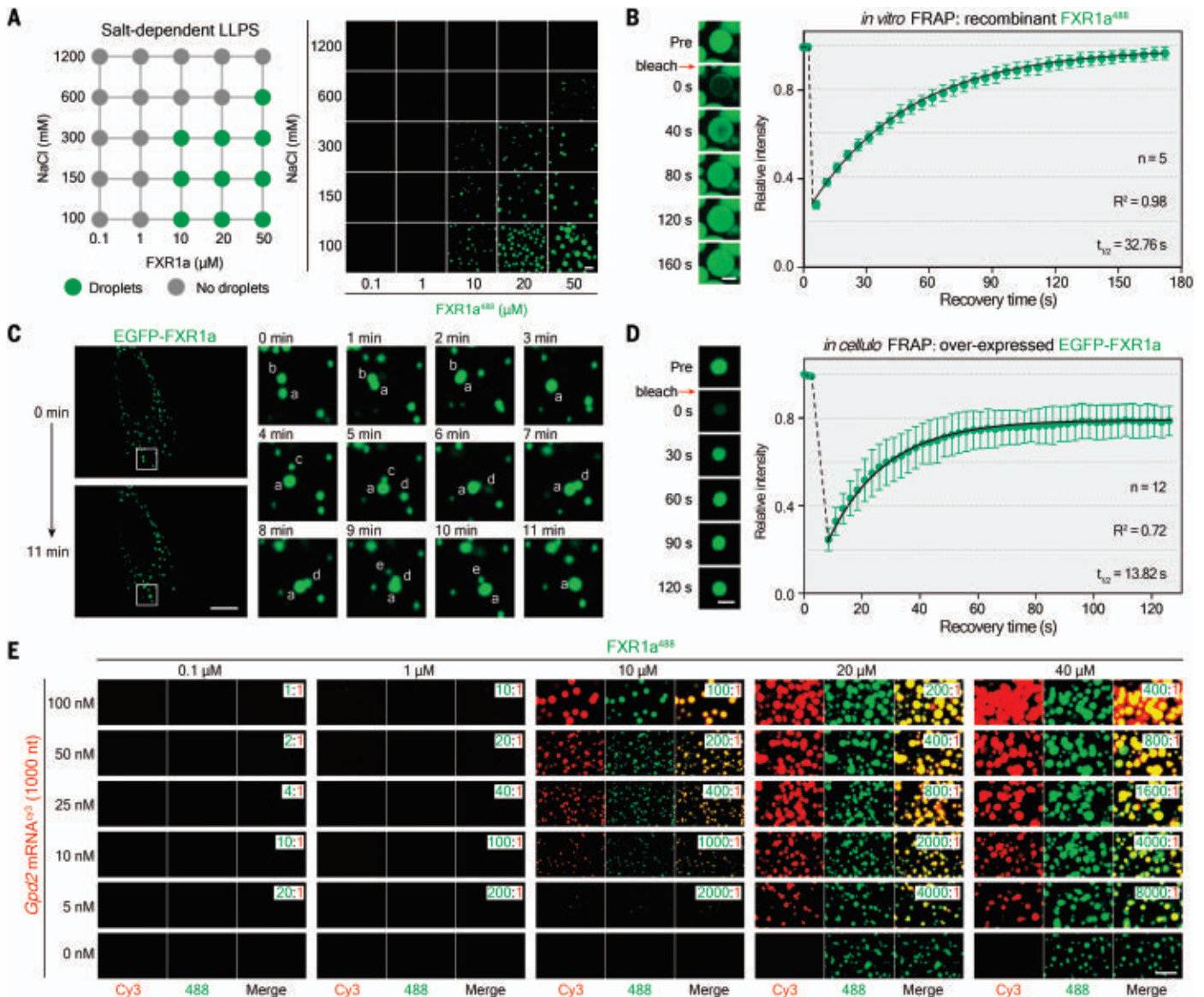


Fig. 5. FXR1 forms mRNP granules with its target mRNAs through LLPS.

(A) LLPS of purified recombinant FXR1a at various amounts and with increasing concentrations of NaCl under 10% PEG. Left, summary of phase-separation behavior of FXR1a; right, representative fluorescent microscopy images. Scale bar, 5 μm . (B) FRAP analysis of FXR1a droplets in vitro. Scale bar, 2 μm . (C) Live-cell images showing the fusion of smaller EGFP-FXR1a granules

into larger spherical granules in transfected C2C12 cells. Scale bar, 10 μm . (D) FRAP analysis of EGFP-FXR1a in transfected C2C12 cells. Scale bar, 2 μm . (E) LLPS of purified recombinant FXR1a at various amounts and with increasing concentrations of *Gpd2* mRNA (~1000 nucleotides) in 150 mM NaCl without PEG. Scale bar, 5 μm . Results of three independent experiments are shown.

FXR1 granules contained different target mRNAs (fig. S8B). These results support that FXR1 and target mRNAs form mRNP granules in LS.

FXR1 phase separation is essential for translation activation of its targets

We used a TRICK (translating RNA imaging by coat protein knock-off) reporter system as described previously (38, 39) to determine whether FXR1 granules are required for the translation of its targets (fig. S9A). As expected, cotransfection of NLS-PCP-GFP [a PP7 coat protein (PCP) fused to a nuclear localization sequence (NLS) and green fluorescent protein (GFP)] and NLS-MCP-RFP [a MS2 coat protein (MCP) fused to an NLS and RFP] resulted in two-color labeling of nuclear *Gpd2* mRNA in transfected C2C12 cells (Fig. 6A, panel I), indicative of the untranslated state. Coexpression of blue fluorescent protein (BFP)-tagged FXR1a yielded a fraction of *Gpd2* into FXR1a granules (blue) in the cytoplasm (Fig. 6A, panels II and III). *Gpd2* mRNA in FXR1a granules was labeled with RFP but not GFP, whereas puromycin treatment resulted in labeling of the mRNA in FXR1a granules by both GFP and RFP. These results demonstrate that FXR1a recruits target mRNA into FXR1a granules to trigger its translation.

We mapped the leucine residue at position 351 in the KH2 domain of FXR1a as an essential residue for phase separation (fig. S9, B to D). However, the FXR1a^{L351P} mutant maintained normal RNA-binding capability (fig. S9, E and F), whereas a L351P mutation impaired FXR1 phase separation likely by altering the hydrophobic pocket and conformation of Y353 and Y357 in the KH2 domain (fig. S9, G and H). As expected, coexpression of BFP-FXR1a^{L351P} led to the recruitment of *Gpd2* mRNA in the cytoplasm, but the mRNA was untranslated, because it was labeled with both GFP and RFP with or without puromycin treatment (Fig. 6A, panels IV and V). These results indicate that FXR1a^{L351P} can recruit target mRNA to the cytoplasm but cannot activate its translation, supporting that FXR1 relies on its phase-separation ability to activate target translation. Supporting this, the addition of FUS N-terminal IDR region (IDR^{FUS}) to the FXR1a^{L351P} mutant (FXR1a^{L351P}-IDR^{FUS} mutant) resulted in the formation of FXR1a/*Gpd2* granules in the cytoplasm and active translation of *Gpd2* mRNA in the granules (Fig. 6A, panels VI and VII). Western blot also confirmed that FXR1a^{L351P} lost the stimulatory effect on *Gpd2* translation considerably, whereas its phase separation-restored mutant rescued GPD2 expression (Fig. 6B). These data support that phase separation is essential for FXR1-mediated target translation.

To further investigate whether phase separation is critical for FXR1 function in translation activation in mouse spermatids, we

constructed FXR1-internal ribosome entry site (IRES)-cytochrome oxidase IV N-terminal (Cyto IV)-fused enhanced GFP (EGFP) lentiviral vectors (fig. S9I) to ectopically express wild-type FXR1, FXR1^{L351P}, or FXR1^{L351P}-IDR^{FUS} in *Fxr1*^{cko} testes (40). We found that wild-type FXR1a, FXR1d, or their FXR1^{L351P}-IDR^{FUS} mutants, but not FXR1^{L351P} mutants, rescued FXR1 granules (Fig. 6C) in *Fxr1*-null LS and protein expression of target genes without altering their mRNA levels (Fig. 6D).

FXR1 phase separation is required for sperm formation in mice

We found that the ectopic FXR1 protein rescued sperm formation in transduced *Fxr1*^{cko} mice. The coexpressed Cyto IV-tagged EGFP anchors to the mitochondria of sperm and thus allowed us to track the spermatozoa progressed from transduced spermatids (fig. S9I). As shown by differential interference contrast microscopy, spermatozoa were absent in epididymides from untransduced or control Cyto IV-EGFP-transduced *Fxr1*^{cko} mice, whereas spermatozoa were detectable in epididymides from wild-type FXR1a, FXR1d, or their FXR1^{L351P}-IDR^{FUS} mutant-transduced mice, but not FXR1^{L351P}-transduced mice (Fig. 6E). As expected, all of these spermatozoa were GFP⁺, indicating that they had progressed from the transduced *Fxr1*^{cko} spermatids. These spermatozoa fertilized oocytes in intracytoplasmic sperm injection to produce blastocysts (Fig. 6F), indicating that they maintain the genome integrity required to support embryonic development. Moreover, similar to the active controls (movie S1), all GFP⁺ spermatozoa swam actively (movies S2 to S5), suggesting that they are functional spermatozoa. We conclude that phase separation of FXR1 is functionally required for spermatid development in mice.

FXR1 phase separation is indispensable to male fertility in mice

We generated germline-specific *Fxr1*^{L351P} knock-in (referred to as *Fxr1*^{cki}) mice (fig. S10, A and B), and found that translation of tested FXR1 targets was reduced in *Fxr1*^{cki} LS (fig. S10C). *Fxr1*^{cki} males were sterile (fig. S10D) with condensed spermatids at steps 15 and 16 (stages IV to VIII) that were substantially reduced in the seminiferous tubules, apoptotic signals were detected in LS (fig. S10, E to G), and spermatozoa were barely present in epididymides (fig. S10, H and I). These results show that *Fxr1*^{cki} males highly phenocopy *Fxr1*^{cko} mice, supporting the functional requirement of FXR1 phase separation for spermatid development and male fertility in mice.

Discussion

Our results uncovered FXR1 as a key translation activator in LS and identified a large

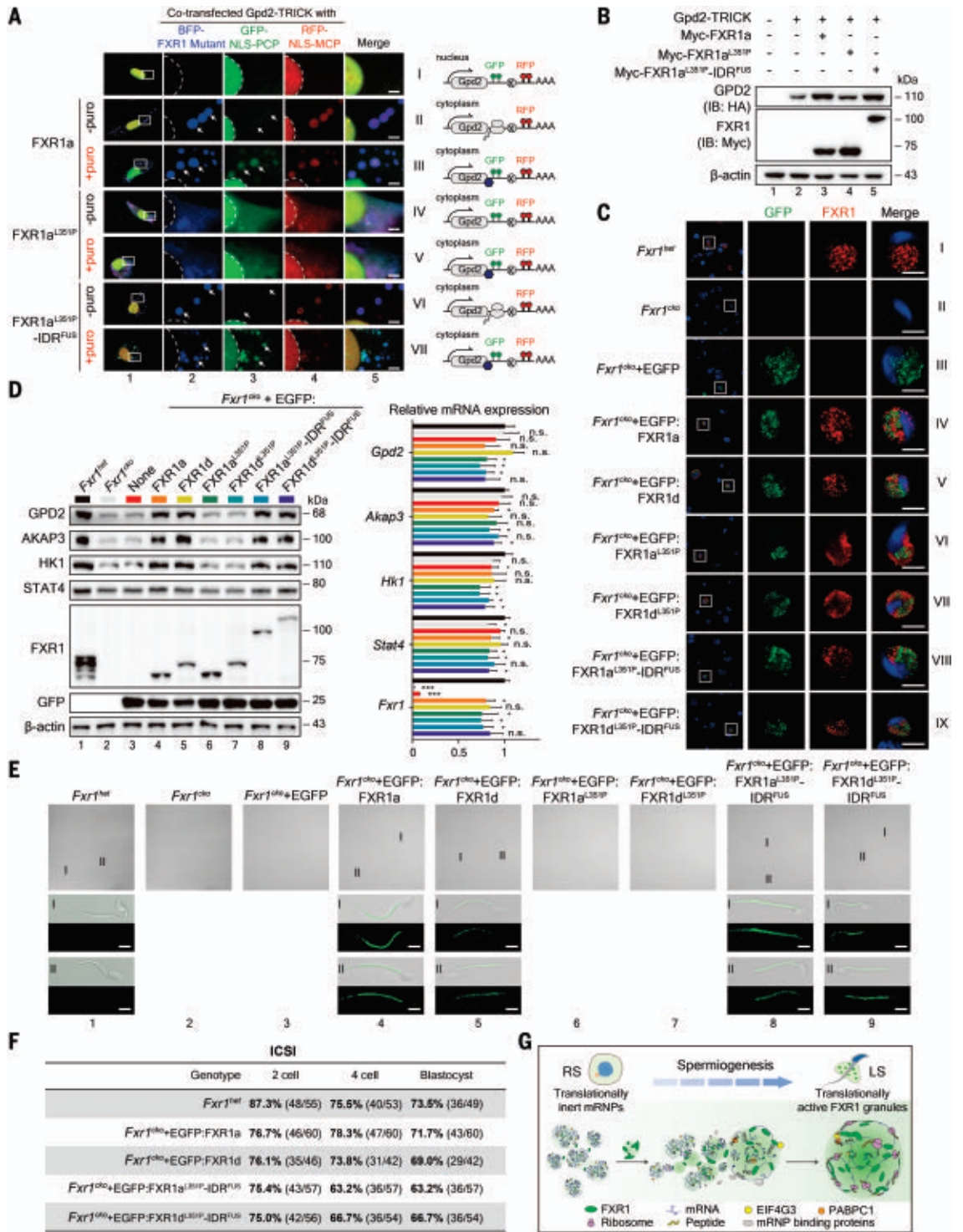
group of spermiogenic genes as likely being direct FXR1-activated targets. Our mechanistic studies revealed that the FXR1 protein undergoes LLPS to form mRNP granules and recruit the translational machinery to translationally activate stored mRNAs within them (Fig. 6G). In RS, where FXR1 is expressed at a low level, the newly transcribed FXR1 target mRNAs may be assembled into translationally inert mRNPs by association with a low level of FXR1 and other cytoplasmic RBPs and/or intermolecular RNA-RNA interactions. In LS, elevated FXR1 protein may “invade” these mRNPs using phase separation to gradually disrupt existing RNA-RNA interactions and/or replace other RBPs to transform the stored mRNAs into new FXR1-dominated mRNP granules. During this transformation process, FXR1 also recruits the translation machinery through its direct interaction with EIF4G3, thereby converting its target mRNAs into a translationally activated state. The proposed mechanism suggests a key role of FXR1 granules in enriching both mRNA transcripts and the translational machinery, whereby it induces the translation of spermiogenic mRNAs in LS.

FXR1 is required for muscle development in mice (41). Similar to our observations, a recent study showed that FXR1 also formed biomolecular condensates in muscle cells (42), but the function of FXR1 phase separation in translational control and muscle development remains to be determined. Recessive mutations in exon 15 of FXR1 are associated with congenital multi-minicore myopathy, and a patient carrying such mutations also showed hypoplastic genitalia and cryptorchidism (43). These findings suggest that defective FXR1 functions in humans might induce abnormal development in both muscles and the male genital system.

In contrast to the well-documented roles of FXR family members in translational repression, FXR1 enhances translation in mouse spermatids. Similar to our observations, an early report showed a positive role of FXR1 in Ago2 and/or miRNA-mediated translation activation in quiescent conditions (44). In addition, a recent study report that FXR1 promotes cMYC translation in cancer cells by recruiting the eIF4F complex to the translation start site (45). Our finding is also reminiscent of a recent study in *Drosophila* oocytes, in which FMRI/FMRP preferentially stimulated the translation of a large set of stored mRNAs (18). Given the lack of binding specificity for either FXR1 or FMRI/FMRP, we suspect that FMRI/FMRP may use a similar phase-separation mechanism to activate its target mRNAs in *Drosophila* oocytes. Thus, the current findings suggest a multifaceted function of FXR family proteins, likely in a context-dependent manner, providing important implications for the function of FXR family members in critical development and disease processes.

Fig. 6. Phase separation is essential for FXR1 function in translation activation and spermatid development in mice.

(A) Live-cell imaging analysis of the effect of wild-type FXR1a, FXR1a^{L351P}, or FXR1a^{L351P}-IDR^{FUS} mutants on *Gpd2* translation using the TRICK reporter system in cotransfected C2C12 cells. Left, representative live-cell images, with the nucleus and FXR1 granules indicated by dotted lines and arrows, respectively. Scale bar, 2 μm. Right, schematic illustrating the translation state of *Gpd2*. **(B)** Western blotting of GPD2 expression in the indicated cotransfected C2C12 cells, with β-actin serving as an internal control. **(C)** Immunostaining of FXR1 (red) in mouse spermatids isolated from the indicated mouse testes. Nuclei are counterstained with DAPI (blue). EGFP (green) indicates the transduced cells, with representative enlarged views of white framed LS shown on the right. Scale bar, 10 μm. **(D)** Western blotting (left) and RT-qPCR analyses (right) of FXR1 target expression in EGFP⁺ spermatids sorted from transduced *Fxr1*^{cko} testes. **(E)** Differential interference contrast microscopy images of epididymal sperm from indicated lentiviral expression vector-transduced *Fxr1*^{cko} mice, with enlarged views of two representative sperm shown on the bottom. Scale bar, 10 μm. **(F)** The resulting sperm maintained the genome integrity required to fertilize oocytes and support embryonic development by intracytoplasmic sperm injection assay. **(G)** Schematic model showing that FXR1 activates the translation of spermiogenic mRNAs in LS through phase separation and by interacting with components of translation machinery. Results shown are representative of three independent experiments. Significant differences by Student's *t* test are marked with asterisks: **P* < 0.05, ***P* < 0.01, ****P* < 0.001, n.s., not significant.



Materials and Methods

Mouse

All animal studies were approved by the Shanghai Institute of Biochemistry and Cell Biology (SIBCB) Institutional Animal Care and

Research Advisory Committee. All mouse lines used in this study were housed under specific pathogen-free conditions. *Fxr1*^{fl/fl} mice were obtained from D. L. Nelson's laboratory at Baylor College of Medicine (41), and *Stra8-Cre*

mice were obtained from M. Tong's laboratory at SIBCB (46). The germline-specific *Fxr1* knockout mice (referred to as *Fxr1*^{cko}) were generated by crossing *Fxr1*^{fl/fl} mice to *Stra8-Cre* mice (fig. S3A). The N-terminal

Flag-tagged *Fxr1* knock-in mice and *Fxr1*^{L351P} knock-in mutant mice were generated in the present study using the CRISPR-Cas9 system combined with semi-cloning technology by Genome Tagging Project (GTP) Center, SIBCB (figs. S6B and S10A). The germline-specific *Fxr1*^{L351P} knock-in mutant mice (referred to as *Fxr1*^{ckⁱ}) were generated with germline-specific *Fxr1* knockout mice by *Stra8-Cre*-mediated recombination (fig. S10B). Four *Fxr1*^{ck^o}, *Fxr1*^{ckⁱ}, or control *Fxr1*^{het} male mice from two litters were scored in the fertility study, with 8 to 12 months mating continued. Briefly, *Fxr1*^{ck^o}, *Fxr1*^{ckⁱ}, or control male mice at 3 months of age were mated to 2-month-old wild-type C57BL/6 female mice. The female was replaced with a 2-month-old female mouse every 2 to 3 months. Cages were monitored daily for the birth and number of newborn pups.

Cell culture and transfections

The mouse myoblast cell line C2C12 was purchased from the cell bank of the Type Culture Collection of Chinese Academy of Science (TCCAS). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS). Transfections were performed with Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. All DNA plasmids were cleared of endotoxin. For live-cell imaging experiment, 1.5 µg of plasmids of GFP-tagged FXR1 or mutants were transfected in a 35-mm dish with a glass-like polymer bottom (Cellvis, D35-20-1.5P). For the TRICK reporter assay, 1.0 µg of GFP-NLS-PCP expression vector, 1.0 µg of RFP-NLS-MCP expression vector, 1.5 µg of BFP-FXR1a or its mutant expression vector, and 1.5 µg of *Gpd2*-TRICK vector were cotransfected in C2C12 cells in a 35-mm dish. For the protein interaction-mapping experiment, 1.5 µg of expression vectors for each protein were cotransfected in 293T cells in each well of a six-well plate.

Plasmids and RNA oligonucleotides

The TRICK reporter pCMV-*Gpd2*-TRICK was derived from pMax-*pona*-12×TRICK-24×MS2SL (Addgene, catalog #64542), with the original promoter replaced with the CMV promoter and the *pona* gene replaced by a *Gpd2*-coding region. The promoters of pHsp83-NLS-HA-2×PCP-2×GFP (Addgene, catalog #71243) and pHsp83-NLS-HA-2×MCP-2×TagRFP-T (Addgene, catalog #71242) were replaced with the CMV promoter. pCMV-Myc-FXR1 and pCMV-EGFP-FXR1 vectors were constructed by insertion of the coding region of the indicated mouse *Fxr1* isoforms into pCMV-Myc and pCMV-EGFP-C1 (Clontech), respectively. pCMV-BFP-FXR1 were constructed by replacement of the EGFP of pCMV-EGFP-FXR1 with BFP. The pCMV-Flag-EIF4G isoform vectors were constructed by the insertion of the coding regions of mouse EIF4G isoforms into pCMV-3×Flag

(Clontech). The lentiviral vectors pLV-Ef1α-FXR1-IRES-Cyto IV-EGFP were generated as we recently described (47, 48). The coding regions of mouse *Fxr1* isoforms were cloned into the pET51b vector for their expression in *E. coli*. All FXR1 mutants were generated using the Site-Directed Mutagenesis Kit (New England Biolabs). All constructs were verified by DNA sequencing.

Protein expression and purification

Bacterial expression and purification of proteins were performed as we described recently (49). Recombinant FXR1 proteins were expressed in *E. coli* strain BL21. Cultures were grown in Luria broth medium at 37°C to an optical density at 600 nm (OD₆₀₀) of 0.6. After induction with 0.2 mM isopropyl-β-D-thiogalactopyranoside, the cells were grown at 18°C overnight, harvested by centrifugation at 4°C, and then lysed with a French press (JNBio) in buffer A (20 mM Tris-HCl, pH 8.0, 500 mM imidazole, and 1 mM phenylmethanesulfonyl fluoride) for Strep-FXR1-10×His proteins. Bacterial lysate was clarified by centrifugation at 16,000g for 1 hour. The supernatant of recombinant FXR1 protein was loaded onto a HisTrap HP column (Cytiva). After extensive washing with buffer A, the target protein was eluted in a buffer with a gradient of imidazole from 25 to 500 mM. Proteins were further loaded onto a HiTrap Q HP anion-exchange chromatography column (Cytiva). The target protein was eluted in a buffer with a gradient of NaCl from 100 mM to 1 M and dialyzed to the final buffer containing 500 mM NaCl and 20 mM HEPES, pH 7.5. Each protein was then labeled by Abberior STAR 488 (Sigma-Aldrich, catalog #61048) before the in vitro phase-separation assay.

TRICK reporter assay

TRICK reporter assay was performed as described previously (38). In brief, the *Gpd2* TRICK reporter was constructed by engineered PP7 binding sites (PBS) or MS2 binding sites (MBS), the binding motifs for PP7 coat protein (PCP) and MS2 coat protein (MCP) (39), into the coding and 3' untranslated regions of *Gpd2*, respectively (fig. S9A). Because translating ribosomes can displace RBP binding to the coding region, PCP targeted to the coding region through the inserted PBS motif thus can act as a biosensor to distinguish translated mRNAs from untranslated ones (39) (fig. S9A). For the *Gpd2* TRICK reporter assay, *Gpd2* TRICK reporter was cotransfected with the expression vectors for GFP-NLS-PCP, RFP-NLS-MCP, and BFP-FXR1a or its mutant in C2C12 cells.

Electrophoretic mobility shift assay

Before the electrophoresis, mRNA and protein were prepared as below. About 1000 nu-

cleotides of *Gpd2* mRNA was transcribed by T7 RNA polymerase (Promega, catalog #P1300) and purified using the MEGAclean kit (ThermoFisher, catalog #AM1908). Wild-type or mutant FXR1 proteins were expressed and purified as described above. *Gpd2* mRNA and FXR1 proteins were then mixed at the indicated concentrations in the binding buffer containing 100 mM NaCl and 20 mM Tris-HCl, pH 7.5, and incubated on ice for 20 min. In the meantime, 1% agarose gel (BIOWEST) containing 0.5× TBE buffer (45 mM Tris-Borate, 1 mM EDTA) supplied with 1× SYBR Gold (ThermoFisher, catalog #S11494) was prepared. The incubated mixture was then subjected to the gel and run in 0.5× TBE buffer at 5V/cm for ~40 min. After the electrophoresis, the gel was scanned using a multi-automatic fluorescence image analysis system (Tanon 5200 Multi). The gray intensity of free and FXR1-bound *Gpd2* mRNA was measured with ImageJ, and GraphPad Prism 9 software was used to fit the binding curve of each reaction with a nonlinear regression method based on the Hill equation

$$f = \left[\frac{1}{1 + \left(\frac{K_D}{[P_i]} \right)^n} \right]$$

where n is the Hill coefficient.

In vitro phase-separation assay

The phase-separation assay was performed on a homemade glass slide covered with tape-sealed coverslips. The purified FXR1 protein was assembled by diluting the protein from a high-salt-containing storage buffer to a varying concentration buffer containing 20 mM HEPES, pH 7.5, 75 to 500 mM NaCl, ± 10% polyethylene glycol (PEG) at the indicated protein concentration. For co-phase separation assay of FXR1 and mRNA, *Gpd2* mRNA (200, 1000, or 2000 nucleotides) was first transcribed by T7 RNA polymerase (Promega, catalog #P1300), with the incorporation of Cy3-labeled UTP (APEBio, catalog #B8330) into the reaction system, and purified with the MEGAclean kit (Thermo Fisher, catalog #AM1908). The FXR1 protein and *Gpd2* mRNA^{Cy3} were then mixed gently at the indicated concentrations and diluted from a high-salt-containing storage buffer to a physiological buffer containing 20 mM HEPES, pH 7.5, and 150 mM NaCl. After incubation for 5 min, images were captured with a FV3000 confocal microscope (Olympus) equipped with a superresolution module.

FRAP assay

FRAP experiments were performed on a white light laser confocal microscopes (Leica TCS SP8 WLL) equipped with immersion objectives of 63× magnification driven by the imaging software provided by the manufacturers. FRAP was performed by photobleaching the region of

interest with a laser at 488 nm (five repeats, 30% intensity, dwell time 1.5 s), and images were captured every 5 s. For FRAP in cultured cells, cells were plated on a 35-mm dish with a glass-like polymer bottom (Cellvis, catalog #D35-20-1.5P). Before image capture, phenol red-free DMEM (Gibco, catalog #21063-029) supplemented with 10% FBS was used in lieu of the cell culture. The images were also taken every 5 s after the region of interest was photobleached. The fluorescence intensity of the bleaching region was analyzed and determined using the manufacturer's software. The data were then normalized and plotted to the time point before the bleaching.

IP, immunoblotting, and immunofluorescent microscopy

For IP microscopy, mouse testicular tissue or the indicated cells were homogenized in ice-cold lysis buffer containing 50 mM Tris-HCl, pH 7.4, 1% Triton X-100, 150 mM NaCl, 5 mM EDTA, and protease inhibitor cocktail, and then the tissue or cell extracts were incubated with antibody-coupled Protein A/G beads (Sigma-Aldrich) according to the manufacturer's instructions. For immunoblotting, IP pellets or tissue or cell extracts were mixed with SDS loading buffer and subjected to the standard SDS-polyacrylamide gel electrophoresis and Western blotting procedures, and immunoblotting images were captured with a model 5200 chemiluminescent imaging system (Tanon). For immunofluorescent microscopy, enriched male germ cells or testicular sections were fixed with 4% paraformaldehyde, permeabilized with 0.5% Triton X-100 in PBS, and then incubated with primary antibodies and Alexa Fluor 488- or Cy3-conjugated secondary antibodies and/or with peanut agglutinin (PNA) conjugated with Alexa Fluor 488/647 (Molecular Probes), with nuclei counterstained with 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories). Images were captured using an FV3000 confocal microscope (Olympus), and immunostaining density was quantified using the FV31S-SW Viewer and ImageJ software.

Histological, TEM, and apoptosis assays

For histological assay, 5- μ m paraffin-embedded mouse testicular or epididymal sections were prepared as described previously (50) and then stained with hematoxylin and eosin (H&E). The stages of the seminiferous epithelium cycle and of spermatid development were determined as described previously (50). For TEM, testicular sections were fixed in 4% paraformaldehyde containing 0.05% glutaraldehyde in 0.1 M phosphate buffer and then postfixed in 1% osmium tetroxide. Dehydration was performed in ethanol, and the samples were embedded in Epon 812. Ultrathin sections were counterstained with uranyl ace-

tate and lead citrate and examined with a TEM (FEI Tecnai G2 Spirit). Apoptosis analysis was performed using the In Situ Cell Death Detection Kit, TMR red (Roche) according to the manufacturer's instructions.

Fluorescence in situ hybridization coupled with immunofluorescent assays

In situ hybridization was performed with a fluorescence in situ hybridization (FISH) kit (RiboBio, Guangzhou, China). The isolated male germ cells were fixed with 4% paraformaldehyde for 10 min, followed by three washes with 1 \times PBS and incubation with prechilled permeabilization buffer containing 0.5% Triton X-100 in PBS at 4°C for 5 min. After three washes with 1 \times PBS for 5 min, cells were incubated with prewarmed prehybridization solution at 37°C for 30 min. The *Gpd2*, *Akap3*, or *Tbpl1* antisense oligodeoxynucleotide probes (RiboBio) were then individually diluted in the hybridization solution and incubated with the cells at 37°C overnight in the dark. The next morning, the cells were washed with elution buffer I (4 \times SSC and 0.1% Tween-20) three times, elution buffer II (2 \times SSC) once, elution buffer III (1 \times SSC) once, and 1 \times PBS once for 5 min at 42°C. For an additional immunofluorescent assay, cells were rinsed with 1 \times PBS for 15 min and then subjected to immunofluorescent staining with anti-FXR1 and anti-DDX4 antibodies at 4°C overnight. After three washes with 1 \times PBS, cells were incubated with Alexa Fluor 488-labeled anti-mouse and 647-labeled anti-rabbit secondary antibody and DAPI in blocking buffer containing 10 mg/ml bovine serum albumin (BSA), 0.5% normal donkey serum (NDS), and 1 \times PBS at room temperature for 2 hours, and then washed with 1 \times PBS three times for 15 min each. Images were captured with an FV3000 confocal microscope (Olympus). Immunostaining density was quantified using the FV31S-SW Viewer and ImageJ software.

RNA isolation, RT-qPCR, Absolute qPCR, and RIP assays

The assays were performed as we described previously (47, 51). Briefly, total RNAs were extracted with RNAiso Plus (Takara) from the indicated cells, tissues, or polysome fractions, and reverse transcription (RT)-qPCR was performed using the PrimeScript RT and SYBR Green kits (Takara) on a QuantStudio 3 system (Applied Biosystems) according to the manufacturer's instructions. For absolute quantification of specific mRNA, a standard curve method was adopted. Plasmid containing the *Fxr1*-coding region was linearized. The molecular number of plasmid was calculated according to the following formula: copy number (copies/ μ L) = 6.02×10^{23} (copies/mol) \times plasmid concentration (g/ μ L)/plasmid molecular weight (g/mol). The plasmid was then diluted into

eight gradients of $\sim 10^{10}$ to 10^3 copies/ μ L. qPCR of standards and extracted total RNAs were performed using the PrimeScript RT and SYBR Green kits (Takara) on a QuantStudio 3 system (Applied Biosystems). The standard curve was plotted, and the absolute amount of specific mRNA in total RNA was extrapolated according to the standard curve. For the RIP assay, cells or tissues were homogenized in lysis buffer containing 100 mM KCl, 5 mM MgCl₂, 10 mM HEPES, 0.5% and NP-40 containing 10 U/ml RNase inhibitor (Takara) and a protease inhibitor cocktail (Roche), and extracts were incubated with anti-FXR1 antibody or control immunoglobulin G (IgG)-coupled Protein A/G for 4 hours at 4°C. After the beads were stringently washed with washing buffer containing 50 mM Tris-HCl, pH 7.4, 0.1% Triton X-100, 500 mM NaCl, 5 mM EDTA, protease inhibitor cocktail, and RNase inhibitor, RNAs were extracted from the beads with RNAiso Plus (Takara) and then subjected to the RT-qPCR assay. The primer sequences for RT-qPCR are provided in data S6. The results were first analyzed as relative mRNA levels of the cycle threshold (CT) values and then converted as fold changes.

Sucrose gradient analysis

Sucrose gradient analysis of testicular extracts was performed as we described recently (48). In brief, mouse testicular tissues were homogenized in extraction buffer containing 100 mM NaCl, 50 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1% Triton X-100, and 100 μ g/ml cycloheximide, and then subjected to centrifugation at 13,000g for 2 min to remove nuclear pellets. The obtained extracts were layered over the 10 to 60% sucrose gradients and centrifuged at 38,000 rpm on the SW41 rotor for 2 hours at 4°C. For puromycin treatment, tissues were homogenized in extraction buffer and then incubated at 30°C for 30 min with 500 μ M puromycin. Twelve fractions were collected by a Piston Gradient Fractionator (Biocomp, Fredericton, Canada), and each fraction was analyzed by Western blot and RT-qPCR for the indicated proteins and RNAs.

Spermatogenic cell isolation

RS and elongating spermatids were isolated from adult C57BL/6 mouse testes using the STA-PUT method described previously (52). Briefly, the testicular tissues were dissected from the albuginea membrane and then washed three times with DMEM (Gibco). Tissues were then digested with 1 mg/ml collagenase IV (Sigma-Aldrich) at 37°C for 10 min and centrifuged at 500g. The pellet was subsequently digested with 0.25% trypsin (Gibco) containing 1 mg/ml DNase I (Sigma Aldrich) at 37°C for 10 min and centrifuged at 500g. The pellet was then resuspended in DMEM and filtered to prepare a single-cell suspension that was

further injected into the sample chamber of the STA-PUT apparatus (ProScience, Canada); a linear BSA gradient (2% BSA and 4% BSA in DMEM) was followed. After harvesting the cell fractions within 2.5 hours of sedimentation, RS and elongating spermatids were determined on the basis of their morphological characteristics and DAPI and PNA staining pattern under the microscope.

Quantification of FXR1 in spermatids

Cytoplasmic concentrations of FXR1 in spermatids were determined by quantitative Western blotting using recombinant purified FXR1d protein as a reference. After washing twice with PBS, enriched spermatids were lysed with lysis buffer containing 50 mM Tris-HCl, pH 7.4, 1% Triton X-100, 150 mM NaCl, 5 mM EDTA, and protease inhibitor cocktail. Cell lysates from defined cell counts were subjected to immunoblotting with anti-FXR1 antibody (Novus, catalog #NBP2-22246) using various amounts of purified FXR1d protein for plotting the calibration curve. The amount of FXR1 protein per cell was calculated by the calibration curve and cell counts.

Lentivirus packaging and testis transduction

Lentivirus packaging and testis transduction were performed as we recently described (57). Briefly, lentiviral vectors and cognate packaging plasmids were cotransfected in 293T cells and packaged into pseudoviruses following standard procedures. High-titer lentivirus ($>10^8$ transduction units/ml) was prepared through ultracentrifugation, and ~ 10 μ l of high-titer lentivirus was injected into the seminiferous tubule through the efferent duct using a sharp glass capillary with a tip diameter of 50 μ m. One week after transduction of lentiviral vectors, the testes were harvested for examining the transduced spermatids, and epididymides were harvested 2 weeks after transduction for examining the sperm progressed from transduced spermatids.

Proteomics and protein profiling analysis

To identify FXR1-binding proteins in mouse testes, testicular tissues were homogenized in lysis buffer containing 50 mM Tris-HCl, pH 7.4, 1% Triton X-100, 150 mM NaCl, 5 mM EDTA, and 1 \times Complete Mini Proteinase Inhibitor cocktail (Roche) and then centrifuged for 10 min at 10,000g and 4°C. Anti-FXR1 or anti-IgG coupled Protein A/G beads were added to the testicular extract and incubated for 3 hours at 4°C. After washing the beads with washing buffer containing 50 mM Tris-HCl, pH 7.4, 0.1% Triton X-100, 500 mM NaCl, 5 mM EDTA, and proteinase inhibitor cocktail, IP pellets were subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis with standard protocols for proteomic analysis. To identify the proteins in polysome fractions

of mouse testes, polysome fractions prepared through sucrose gradients were subjected to LC-MS/MS analysis with standard protocols for proteomic analysis. The protein intensities of all identified proteins were normalized by total intensity and then used for differential protein expression analysis using DEqMS (53). The proteins in the differential expression analysis were required to have values greater than zero in two replicates for both IP and control groups to calculate the fold change and false discovery rate (FDR) values. For proteins with only one value greater than zero in either group, the protein intensity equal to zero was replaced by the minimum value of this sample to calculate the fold change without FDR.

RNA-seq and polysome profiling

Mouse testicular tissues were homogenized and extracted with RNAiso Plus (Takara) according to the manufacturer's protocol. For polysome profiling, testicular extracts were subjected to sucrose gradient analysis, and the obtained polysomes were subjected to RNA extraction with standard protocols. The extracted RNA samples were then subjected to RNA-seq using Illumina sequencing technology. Clean reads were first mapped to ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), and those unmapped reads were then mapped to the mouse genome (mm10) with gene annotation GENECODE V23 using STAR (v2.5.3a) (54) under the default parameters except for the setting “-outFilterMatchNmin 40.” The FeatureCounts program from Rsubreads package (55) was used to count mapped reads for all annotated genes for all samples. The reads counts of all genes in different sample tables were used for differential gene expression analysis with DESeq2 (56). The cutoff for differentially expressed genes was set as FDR < 0.05 and fold change > 1.5.

The polysome profiling sequencing data were preprocessed in the same manner as RNA-seq data, and the read counts of two groups were used for differential translational efficiency analysis using the R package Xtail (1.15) (57). The cutoff for call differential translational efficiency genes was set as FDR < 0.01 and fold change > 1.5.

eCLIP-seq and sequencing data analysis

eCLIP-seq was performed as previously described (21, 58). In brief, testicular lysates prepared from adult mouse testes were subjected to UV cross-linked treatment (254 nm, 400 mJ/cm²) three times, and then a limited RNase I digestion was performed before FXR1 immunoprecipitation. After dephosphorylation, a barcoded RNA adapter was ligated to the 3' end of anti-FXR1 IPed RNAs. Samples were then run on standard protein gels, transferred to nitrocellulose membranes, and RNA-protein bands ranging from 75 to 150 kDa were cut

for extracting RNAs. The obtained RNAs were then reverse transcribed, and a second DNA adapter was ligated to the cDNA fragment 3' end. After cleanup, an aliquot of each sample was first subjected to qPCR (to identify the proper number of PCR cycles) and then the remaining cDNA samples were PCR amplified and size selected by agarose gel electrophoresis. Samples were subjected to RNA-seq analysis using Illumina sequencing technology. Briefly, eCLIP-seq libraries with distinct inline barcodes were demultiplexed, and the adapters in reads were trimmed with cutadapt (v1.18). The reads ≥ 18 nucleotides were discarded. Reads were first mapped to rRNAs and tRNAs, and those unmapped reads were then mapped to the mouse genome (mm10) with gene annotation GENECODE V23 using STAR (54). For each sample, uniquely mapped reads for annotated genes using the FeatureCounts program (55). The read distribution across the genome was calculated using the read distribution.py module in the RSeQC program (59). The FXR1 targeting genes over input controls were identified using DESeq2 (56) with the cutoffs set as FDR < 0.05 and fold change > 1.5.

GO analysis

GO term enrichment analysis was performed using the R package clusterProfiler (60) with a Benjamini-Hochberg multiple testing adjustment under an FDR cutoff of 0.05. Enriched GO terms of biological process were plotted for visualization.

Statistical analysis and data visualization

The statistical tests were performed using the ggpubr package in R. The figures from computational analysis were made using the ggplot2 (61) and ComplexHeatmap (62) packages, and then the Venn diagrams were made using the eulerr package in R.

REFERENCES AND NOTES

- M. L. Meistrich, R. A. Hess, in *Spermatogenesis: Methods and Protocols*, D. T. Carrell, K. I. Aston, Eds. (Humana, 2013), pp. 299–307.
- P. Sassone-Corsi, Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science* **296**, 2176–2178 (2002). doi: 10.1126/science.1070963; pmid: 12077401
- R. K. Idler, W. Yan, Control of messenger RNA fate by RNA-binding proteins: An emphasis on mammalian spermatogenesis. *J. Androl.* **33**, 309–337 (2012). doi: 10.2164/jandrol.111.014167; pmid: 21757510
- S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017). doi: 10.1038/nrm.2017.7; pmid: 28225081
- S. Alberti, A. Gladfelder, T. Mittag, Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019). doi: 10.1016/j.cell.2018.12.035; pmid: 30682370
- S. C. Weber, C. P. Brangwynne, Getting RNA and protein in phase. *Cell* **149**, 1188–1191 (2012). doi: 10.1016/j.cell.2012.05.022; pmid: 22682242
- Y. Lin, D. S. W. Protter, M. K. Rosen, R. Parker, Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015). doi: 10.1016/j.molcel.2015.08.018; pmid: 26412307

8. A. Mollieux *et al.*, Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015). doi: 10.1016/j.cell.2015.09.015; pmid: 26406374
9. A. Patel *et al.*, A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015). doi: 10.1016/j.cell.2015.07.047; pmid: 26317470
10. B. Van Treec, R. Parker, Emerging roles for intermolecular RNA-RNA interactions in RNP assemblies. *Cell* **174**, 791–802 (2018). doi: 10.1016/j.cell.2018.07.023; pmid: 30096311
11. T. Lehtiniemi, N. Kotaja, Germ granule-mediated RNA regulation in male germ cells. *Reproduction* **155**, R77–R91 (2018). doi: 10.1530/REP-17-0356; pmid: 29038333
12. O. Meikar *et al.*, An atlas of chromatoid body components. *RNA* **20**, 483–495 (2014). doi: 10.1261/rna.043729.113; pmid: 24554440
13. L. L. Kirkpatrick, K. A. Mcllwain, D. L. Nelson, Comparative genomic sequence analysis of the FXR gene family: FMRL, FXR1, and FXR2. *Genomics* **78**, 169–177 (2001). doi: 10.1006/geno.2001.6667; pmid: 11735223
14. A. Banerjee, M. F. Ifrim, A. N. Valdez, N. Raj, G. J. Bassell, Aberrant RNA translation in fragile X syndrome: From FMRP mechanisms to emerging therapeutic strategies. *Brain Res.* **1693** (Pt A), 24–36 (2018). doi: 10.1016/j.brainres.2018.04.008; pmid: 29653083
15. A. L. Bhakar, G. Dolen, M. F. Bear, The pathophysiology of fragile X (and what it teaches us about synapses). *Annu. Rev. Neurosci.* **35**, 417–443 (2012). doi: 10.1146/annurev-neuro-060909-153138; pmid: 22483044
16. R. Mazroui *et al.*, Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Hum. Mol. Genet.* **11**, 3007–3017 (2002). doi: 10.1093/hmg/11.24.3007; pmid: 12417522
17. N. E. Patzlaff, K. M. Nemes, S. G. Malone, Y. Li, X. Zhao, Fragile X related protein 1 (FXR1P) regulates proliferation of adult neural stem cells. *Hum. Mol. Genet.* **26**, 1340–1352 (2017). doi: 10.1093/hmg/ddx034; pmid: 28204491
18. E. J. Greenblatt, A. C. Spradling, Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science* **361**, 709–712 (2018). doi: 10.1126/science.aas9963; pmid: 30115809
19. J. R. Warner, P. M. Knopf, The discovery of polyribosomes. *Trends Biochem. Sci.* **27**, 376–380 (2002). doi: 10.1016/S0968-0004(02)02126-6; pmid: 12114027
20. J. Zhong, A. H. F. M. Peters, K. Lee, R. E. Braun, A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. *Nat. Genet.* **22**, 171–174 (1999). doi: 10.1038/9684; pmid: 10369260
21. E. L. Van Nostrand *et al.*, Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016). doi: 10.1038/nmeth.3810; pmid: 27018577
22. J. C. Darnell *et al.*, FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011). doi: 10.1016/j.cell.2011.06.013; pmid: 21784246
23. J. M. Weitzel *et al.*, Testis-specific expression of rat mitochondrial glycerol-3-phosphate dehydrogenase in haploid male germ cells. *Biol. Reprod.* **68**, 699–707 (2003). doi: 10.1095/biolreprod.102.008540; pmid: 12533437
24. K. Xu, L. Yang, L. Zhang, H. Qi, Lack of AKAP3 disrupts integrity of the subcellular structure and proteome of mouse sperm and causes male sterility. *Development* **147**, dev181057 (2020). doi: 10.1242/dev.181057; pmid: 31969357
25. N. Nakamura, H. Shibata, D. A. O'Brien, C. Mori, E. M. Eddy, Spermatogenic cell-specific type 1 hexokinase is the predominant hexokinase in sperm. *Mol. Reprod. Dev.* **75**, 632–640 (2008). doi: 10.1002/mrd.20791; pmid: 17924400
26. K. A. Burton, B. Treash-Osio, C. H. Muller, E. L. Dunphy, G. S. McKnight, Deletion of type IIa/pa regulatory subunit delocalizes protein kinase A in mouse spermatocytes without affecting motility or fertilization. *J. Biol. Chem.* **274**, 24131–24136 (1999). doi: 10.1074/jbc.274.34.24131; pmid: 10446185
27. G. Herrada, D. J. Wolgemuth, The mouse transcription factor Stat4 is expressed in haploid male germ cells and is present in the perinuclear theca of spermatozoa. *J. Cell Sci.* **110**, 1543–1553 (1997). doi: 10.1242/jcs.110.14.1543; pmid: 9247188
28. J. Swiston, A. Hubberstey, G. Yu, D. Young, Differential expression of CAP and CAP2 in adult rat tissues. *Gene* **165**, 273–277 (1995). doi: 10.1016/0378-1119(95)00522-8; pmid: 8522189
29. H. Hess, H. Heid, W. W. Franke, Molecular characterization of mammalian cylicin, a basic protein of the sperm head cytoskeleton. *J. Cell Biol.* **122**, 1043–1052 (1993). doi: 10.1083/jcb.122.5.1043; pmid: 8354692
30. M. W. Hentze, eIF4G: A multipurpose ribosome adapter? *Science* **275**, 500–501 (1997). doi: 10.1126/science.275.5299.500; pmid: 9019810
31. W. C. Merrick, eIF4F: A retrospective. *J. Biol. Chem.* **290**, 24091–24099 (2015). doi: 10.1074/jbc.R115.675280; pmid: 26324716
32. F. Sun, K. Palmer, M. A. Handel, Mutation of Eif4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes. *Development* **137**, 1699–1707 (2010). doi: 10.1242/dev.043125; pmid: 20430745
33. M. Kimura, K. Ishida, S. Kashiwabara, T. Baba, Characterization of two cytoplasmic poly(A)-binding proteins, PABPC1 and PABPC2, in mouse spermatogenic cells. *Biol. Reprod.* **80**, 545–554 (2009). doi: 10.1095/biolreprod.108.072553; pmid: 19020299
34. H. Imataka, A. Gradi, N. Sonenberg, A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* **17**, 7480–7489 (1998). doi: 10.1093/emboj/17.24.7480; pmid: 9857202
35. T. H. Kim *et al.*, Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science* **365**, 825–829 (2019). doi: 10.1126/science.aax4240; pmid: 31439799
36. L. L. Kirkpatrick, K. A. Mcllwain, D. L. Nelson, Alternative splicing in the murine and human FXR1 genes. *Genomics* **59**, 193–202 (1999). doi: 10.1006/geno.1999.5868; pmid: 10409431
37. D. T. McSwiggan, M. Mir, X. Darzacq, R. Tjian, Evaluating phase separation in live cells: Diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019). doi: 10.1101/gad.331520.119; pmid: 31594803
38. J. M. Halstead *et al.*, Translation. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science* **347**, 1367–1671 (2015). doi: 10.1126/science.aaa3380; pmid: 25792328
39. S. Hocine, P. Raymond, D. Zenklusen, J. A. Chao, R. H. Singer, Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat. Methods* **10**, 119–121 (2013). doi: 10.1038/nmeth.2305; pmid: 23263691
40. L.-T. Gou *et al.*, Ubiquitination-deficient mutations in human Piwi cause male infertility by impairing histone-to-protamine exchange during spermiogenesis. *Cell* **169**, 1090–1104.e13 (2017). doi: 10.1016/j.cell.2017.04.034; pmid: 28552346
41. E. J. Mientjes *et al.*, Fxr1 knockout mice show a striated muscle phenotype: Implications for Fxr1p function in vivo. *Hum. Mol. Genet.* **13**, 1291–1302 (2004). doi: 10.1093/hmg/ddh150; pmid: 15128702
42. J. A. Smith *et al.*, FXR1 splicing is important for muscle development and biomolecular condensates in muscle cells. *J. Cell Biol.* **219**, e201911129 (2020). doi: 10.1083/jcb.201911129; pmid: 32328638
43. M. C. Estañ *et al.*, Recessive mutations in muscle-specific isoforms of FXR1 cause congenital multi-minicore myopathy. *Nat. Commun.* **10**, 797 (2019). doi: 10.1038/s41467-019-08548-9; pmid: 30770808
44. S. Vasudevan, J. A. Steitz, AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell* **128**, 1105–1118 (2007). doi: 10.1016/j.cell.2007.01.038; pmid: 17382880
45. J. George *et al.*, RNA-binding protein FXR1 drives cMYC translation by recruiting eIF4F complex to the translation start site. *Cell Rep.* **37**, 109934 (2021). doi: 10.1016/j.celrep.2021.109934; pmid: 34731628
46. Z. Lin *et al.*, Mettl3-/Mettl14-mediated mRNA N⁶-methyladenosine modulates murine spermatogenesis. *Cell Res.* **27**, 1216–1230 (2017). doi: 10.1038/cr.2017.117; pmid: 28914256
47. L.-T. Gou *et al.*, Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* **24**, 680–700 (2014). doi: 10.1038/cr.2014.41; pmid: 24787618
48. P. Dai *et al.*, A translation-activating function of MIWI/piRNA during mouse spermiogenesis. *Cell* **179**, 1566–1581.e16 (2019). doi: 10.1016/j.cell.2019.11.022; pmid: 31835033
49. Y. Zhang *et al.*, Structural insights into the sequence-specific recognition of Piwi by *Drosophila* Papi. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3374–3379 (2018). doi: 10.1073/pnas.1717116115; pmid: 29531043
50. L. D. Russell, T. Weiss, J. C. Goh, J. L. Curl, The effect of submandibular gland removal on testicular and epididymal parameters. *Tissue Cell* **22**, 263–268 (1990). doi: 10.1016/0040-8166(90)90001-P; pmid: 2237906
51. S. Zhao *et al.*, piRNA-triggered MIWI ubiquitination and removal by APC/C in late spermatogenesis. *Dev. Cell* **24**, 13–25 (2013). doi: 10.1016/j.devcel.2012.12.006; pmid: 23328397
52. S. Bai *et al.*, Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. *Development* **145**, dev164855 (2018). doi: 10.1242/dev.164855; pmid: 29866902
53. Y. Zhu *et al.*, DEqMS: A Method for Accurate Variance Estimation in Differential Protein Expression Analysis. *Mol. Cell. Proteomics* **19**, 1047–1057 (2020). doi: 10.1074/mcp.T1R19.001646; pmid: 32205417
54. A. Dobin *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bt563; pmid: 23104886
55. Y. Liao, G. K. Smyth, W. Shi, The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47 (2019). doi: 10.1093/nar/gkz114; pmid: 30783653
56. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
57. Z. Xiao, Q. Zou, Y. Liu, X. Yang, Genome-wide assessment of differential translations with ribosome profiling data. *Nat. Commun.* **7**, 11194 (2016). doi: 10.1038/ncomms11194; pmid: 27041671
58. Q. Xu *et al.*, Enhanced crosslinking immunoprecipitation (eCLIP) method for efficient identification of protein-bound RNA in mouse testis. *J. Vis. Exp.* **147**, e59681 (2019). doi: 10.3791/59681; pmid: 31132071
59. L. Wang, S. Wang, W. Li, RSeQC: Quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012). doi: 10.1093/bioinformatics/bts356; pmid: 22743226
60. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012). doi: 10.1089/omi.2011.0118; pmid: 22455463
61. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
62. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016). doi: 10.1093/bioinformatics/btw313; pmid: 27207943

ACKNOWLEDGMENTS

We thank Y. Zheng from UT Southwestern and members of the M.-F.L. laboratory for helpful comments; M. Tong from SIBCB of the Chinese Academy of Sciences for the *Stras8-Cre* knock-in mice; and J. Sha and K. Zheng from Nanjing Medical University and L. Ge and J. He from the Cell Biology Core Facility at SIBCB for experimental assistance. **Funding:** This work was supported by the National Key R&D Program of China (grants 2017YFA0504400 and 2021YFC2700200), the National Natural Science Foundation of China (grants 91940305, 31830109, 31821004, 32171186, 91940302, 31961133022, 32170815, and 91640201), the Chinese Academy of Sciences (Strategic Priority Research Program grant XDB19010203), the Science and Technology Commission of Shanghai Municipality (grants 2017SHZDX01, 19JC1410200, 17JC1420100, 21PJ1413800, and 21ZR1470200), the Innovative Research Team of High-level Local Universities in Shanghai (grant SHSMU-ZDCX20210902), the National Postdoctoral Program for Innovative Talent (grants BX20180331 and BX20190081), the China Postdoctoral Science Foundation (grants 2018M642018 and 2020M670986), the Open Fund of State Key Laboratory of Reproductive Medicine (grant SKLRM-K202101), and the Foundation of Key Laboratory of Gene Engineering of the Ministry of Education. **Author contributions:** M.-F.L. supervised the overall work and Y.H. was responsible for phase separation studies. J.-Y.K., Z.W., D.P., Y.Z., A.Z., Y.-C.W., Y.C., X.Z., P.-C.K., J.G., M.-M.H., and X.G. performed the experiments. Z.Y., Y.Z., and J.-Y.K. analyzed the MS and high-throughput sequencing data. D.L.N. and R.Z. provided *Fxr1^{fllox/fllox}* mice. Q.L., Y.-Y.W., J.L., and J.-Y.K. generated *Fxr1^{L351P}* and Flag-tagged FXR1 mice. J.G., P.-C.K., and L.-T.G. performed intracytoplasmic sperm injection. M.-F.L., Y.H., J.-Y.K., and Z.W. wrote the manuscript with helpful input from X.-D.F., J.L., D.L., B.L., H.-J.S., L.-T.G., and D.L.N. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials**

availability: All data presented are available in the main text or the supplementary materials. The accession number for high-throughput sequencing reported herein is PRJNA719255, deposited in BioProject of the National Center for Biotechnology Information (NCBI). The *Fxr1* conditional knockout mice used in the study are available to researchers who agree to the terms of the standard materials transfer agreement provided by Baylor College of Medicine. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American

Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj6647
Figs. S1 to S10
Reference (63)

Movies S1 to S5
Data S1 to S6
MDAR Reproducibility Checklist

Submitted 25 May 2021; resubmitted 23 March 2022
Accepted 12 July 2022
[10.1126/science.abj6647](https://doi.org/10.1126/science.abj6647)

Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE



Since 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

cut here ✂

“As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all.”

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science

Federal Tax ID Number: 53-0196568

Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at www.aaas.org/1848Society.



Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:

AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLES

CORONAVIRUS

Broadly neutralizing antibodies target the coronavirus fusion peptide

Cherrelle Dacon^{1†}, Courtney Tucker^{1,2†}, Linghang Peng^{3†}, Chang-Chun D. Lee^{4†}, Ting-Hui Lin⁴, Meng Yuan⁴, Yu Cong⁵, Lingshu Wang⁶, Lauren Purser¹, Jazmean K. Williams⁷, Chul-Woo Pyo⁸, Ivan Kosik⁹, Zhe Hu⁹, Ming Zhao¹⁰, Divya Mohan¹, Andrew J. R. Cooper¹, Mary Peterson¹¹, Jeff Skinner¹¹, Saurabh Dixit⁵, Erin Kollins⁵, Louis Huzella⁵, Donna Perry⁵, Russell Byrum⁵, Sanae Lembirik⁵, David Drawbaugh⁵, Brett Eaton⁵, Yi Zhang⁶, Eun Sung Yang⁶, Man Chen⁶, Kwanyee Leung⁶, Rona S. Weinberg¹², Amarendra Pegu⁶, Daniel E. Geraghty⁸, Edgar Davidson⁷, Iyadh Douagi¹³, Susan Moir¹⁴, Jonathan W. Yewdell⁹, Connie Schmaljohn⁵, Peter D. Crompton¹¹, Michael R. Holbrook⁵, David Nemazee³, John R. Mascola⁶, Ian A. Wilson^{4,15}, Joshua Tan^{1*}

The potential for future coronavirus outbreaks highlights the need to broadly target this group of pathogens. We used an epitope-agnostic approach to identify six monoclonal antibodies that bind to spike proteins from all seven human-infecting coronaviruses. All six antibodies target the conserved fusion peptide region adjacent to the S2' cleavage site. COV44-62 and COV44-79 broadly neutralize alpha- and betacoronaviruses, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Omicron subvariants BA.2 and BA.4/5, albeit with lower potency than receptor binding domain-specific antibodies. In crystal structures of COV44-62 and COV44-79 antigen-binding fragments with the SARS-CoV-2 fusion peptide, the fusion peptide epitope adopts a helical structure and includes the arginine residue at the S2' cleavage site. COV44-79 limited disease caused by SARS-CoV-2 in a Syrian hamster model. These findings highlight the fusion peptide as a candidate epitope for next-generation coronavirus vaccine development.

Coronaviruses consist of four genera that infect birds and mammals (1). Seven coronaviruses are known to cause human disease: the alphacoronaviruses HCoV-229E (human coronavirus 229E) and HCoV-NL63 (human coronavirus NL63), as well as the betacoronaviruses HCoV-OC43 (human coronavirus OC43), HCoV-HKU1 (human coronavirus HKU1), SARS-CoV (severe acute respiratory syndrome coronavirus), MERS-CoV (Middle East respiratory syndrome coronavirus), and SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). Whereas the first four coronaviruses generally cause mild disease, the latter three have brought about serious outbreaks in recent years. In particular, the ongoing COVID-19 pandemic caused by SARS-CoV-2 has resulted in more than 6 million deaths since the first cases were identified in 2019 (2). The currently dominant SARS-CoV-2 Omicron BA.2, BA.2.12.1, BA.4, and BA.5 subvariants are at

least partially resistant to most available vaccines and antibody therapeutics (3–6). Furthermore, two coronaviruses previously linked only to animal infection were recently detected in individuals with flu-like symptoms (7, 8). These developments highlight the importance of targeting conserved and functionally essential sites on coronaviruses.

Coronavirus infection is a multistep process that involves enzymatic cleavage and rearrangement of the surface spike protein (9). The SARS-CoV-2 spike contains two cleavage sites: a furin cleavage site at the boundary of the S1 and S2 subunits, and an S2' site that is conserved in coronaviruses. The spike protein is thought to be cleaved at the S1-S2 site during virus assembly, leaving the S1 and S2 subunits non-covalently linked. During entry, the SARS-CoV-2 spike protein uses the receptor binding domain (RBD) on the S1 subunit to engage angiotensin-converting enzyme 2 (ACE2) on target cells. After receptor binding, the S1 subunit is shed

and the S2' site is cleaved by the membrane enzyme transmembrane serine protease 2 (TMPRSS2) or endosomal cathepsins (1), leading to insertion of the fusion peptide into the cell membrane and viral fusion.

Much of the protection provided by COVID-19 vaccines arises from neutralizing antibodies that target the RBD (10). Likewise, all currently available therapeutic monoclonal antibodies (mAbs) target this domain (3). However, spike elements that participate in the subsequent stages of infection involve the more complex S2 fusion machinery with many moving parts, and these elements are more conserved than the RBD, which so far has been capable of retaining or even increasing binding to ACE2 despite a variety of mutations (11). Therefore, these sites are worth exploring as targets for novel COVID-19 vaccines and therapeutics that retain efficacy against new variants and protect against a wider range of coronaviruses. Progress in this direction has started with recent studies identifying several mAbs that target the conserved stem helix (12–16) and using unbiased approaches to screen for mAbs of interest (17–20). In this study, we carried out a large-scale survey of the binding landscape of broadly reactive mAbs against coronaviruses.

Identification of broadly reactive mAbs from COVID-19-convalescent donors

To identify individuals who are likely to harbor B cells that produce broadly reactive mAbs, we used a multiplex bead-based assay to examine plasma samples of 142 donors from a previously described cohort of COVID-19-convalescent individuals (20). We assessed plasma immunoglobulin G (IgG) reactivity toward spike glycoproteins of the seven human coronaviruses: SARS-CoV-2 (Wuhan-Hu-1 strain), SARS-CoV, MERS-CoV, HCoV-HKU1, HCoV-OC43, HCoV-NL63, and HCoV-229E. Nineteen donors were selected for mAb isolation and characterization on the basis of plasma IgG reactivity to the spike proteins of SARS-CoV-2 and at least two other betacoronaviruses (fig. S1A).

We next investigated human IgG⁺ memory B cells (MBCs) from the selected donors by using a two-stage screen to prioritize isolating mAbs with the greatest possible breadth of reactivity. First, we screened supernatants from 673,671 stimulated IgG⁺ B cells for binding

¹Antibody Biology Unit, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA. ²Department of Biology, The Catholic University of America, Washington, DC 20064, USA. ³Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA. ⁴Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. ⁵Integrated Research Facility, Division of Clinical Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, MD 21702, USA. ⁶Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ⁷Integral Molecular, Philadelphia, PA 19104, USA. ⁸Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. ⁹Cellular Biology Section, Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ¹⁰Protein Chemistry Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA. ¹¹Malaria Infection Biology and Immunity Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA. ¹²New York Blood Center, Lindsley F. Kimball Research Institute, New York, NY 10065, USA. ¹³Flow Cytometry Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ¹⁴B Cell Immunology Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ¹⁵The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

*Corresponding author. Email: tanj4@nih.gov

†These authors contributed equally to this work.

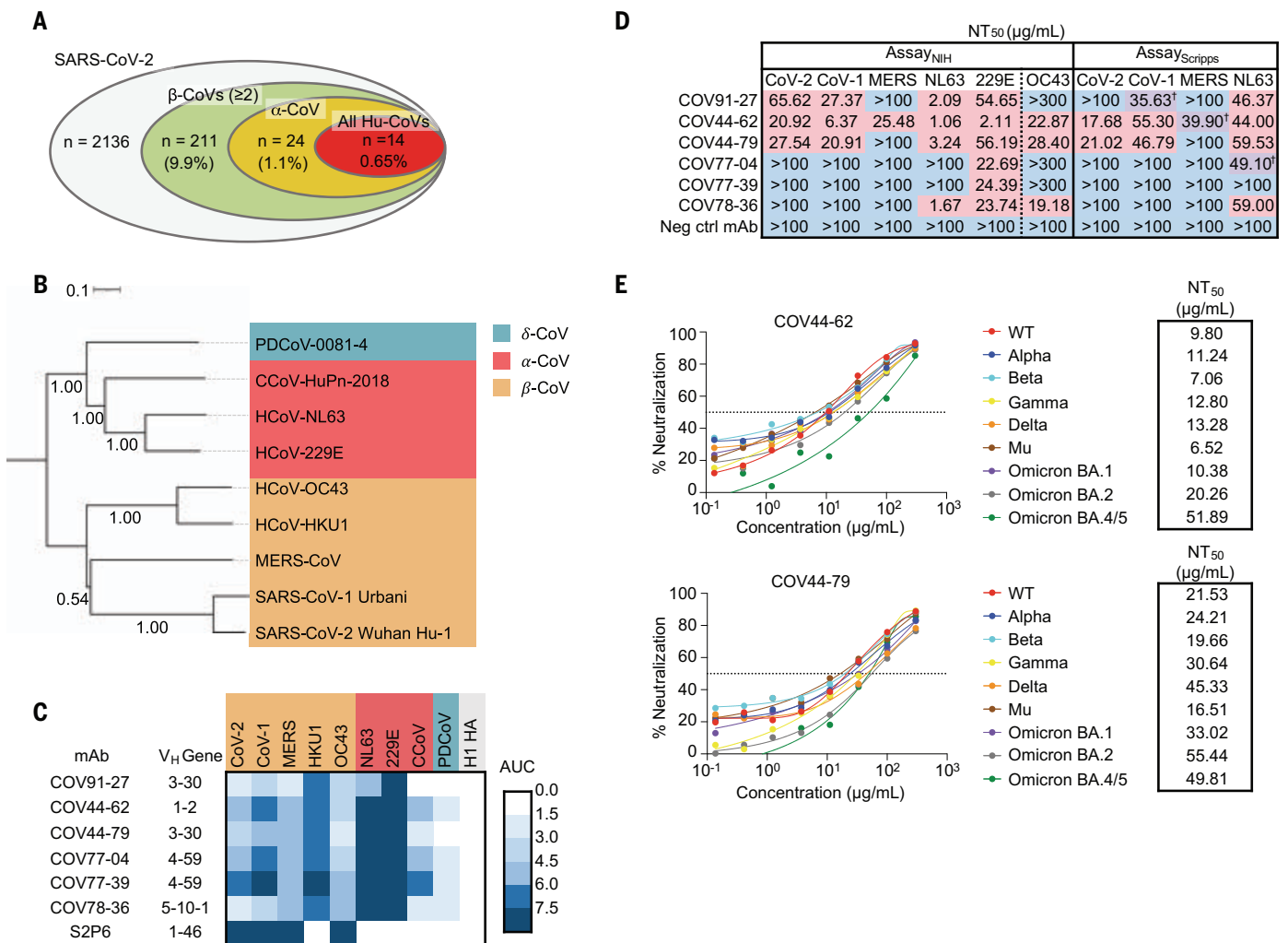


Fig. 1. Broadly neutralizing antibodies target coronaviruses associated with human disease. (A) Analysis of the frequency of MBCs expressing broadly reactive antibodies from 19 donors. Values in parentheses represent the percentage of SARS-CoV-2-reactive supernatants that also bind the specified subsets of nonsarbecovirus coronavirus spikes. A total of 10,356 MBC culture supernatants (50 to 100 B cells per well) were screened. *n*, number of MBC culture wells. (B) Phylogenetic relationships across the coronavirus spike proteins targeted by the broadly reactive mAbs were inferred by the neighbor-joining method in MEGA11 using full-length amino acid sequences of CoV spike proteins. Branch lengths are drawn to scale, and bootstrap values from 500 samplings are shown on the branches. The scale bar represents the number of amino acid substitutions per site. (C) Heatmap representing the binding of broadly reactive mAbs to spike proteins from coronaviruses across the *alpha*-, *beta*-, and *delta* coronavirus genera. H1 hemagglutinin was included as a negative

control for mAb binding experiments, and area under the curve (AUC) values for each antigen are shown after subtraction with values for the negative control antigen CD4. (D) Values represent antibody titer at 50% neutralization (NT₅₀) against SARS-CoV-2 (Wuhan-Hu-1 strain), SARS-CoV (indicated as “CoV-1” in the figure), MERS-CoV, HCoV-NL63, and HCoV-229E envelope-pseudotyped lentivirus, as well as authentic HCoV-OC43. For Assay_{Scripps}, values are average of two experiments. For values with the † symbol, one NT₅₀ was determinable and one was not (i.e., >100 μg/ml), and the determinable NT₅₀ is shown. Negative control mAbs were anti-CoV-2 RBD CV503 (for OC43 assay) (20), anti-influenza HA CR9114 (for Assay_{NIH} except OC43) (42), and anti-dengue DEN3 (for Assay_{Scripps}) (43). NT₅₀ values were calculated using the dose-response-inhibition model with five-parameter Hill slope equation in GraphPad Prism 9. (E) Neutralization of SARS-CoV-2 variants of concern (pseudovirus, Assay_{Scripps}) by COV44-62 and COV44-79.

to the coronavirus spike panel used in the plasma screen. Supernatants from only 2% (*n* = 211) of the MBC culture wells met our criteria for broad reactivity by binding to at least three betacoronavirus spike proteins (Fig. 1A). Next, we developed an optofluidics assay to isolate individual MBCs of interest using the Berkeley Lights Beacon system (fig. S1B). Candidate MBCs identified in the supernatant screen were sorted individually

into nanoliter-volume pens and assessed in real time for secretion of mAbs that bound to beads coated with a cocktail of MERS-CoV and HCoV-OC43 spikes, followed by beads coated with SARS-CoV-2 spike. Double-positive MBCs were exported for single-cell reverse transcription polymerase chain reaction and antibody expression as recombinant IgG1. In total, we obtained 60 IgG mAbs with reactivity to at least three coronaviruses.

To fully examine their breadth, we tested the 60 mAbs for binding to spikes from the seven human coronaviruses. Only six mAbs—COV91-27, COV44-62, COV44-79, COV77-04, COV77-39, and COV78-36—bound to spike proteins from all seven coronaviruses (Fig. 1, B and C). Notably, four of the six also bound to spike from two new coronaviruses that have recently been associated with human disease: canine CoV HuPn-2018 (CCoV-HuPn-2018) and

porcine deltacoronavirus 0081-4 (PDCoV-0081-4) (7, 8) (Fig. 1C). The six broadly reactive mAbs were isolated from four different donors and were encoded by four different heavy-chain variable (VH) genes (VH1-2, VH3-30, VH4-59, and VH5-10-1) (table S1). Five mAbs were highly mutated, with VH nucleotide mutation frequencies ranging from 10 to 13% (fig. S1C). Given that these mAbs were isolated from COVID-19-convalescent individuals in New York ~1 month after the first outbreak in March 2020, these mutation levels suggest that the B cells were primed during an earlier seasonal coronavirus infection and possibly reactivated during SARS-CoV-2 infection.

COV44-62 and COV44-79 broadly neutralize coronaviruses

We assessed the neutralizing potency of the six mAbs against SARS-CoV-2, SARS-CoV, MERS-CoV, HCoV-NL63, and HCoV-229E envelope pseudotyped viruses, as well as authentic HCoV-OC43. COV44-62 and COV44-79 showed the broadest functional reactivity, neutralizing the betacoronaviruses SARS-CoV-2, SARS-CoV, and HCoV-OC43, as well as the alphacoronaviruses HCoV-NL63 and HCoV-229E (Fig. 1D and fig. S1D). Moreover, both mAbs neutralized SARS-CoV-2 variants of concern, including the Omicron BA.2 and BA.4/5 sub-variants, as well as authentic SARS-CoV-2

(Fig. 1E and fig. S1E). COV44-62 also neutralized MERS-CoV, whereas no other mAbs neutralized this virus within the concentrations tested.

Broadly reactive mAbs target the coronavirus fusion peptide

To determine the domain of SARS-CoV-2 spike that was targeted by the six broadly reactive mAbs, we assessed mAb binding to the SARS-CoV-2 S2 subunit as well as the RBD and N-terminal domain (NTD) of the S1 subunit. All six mAbs bound only to the S2 subunit (Fig. 2A). VJ germline-reverted versions of the broadly neutralizing mAbs COV44-62 and COV44-79

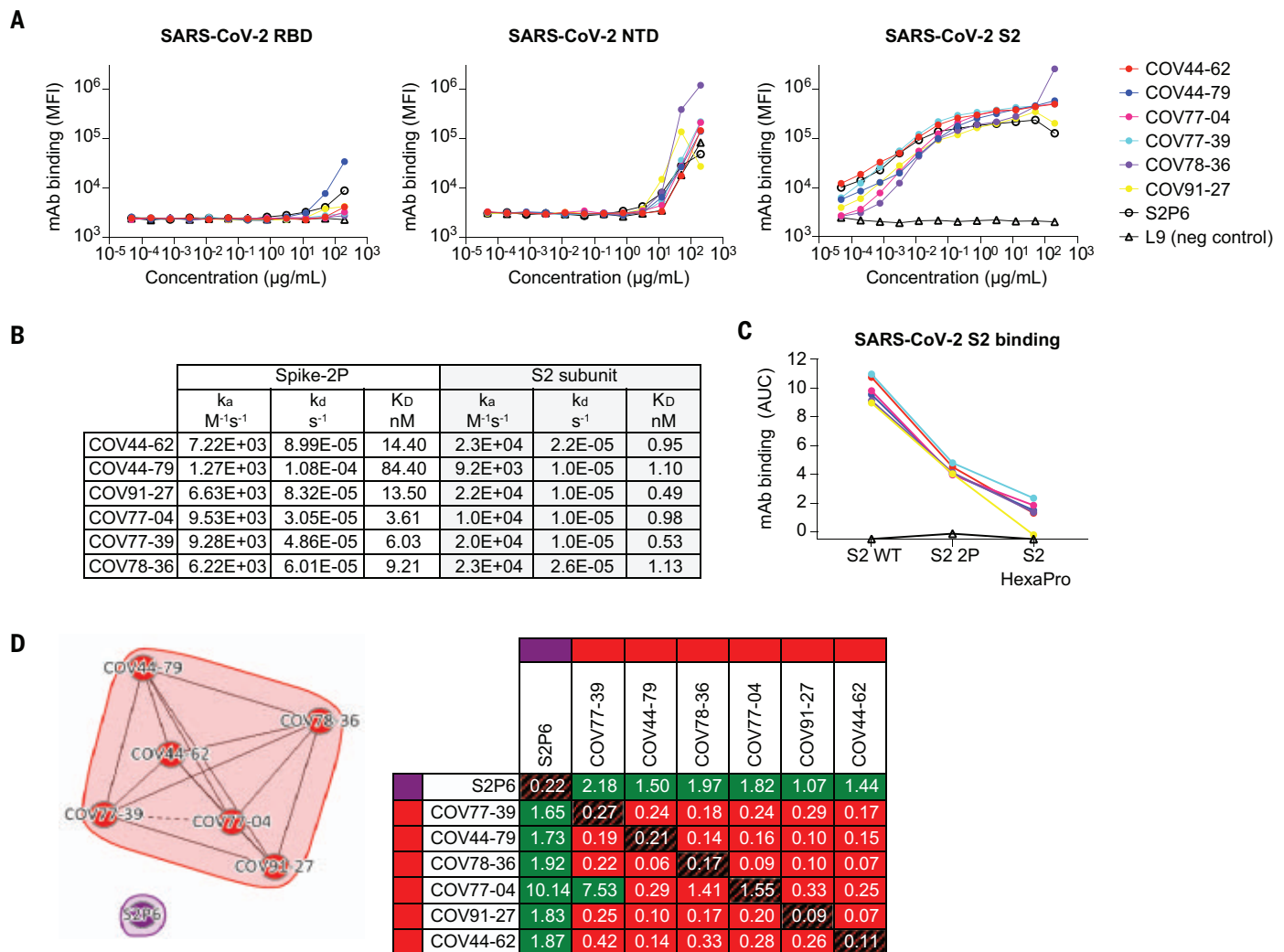


Fig. 2. Broadly reactive mAbs target the same region within the SARS-CoV-2 S2 subunit. (A) Titration curves for mAb binding to selected regions within the SARS-CoV-2 spike protein: the RBD, NTD, and S2 subunit. Interconnected data points are shown without curve fitting. L9 is a malaria-specific mAb used as a negative control (44). MFI, median fluorescence intensity. (B) On rates, off rates, and dissociation constants of the six fusion peptide Fabs for binding to SARS-CoV-2 prefusion-stabilized spike (2P) with an unmodified furin cleavage site and the nonstabilized S2 subunit. k_a , association rate constant; k_d , dissociation rate constant;

K_D , equilibrium dissociation constant. (C) Fusion peptide mAb binding (AUC) to wild-type (WT) SARS-CoV-2 S2 subunit and S2 subunit constructs modified with two (2P) or six (HexaPro) stabilizing proline mutations. (D) Epitope binning of broadly reactive antibodies versus the S2 stem-helix targeting mAb S2P6. All included antibodies were tested as both ligands and analytes. Solid lines indicate two-way competition, whereas the dashed line indicates one-way competition. Red boxes indicate competing antibody pairs, green boxes indicate noncompeting antibody pairs, and hatched boxes indicate self-competition.

showed weaker binding to S2 than the mutated versions (fig. S2A), suggesting that somatic mutations were important for improving binding to the target site. A surface plasmon resonance (SPR) kinetics assay determined the binding affinity of antigen-binding fragments (Fabs) derived from these mAbs for prefusion-stabilized whole SARS-CoV-2 spike (2P, with an intact S1-S2 cleavage site) and the unmodified S2 subunit. The Fabs bound with low-to-moderate nanomolar affinity to both proteins, but their affinity for the S2 subunit was 3- to 76-fold higher than their affinity for whole spike (Fig. 2B and fig. S2B). There were no substantial differences between the six Fabs in their affinity for the S2 subunit. The mAbs showed reduced binding to a form of S2 that had been stabilized with two proline mutations (S2-2P) and bound more poorly still to a further stabilized version with six proline mutations (S2-HexaPro) (Fig. 2C). SPR-based competition showed that the six mAbs competed for the same binding site on the S2 subunit (Fig. 2D), but none competed with S2P6, a control mAb targeting the stem helix region (I2), indicating their specificity for a distinct site on S2.

To further investigate the specificity of these mAbs, we performed SPR-based peptide mapping using an array of 15-amino acid (15-mer) overlapping peptides that spanned the entire SARS-CoV-2 S2 subunit (S686 to K1211, accession no. YP_009724390.1). All six mAbs bound to peptides 42 to 44, which share the $_{815}\text{RSFIEDLLF}_{823}$ motif (Fig. 3A). This motif is located within the SARS-CoV-2 fusion peptide region, directly C-terminal to the S2' cleavage site. To determine the diversity of this region across coronaviruses, we selected 34 viral isolates representing each of the four coronavirus genera (Fig. 3B). Nearly all amino acid positions in the $_{815}\text{RSFIEDLLF}_{823}$ motif were conserved in >85% of viruses selected; the exception was F₈₁₇, which was conserved in <50% of isolates examined (Fig. 3, B and C). The fusion peptide appears partially surface-exposed in a range of coronavirus spike proteins, including those of SARS-CoV-2, SARS-CoV, MERS-CoV, and MHV (mouse hepatitis virus) (21, 22) (Fig. 3C and fig. S3A). However, antibody access to this site may be partially occluded by the S1 subunit on an adjacent protomer, consistent with stronger binding to the S2 subunit relative to the SARS-CoV-2 spike (Fig. 2B and fig. S3B).

The mAb specificity toward the fusion peptide is consistent with their reduced binding to the HexaPro S2 construct (Fig. 2C), which includes a nonconservative F817P mutation at this site (23). To identify key amino acids for mAb binding, we performed an alanine scan on a peptide encompassing residues 810 to 830 and focused on residues targeted by the broadly neutralizing mAbs COV44-62 and

COV44-79 (Fig. 3D). Four amino acids—E819, D820, L822, and F823—were important for binding of COV44-62, where mutation of the F823 residue abolished binding. Similarly, residues critical for the binding of COV44-79 were E819, D820, and F823 but also included R815 at the S2' cleavage site (Fig. 3D). All five residues identified as important for COV44-62 or COV44-79 binding are among the most conserved residues in the coronavirus spike protein. D820 and L822 are completely conserved, whereas R815, E819, and F823 are conserved in 33 of 34 coronaviruses (Fig. 3, B and E). Amino acid mutations at the peptide level may have different effects from mutations in the intact spike protein, where modified interactions with surrounding residues may also affect antibody binding. Therefore, we screened the six mAbs using a shotgun alanine mutagenesis approach, whereby every amino acid in the S2 subunit of intact spike was individually mutated to generate a panel of spike mutants (Fig. 3F and fig. S3, C and D). In general, this assay identified a greater number of residues as important for mAb binding, including some with a more intermediate phenotype. For COV44-62, D820, L822, and F823 were again crucial for binding, and K825, D830, and R815 were also identified as important (Fig. 3F). For COV44-79, the results closely matched the peptide alanine scan, with the same four amino acids (R815, E819, D820, and F823) identified as the most critical. When the six mAbs were analyzed as a group, we found that only the four broadest neutralizing mAbs were negatively affected by the R815A mutation (fig. S4, A and B), suggesting that binding to the S2' cleavage site may be a distinguishing property of broadly neutralizing mAbs against this site.

Crystal structures of anti-fusion peptide antibodies

To elucidate the molecular characteristics of anti-fusion peptide antibodies that neutralize SARS-CoV-2, the Fabs of the three broadest neutralizing mAbs (COV44-62, COV44-79, and COV91-27) were complexed with 15-mer peptides containing the fusion peptide sequence (Fig. 4). Crystal structures were determined to 1.46-, 2.8-, and 2.3-Å resolution, respectively (Fig. 4, fig. S5, and table S2). Fourteen of the 15 peptide residues were visible in the electron density map for COV44-62 (fig. S5A), 13 of which have a buried surface area (BSA) >0 Å² in complex with antibody. For COV44-79, 12 of the 15 peptide residues were visible (fig. S5A), 10 of which have a BSA >0 Å². Similarly, in COV91-27, 12 peptide residues had interpretable density (Fig. 4C and fig. S5A), with nine exhibiting a BSA >0 Å². The fusion peptide forms a helix as in the prefusion state of the SARS-CoV-2 spike (fig. S5B). All three complementarity-determining regions (CDRs) of the heavy chain

(HC) of all three Fabs are involved in peptide recognition, whereas CDR1 and CDR3 of the light chain (LC) of COV44-62, and only LCDR3 of COV44-79 and COV91-27, contact the peptide (Fig. 4, A to C). The BSA on each Fab is dominated by the HC and is 791 Å² for COV44-62 (627 Å² by HC and 164 Å² by LC), 634 Å² for COV44-79 (505 Å² by HC and 129 Å² by LC), and 573 Å² for COV91-27 (447 Å² by HC and 126 Å² by LC). The fusion peptide makes side-chain and backbone H-bonds and salt bridges with COV44-62 mainly through K814, R815, E819, D820, L822, F823, and N824, and hydrophobic interactions through I818, L822, and F823 (Fig. 4D). These residues include the key residues (E819, D820, L822, and F823) identified by site-directed mutagenesis (Fig. 3D). The fusion peptide did not form as many interactions with COV44-79 and COV91-27 (Fig. 4, E and F). However, R815, E819, and D820 contributed H-bonds and salt bridges, and I818, L822, and F823 made hydrophobic interactions. In all three antibodies, R815, S816, I818, E819, D820, L822, and F823 contributed the most BSA to the interaction (Fig. 4, G to I). There was partial overlap between the residues of COV44-62 and COV44-79 that interacted with the fusion peptide and those that were mutated from the germline (fig. S6), consistent with the reduction in binding of germline-reverted versions of these mAbs (fig. S2A). The structural results are consistent with the peptide and spike protein mutagenesis data that identify the key binding residues (Fig. 3, D to F, and fig. S3D). Notably, the arginine at the S2' cleavage site is involved in recognition by these anti-fusion peptide antibodies. Although the antibodies all interact with one face of the fusion peptide helical structure (fig. S5D), their approach angles to the peptide differ. Superimposition of the fusion peptide structures onto an intact SARS-CoV-2 spike trimer structure in the prefusion state showed a potential clash with the S protein, suggesting that a conformational change or conformational dynamics around the fusion peptide is required to accommodate antibody targeting (fig. S5C). Nevertheless, these antibodies have neutralization activity against SARS-CoV-2 and therefore are able to interact with the fusion peptide on the virus (Fig. 1D).

Response to the fusion peptide after vaccination and infection

We compared the binding of polyclonal IgG from mRNA-1273-vaccinated donors (fig. S7A), COVID-19-convalescent individuals, and COVID-19-naïve individuals to the SARS-CoV-2 fusion peptide (peptide 43) (Fig. 3A and fig. S7B). All COVID-19-naïve donors showed minimal binding to the peptide, indicating a minor contribution by previous seasonal coronavirus infections to circulating fusion peptide-specific IgG. Although there was an increased response in

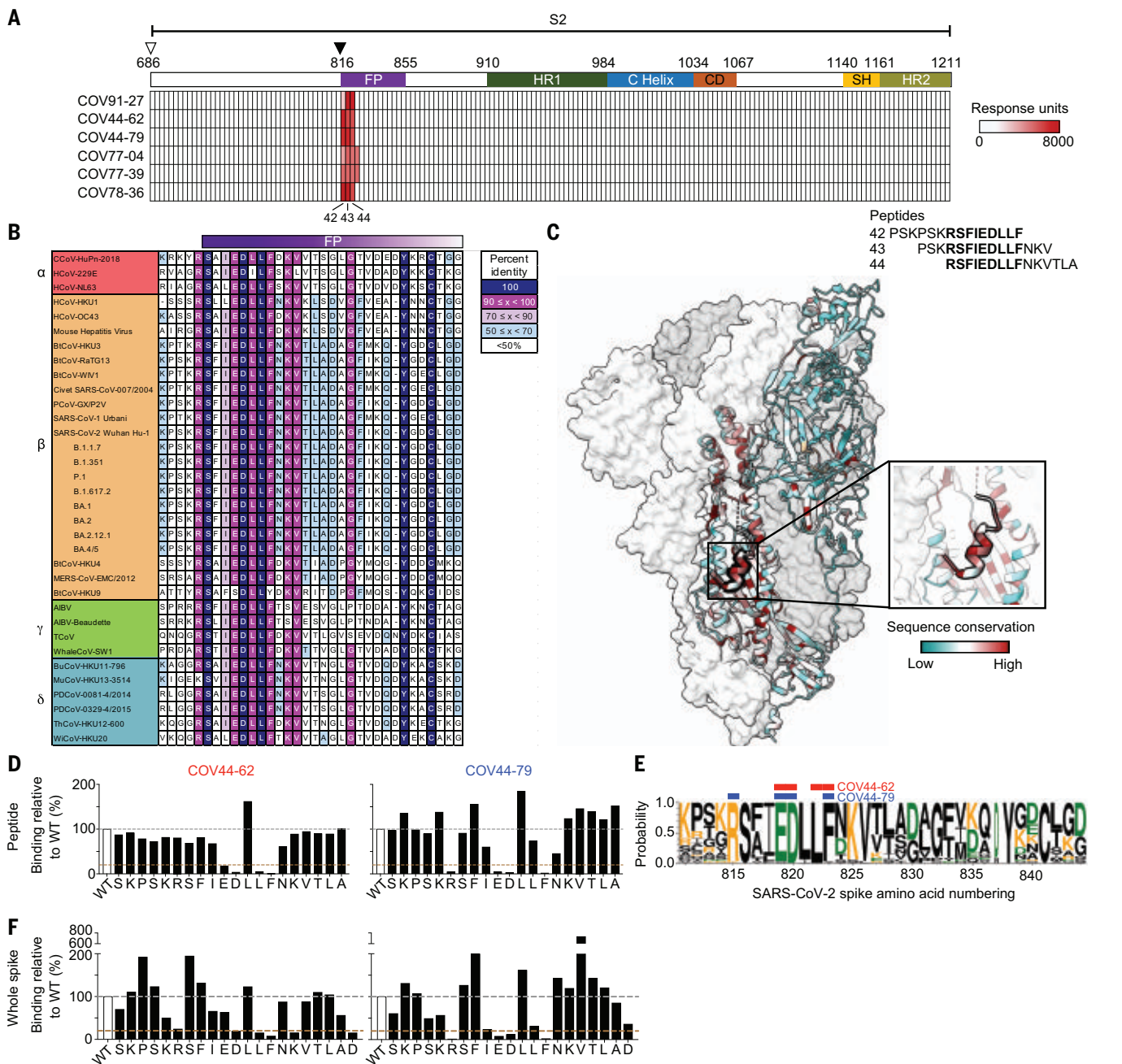


Fig. 3. Broadly neutralizing antibodies target the conserved fusion peptide. (A) Heatmap of SARS-CoV-2 S2 peptide array. Binding responses were assessed by SPR using a 15-mer peptide array with 12-amino acid overlap covering the entire S2 subunit. Each column within the map represents a single peptide. The white triangle denotes the S1-S2 cleavage site, and the black triangle indicates the S2' cleavage site. FP, fusion peptide; HR1, heptad repeat 1; C Helix, central helix; CD, connector domain; SH, stem helix; HR2, heptad repeat 2. (B) Sequence alignment of the fusion peptide from 34 viral isolates representing a diverse group of coronaviruses across four genera. This analysis was performed using MAFFT v7 with a BLOSUM62 scoring matrix and the L-INS-I algorithm. Single-letter abbreviations for amino acids: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (C) Sequence conservation of prefusion SARS-CoV-2 spike protein (PDB ID 6VSB); the fusion peptide (amino acids 816 to 843) is outlined in black. The inset shows a magnified view of this region. (D) Alanine scan evaluating the

binding of COV44-62 and COV44-79 to the SARS-CoV-2 fusion peptide. Responses were normalized to the wild-type sequence. A cutoff of 20% (brown dashed line) was used to identify residues that were critical for binding. (E) Sequence logo plot of diversity within the fusion peptide region of coronaviruses from 34 isolates, built using WebLogo 3. Letter height is proportional to the probability of an amino acid at a given position, and amino acid residues are colored by charge. Narrow stacks (amino acids) indicate deletions or gaps in the sequences. Numbering is based on the SARS-CoV-2 Wuhan-Hu-1 sequence. The key residues in the epitope footprints of mAbs COV44-62 (red) and COV44-79 (blue), based on peptide alanine scanning, are highlighted above the logo plot. (F) Amino acid scan for the binding of COV44-62 and COV44-79, identified by shotgun alanine mutagenesis of S2 residues on whole spike protein. Only fusion peptide residues are shown. Key residues were identified on the basis of a <20% signal relative to wild-type spike (brown dashed line), with no corresponding loss of signal for a control mAb, which targets the spike protein but does not bind to this site (fig. S3C).

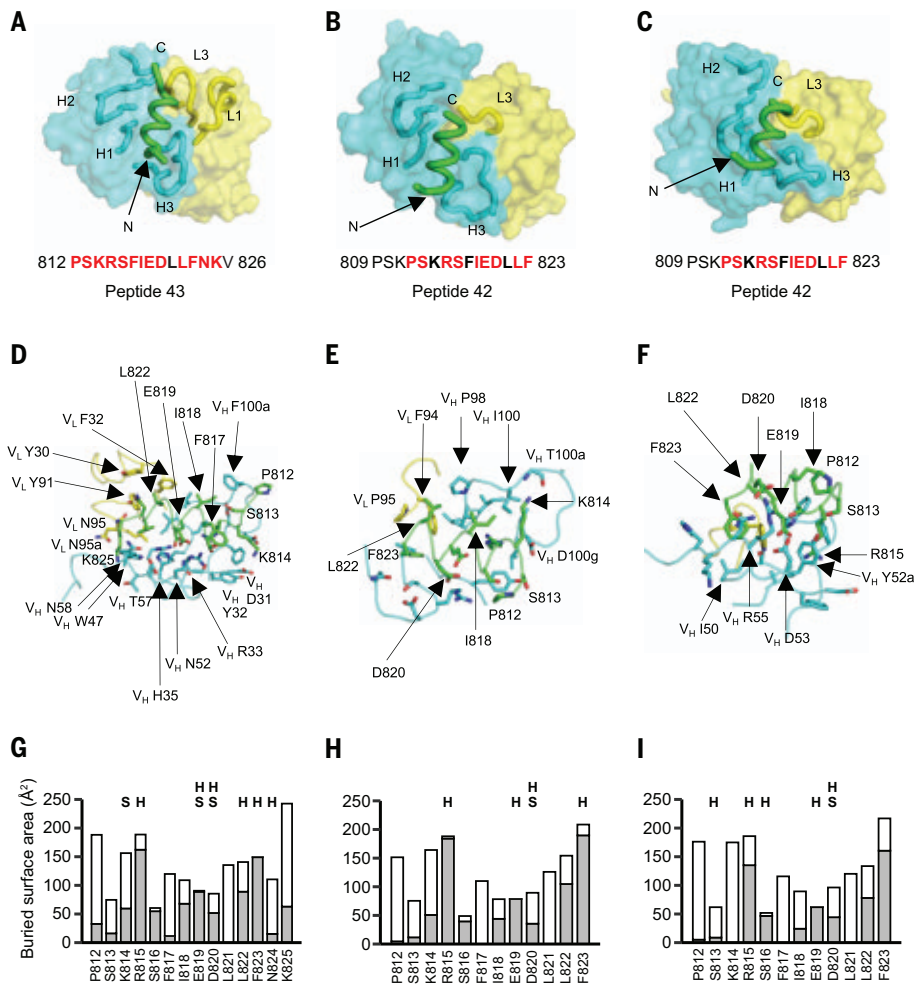


Fig. 4. Crystal structures of COV44-62, COV44-79, and COV91-27 in complex with SARS-CoV-2 fusion peptide. (A to C) Overall interactions of (A) COV44-62, (B) COV44-79, and (C) COV91-27 with the fusion peptide. Fabs are shown in molecular surface representation, and the CDRs and peptides are represented as tubes. Cyan and yellow denote the heavy and light chains of the Fabs, respectively. Peptides are shown in green. H1, H2, H3, L1, and L3 denote CDRs in the heavy (H) and light (L) chains. The resolutions of the crystal structures are 1.46, 2.8, and 2.3 Å for the COV44-62, COV44-79, and COV91-27 complexes, respectively. Peptide residues observed in the crystal structures are in bold; residues that interact with antibodies (BSA >0 Å²) are in red. (D to F) Details of the interactions between (D) COV44-62, (E) COV44-79, and (F) COV91-27 with the fusion peptide. V_H and V_L indicate the variable domains of the heavy and light chains, respectively. Kabat numbering was used for the Fabs; numbering in the native spike protein was used for the fusion peptide. Colors for the heavy chain, light chain, and fusion peptide are as in (A). (G to I) BSA (gray) and accessible surface area (white) of each residue of the fusion peptide in complex with antibody are shown in the stacked bar chart. Residues that form polar interactions with COV44-62, COV44-79, and COV91-27 are denoted atop the corresponding bar with “H” for a hydrogen bond or “S” for a salt bridge. Buried and accessible surface areas were calculated with PISA (45).

several vaccine recipients after the second dose ($P = 0.025$), this was not enhanced by administration of a booster dose. The COVID-19-convalescent donors did not have significantly higher responses than vaccinated donors after the second dose ($P = 0.864$). However, several convalescent donors had the highest responses in all three cohorts, suggesting that SARS-CoV-2 infection triggers a strong fusion peptide-specific antibody response in some individuals.

COV44-62 and COV44-79 inhibit membrane fusion

Spike-mediated cell fusion relies on insertion of the fusion peptide into the target cell membrane and thus might be inhibited by antibody binding to the fusion peptide. Consistent with this, COV44-62 and COV44-79 inhibited the fusion of cells expressing SARS-CoV-2 spike and cells expressing the ACE2 receptor in an imaging-based assay (Fig. 5A). We further tested the six fusion peptide-specific mAbs

in a more quantitative assay wherein fusion would trigger the release of an enzyme that cleaves a chromogenic substrate. Only the mAbs that neutralized SARS-CoV-2 were able to strongly inhibit fusion (Fig. 5B).

COV44-79 limits disease in the Syrian hamster model

We evaluated the in vivo efficacy of COV44-62 and COV44-79 against SARS-CoV-2 infection in the Syrian hamster model, a well-established model that recapitulates features of moderate-to-severe COVID-19 in humans (24–26). We converted the Fc regions of the two mAbs to hamster IgG2 to allow optimal Fc function. The mAbs were administered intraperitoneally at 16 mg per kilogram of body weight (mg/kg), followed by intranasal administration of 5 log₁₀ plaque-forming units of SARS-CoV-2 WAI 24 hours later (Fig. 5, C and D, and fig. S8). Hamsters treated with COV44-79 and, to a lesser extent, COV44-62 had a smaller decrease in body weight and recovered more quickly than untreated hamsters ($P < 0.01$ from days 3 to 7 for COV44-79; $P < 0.05$ from days 5 to 7 for COV44-62) (Fig. 5C). Similar results were observed in a second experiment in which COV44-79 was tested in comparison with a hamster IgG2 isotype control (fig. S8B). Furthermore, semiquantitative scoring revealed that hamsters treated with COV44-79 had less interstitial pneumonia than untreated hamsters on day 7 ($P < 0.05$) (Fig. 5D). COV44-79 was also able to slightly reduce lung viral titers relative to control hamsters, on the basis of subgenomic RNA quantification and plaque assay analysis (fig. S8C).

Discussion

The broad conservation and functional importance of the fusion peptide highlight the potential of this site as a candidate for coronavirus vaccine development. These findings have parallels to studies of HIV-1 gp120, where the surface-exposed fusion peptide was identified as a target of neutralizing mAbs (27). This discovery led to the investigation of the HIV-1 fusion peptide as a candidate immunogen to elicit broadly neutralizing antibodies in animals (28); subsequent animal studies have increased the potency and breadth of the fusion peptide-targeting antibody response (29). Despite these potential advantages, the coronavirus fusion peptide has not been a major focus for development of therapeutic mAbs and COVID-19 vaccines. The main drawback of the mAbs described here is their comparatively low in vitro neutralization potency. These mAbs fit into a wider trend of a trade-off between potency and breadth: Highly potent mAbs that target the RBD are restricted to sarbecoviruses and most do not neutralize all variants of SARS-CoV-2 (3, 5, 6, 20, 30), whereas mAbs that target the stem helix (12–16) and

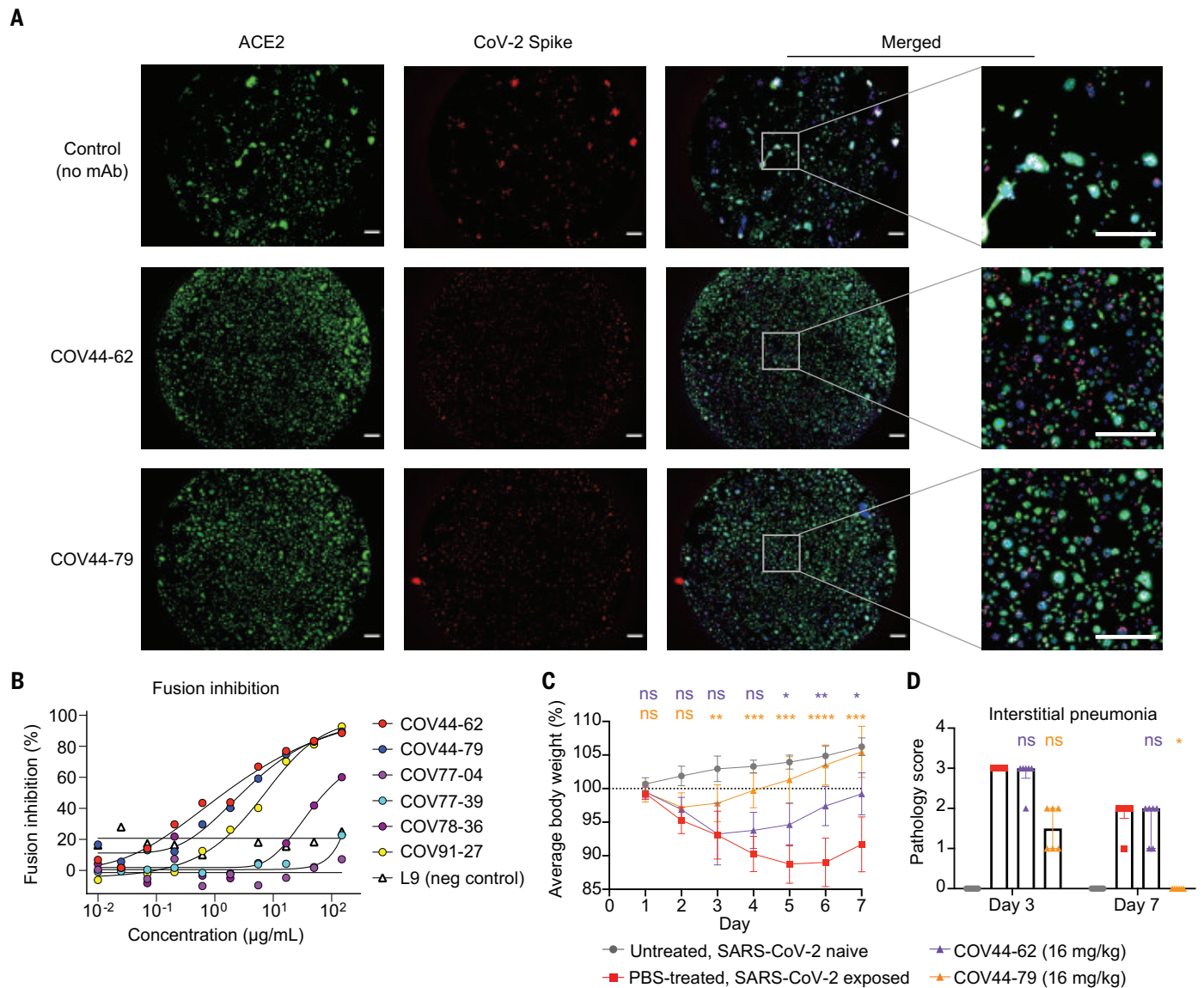


Fig. 5. COV44-62 and COV44-79 inhibit SARS-CoV-2 spike-mediated fusion and COV44-79 limits disease in a Syrian hamster model. (A) Images of fusion between HeLa cells stably expressing SARS-CoV-2 spike (red fluorescent protein) and HeLa cells stably expressing the ACE2 receptor (green fluorescent protein) after counterstaining with Hoechst (blue). Cells were cocultured in the presence of COV44-62 or COV44-79, or without a mAb (control). Scale bars, 500 μm . (B) Fusion inhibition of six fusion peptide-specific mAbs in a quantitative assay. (C) Weight change for SARS-CoV-2-naive animals versus virus-exposed animals that were mock-treated or treated with 16 mg/kg of mAb. Statistical significance for average body weight was analyzed

across the 7-day time course using a mixed-effects repeated measures model with Dunnett's post-test multiple comparison ($n = 12$ animals from days 0 to 3; $n = 6$ animals from days 4 to 7). Error bars show mean \pm SD. (D) Pathology scores for SARS-CoV-2-naive animals versus virus-exposed animals that were mock-treated or treated with 16 mg/kg of mAb. Scores for interstitial pneumonia pathology (days 3 and 7) based on gross pathology observations were statistically analyzed by a Kruskal-Wallis test with Dunn's post-test multiple comparison ($n = 6$ to 12 animals per condition) between the mAb-treated and mock-treated groups on each day. Bars show median and interquartile range. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$; ns, not significant.

those identified here have greater breadth but are less potent. However, as previously reported for at least one antistem helix mAb (15), COV44-79 performed better than expected in the hamster model, which suggests that it may function in a way that is not captured effectively in the neutralization assay. For instance, Fc effector functions may enhance activity, as observed with anti-SARS-CoV-2 antibodies with other specificities (31–33).

Moreover, there is substantial scope for improvement for mAbs with this specificity, and subsequent studies may discover mAbs with higher potency, as with the HIV-1 fusion peptide (34). Techniques to improve antibody affinity and potency could also be useful (35, 36).

Additionally, vaccination with fusion peptide constructs may trigger a polyclonal response of greater magnitude and potency. We found that three doses of the mRNA-1273 vaccine did not

produce strong antibody responses to the fusion peptide, although several COVID-19-convalescent individuals exhibited strong antibody responses to this site. This observation is consistent with greater exposure of the S2 subunit to B cells during natural infection as a result of S1 uncoupling, which likely occurs less frequently with prefusion-stabilized spike protein. Furthermore, depletion of fusion peptide-specific antibodies from the serum

of COVID-19–convalescent patients resulted in a 20% reduction in SARS-CoV-2 neutralization (37), and reactivity to the fusion peptide was correlated with neutralization titer (38), indicating that polyclonal antibodies to the fusion peptide can play an appreciable role. Thus, our findings are consistent with previous studies that have highlighted the potential utility of the fusion peptide as a target epitope (37–41), and the fusion peptide–targeted mAbs provide additional tools to combat COVID-19 and enhance pandemic preparedness.

REFERENCES AND NOTES

- P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, V. Thiel, *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
- E. Dong, H. Du, L. Gardner, *Lancet Infect. Dis.* **20**, 533–534 (2020).
- S. Iketani et al., *Nature* **604**, 553–556 (2022).
- N. Andrews et al., *N. Engl. J. Med.* **386**, 1532–1546 (2022).
- E. Takashita et al., *N. Engl. J. Med.* 10.1056/NEJM2207519 (2022).
- L. A. VanBlargan et al., *Nat. Med.* **28**, 490–495 (2022).
- J. A. Lednický et al., *Nature* **600**, 133–137 (2021).
- A. N. Vladnina et al., *Clin. Infect. Dis.* **74**, 446–454 (2022).
- C. B. Jackson, M. Farzan, B. Chen, H. Choe, *Nat. Rev. Mol. Cell Biol.* **23**, 3–20 (2022).
- J. S. Tregoning, K. E. Flight, S. L. Higham, Z. Wang, B. F. Pierce, *Nat. Rev. Immunol.* **21**, 626–636 (2021).
- T. N. Starr et al., *Cell* **182**, 1295–1310.e20 (2020).
- D. Pinto et al., *Science* **373**, 1109–1116 (2021).
- M. M. Sauer et al., *Nat. Struct. Mol. Biol.* **28**, 478–486 (2021).
- C. Wang et al., *Nat. Commun.* **12**, 1715 (2021).
- P. Zhou et al., *Sci. Transl. Med.* **14**, eabi9215 (2022).
- W. Li et al., *Cell Rep.* **38**, 110210 (2022).
- F. Amanat et al., *Cell* **184**, 3936–3948.e10 (2021).
- K. A. Huang et al., *PLOS Pathog.* **17**, e1009352 (2021).
- W. N. Voss et al., *Science* **372**, 1108–1112 (2021).
- H. Cho et al., *Sci. Transl. Med.* **13**, eabj5413 (2021).
- A. C. Walls et al., *Nature* **531**, 114–117 (2016).
- Y. Yuan et al., *Nat. Commun.* **8**, 15092 (2017).
- C. L. Hsieh et al., *Science* **369**, 1501–1505 (2020).
- J. F. Chan et al., *Clin. Infect. Dis.* **71**, 2428–2446 (2020).
- M. Imai et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16587–16595 (2020).
- S. F. Sia et al., *Nature* **583**, 834–838 (2020).
- R. Kong et al., *Science* **352**, 828–833 (2016).
- K. Xu et al., *Nat. Med.* **24**, 857–867 (2018).
- R. Kong et al., *Cell* **178**, 567–584.e19 (2019).
- M. A. Tortorici et al., *Nature* **597**, 103–108 (2021).
- I. Ullah et al., *Immunity* **54**, 2143–2158.e15 (2021).
- R. Yamin et al., *Nature* **599**, 465–470 (2021).
- Y. C. Bartsch et al., *Nat. Med.* **27**, 454–462 (2021).
- C. H. Shen et al., *Cell Host Microbe* **27**, 531–543.e6 (2020).
- S. Kratochvíl et al., *Immunity* **54**, 2859–2876.e7 (2021).
- A. Wellner et al., *Nat. Chem. Biol.* **17**, 1057–1064 (2021).
- C. M. Poh et al., *Nat. Commun.* **11**, 2806 (2020).
- W. R. Morgenlander et al., *J. Clin. Invest.* **131**, e146927 (2021).
- X. Sun et al., *Nat. Microbiol.* **7**, 1063–1074 (2022).
- N. Vanderheijden et al., *Front. Immunol.* **12**, 790415 (2022).
- D. L. N. F. Maeda et al., *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025622118 (2021).
- C. Dreyfus et al., *Science* **337**, 1343–1348 (2012).
- T. F. Rogers et al., *Science* **369**, 956–963 (2020).
- L. T. Wang et al., *Immunity* **53**, 733–744.e8 (2020).
- E. Krissinel, K. Henrick, *J. Mol. Biol.* **372**, 774–797 (2007).

ACKNOWLEDGMENTS

We thank the blood sample donors at the New York Blood Center; S. Bangaru, G. Ozorowski, A. Torrents de la Peña, and A. Ward for providing HCoV-OC43 and MERS-CoV; G. Wright and N. Muller-Siennerth (Wright lab, University of York) for providing recombinant CD4; L. Wang and R. Seder for providing the L9 antibody; the Burton lab (The Scripps Research Institute) for providing BA.2 and BA.4/5 gene fragments; and M. Cohen and J. Laux for assistance with cell sorting. We also thank N. Vaughan, K. Cooper, R. Reeder, M. St. Claire, K. Hadley, D. Drawbaugh, A. Hischak, R. Hart, N. Isic, M. Murphy, E. Postnikova, M. Anantpadma, and E. Eudy for assistance with hamster experiments. We are grateful to the staff of Advanced Photon Source and Stanford Synchrotron Radiation Lightsource (SSRL) Beamline 12-1 for assistance with synchrotron data collection. **Funding:** This work was supported by the

Division of Intramural Research and the Vaccine Research Center, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH) (I.D., S.M., J.W.Y., P.D.C., M.R.H., J.R.M., and J.T.). This project has been funded in whole or in part with federal funds from the NIAID, NIH, US Department of Health and Human Services (DHHS), under contract HHSN272201800013C. This work was also supported by NIH grant R01AI132317 (D.N. and L.P.), HHSN contract 75N93019C00073 (J.K.W. and E.D.), and Bill and Melinda Gates Foundation grant INV-004923 (I.A.W.). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science user facility operated for the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357. GM/CA@APS has been funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006 and P30GM138396). Extraordinary facility operations were supported in part by the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on the response to COVID-19, with funding provided by the Coronavirus CARES Act. Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the DOE, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institute of General Medical Sciences of the NIH (P30GM133894). **Author contributions:** Conceptualization: J.T. and C.D. Methodology: C.D., C.T., L.P., C.-C.D.L., T.-H.L., M.Y., Y.C., L.W., D.D., B.E., I.K., Z.H., R.S.W., A.P., E.D., D.E.G., I.D., and S.M. Formal analysis: C.D., C.T., L.P., M.Y., Y.C., L.W., C.-C.D.L., T.-H.L., C.-W.P., J.S., and E.D. Data curation: M.P. Investigation: C.D., C.T., L.P., M.Y., L.W., C.-C.D.L., L.P., C.-W.P., J.K.W., T.-H.L., M.Z., M.P., D.M., A.J.R.C., S.D., E.K., L.H., D.P., R.B., S.L., D.D., B.E., Y.Z., E.S.Y., M.C., and K.L. Resources: R.S.W. and S.M. Writing – original draft: C.D., C.T., M.Y., and J.T. Writing – review & editing: all authors. Visualization: C.D., C.T., L.P., M.Y., Y.C., L.W., C.-C.D.L., J.K.W., E.D., and J.T. Supervision: E.D., D.E.G., S.M., J.W.Y., C.S., P.D.C., D.N., M.R.H., J.R.M., I.A.W., and J.T. Funding acquisition: C.S., P.D.C., D.N., M.R.H., J.R.M., I.A.W., and J.T. **Competing interests:** J.T. and C.D. are coinventors on a

provisional patent (US Patent Application no. 63/308,898) filed on the mAbs described in this study. J.K.W. and E.D. are employees of Integral Molecular. Y.C., S.D., E.K., L.H., D.P., R.B., S.L., D.D., B.E., and M.R.H. performed this work as employees of Lualaba Government Solutions, LLC. The content of this publication does not necessarily reflect the views or policies of the DHHS or of the institutions and companies with which the authors are affiliated. All other authors declare no competing interests. **Data and materials availability:** All data associated with this manuscript are available in the main text or the supplementary materials. Crystal structures have been deposited into the Protein Data Bank (PDB IDs 8D36 for COV44-62, 8DAO for COV44-79, and 8D6Z for COV91-27). Antibody sequences have been deposited in GenBank (accession numbers ON695901 to ON695912). Materials described in this manuscript will be available through a material transfer agreement with the NIAID. **License information:** This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq3773
Materials and Methods
Figs. S1 to S8
Tables S1 and S2
References (46–71)
MDAR Reproducibility Checklist

Submitted 4 April 2022; accepted 6 July 2022
Published online 12 July 2022
10.1126/science.abq3773

CORONAVIRUS

ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies

Jun Siong Low^{1,2*}†, Josipa Jerak^{1,2†}, M. Alejandra Tortorici^{3†}, Matthew McCallum^{3†}, Dora Pinto^{4†}, Antonino Cassotta¹, Mathilde Foglierini¹, Federico Mele¹, Rana Abdelnabi⁵, Birgit Weynand⁶, Julia Noack⁷, Martin Montiel-Ruiz⁷, Siro Bianchi⁴, Fabio Benigni⁴, Nicole Sprugasci⁴, Anshu Joshi³, John E. Bowen³, Cameron Stewart³, Megi Rexhepaj³, Alexandra C. Walls^{3,8}, David Jarrossay¹, Diego Morone¹, Philipp Paparoditis¹, Christian Garzoni⁹, Paolo Ferrari^{10,11,12}, Alessandro Ceschi^{10,13,14,15}, Johan Neyts^{5,16}, Lisa A. Purcell¹⁷, Gyorgy Snell¹⁷, Davide Corti⁴, Antonio Lanzavecchia^{4,17}§, David Veasley^{3,8*}§, Federica Sallusto^{1,2*}§

The coronavirus spike glycoprotein attaches to host receptors and mediates viral fusion. Using a broad screening approach, we isolated seven monoclonal antibodies (mAbs) that bind to all human-infecting coronavirus spike proteins from severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) immune donors. These mAbs recognize the fusion peptide and acquire affinity and breadth through somatic mutations. Despite targeting a conserved motif, only some mAbs show broad neutralizing activity in vitro against alpha- and betacoronaviruses, including animal coronaviruses WIV-1 and PDF-2180. Two selected mAbs also neutralize Omicron BA.1 and BA.2 authentic viruses and reduce viral burden and pathology in vivo. Structural and functional analyses showed that the fusion peptide–specific mAbs bound with different modalities to a cryptic epitope hidden in prefusion stabilized spike, which became exposed upon binding of angiotensin-converting enzyme 2 (ACE2) or ACE2-mimicking mAbs.

The *Orthocoronavirinae* subfamily of coronaviruses comprises four genera: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus*. Alphacoronaviruses NL63 and 229E and betacoronaviruses OC43, HKU1, severe acute

respiratory syndrome coronavirus 2 (SARS-CoV-2), SARS-CoV, and Middle East respiratory syndrome (MERS)-CoV account for almost all human coronavirus (hCoV) infections, with sporadic infections attributed to canine alphacoronavirus CCov-HuPn-2018

and porcine deltacoronavirus (PDCoV) (1–3). The spike (S) glycoprotein, which binds to various host receptors for viral entry (table S1), is composed of the S₁ and S₂ subunits, is highly divergent, with only ~30% sequence identity between alpha- and betacoronaviruses, and is the main target of neutralizing antibodies (4–7). Previous studies have described neutralizing monoclonal antibodies (mAbs) that cross-react among sarbecoviruses by targeting the receptor-binding domain (RBD) in the S₁ subunit (8–14) or more broadly across betacoronaviruses by targeting the stem helix in the S₂ subunit (15–20). However, mAbs neutralizing both alpha- and betacoronaviruses have not been reported. Broadly neutralizing mAbs, as exemplified by those targeting influenza viruses or HIV-1 (21–26), can potentially be used for prophylaxis or therapy and to guide the design of vaccines eliciting broadly protective immunity (27, 28).

Isolation of broadly reactive coronavirus mAbs from convalescent and vaccinated individuals

To search for antibodies that cross-react with human-infecting alpha- and betacoronaviruses, we stimulated, under limiting conditions, total peripheral blood mononuclear cells (PBMCs) from SARS-CoV-2 immune donors in the presence of the TLR7/8 agonist R848 and interleukin-2 (IL-2), which selectively induce the proliferation and differentiation of memory B cells (29). On day 12, the specificities of immunoglobulin G (IgG) secreted in the culture supernatants were tested by enzyme-linked immunosorbent assay (ELISA) against a panel of recombinant S proteins from alpha and beta hCoVs (Fig. 1, A to C). The number of SARS-CoV-2 S IgG-positive cultures was gener-

ally higher in COVID-19 convalescent patients and in SARS-CoV-2 vaccinees with prior infection (preimmune) compared with vaccinees without prior infection (naïve). Most SARS-CoV-2 S IgG-positive cultures were either monospecific or cross-reactive with the closely related SARS-CoV S, whereas a small fraction was also reactive with OC43 and HKU1 S (Fig. 1, A to D), consistent with previous serological analyses (17, 30, 31). Six cultures (of >4000) from five individuals (of 43) cross-reacted with all alpha and beta hCoV S proteins tested (Fig. 1, A to D), suggesting that memory B cells producing broadly reactive coronavirus antibodies might exist at very low frequency.

To isolate broadly reactive coronavirus mAbs, we combined the screening of polyclonally activated memory B cells with the sorting and cloning of antibody-secreting cells to retrieve paired heavy- and light-chain sequences (fig. S1). Using this approach, we isolated 16 SARS-CoV-2 S-specific mAbs that cross-reacted with various hCoV S proteins (Fig. 1E and fig. S2A). Six mAbs (group 1) cross-reacted with the betacoronaviruses SARS-CoV, SARS-CoV-2, OC43, and HKU1 S proteins, with median effective concentration (EC₅₀) values ranging from 45 to 3000 ng/ml. Three mAbs (group 2) cross-reacted with high affinity to all beta hCoV S proteins, with EC₅₀ values ranging from 18 to 40 ng/ml (Fig. 1E). These mAbs were found to target the stem helix region and have been described in a separate study (15). The remaining seven mAbs (group 3) exhibited the broadest cross-reactivity to both alpha and beta hCoV S proteins, with EC₅₀ values ranging from 29 to 800 ng/ml (Fig. 1E). These broadly reactive mAbs, which are the focus of the present study, were isolated from convalescent or vaccinated individuals, used different V genes (except for C13B8 and C13A7, which are clonally related), and displayed a high load of somatic mutations (7 to 14% in VH, 2 to 8% in VL at the nucleotide level; table S2). These results illustrate the utility of a simple high-throughput method based on multiple parallel screening steps of memory B cells to isolate broadly reactive coronavirus mAbs.

Broadly reactive mAbs bind to the fusion peptide and acquire affinity and breadth through somatic mutations

Using SARS-CoV-2 S₁, RBD, and S₂ protein domains, as well as 15-mer linear peptides covering the entire S sequence, the specificities of all seven group 3 mAbs were mapped to the K811PSKRSFIEDLLFNK825 sequence in the SARS-CoV-2 S₂ subunit (Fig. 2A). This sequence spans the S₂' cleavage site (R815) and the fusion peptide N-terminal region, which is essential for membrane fusion (32) and is highly conserved among all genera of the *Orthocoronavirinae* subfamily, including deltacoronavirus PDCoV and gammacoronavirus IBV, as well as all SARS-CoV-2 variants sequenced to date (fig.

S3, A and B). Most of the seven group 3 mAbs bound more tightly to the prefusion SARS-CoV-2 S trimer without stabilizing proline substitution in the fusion peptide (PentaPro) relative to HexaPro (F817P) (33), and all of them bound with greater avidity to prefusion SARS-CoV-2 S₂ compared with postfusion S₂ (Fig. 2A and fig. S2D).

To explore the ontogeny of the isolated fusion peptide-specific mAbs, we compared the binding properties of mature mAbs to their unmutated common ancestors (UCAs) (Fig. 2B and fig. S2B). The UCAs of C13B8, C13A7 (the two clonally related mAbs), and VN01H1 exhibited broad reactivity with hCoV S proteins but had lower affinity than their mature counterparts. By contrast, the C13C9 UCA bound only to the beta hCoVs OC43 and MERS-CoV S, whereas the VP12E7 UCA bound only to the alpha hCoV NL63 S, with low affinity in both cases. Finally, the UCAs of C28F8 and C77G12 did not bind to any hCoV S proteins tested. Given the lack of a common V gene usage, these findings suggest that broadly reactive fusion peptide-specific mAbs can mature through multiple pathways and acquire high affinity and cross-reactivity through somatic mutations, possibly because of priming by endemic coronavirus infection followed by SARS-CoV-2 boosting upon infection or vaccination.

Considering the high conservation of the fusion peptide region, we next assessed whether broad reactivity is a property shared by most fusion peptide-specific mAbs. IgGs secreted upon polyclonal activation of memory B cells from 71 convalescent individuals were screened for binding to a pool of alpha and beta hCoV fusion peptides, as well as to SARS-CoV-2 S protein. Cultures producing fusion peptide-specific mAbs were detected at low frequency and only in 19 individuals (fig. S4A). Although nearly all fusion peptide-reactive mAbs bound to SARS-CoV-2 S, only nine of 30 were broadly reactive, and the remaining mAbs showed different degrees of cross-reactivity (fig. S4B). Thus, although the coronavirus S fusion peptide is conserved (34–39), broad reactivity is the property of a minority of fusion peptide-specific mAbs, consistent with the observation that only three of the seven group 3 mAbs cross-reacted with the PDCoV and IBV fusion peptides (fig. S3C).

Fusion peptide-specific mAbs have heterogeneous neutralizing activities and can reduce viral burden in vivo

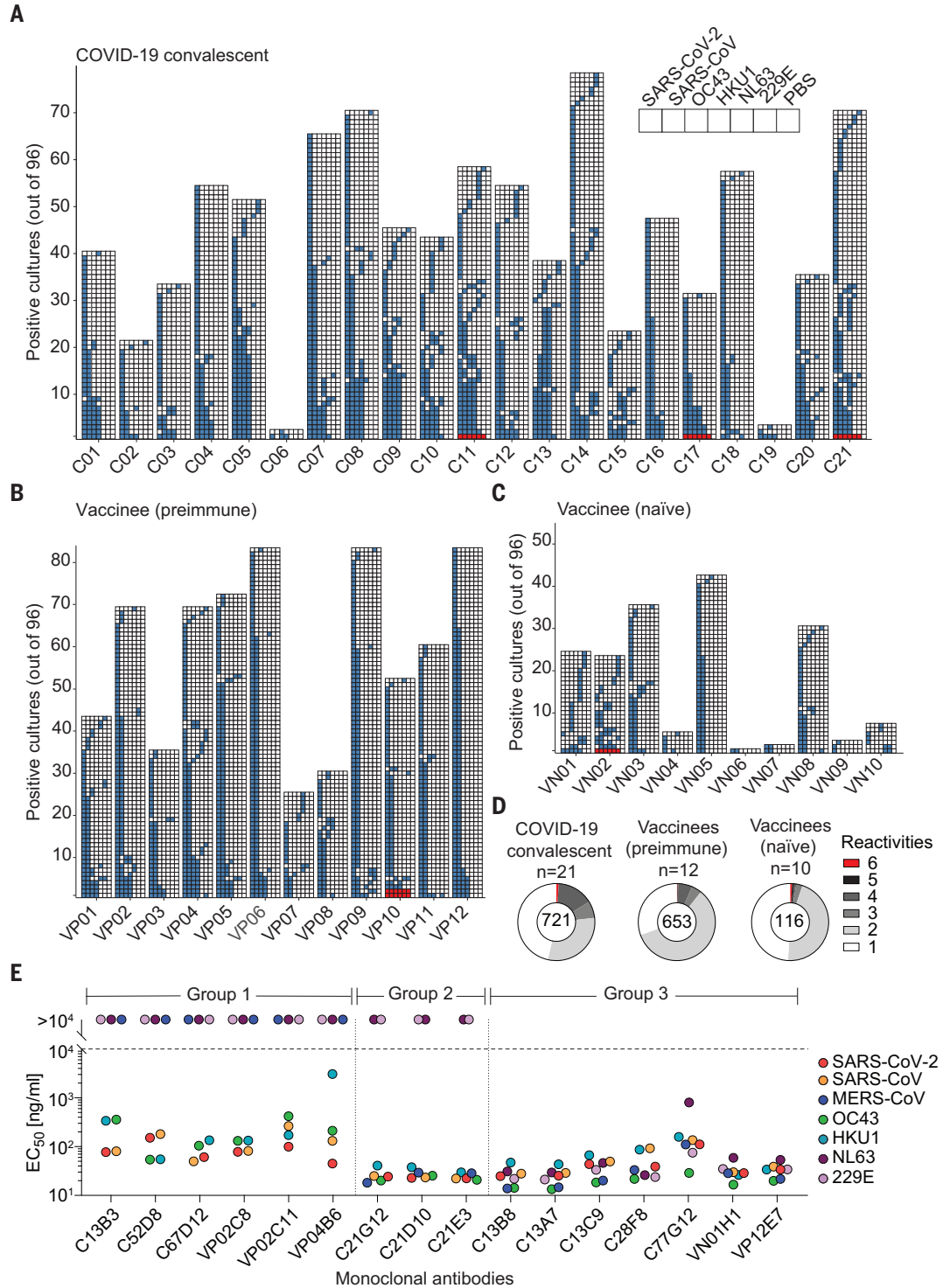
We next tested the neutralizing activity of all seven group 3 mAbs against pseudotyped viruses carrying S proteins of alpha- and betacoronaviruses. Despite binding to the same motif, these mAbs exhibited distinct neutralizing potencies. Most notably, VN01H1 and VP12E7 neutralized all hCoV S-pseudotyped viruses tested (SARS-CoV-2, SARS-CoV, MERS-CoV, NL63, and 229E), as

¹Institute for Research in Biomedicine, Università della Svizzera Italiana, 6500 Bellinzona, Switzerland. ²Institute of Microbiology, ETH Zürich, 8093 Zürich, Switzerland. ³Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. ⁴Humabs BioMed SA (subsidiary of Vir Biotechnology), 6500 Bellinzona, Switzerland. ⁵KU Leuven Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, B-3000 Leuven, Belgium. ⁶KU Leuven Department of Imaging and Pathology, Translational Cell and Tissue Research, B-3000 Leuven, Belgium. ⁷Vir Biotechnology, San Francisco, CA 94158, USA. ⁸Howard Hughes Medical Institute, Seattle, WA 98195, USA. ⁹Clinic of Internal Medicine and Infectious Diseases, Clinica Luganese Moncucco, 6900 Lugano, Switzerland. ¹⁰Faculty of Biomedical Sciences, Università della Svizzera Italiana, 6900 Lugano, Switzerland. ¹¹Department of Internal Medicine, Ente Ospedaliero Cantonale, 6500 Bellinzona, Switzerland. ¹²Prince of Wales Hospital Clinical School, University of New South Wales, Sydney, New South Wales 2052, Australia. ¹³Division of Clinical Pharmacology and Toxicology, Institute of Pharmacological Sciences of Southern Switzerland, Ente Ospedaliero Cantonale, 6900 Lugano, Switzerland. ¹⁴Clinical Trial Unit, Ente Ospedaliero Cantonale, 6500 Bellinzona, Switzerland. ¹⁵Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, 8091 Zurich, Switzerland. ¹⁶Global Virus Network, Baltimore, MD 21201, USA. ¹⁷National Institute of Molecular Genetics, 20122 Milano, Italy. *Corresponding author. Email: sallustf@ethz.ch (F.S.); junsiong.low@biol.ethz.ch (J.S.L.); dveesler@uw.edu (D.V.) †These authors contributed equally to this work. ‡These authors contributed equally to this work. §These authors contributed equally to this work.

Fig. 1. Rare broadly reactive memory B cells are elicited upon natural infection or vaccination. (A to D) Total PBMCs from COVID-19 convalescents

(C) (A), vaccinees with prior infection (preimmune) (VP) (B), and vaccinees without prior infection (naïve) (VN) (C) (table S5) were plated in replicate 96 wells (10^4 cells/well) and stimulated with the TLR7/8 agonist R848 ($2.5 \mu\text{g/ml}$) in the presence of IL-2 (1000 U/ml). Twelve days later, the supernatant of each culture was screened, in parallel, for the specificities of the secreted IgG antibodies to commercially available hCoV S proteins from SARS-CoV-2, SARS-CoV, OC43, HKU1, NL63, and 229E by ELISA. The skyline plot provides a detailed view of the specificities of each culture well (represented in rows) to the respective antigens (in subcolumn) from each donor (in column). The order of the antigens is indicated in the legend and uncoated plates (phosphate-buffered saline, PBS) were used as control. If the optical density at 405 nm (OD_{405}) value exceeded the cutoff value determined by average OD_{405} of PBS wells ± 4 standard deviation, the culture was considered reactive to the antigen and is indicated here with colored cells. Only cultures exhibiting reactivity to at least one antigen are shown. Red cells highlight the six cultures exhibiting reactivity to all alpha and beta hCoV S proteins tested.

(D) Cross-reactivity patterns of SARS-CoV-2 S-positive cultures from (A) to (C) are shown as pie charts. The total number of SARS-CoV-2 S-positive cultures from each cohort is indicated at the center of the pie. (E) Using a two-step screening strategy as depicted in fig. S1, 16 mAbs that cross-reacted with multiple hCoV S proteins were isolated from 10 donors. Shown are EC_{50} values to the respective commercially available hCoV S proteins measured by ELISA. mAbs are grouped on the basis of reactivity patterns. Shown is one representative experiment out of at least two performed. Group 2 mAbs C21G12, C21D10, and C21E3 were described in a separate study (15) (designated P34G12, P34D10, and P34E3, respectively).



well as bat sarbecovirus WIV-1 and merbecovirus PDF-2180 (40) S-pseudotyped viruses (Fig. 2C and fig. S5, A to C). C77G12 neutralized all betacoronavirus S pseudotype viruses tested with the highest relative potency for SARS-CoV-2 (Fig. 2C and fig. S5, A

to C). The SARS-CoV-2-neutralizing activity of VN01H1 and C77G12 was reduced by the addition of soluble fusion peptide (fig. S5D). Furthermore, VN01H1, C77G12, and C28F8 also inhibited SARS-CoV-2 S-mediated cell fusion (Fig. 2D), suggesting that fusion

peptide-specific mAbs could prevent S proteolytic activation or fusogenic rearrangements, thereby inhibiting membrane fusion and viral entry.

We next assessed the protective efficacy of the VN01H1 and C77G12 mAbs in vivo. In the Syrian hamster model of SARS-CoV-2 P.1

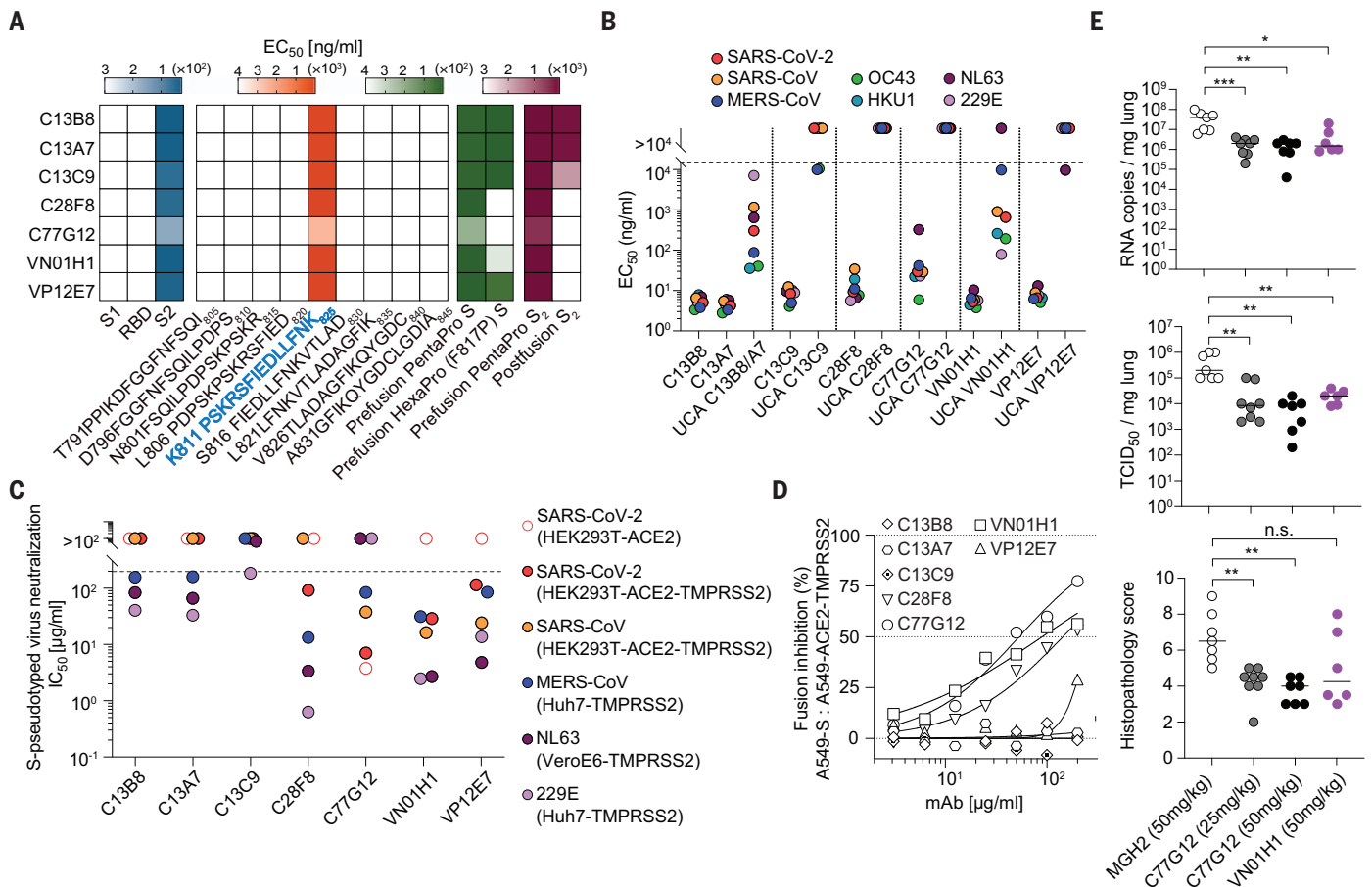


Fig. 2. Broadly reactive fusion peptide-specific mAbs acquire affinity and breadth through somatic mutations. (A) Binding profiles of the seven group small Group 3 mAbs. Epitope mapping was performed using commercially available SARS-CoV-2 protein domains (S₁, RBD, and S₂), 15-mer overlapping synthetic peptides with the indicated sequences, prefusion HexaPro (F817P) S, prefusion PentaPro (F817) S, prefusion PentaPro (F817) S₂, and postfusion S₂ coated onto plastic and assessed by ELISA (log₁₀ EC₅₀ shown). The mAbs all recognized the fusion peptide sequence highlighted in blue. (B) Binding evaluation of group 3 mAbs and their respective germline reverted UCAs to different hCoV S proteins. C13B8 and C13A7 are sister clones and thus share a single UCA. One representative experiment of at least two is shown. (C) Half-maximum inhibitory concentrations (IC₅₀s) of group 3 mAbs against SARS-CoV-2, SARS-CoV, MERS-CoV, NL63, and 229E S-pseudotyped viruses in the indicated target

cell lines. For each hCoV pseudotyped assay, all mAbs were compared in parallel. One representative experiment of at least two is shown. (D) Group 3 mAbs were assessed for their ability to inhibit the fusion of SARS-CoV-2 S-expressing (A549-S) and ACE2-TMPRSS2-expressing (A549-ACE2-TMPRSS2) A549 cells. Inhibition of fusion values are normalized to the percentage of fusion without mAb (100%) and to that of fusion of non-transfected cells (0%). One representative experiment of two is shown. (E) The prophylactic efficacy of VN01H1 (50 mg/kg), C77G12 (25 and 50 mg/kg), and the negative control malaria-specific mAb MGH2 (50 mg/kg) were tested in hamsters challenged with the SARS-CoV-2 P.1 (Gamma) variant of concern. Viral RNA loads (top), replicating virus titers (middle), and histopathological scores (bottom) are shown. $P > 0.05$ (non-significant, n.s.), $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$, Mann-Whitney test corrected using Bonferroni multiple comparisons.

(Gamma) infection, prophylactic administration of either mAb at high doses reduced viral RNA copies and lung titers and ameliorated lung pathology at statistically significant levels (Fig. 2E). These findings demonstrate that fusion peptide-specific mAbs can reduce viral burden in vivo, albeit with moderate potency.

Structural basis for fusion peptide recognition

To gain insight into the epitope recognized by fusion peptide-specific broadly reactive coronavirus mAbs, we performed substitution scan and structural analysis on the six clonally unrelated mAbs. All mAbs bound to a core motif comprising I818EDLLFNK825 (fig. S6A), with C28F8, C77G12, VN01H1, and VP12E7 having a

three-amino acid expanded footprint spanning the N-terminal R815SF817 residues, which includes the S₂' cleavage site at position R815.

We determined the crystal structures of five Fabs (C13B8, C13C9, C77G12, VN01H1, and VP12E7) in complex with the K811PSKRSFIEDLLFNK825 fusion peptide at 2.1 Å, 2.1 Å, 1.7 Å, 1.86 Å, and 2.5 Å, resolution, respectively (Fig. 3; fig. S6, B to D; and table S3). All five Fabs bind to overlapping epitopes in the fusion peptide through interactions involving the heavy- and light-chains. The strict conservation and the conservative substitution of key residues involved in mAb recognition (R815, S816, I818, E819, D820, L821, L822, F823, N824, and K825) across the *Orthocoronavirinae* subfamily ex-

plains the broad cross-reactivity of these fusion peptide-specific mAbs (figs. S3 and S6 and table S4). The overall architecture of the fusion peptide in the C13C9-bound complex structure is most similar to that observed in the prefusion S trimer (PDB 6VXX) (4) (Fig. 3C and fig. S6B). In the two structures determined in complex with VN01H1 or VP12E7, the SARS-CoV-2 S fusion peptide adopts a similar conformation, which is distinct from the conformation observed in prefusion SARS-CoV-2 S trimeric structures (4) (Fig. 3, B and C, and fig. S6C). Specifically, residues 813SKR815 refold from an extended conformation in prefusion S to an α -helical conformation in the two Fab-bound peptide structures, thereby

extending the α -helix found at the N-terminal region of the fusion peptide. In the C77G12-bound complex structure, the fusion peptide residues P812 to R815 adopt an extended conformation, distinct from prefusion S (Fig. 3, A and C), whereas these residues are disordered in the C13B8-bound structure (Fig. 3C and fig. S6D). The conserved residue R815, which is the S₂' site of proteolytic processing upon receptor binding for membrane fusion activation, is engaged in electrostatic interactions with the C13C9, VN01H1, VP12E7, and C77G12 Fabs and is therefore buried at the interface with their paratopes (table S4). Because preincubation of a soluble, native-like SARS-CoV-2 S ectodomain trimer with fusion peptide-specific mAbs did not prevent S2X58-induced triggering of fusogenic conformational changes (41, 42) (fig. S7), these mAbs likely inhibit TMPRSS2 cleavage of the S₂' site (through steric hindrance) and in turn activation of membrane fusion. Although residue F817 is not part of the epitope (table S4), the F817P substitution present in the HexaPro construct likely prevents the adoption of the extended α -helical conformation observed in the structures bound to VN01H1 or VP12E7 (because of restricted backbone torsion angles), resulting in dampened binding in ELISA (Fig. 2A and fig. S2D).

Superimposition of the Fab-fusion peptide complexes with available prefusion S structure (PDB 6VXX) (4, 5) revealed that the targeted epitope is buried toward the core of the S trimer and is therefore inaccessible (Fig. 3, D and E, and fig. S6, B to D), likely explaining the lack of detectable complexes of Fabs with prefusion S trimers during single-particle electron microscopy analysis. Taken together, these findings suggest that the epitope recognized by fusion peptide-specific mAbs is cryptic and may become accessible only transiently (43).

The fusion peptide is unmasked by ACE2 binding

To investigate the neutralization mechanism of fusion peptide-specific mAbs, we transfected human embryonic kidney 293T (HEK293T) cells to express coronavirus S proteins on their cell surface. All fusion peptide-specific mAbs showed only marginal binding to SARS-CoV-2 S-expressing HEK293T cells compared with control mAbs targeting the RBD (C94) (fig. S2C) or the stem helix (C21E3) (Fig. 4A and fig. S8A). However, the addition of soluble ACE2 enhanced binding of all fusion peptide-specific mAbs to SARS-CoV-2 S to levels comparable to that of the C94 and C21E3 control mAbs, suggesting that receptor engagement induces a conformational change that exposes the cryptic fusion peptide epitope (Fig. 4, A and B, and fig. S8A). This ACE2-dependent enhancement of binding was also observed with SARS-CoV-2 S harboring a deletion in the polybasic S₁/S₂

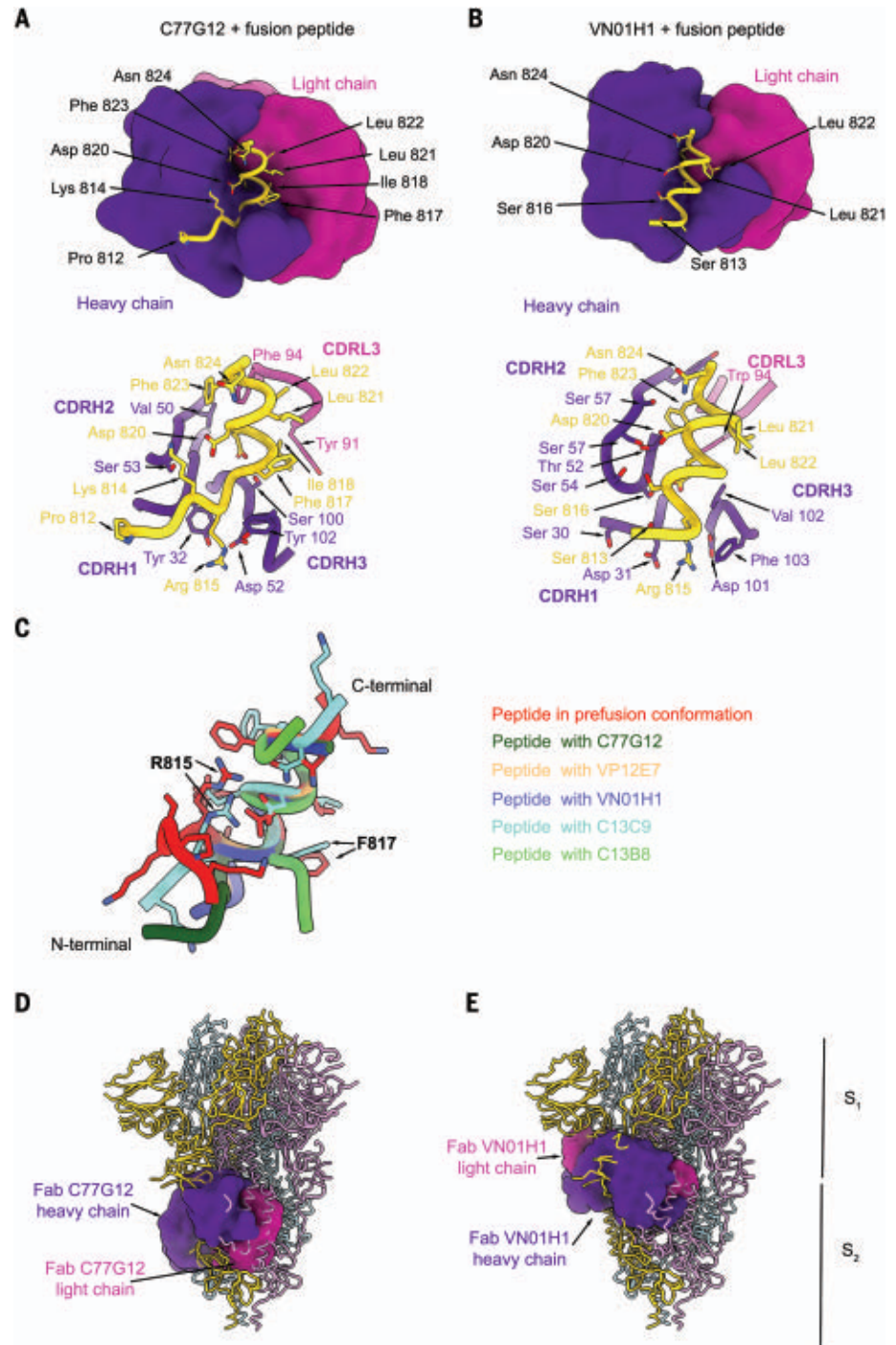


Fig. 3. Fusion peptide-specific antibodies target a cryptic epitope. (A and B) (Top) Surface representation of the crystal structures of the C77G12 (A) and VN01H1 (B) Fabs in complex with SARS-CoV-2 fusion peptide epitope. (Bottom) Ribbon representation of the corresponding structures highlighting the interactions of Fab heavy- and light-chain complementarity-determining regions with the fusion peptide (selected regions are shown for clarity). (C) Ribbon representation of the fusion peptides in the Fab-bound complexes superimposed with the fusion peptide in prefusion SARS-CoV-2 S (PDB 6VXX). (D and E) Superimposition of the C77G12-bound (D) or VN01H1-bound (E) fusion peptide structures to prefusion SARS-CoV-2 S uncovering the cryptic nature of the epitope. The Fabs are shown as surfaces, and S is rendered as ribbons. Each SARS-CoV-2 S protomer is colored distinctly (light blue, pink, and gold), and Fab heavy and light chains are colored purple and magenta, respectively.

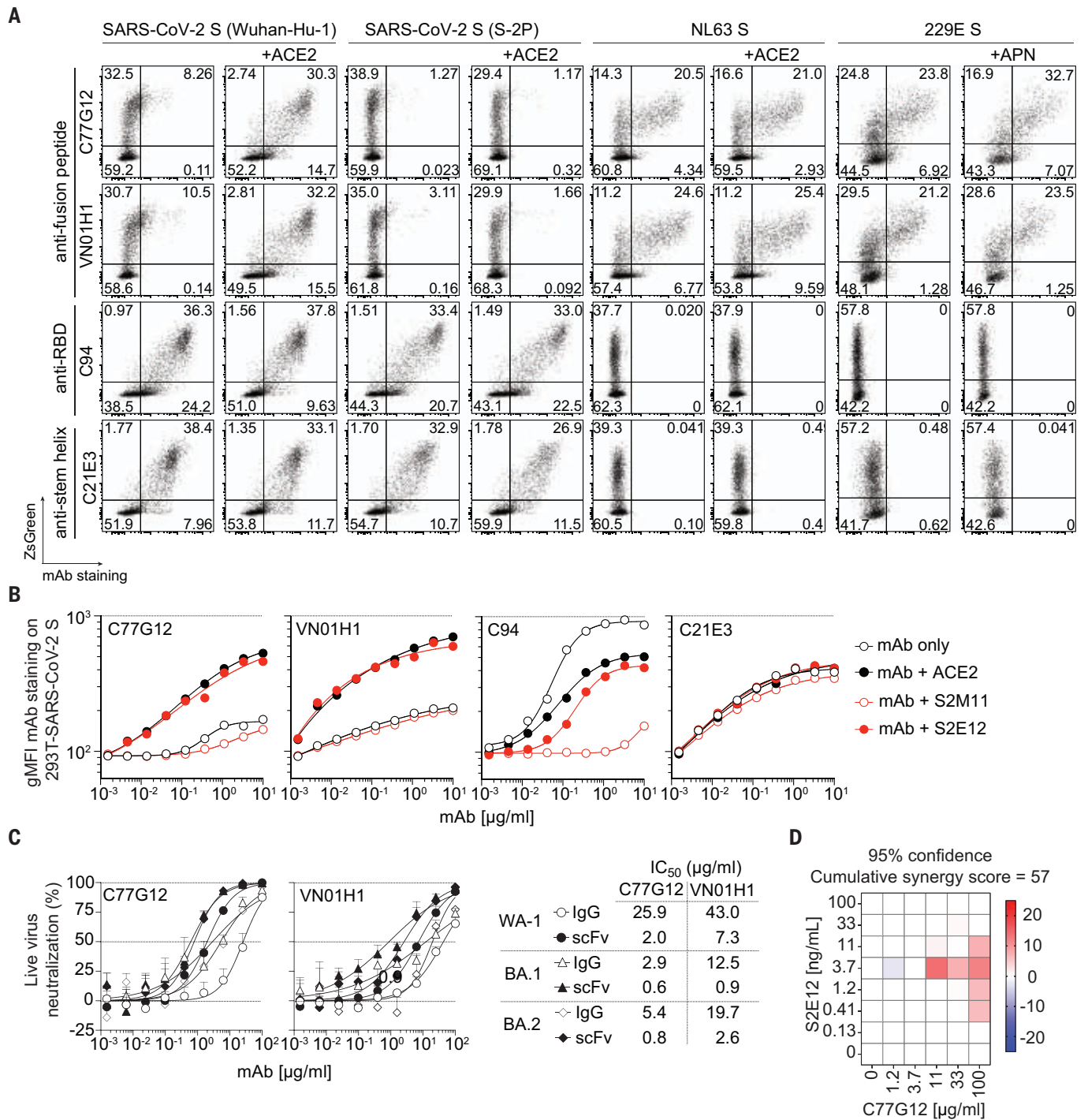


Fig. 4. The SARS-CoV-2 fusion peptide is unmasked after ACE2 receptor engagement. (A) Binding of fusion peptide-specific mAbs VN01H1 and C77G12 (8 µg/ml), as well as RBD-specific C94 and stem helix-specific C21E3 mAbs (8 µg/ml) on HEK293T cells transiently cotransfected with plasmid encoding ZsGreen and plasmids encoding full-length SARS-CoV-2 S Wuhan-Hu-1 (GenBank: NC_045512), SARS-CoV-2 S-2P (K986P, V987P), NL63 S (GenBank: APF29071.1), or 229E S (GenBank: APT69883.1) in the presence or absence of respective recombinant receptors ACE2 or APN, as measured by flow cytometry. (B) Titrating doses of fluorescently labeled mAbs were coincubated with HEK293T cells expressing full-length SARS-CoV-2 S for 2 hours at room temperature in the presence or absence of recombinant ACE2-mFc (27 µg/ml), ACE2-mimicking mAb S2E12 (20 µg/ml), or S2M11 (20 µg/ml). The latter mAb

locks the S trimer in a closed conformation. The line in the scatter plot is a reference for the maximum geometric mean fluorescence intensity (gMFI) of staining with control anti-RBD C94 mAb. One representative experiment of two is shown. (C) VN01H1 and C77G12 in full-length IgG or scFv formats were tested for their neutralization of authentic SARS-CoV-2 Washington-1 and Omicron BA.1 and BA.2 variants. IC₅₀ values are displayed as a table. (D) Synergy experiment performed against SARS-CoV-2 (Wuhan-Hu-1) S-pseudotyped virus (0.1 MOI) on VeroE6-TMPRSS2 at the indicated concentrations of S2E12 and C77G12 mAbs alone or in combination in a checkerboard manner. Analysis of mAb synergy was performed using MacSynergy II (70). Synergy plots adjusted with Bonferroni correction at 95% confidence were used for reporting. A synergy score between 50 and 100 was defined as moderate.

cleavage site (Δ PRRA) (44, 45) (fig. S8A), suggesting that these mAbs can bind before S₁ shedding or S fusogenic refolding to post-fusion state, consistent with the preferential binding to prefusion S₂ relative to postfusion S₂ (fig. S2A). However, the 2P (K986P, V987P) prefusion-stabilizing mutations (4, 5) that lie outside of the epitope (Fig. 4A and fig. S8A) inhibited ACE2-mediated enhancement of mAb binding, implying an impediment of receptor-induced allosteric conformational changes, consistent with recent findings (46). Enhanced binding of fusion peptide-specific mAbs was also observed with SARS-CoV S- and MERS-CoV S-expressing HEK293T cells in the presence of their respective receptors ACE2 (47) and DPP4 (48). By contrast, fusion peptide-specific mAbs bound efficiently to HEK293T cells displaying the alphacoronaviruses NL63 and 229E S independently of receptor engagement by ACE2 (49) or APN (50), respectively (Fig. 4A and fig. S8A). We observed that VN01H1 and C77G12 neutralized authentic SARS-CoV-2 Omicron BA.1 and BA.2 more potently than Washington-1, likely because of increased accessibility of the fusion peptide in Omicron S (46). Furthermore, the neutralization potencies of VN01H1 and C77G12 could be further improved by engineering mAbs with smaller formats (such as scFv), suggesting possible limited accessibility of the fusion peptide region (Fig. 4C).

It was previously shown that certain mAbs can mimic ACE2 binding and trigger fusogenic activity (51, 52). Indeed, the addition of S2E12, an ACE2-mimicking mAb, but not S2M11, a mAb that locks SARS-CoV-2 S trimer in the closed state (51), enhanced binding of the fusion peptide-specific mAbs (Fig. 4B). Consistent with this finding, we observed a synergy in pseudovirus neutralization when C77G12 was used in combination with S2E12 (Fig. 4D), suggesting that fusion peptide-specific mAbs may be more effective in the context of a polyclonal antibody response against the RBD. Collectively, these results indicate that the fusion peptide epitope only becomes accessible upon receptor-induced S conformational changes in sarbecovirus SARS-CoV-2, SARS-CoV, and merbecovirus MERS-CoV, whereas it is more readily accessible in the alphacoronaviruses NL63 and 229E, possibly because of molecular breathing to different extents in these S trimers.

Discussion

Using a high-throughput screening of memory B cells from immune donors with a panel of S proteins, we identified a new class of mAbs that target the fusion peptide region. These mAbs have different binding modalities, some of which exhibit unprecedented cross-reactivity to the fusion peptide of all four coronavirus genera. The findings that the UCAs of some of these mAbs bind preferentially to common cold coronaviruses and that the mAbs acquire

affinity and breadth through somatic mutations suggest that their elicitation may require multiple rounds of selection, possibly by heterologous coronavirus infections. A complex developmental pathway has been reported for HIV-1-neutralizing mAbs and may be a general requirement for mAbs that recognize highly conserved epitopes (53, 54).

Previous studies have identified serum antibodies binding to the fusion peptide of SARS-CoV-2 and showed, through depletion or peptide inhibition experiments, that such antibodies can contribute to the serum-neutralizing activity in a polyclonal setting (34–38). We show here that some fusion peptide-specific mAbs have direct neutralizing activity in vitro against alpha- and beta-coronaviruses and can ameliorate pathology and viral burden in vivo at high doses. Although the neutralizing activity of these mAbs is low when used alone, it is possible that in the context of a polyclonal response, they may synergize with other antibodies that favor the exposure of the fusion peptide region, as shown here with S2E12.

The biochemical events associated with S₂' proteolytic processing for SARS-CoV, SARS-CoV-2, and MERS-CoV that are critical for unmasking the fusion peptide region have been described previously (55–59). The mAbs isolated in this study provide tools to characterize allosteric conformational changes in S triggered by receptor binding or by receptor-mimicking mAbs (51, 52), a process that has thus far eluded structural definition. These events lead to exposure of the fusion peptide region for proteolytic cleavage by TMPRSS2 or endosomal cathepsins preceding membrane fusion (32). The finding that this region is readily accessible in alphacoronaviruses is consistent with the higher neutralizing activity of some fusion peptide-specific mAbs against NL63 and 229E and proposes that conformational changes might occur spontaneously as a consequence of molecular breathing.

Broadly neutralizing mAbs recognizing the fusion peptide region of influenza virus hemagglutinin and HIV-1 gp120 have guided the design of universal vaccine candidates against these highly variable pathogens (60–66). Similarly, the conservation of the fusion peptide region across the *Orthocoronavirinae* subfamily provides a rationale to develop a vaccine capable of inducing fusion peptide-specific responses that may be less prone to viral escape because of the potential fitness loss of escape mutants (67, 68). The finding that prefusion stabilizing 2P mutations (4, 5) abolish receptor-induced exposure of the fusion peptide suggests that current mRNA vaccines may be suboptimal in inducing antibody responses against this region. The broadly neutralizing coronavirus mAbs described in this study may be used as probes to design immunogens that can better

unmask the fusion peptide region and elicit broadly protective antibody responses. Notably, the fusion peptide region was also found to stimulate broadly reactive CD4 T cells (69), providing a cue for intramolecular help in the generation of such antibodies. In conclusion, our study identifies neutralizing mAbs with unprecedented breadth and sheds light on a receptor-triggered conformational change that limits the immunogenicity of this conserved region, with potential impacts on universal coronavirus vaccine design.

REFERENCES AND NOTES

1. A. N. Vlasova et al., *Clin. Infect. Dis.* **74**, 446–454 (2022).
2. J. A. Lednický et al., *Nature* **600**, 133–137 (2021).
3. M. A. Tortorici et al., *Cell* **185**, 2279–2291.e17 (2022).
4. A. C. Walls et al., *Cell* **181**, 281–292.e6 (2020).
5. D. Wrapp et al., *Science* **367**, 1260–1263 (2020).
6. L. Piccoli et al., *Cell* **183**, 1024–1042.e21 (2020).
7. A. J. Greaney et al., *Cell Host Microbe* **29**, 44–57.e9 (2021).
8. A. Z. Wec et al., *Science* **369**, 731–736 (2020).
9. D. Pinto et al., *Nature* **583**, 290–295 (2020).
10. T. N. Starr et al., *Nature* **597**, 97–102 (2021).
11. C. A. Jette et al., *Cell Rep.* **36**, 109760 (2021).
12. M. A. Tortorici et al., *Nature* **597**, 103–108 (2021).
13. Y.-J. Park et al., *Science* **375**, 449–454 (2022).
14. D. R. Martínez et al., *Sci. Transl. Med.* **14**, eabj7125 (2022).
15. D. Pinto et al., *Science* **373**, 1109–1116 (2021).
16. C. Wang et al., *Nat. Commun.* **12**, 1715 (2021).
17. G. Song et al., *Nat. Commun.* **12**, 2938 (2021).
18. M. M. Sauer et al., *Nat. Struct. Mol. Biol.* **28**, 478–486 (2021).
19. P. Zhou et al., *Sci. Transl. Med.* **14**, eabi9215 (2022).
20. P. Zhou et al., bioRxiv 479488 [Preprint] (2022); <https://doi.org/10.1101/2022.03.04.479488>.
21. D. Corti et al., *Science* **333**, 850–856 (2011).
22. M. M. Sajadi et al., *Cell* **173**, 1783–1795.e14 (2018).
23. H.-X. Liao et al., *Nature* **496**, 469–476 (2013).
24. J. F. Scheid et al., *Science* **333**, 1633–1637 (2011).
25. J. Huang et al., *Nature* **491**, 406–412 (2012).
26. C. Dreyfus et al., *Science* **337**, 1343–1348 (2012).
27. D. Sok, D. R. Burton, *Nat. Immunol.* **19**, 1179–1188 (2018).
28. D. Corti, A. Lanzavecchia, *Annu. Rev. Immunol.* **31**, 705–742 (2013).
29. D. Pinna, D. Corti, D. Jarrossay, F. Sallusto, A. Lanzavecchia, *Eur. J. Immunol.* **39**, 1260–1270 (2009).
30. A. C. Walls et al., *Cell* **185**, 872–880.e3 (2022).
31. E. M. Anderson et al., *Cell* **184**, 1858–1864.e10 (2021).
32. M. Hoffmann et al., *Cell* **181**, 271–280.e8 (2020).
33. C.-L. Hsieh et al., *Science* **369**, 1501–1505 (2020).
34. E. Shrock et al., *Science* **370**, eabd4250 (2020).
35. K. W. Ng et al., *Science* **370**, 1339–1343 (2020).
36. C. M. Poh et al., *Nat. Commun.* **11**, 2806 (2020).
37. C. Daniel et al., *J. Virol.* **67**, 1185–1194 (1993).
38. H. Zhang et al., *J. Virol.* **78**, 6938–6945 (2004).
39. A. C. Walls et al., *Nature* **531**, 114–117 (2016).
40. Q. Xiong et al., bioRxiv 477490 [Preprint] (2022); <https://doi.org/10.1101/2022.01.24.477490>.
41. F. A. Lempp et al., *Nature* **598**, 342–347 (2021).
42. H. V. Dang et al., *Nat. Struct. Mol. Biol.* **26**, 980–987 (2019).
43. A. C. Walls et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–11162 (2017).
44. B. A. Johnson et al., *Nature* **591**, 293–299 (2021).
45. M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, *Mol. Cell* **78**, 779–784.e5 (2020).
46. S. M.-C. Gobeil et al., *Mol. Cell* **82**, 2050–2068.e6 (2022).
47. W. Li et al., *Nature* **426**, 450–454 (2003).
48. V. S. Raj et al., *Nature* **495**, 251–254 (2013).
49. H. Hofmann et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7988–7993 (2005).
50. C. L. Yeager et al., *Nature* **357**, 420–422 (1992).
51. M. A. Tortorici et al., *Science* **370**, 950–957 (2020).
52. A. C. Walls et al., *Cell* **176**, 1026–1039.e15 (2019).
53. D. R. Burton, *Nat. Rev. Immunol.* **19**, 77–78 (2019).
54. M. Bonsignori et al., *Immunol. Rev.* **275**, 145–160 (2017).
55. J.-E. Park et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12262–12267 (2016).
56. J. K. Millet, G. R. Whittaker, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15214–15219 (2014).
57. P. V. Raghuvamsi et al., *eLife* **10**, e63646 (2021).
58. S. Yu et al., *Proc. Natl. Acad. Sci. U.S.A.* **119**, e211199119 (2022).
59. E. Qing et al., *Cell Rep.* **39**, 110786 (2022).
60. C. Cheng et al., *PLOS ONE* **14**, e0215163 (2019).
61. K. Xu et al., *Nat. Med.* **24**, 857–867 (2018).

62. R. Kong *et al.*, *Science* **352**, 828–833 (2016).
 63. L. Ou *et al.*, *Sci. Rep.* **10**, 3032 (2020).
 64. R. Nachbagauer *et al.*, *Nat. Med.* **27**, 106–114 (2021).
 65. H. M. Yassine *et al.*, *Nat. Med.* **21**, 1065–1070 (2015).
 66. A. Impagliazzo *et al.*, *Science* **349**, 1301–1306 (2015).
 67. S. Belouzard, V. C. Chu, G. R. Whittaker, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5871–5876 (2009).
 68. I. G. Madu, S. L. Roth, S. Belouzard, G. R. Whittaker, *J. Virol.* **83**, 7411–7421 (2009).
 69. J. S. Low *et al.*, *Science* **372**, 1336–1341 (2021).
 70. M. N. Prichard, C. Shipman Jr., *Antiviral Res.* **14**, 181–205 (1990).

ACKNOWLEDGMENTS

We thank all study participants who donated blood and devoted time to our research; M. Biggiogero, A. Franzetti Pellanda, E. Picciocchi, T. Terrot, S. Tettamanti, T. Urbani, L. Vicari, and all personnel at the hospitals and nursing homes for providing blood samples; D. Vaqueirinho, S. Jovic, I. Giacchetto Sasselli, R. Schmidt, and X. Xi from the Sallusto laboratory for help with blood processing; M. Kopf (ETH Zurich) for providing the pLVX-puro-ACE2 transfer plasmid; H. Tani (University of Toyama) for providing the reagents necessary for preparing VSV pseudotyped viruses; M. Weisshaar (ETH Zurich) for his artistic rendition of fig. S1; and Y. Z. Tan from the Rubinstein laboratory for helpful discussion. **Funding:** The work was supported by the Henry Kreter Foundation (F.S.); the European Research Council (AdG no. 885539 ENGRAB to A.L.); the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (grants DP1AI158186 and HHSN272201700059C to D.V.); a Pew Biomedical Scholars Award (D.V.); Investigators in the Pathogenesis of Infectious Disease Awards from the Burroughs Wellcome Fund (D.V.); Fast Grants (D.V.); the Natural Sciences and Engineering Research Council of Canada (M.M.); University of Washington Arnold and Mabel Beckman cryoEM center; the National Institutes of Health (grant S100D032290 to D.V.); Beamline 5.0.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory; and EOC research funds. D.V. is an investigator of the Howard Hughes Medical Institute. F.S. and the Institute for Research in Biomedicine are supported by the Helmut Horten Foundation. **Competing interests:** J.S.L., J.J., F.S., A.L., and A.Ca. are currently listed as inventors on multiple patent applications, which disclose the subject matter described in this manuscript. A.L., D.C., F.S., J.N., M.M.-R., L.A.P., G.S., and D.P. hold shares in Vir Biotechnology. The Veselir laboratory and the Sallusto laboratory have received sponsored research agreements from Vir Biotechnology Inc. The remaining authors declare no competing interests. **Author contributions:** Conceptualization: J.S.L., M.A.T., A.L., D.V., F.S.; Funding acquisition: F.S., D.V., A.L.; Investigation: J.S.L., J.J., M.A.T., M.M., D.P., A.C., S.B., J.E.B., A.J., C.S., M.R., A.C.W., D.M., F.M., P.P., D.J., M.F., R.A., B.W., J.No., J.Ne., M.M., L.A.P., C.G., P.F., A.Ce.; Methodology: J.S.L., J.J., M.A.T., M.M., D.P., A.C., S.B., J.E.B., A.J., A.C.W., D.M., F.M., P.P., D.J., M.F., R.A., B.W., J.No.; Supervision: J.S.L., D.C., D.V., A.L., F.S.; Writing, review and editing: J.S.L., M.A.T., G.S., D.C., D.V., A.L., F.S. **Data and materials availability:** All data associated with this manuscript are available in the main text or the supplementary materials, including the FACS data gating strategy. The crystal structures were deposited to the Protein Data Bank with accession numbers listed in table S3. All further relevant source data that support the findings of this study are available from the corresponding authors upon reasonable request. Materials are available through materials transfer agreements (UBMTAs or similar agreements) with the Institute for Research in Biomedicine and the University of Washington. **License information:** This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq2679
 Materials and Methods
 Figs. S1 to S8
 Tables S1 to S5
 References (71–106)
 MDAR Reproducibility Checklist

Submitted 28 March 2022; accepted 3 July 2022
 10.1126/science.abq2679

PHYSICAL CHEMISTRY

Proton-coupled energy transfer in molecular triads

Belinda Pettersson Rimgard^{1†}, Zhen Tao^{2†‡}, Giovanni A. Parada^{2,3}, Laura F. Cotter², Sharon Hammes-Schiffer^{2*}, James M. Mayer^{2*}, Leif Hammarström^{1,*}

We experimentally discovered and theoretically analyzed a photochemical mechanism, which we term proton-coupled energy transfer (PCEnT). A series of anthracene-phenol-pyridine triads formed a local excited anthracene state after light excitation at a wavelength of ~400 nanometers (nm), which led to fluorescence around 550 nm from the phenol-pyridine unit. Direct excitation of phenol-pyridine would have required ~330-nm light, but the coupled proton transfer within the phenol-pyridine unit lowered its excited-state energy so that it could accept excitation energy from anthracene. Singlet-singlet energy transfer thus occurred despite the lack of spectral overlap between the anthracene fluorescence and the phenol-pyridine absorption. Moreover, theoretical calculations indicated negligible charge transfer between the anthracene and phenol-pyridine units. We construe PCEnT as an elementary reaction of possible relevance to biological systems and future photonic devices.

Coupling of proton transfer to electronic transitions exerts a great influence on the energetics and kinetics of these processes. Photoacids are well-studied examples, where light excitation leads to a rapid proton release and, in some cases, even produces long-lived pH changes and pH gradients (1, 2). Another example is excited-state intramolecular proton transfer (ESIPT), which generates an excited tautomer state of the light-absorbing molecule; systems where this state is fluorescent have applications for optoelectronic materials (3). Often, ESIPT instead allows for ultrafast dissipation of excess electronic energy as heat, and is therefore an important photostabilizing mechanism in plastics (4) as well as in biomolecules such as DNA (5).

More recently, the charge transfer mechanism of proton-coupled electron transfer (PCET) has generated great interest (6). PCET is a key reaction behind biological energy conversion in photosynthesis, respiration, and nitrogen fixation (7), as well as in synthetic chemical systems for photoredox catalysis and renewable energy conversion and storage (8–11). It is also involved in DNA synthesis and repair, photo-receptor function, and many other radical enzyme reactions (12–15). Of particular interest is the concerted PCET mechanism, where the electron and proton are transferred without the formation of high-energy intermediates, thereby offering a low-barrier reaction pathway.

In all the above reactions, the coupled transfer of the proton changes the energy landscape of the overall process, which obviously affects its dynamics and rate. There are also important effects of the much greater mass of the proton, whose wave function is more localized than that of the electrons involved, which may impose a greater sense of directionality on the process. Understanding and analyzing such effects, theoretically and experimentally, is a fundamentally interesting challenge and of practical importance.

In 2019, our groups used a series of anthracene-phenol-pyridine (An-PhOH-Py) triads (Fig. 1) (16) to demonstrate the first example of concerted PCET occurring in the Marcus inverted region (17). This behavior, where the rate decreases with increasing driving force, is important for slowing recombination after photochemical charge separation, but was previously considered unlikely for concerted PCET reactions because of the nearly barrierless tunneling to higher vibrational product states (18). Nonetheless, photoexcitation of the An subunit of the triads results in the formation of a local excited state (LES) on the anthracene (Eq. 1) (16, 19). In tenths of picoseconds, the LES forms the charge-separated state (CSS) by electron transfer from the PhOH group to the ¹*An concerted with proton transfer to the pyridine

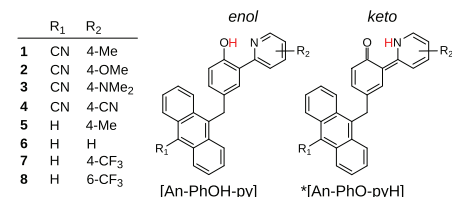


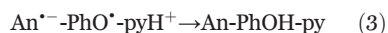
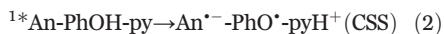
Fig. 1. Structures of the studied anthracene-phenol-pyridine ([An-PhOH-py]) triads. The enol form is the stable ground-state form, and the electronically excited keto form is the local electron-proton transfer (LEPT) state.

¹Department of Chemistry - Ångström Laboratory, Uppsala University, SE 75120 Uppsala, Sweden. ²Department of Chemistry, Yale University, New Haven, CT 06520, USA. ³Department of Chemistry, College of New Jersey, Ewing, NJ 08628, USA.

†These authors contributed equally to this work. ‡Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA.

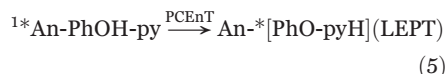
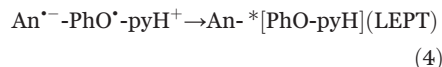
*Corresponding author. Email: sharon.hammes-schiffer@yale.edu (S.H.-S.); james.mayer@yale.edu (J.M.M.); leif.hammarstrom@kemi.uu.se (L.H.)

(Eq. 2). The charges subsequently recombine to the ground-state reactants (Eq. 3), once again in a concerted electron-proton transfer reaction but with a large driving force (>2.0 eV in CH_3CN). This process consequently places the reaction in the inverted region, as was shown by the rate dependence on driving force and solvent polarity.



The CSS intermediate was observed only in triads **1** to **3**, however; in triads **4** to **8**, a similar quenching of the LES did not result in an observable CSS (16). This result is puzzling, as the charge recombination driving force is larger for **4** to **8**, meaning the reaction would be more deeply in the inverted region and thus even slower than for **1** to **3**. Theoretical investigations have suggested the involvement of another state that would lie lower than the CSS in **4** to **8** (20, 21). This is the so-called local electron-proton transfer (LEPT) state, where the PhOH-py fragment is in a local, electronically excited tautomeric (keto) state, ${}^*\text{[PhO-pyH]}$, in which the phenolic proton

has transferred to the pyridine. The LEPT state is the same excited tautomeric state as that formed by ESIPT after excitation of the separate phenol-pyridine compound (22), and similar to other ESIPT compounds (3). This state was proposed to be populated after excitation in **4** to **8** (20, 21); it can possibly be populated from the CSS (Eq. 4) or be formed directly from the LES (Eq. 5) by a process here denoted proton-coupled energy transfer, PCEnT (see below).



Here, we present experimental and theoretical evidence for direct formation of the fluorescent LEPT state from the LES at 77 K in a butyronitrile glass, in which formation of the CSS is inhibited. In this experimental example of PCEnT, excitation energy was transferred from ${}^1\text{*An}$ to ${}^1\text{*[PhO-pyH]}$, coupled to proton transfer within the PhOH-py unit (Eq. 5). This finding was surprising, as it is a singlet-singlet transition without any detectable spectral overlap between the donor fluorescence and acceptor absorption. Moreover, PCEnT occurred with

negligible charge transfer between the An and PhOH-py units, which is in sharp contrast to PCET reactions. This result significantly expands the scope of concerted processes where proton transfer controls reactivity. Many excited-state charge and energy transfer reactions in proteins and DNA occur between hydrogen-bonded pigments, making PCEnT possibly important in natural systems. Yet such a reaction has, to the best of our knowledge, not previously been reported.

Spectroscopic observation of LEPT state

The lack of observed CSS formation in triads **4** to **8**, together with the computationally predicted excited-state keto tautomer (LEPT) (20), laid the basis for further investigation. The triad absorption spectra agree well with a sum of contributions from the separate components: a vibrationally structured anthracene band around 400 nm and a PhOH-py band around 330 nm (Fig. 2B and fig. S1). The LEPT state should fluoresce around 500 nm in aprotic, rigid media, as has been shown for the excited keto state, ${}^*\text{[PhO-pyH]}$ of the separate phenol-pyridine compound (22). Therefore, we examined the triads (**1**, **2**, **4** to **7**) by fluorescence spectroscopy in rigid butyronitrile glass at $T = 77$ K. We also studied these triads computationally and found that the minimum-

Fig. 2. Steady-state absorption (room temperature) and fluorescence (77 K) spectra in butyronitrile. (A) Schematic Jablonski diagram of the possible transitions at 77 K for the triads:

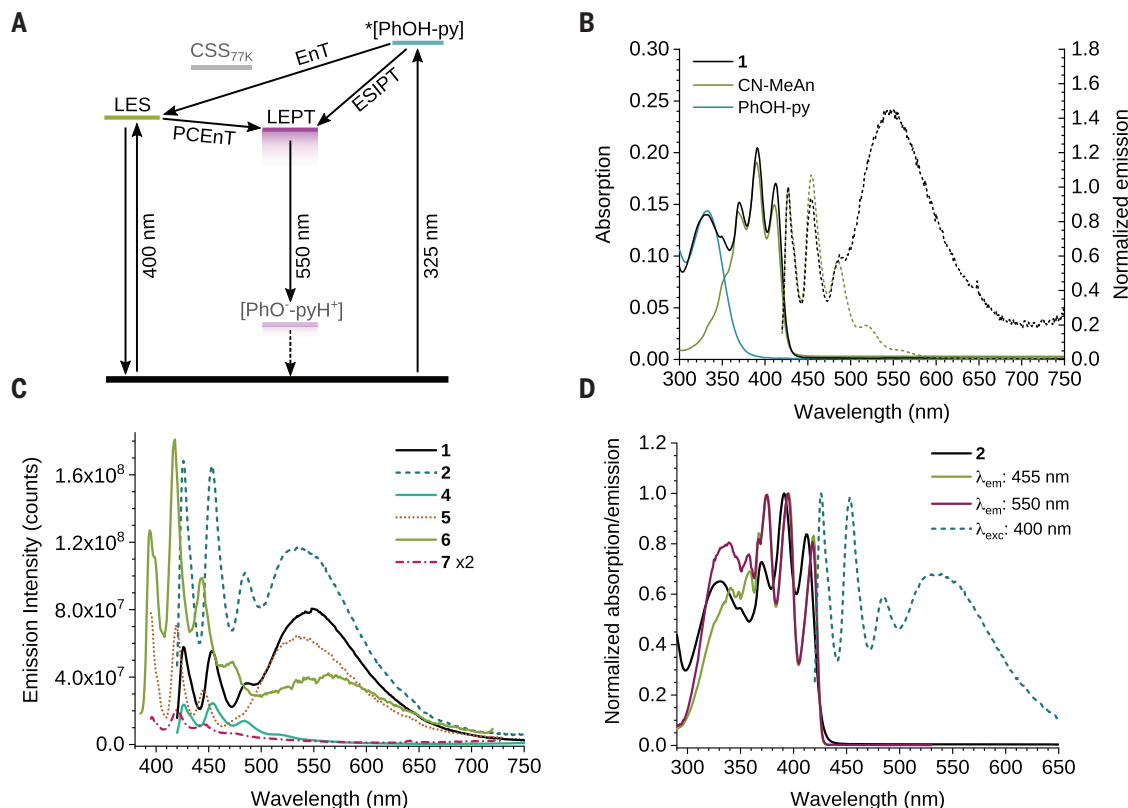
radiative decay (downward solid line) and nonradiative decay. The CSS energy depends strongly on the chemical environment (16, 19) and is higher than the LES at 77 K.

(B) Absorption (solid line) and fluorescence spectra normalized to the first An emission peak (dashed line, λ_{exc} : 400 nm) of **1** (black) and the reference compounds of similar concentrations (4 to 7 μM): 2,4-di-*tert*-butyl-6-(pyridin-2-yl)phenol (PhOH-py, blue) and 9-cyano-10-methylanthracene (CN-MeAn, green).

(C) Fluorescence spectra of triads **1**,

2, **4**, **5**, **6**, and **7** with

excitation at 400 nm (**1** and **2**), 365 nm (**6**), and 375 nm (**5** and **7**). Note that **7** was measured with twice the absorption at the excitation wavelength. (D) Absorption spectrum (black line), fluorescence excitation spectrum ($\lambda_{\text{em}} = 455$ and 550 nm, solid line) and fluorescence spectrum ($\lambda_{\text{exc}} = 400$ nm, dashed line) of **2**.



energy LEPT geometry exhibits a $\sim 90^\circ$ twist between the phenol and pyridine fragments, whereas the planar LEPT geometry is a first-order saddle point. This twisting, which is known to quench the fluorescence (22, 23), is expected to be hindered in the glass at 77 K. Thus, the computational studies were performed for both the planar and twisted LEPT geometries.

The 77 K fluorescence spectrum of **1**, after selective excitation of the An unit at 400 nm, showed a structured ^1An fluorescence at 420 to 580 nm, accompanied by a broad band with a maximum around 550 nm (Fig. 2B). The

latter is expected for the LEPT state (22) and was similarly observed for all measured triads with an electron-donating group on the pyridine: **1**, **2**, **5**, and **6** (Fig. 2, B and C). The LEPT emission shifted to lower energy when the electron-donating group was weaker ($2 > \mathbf{1} > \mathbf{6}$), consistent with the computational results in tables S8 and S12. Because **1** and **5** share the same functional group ($R_2 = 4\text{-Me}$) on the pyridine, the LEPT energy should be similar. Still, the peak for **5** was blueshifted by 11 nm, which indicated a small but significant coupling to the An unit, which differs in its functional group (CN in **1**; H in **5**) between these

triads. The relative intensity of the ^1An and LEPT fluorescence was independent of triad concentration (fig. S2), and control experiments with 9-CN-10-Me-An showed only the structured ^1An fluorescence (Fig. 2B). This excluded any excimer or other anthracene complex as responsible for the broad fluorescence around 550 nm.

Fluorescence excitation spectroscopy provided unambiguous evidence for LEPT state formation from excited ^1An (Fig. 2D). The shape of the excitation spectrum of the 550-nm LEPT emission peak (purple line) matched that of the absorption spectrum (black line), with

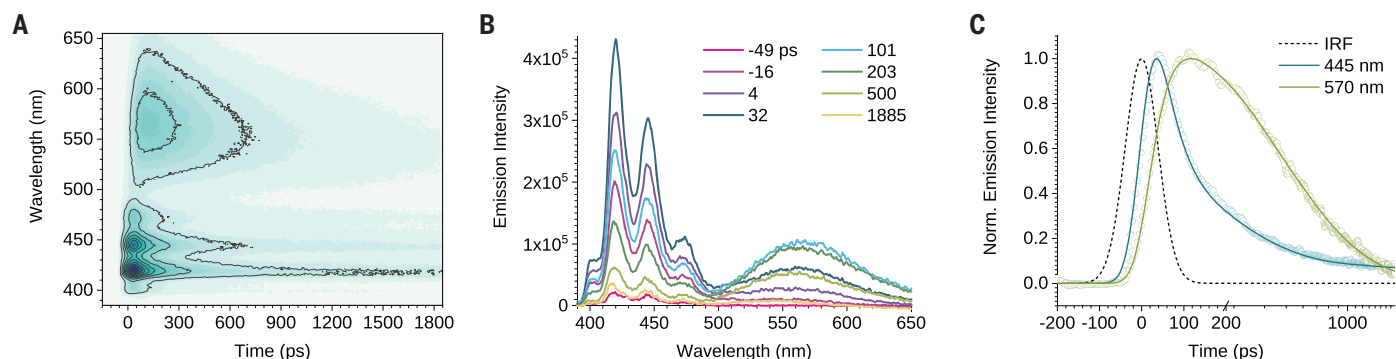


Fig. 3. Transient fluorescence emission spectra in butyronitrile at 77 K.

(A) Contour map of the transient fluorescence emission intensity counts (dark blue = highest intensity) of **6** in butyronitrile (0.2 mM) between 390 and 640 nm within a 2-ns time window. (B) Emission spectra at given times after excitation with 370-nm light. The ^1An fluorescence reaches a maximum at

~ 30 ps (instrumental response function $\text{IRF}_{\text{FWHM}} = 65$ ps) and decreases simultaneously with the rise of LEPT fluorescence that maximizes after ~ 140 ps. (C) Time-resolved fluorescence traces (lines, fit; circles, data) within the LES (445 nm, blue line) and LEPT (570 nm, green line) bands; the IRF is shown as a dashed line.

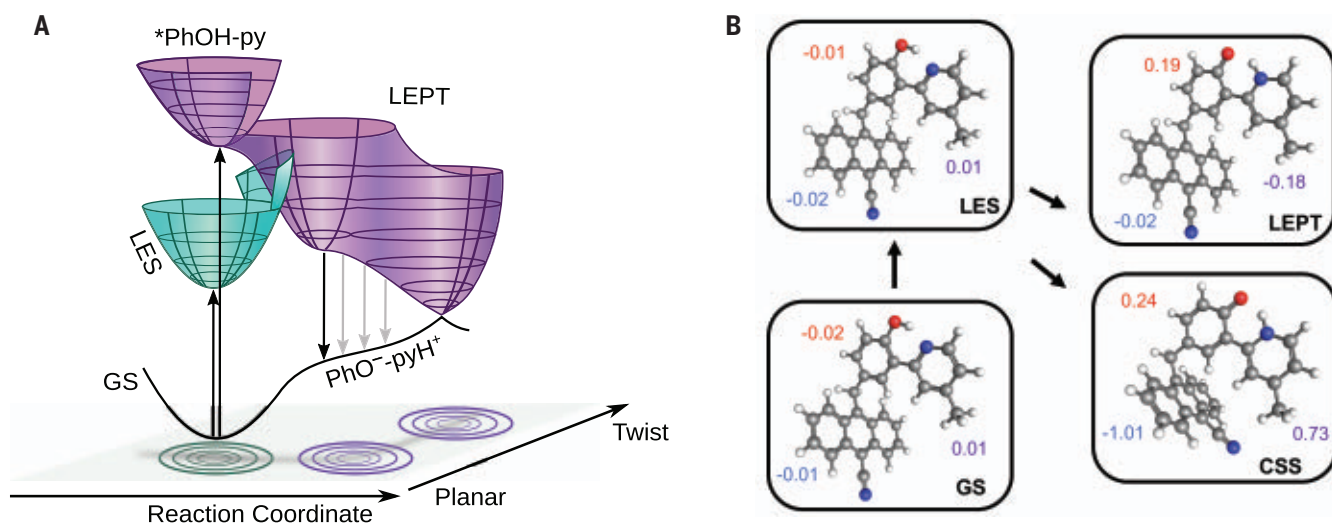


Fig. 4. Potential energy surfaces and optimized structures for the electronic states discussed.

(A) Schematic potential energy surfaces for the ESIP and PCEnT reactions, color-coded to mark the electronically excited fragment (green, An; purple, PhOH-py). Note that ESIP is barrierless on the excited-state potential energy surface (purple), whereas PCEnT follows the reorganization of the heavy nuclei including the solvent (green to purple). In fluid solution, the LEPT state will twist along the PhOH-py bond, where 90° leads to a conical intersection, and nonradiative decay. This twist is impeded in the 77 K butyronitrile

glass, although a partial twist could occur to lower the energy of the LEPT state relative to the LES (table S11). GS, ground state. (B) The optimized structures and the natural bond orbital [NBO (29)] charges on the anthracene (blue), phenol (red), and pyridine (purple) subunits of **1** for the GS, LES, CSS, and planar (not twisted, first-order saddle point) LEPT state in the gas phase. The charges reflect the electron density, where a more negative number equals a more electron-dense fragment (the methylene spacer between anthracene and phenol is not included in any of the fragments, and therefore those charges are excluded in this analysis).

contributions from both the PhOH-py (330 nm) and the An (300 to 430 nm) absorption bands. The excitation spectrum of the 450-nm ^1An emission (green line) showed instead a weaker relative contribution from the 330-nm PhOH-py band. The 550-nm LEPT state could thus be formed either by exciting the PhOH-py unit, which would cause formation of the LEPT state by the well-established ESIPT mechanism, or directly via excitation of the An unit followed by energy transfer from ^1An .

Evidence for proton-coupled energy transfer mechanism

The excitation energy transfer from the anthracene donor to the phenol-pyridine unit to yield the LEPT state is a type of mechanism that we denote proton-coupled energy transfer, PCEnT (Fig. 2A). Energy transfer directly to the $^*\text{[PhOH-py]}$ state would be strongly uphill, as is seen from its low-wavelength absorption spectrum around 330 nm, but proton transfer lowers the phenol-pyridine excited-state energy to make the PCEnT process energetically feasible. The quantum yield for the LEPT formation was deemed to be similar irrespective of which fragment was excited, on the basis of the good agreement between the shapes of the absorption spectrum (Fig. 2D) and the 550-nm excitation spectrum, where the respective PhOH-py and An amplitudes were fairly equal in the two spectra. This fact shows that PCEnT from ^1An was about as efficient as ESIPT from $^*\text{[PhOH-py]}$.

To follow the formation and decay of the LEPT state in real time, we performed ultra-

fast transient fluorescence spectroscopy measurements of **6** with excitation at 370 nm, in the mid- to red part of the An absorption band (Fig. 3). The observed features matched those seen in the steady-state experiments, and global fits were performed within the peak regions of the LES (380 to 490 nm) and LEPT (520 to 655 nm) fluorescence (Fig. 3B). The rise time of the LEPT peak was fitted to ~ 44 ps and a similarly short decay component could be found for the LES (46 ps), which reinforced the notion that the LEPT state could be populated directly from the LES (as depicted in the overall reaction scheme in Fig. 2A).

We can exclude LEPT formation via a sequential LES-CSS-LEPT reaction at 77 K, because the CSS would be substantially destabilized in a solvent glass (Fig. 2A) as a result of its greater dipole moment relative to both the LES and LEPT (Fig. 4B). Moreover, the solvent glass at 77 K prevented the large degree of twisting of the An group to reach the CSS optimized structure (Fig. 4B and table S1), which further destabilized the CSS energy. The CSS lies less than 0.2 eV below the LES at room temperature (16), and small-molecule charge separation is typically destabilized by more than 0.5 eV by freezing out solvent repolarization in polar solvents (24). Our gas-phase calculations in table S6 indeed placed the CSS higher than the LES. Moreover, the LEPT fluorescence was formed faster in **6** (~ 44 ps) than in **1** (~ 69 ps), despite greater room-temperature driving force for CSS formation (by ~ 0.5 eV) in the latter (16); this is clear experimental proof against a reaction involving the CSS.

The photophysics of triads **1** to **8** have been previously studied by transient absorption (TA) at ambient temperature (16), and **6** was also examined by time-resolved fluorescence spectroscopy at various temperatures (25). In the TA study at room temperature, the initial anthracene LESs in **1** to **3** decayed to a long-lived state (up to $\tau = 755$ ps) with spectral features in agreement with the CSS being formed by PCET, including obvious features of the anthracene anion. The same mechanism was assigned to the decay of the LES of **4** to **8** on the basis of the close correlation of LES decay rates with the free energies of PCET for each triad. These triads did not show a long-lived transient in the TA spectra at ambient temperatures. However, computations later placed the LEPT state below the CSS for **4** and **6** (20), indicating that rapid CSS-to-LEPT conversion and/or a contribution of PCEnT from the LES would be reasonable.

The previously reported spectrally resolved ultrafast fluorescence spectra of **6** at ambient temperatures showed an H/D kinetic isotope effect (KIE) = 4 and did not show any evidence of emission from a LEPT state. At lower temperatures, in a 2-methyl-tetrahydrofuran (2-Me-THF) glass or a polymethylmethacrylate (PMMA) film, the subnanosecond integrated fluorescence followed a single exponential with $\tau \approx 150$ ps, with an unexplained long-time scale (\gg ns) component. The data in 2-Me-THF were obtained in the presence of 3% methanol, which was observed to quench the LEPT emission in the current study. Therefore, the two studies are not directly comparable. In retrospect,

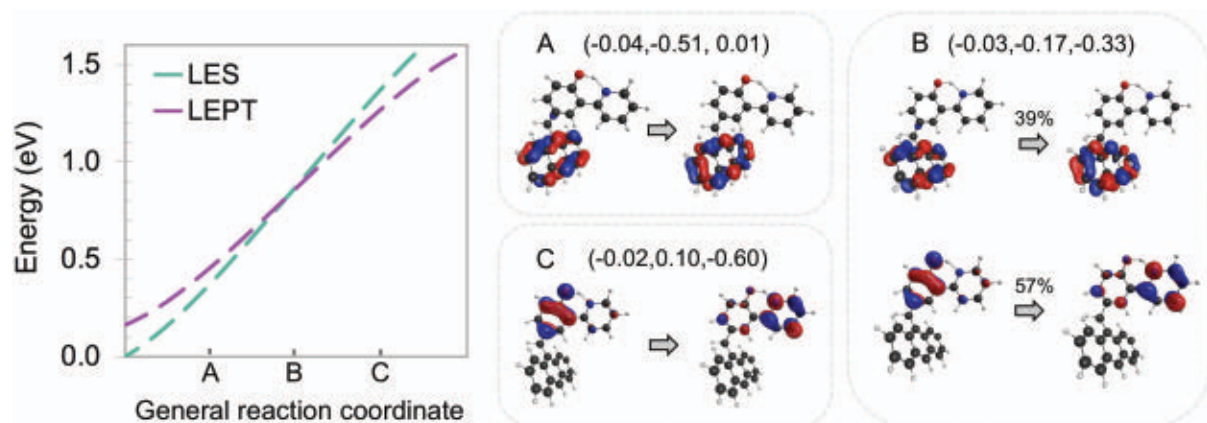


Fig. 5. Avoided-crossing region between the S_1 and S_2 adiabatic electronic states for triad **6, corresponding to a crossing between the LES and LEPT diabatic electronic states, and analysis of the S_1 excited state for relevant configurations.** The energy curves were obtained from an excited-state adiabatic molecular dynamics trajectory and are plotted along a unitless general reaction coordinate rather than time; this trajectory is far from equilibrium and is not directly comparable to experiment because of the reaction conditions (supplementary text). (A to C) Configurations obtained from this trajectory. The dominant natural transition orbitals (NTOs), corresponding to the transitions from occupied to

virtual orbitals that characterize the S_1 excited state, are shown for these three configurations. For (A), the NTOs are localized on the anthracene (LES), whereas for (C), the NTOs are localized on the phenol-pyridine fragment (LEPT state), where the proton has mostly transferred. For (B), an approximately equal mixture of LES and LEPT is observed, as indicated by the two NTOs. Above the orbital drawings of each configuration, the numbers are the charges (in a.u., the elementary charge of an electron) of the anthracene, phenol, and pyridine fragments for the S_1 state, in that order (excluding the charge of the CH_2 bridging group and the transferring hydrogen, which has a charge of ~ 0.5 a.u.).

these data could be consistent with the LES decaying either by PCET or by the new PCEnT pathway reported here in butyronitrile glasses; additional studies are needed.

Physical principles underlying proton-coupled energy transfer

The observation that PCEnT directly formed the excited tautomeric LEPT state was surprising and needs mechanistic clarification. As both the ^1An and the LEPT state are fluorescent, PCEnT is clearly a singlet-singlet process. However, there was no detectable spectral overlap between the donor fluorescence and acceptor ground-state absorption (Fig. 2B), which is a requirement for Förster (dipole-dipole) energy transfer. The absence of low-energy absorption of the keto form $[\text{PhO}^-\text{pyH}^+]$ is because the ground state, at the LEPT geometry, lies ~ 0.5 eV higher than the enolic ground state (table S4) and was thus not populated to any appreciable extent. Instead, as the reaction occurred at quite a small donor-acceptor distance (~ 6 Å between the central rings of An and PhOH), the coupling was presumably dominated by exchange and/or penetration terms (26). Any stronger coupling mechanism was probably prevented by the methylene spacer between the subunits.

We propose, therefore, that PCEnT could be described by a nonadiabatic surface crossing (Fig. 4A), in analogy to Marcus-type theoretical descriptions of PCET (6, 7). Thermal fluctuations of heavy nuclei lead to a configuration where electronic energy transfer and proton tunneling occur with energy conservation (crossing point of the reactant and product free energy curves). It is different from PCET, however, in that there is no charge transfer between donor and acceptor. The product state is a fluorescent excited state when phenol-pyridine twisting is hindered, and can thus be readily detected. Moreover, this mechanism differs from an ESIPT reaction, where proton transfer is driven by intramolecular charge redistribution, and the electronic coupling is usually strong, which leads to ultrafast reactions. The PCEnT reaction described here was instead driven by a remote, weakly coupled electronically excited state.

The rearrangement from the enol to the keto form of the phenol-pyridine acceptor unit is much greater than just the transfer of a proton, as it involves substantial charge redistribution and changes in bond lengths and angles. These combined effects change the excitation energy of the phenol-pyridine acceptor unit and allow it to come into resonance with that of the anthracene. This behavior is the fundamental basis for coupling of proton transfer to energy transfer and distinguishes PCEnT from typical energy transfer reactions. The details of this coupling and the role of coupling with

different proton vibrational states remain to be further explored.

Figure 5 demonstrates the PCEnT process observed in a molecular dynamics trajectory for triad 6 propagated on the S_1 state, which changed from LES to LEPT character as the proton transferred from the phenol to the pyridine. Representative configurations along this trajectory, along with an analysis of the fragment charges and dominant natural transition orbitals, show that the character of the S_1 state changed from LES to a mixture of LES and LEPT and then to LEPT. The nonadiabatic coupling between the S_1 and S_2 adiabatic electronic states (27, 28) exhibited a peak at the avoided crossing, providing evidence that this process is nonadiabatic. Throughout this process, the charge on the anthracene fragment remained virtually zero, indicating that the LES-to-LEPT transition did not involve charge transfer between anthracene and the phenol-pyridine group, consistent with the PCEnT mechanism. Moreover, the net charge transfer from the phenol to the pyridine was also nearly zero after accounting for the charge of the transferring proton. In the framework of the vibronically nonadiabatic PCET theory (6), the transferring proton would be treated quantum mechanically and would tunnel during the LES-to-LEPT transition.

Conclusions and outlook

We have experimentally proved the involvement of the excited $^*[\text{PhO-pyH}]$ (LEPT) state, in the photochemical reactions of these triad molecules. This mechanism competes with PCET from the LES and could explain why the CSS was not observed in 4 to 8. The LEPT state was formed by a novel proton-coupled energy transfer (PCEnT) mechanism. This is a singlet-singlet exchange energy transfer, coupled to proton transfer in the acceptor unit, that can be described as a nonadiabatic transition to the excited product state. Its free energy barrier in turn is determined by the reorganization energy associated with the solvent and other heavy nuclei, similar to Marcus ET theory or PCET theories (6, 17). In contrast to PCET, there is negligible charge transfer between donor and acceptor.

This mechanism can have several intriguing implications for both natural and synthetic systems. PCEnT may allow for the photosensitization of otherwise dark states, such as seen here by the LEPT formation from the low energy excitation of the LES, rather than the direct excitation of the PhOH-py unit at higher energy (ca 330 nm). Moreover, as energy transfer occurs over long distances but proton transfer is short-range, we suggest that this mechanism may be of importance to direct energy migration in e.g., light-harvesting systems and photonic materials, where the PCEnT acceptor would form a low-energy site. PCEnT

could also form the basis of a molecular switch for energy transfer, where proton transfer could be controlled independently by hydrogen-bonding or (de)protonation, thus effectively turning energy transfer on or off. We speculate that these processes could already be operational in natural systems with hydrogen-bonded light-absorbers, such as porphyrins in photosynthetic assemblies, DNA and flavin proteins, but have yet to be identified.

REFERENCES AND NOTES

- Y. Liao, *Acc. Chem. Res.* **50**, 1956–1964 (2017).
- L. Schulte, W. White, L. A. Renna, S. Ardo, *Joule* **5**, 2380–2394 (2021).
- A. P. Demchenko, K.-C. Tang, P.-T. Chou, *Chem. Soc. Rev.* **42**, 1379–1408 (2013).
- J. Catalan *et al.*, *J. Am. Chem. Soc.* **114**, 5039–5048 (1992).
- A. L. Sobolewski, W. Domcke, C. Hättig, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17903–17906 (2005).
- S. Hammes-Schiffer, A. A. Stuchebrukhov, *Chem. Rev.* **110**, 6939–6960 (2010).
- A. Migliore, N. F. Polizzi, M. J. Therien, D. N. Beratan, *Chem. Rev.* **114**, 3381–3465 (2014).
- J. J. Warren, T. A. Tronic, J. M. Mayer, *Chem. Rev.* **110**, 6961–7001 (2010).
- P. R. D. Murray *et al.*, *Chem. Rev.* **122**, 2017–2291 (2022).
- N. S. Lewis, D. G. Nocera, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729–15735 (2006).
- R. Tyburski, T. Liu, S. D. Glover, L. Hammarström, *J. Am. Chem. Soc.* **143**, 560–576 (2021).
- J. Stubbe, D. G. Nocera, C. S. Yee, M. C. Y. Chang, *Chem. Rev.* **103**, 2167–2201 (2003).
- C. Aubert, M. H. Vos, P. Mathis, A. P. M. Eker, K. Brettel, *Nature* **405**, 586–590 (2000).
- F. Lacombat *et al.*, *J. Am. Chem. Soc.* **141**, 13394–13409 (2019).
- S. Hu, A. V. Soudackov, S. Hammes-Schiffer, J. P. Klinman, *ACS Catal.* **7**, 3569–3574 (2017).
- G. A. Parada *et al.*, *Science* **364**, 471–475 (2019).
- R. Marcus, N. Sutin, *Biochim. Biophys. Acta Rev. Bioenerg.* **811**, 265–322 (1985).
- S. J. Edwards, A. V. Soudackov, S. Hammes-Schiffer, *J. Phys. Chem. B* **113**, 14545–14548 (2009).
- L. F. Cotter, B. Pettersson Rimgard, G. A. Parada, J. M. Mayer, L. Hammarström, *J. Phys. Chem. A* **125**, 7670–7684 (2021).
- E. R. Sayfutyarova, S. Hammes-Schiffer, *J. Am. Chem. Soc.* **142**, 487–494 (2020).
- E. R. Sayfutyarova, S. Hammes-Schiffer, *J. Phys. Chem. Lett.* **11**, 7109–7115 (2020).
- D. LeGourriérec, V. Kharlanov, R. G. Brown, W. Rettig, *J. Photochem. Photobiol. Chem.* **117**, 209–216 (1998).
- S. Kim, J. Seo, S. Y. Park, *J. Photochem. Photobiol. Chem.* **191**, 19–24 (2007).
- G. L. Gaines III, M. P. O’Neil, W. A. Svec, M. P. Niemczyk, M. R. Wasielewski, *J. Am. Chem. Soc.* **113**, 719–721 (1991).
- M. A. Bowring *et al.*, *J. Am. Chem. Soc.* **140**, 7449–7452 (2018).
- G. D. Scholes, K. P. Ghiggino, *J. Phys. Chem.* **98**, 4580–4590 (1994).
- S. Fatehi, E. Alguire, Y. Shao, J. E. Subotnik, *J. Chem. Phys.* **135**, 234105–234121 (2011).
- Q. Ou, G. D. Bellchambers, F. Furche, J. E. Subotnik, *J. Chem. Phys.* **142**, 064114–14 (2015).
- A. E. Reed, L. A. Curtiss, F. Weinhold, *Chem. Rev.* **88**, 899–926 (1988).
- B. Pettersson Rimgard *et al.*, Data for “Proton-coupled energy transfer in molecular triads,” Open Science Framework. DOI: 10.17605/OSF.IO/JY46X.

ACKNOWLEDGMENTS

We acknowledge N. Kaul, L. D’Amario, and E. Sayfutyarova for their expertise and helpful discussions. **Funding:** Supported by Swedish Research Council grant 2020-05246 and NIH grants 2R01GM50422 and R35GM139449. **Author contributions:** B.P.R. and L.H. conceived the project and, together with Z.T. and S.H.-S., they wrote the paper. B.P.R., L.F.C., and G.A.P. collected the data, and B.P.R. analyzed all the data. Z.T. and S.H.S. performed

all the computational work. All authors participated in active discussions and reviewed the manuscript. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. An Excel file with steady-state and time-resolved absorption and fluorescence data, as well as the optimized geometry for triads **1** and **6**, and their NBO charges, has been deposited in the Open Science Framework (30). **License information:** Copyright © 2022

the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq5173
Materials and Methods

Supplementary Text
Figs. S1 to S11
Tables S1 to S13
References (31–50)
Data S1

Submitted 13 April 2022; accepted 7 July 2022
Published online 21 July 2022
10.1126/science.abq5173

PLANT SCIENCE

Synthetic genetic circuits as a means of reprogramming plant roots

Jennifer A. N. Brophy^{1,2*}, Katie J. Magallon¹, Lina Duan¹, Vivian Zhong², Prashanth Ramachandran¹, Kiril Kniazhev¹, José R. Dinnyen^{1*}

The shape of a plant's root system influences its ability to reach essential nutrients in the soil and to acquire water during drought. Progress in engineering plant roots to optimize water and nutrient acquisition has been limited by our capacity to design and build genetic programs that alter root growth in a predictable manner. We developed a collection of synthetic transcriptional regulators for plants that can be compiled to create genetic circuits. These circuits control gene expression by performing Boolean logic operations and can be used to predictably alter root structure. This work demonstrates the potential of synthetic genetic circuits to control gene expression across tissues and reprogram plant growth.

Recent advances in our understanding of the molecular mechanisms driving plant development have identified key regulators of agronomic traits, but our limited ability to control gene expression in plants is a barrier to applying this knowledge (1, 2). Indeed, it remains challenging to express genes in specific patterns in plants, especially if those patterns are outside the reach of native promoters (3–6). One attractive solution is synthetic genetic circuits, which have been constructed in a wide array of prokaryotic organisms and eukaryotic cell lines (7–10). However, circuit technology has been difficult to implement in plants because of the long time scales required for producing transgenic lines and the difficulty of tuning circuit activity across heterogeneous cell types (11). Here, we show that quantitative transient expression assays can be used to aid the construction of circuits in plants by providing a platform for testing and tuning circuit designs. We show how these assays can be leveraged to improve the performance of circuits designed to produce specific spatial patterns of gene expression across root tissues, and we use one of the circuits to precisely modify plant root architecture.

Constructing synthetic genetic circuits in plants

To construct genetic circuits in plants, we first generated a collection of synthetic transcriptional regulators. We designed simple transcriptional activators [composed of bacterial DNA binding proteins, VP16-activation domains, and SV40 nuclear localization signals (NLSs)], and synthetic repressors that rely on steric hindrance to achieve repression (composed of only DNA binding proteins and NLSs) (Fig. 1, A and B) (12–16). Similar to previous synthetic promoter designs, activatable plant promoters were created by fusing six copies of the DNA sequence (operator) bound by these transcription factors (TFs) to a minimal plant promoter [positions –66 to +18 of the cauliflower mosaic virus (CaMV) 35S promoter] (Fig. 1A) (17, 18). Repressible promoters were built by placing one operator sequence at the 3' end of a full-length CaMV 35S promoter (Fig. 1B) (19). This design was selected to avoid disrupting 35S promoter activity when adding operators.

Synthetic TF activity was measured in plants by delivering DNA encoding the proteins and their target promoters to *Nicotiana benthamiana* leaves using *Agrobacterium tumefaciens*. To enable quantitative measurements of gene expression, we used the synthetic promoters to drive expression of green fluorescent protein (GFP) and normalized GFP expression to a constitutively expressed mCherry encoded on the same T-DNA (see supplementary materials). We also introduced an intron to *GFP* to prevent

Agrobacterium from expressing the reporter and confounding fluorescence measurements (fig. S1). Nine of the 10 synthetic activators were able to turn on their target promoters in *N. benthamiana* (factor of 3 to 45 increase relative to promoter-only controls), and seven of these were specific to the promoters they were designed to regulate (Fig. 1, C and D). Specificity is needed to construct circuits free of cross-talk between components (20). In contrast to the activators, only four of the synthetic repressors generated more than a factor of 2 change in gene expression, and none achieved complete repression (Fig. 1, E and F). Thus, many of the repressors required further optimization before they could be used to construct circuits.

To achieve greater control over expression, we modified the synthetic TFs. Replacing the human herpesvirus-derived VP16 activation domain with the activation domain from *Arabidopsis* ETHYLENE RESPONSE FACTOR 2 (ERF2^{AD}) improved activity of the PhIF-based synthetic activator by nearly a factor of 6 (Fig. 1G) (21). A small fragment of ERF2 (78 amino acids) was used to reduce the possibility of regulation by native plant components, such as small RNAs, kinases, etc. (21). In contrast to the activators, adding a repressor domain to the AmtR-based synthetic repressor either had no effect [TOPLESS (TPLRD)] or only modestly increased repression (relative to NLS only) by 20% [ETHYLENE RESPONSE FACTOR 4 (ERF4RD)] or 33% [SALT TOLERANCE ZINC FINGER (ZAT10RD)] (Fig. 1H) (22, 23); thus, swapping the activation domain led to increased expression, whereas adding repressive domains had a more modest effect.

We also engineered the synthetic promoters to tune expression levels and increase dynamic range. Changing the number of TF binding sites in the AmtR-activatable promoter resulted in a collection of promoters of varying strength. Promoters with three or more operators amplified expression relative to the constitutive G1090 promoter used to drive expression of AmtR-VP16 (Fig. 1I, compare 6× to 3× with control) (24). In contrast, promoters with two operators roughly replicated G1090's expression level and one operator produced less expression (Fig. 1I). These data suggest that synthetic promoters can be used to tune gene expression levels by either amplifying or

¹Department of Biology, Stanford University, Stanford, CA, USA. ²Department of Bioengineering, Stanford University, Stanford, CA, USA.

*Corresponding author. Email: jbrophy@stanford.edu (J.A.N.B.); dinnyen@stanford.edu (J.R.D.)

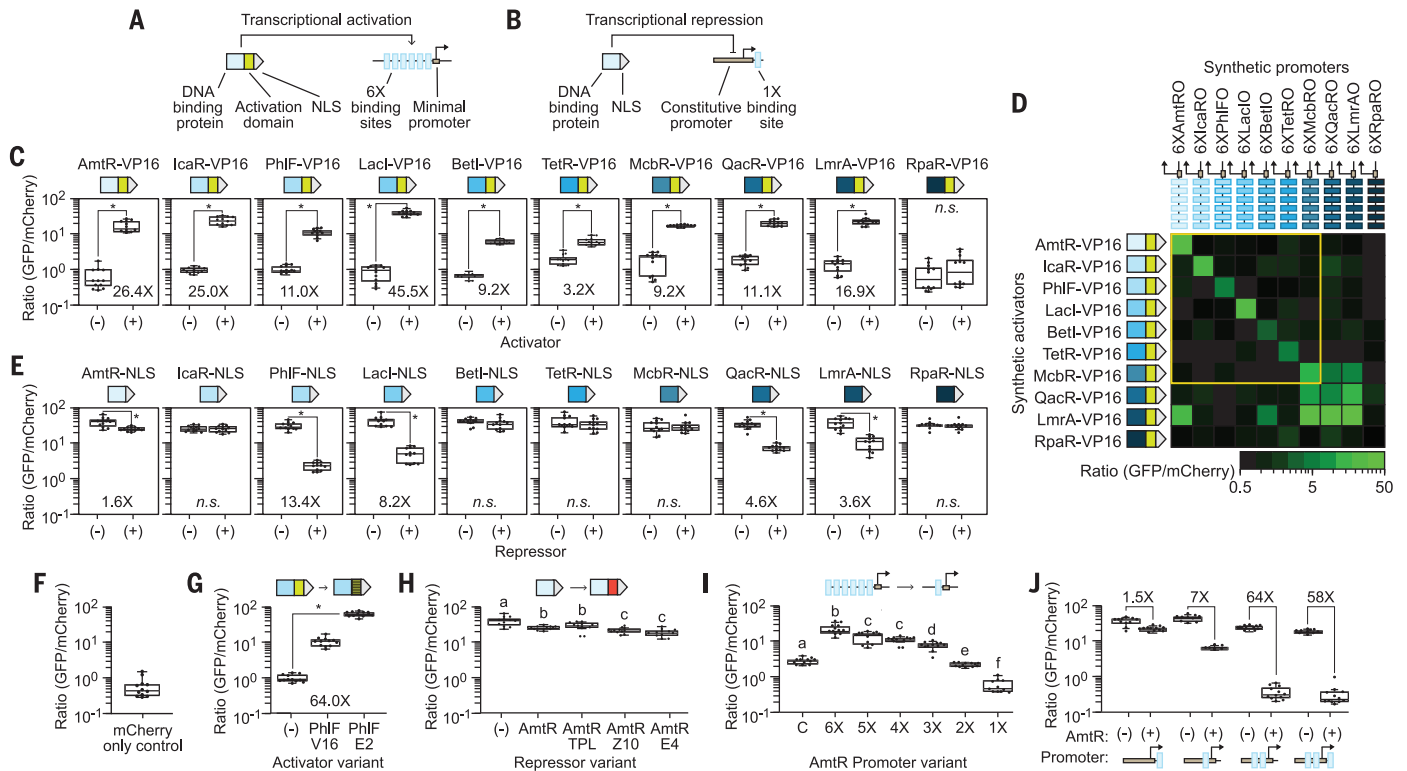


Fig. 1. Basic building blocks for constructing synthetic genetic circuits.

(A and B) Schematics of the synthetic transcriptional activators (A) and repressors (B) built to control gene expression in plants. Small bent arrow denotes the transcription start site. Activity (C) and specificity (D) of the synthetic activators. Yellow box highlights orthogonal activators. (E) Activity of the synthetic repressors. (F) Fluorescence of leaves infiltrated with a plasmid that does not encode GFP. (G) Comparison of PhIF-based activators containing either the VP16 (V16) or ERF2 (E2) activation domain. (H) Comparison of AmtR-based repressors containing either no repressive domain or a repressive domain from TOPLESS (TPL), ZAT10 (Z10), or ERF4 (E4). Control (–) is promoter alone. (I) Activity of synthetic activatable promoters

containing one (1×) to six (6×) AmtR operators co-infiltrated with a plasmid encoding *proG0190::AmtR-VP16*. Control construct (C) is *proG0190::GFP*. (J) Engineered synthetic promoters' response to AmtR-NLS. In all panels, expression is measured in *N. benthamiana* leaves. TFs are expressed using the constitutive G1090 promoter. Box plots show the median of 12 leaf punches collected from three leaves that were infiltrated and measured on different days. Box plot hinges indicate the first and third quartiles. Dots show individual data points. Stars denote significant differences in normalized GFP expression and letters denote significance groups ($P < 0.01$, Student's two-tailed t test); n.s., not significant. Relative change for significantly active transcription factors was calculated by dividing the average ON state by the average OFF state.

dampening the transcriptional signal produced by an input promoter.

In line with a previous study, we found that moving the location of the operator to a region between the TATA and CAAT boxes within the 35S promoter improved repression and dynamic range (Fig. 1J) (10). Adding a second operator between the TATA box and the transcription start site further improved repression and dynamic range (Fig. 1J). However, adding a third binding site at the 3' end of the promoter reduced its ON state without significantly reducing the repressed state, resulting in a smaller dynamic range than the two-operator promoter (Fig. 1J, 58× versus 64×). These design features were transferrable to other synthetic repressor-promoter pairs (fig. S2), demonstrating that the number and location of operators can be adjusted to tune promoter activity.

Using these synthetic regulators, we constructed circuits that perform Boolean logic operations. Synthetic TFs built with the AmtR

and PhIF DNA binding proteins, which demonstrated strong activation and repression in our initial TF designs, served as the inputs to all circuits, and GFP served as the output. Circuit activity was measured in *N. benthamiana* leaves, which were infiltrated with multiple *Agrobacterium* strains, each containing one plasmid that encoded either an input TF or the output (Fig. 2, A and B).

Computation was performed by synthetic promoters that responded to the input TFs in unique ways. Simple gates, like the A and B BUFFER gates, used synthetic promoters that responded to one input TF; more complex gates required synthetic promoters that responded to multiple inputs (Fig. 2C). To create functional circuits, both the promoter architecture and the synthetic TFs needed to be optimized (fig. S3). We found that the arrangement of operators in the OR promoter affected fold change, with the best design containing alternating pairs of operators (Fig. 2C and fig. S3). Additionally, synthetic repressors needed

to include repressive domains, such as ERF4RD, to prevent synthetic activators from initiating transcription at composite promoters (1-1 state, A NIMPLY B; Fig. 2C and fig. S3). However, this requirement could be overcome by increasing the number and location of repressor operators within the synthetic promoter (B NIMPLY A; fig. S3). Most circuits involved several design-build-test cycles, which were facilitated by the rapid *N. benthamiana*-based assays and the modular nature of the synthetic biology parts generated here. Additional tuning information is available in figs. S3 and S4.

Layered logic gates in which AmtR and PhIF do not directly control expression of GFP, but instead modulate the expression of other synthetic TFs, occasionally encountered problems where expression cassettes were insufficiently insulated from each other. The A IMPLY B gate, which worked well when its output genes were encoded on separate plasmids, had an erroneously reduced “no input” state when

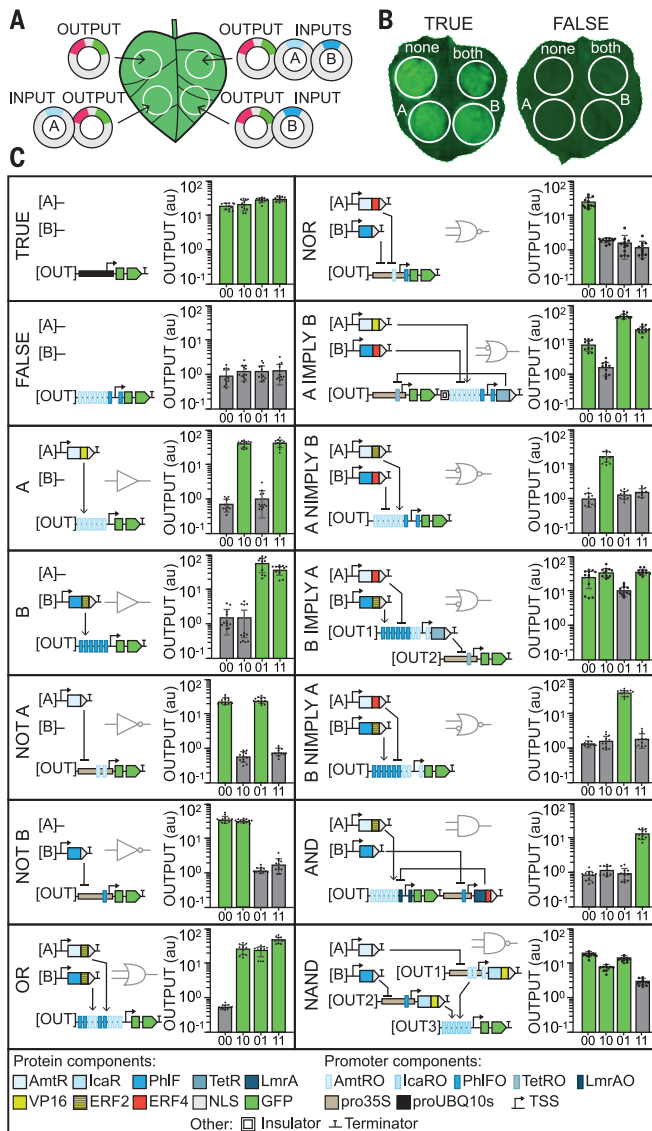


Fig. 2. Logic gates in *N. benthamiana* leaves. (A) Schematic for testing circuit activity in *N. benthamiana* using a transient expression assay (see supplementary materials). *N. benthamiana* leaves were infiltrated with *Agrobacterium* strains containing plasmids that encode either one input transcription factor (A or B) or the output gene *GFP*. mCherry was encoded on the output plasmid and used to normalize fluorescence measurements. (B) Example images of leaves infiltrated with TRUE and FALSE gate components. Circles encompass areas of infiltration. (C) Gate behavior in *N. benthamiana*. Bar charts show output of each circuit, reported as the ratio of GFP to mCherry. Green bars indicate gate states that should be ON; gray bars indicate states that should be OFF to implement correct logic. Data are mean and SD of 12 leaf punches collected from three leaves infiltrated and measured on different days. Dots show individual data points. Data for individual GFP and mCherry channels are provided in the supplementary materials.

both output genes were encoded on the same plasmid (Fig. 2C and fig. S4). In an attempt to fix this problem, we added one or two copies of the previously characterized insulator *Arabidopsis thaliana* MATRIX ATTACHMENT REGION 10 (MAR10) between the two

A IMPLY B output genes (25). However, insulation only improved gate activity by a factor of 2 (fig. S4, version 2 versus versions 3 and 4). Consequently, the other IMPLY gate (B IMPLY A) was built with two output plasmids that were co-delivered to plant cells (Fig. 2C).

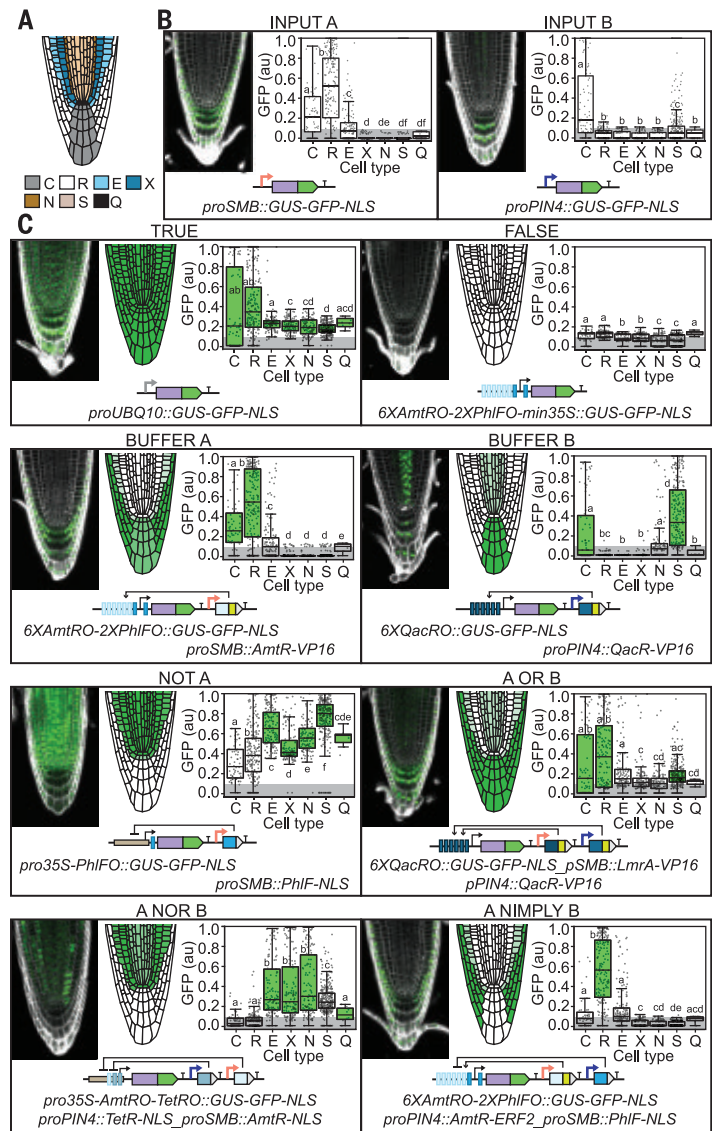


Fig. 3. Patterning gene expression using logic gates. (A) Cell types in the *Arabidopsis* root tip: columella (C), lateral root cap (R), epidermis (E), cortex (X), endodermis (N), stele (S), quiescent center (Q). (B) Confocal images show the expression pattern of input promoters *SOMBRERO* (SMB) (input A) and *PIN-FORMED 4* (PIN4) (input B). (C) Confocal images (left) and expected expression patterns (right) of logic gates that use the SMB and PIN4 promoters to express input TFs. Output of each gate is nuclear localized GUS-GFP fusion protein. In all panels, T-DNA schematics show the identity and arrangement of circuit components. Root images were taken 5 days after sowing. Box-and-whisker plots show quantified reporter expression for three T2 plants from a single transgenic line. Green boxes indicate cell layers that should be ON and gray bars indicate states that should be OFF to implement correct logic. Expression levels of individual cells in each root layer are shown as black dots. Box plot hinges indicate the first and third quartiles. Letters denote significant differences in expression ($P < 0.01$, Student's two-tailed t test). Additional independent lines in fig S6.

Engineering spatial patterns of gene expression

Functional plant circuits were transferred to the model plant *A. thaliana* to test their capacity to generate specific spatial patterns of gene expression across root tissues. The tissue-specific

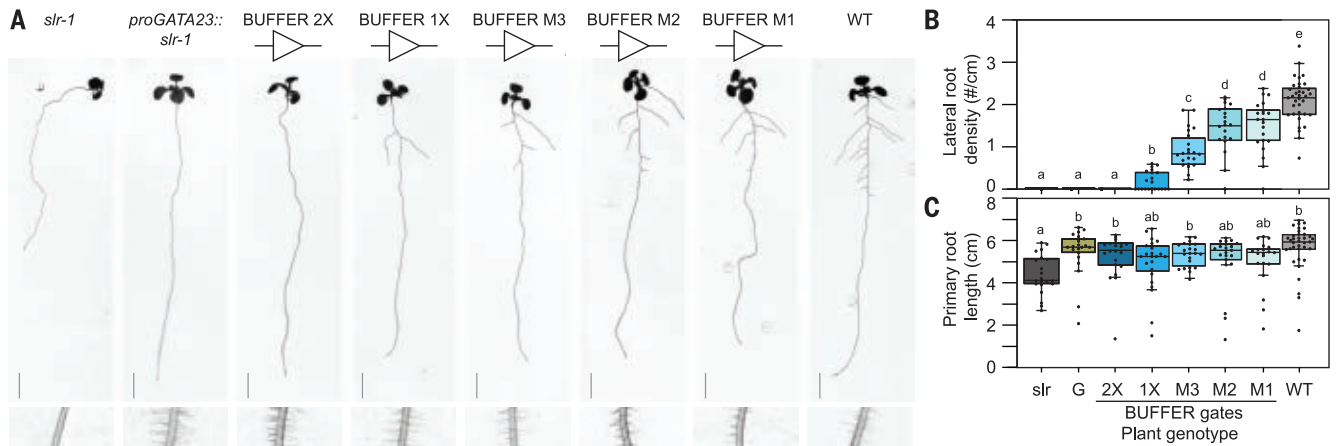


Fig. 4. Engineering root branch density using BUFFER gates. (A) Root structure (top) and root hairs (bottom) of *Arabidopsis* plants engineered to modify root branch density. BUFFER gate output promoters contain either one (1×) or two (2×) copies of the consensus AmtR operator or one copy of a mutated AmtR operator (M1 to M3). Scale bar, 1 cm. (B) Lateral root density calculated as the number of emerged lateral roots divided by total primary root

length. (C) Primary root length measured from root-hypocotyl junction to the root tip. Control samples are labeled as follows: *slr-1* (slr), *proGATA23::slr* (G), wild type (WT). In (B) and (C), box plots show the median of at least 20 T1 plants. Box plot hinges indicate the first and third quartiles. Letters denote significant differences in LR density or primary root length ($P < 0.01$, Student's two-tailed t test). All measurements and images were taken 10 days after sowing.

promoters of *SOMBRERO* (*proSMB*, expressed in the entire root cap) and *PIN-FORMED 4* (*proPIN4*, expressed in columella, root cap, and stele) were used to drive expression of our input TFs (Fig. 3, A and B) (26, 27). By combining the activity of these input promoters using logic gates, we expected to generate several different spatial patterns of gene expression (Fig. 3C, cartoons).

We tested eight different logic gates in *Arabidopsis*, four of which generated the expected expression pattern upon first attempt (TRUE, FALSE, A BUFFER, NOT A; Fig. 3C). Successful gates qualitatively matched the expected expression patterns and produced a significant difference between GFP expression in tissues expected to be ON versus those expected to be OFF. By this definition of success, TRUE, FALSE, A BUFFER, and NOT A gates were successful. For NOT A, the difference between the lowest ON (cortex) and highest OFF (root cap) states was only 1.2×; this finding suggests that further optimization may be necessary for applying the circuit in other contexts (Fig. 3C).

The gates that did not work revealed differences in circuit behavior between *Arabidopsis* and *N. benthamiana*. For example, the first B BUFFER gate produced a spatially expanded expression pattern relative to the input promoter, with aberrant expression in the quiescent center (QC) and neighboring initial cells (figs. S5 and S6). Similarly, the A NIMPLY B pattern was incorrect in *Arabidopsis*; expression was missing in several root cap cell layers (figs. S5 and S6).

We tuned the B BUFFER, A NIMPLY B, OR, and NOR gates to improve performance. For B BUFFER gate, we changed the input TF from AmtR-VP16 to QacR-VP16, another highly active synthetic activator in *N. benthamiana* (Fig. 1C,

Fig. 3C, and fig. S5). This change was made to eliminate potential interactions between the promoter (*proPIN4*) and the downstream gene. For A NIMPLY B, we removed the ERF4RD from the input TF (*PhlF-ERF4*). ERF4 is believed to recruit histone-modifying enzymes and may maintain a repressed state in cells that are not actively expressing the repressor (23). Successful A NIMPLY B gates were built with repressors that lack the repression domain (*PhlF-NLS*) (Fig. 3C and fig. S5). Tuning of the remaining gates (OR and NOR) and a description of our troubleshooting process are outlined in figs. S5 to S7.

After tuning, all gates qualitatively matched the expected expression patterns (Fig. 3C). However, three had expression in a single cell layer that was either aberrantly high (B BUFFER, endodermis) or low (OR, stele; NOR, QC). Thus, quantitative analysis highlights the challenge of implementing circuits across cell types in heterogeneous tissue and additional optimization would be required for these gates to achieve significant differences across every tissue layer.

Modifying root structure

To demonstrate how precise spatial control over gene expression may be used to engineer development, we modified a key aspect of *Arabidopsis* root architecture: lateral root branch density. Lateral roots allow plants to radially sample soil, and the number of lateral roots that a plant generates affects its ability to search for water and essential nutrients in the environment (28). The close relationship between root growth and plant fitness has led to proposals of ideal root architectures for plant growth in specific environments (called root ideotypes) (29). Ideotype hypothe-

ses have been difficult to test directly because genetic changes that affect root structure often have pleiotropic effects. For example, a gain-of-function mutation in the developmental regulator *INDOLE-3-ACETIC ACID INDUCIBLE 14* (*IAA14*) called *solitary root* (*slr-1*) eliminates root branching, but also hinders root gravitropism, root hair development, and primary root growth (Fig. 4 and fig. S8, *slr-1*) (30). To disentangle root branching from other developmental processes, we expressed the *slr-1* mutant gene using a tissue-specific promoter that is only ON in lateral root stem cells, *proGATA23*²³. When expressed from *proGATA23*, *slr-1*'s impact on development is restricted to lateral roots. In these plants, no lateral roots form, but gravitropism, root hair development, and primary root growth are normal (Fig. 4 and fig. S8, *proGATA23::slr-1*).

We designed BUFFER gates to express *slr-1* at varying levels in lateral root stem cells to determine whether lateral root branch development could be quantitatively controlled. These BUFFER gates use the *GATA23* promoter to drive expression of the AmtR-VP16 synthetic TF, which then activates a synthetic promoter with one, two, four, or six copies of the AmtR operator to drive expression of *slr-1*. Because plant transformation randomly inserts transgenes into the genome, a single construct can create a range of expression. Although this variation potentially alleviates the need for promoters of varying strength, none of the plants (>20 independent lines per construct) containing *proGATA23::slr-1* or a BUFFER gate with two, four, and six copies of the AmtR operator produced lateral roots; this suggests that the range of expression conferred by these constructs was above the threshold for fully blocking lateral root development (Fig. 4, A

and B; data not shown for 4× and 6× promoters). To further reduce the strength of the output promoter, we mutated key residues in the AmtR binding site (fig. S9) (31). Weaker synthetic promoters resulted in BUFFER gates that produced a gradient of lateral root densities, roughly correlated with the strength of the synthetic promoter (Fig. 4 and fig. S10). None of the BUFFER gates affected primary root growth, root hair development, or gravitropism (Fig. 4, B and C, and figs. S8 and S10). Thus, BUFFER gates enabled specific tuning of lateral root density. Our results show that the difference in regulatory capacity conferred by our synthetic system has a greater effect on root branching than the T-DNA insertion site, but they do not define how a specific dosage of *str-1* determines lateral root development; the resources generated here can be used to explore this relationship further.

Our work presents a roadmap for the design and implementation of synthetic genetic circuits that program gene expression across cell types in plants. This approach expands the impact that a handful of characterized promoters can have on our ability to express genes in specific cells and presents an alternative to searching genomes for tissue-specific promoters with desired expression patterns. The ability to modulate expression in a tissue-specific manner provides an opportunity to probe gene dosage effects in a tissue-specific manner, which previously lacked a reliable framework in plants. As circuit technology matures in plants, it could be applied to reprogram responses to the environment or to better understand gene expression control in plants, as has been done recently in *Drosophila* (32, 33).

Reprogramming crops using synthetic genetic circuits will require careful tuning. As evidenced here, controlling expression levels across tissues can be challenging, and although some applications may tolerate aberrant expression in a few cell layers, others will be extremely sensitive to off-target expression. To efficiently engineer genetic circuits that meet specific design requirements, researchers can leverage the growing number of transient expression methods available for plants (34, 35). Advances in single-cell measurements and spatially resolved omics will also support the construction of circuits in multicellular organisms by enabling more precise measurements (36, 37). Ultimately, methods for programming novel traits in plants will become increasingly useful as climate challenges grow and new agricultural solutions are needed.

REFERENCES AND NOTES

1. S. J. Gurr, P. J. Rushton, *Trends Biotechnol.* **23**, 275–282 (2005).
2. A. E. Richardson et al., *Science* **374**, 1377–1381 (2021).
3. S. Ali, W.-C. Kim, *Front. Plant Sci.* **10**, 1433 (2019).
4. J. Bai et al., *Plant Biotechnol. J.* **18**, 668–678 (2020).
5. R. Wang et al., *Sci. Rep.* **5**, 18256 (2015).
6. L. Chen et al., *Agriculture* **11**, 1195 (2021).
7. A. A. K. Nielsen et al., *Science* **352**, aac7341 (2016).
8. D. Mishra et al., *Science* **373**, eaav0780 (2021).

9. X. J. Gao, L. S. Chong, M. S. Kim, M. B. Elowitz, *Science* **361**, 1252–1258 (2018).
10. K. A. Schaumberg et al., *Nat. Methods* **13**, 94–100 (2016).
11. T. K. Kassaw, A. J. Donayre-Torres, M. S. Antunes, K. J. Morey, J. I. Medford, *Plant Sci.* **273**, 13–22 (2018).
12. A. J. Meyer, T. H. Segall-Shapiro, E. Glassey, J. Zhang, C. A. Voigt, *Nat. Chem. Biol.* **15**, 196–204 (2019).
13. B. C. Stanton et al., *Nat. Chem. Biol.* **10**, 99–105 (2014).
14. I. Sadowski, J. Ma, S. Triezenberg, M. Ptashne, *Nature* **335**, 563–564 (1988).
15. B. C. Stanton et al., *ACS Synth. Biol.* **3**, 880–891 (2014).
16. C. Dingwall, R. A. Laskey, *Trends Biochem. Sci.* **16**, 478–481 (1991).
17. R. X. Fang, F. Nagy, S. Sivasubramanian, N. H. Chua, *Plant Cell* **1**, 141–150 (1989).
18. L. Laplace et al., *J. Exp. Bot.* **56**, 2433–2442 (2005).
19. J. T. Odell, F. Nagy, N.-H. Chua, *Nature* **313**, 810–812 (1985).
20. J. B. Lucks, L. Qi, W. R. Whitaker, A. P. Arkin, *Curr. Opin. Microbiol.* **11**, 567–573 (2008).
21. J. Li et al., *Plant Biotechnol. J.* **11**, 671–680 (2013).
22. E. Pierre-Jerome, S. S. Jang, K. A. Havens, J. L. Nemhauser, E. Klavins, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9407–9412 (2014).
23. M. Ohta, K. Matsui, K. Hiratsui, H. Shinshi, M. Ohme-Takagi, *Plant Cell* **13**, 1959–1968 (2001).
24. F. Ishige, M. Takaichi, R. Foster, N.-H. Chua, K. Oeda, *Plant J.* **18**, 443–448 (1999).
25. A. Pérez-González, E. Caro, *Sci. Rep.* **9**, 8474 (2019).
26. M. Kamiya et al., *Development* **143**, 4063–4072 (2016).
27. M. D. M. Marqués-Bueno et al., *Plant J.* **85**, 320–333 (2016).
28. R. Rellán-Álvarez, G. Lobet, J. R. Dinnyen, *Annu. Rev. Plant Biol.* **67**, 619–642 (2016).
29. J. P. Lynch, *New Phytol.* **223**, 548–564 (2019).
30. H. Fukaki, S. Tameda, H. Masuda, M. Tasaka, *Plant J.* **29**, 153–168 (2002).
31. K. Hasselt, S. Rankl, S. Worsch, A. Burkovski, *J. Biotechnol.* **154**, 156–162 (2011).
32. J. Crocker, G. R. Ilsley, D. L. Stern, *Nat. Genet.* **48**, 292–298 (2016).
33. J. Crocker, A. Tsai, D. L. Stern, *Cell Rep.* **18**, 287–296 (2017).

34. A. Khakhar, D. F. Voytas, *Front. Plant Sci.* **12**, 668580 (2021).
35. J. W. Wang et al., *Mol. Plant* **12**, 1037–1040 (2019).
36. B. Cole et al., *Commun. Biol.* **4**, 962 (2021).
37. S. Vickovic et al., *Nat. Commun.* **13**, 795 (2022).

ACKNOWLEDGMENTS

We thank T. Beekman for the *Arabidopsis slr* mutant and T. Velloso for assistance in establishing imaging protocols. **Funding:** Burroughs Wellcome Fund CASI (J.A.N.B.); Chan Zuckerberg Biohub (J.A.N.B., J.R.D.); US Department of Energy Biological and Environmental Research Program grant DE-SC0008769 (J.R.D.); Faculty Scholar grant from Howard Hughes Medical Institute and the Simons Foundation 55108515 (J.R.D.). **Author contributions:** Conceptualization, J.A.N.B., J.R.D.; data acquisition, J.A.N.B., K.J.M., L.D., V.Z., K.K.; visualization, J.A.N.B.; funding acquisition, J.A.N.B., J.R.D.; project administration, J.R.D.; supervision, J.A.N.B., J.R.D.; writing—original draft, J.A.N.B., J.R.D.; writing—review and editing, J.A.N.B., V.Z., J.R.D. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All plasmid materials and bacterial strains will be made available through AddGene. Sequence files are available at doi.org/10.5281/zenodo.6949657 and raw data are available as supplementary materials. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo4326
Materials and Methods
Figs. S1 to S10
References (38–54)
Data S1 and S2

Submitted 3 February 2022; accepted 23 June 2022
10.1126/science.abo4326

REPORTS

BIO MATERIALS

Controlled tough bioadhesion mediated by ultrasound

Zhenwei Ma¹, Claire Bourquard², Qiman Gao³, Shuaibing Jiang¹, Tristan De lure-Grimmel⁴, Ran Huo¹, Xuan Li¹, Zixin He¹, Zhen Yang¹, Galen Yang⁵, Yixiang Wang⁶, Edmond Lam^{5,7}, Zu-hua Gao⁸, Outi Supponen^{2*}, Jianyu Li^{1,9*}

Tough bioadhesion has important implications in engineering and medicine but remains challenging to form and control. We report an ultrasound (US)-mediated strategy to achieve tough bioadhesion with controllability and fatigue resistance. Without chemical reaction, the US can amplify the adhesion energy and interfacial fatigue threshold between hydrogels and porcine skin by up to 100 and 10 times. Combined experiments and theoretical modeling suggest that the key mechanism is US-induced cavitation, which propels and immobilizes anchoring primers into tissues with mitigated barrier effects. Our strategy achieves spatial patterning of tough bioadhesion, on-demand detachment, and transdermal drug delivery. This work expands the material repertoire for tough bioadhesion and enables bioadhesive technologies with high-level controllability.

Bioadhesive technologies find frequent use in wearable electronics, biomedical implants, wound management, anastomosis, regenerative medicine, and drug delivery (1–5). However, their use has long been hindered by the barrier effects of biological tissues, such as low permeability and limited functional groups. Skin, for instance, contains dense stratum corneum, limiting the penetration and bonding of bioadhesive agents (Fig. 1A). These issues are

particularly critical for bioadhesives based on physical interactions such as polymer interpenetration; the polymers are too slow and even impossible to diffuse and entangle with tissues (6), resulting in poor bioadhesion (7). Chemical strategies have thus far been primarily used for tough bioadhesion. Despite achieving high adhesion energy, they do not enable high-level control over bioadhesion in space and time. Exceptions require sophisticated surface patterning, exogenous

chemicals, and an external apparatus to remove interfacial bonding (8, 9). Other disadvantages include interference with payloads for drug delivery (10), low fatigue threshold (caused by limited functional groups on tissue surfaces) (11), and acute and/or chronic toxicity (caused by chemical reactions and their reagents) (12).

To form and control tough bioadhesion, our strategy leverages ultrasound (US) and anchoring agents such as polymers and nanoparticles. US has been extensively used in the clinical setting for imaging (13), monitoring (14), tumor ablation (15), and drug delivery (16). The anchoring agents can form a physically cross-linked network but do not pene-

trate tissues by passive diffusion (Fig. 1A) (6). We hypothesized that US could propel and immobilize the anchoring agents deep into tissues, which would form a bridging network at the interface when a hydrogel patch is applied (Fig. 1B). This strategy of US-mediated bioadhesion differentiates from existing strategies based on passive diffusion, chemical reactions, or invasive mechanical interlocking, including those in medical devices such as suture anchors and swellable microneedles (17) (fig. S1).

US-mediated bioadhesion was achieved in two steps. First, we used an ultrasonic transducer (20 kHz) to apply US to a primer solution or suspension of anchoring agent spread on a tissue substrate (e.g., freshly excised porcine skin). Second, we covered the treated area with a hydrogel patch, which triggers gelation of the anchoring agent at the interface (Fig. 1B). As a model system, we deployed a chitosan solution and a polyacrylamide-alginate (PAAm-alg) hydrogel as the primer solution and the hydrogel patch, respectively. We measured the adhesion energy between the hydrogel and the tissue with peeling tests (Fig. 1C). With US treatment (116 W cm^{-2} , 1 min), the adhesion energy obtained on porcine skin was $\sim 1750 \text{ J m}^{-2}$, >15 times that of the no-US control (Fig. 1D). The adhesion energy reached $\sim 100 \text{ J m}^{-2}$ within a minute and then plateaued in 10 min (fig. S2). The anchoring agent is critical for tough bioadhesion because merely

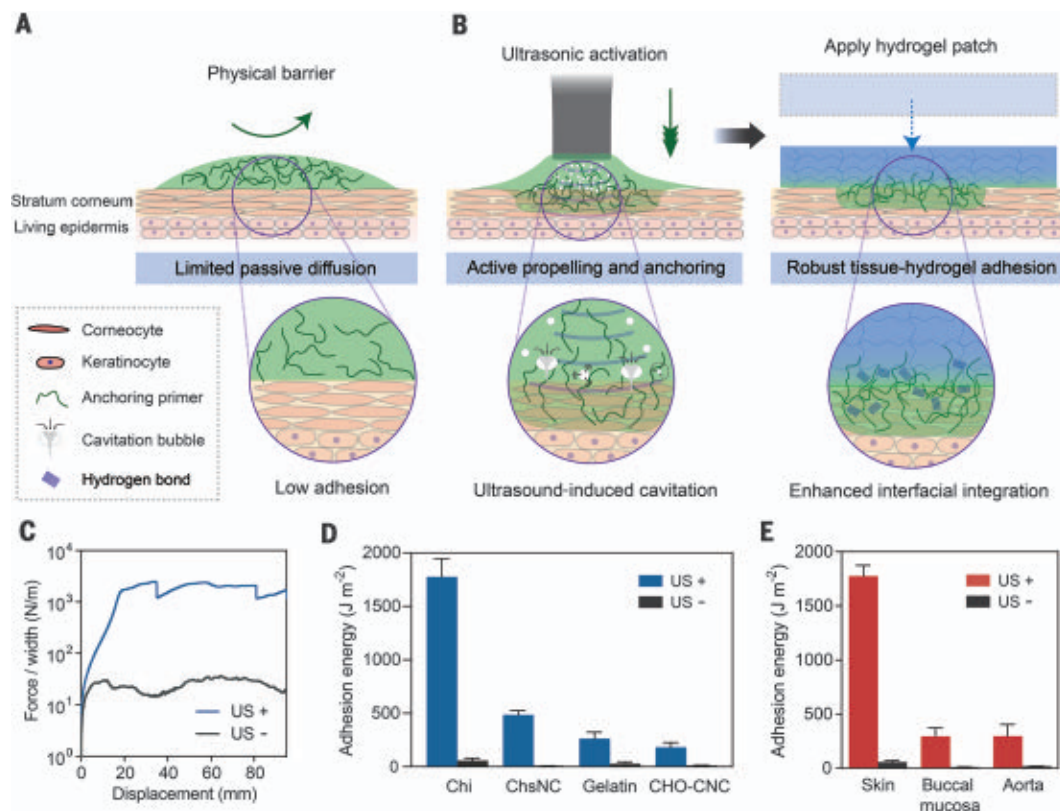
replacing it with phosphate-buffered saline yielded weak bioadhesion, even with US treatment (fig. S3). The adhesion performance was dependent on the pH difference between the hydrogel patch and the primer solution (fig. S4) but independent of blood exposure on tissue substrates (fig. S5). We also used an ultrasonic scaler (20 to 35 kHz) like the type used in dental clinics to obtain tough adhesion ($\sim 800 \text{ J m}^{-2}$) on porcine skin with US treatment (fig. S6).

Our strategy is applicable to a large repertoire of materials. The adhesion enhancement by US was confirmed with different hydrogels, including a double-network poly(N-isopropylacrylamide)-alginate (PNIPAm-alg) hydrogel and a single-network PAAm hydrogel; the variance of adhesion energy can be linked with the bulk toughness of the hydrogel (fig. S7). We also demonstrated US-mediated bioadhesion with various anchoring agents, including gelatin, ChsNC, and CHO-CNC (Fig. 1D), and the different adhesion performances could be related to the charge density and structure of anchoring agents. Gelatin could lead to $>200 \text{ J m}^{-2}$ of adhesion energy with US treatment, markedly higher than that without US ($\sim 26 \text{ J m}^{-2}$). CHO-CNC yielded high adhesion energy ($\sim 180 \text{ J m}^{-2}$) with US treatment, whereas the condition without US yielded weak adhesion ($\sim 12 \text{ J m}^{-2}$). ChsNC could not penetrate deep into skin without the aid of US (fig. S8). The adhesion of ChsNC without US remained weak even

¹Department of Mechanical Engineering, McGill University, Montréal, Quebec H3A 0C3, Canada. ²Institute of Fluid Dynamics, Department of Mechanical and Process Engineering, ETH Zürich, 8092 Zürich, Switzerland. ³Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, Quebec H3A 1G1, Canada. ⁴Department of Bioengineering, McGill University, Montréal, Quebec H3A 0E9, Canada. ⁵Department of Chemistry, McGill University, Montréal, Quebec H3A 0B8, Canada. ⁶Department of Food Science and Agricultural Chemistry, McGill University, Sainte-Anne-De-Bellevue, Quebec H9X 3V9, Canada. ⁷Aquatic and Crop Resource Development Research Centre, National Research Council of Canada, Montréal, Quebec H4P 2R2, Canada. ⁸Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC V6T 1Z7, Canada. ⁹Department of Biomedical Engineering, McGill University, Montréal, Quebec H3A 2B4, Canada. *Corresponding author. Email: jianyu.li@mcgill.ca (J.L.); outis@ethz.ch (O.S.)

Fig. 1. Robust and versatile US-mediated tough bioadhesion.

(A) Schematic of skin with barrier effects limiting passive diffusion and impairing bioadhesion. (B) US actively propels and anchors primer agents into a tissue substrate, causing spatially confined tough adhesion between hydrogel and tissue. (C) Representative force-displacement curves of hydrogel-tissue (porcine skin) hybrids with or without US treatment in peeling tests. (D) US enables diverse anchoring agents for tough bioadhesion on skin. Chi, chitosan; ChsNC, chitosan nanocrystal; CHO-CNC, aldehyde functionalized cellulose nanocrystal. (E) Adhesion enhancement with US on diverse biological tissues, including skin, buccal mucosa, and aorta. Chitosan was used as the anchoring primer. Data are shown as means \pm SD for $n = 3$ independent experiments.



when they formed amide bonds with tissues (fig. S9). When the US was applied, the adhesion energy increased to 500 J m^{-2} , two orders of magnitude higher than the same material without US treatment ($\sim 5 \text{ J m}^{-2}$).

In addition to the skin, our strategy is applicable to other biological tissues, including buccal mucosa and aorta (Fig. 1E). The measured adhesion energy was $\sim 295 \text{ J m}^{-2}$ for buccal mucosa and $\sim 297 \text{ J m}^{-2}$ for aorta; the values for no-US conditions were ~ 12 and $\sim 17 \text{ J m}^{-2}$, respectively. Tough bioadhesion was evidenced by the debonding of buccal membrane from the underlying tissues during peeling (fig. S10), indicative of strong adhesion.

The tissue-dependent adhesion performance can be related to the mechanics and chemistry of specific tissues (18). The versatility of US-mediated bioadhesion and its indispensable role in nanoparticle bioadhesion unlock the potential of various materials for tough bioadhesion (fig. S11) (19).

We investigated the mechanism underlying US-mediated bioadhesion as follows. The US exerts mainly thermal and mechanical effects on the primer solution and the tissue substrate. Given the US intensity used in this study, we detected minimal change of the US-treated chitosan in terms of structure and gelling behavior (fig. S12) (20). To evaluate the

thermal effect, we monitored the temperature change using a thermal camera. After a 1-min application of 16 W cm^{-2} US, the surface temperature of porcine skin increased by $\sim 10^\circ\text{C}$ (fig. S13). To determine the effects of this, we replicated the temperature change with a temperature-controlled oven to incubate hydrogel-primer-tissue hybrids. Because no significant change of bioadhesion was found (fig. S14), we excluded the thermal effect. We then focused on the mechanical effects of US such as cavitation, viscous stress, and acoustic radiation force. Of these, cavitation was hypothesized to play a key role (21) because viscous stress and acoustic radiation force are less

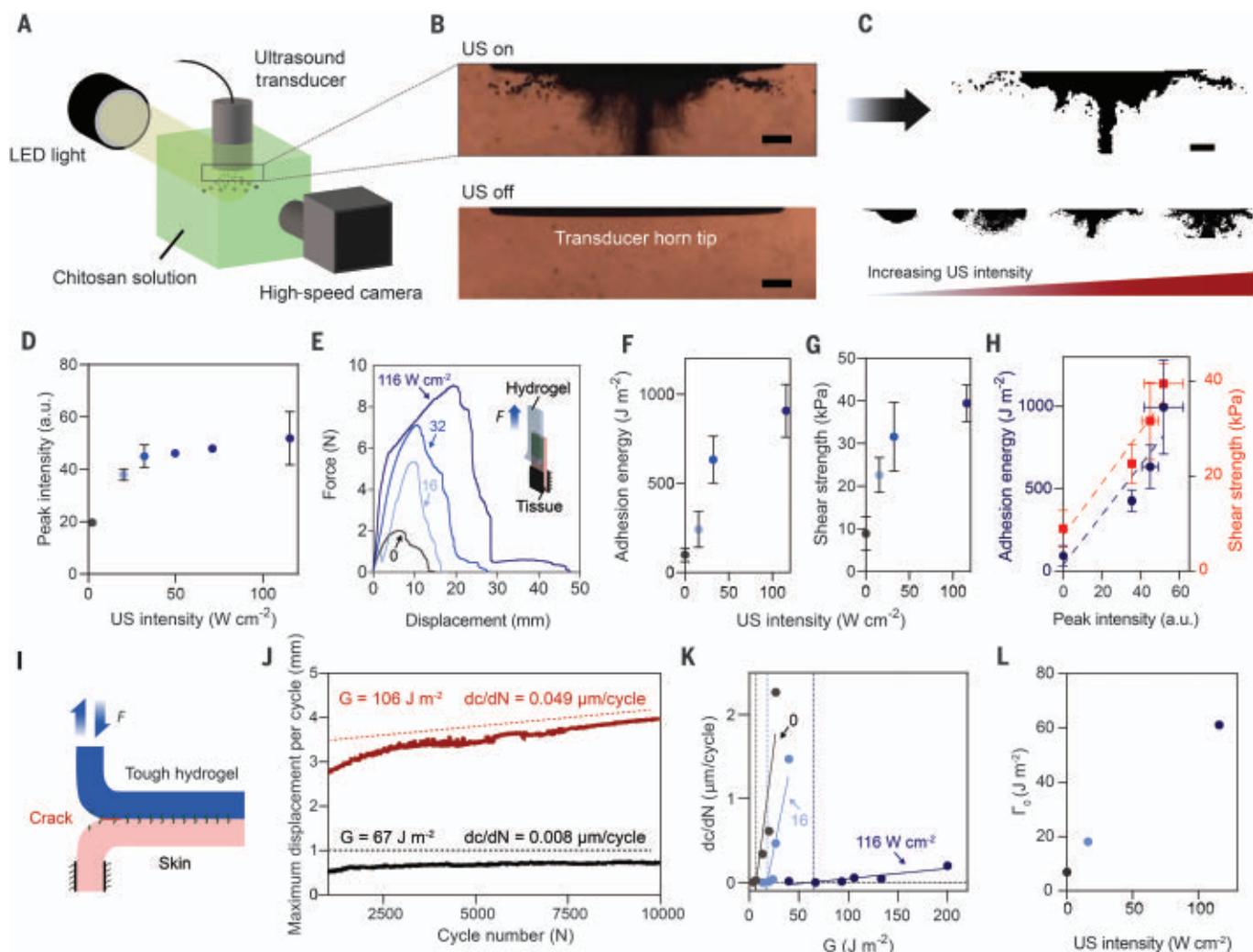


Fig. 2. US-induced cavitation regulates bioadhesion. (A) Experimental setup for characterizing US-induced microbubble cavitation. (B) Digital images of the microbubble cloud at peak intensity in a cycle captured by the high-speed camera. Scale bar, $100 \mu\text{m}$. (C) Processed binary images of the bubble clouds. (D) Normalized peak intensity of the induced microbubble clouds as a function of the US intensity. (E) Representative force-displacement curves of hydrogel-tissue hybrids in lap-shear tests. (F and G) Correlation between applied US intensity and interfacial adhesion energy (F) and shear strength (G). (H) Correlation between peak intensity of

bubble clouds, interfacial adhesion energy, and shear strength. (I) Schematic illustration of the fatigue test of hydrogel-tissue hybrid. Data are shown as means \pm SD for $n = 3$ independent experiments. (J) Representative curves of the cycle number and maximum displacement per cycle at two energy release rates for samples with US treatment. (K) Crack extension rate (dc/dN) versus applied energy release rate $G = F/W$ for hydrogel adhesion on tissues after no US or exposure to low (16 W cm^{-2})– or high (116 W cm^{-2})–intensity US. The linear extrapolation to the G -axis (solid lines) gives the fatigue threshold Γ_0 . (L) Correlation of US intensity and obtained fatigue threshold Γ_0 .

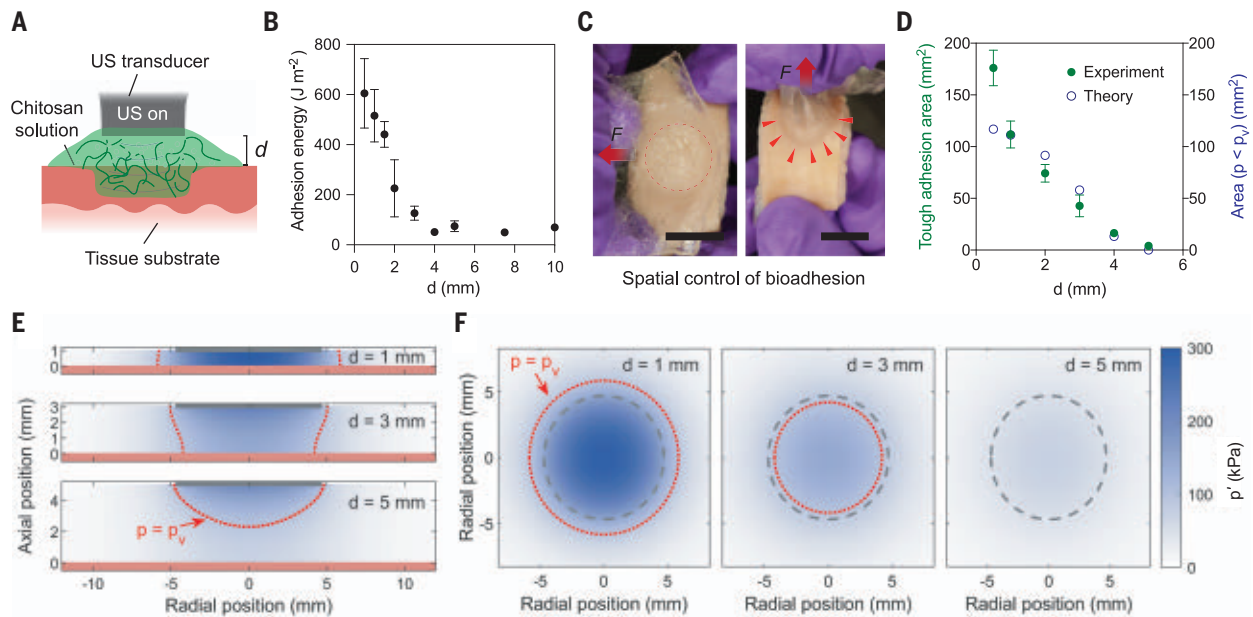


Fig. 3. Spatial control of US-mediated bioadhesion. (A) Schematics of US application at varying distances d between the US horn and the tissue substrate. (B) Correlation of d and the adhesion energy characterized by lap-shear tests. (C) Images of porcine skin where a clear boundary between tough adhesion and nonadhesion regions can be observed. Scale bar, 1 cm. (D) Correlation of d and tough adhesion area measured experimentally (green dot) and

theoretically (blue circle). Data are shown as means \pm SD for $n = 3$ independent experiments. (E) Side view of theoretically computed fields. The area where $p < p_v$ during part of the acoustic cycle is shown by a red dashed line for $d = 1, 3,$ and 5 mm. (F) Corresponding top view at substrate level for $d = 1, 3,$ and 5 mm. Red arrowhead indicates the boundary of tough adhesion area. Blue areas are where total pressure p goes below vapor pressure p_v according to theory.

significant for the low-frequency US used here (see the supplementary text and fig. S15). To test this hypothesis, we combined experiments and theoretical modeling to study the link between cavitation and bioadhesion.

We first characterized cavitation and bioadhesion as a function of US intensity. The cavitation was manifested with dynamic microbubble clouds present under the US transducer (Fig. 2, A and B). A high-speed camera revealed the pattern and geometry of oscillating vapor bubble clouds. The peak cavitation intensity was quantified with the maximum area of microbubble cloud in each cycle (Fig. 2C). The US substantially enhanced the peak intensity of cavitation, but the effect diminished when the intensity was high (Fig. 2D). A similar trend in interfacial adhesion energy was obtained in modified lap-shear tests (Fig. 2E). The adhesion energy and shear strength between hydrogel and porcine skin increased with the US intensity (Fig. 2, F and G), and both were shown to correlate linearly with the peak bubble intensity (Fig. 2H). The adhesion energy scaled almost linearly with the duration of US treatment (fig. S16), further corroborating the correlation between cavitation and bioadhesion.

This correlation can be explained as follows. When the bubbles collapse and jet, the anchoring agents are propelled and anchored into tissue substrates, enabling interfacial bonding with the hydrogel patch. In support of this point, we visualized the single-bubble

dynamics with laser-induced cavitation experiments, showing that the laser-induced bubble collapses and jets toward the tissue substrate (fig. S17 and movie S1); US-induced microbubble jetting was suggested by pit generation in aluminum foil (fig. S18). There might also be contributions from other mechanical effects of the US such as viscous stress, which requires further investigation.

To further study the US effect on interfacial bonding, we performed fatigue fracture tests. Such tests output interfacial fatigue threshold or intrinsic work of adhesion Γ_0 , which scales with the density of interfacial bonds according to the Lake-Thomas theory (18). In fatigue tests under 180° peeling configuration (Fig. 2I), we varied the magnitudes of cyclic loads applied onto the specimens for different energy release rates (G) and monitored the crack extension over cycles (dc/dN) (Fig. 2J). The crack growth rate increased with the loading G (Fig. 2K). The linear regression of the G- dc/dN curves informed the intrinsic work of adhesion (Γ_0). The US treatment raised Γ_0 from $\sim 5 \text{ J m}^{-2}$ (no-US) to $\sim 65 \text{ J m}^{-2}$ (Fig. 2L). With the absence of covalent bonding at the interface, the high Γ_0 exceeded the case of forming interfacial amide bonds through carbodiimide chemistry ($\sim 25 \text{ J m}^{-2}$) (11). These results substantiate the existence of strong interfacial interactions, which are typically only observed with covalent bonding, in contrast to the often weak physical interactions resulting from interdiffusion, such as entanglement.

We next showed that the US treatment mediates tough bioadhesion. Because the US effects scaled with the distance between the transducer and the tissue (d), simply maneuvering the US transducer could control the bioadhesion in magnitude and space (Fig. 3A). The adhesion energy decreased with increasing d and eventually to the level without US when $d > 4$ mm (Fig. 3B). The distance d also mediated the area of bioadhesion (Fig. 3C). The tough adhesion area decreased with the gap between the transducer and the tissue (Fig. 3D). To understand and predict the spatially controlled bioadhesion, we conducted theoretical modeling on the acoustic field produced by the US transducer between the horn and the substrate (see the supplementary text); the experimentally measured displacement of the US transducer was used in the model (fig. S19). We extracted the area on the substrate where the absolute pressure dropped below the vapor pressure in every acoustic cycle, thereby enabling the formation of cavitation bubbles (Fig. 3, E and F). At various values of d , we obtained substantially different pressure profiles on the tissue substrate from which the regions affected by cavitation were estimated (Fig. 3D). The computational results agreed with the experimental measurements. The spatial resolution of bioadhesion could be further improved with focused-US technology (22). In addition to demonstrating spatial control, we demonstrated that US could enable temporal control over bioadhesion by using

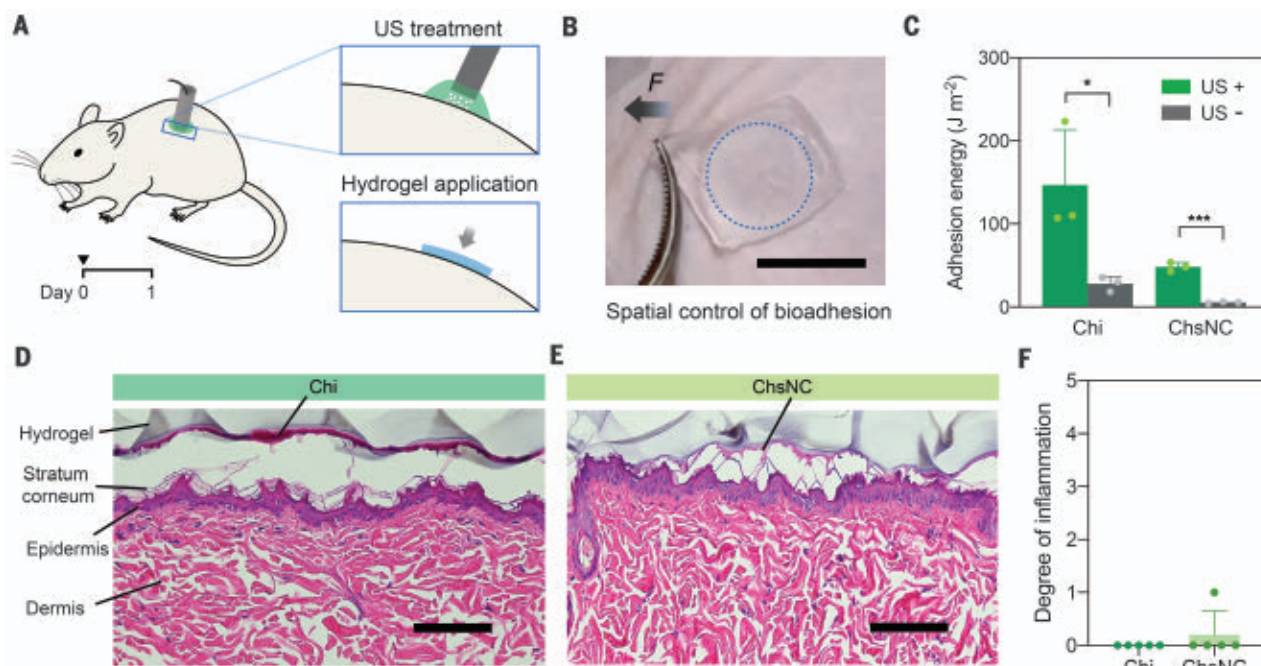


Fig. 4. In vivo assessments. (A) Schematics of the procedures of US-mediated bioadhesion with a rodent model. (B) Spatial control of bioadhesion on rat skin. Blue dotted circle indicates the adhesion region. Scale bar, 1 cm. (C) Adhesion energy between rat skin and hydrogel without or with US treatment (16 W cm^{-2}). Data are shown as means \pm SD for $n = 3$ independent experiments. (D and E) Representative hematoxylin and eosin-stained images of the hydrogel-rat skin

hybrid using chitosan (D) or ChsNC (E) as the anchoring primer, respectively. Scale bar, $100 \mu\text{m}$. (F) Degree of inflammation after a 1-day hydrogel attachment mediated by US; 0 indicates no inflammation; 1, very mild inflammation; 2, mild inflammation; 3, moderate inflammation; 4, severe inflammation; and 5, very severe inflammation. Statistical significance and P values were determined by two-sided Student's t test. * $P < 0.05$; *** $P < 0.001$.

the heating effect of US and a thermo-gelling gelatin as the anchoring primer (see the supplementary text and fig. S20). The controllability is particularly desired because current bioadhesives are limited by homogeneous adhesiveness, poor control over the diffusion of adhesive agents, and complications associated with patterning adhesives (2, 8, 9).

Finally, we validated the safety and efficacy of US-mediated bioadhesion with a rodent model in vivo (Fig. 4A). This animal study is critical because the mechanical index for the high-frequency US safety criteria for imaging applications is not applicable to the low-frequency US used in this study (16). We tested both chitosan and ChsNCs to demonstrate in vivo applicability. Tough bioadhesion formed within minutes selectively on a circular region treated with US, indicative of spatially controlled adhesion (Fig. 4B). By testing freshly excised rat skin, we confirmed the high adhesion energy achieved with US in vivo (Fig. 4C). Histological assessments concluded that there was no marked tissue damage or acute inflammation caused by US (16 W cm^{-2}) or bioadhesives (Fig. 4, D to F). Rodent skin is more sensitive to the heating of high-intensity US compared with porcine skin (see the supplementary text and fig. S21). Future study is needed to establish the US conditions (e.g., intensity and profile) for clinical translation.

Here, we report US-mediated bioadhesion to precisely control hydrogel bioadhesion in space and time. Along with drug-eluting hydrogels, our strategy enables tough bioadhesion and transdermal delivery of protein concurrently (see the supplementary text and fig. S22). The universal applicability of our strategy promises impacts in broad areas ranging from wearable devices to drug delivery.

REFERENCES AND NOTES

- S. Nam, D. Mooney, *Chem. Rev.* **121**, 11336–11384 (2021).
- Z. Ma, G. Bao, J. Li, *Adv. Mater.* **33**, e2007663 (2021).
- J. Li et al., *Science* **357**, 378–381 (2017).
- H. Yuk et al., *Nature* **575**, 169–174 (2019).
- K. Liu et al., *Bioact. Mater.* **13**, 260–268 (2021).
- J. Yang, R. Bai, Z. Suo, *Adv. Mater.* **30**, e1800671 (2018).
- H. R. Brown, T. P. Russell, *Macromolecules* **29**, 798–800 (1996).
- A. Mahdavi et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2307–2312 (2008).
- X. Chen, H. Yuk, J. Wu, C. S. Nabzdyk, X. Zhao, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15497–15503 (2020).
- J. Li, D. J. Mooney, *Nat. Rev. Mater.* **1**, 16071 (2016).
- X. Ni, C. Chen, J. Li, *Extreme Mech. Lett.* **34**, 100601 (2020).
- N. Artzi, T. Shazly, A. B. Baker, A. Bon, E. R. Edelman, *Adv. Mater.* **21**, 3399–3403 (2009).
- C. Demené et al., *Nat. Biomed. Eng.* **5**, 219–228 (2021).
- C. Wang et al., *Nat. Biomed. Eng.* **2**, 687–695 (2018).
- C. Lovegrove, *Nat. Clin. Pract. Oncol.* **3**, 8–9 (2006).
- C. M. Schoellhammer et al., *Sci. Transl. Med.* **7**, 310ra168 (2015).
- S. Y. Yang et al., *Nat. Commun.* **4**, 1702 (2013).
- Z. Yang, Z. Ma, S. Liu, J. Li, *Mech. Mater.* **157**, 103800 (2021).
- S. Rose et al., *Nature* **505**, 382–385 (2014).
- G. Kim et al., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10214–10222 (2019).
- F. Duck, T. Leighton, *J. Acoust. Soc. Am.* **144**, 2490–2500 (2018).
- A. R. Rezaei et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 9180–9182 (2020).

23. C. Bourquard, O. Supponen, Scripts for: Controlled tough bioadhesion mediated by ultrasound, Zenodo (2022); <https://zenodo.org/record/6718010>.

ACKNOWLEDGMENTS

We thank L. Mongeau, A. Higgins, A. Moores, and D. Kurdyla for providing materials and equipment; the Life Sciences Complex Advanced Biolmaging Facility (ABIF) at McGill University for fluorescence image collection and processing; and X. Yang at the University of Kansas for helpful discussions. **Funding:** This work was supported by the Natural Sciences and Engineering Research Council of Canada (grant RGPIN-2018-04146 to J.L. and grant RGPIN-2019-04498 to Y.W.); the Canada Foundation for Innovation (grant 37719 to J.L.); Fonds de Recherche du Québec - Nature et Technologies (grant FRQ-NT NC-270740 to J.L. and B2X doctoral research scholarship 268520 to Z.M.); the Canada Research Chair Program (J.L.); ETH Zürich (O.S. and C.B.); and the National Research Council Canada Industrial Biotechnology program (E.L.). **Author contributions:** Conceptualization: Z.M., J.L.; Investigation: Z.M., C.B., Q.G., S.J., T.D.G., R.H., X.L., Z.H., Z.Y., G.Y., Z.G.; Methodology: Z.M., J.L., C.B., O.S.; Resources: J.L., O.S., Y.W., E.L.; Supervision: J.L., O.S.; Writing – original draft: Z.M., J.L., C.B., O.S.; Writing – review and editing: Q.G., S.J., T.D.G., R.H., X.L., Z.H., Y.W., E.L., Z.G., Z.M., J.L., C.B., O.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials or have been deposited at Zenodo (23). **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn8699
Materials and Methods
Supplementary Text
Figs. S1 to S22
References (24–33)
Movie S1

Submitted 3 January 2022; accepted 22 June 2022
10.1126/science.abn8699

ORGANIC CHEMISTRY

Electron in a cube: Synthesis and characterization of perfluorocubane as an electron acceptor

Masafumi Sugiyama¹, Midori Akiyama^{1*}, Yuki Yonezawa¹, Kenji Komaguchi², Masahiro Higashi³, Kyoko Nozaki¹, Takashi Okazoe^{1,4}

Fluorinated analogs of polyhedral hydrocarbons have been predicted to localize an electron within their cages upon reduction. Here, we report the synthesis and characterization of perfluorocubane, a stable polyhedral fluorocarbon. The key to the successful synthesis was the efficient introduction of multiple fluorine atoms to cubane by liquid-phase reaction with fluorine gas. The solid-state structure of perfluorocubane was confirmed using x-ray crystallography, and its electron-accepting character was corroborated electrochemically and spectroscopically. The radical anion of perfluorocubane was examined by matrix-isolation electron spin resonance spectroscopy, which revealed that the unpaired electron accepted by perfluorocubane is located predominantly inside the cage.

For organic chemists, polyhedral molecules such as cubane (1), dodecahedrane (2), and buckminsterfullerene (3) represent attractive synthetic targets beyond their arguably subjective elegance. After the syntheses and structural characterizations of these polyhedral cages were achieved, interest quickly shifted to the possibility of encapsulating guests within their internal cavities. Early examples included the encapsulation of a variety of atoms in fullerene (4) and dodecahedrane (5) as well as the encapsulation of a molecule in fullerene (6, 7). Meanwhile, polyhedral fluorocarbons (CF)_n—i.e., the perfluorinated analogs of polyhedranes—have

become a focal point of attention and the subject of numerous theoretical studies (Fig. 1A). One of their most interesting features is their electron-accepting character inside their cage, which arises from a stabilized vacant orbital within the cage derived from multiple σ* orbitals of C–F bonds (8–10). This internal localization of electrons stands in stark contrast to common π-conjugated electron acceptors, which usually host electrons on their molecular surfaces. Although some experimental evidence that supports the formation of (CF)_n (where n = 20 or 60) has been reported (11, 12), isolation as a single isomer has not yet been achieved because of the difficulties associated

with the exhaustive fluorination of all cage vertices and the latent instability of (CF)_n on account of the overcrowded fluorine atoms (13–15). By contrast, perfluorocubane [(CF)₈, **1**], whose structure was proposed in 2004 (16), should be sufficiently stable for isolation given that the vicinal fluorine atoms are sterically less hindered than those in larger polyhedranes. Moreover, a theoretical study by Irikura (8) and our own density functional theory (DFT) calculations have suggested that **1** can be expected to exhibit pronounced electron-accepting character [electron affinity = 1.6 eV (8); energy of the lowest unoccupied molecular orbital (E_{LUMO}) = –2.8 eV; Fig. 1B]. Here, we report the synthesis, isolation, and characterization of **1** as a polyhedral fluorocarbon that can accept and store an electron within its internal cubic cavity (i.e., electron in a cube).

As a synthetic route to **1**, the stepwise introduction of eight F atoms onto cubane is impractical; in fact, only two F atoms have been successfully introduced into a cubane scaffold

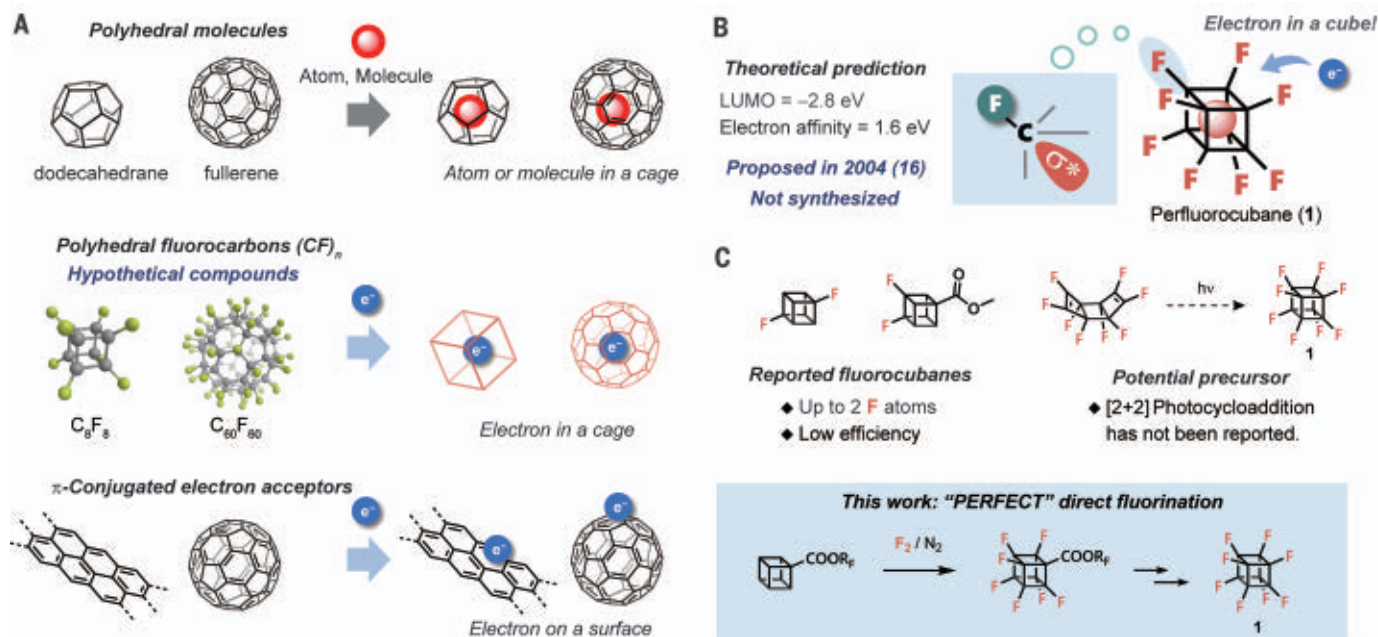


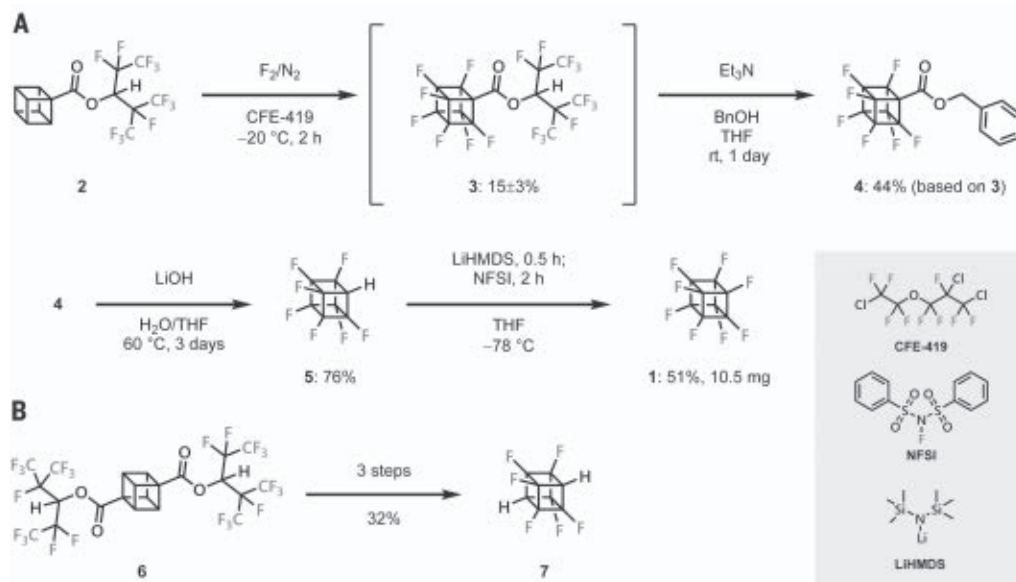
Fig. 1. A perfluorinated cube-shaped molecule accepts an electron inside its cage. (A) Schematic representation of the encapsulation of electrons, atoms, or molecules inside cage-shaped molecules. Polyhedral fluorocarbons can accept an electron inside their cage, whereas common π-conjugated electron acceptors usually store electrons on their surface. (B) Structure of perfluorocubane (**1**) and its predicted properties. The energy level of **1** was calculated at the B3LYP/6-311++G(d,p) level of theory, and the predicted electron affinity has been reported previously (8). (C) Previously reported related molecules and synthetic strategy for this study.

¹Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan. ²Department of Applied Chemistry, Graduate School of Advanced Science and Engineering, Hiroshima University, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8527, Japan. ³Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan. ⁴AGC Inc., Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan.

*Corresponding author. Email: akiyama@moleng.kyoto-u.ac.jp

Fig. 2. Synthetic routes to fluorinated cubanes.

(A) Preparation of heptafluorinated compound **5** and octafluorinated **1**. The yield for **3** in the direct fluorination represents an average value of six experiments, and the error ($\pm\sigma$) was obtained from the standard deviation (σ). THF, tetrahydrofuran; LiHMDS, lithium bis(trimethylsilyl)amide; NFSI, *N*-fluorobenzenesulfonimide; Bn, benzyl. **(B)** Preparation of hexafluorinated compound **7**.



so far (Fig. 1C) (17, 18). Several per-substituted cubanes have been synthesized through the [2+2] photocycloaddition of the corresponding octa-substituted *syn*-tricyclo[4.2.0.0^{2,5}]octa-3,7-dienes (19, 20). Although it is feasible to conceive the synthesis of **1** from *syn*-octafluorotricycloocta-3,7-diene, its conversion to **1** has not yet been reported (21, 22). Against this background, the introduction of multiple fluorine atoms through radical C-H fluorination using fluorine gas might seem the most promising strategy to synthesize **1**. However, a reported attempt to fluorinate cubane using fluorine gas resulted in ring opening owing to the highly strained nature of the cubic skeleton (23). We envisaged that the C-H fluorination of cubane could be achieved using a modified version of the liquid-phase direct fluorination designated the “PERFECT” method, in which the direct fluorination of a partially fluorinated ester with an excess of fluorine gas in a fluorous solvent affords the corresponding perfluorinated product with marked suppression of C-C bond cleavage (24, 25). This synthetic approach enabled the systematic preparation of octa-, hepta-, and hexafluorocubanes, which enabled study of the effect of the degree of fluorination of cubane on its electron affinity.

The synthetic route to **1** is shown in Fig. 2A. A solution of cubanemonoester **2** in 1,2-dichloro-3-(2-chloro-1,1,2,2-tetrafluoroethoxy)-1,1,2,3,3-pentafluoropropane (CFE-419) was treated at -20°C with an excess of fluorine gas. After the reaction, three new peaks were observed in the ^{19}F nuclear magnetic resonance (NMR) spectrum with an area ratio of 3:3:1, which indicated the formation of the heptafluorocubane monoester (**3**) with an average NMR yield of 15% (for details, see the supplementary

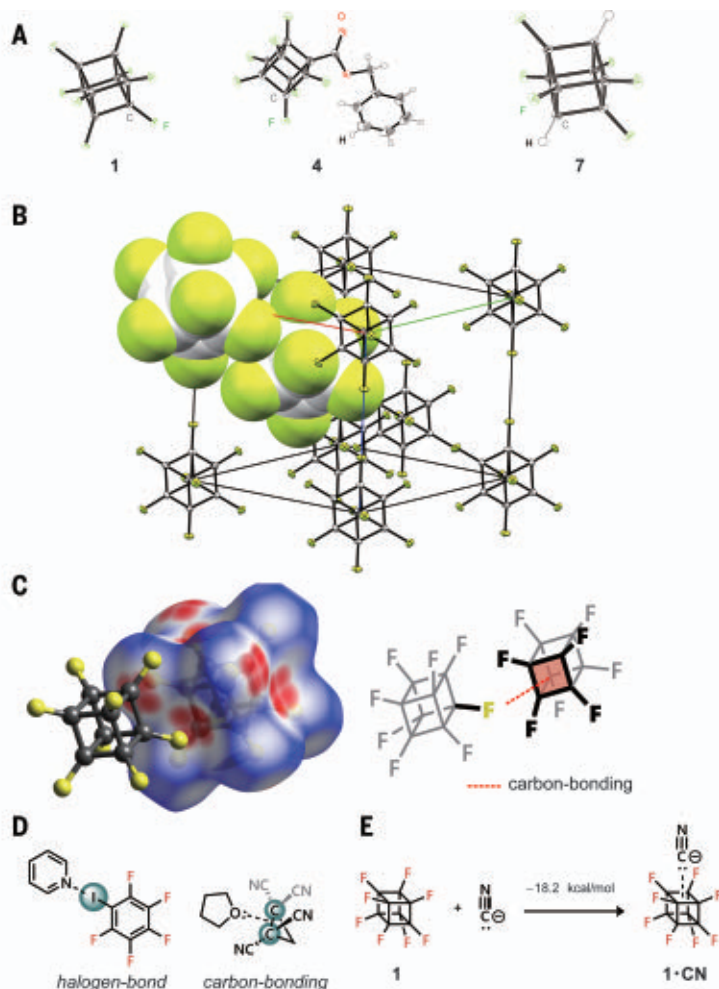
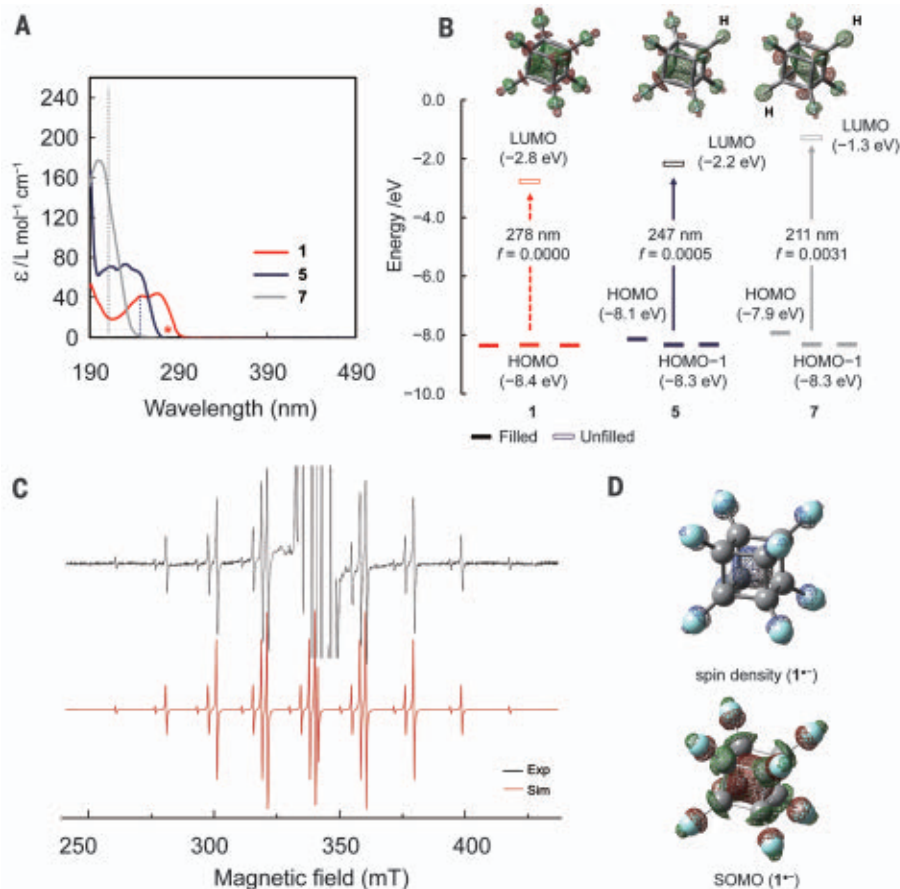


Fig. 3. Crystal structures of fluorocubanes. (A) Thermal ellipsoid plots of **1**, **4**, and **7** (ellipsoid probability, 50%). (B) Packing structure of **1** in the unit cell. Two molecules in the unit cell are depicted using a space-filling model to show the manner of interaction (carbon, white; fluorine, yellow). (C) Hirshfeld surface for dimers of **1** in the crystal structure mapped using d_{norm} (left), together with a hypothetical illustration of the interactions in the dimers (right). (D and E) Schematic illustration of noncovalent σ -hole interactions.

Fig. 4. Properties of fluorocubanes. (A) UV-vis absorption spectra of **1** (red), **5** (blue), and **7** (gray) in MeCN; the vertical dashed lines represent the theoretically predicted absorption coefficients and wavelengths calculated at the B3LYP/6-311++G(d,p) level of theory. The red asterisk indicates the predicted absorption wavelength of **1**, for which the predicted oscillator strength f is 0.00 given that it is symmetry forbidden. (B) LUMOs and the energy levels of frontier orbitals for **1**, **5**, and **7**, calculated at the B3LYP/6-311++G(d,p) level of theory. Red and green areas depict different phases of the plotted orbitals. (C) ESR spectrum of **1** in a hexamethylethane matrix after γ -irradiation at 77 K (black) and simulated ESR spectrum (red) using $g = 1.9985$ and $a(^{19}\text{F}) = 19.62$ mT (8F). Exp, experimental; Sim, simulated. (D) Spin density (contour = 0.005; top) and singly occupied molecular orbital (SOMO) (isovalue = 0.05; bottom) in the radical anion of **1**, calculated at the UB3LYP/6-311++G(d,p) level of theory.



materials). Ester **3** was subsequently subjected to transesterification without isolation, and benzyl heptafluorocubane carboxylate (**4**) was successfully isolated. Upon hydrolysis of **4**, a subsequent decarboxylation furnished heptafluorocubane (**5**; 76% yield). The decarboxylation of perfluorinated tertiary carboxylic acids is known to occur without requiring any other reagents (26). The reaction of **5** with lithium bis(trimethylsilyl)amide (LiHMDS) and *N*-fluorobenzenesulfonimide (NFSI) gave **1** in 51% yield. The ^{19}F and the ^{19}F -decoupled ^{13}C NMR spectra of **1** in acetone- d_6 each feature one singlet peak [^{19}F : -197.19 parts per million (ppm); $^{13}\text{C}\{^{19}\text{F}\}$: 103.8 ppm], and a cross peak between these two peaks was observed in the ^{19}F - ^{13}C heteronuclear multiple quantum coherence (HMQC) spectrum. The Fourier transform infrared (FTIR) spectrum in the range 400 to 4000 cm^{-1} contained only two peaks at 1371 and 798 cm^{-1} , which reflects the high symmetry (O_h point group) of **1**. In the microscopic Raman spectrum, the characteristic peak corresponding to the breathing vibration mode of **1** was observed at 552 cm^{-1} . Hexafluorocubane (**7**) was also synthesized from cubanediester **6** (32%, in three steps) using a similar procedure (Fig. 2B and the supplementary materials).

The structures of fluorocubanes **1**, **4**, and **7** were unambiguously determined using single-crystal x-ray diffraction analysis (Fig. 3A). The identical lengths (1.570 Å) of the 12 C-C bonds in **1** clearly indicate a nondistorted cubic structure in the solid state, and this length was almost identical to a previously reported theoretical value that was obtained from high-level DFT calculations (27). The bond length was also almost identical to that of the parent cubane C_8H_8 (1.572 Å) (28). According to Bent's rule, atoms tend to direct the bonding of hybrid orbitals of greater p character toward electronegative substituents, such as fluorine. As a result, the s character of the C-C bond is increased, and the bond is shortened. By contrast, the average C-C bond expands with increasing numbers of fluorine atoms, which is the result of the repulsion between vicinal fluorine atoms (29). These two opposing effects appear to cancel each other out, and hence **1** shows a C-C bond length that is similar to that of the parent cubane (supplementary materials). In the crystal structure, **1** exhibits a characteristic intermolecular interaction. One fluorine atom of **1** interacts with the center of the cyclobutane ring of the nearest molecule, as depicted using a space-filling model in Fig. 3B; this interaction is also clearly observed

in the Hirshfeld surface analysis (Fig. 3C) (30). The interaction energy was estimated to be 3.5 kcal/mol using natural energy decomposition analysis (31). This type of interaction has recently been recognized as a noncovalent carbon-bonding interaction, which is a subset of tetrel bonding. Specifically, carbon bonding represents an interaction between a Lewis-acidic σ -hole centered on a carbon atom and a Lewis base, akin to halogen bonds (Fig. 3D) (32). The only reported example of a synthesized carbon-bonding donor is tetracyanocyclopropane (33). Bauzá *et al.* have predicted that **1** can be expected to act as a strong carbon-bonding donor that interacts with anions such as cyanide (Fig. 3E) (34), and the packing structure of **1** in the single crystal obtained in the present study strongly supports this prediction.

Fluorocubanes **1**, **5**, and **7** showed ultraviolet-visible (UV-vis) absorption peaks at wavelengths longer than 160 nm owing to their lower-lying LUMOs compared with those of other fluoroalkanes, which are transparent at 160 nm (35). As shown in Fig. 4A, **1** exhibits the longest absorption wavelength edge among the three fluorocubanes, and the edges of **5** and **7** shift hypsochromically with decreasing numbers of fluorine atoms. DFT calculations at the B3LYP/

6-311++G(d,p) level of theory revealed that the LUMO of each fluorocubane comprises the σ^* orbitals of the C–F bonds and that the energy of the LUMO decreases with increasing numbers of fluorine atoms (Fig. 4B). The results of time-dependent DFT calculations suggest that the absorptions of **5** and **7** should most likely be attributed to the transition from the doubly degenerate the second-highest occupied molecular orbital (HOMO–1), which consists of the C–C σ bonds of the cubane scaffold, to the LUMO. The bathochromic shift of the absorption peak of **5** (231 nm) compared with that of **7** (200 nm) is the result of the lower LUMO level of **5** relative to that of **7** and consistent with the calculated values (**5**: 247 nm; **7**: 211 nm). Perfluorocubane (**1**) has a triply degenerate HOMO because of its highly symmetric structure, and the predicted oscillator strength of the HOMO–LUMO transition is 0.00 (Fig. 4B, red asterisk). This can be rationalized in terms of the point group of **1** (O_h), which renders the transition from the HOMO (t_{2u}) to the LUMO (a_{1g}) symmetry forbidden (16). In the experimental UV-vis absorption spectrum of **1**, a weak absorption was observed at a wavelength similar to the calculated HOMO–LUMO transition (278 nm), which can be caused by vibronic coupling (fig. S11 and the supplementary materials).

The reduction potentials of **1** and **5** (**1**: –2.1 V; **5**: –2.5 V) were experimentally determined using differential pulse voltammetry (DPV) measurements (fig. S14). The LUMO levels of **1** and **5** were estimated to be –2.7 and –2.4 eV, respectively, corrected by comparison with the HOMO level of ferrocene (–4.8 eV relative to the vacuum level). Both values are in good agreement with the corresponding calculated values (**1**: –2.8 V eV; **5**: –2.2 eV). In cyclic voltammetry (CV) measurements, both **1** and **5** showed an irreversible reduction wave in the range of –2.1 to –2.9 V (fig. S15), which likely indicates that **1** and **5** accept an electron and that the resulting radical anions are unstable under the applied conditions and decompose.

To confirm the structure of the radical anion generated upon accepting an electron into the LUMO of **1**, derived from the eight C–F σ^* orbitals, we used matrix-isolation electron spin resonance (ESR) spectroscopy in combination with γ -ray radiolysis at low temperatures. This technique has already been used to observe unstable radical anions of perfluorocycloalkanes (36). As shown in Fig. 4C, the ESR spectrum of **1**^{•–} in a hexamethylethane (HME) matrix at 77 K is well resolved. Although the innermost peaks overlap with the intense signals of the HME matrix radicals in the congested central region, the other eight peak bands were observed in accordance with the total nuclear spin quantum number, $M_I = \pm 4, \pm 3, \pm 2$, and ± 1 . Because of the large $a(^{19}\text{F})$ hyperfine

coupling constant (hfcc), the theoretically predicted hyperfine structure characteristic of second-order splitting could be confirmed (37). The isotropic line shape was well simulated by assuming identical coupling of the eight equivalent ^{19}F nuclei. This result suggests that the observed radical anion of **1** rotates rapidly even in the glass-state matrix at 77 K. The hfcc (19.62 mT) is in accord with the value based on the calculated structure (19.24 mT), in which the spin density is mainly distributed inside the cubane cage (Fig. 4D).

Previous polyhedral molecules have found widespread applications in functional materials after their synthesis. Thus, our study, which demonstrates that perfluorinated cage compounds act as electron acceptors, can be expected to pave the way for the molecular design of distinctive functional organic materials [a connection to the environmental accumulation of per- and polyfluoroalkyl substances (PFAS) cannot be discarded at present, and attention should be paid to this issue when devising applications of such fluorinated cubanes] (38).

REFERENCES AND NOTES

- P. E. Eaton, T. W. Cole, *J. Am. Chem. Soc.* **86**, 3157–3158 (1964).
- R. J. Ternansky, D. W. Balogh, L. A. Paquette, *J. Am. Chem. Soc.* **104**, 4503–4504 (1982).
- H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, R. E. Smalley, *Nature* **318**, 162–163 (1985).
- A. A. Popov, S. Yang, L. Dunsch, *Chem. Rev.* **113**, 5989–6113 (2013).
- R. J. Cross, M. Saunders, H. Prinzbach, *Org. Lett.* **1**, 1479–1481 (1999).
- K. Komatsu, M. Murata, Y. Murata, *Science* **307**, 238–240 (2005).
- K. Kurotobi, Y. Murata, *Science* **333**, 613–616 (2011).
- K. K. Irikura, *J. Phys. Chem. A* **112**, 983–988 (2008).
- O. V. Boltalina, in *New Fluorinated Carbons: Fundamentals and Applications*, O. V. Boltalina, T. Nakajima, Eds. (Elsevier, 2017), chap. 1.
- Y.-F. Wang *et al.*, *Theor. Chem. Acc.* **127**, 641–650 (2010).
- F. Wahl *et al.*, *Chem. Eur. J.* **12**, 6255–6267 (2006).
- J. H. Holloway *et al.*, *J. Chem. Soc. Chem. Commun.* **1991**, 966–969 (1991).
- R. Taylor, *J. Fluor. Chem.* **125**, 359–368 (2004).
- J. Jia, C. Liu, H.-S. Wu, P. von Ragué Schleyer, H. Jiao, *J. Phys. Chem. C* **113**, 8077–8084 (2009).
- R. Taylor *et al.*, *Nature* **355**, 27–28 (1992).
- T. Kato, T. Yamabe, *J. Chem. Phys.* **120**, 1006–1016 (2004).
- E. W. Della, N. J. Head, *J. Org. Chem.* **60**, 5303–5313 (1995).
- H. Irrgartinger, S. Strack, *J. Am. Chem. Soc.* **120**, 5818–5819 (1998).
- H. H. Freedman, D. R. Petersen, *J. Am. Chem. Soc.* **84**, 2837–2838 (1962).
- K. F. Biegasiewicz, J. R. Griffiths, G. P. Savage, J. Tsanaktsidis, R. Priefer, *Chem. Rev.* **115**, 6719–6745 (2015).
- M. J. Gerace, D. M. Lemal, H. Ertl, *J. Am. Chem. Soc.* **97**, 5584–5586 (1975).
- A. C. Barefoot III, W. D. Saunders, J. M. Buzby, M. W. Grayston, D. M. Lemal, *J. Org. Chem.* **45**, 4292–4295 (1980).
- J. L. Adcock, H. Zhang, *J. Org. Chem.* **61**, 1975–1977 (1996).
- T. Okazoe, *J. Fluor. Chem.* **174**, 120–131 (2015).
- G. Sandford, *J. Fluor. Chem.* **128**, 90–104 (2007).
- K. Murata, S.-Z. Wang, Y. Morizawa, K. Oharu, *J. Fluor. Chem.* **127**, 1125–1129 (2006).
- S. Berski, A. J. Gordon, Z. Latajka, *J. Phys. Chem. A* **118**, 4147–4156 (2014).

- R. J. Doedens, P. E. Eaton, E. B. Fleischer, *Eur. J. Org. Chem.* **2017**, 2627–2630 (2017).
- J. Cioslowski, L. Edgington, B. B. Stefanov, *J. Am. Chem. Soc.* **117**, 10381–10384 (1995).
- M. A. Spackman, D. Jayatilaka, *CrystEngComm* **11**, 19–32 (2009).
- E. D. Glendening, A. Streitwieser, *J. Chem. Phys.* **100**, 2900–2909 (1994).
- A. Bauzá, T. J. Mooibroek, A. Frontera, *Angew. Chem. Int. Ed.* **52**, 12317–12321 (2013).
- A. Frontera, *C* **6**, 60 (2020).
- A. Bauzá, T. J. Mooibroek, A. Frontera, *Phys. Chem. Chem. Phys.* **16**, 19192–19197 (2014).
- G. Bélanger, P. Sauvageau, C. Sandorfy, *Chem. Phys. Lett.* **3**, 649–651 (1969).
- A. M. ElSohly, G. S. Tschumper, R. A. Crocombe, J. T. Wang, F. Williams, *J. Am. Chem. Soc.* **127**, 10573–10583 (2005).
- R. W. Fessenden, *J. Chem. Phys.* **37**, 747–750 (1962).
- M. G. Evich *et al.*, *Science* **375**, eabg9065 (2022).

ACKNOWLEDGMENTS

We thank R. Takita (The University of Tokyo) for atmospheric pressure chemical ionization mass spectrometry measurements; H. Takiguchi (The University of Tokyo) for microscopic Raman measurements; T. Tajima (Shimadzu) for solid-state UV-vis absorption measurements and analysis; K. Oshida (The University of Tokyo) for melting point measurements; and T. Aida (The University of Tokyo), Y. Ito (The University of Tokyo), H. Sato (RIKEN), K. Aikawa (The University of Tokyo), Y. Ishibashi (AGC Inc.), S. Kusumoto (The University of Tokyo), Y. Takemasa (The University of Tokyo), Y. Nishibayashi (The University of Tokyo), S. Kuriyama (The University of Tokyo), H. Shinokubo (Nagoya University), N. Fukui (Nagoya University), and Y. Segawa (Institute for Molecular Science) for assistance with experiments and fruitful discussions. We also thank U. Mayer at www.mayerscientificediting.com for proofreading. M.S. is grateful to the Program for Leading Graduate Schools (MERIT-WINGS) and to the JSPS for a fellowship for young scientists. A part of this work was conducted in the Research Hub for Advanced Nano Characterization (The University of Tokyo), which is supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT; Japan), and in the One-stop Sharing Facility Center for Future Drug Discoveries in the Graduate School of Pharmaceutical Sciences (The University of Tokyo). Part of the theoretical calculations was carried out at the Research Center for Computational Science, Okazaki, Japan. **Funding:** This work was financially supported by JSPS KAKENHI grants JP19K15532 (M.A.), JP21K14608 (M.A.), JP21J21713 (M.S.), JP20H05839 (M.H.), and JP21H05385 (M.A.); the Toyota Riken Scholar Program; the Hattori Hokokai Foundation; the Iketani Science and Technology Foundation; and the Koyanagi Foundation (M.A.). AGC Inc. is acknowledged for supporting this work. **Author contributions:** M.A. conceived the concept and directed the project. M.S., M.A., and Y.Y. conducted experiments. M.S. synthesized **1**. M.A. and M.S. performed the crystallographic studies and the DFT calculations. K.K. performed the matrix-isolation ESR measurements. M.H. performed Frank-Condon and Herzberg-Teller analysis. All authors analyzed and discussed the results. M.S. and M.A. prepared the manuscript with feedback from all authors. K.N. and T.O. supervised the project. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The x-ray data have been deposited in the Cambridge Crystallographic Data Centre (CCDC) under reference numbers 2144167 to 2144171. All other data are presented in the main text or the supplementary materials. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq0516
Materials and Methods
Supplementary Text
Figs. S1 to S67
Tables S1 to S6
References (39–56)

Submitted 15 March 2022; accepted 7 June 2022
10.1126/science.abq0516

SPEECH EVOLUTION

Evolutionary loss of complexity in human vocal anatomy as an adaptation for speech

Takeshi Nishimura^{1,2*}, Isao T. Tokuda³, Shigehiro Miyachi^{1,2}, Jacob C. Dunn^{4,5,6}, Christian T. Herbst^{1,6}, Kazuyoshi Ishimura³, Akihisa Kaneko^{1,2}, Yuki Kinoshita^{1,2}, Hiroki Koda^{1,†}, Jaap P. P. Saers⁵, Hirohiko Imai⁷, Tetsuya Matsuda⁷, Ole Næsbye Larsen⁸, Uwe Jürgens⁹, Hideki Hirabayashi¹⁰, Shozo Kojima¹, W. Tecumseh Fitch^{6,11*}

Human speech production obeys the same acoustic principles as vocal production in other animals but has distinctive features: A stable vocal source is filtered by rapidly changing formant frequencies. To understand speech evolution, we examined a wide range of primates, combining observations of phonation with mathematical modeling. We found that source stability relies upon simplifications in laryngeal anatomy, specifically the loss of air sacs and vocal membranes. We conclude that the evolutionary loss of vocal membranes allows human speech to mostly avoid the spontaneous nonlinear phenomena and acoustic chaos common in other primate vocalizations. This loss allows our larynx to produce stable, harmonic-rich phonation, ideally highlighting formant changes that convey most phonetic information. Paradoxically, the increased complexity of human spoken language thus followed simplification of our laryngeal anatomy.

Speech is the dominant mode of human linguistic expression, and for most of our evolutionary history, until the emergence of signed and written languages, speech provided the sole communicative modality for language. Speech production is based on the same acoustic and physiological principles as vocal production in other terrestrial vertebrates (1–4) but nonetheless possesses distinctive attributes. First, our vocal source (produced by laryngeal vocal fold oscillations modulating air flow from the lungs; Fig. 1, A and B) is uncharacteristically stable, and nonpathological adult speech completely avoids the nonlinear phenomena and bifurcations to chaos—i.e., irregular oscillations and abrupt frequency transitions—commonly seen in most other mammals (5, 6). This stability, combined with enhanced neural control of laryngeal muscles (7), yields a highly reliable fundamental frequency (f_0 or “pitch”) and rich array of harmonics (5, 8, 9). The predictable, broadband acoustic energy in this source sig-

nal provides the backdrop upon which rapid modulations of our vocal tract filter yield a dynamic pattern of formant frequencies conveying phonetic information. Together, under fine neural control, this human source-filter system provides the high-bandwidth communicative signal needed to rapidly encode complex linguistic information.

Speech-related specializations of the human vocal tract, including the descent of the tongue root into the pharynx, are well documented (10–12), and their acoustic effects well understood (5, 8–11), but evolutionary changes in our larynx have been relatively neglected. First, humans lost the laryngeal air sacs seen in other great apes (6, 13), and which were probably still present in *Australopithecus* (14). Second, we show here that vocal membranes (also known as “vocal lips”)—thin upward projections of the vocal folds (Fig. 1C)—are typical laryngeal features in primates but were lost in humans (Fig. 1B). Using a combination of multidisciplinary methods, we show that vocal membranes increase nonlinearities, yielding vocal instability. This leads to the surprising conclusion that the increased stability of human phonation results from an evolutionary loss of anatomical complexity. Although fossil indicators of vocal fold anatomy are unavailable, our comparative data indicate that this simplification through loss must have occurred since our divergence from chimpanzees roughly 6 million years ago.

We show that a vocal membrane is a key anatomical feature shared by all nonhuman anthropoid primates (“primates” hereafter), including hominoids or apes, cercopithecids or Old World monkeys, and platyrrhines or New World monkeys (see Fig. 1D for phyletic relationships). We used magnetic resonance imaging (MRI) and computed tomography

(CT) to examine larynges from 25 genera and 43 species of primates (table S1). All non-human species possess a vocal membrane and often exhibit a shallow sulcus separating the vocal membrane from the vocal fold (Fig. 1C). Platyrrhines usually have a tall vocal membrane (Fig. 1, D to G, and fig. S1, A to G), whereas cercopithecids have a shorter lip-like membrane (Fig. 1, D and H to J, and fig. S1, H to L). This feature can vary even within a single species in hominoids (Fig. 1, C, D, and K to M; and fig. S1, M to P). In the gibbons and siamang (hylobatids), the vocal membrane extends from the lateral wall of the laryngeal cavity and has become disconnected from the vocal fold (Fig. 1K and fig. S1, M to O).

Given these anatomical data, the most parsimonious evolutionary conclusion is that the vocal membrane is an ancestral primate feature, and that this feature was lost in the human lineage to yield the rounded vocal fold typical of humans (Fig. 1B). Thus, we argue that the absence of a vocal membrane in the human larynx is an evolutionarily derived feature.

Turning to function, the anatomical features of the vocal membrane suggest that it plays a role in phonation. When the glottis (the air space between the vocal folds) is closed for phonation by adducting the arytenoid cartilages, both the vocal folds and membranes move toward the glottal midline, implying that the vocal membranes should also vibrate during phonation in primates.

Supporting this hypothesis, vocal membrane vibration was documented *in vivo* during phonation in a chimpanzee, the closest phylogenetic relative of humans. Hayama *et al.* examined reflex glottal closure using a transnasal fiberoptic in an adult male chimpanzee under anesthesia (15). Serendipitously, these video data also documented phonation during 375 grunts or growls as the chimpanzee was awakening from anesthesia. We reexamined these data and found that the vocal membranes always vibrate during vocalization, colliding to close the glottis (Fig. 2A and movie S1). These observations suggest that, in chimpanzees, the vocal membrane plays a central role in phonation even during low-frequency calls and is not simply an accessory feature of the vocal folds subserving high-pitch vocalization as previously thought (5, 16, 17).

Because such *in vivo* evidence is very challenging to obtain in chimpanzees, we next quantitatively investigated the role of vocal membranes in an *ex vivo* setting, using excised larynges from three chimpanzees. Each larynx was mounted on a vertical tube supplying airflow, and the adductory gestures observed *in vivo* were reproduced by positioning the arytenoid cartilages with adjustable prongs. Adductory conditions were further modified by pulling the thyroid cartilage antero-inferiorly. Vocal membrane and fold vibrations driven by

¹Primate Research Institute, Kyoto University, Inuyama, Aichi 484-8506, Japan. ²Center for the Evolutionary Origins of Human Behavior, Kyoto University, Inuyama, Aichi 484-8506, Japan.

³Department of Mechanical Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan. ⁴Behavioural Ecology Research Group, School of Life Science, Anglia Ruskin University, Cambridge CB1 1PT, UK. ⁵Department of Archaeology, University of Cambridge, Cambridge CB2 3DZ, UK. ⁶Department of Behavioral and Cognitive Biology, University of Vienna, 1030 Vienna, Austria. ⁷Department of Systems Science, Graduate School of Informatics, Kyoto University, Sakyo, Kyoto 606-8501, Japan. ⁸Department of Biology, University of Southern Denmark, DK-5230 Odense M, Denmark. ⁹Section of Neurobiology, German Primate Center, D-37077 Göttingen, Germany. ¹⁰Dokkyo Medical University, Mibu, Tochigi 321-0293, Japan. ¹¹Cognitive Science Hub, University of Vienna, Vienna, Austria.

*Corresponding author. Email: nishimura.takeshi.2r@kyoto-u.ac.jp (T.N.); tecumseh.fitch@univie.ac.at (W.T.F.)

†Present address: Department of Life Science, Graduate School of Arts and Sciences, The University of Tokyo, Meguro, Tokyo 153-8902, Japan.

varying air pressure were documented by using high-speed video, and phonatory dynamics were simultaneously documented by using time-synchronized acoustic and electroglot-tographic recordings.

The vocal membranes always participated in vibration in these chimpanzee ex vivo experiments (Fig. 2, B to G, and movie S2, A to C) and oscillated in a wave-like fashion in the absence of vocal fold vibration (Fig. 2, B and C, and movie S2A). The vocal membranes also vibrated separately from—but in phase with—the vocal folds (Fig. 2, D and E, and movie S2B). In such cases, the vocal membranes usually collided, but the vocal folds rarely did so. Only when very firmly adducting and lowering the arytenoid vocal process did we observe vocal fold collision resembling that of humans. In such cases, there was sometimes a phase-delayed oscillatory pattern, in which the vocal folds first collided and the vocal membranes followed (Fig. 2, F and G, and movie S2C). Thus, adducting the arytenoid always resulted in vocal membrane vibration and often collision. By contrast, vibrations of the vocal folds alone, without the vocal membrane, were never observed in these experiments. In sharp contrast, vocal fold vibrations always play a central role in phonation in human speech.

Further in vivo experiments with two rhesus macaques (cercopithecids) and two squirrel monkeys (platyrrhines) corroborated the observations in chimpanzees. We successfully induced vocalization by electrical stimulation to the periaqueductal gray and surrounding areas of the midbrain in these monkeys under anesthesia. In both species, high-speed video recordings showed that the vocal membranes always vibrated and, in many cases, collided during phonation (Fig. 2, H and I, and movies S3 to S4). In macaques, as in chimpanzees, the vocal folds did not always vibrate (fig. S2, A and B, and movie S5A), but the second subject did show simultaneous vibrations of both vocal membranes and folds, either in phase (fig. S2, C and D, and movie S5B) or phase delayed (fig. S2, E and F, and movie S5C). Further ex vivo experiments with six macaque larynges reproduced the same patterns (fig. S2, G to L, and movie S6). Humanlike vibration of the vocal folds alone was not observed in macaques ex vivo. These physiological data, combined with previous work, provide clear empirical evidence that vocal membranes constitute the predominant vocal source generator in primates, always vibrating and typically colliding, whereas the contributions of the vocal folds are limited or even absent.

Taken together, the evolutionary loss of the vocal membrane transformed the predominant source generator from the vocal membranes to the vocal folds in the human lineage. However, understanding the adaptive importance of the derived human condition re-

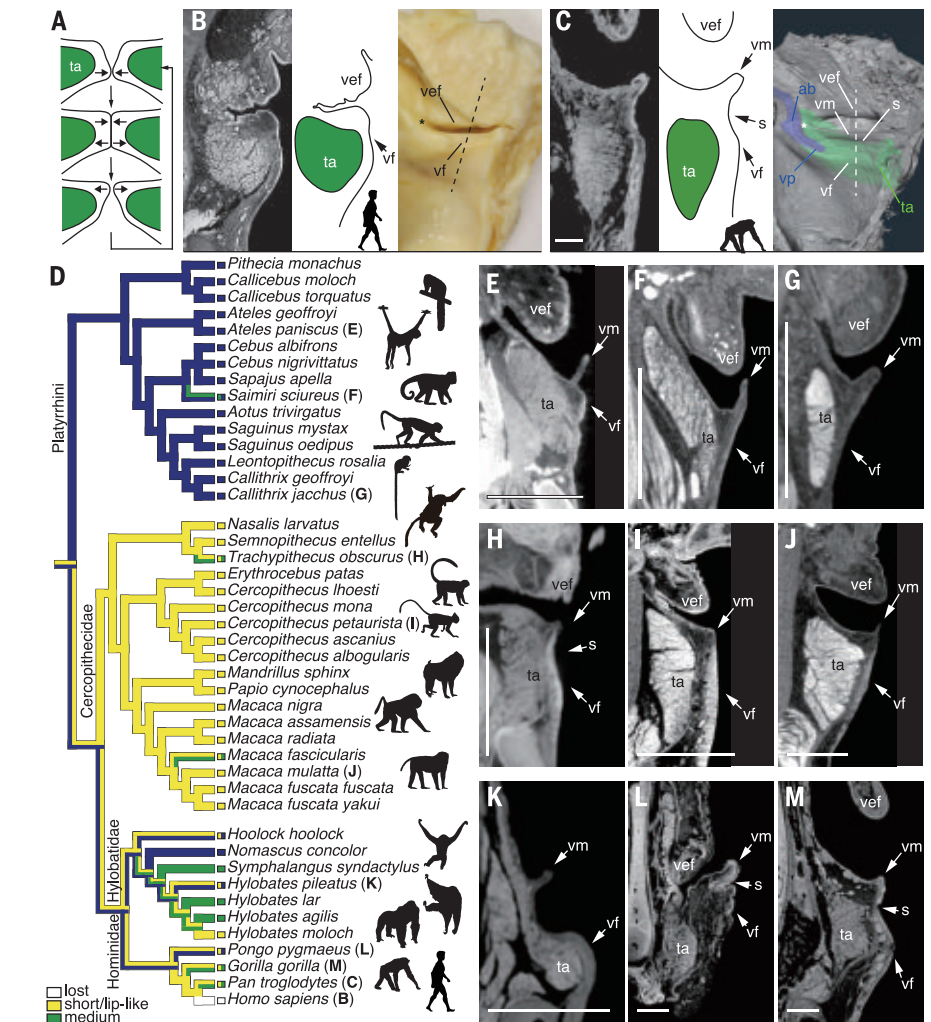


Fig. 1. Vocal anatomy and its phylogeny in anthropoids. (A) The cycle of vocal fold vibration in wave fashion in humans. (B) Frontal section (gradation inversion, left) and a corresponding line drawing (middle) at the level of the dashed line on an excised and formalin-fixed larynx (right) in humans (courtesy of K. Sato). (C) Frontal MRI scan (left) and a corresponding line drawing (middle) at the level of the dashed line on vocal anatomy reconstructed from MRI scans (right) for a chimpanzee, *Pan troglodytes*. Blue indicates the underlying arytenoid cartilage and green the thyroarytenoid muscle. (D) Variation and phylogeny of the vocal membrane and (E to M) MRI-CT frontal scans in the species labeled on (D). Transverse scale bar, 5 mm; vertical scale bar, 2.5 mm. Key: ab, body of the arytenoid cartilage; s, sulcus; ta, thyroarytenoid muscle; vef, ventricular fold; vf, vocal fold; vm, vocal membrane; and vp, vocal process of the arytenoid cartilage. The asterisk (*) indicates a posterior commissure of the vocal membrane and fold.

quired further investigation of vocal membrane function. Although many primate species can produce clear harmonically structured calls [e.g., marmoset pheeas (17), macaque coos (18), or chimpanzee hooos (19)], even these call types often bifurcate to subharmonics or chaos at higher intensities (6, 17). Our findings below indicate that the primate vocal membrane plays an important acoustic role by increasing susceptibility to nonlinear phenomena.

Subharmonics and chaos were observed in vivo in both our rhesus macaque (fig. S3, A to C, and movie S5, C and D) and squirrel monkey data (fig. S3, D and E, and movies S7

and S8), and also ex vivo in our chimpanzee (Fig. 2, F and G, and movie S2C) and rhesus macaque data (fig. S2, I and J, and movie S6B). The chaotic episodes were associated with highly irregular vibrations and collisions of the vocal membranes, superimposed upon nearly periodic vibrations of the vocal folds in chimpanzees ex vivo (Fig. 3, B and C, and movies S2D and S3E) and macaques in vivo (fig. S3, A to C, and movie S5, C and D). This suggests that vocal membrane vibrations play a crucial role in generating the high proportion of subharmonics and chaos empirically observed in nonhuman primate vocalization.

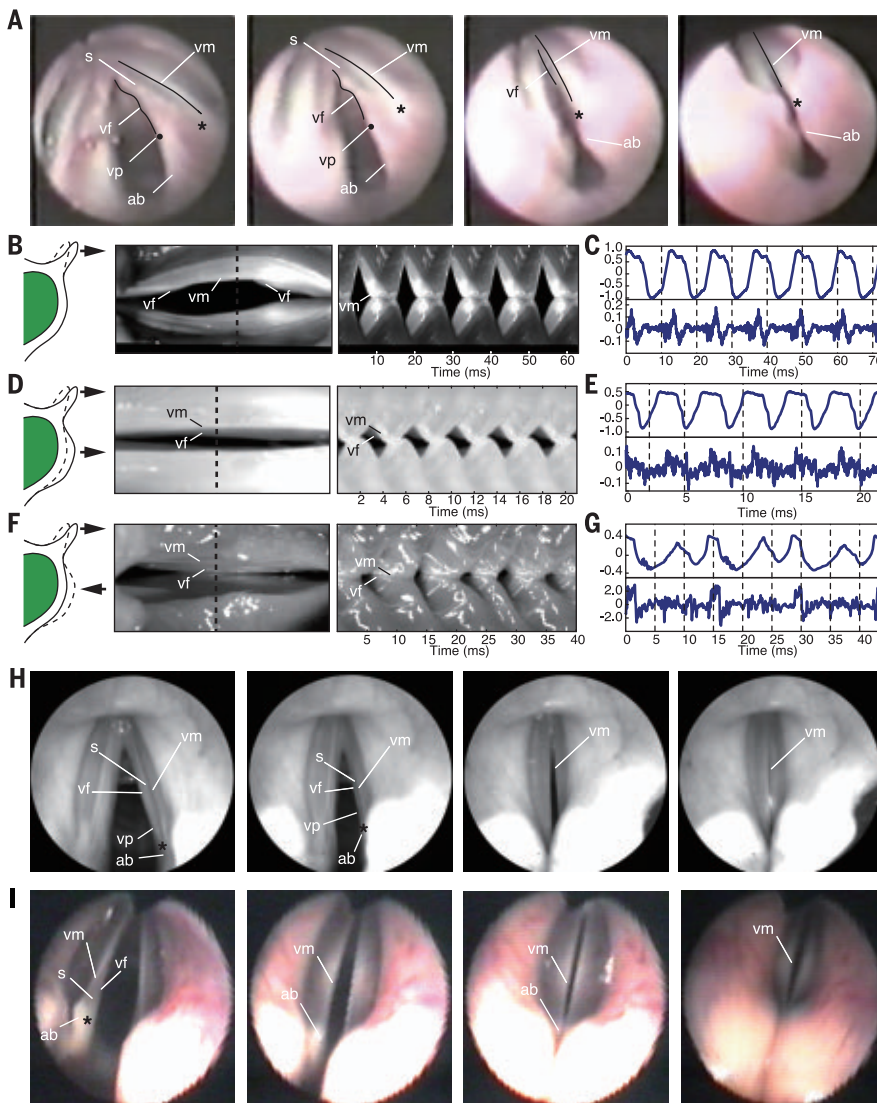


Fig. 2. Vocal membrane and vocal fold vibrations. (A) Images extracted from video recordings in vivo in a chimpanzee. From left to right: The vocal membranes and folds are maximally separated, the glottis is closing, and the vocal membranes close the glottal space and vibrate. (B to G) Vibration of the vocal membrane and vocal fold ex vivo in chimpanzees. (B, D, and F) A top-view image of the glottis extracted from video recordings with the dotted line for creating the kymograph (left) and kymograph showing a time series of glottal vibration (right). (C, E, and G) Electroglottograph (upper panel) and acoustic (lower panel) signals. (B and C) Only the vocal membranes vibrate, whereas the vocal folds do not; and (D and E) vocal membrane and fold both vibrate in-phase and (F and G) out-of-phase. (H and I) Images extracted from video recordings in vivo in a rhesus macaque, *Macaca mulatta* (H) and a common squirrel monkey, *Saimiri sciureus* (I). See Fig. 1 for the anatomical key.

We used a mathematical model to determine the acoustic effects of vocal membrane vibration in primates compared with vocal fold vibration alone, as typifies humans. In our model, the vocal fold is represented by two masses coupled with springs (3, 16, 20–22) (Fig. 3D). The vocal membrane is modeled as an additional reed-like plate that can vibrate independently, attached to the upper mass (21) (Fig. 3D). We first confirmed that all three empirically observed vibration patterns can be reproduced by the numerical simulations in the nonhuman model with vocal membranes (fig. S3F). Next, we generated bifurcation diagrams by adducting the vocal membranes and folds in the model (Fig. 3, E and F). Notably, the phonation threshold pressure observed at 0.14 kPa in the nonhuman model was much lower than that of the human model lacking vocal membranes (0.19 kPa). This is consistent with previous theoretical models of a fixed vocal membrane (16) and

empirical data from marmosets (17), suggesting that the vocal membrane provides improved efficiency in phonation and/or allows louder and higher-frequency vocalizations in nonhuman primates.

In bifurcation diagrams, we showed that the nonhuman model gives rise to various nonlinear phenomena, including subharmonics and chaos, with increasing adduction (Fig. 3E), reproducing the chaotic episode observed in chimpanzees ex vivo (Fig. 3, B and C, and movies S2D and S3E). Similar bifurcations were observed with increasing subglottal pressure in the nonhuman model (fig. S3G). This contrasts with the situation in our human model, which generated only stable periodic vibrations regardless of increasing adduction (Fig. 3F) and pressure (fig. S3H).

The increased instability seen in the nonhuman model is not surprising, because the addition of mechanically coupled vocal membranes contributes additional degrees of

freedom to laryngeal biomechanics. This dynamically generates multiple routes to voice instabilities, as seen in many coupled nonlinear systems (22–24). Thus, the simulations confirmed our empirical observations, indicating that the primate vocal membrane, combined with vocal fold vibration, destabilizes the vocal source. Their interaction can spontaneously lead to subharmonics and chaos in response to simple linear variations in glottal and respiratory parameters. The loss of the vocal membrane in humans therefore reduces the risk of contaminating the stable vocal fold oscillations used in human speech or singing with chaotic irregularities and noise.

In summary, the primate vocal membrane interacts with the vocal fold beneath it to efficiently generate phonation but simultaneously readily generates nonlinear phenomena that destabilize the vocal source. The resulting spontaneous phenomena are unlikely to be under the animal's volitional control, but this

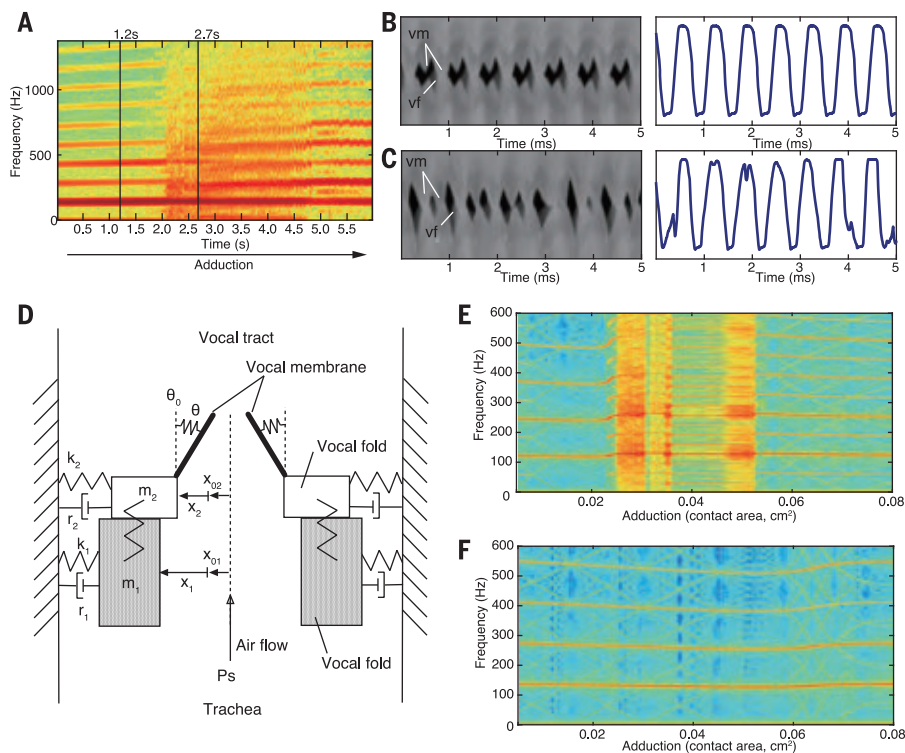


Fig. 3. Bifurcation observed and numerical simulation of mathematical model. (A) Sound spectrograph showing bifurcation to chaos ex vivo in a chimpanzee. (B) Kymographs (left) and electroglottograph signals (right) for a stable phonation at 1.2 s of (A); and (C) for the chaotic phonation at 2.7 s of (A). (D) Schematic illustration of the nonhuman model with a vocal membrane. (E) Bifurcation diagram created by slowly adducting (increasing the contact area) for the nonhuman and (F) the human models. See Fig. 1 for the anatomical key.

more complex laryngeal anatomy enables non-human primates to efficiently generate a diverse set of complex vocal signals without the need for correspondingly complex motor control.

By contrast, despite its disadvantages for efficiency, we suggest that the stable low-pitched phonation resulting from the loss of the vocal membranes in human evolution enhances the detectability and salience of formant frequencies, which carry most of the phonetic information in human speech. Chaos and subharmonics generate an irregular and changing spectrum thought to obscure the conformation of the vocal tract filter (18, 25) by introducing spurious spectral peaks or “pseudo-formants” (26). Although nonlinear phenomena can occur in humans—e.g., in screams and baby cries (27)—they are rarely observed during speech except in pathological cases (28). We conclude that, in addition to changes in vocal tract anatomy and neural control, this previously unnoticed simplification of human laryngeal anatomy allowed our species to easily generate and regulate complex, high-bandwidth vocalizations under cognitive control in speech. Thus, seemingly paradoxically, a loss of complexity in human laryngeal anatomy, coupled with enhanced

neural control mechanisms, represented a crucial evolutionary transition supporting the sophistication and complexity of human spoken language.

REFERENCES AND NOTES

- A. M. Taylor, D. Reby, *J. Zool.* **280**, 221–236 (2010).
- C. T. Herbst *et al.*, *Science* **337**, 595–599 (2012).
- H. Koda, I. T. Tokuda, M. Wakita, T. Ito, T. Nishimura, *J. Acoust. Soc. Am.* **137**, 3068–3076 (2015).
- C. P. Elemans *et al.*, *Nat. Commun.* **6**, 8978 (2015).
- W. T. Fitch, *Trends Cogn. Sci.* **4**, 258–267 (2000).
- W. T. Fitch, J. Neubauer, H. Herzel, *Anim. Behav.* **63**, 407–418 (2002).
- U. Jürgens, *Neurosci. Biobehav. Rev.* **26**, 235–258 (2002).
- I. R. Titze, *Principles of Voice Production* (Prentice Hall, 1994).
- K. N. Stevens, *Acoustic Phonetics* (MIT Press, 1998).
- P. H. Lieberman, D. H. Klatt, W. H. Wilson, *Science* **164**, 1185–1187 (1969).
- P. Lieberman, *Curr. Anthropol.* **48**, 39–66 (2007).
- T. Nishimura, A. Mikami, J. Suzuki, T. Matsuzawa, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6930–6933 (2003).
- S. Hayama, *J. Anthropol. Soc. Jpn.* **78**, 274–298 (1970).
- Z. Alemseged *et al.*, *Nature* **443**, 296–301 (2006).
- S. Hayama, in *Primate Today*, A. Ehara, T. Kimura, O. Takenaka, M. Iwamoto, Eds. (Elsevier, 1991), pp. 491–492.
- P. Mergell, W. T. Fitch, H. Herzel, *J. Acoust. Soc. Am.* **105**, 2020–2028 (1999).
- Y. S. Zhang, D. Y. Takahashi, D. A. Liao, A. A. Ghazanfar, C. P. H. Elemans, *Nat. Commun.* **10**, 4592 (2019).
- D. Rendall, M. J. Owren, P. S. Rodman, *J. Acoust. Soc. Am.* **103**, 602–614 (1998).

- P. Marler, R. Tenaza, in *How Animals Communicate*, T. A. Sebeok, Ed. (Indiana Univ. Press, Bloomington, 1977), pp. 965–1033.
- K. Ishizaka, J. L. Flanagan, *Bell Syst. Tech. J.* **51**, 1233–1268 (1972).
- J. Neubauer, “Nonlinear dynamics of the voice: Bifurcations and mode analysis of complex spatio-temporal signals,” thesis, Humboldt University of Berlin (2004).
- I. T. Tokuda, *Anthropol. Sci.* **126**, 35–41 (2018).
- P. Berg, Y. Pomeau, C. Vidal, *Order within Chaos: Towards a Deterministic Approach to Turbulence* (Wiley, 1986).
- L. Glass, M. C. Mackey, *From Clocks to Chaos: The Rhythms of Life* (Princeton Univ. Press, 1988).
- M. J. Owren, D. Rendall, *J. Comp. Psychol.* **117**, 380–390 (2003).
- W. T. Fitch, J. B. Fritz, *J. Acoust. Soc. Am.* **120**, 2132–2141 (2006).
- W. Mende, H. Herzel, K. Wermke, *Phys. Lett. A* **145**, 418–424 (1990).
- H. Herzel, D. Berry, I. R. Titze, M. Saleh, *J. Speech Hear. Res.* **37**, 1008–1019 (1994).

ACKNOWLEDGMENTS

We are grateful to the late Sugio Hayama for conceiving and conducting in vivo chimpanzee experiments and for establishing the collection of excised larynges at the Japan Monkey Centre (JMC). We thank the JMC, Nagoya Higashiyama Zoo, Fukuoka City Zoological Garden, Kagoshima City Hirakawa Zoological Park, Yokohama Zoological Gardens, and the Wildlife Research Center, the Laboratory of Physical Anthropology, and the Primate Research Institute (PRI) of Kyoto University for providing samples. We thank T. Udono, Y. Shintaku, K. Matsui, H. Imai, S. Tsubouchi, and A. Kataoka for preparing or scanning the specimens; Y. Nomura and Z. Chen for examining laryngeal anatomy; K. Migimatsu, T. Matsumoto, M. Kanaya, and R. Miyazaki for conducting ex vivo experiments; C. R. Larson for providing suggestions for in vivo experiments of macaques; and T. Miyabe-Nishiwaki, Y. Sawada, B. Bach Andersen, and present and past staff of the PRI and the German Primate Center for help conducting in-vivo experiments and/or daily care of subjects. We appreciate the photograph courtesy of K. Sato at the Kurume University. This research was in part supported by the Kyoto University–University of Vienna Strategic Partnership Program. **Funding:** Japan Society for the Promotion of Science grants 16H04848 and 19H01002 (T.N.), 17H06313 and 20K11875 (I.T.), and 18H03503 (H.K.) Research Units for Exploring Future Horizons through the Kyoto University Research Coordination Alliance (T.N., C.T.H.) Rhinology and Laryngology Research Fund UK (J.C.D.) Ministry of Education, Culture, Sports, Science and Technology Grant-in-aid for Scientific Research on Innovative Areas no. 4903 (Evolinguistics) grant 17H06380 (H.K.) Ministry of Education, Culture, Sports, Science and Technology Grant-in-aid for Scientific Research on Innovative Areas no. 4903 (Evolinguistics) (W.T.F.) Austrian Science Fund DK Grant “Cognition & Communication 2” grant W1262-B29 (W.T.F.). **Author contributions:** Conceptualization: T.N., W.T.F. Methodology: T.N., C.T.H., I.T.T., S.M., J.C.D., U.J., W.T.F., O.N.L. Formal analysis: T.N., I.T.T., W.T.F. Investigation: T.N., W.T.F., I.T.T., S.M., J.C.D., C.T.H., K.I., A.K., Y.K., H.K., J.P.P.S., H.I., T.M., O.N.L., U.J., H.H., S.K. Resources: T.N., A.K., U.J., S.K., O.N.L. Writing – original draft: T.N., I.T.T., S.M., J.C.D., C.T.H., W.T.F. Writing – review and editing: T.N., I.T.T., W.T.F. Project administration: T.N. Funding acquisition: T.N., I.T.T., J.C.D., H.K., W.T.F. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Laryngeal specimens from the subjects named Baran, Keiko, Willie, and Yuri were available from the zoos under a material transfer agreement with the Kyoto University. Some specimens were provided through the Collaborative Research program of the JMC (no. 2018017) and the Great Ape Information Network under the National BioResource Project (NBRP) of Japan. All data are available in the main text or the supplementary materials. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm1574
Materials and Methods
Supplementary Text
Figs. S1 to S4
Table S1
References (29–35)
Movies S1 to S8

Submitted 30 August 2021; accepted 30 June 2022
10.1126/science.abm1574

INSECT MIGRATION

Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth

Myles H. M. Menz^{1,2,3*}, Martina Scacco^{1,3}, Hans-Martin Bürki-Spycher⁴, Hannah J. Williams^{1,3}, Don R. Reynolds^{5,6}, Jason W. Chapman^{7,8,9*}, Martin Wikelski^{1,3,10*}

Each year, trillions of insects make long-range seasonal migrations. These movements are relatively well understood at a population level, but how individual insects achieve them remains elusive. Behavioral responses to conditions en route are little studied, primarily owing to the challenges of tracking individual insects. Using a light aircraft and individual radio tracking, we show that nocturnally migrating death's-head hawkmoths maintain control of their flight trajectories over long distances. The moths did not just fly with favorable tailwinds; during a given night, they also adjusted for head and crosswinds to precisely hold course. This behavior indicates that the moths use a sophisticated internal compass to maintain seasonally beneficial migratory trajectories independent of wind conditions, illuminating how insects traverse long distances to take advantage of seasonal resources.

Insect migration takes place on an enormous scale, with trillions of individuals performing bidirectional seasonal movements that have important impacts on ecosystem function and provision of essential services (1–5). However, the navigational mechanisms and behavioral strategies used by night-flying migrants, especially larger nocturnal lepidopterans (macromoths), during these long-range journeys have been unknown for more than 100 years.

The view in the first half of the 20th century, promoted by C. B. Williams, was that migrant moths controlled their movement direction irrespective of the wind and maintained straight flightpaths over long distances (6, 7). Empirical evidence of persistent, self-directed tracks was lacking, however, and by the second half of the 20th century, C. G. Johnson and L. R. Taylor downplayed the importance of orientation behavior and emphasized the role of wind in determining migratory trajectories (8, 9). The modern view has swung back again, because radar observations of free-flying migrants (10–12) and experimental manipulation of tethered individuals (13, 14) have both clearly demonstrated that nocturnally migrating moths can select adaptive headings

and modify them with respect to ambient wind conditions. However, owing to the methodological constraints of tracking such small animals over long distances at night (15), individual moths have never been tracked throughout their migration, and so the capability of these migrants to maintain straight flight paths, over long distances and in seasonally beneficial directions, is unknown.

We used animal-borne radio telemetry to record complete tracks of individually tagged moths over a full night during autumn migration, within the context of the fine-scale wind fields experienced as they migrated southward through the Alps of central Europe. Our study species, the death's-head hawkmoth (*Acherontia atropos*, Sphingidae; Fig. 1A), is Europe's largest lepidopteran, with a rich folklore that stems from its sinister skull-like thoracic markings, unusual habit of raiding beehives to steal honey, and startling acoustic capabilities (16, 17). *A. atropos* is a long-distance Afro-Palaearctic migrant, arriving to breed in Europe north of the Alps each spring. The subsequent generation returns south the following autumn to winter-breeding regions in the Mediterranean Basin and likely also sub-Saharan Africa (16, 17), covering a distance of up to 4000 km. The moths are extremely large for flying insects, weighing up to 3.5 g [2.65 ± 0.15 g (mean \pm SE), $n = 14$] and capable of carrying tiny very high frequency (VHF) radio transmitters. We used a light aircraft (Cessna 172) to track hawkmoths fitted with transmitters (Fig. 1A) and recorded precise (± 150 m) GPS locations from the aircraft (18) at regular intervals throughout their migration (every 5 to 15 min, when possible).

We recorded nocturnal migratory flights of 14 moths, with eight at high spatiotemporal frequency, as they migrated toward the Mediterranean (Fig. 1, B and C, and table S1). Moths initiated migration at a similar time after sunset (62 ± 4.9 min, range 42 to 81 min,

$n = 8$) and were then followed for a minimum of 1 hour and up to 3.65 hours (2.5 ± 0.30 hours, $n = 8$; table S1). The moths were followed for a mean distance of 62.7 ± 6.7 km ($n = 8$) and up to 89.6 km (Fig. 1C and tables S1 and S2), the longest distance over which any insect has been continuously tracked in the field. The overall migration direction was toward the south-southwest [Rayleigh test: $208.70^\circ \pm 0.42^\circ$ (mean \pm SD), mean vector length (r) = 0.917, $P \leq 0.001$, $n = 14$; Fig. 1D]. This track direction is very similar to the preferred headings of a range of migratory insects (moths, butterflies, and hoverflies) that were observed with radar in Western Europe (2, 10, 19, 20), including hawkmoths (10), all of which likely follow a similar western route to the Mediterranean or northwest Africa.

We obtained detailed tracks for seven of these moths, each with three or more locations in a single night (table S2). Moths traveled with a mean ground speed of 9.4 ± 0.4 m/s (33.8 km/hour; $n = 99$ segments; Fig. 1E) and a maximum recorded ground speed of 19.4 m/s (69.7 km/hour). The mean ground speed recorded (Fig. 1E) is consistent with what we expect the upper limit of self-powered flight in *A. atropos* to be (21), suggesting that moths modulated their self-powered airspeed and/or received relatively modest wind assistance. Although there was variation in individual migration direction, all moths maintained straight tracks (straightness index: mean = 0.95, range = 0.80 to 0.99, $n = 7$; Fig. 1C and table S2) along their entire flight paths, which lasted many tens of kilometers, despite being subjected to winds of varying strength and direction throughout their course (Fig. 2). Two of the seven moths evidently crossed the Alps during a single night, because they were relocated south of the Alps during searches early the following morning. Their locations were consistent with their individual trajectories recorded the preceding night, suggesting that they had maintained straight tracks even while transiting the Alps [covering distances of 173.9 and 161.8 km from the release point (Fig. 1C and table S1)].

To answer the question of how moths are able to maintain straight tracks relative to the ground while exposed to varying winds, we calculated the distribution of the angle of deviation, β (the difference between the track and the downwind direction), to determine the extent to which the self-powered heading influenced the trajectory (22). The analysis revealed that moths used three distinct behavioral strategies, which resulted in the flight paths of the moths grouping into three directional clusters (Fig. 1, C and D). These clusters appeared to be partly determined by the ambient wind conditions experienced along the flight path (Fig. 2) and partly by the topography of the landscape (Fig. 1C).

¹Department of Migration, Max Planck Institute of Animal Behavior, 78315 Radolfzell, Germany. ²College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia. ³Department of Biology, University of Konstanz, 78464 Konstanz, Germany. ⁴Independent Researcher, Promenadenstrasse 2, 3076 Worb, Switzerland. ⁵Natural Resources Institute, University of Greenwich, Chatham, Kent ME4 4TB, UK. ⁶Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK. ⁷Centre for Ecology and Conservation, University of Exeter, Penryn, Cornwall TR10 9FE, UK. ⁸Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall TR10 9FE, UK. ⁹Department of Entomology, College of Plant Protection, Nanjing Agricultural University, Nanjing 210095, China. ¹⁰Centre for the Advanced Study of Collective Behaviour, University of Konstanz, 78464 Konstanz, Germany.

*Corresponding author. Email: myles.menz@jcu.edu.au (M.H.M.M.); wikelski@ab.mpg.de (M.W.); j.chapman2@exeter.ac.uk (J.W.C.)

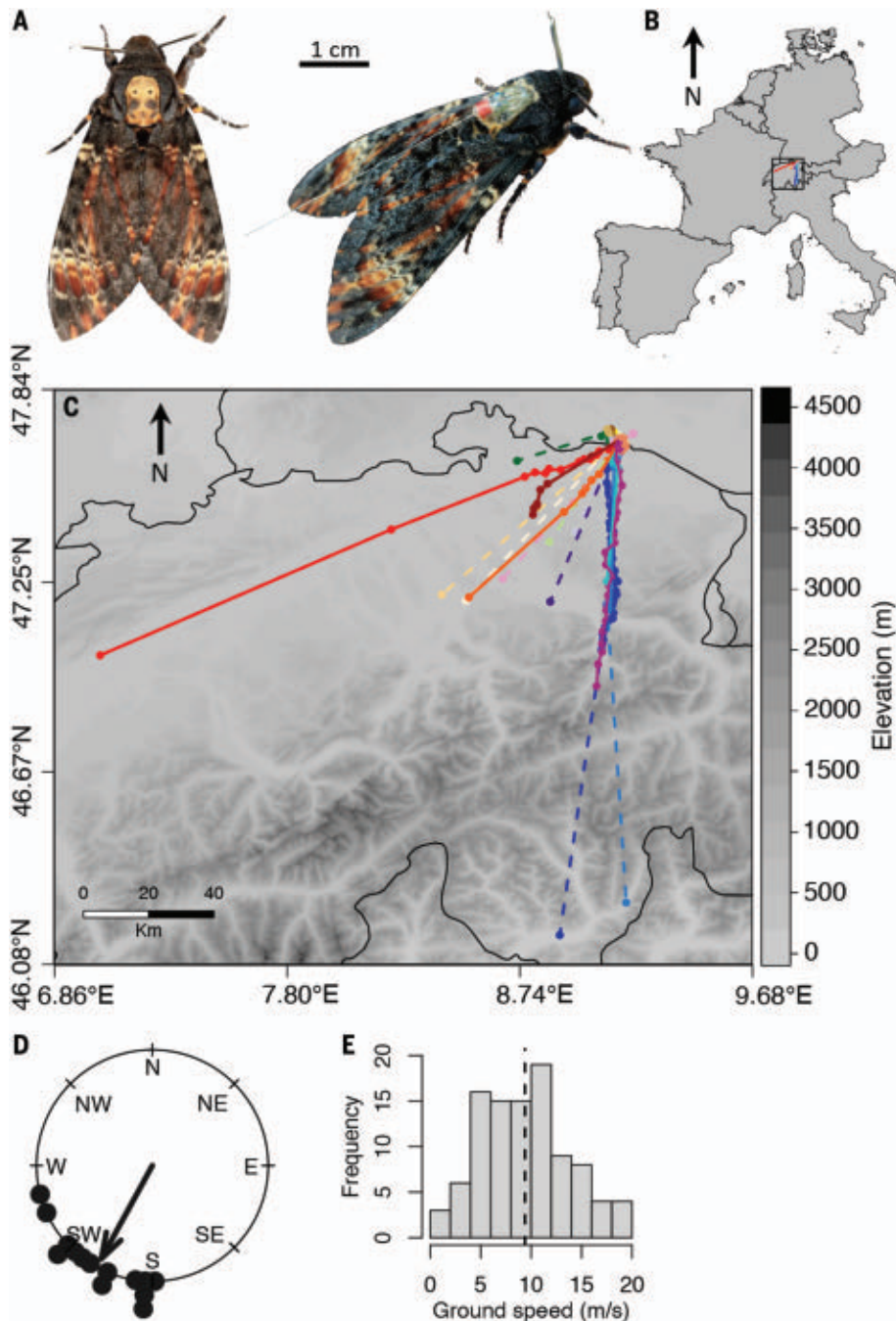


Fig. 1. Individual tracking reveals consistently straight flight paths in migrating hawkmoths. (A) Death's-head hawkmoth (*A. atropos*) showing the characteristic skull-like marking (left) and fitted with a miniaturized VHF radio transmitter (weighing 240 mg) (right). (B) Map of Europe with the tracking region outlined [magnified view shown in (C)]; the release site was ~50 km north of the Swiss Alps. (C) Nighttime tracks of migrating hawkmoths showing persistently straight trajectories throughout a night's flight. Solid lines indicate moths that were tracked continuously throughout a night, and dashed lines indicate presumed tracks of moths to their relocation position during searches in the following day(s). Colors represent different individuals and are consistent between figures. (D) Mean track directions of the 14 moths that demonstrated migratory behavior. Each point represents the track direction of an individual moth. The arrow indicates the overall mean direction (208.70°), and arrow length indicates the directedness ($r = 0.917$). (E) Frequency distribution of track segments in each of 10 ground-speed categories (each bar represents a range of 2 m/s) for the seven moths that were tracked continuously on migration ($n = 99$ segments). The dashed line indicates the mean of 9.4 m/s. [Photo credit: Christian Ziegler]

The first strategy was used under opposing wind directions and resulted in moths taking the most direct route to the wintering grounds by maintaining a constant southward track (Fig. 2, A and B). Under this strategy, the moths continuously adjusted their headings so that distributions of β had 95% confidence intervals (CIs) that overlapped 180° and had a mean β close to that value (Fig. 3A and table S2), resulting in more-or-less upwind flight (Fig. 4, A to D). Examination of ground speeds and wind speeds along the track (figs. S1 and S2) (21) indicated that moths that used this orientation strategy must have flown close to the ground (50 m or lower), that is, within their “flight boundary layer” [the lower-most layer of the atmosphere within which the insect's self-powered flight speed exceeds the wind speed, allowing control of their trajectory (4, 22)].

The second and third strategies were both used under favorable wind directions (i.e., occasions when southward flight would expose moths to some degree of tailwind assistance). We predicted that moths using tailwind assistance would fly in the layer where winds were fastest, as previously observed in studies of noctuid moths (10, 19). However, examination of ground speeds and airspeeds on these occasions indicated that hawkmoths that used these strategies flew about 300 m above the ground, considerably lower than the wind-speed maxima available (figs. S1 and S2) but high enough to receive some wind assistance (Fig. 4, E to G). Under these conditions, moths appear to balance speed with direction, as seen in other migrant moths (10). The second orientation strategy, which was used by a single individual (moth 5), involved flying relatively close to the south-westward downwind direction (Fig. 2, C and D), but individuals modified their heading to achieve a straight trajectory lying somewhat further south of the strongest wind (as supported by values of β around -30° to -50° and for which the 95% CIs do not overlap with 0° ; Fig. 3B and table S2). The final orientation strategy, which was used by a single individual (moth 5), involved flying directly downwind (as indicated by the 95% CI of β overlapping 0° ; Fig. 3C and table S2), resulting in a track toward the west-southwest (Fig. 2D) with a higher ground speed than any other moth (Fig. 4H).

In general, there was a negative relationship between airspeed and wind assistance, with airspeed increasing in headwinds and decreasing in tailwinds (Fig. 4). Furthermore, median ground speed was relatively similar across the orientation strategies (Fig. 4H). Thus, moths modulated their ground speed by varying their self-powered flight vector under different wind conditions to achieve a preferred ground speed, similar to that documented in many insects (23), which may be beneficial in the trade-off between energy consumption and travel speed (22).

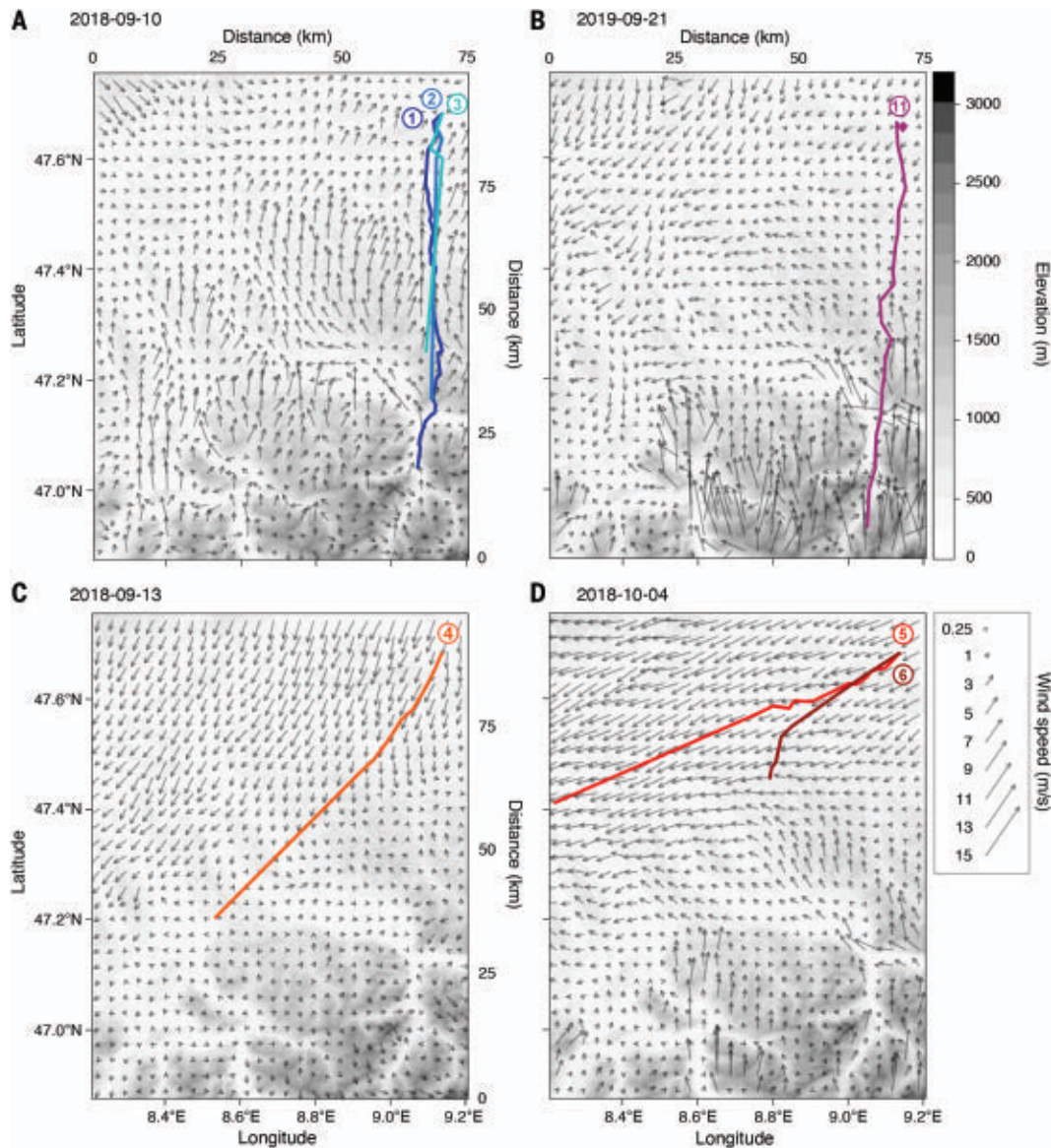


Fig. 2. Migrating hawkmoths continuously compensate for wind to maintain straight flight paths. (A to D) Tracks of migrating *A. atropos* in relation to wind direction and speed (length of the arrows). The moths exhibit different strategies under different wind conditions, traveling due south through the Alps when primarily encountering headwinds [(A) and (B)] but traveling toward the

southwest and thus skirting the Alps under tailwind conditions [(C) and (D)]. Colors represent different individuals and are consistent between figures. Wind layers are derived from the COSMO-1 model and represent conditions at 50 m above ground level [(A) and (B)] and 300 m above ground level [(C) and (D)], the estimated altitudes at which the moths were flying in the corresponding cases.

The maintenance of consistently straight tracks and regulation of ground speed throughout the night under variable wind conditions strongly suggests that *A. atropos* has an internal compass mechanism. Flight simulator studies have demonstrated that migrating Bogong moths (*Agrotis infusa*) use a combination of visual landmarks and Earth's magnetic field to navigate toward a goal (13). This has yet to be demonstrated in free-flying migratory insects, but we predict that migrating hawkmoths, which have excellent nocturnal vision (24), use a similar suite of sensory modalities to navigate over very large spatial scales during migration (although nothing is

yet known of the capability of hawkmoths to detect magnetic fields). At the landscape scale, we propose that the moths used topographical cues to visually navigate, because magnetic cues are unlikely to be accurate enough to maintain such straight trajectories. Overlaying the straight tracks on a topographical map (Fig. 1C) shows that the three orientation strategies, and their directional clusters, are each clearly aligned with a topographical feature that would also result in avoiding the highest elevations of the Alps (high-altitude passes running due south and southwest through the Alps and a wide valley running west-southwest that would enable circumventing the Alps altogether).

Here, we provide evidence that large night-flying insects actively select an orientation strategy in response to environmental conditions, at least for some part of their migratory journey. To maintain such straight trajectories over long periods of time, as seen here, the moths must regularly update their position relative to whichever navigational cues they rely on. However, complete compensation has not been previously documented in a long-range migratory insect and is generally an unusual and very rare strategy in long-range migrants (25). Our results show that complex migratory strategies are not limited to vertebrates.

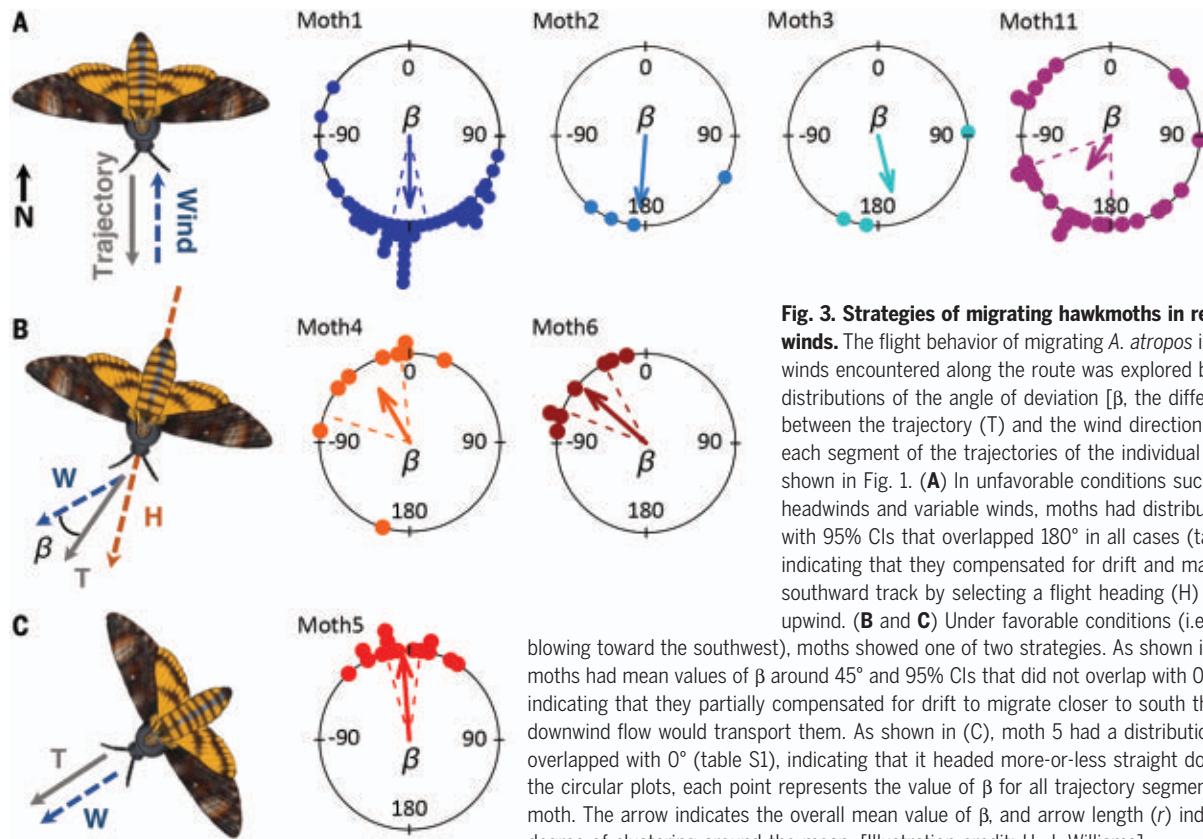


Fig. 3. Strategies of migrating hawkmoths in relation to winds. The flight behavior of migrating *A. atropos* in relation to winds encountered along the route was explored by analyzing distributions of the angle of deviation [β , the difference between the trajectory (T) and the wind direction (W)] for each segment of the trajectories of the individual moths shown in Fig. 1. (A) In unfavorable conditions such as headwinds and variable winds, moths had distributions of β with 95% CIs that overlapped 180° in all cases (table S1), indicating that they compensated for drift and maintained a southward track by selecting a flight heading (H) directly upwind. (B and C) Under favorable conditions (i.e., winds blowing toward the southwest), moths showed one of two strategies. As shown in (B), some moths had mean values of β around 45° and 95% CIs that did not overlap with 0° (table S1), indicating that they partially compensated for drift to migrate closer to south than the downwind flow would transport them. As shown in (C), moth 5 had a distribution of β that overlapped with 0° (table S1), indicating that it headed more-or-less straight downwind. On the circular plots, each point represents the value of β for all trajectory segments of each moth. The arrow indicates the overall mean value of β , and arrow length (r) indicates the degree of clustering around the mean. [Illustration credit: H. J. Williams]

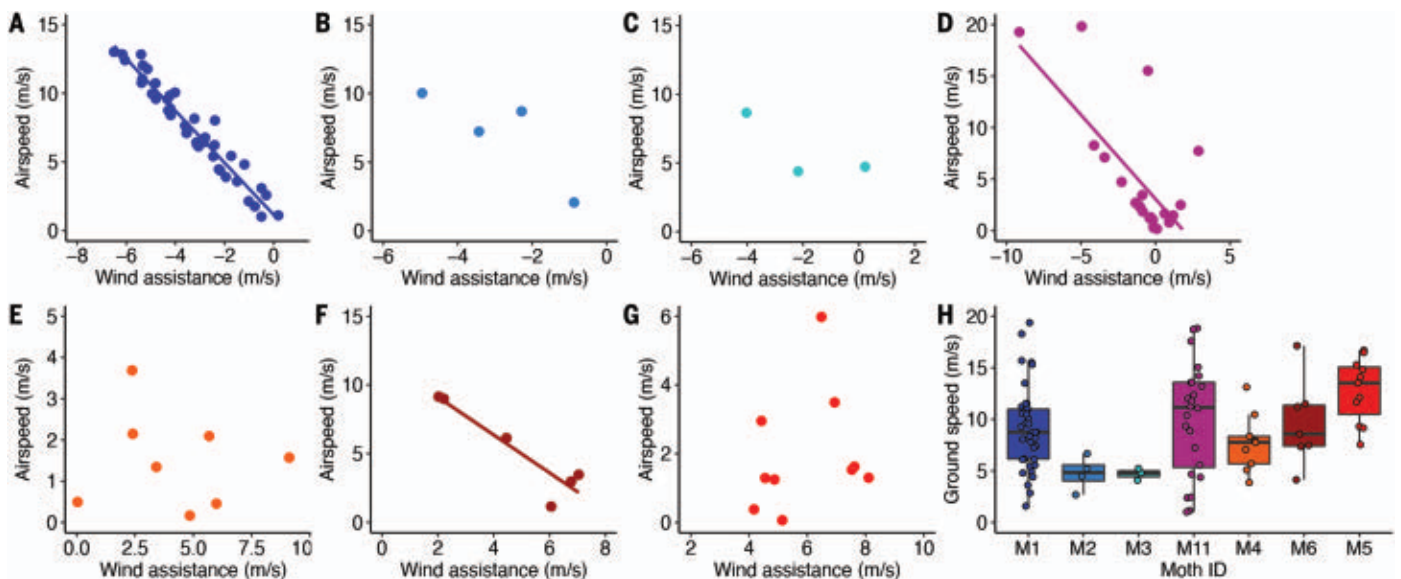


Fig. 4. Migrating hawkmoths modulate their airspeed in relation to wind assistance. (A to G) Airspeed (m/s) of seven migrating *A. atropos* individuals [(A) to (G)] in relation to wind assistance (m/s) en route. Wind assistance was calculated as the wind vector at the location of the moth in the direction of travel toward its next location, with positive values indicating tailwind and negative values indicating headwind. Moths generally increased their airspeed under headwind conditions [(A) to (D)] and reduced their airspeed in more favorable tailwind conditions [(E) to (G)]. (H) Ground speed per segment for each of the seven individuals. Boxes

represent the 25th and 75th percentiles [interquartile range (IQR)], and whiskers indicate data within ± 1.5 times the IQR. The central bars represent median values. Colors represent different individuals and are consistent between figures. Regression lines from linear models (LMs) are presented for significant relationships. LMs were performed for individuals with more than five data points. Significance ($P < 0.05$) was based on likelihood-ratio tests: (A) moth 1, $F = 651.91$, $P < 0.001$; (D) moth 11, $F = 20.66$, $P < 0.001$; (E) moth 4, $F = 0.065$, $P = 0.807$; (F) moth 6, $F = 24.69$, $P = 0.008$; and (G) moth 5, $F = 0.465$, $P = 0.514$.

REFERENCES AND NOTES

- R. A. Holland, M. Wikelski, D. S. Wilcove, *Science* **313**, 794–796 (2006).
- G. Hu *et al.*, *Science* **354**, 1584–1587 (2016).
- K. R. Wotton *et al.*, *Curr. Biol.* **29**, 2167–2173.e5 (2019).
- J. W. Chapman, D. R. Reynolds, K. Wilson, *Ecol. Lett.* **18**, 287–302 (2015).
- D. A. Satterfield, T. S. Sillett, J. W. Chapman, S. Altizer, P. P. Marra, *Front. Ecol. Environ.* **18**, 335–344 (2020).
- C. B. Williams, *Insect Migration* (Collins, 1958).
- C. B. Williams, *Proc. R. Entomol. Soc. London* **13**, 70–84 (1949).
- C. G. Johnson, *Migration and Dispersal of Insects by Flight* (Methuen, 1969).
- L. R. Taylor, R. A. French, E. D. M. Macaulay, *J. Anim. Ecol.* **42**, 751–760 (1973).
- J. W. Chapman *et al.*, *Science* **327**, 682–685 (2010).
- J. W. Chapman *et al.*, *Curr. Biol.* **18**, R908–R909 (2008).
- J. W. Chapman *et al.*, *J. Anim. Ecol.* **85**, 115–124 (2016).
- D. Dreyer *et al.*, *Curr. Biol.* **28**, 2160–2166.e5 (2018).
- D. Dreyer *et al.*, *J. Exp. Biol.* **221**, jeb179218 (2018).
- W. D. Kissling, D. E. Pattenmore, M. Hagen, *Biol. Rev. Camb. Philos. Soc.* **89**, 511–530 (2014).
- A. R. Pittaway, *The Hawkmoths of the Western Palaearctic* (Harley Books, 1993).
- P. Howse, *Bee Tiger: The Death's Head Hawk-Moth Through the Looking-Glass* (Brambleby Books, 2021).
- G. F. McCracken *et al.*, *R. Soc. Open Sci.* **3**, 160398 (2016).
- J. W. Chapman *et al.*, *Curr. Biol.* **18**, 514–518 (2008).
- B. Gao *et al.*, *Proc. Biol. Sci.* **287**, 20200406 (2020).
- Materials and methods are available as supplementary materials.
- R. B. Srygley, R. Dudley, *Integr. Comp. Biol.* **48**, 119–133 (2008).
- K. J. Leitch, F. V. Ponce, W. B. Dickson, F. van Breugel, M. H. Dickinson, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2013342118 (2021).
- A. L. Stöckl, D. C. O'Carroll, E. J. Warrant, *Curr. Biol.* **26**, 821–826 (2016).
- J. W. Chapman *et al.*, *Curr. Biol.* **21**, R861–R870 (2011).
- M. H. M. Menz *et al.*, Data from: Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth. Movebank Data Repository (2022); <https://doi.org/10.5441/001/1.f4r24r5r>.

ACKNOWLEDGMENTS

Special thanks to C. Ziegler for assistance in the field and providing images; the Swiss Ornithological Institute for providing access to the field site at Col de Bretolet; M. Thoma for discussion regarding the migration of *A. atropos*; D. Dreyer and R. Massy for discussion of circular statistics; I. Couzin for useful discussion of the project; and D. Dechmann, K. Safi, and two anonymous reviewers for constructive comments on the manuscript. **Funding:** This project has received funding from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement no. 795568 to M.H.M.M. and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2117 – 422037984 to M.W. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the UK. **Author contributions:** Conceptualization: M.H.M.M., M.W.; Data collection: M.H.M.M., H.-M.B.-S., M.W.; Data analysis: M.H.M.M., M.S.; Writing – original draft: M.H.M.M., J.W.C., M.W., H.J.W.; Writing – review and editing: M.H.M.M., H.J.W., H.-M.B.-S., D.R.R., M.S., J.W.C., M.W. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The datasets generated and analyzed during the current study are available in the Movebank Data Repository (26). Custom code for extraction of wind variables is available in data S1. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1663

Materials and Methods

Figs. S1 and S2

Tables S1 and S2

References (27–41)

Data S1

MDAR Reproducibility Checklist

Submitted 22 November 2021; accepted 14 July 2022

10.1126/science.abn1663

FLUID DYNAMICS

Dynamics of active liquid interfaces

Raymond Adkins^{1†}, Itamar Kolvin^{1*†}, Zhihong You^{1†}, Sven Witthaus¹, M. Cristina Marchetti^{1,2*}, Zvonimir Dogic^{1,2*}

Controlling interfaces of phase-separating fluid mixtures is key to the creation of diverse functional soft materials. Traditionally, this is accomplished with surface-modifying chemical agents. Using experiment and theory, we studied how mechanical activity shapes soft interfaces that separate an active and a passive fluid. Chaotic flows in the active fluid give rise to giant interfacial fluctuations and noninertial propagating active waves. At high activities, stresses disrupt interface continuity and drive droplet generation, producing an emulsion-like active state composed of finite-sized droplets. When in contact with a solid boundary, active interfaces exhibit nonequilibrium wetting transitions, in which the fluid climbs the wall against gravity. These results demonstrate the promise of mechanically driven interfaces for creating a new class of soft active matter.

Liquid-liquid phase separation (LLPS) is a ubiquitous phase transition, with examples abounding throughout material science, biology, and everyday life (1, 2). Immiscible liquid phases are separated by sharp but deformable interfaces that strongly couple to flows and the input of mechanical energy. For example, gentle shaking of an oil-water mixture induces gravity-capillary interfacial waves, whereas more vigorous perturbations break up the entire interface, reinitializing the phase separation (3–6). Active matter provides an alternative method of continuously stirring a fluid (7, 8). In such systems, mechanical energy, inputted locally through the motion of microscopic constituents, cascades upward to generate large-scale turbulent-like dynamics (9–11). We studied how active stresses and associated flows perturb soft interfaces and LLPS. Using experiment and theory, we identified universal features of active-LLPS, including giant interfacial fluctuations, traveling interfacial waves, activity-arrested phase separation, and activity-induced wetting transitions. These results demonstrate how active matter drives liquid interfaces to configurations that are not accessible in equilibrium. In turn, active interfaces are elastic probes that provide insight into the forces that drive active fluids—for example, by allowing for the measurement of the active stresses.

The active liquid interfaces we studied belong to a wider class of activity-driven boundaries that includes lipid bilayers, colloidal chiral fluids, and interfaces between motile and immotile bacteria in a swarm (12–16). From a biology perspective, LLPS has emerged as a ubiquitous organizational principle (2, 17).

How cytoskeletal active stresses couple to self-organization of membraneless organelles remains an open question. Studies of simplified systems can shed light on these phenomena. Relatedly, active wetting plays a potential role in the development and shaping of tissues (18).

To explore how activity modifies soft interfaces, we combined poly(ethylene glycol) (PEG) and polysaccharide dextran with stabilized microtubule filaments and clusters of kinesin molecular motors. Above a critical polymer concentration, the passive PEG-dextran mixture phase separated (19). Microtubules and kinesin clusters exclusively partitioned into the dextran phase, in which depletion forces promoted microtubule bundling (Fig. 1, A to C). Streptavidin-bound kinesin clusters (KSA) stepped along adjacent microtubules within a bundle, driving interfilament sliding. The kinesin-powered bundle extensions continuously reconfigured the filamentous network, generating large-scale turbulent-like flows, similar to those previously studied (Fig. 1D) (9). The PEG-dextran interfaces were susceptible to large deformations by active stresses generated within the dextran phase because of their ultralow interfacial tension (<1 μN/m) (19).

We first visualized the phase separation dynamics of active LLPS in ~30-μm-thick horizontal microscopy chambers. In such samples, PEG-dextran interfaces had a nearly flat vertical profile (fig. S1). The quasi-two dimensional (2D) nature of the system was supported by a nearly constant area fraction of the PEG-rich domains (fig. S2). In a passive system with microtubules but no kinesin motors, the droplets coalesced slowly (Fig. 1E and movie S1). The addition of motors altered the coarsening kinetics. At intermediate KSA concentrations, active flows powered droplet motility, which increased the probability of droplets encountering each other and coalescing, thus speeding up coarsening dynamics (Fig. 1E and movie S2). Higher KSA concentrations accelerated buildup of interfacial fluctuations,

¹Department of Physics, University of California at Santa Barbara, Santa Barbara, CA 93106, USA. ²Graduate program in Biomolecular Science and Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106, USA.

*Corresponding author. Email: zdogic@physics.ucsb.edu (Z.D.); cmarchetti@ucsb.edu (M.C.M.); itamar@ucsb.edu (I.K.)

†These authors contributed equally to this work.

leading to an entirely different dynamical state in which droplets incessantly fused and fissioned with each other (Fig. 1, E and G, and movie S3).

To quantify the influence of activity on the PEG-dextran phase separation, we measured the equal-time two-point correlation function $C(\Delta\mathbf{r}, t) = \langle I(\mathbf{r} + \Delta\mathbf{r}, t)I(\mathbf{r}, t) \rangle$, where $I = 1$ in the dextran phase and -1 otherwise (fig. S3). Spatial correlations decayed over a length scale ξ , defined by $C(\xi) = 0.5$, which is comparable with the average droplet size (fig. S4). For passive samples, ξ increased slowly in time (Fig. 1F). Enhanced coarsening at intermediate KSA concentration was reflected by a much faster initial growth of ξ than the passive case. At high motor concentration, ξ peaked at ~ 1 hour and subsequently decayed to a finite plateau, ξ_{steady} . In parallel, average interface curvature κ monotonically grew, reaching a sufficiently large value to cause droplet fission (fig. S5). The steady-state length scale ξ_{steady} was maintained by the balance of droplet fission and fusion events, in which ξ_{steady} was comparable with the inverse of the average interface curvature κ_{steady} . Concomitantly with the plateauing of ξ , active flow speed became constant (Fig. 1F). These results demonstrate activity-suppressed coarsening dynamics, which created an emulsion-like state in which finite-sized droplets continuously merge, break apart, and exchange their content (Fig. 1G and movies S3 and S4). The volume fractions of the active and passive phases were nearly equal (fig. S2). Low volume fraction of active fluid generated similar steady states. Finite-sized domains are reminiscent of theoretical prediction in motility-induced phase separation of isotropic active particles (20). However, in contrast to theory, the active fluid in our experiments is anisotropic and perturbs an underlying equilibrium phase separation.

To gain insight into how active stresses drive interfacial fluctuations, we formed a macroscopic interface through gravity-induced bulk phase separation (Fig. 2A and fig. S6). In equilibrium, molecular motion works against the density difference $\Delta\rho$ and interfacial tension γ to roughen the liquid-liquid interface. Typical disturbances of PEG-dextran interfaces, bereft of activity, are ~ 100 nm in amplitude, resulting in boundaries that appear flat when viewed with our imaging setup (Fig. 2A and fig. S7). When driven out of equilibrium, however, interfaces exhibited giant undulations that were visible with the naked eye (movie S5). As motor concentration increased, interfaces became multivalued with frequent overhangs, indicating that active stresses directly control interface configurations (Fig. 2A and movie S6).

The interplay of activity and capillarity is clarified by measuring the interfacial fluctuation spectrum. To this end, local interface tangent angles $\theta(s, t)$ were sampled at a time

t as a function of the arc-length distance s along the interface (Fig. 2A). Interfacial fluctuations were described by time-averaged power spectra $S(k) = \langle \|\theta_k\|^2 \rangle_t$, with $\theta_k = \int ds \theta(s, t) e^{-iks}$. Because of equipartition of thermal energy among Fourier modes, the spectrum of equilibrium interfaces is $S(k) \sim Tk^2 / (k^2 + k_c^2)$, where T denotes temperature. The capillary wave number $k_c = \sqrt{\Delta\rho g / \gamma}$ sets a crossover from a gravity-dominated regime at large scales $S(k) \sim k^2$ to a plateau at small scales, where surface tension attenuates fluctuations. Active interfacial fluctuations were markedly different. Active spectra $S(k)$ increased for small wave numbers (Fig. 2B and figs. S7 and S8). After reaching a maximum for $k_m \sim 30 \text{ mm}^{-1}$, it decayed as $S(k) \sim k^{-3}$, instead of plateauing as in equilibrium. Whereas the shape of $S(k)$ remained the same for all KSA concentrations, the root mean square tangent angle θ_{rms} increased linearly with activity (Fig. 2B, inset). Using the crossover at k_m as a determinant of the fluctuation amplitude, it would take an effective temperature of $\sim 10^{11} \text{ K}$ to achieve equi-

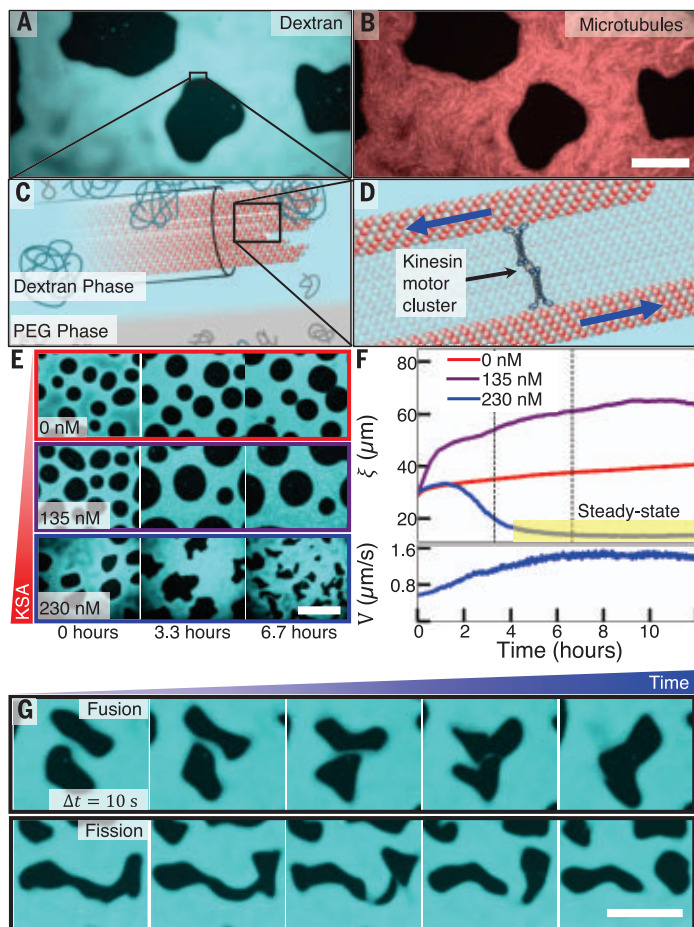
librium interfaces whose roughness is comparable with those measured at the lowest activities.

The dynamics of activity-driven interfacial fluctuations exhibited nontrivial spatiotemporal correlations. To gather sufficient statistics, we imaged ~ 10 -mm-long active interfaces over a 2-hour interval. Space-time maps of local interface height $h(x, t)$ exhibited diagonally streaked crests and troughs that were suggestive of propagating waves (Fig. 2C). These translational modes were also evident in time-lapse movies (movie S6). To characterize these modes, we measured the dynamic structure factor (DSF) of the interface height $D(k, \omega) = \int dx dt e^{ikx + i\omega t} \langle h(x', t') h(x' + x, t' + t) \rangle_{x', t'}$ (Fig. 2D). Over a finite range of wave numbers, the DSF exhibited peaks at finite frequencies ω_p , confirming the presence of propagating modes (Fig. 2E). Increased KSA concentration resulted in higher ω_p for the same wave numbers; thus, activity controlled the phase velocity (Fig. 2F).

The giant nonequilibrium fluctuations and propagating wave modes result from the interaction of active flows in the bulk dextran

Fig. 1. Active LLPS.

(A) Coexisting PEG-rich (dark) and dextran-rich (cyan) domains. (B) Labeled microtubules (red) are dissolved in the dextran phase. Scale bar, 75 μm . (C) Microscopic-scale depiction of phase separation. Minority PEG polymers (gray) in the dextran-rich phase induce microtubule bundling. (D) Kinesin clusters drive interfilament sliding. (E) Time evolution of the active LLPS at three KSA concentrations. Scale bar, 350 μm . (F) (Top) Correlation length evolution $\xi(t)$ for three KSA concentrations. For 230 nM KSA, ξ plateaus at long time (yellow highlight). (Bottom) Root mean square velocity of turbulent flows in the dextran phase at 230 nM KSA. (G) Fusion and fission of PEG droplets. Sample chamber thickness, 30 μm . Scale bar, 100 μm .



phase with interfacial elasticity. To elucidate the processes that drive active interfaces, we numerically integrated 2D hydrodynamic equations that describe a bulk phase-separated active fluid (21, 22). The two coexisting phases were modeled as incompressible Newtonian fluids under gravity that experience confinement friction in the low Reynolds number limit (23). The top phase was passive, whereas the velocity of the bottom phase \mathbf{v} was governed by

$$\gamma_v \mathbf{v} - \eta \nabla^2 \mathbf{v} = -\nabla P + \nabla \cdot \boldsymbol{\sigma} \quad (1)$$

where P is the pressure, η is the viscosity, and γ_v is the confinement friction. The stresses $\boldsymbol{\sigma}$ driving the flows were assumed to be generated by an active liquid crystal producing extensile active stresses, $\boldsymbol{\sigma}^a = \alpha \mathbf{Q}$, with $\alpha < 0$. The local orientational order was quantified by a traceless tensor $Q_{ij} = \langle \hat{n}_i \hat{n}_j - \delta_{ij}/2 \rangle$ averaged over molecular orientations $\hat{\mathbf{n}}$. Active shear flows engendered orientational order through flow-induced alignment. These assumptions are summarized in the continuum equation

$$\frac{D\mathbf{Q}}{Dt} + [\boldsymbol{\omega}, \mathbf{Q}] = \lambda \mathbf{u} + \frac{1}{\lambda_Q} \mathbf{H} \quad (2)$$

where $\boldsymbol{\omega}$ is the vorticity tensor, \mathbf{u} is the strain rate, and λ is the flow alignment parameter. \mathbf{H} denotes elastic forces that arise from the liquid crystal free energy, and γ_Q is the rotational viscosity of microtubule bundles (23). In the absence of activity, the liquid crystal is in the isotropic phase, which is consistent with the microtubule density used in experiments.

The hydrodynamic model reproduces key experimental observations. Finite-sized chaotic flows, driven by active stresses, continuously deform the liquid-liquid interface (Fig. 3A and movie S7). Similarly to experiments, the interfacial power spectra showed a crossover between growth at small wave numbers and decay at large wave numbers, whereas the root mean square tangent angle increased linearly with $|\alpha|$ (Fig. 3B). The numerically obtained DSF also exhibited signatures of active traveling waves (Fig. 3, C to E), as in the experiment. The wave frequencies $\omega_p(k)$ increased with activity, showing active-stress-dependent wave velocity. Our numerical model also suggests a noninertial mechanism of active waves (12), which differs from conventional inertia-dominated capillary waves (24, 25). In the context of our hydrodynamic model, the active waves arise from the coupling between the interface vertical displacement h and orientational order Q in the interfacial region (22). Stress balance at the interface predicts that the orientational order drives interfacial deformation as $\partial_t h \sim v_{\perp}^a - \alpha |k| Q$, where v_{\perp}^a is the active contribution to the flow velocity normal to the interface. In turn, passive flows v_{\perp}^p relaxing the

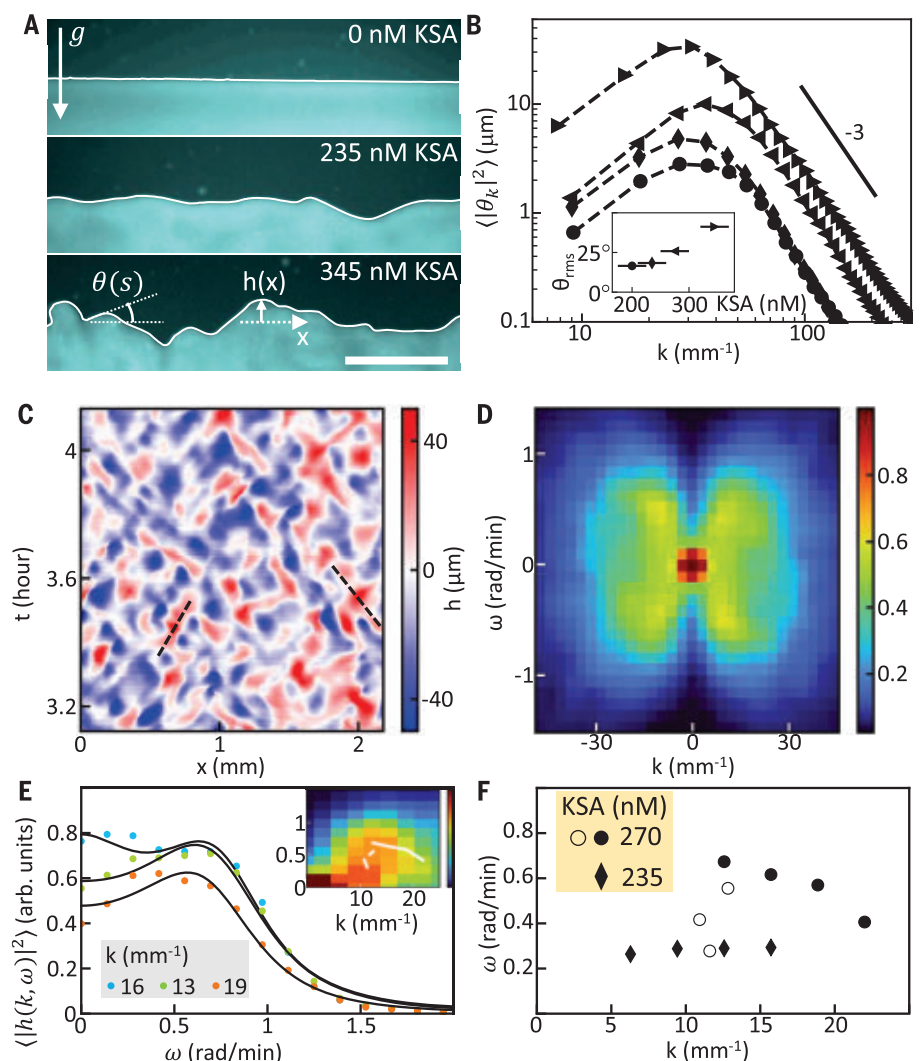


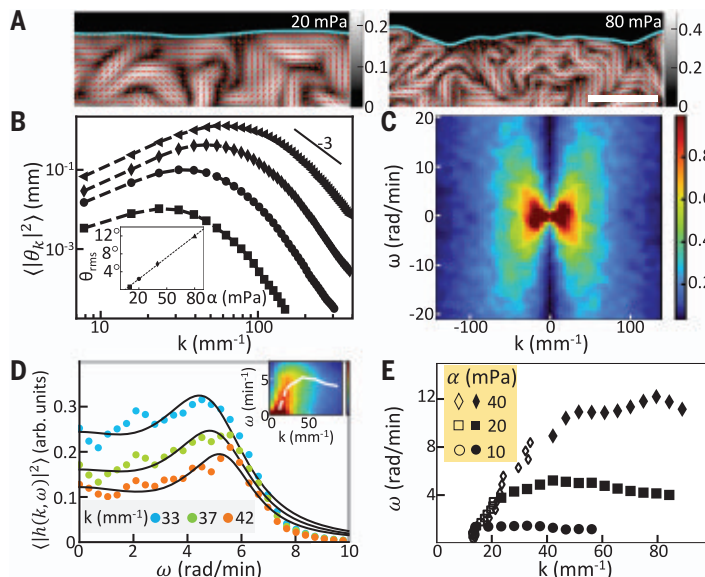
Fig. 2. Active propagating waves. (A) Conformations of bulk phase-separated interfaces with increasing motor (KSA) concentration. $\theta(s)$ defines the local interface tangent angle as a function of arc length. $h(x)$ is the local interface height. Chamber thickness, 60 μm . Scale bar, 150 μm . (B) θ power spectra obtained by averaging over ~ 3 -mm interface length, 6 to 8 hours after sample preparation. (Inset) Root mean square θ versus KSA concentration. (C) Space-time map of h . Disturbances propagating along the interface (dashed lines). Interfaces were corrected for drift and tilt. (D) The square-root DSF averaged 4 to 6 hours after sample preparation. Maximum lag distance, 1000 μm ; maximum lag time 1800 sec. (E) DSF sections at constant wave number (colored dots). Black lines are best approximations with $F(\omega) = a [(\omega/\omega_0)^2 + 1]^{-1} + b \left\{ \left[(\omega/\omega_0)^2 - 1 \right]^2 + (\omega \Delta \omega / \omega_0^2)^2 \right\}^{-1}$, where a , b , ω_1 , ω_0 , and $\Delta \omega$ are adjustable parameters. Data were taken over 2 to 4 hours after sample preparation. (Inset) Frequency peaks $\omega_p = \sqrt{\omega_0^2 \Delta \omega^2 / 2}$ overlaid on DSF intensity (full line). DSF maxima for constant ω (dashed line). For (C) to (E), KSA concentration, 270 nM. (F) Peak frequencies of the propagating modes ω_p (solid symbols). DSF maxima at constant frequency sections (open symbols).

interface at a wave number dependent rate $v(k)$ feed back to induce local liquid crystalline order $\partial_t Q \sim \lambda i k v_{\perp}^p - \lambda i k v(k) h$. Consequently, interface height obeys a wave equation $\partial_t^2 h \sim -\alpha \lambda i k |k| v(k) h$ (23). Accordingly, traveling wave velocities increase with active stress, which is in agreement with both experiments and numerics.

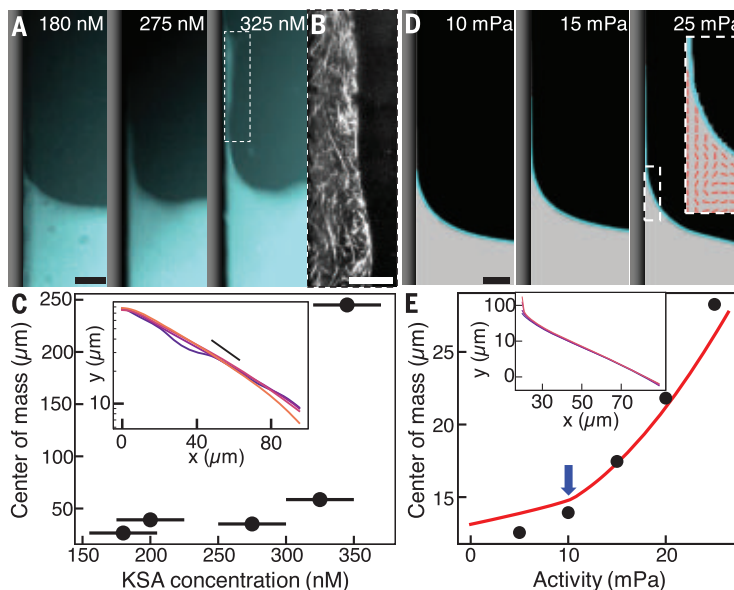
Propagating waves might be a generic feature of active boundaries (12, 13). More broadly, the active-stress-dependent wave dispersion mirrors those of elastic waves in entangled polymer solutions (26). Although the numerical hydrodynamics reproduced qualitative features of the experimentally observed active fluctuations and waves, there were important

Fig. 3. Numerical hydrodynamics of active interfaces.

(A) Phase boundary (cyan) separating a passive fluid (black) and an active isotropic liquid crystal. The latter are depicted with local order parameter (grayscale) and orientation (red lines). Legends denote activities $|\alpha|$. Scale bar, 100 μm . **(B)** θ power spectra. Simulation interval, 2 mm. Correlation maximum lag distance, 400 μm . (Inset) Root mean square θ versus $|\alpha|$. **(C)** Square root DSF. $|\alpha| = 40$ mPa. Maximum lag time, 270 s. Maximum lag distance, 670 μm . Total simulation time, 3 hours. Simulation interval, 2 mm. **(D)** Constant wave number sections of DSF intensity (solid circles). Best approximations $F(\omega)$ are as in Fig. 2E (black lines). $|\alpha| = 20$ mPa. (Inset) $\omega_p(k)$ overlay on DSF intensity (solid line) and DSF maxima for constant ω (dashed line). **(E)** Peak frequencies of the propagating modes ω_p (solid symbols) and DSF maxima for constant ω (open symbols). Activities are noted in absolute values.

**Fig. 4. Active wetting transition.**

(A) Active fluid wetting a polyacrylamide-coated vertical boundary at three KSA concentrations. Scale bar, 50 μm . **(B)** Magnified section of the wetting layer at 345 nM KSA showing local orientations of microtubule bundles. Scale bar, 20 μm . **(C)** Average center-of-mass height of active fluid within $5\ell_e$ of the vertical boundary, where $\ell_e = 45$ μm . Zero height is defined as the average bulk interface position. Each point is the mean of two experiments. Uncertainty in sample preparation is indicated with horizontal bars. (Inset) Average wetting height profiles as a function of distance from the vertical boundary. KSA concentrations, 180 nM (blue) to 270 nM (red). Black line indicates $\sim e^{-x/\ell_e}$. **(D)** Wetting profiles from numerical simulations. Equilibrium contact angle is $\theta_e = 10^\circ$. Interface position is in cyan. Scale bar, 50 μm . (Inset) Local liquid crystalline orientation (red lines). **(E)** Center-of-mass height of the numerical wetting profile. Red line indicates prediction of Eq. 3 with $\ell_w = 2.5$ μm . Onset of complete wetting is indicated with the blue arrow. (Inset) Average numerical wetting profiles. Activities are 10 mPa (blue) to 40 mPa (red).



quantitative differences. In particular, with increasing activity, numerics predicted an increase in both interface fluctuations and bulk velocity. By contrast, in experiments active flows remained constant between 200 and 350 nM KSA, whereas interfacial fluctuations increased (fig. S9). Moreover, in numerics, the maximum wave number k_m increased with activity while remaining constant in experiments (Figs. 2B and 3B and fig. S10).

To demonstrate the particular properties of active interfaces, we studied their structure next to a solid boundary (Fig. 4, A and B). In the absence of motors, the interface assumed an exponential profile $h \propto e^{-x/\ell_e}$ with a decay length of $\ell_e \sim 45$ μm , which we identified with the capillary length $\ell_c = \sqrt{\gamma/\Delta\rho g}$. At the wall,

the rise in the interface height was ~ 70 μm , which is close to the maximum capillary rise of $\sqrt{2}\ell_c$, indicating complete wetting (23). At intermediate KSA concentrations, the capillary rise exhibited active fluctuations around the equilibrium exponential height profile, and the time-averaged center-of-mass height of the wetting region increased slowly (Fig. 4C). Above a critical value of 300 nM KSA, activity generated a new interfacial structure. Specifically, we observed formation of a ~ 20 - μm -thick dynamical wetting layer, which climbed several hundred micrometers above the equilibrium capillary rise (Fig. 4A and movie S8). Within this layer, microtubule bundles preferentially aligned with the wall (Fig. 4B). Coinciding with the appearance of

the microtubule-rich wetting layer, the mean capillary rise sharply increased (Fig. 4C). These observations demonstrate an activity-driven wetting transition beyond the complete wetting of a passive fluid.

We performed numerical simulations of the active-interface adjacent to a vertical boundary (23). The liquid crystal director was anchored parallel to the wall, and the surface-liquid energy γ_w corresponded to an equilibrium contact angle 10° (Fig. 4D). Similar to experiments, the average height profile had an exponential decay (Fig. 4E, inset). As the activity α increased from zero, the height of the contact point increased. Furthermore, above $|\alpha| = 10$ mPa, the active fluid generated a thin wall-adjacent layer, indicating a transition from partial to

complete wetting (Fig. 4D and movie S9). The capillary rise was supported by a $\sim 3\text{-}\mu\text{m}$ -thick ℓ_w , vertically aligned liquid crystalline domain, with $Q \sim 1$. This domain generated coherent active stress along the wall $\sigma^a = \alpha$, which supported the interface rise. Balancing the active tension $\gamma_a \equiv |\sigma^a| \ell_w$ at the contact point with wall adhesion γ_w , surface tension $\gamma \cos \theta_a$, and gravity F_g resulted in a boundary condition for the climbing interface (fig. S11) (23).

$$|\sigma^a| \ell_w + \gamma_w = \gamma \cos \theta_a + F_g \quad (3)$$

Predictions of the center-of-mass height of the capillary rise, using Eq. 3, show a crossover from slow to fast growth at $\alpha = 10 \text{ mPa}$, which is in agreement with the partial-to-complete wetting transition seen in simulations (Fig. 4E and fig. S12).

Active interfaces provide a distinct experimental probe with which to estimate the magnitude of the active stress, a critical parameter that governs dynamics of active fluids. To avoid resorting to various assumptions on the numerical model, we analytically solved Eq. 1, treating the stress σ as a random field with correlations $\langle \sigma_{ij}(\mathbf{r}, t) \sigma_{ij}(\mathbf{0}, 0) \rangle = \sigma_{\text{rms}}^2 e^{-|\mathbf{r}|/\ell_a - |t|/\tau_a}$, where correlation length ℓ_a and time τ_a are identified with those of the bulk active flow (fig. S13) (23). The analytical model captured the interface fluctuations spectrum $S(k)$ without fitting parameters, revealing that its nonmonotonic shape resulted from active flows with scale-dependent kinetic energy spectrum (Fig. 5A, fig. S10, and eq. S32) (23, 27). By contrast, fluctuations of equilibrium interfaces are driven by thermal energy $\sim k_B T$, where k_B is the Boltzmann constant, that is equally distributed among all scales (23). By integrating $S(k)$ over all wave numbers, the active stress is predicted to increase proportionally to tangent angle fluctuations $\sigma_{\text{rms}} = p \theta_{\text{rms}}$, with $p \approx 9 \text{ mPa/rad}$ (Fig. 5B and eq. S35).

To independently verify these numbers, the force balance Eq. 3, which is associated with active wetting, provides an alternative method of estimating active stress. For intermediate KSA values, before the appearance of the active wetting layer, the active stress estimated from wetting is comparable with those extracted from interface fluctuations (Fig. 5B). Above 300 nM KSA, active stresses are a few times larger than those obtained from interface fluctuations. These large values might be a consequence of flow-enhanced alignment of microtubule bundles within the thin wetting layer. The formation of the active wetting layer at finite activity, however, is outside the scope of the static stress balance embodied in Eq. 3. A more complete description of the wetting transition would include dynamical considerations, such as active wave propagation and gravitational sedimentation, that appear to be essential for the formation and turnover of the wetting layer.

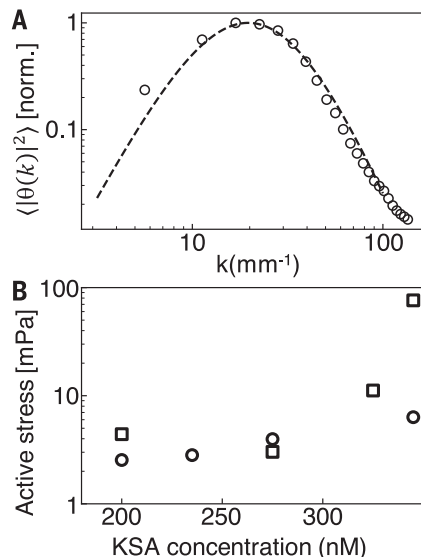


Fig. 5. Active stress estimation from interface fluctuations and wetting. (A) Normalized tangent angle power spectrum. KSA concentration, 200 nM (circles). Analytical hydrodynamic prediction is indicated with the dashed line (eq. S35). (B) Root mean square active stress estimates from interfacial fluctuations $\sigma_{\text{rms}} = p \theta_{\text{rms}}$ (circles). Experimental θ_{rms} are as in Fig. 2B, inset. Analytical hydrodynamic theory (Eq. 1 and eq. S35) predicts $p = 8.8 \text{ mPa/rad}$, by using active length scale $\ell_a = 65 \mu\text{m}$ and time scale $\tau_a = 77 \text{ s}$ that were derived from bulk flow correlations (fig. S13). Active stress estimates from wetting $\sigma^a = 4(h_{\text{cm}} - h_{\text{cm}}^0) \Delta \rho g$ (squares). h_{cm} is the wetting center-of-mass height (Fig. 4C). h_{cm}^0 equals h_{cm} for 180 nM KSA.

We demonstrated a rich interplay between active fluids and soft deformable interfaces. Liquid interfaces provide a quantitative probe that can reveal intrinsic properties of the active fluids, such as active stress. As well, bulk active fluids drive the extreme interfacial deformations that yield intriguing nonequilibrium dynamics, including arrested phase separation, stress-dependent noninertial propagating waves, and activity-controlled wetting transitions. Our findings provide a promising experimental platform by which to design shape-changing adaptable soft materials and machines whose capabilities begin to match those observed in biology (28–30).

REFERENCES AND NOTES

1. A. J. Bray, *Adv. Phys.* **51**, 481–587 (2002).
2. C. P. Brangwynne et al., *Science* **324**, 1729–1732 (2009).
3. J. O. Hinze, *AIChE J.* **1**, 289–295 (1955).
4. P. Tong, W. I. Goldberg, J. Stavans, A. Onuki, *Phys. Rev. Lett.* **62**, 2668–2671 (1989).
5. L. Berthier, J. L. Barrat, J. Kurchan, *Phys. Rev. Lett.* **86**, 2014–2017 (2001).
6. P. Perlekar, R. Benzi, H. J. H. Clercx, D. R. Nelson, F. Toschi, *Phys. Rev. Lett.* **112**, 014502 (2014).

7. R. Aditi Simha, S. Ramaswamy, *Phys. Rev. Lett.* **89**, 058101 (2002).
8. M. C. Marchetti et al., *Rev. Mod. Phys.* **85**, 1143–1189 (2013).
9. T. Sanchez, D. T. N. Chen, S. J. DeCamp, M. Heymann, Z. Dogic, *Nature* **491**, 431–434 (2012).
10. H. H. Wensink, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14308–14313 (2012).
11. S. Zhou, A. Sokolov, O. D. Lavrentovich, I. S. Aranson, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1265–1270 (2014).
12. H. Soni, W. Luo, R. A. Pelcovits, T. R. Powers, *Soft Matter* **15**, 6318–6330 (2019).
13. V. Soni et al., *Nat. Phys.* **15**, 1188–1194 (2019).
14. A. E. Patteson, A. Gopinath, P. E. Arratia, *Nat. Commun.* **9**, 5373 (2018).
15. S. C. Takatori, A. Sahu, *Phys. Rev. Lett.* **124**, 158102 (2020).
16. H. R. Vutukuri et al., *Nature* **586**, 52–56 (2020).
17. C. M. Caragine, S. C. Haley, A. Zidovska, *eLife* **8**, e47533 (2019).
18. C. Pérez-González et al., *Nat. Phys.* **15**, 79–88 (2019).
19. Y. Liu, R. Lipowsky, R. Dimova, *Langmuir* **28**, 3831–3839 (2012).
20. R. Singh, M. E. Cates, *Phys. Rev. Lett.* **123**, 148005 (2019).
21. L. Gioni, A. DeSimone, *Phys. Rev. Lett.* **112**, 147802 (2014).
22. L. Matthew, *Phys. Rev. Lett.* **113**, 1–5 (2014).
23. Materials and methods are available as supplemental materials.
24. D. Langevin, *Light Scattering by Liquid Surfaces and Complementary Techniques* (CRC Press, 1992).
25. D. G. A. L. Aarts, M. Schmidt, H. N. W. Lekkerkerker, *Science* **304**, 847–850 (2004).
26. A. Varshney, V. Steinberg, *Nat. Commun.* **10**, 652 (2019).
27. B. Martínez-Prat et al., *Phys. Rev. X* **11**, 031065 (2021).
28. J.-F. Joanny, S. Ramaswamy, *J. Fluid Mech.* **705**, 46–57 (2012).
29. E. Tjhung, A. Tiribocchi, D. Marenduzzo, M. E. Cates, *Nat. Commun.* **6**, 5420 (2015).
30. Y. -N Young, M. J. Shelley, D. B. Stein, *Math. Biosci. Eng.* **18**, 2849–2881 (2021).
31. R. Adkins, I. Kolvin, Z. You, S. Witthaus, M. C. Marchetti, Z. Dogic, Dynamics of active liquid interfaces. *Dryad* (2022); <https://doi.org/10.25349/D9J56Q>.

ACKNOWLEDGMENTS

We thank F. Caballero for enlightening discussions. **Funding:** Experimental work was supported by the US Department of Energy, Office of Basic Energy Sciences, through award DE-SC0019733 (R.A., I.K., and Z.D.). Theoretical analysis was primarily supported by NSF-DMR-1720256 (iSuperSeed), with additional support from NSF-DMR-2041459 (Z.Y. and M.C.M.). I.K. acknowledges support from the HFSP cross-disciplinary fellowship LT001065/2017-C. R.A. acknowledges support from NSF-GRFP-1650114. We also acknowledge the Brandeis biosynthesis facility, supported by (MRSEC) grant DMR-2011846. **Author contributions:** R.A., S.W., and I.K. conducted experimental research; I.K., S.W., R.A., and Z.Y. analyzed experimental and theoretical data; Z.Y. applied finite volume OpenFOAM solver code to integrate continuum equations; I.K., Z.Y., and M.C.M. conducted theoretical modeling and interpretation of data. I.K., M.C.M., and Z.D. conceived the work. I.K., M.C.M., and Z.D. wrote the manuscript. All authors reviewed the manuscript. **Competing interests:** The authors declare no competing interests. **Data availability:** All data reported in the main text and supplementary information as well as executable code for numerical simulations are available at (31). **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo5423
Materials and Methods
Figs. S1 to S13
References (32–57)
Movies S1 to S9

Submitted 11 February 2022; accepted 27 June 2022
10.1126/science.abo5423

EVOLUTION

The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization

Anthony J. Barley^{1,2,*}, Adrián Nieto-Montes de Oca³, Norma L. Manríquez-Morán⁴, Robert C. Thomson²

Hybridization between diverging lineages is associated with the generation and loss of species diversity, introgression, adaptation, and changes in reproductive mode, but it is unknown when and why it results in these divergent outcomes. We estimate a comprehensive evolutionary network for the largest group of unisexual vertebrates and use it to understand the evolutionary outcomes of hybridization. Our results show that rates of introgression between species decrease with time since divergence and suggest that species must attain a threshold of evolutionary divergence before hybridization results in transitions to unisexuality. Rates of hybridization also predict genome-wide patterns of genetic diversity in whiptail lizards. These results distinguish among models for hybridization that have not previously been tested and suggest that the evolutionary outcomes can be predictable.

Understanding why mechanisms of species diversification change across the tree of life remains one of the most challenging endeavors in biology. Recent studies have demonstrated how the strength and pace of mutation, genetic drift, and selection can generate distinct patterns of diversification (1–3). Gene flow through hybridization is arguably the most poorly understood of these primary mechanisms of evolution, despite an increasing appreciation for its importance during speciation (4). This is driven, in part, by the fact that hybridization induces a network structure in evolutionary history that is vastly less tractable to infer than bifurcating phylogenetic history. Therefore, our understanding of speciation in many groups where hybridization is common remains incomplete.

The diversity of hybridization outcomes in nature is substantial. Hybridization is associated with numerous evolutionary phenomena, including genetic introgression, hybrid speciation, lineage fusion, changes in reproductive mode, and polyploidization (5). These distinct biological outcomes have different effects on diversification. For example, hybrid speciation and lineage fusion generate or eliminate species diversity, respectively. Transitions from sexual to unisexual reproduction are also disproportionately associated with hybridization (6). These phenomena have motivated theories to explain the relationship

between hybridization and the evolution of unisexuality and the effects of hybridization on genetic diversity and microevolution (7, 8). However, our understanding of the manner and extent to which the diverse outcomes of hybridization are predictable remains hampered by a lack of empirical data for testing these theoretical models. This is in part caused by a limited number of biological systems that contain replicated examples of each outcome. In this work, we used the North American whiptail lizard (*Aspidoscelis*) to test whether hybridization produces predictable outcomes.

Whiptails include the highest diversity of unisexual lineages among vertebrates. The clade includes >30 sexual species and ~15 recognized unisexual taxa that reproduce by parthenogenesis. Diploid unisexual lineages in this clade are derived from hybridization events between sexual species that form F1 hybrid offspring that reproduce clonally (9). Triploid unisexual lineages are thought to be derived from backcrossing events between diploid unisexual lineages and sexual species. Even before unisexual reproduction was documented in whiptails, the clade was recognized as being “an ideal object lesson in Nature’s way of species-making” because of the notable patterns of phenotypic variation they exhibit across populations (10). However, these patterns of variation and hybridization also produced myriad complications in defining the boundaries between species that have hindered understanding of the group’s systematics and diversification history despite decades of study (9–11). In this work, we study the tree-like and network-like patterns of evolution in *Aspidoscelis* to understand the mechanisms that underlie their divergent outcomes of hybridization.

Using a reduced-representation genomic dataset (tables S1 and S2 and data S1), we inferred patterns of evolutionary history among all the sexual whiptail species (12). Introgression

tests based on the genome-wide frequencies of site patterns detected widespread evidence of nontree-like patterns of evolution (fig. S1 and data S2 and S3). A substantial proportion (~60 to 70%) of the 33 sexual species appears to have been involved in introgressive hybridization events (figs. S2 and S3). Multiple significant tests were detected for several species, which suggests that they have genomes containing introgressed regions from multiple species (data S2 and S3). Phylogenetic network analyses found that patterns of introgressive hybridization among sexual whiptails could largely be captured by ~5 to 13 reticulation edges (figs. S4 and S5). The results across all hybridization analyses were largely congruent. We used admixture graphs to compare introgression hypotheses in the few cases where they were not (figs. S6 and S7 and tables S3 and S4). We detected no evidence of introgressive hybridization between the most distantly related species, even when they occur in sympatry (figs. S2, S4, and S5). Introgressive hybridization appears to have been common even when phenotypic evolution is rapid across populations, as it is in this clade, where ~70% of species encompass multiple, morphologically distinctive subspecies (table S1). Some whiptail subspecies may reflect instances of lineage fusion, where populations that have not diverged sufficiently to develop reproductive isolation come back into contact (13). Several of the inferred hybridization events include species that are morphologically, ecologically, and evolutionarily divergent (figs. S2 and S8). We lack time estimates for the introgression events but suspect that they span a broad range of historical time frames, given the diversity of the introgression patterns (Table 1).

We used phylogenetic network approaches to infer which of the sexual species are the parental ancestors of the hybrid, diploid unisexual lineages [fig. S9; see (12) for a discussion of terminology]. We used population genetics tools to infer the ancestry of the triploid unisexual lineages (fig. S10 and table S5) (12). We identified 11 unisexual lineages that have distinct combinations of genetic ancestry derived from 10 sexual species (table S6). Six of these unisexual lineages are diploid, and two of these appear to have formed multiple times through hybridization. The other five are triploid, and two of these also appear to have been formed multiple times. Within unisexual whiptails, a primary hybrid speciation process always appears to precede speciation by ploidy elevation through genome addition (Fig. 1).

Our ensemble analysis provides a comprehensive view of the patterns and outcomes of hybridization in whiptails, including both introgressive hybridization and hybrid speciation (Table 1 and Fig. 2) (12). The phenomenon

¹Department of Evolution and Ecology, University of California, Davis, Davis, CA 95616, USA. ²School of Life Sciences, University of Hawai‘i, Honolulu, HI 96822, USA.

³Laboratorio de Herpetología and Museo de Zoología Alfonso L. Herrera, Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad Universitaria, Alcadía Coyoacán, Ciudad de México, México.

⁴Laboratorio de Sistemática Molecular, Centro de Investigaciones Biológicas, Universidad Autónoma del Estado de Hidalgo, Colonia Carboneras, Mineral de la Reforma, Hidalgo, México.

*Corresponding author. Email: ajbarley@asu.edu

†Present address: School of Mathematical and Natural Sciences, Arizona State University–West Campus, Glendale, AZ 85306, USA.

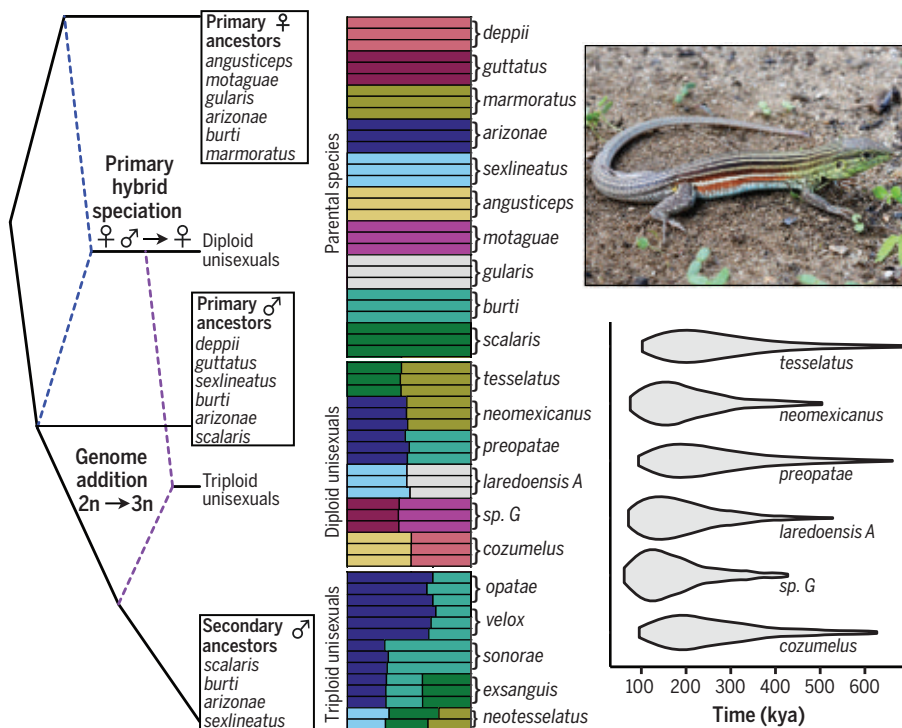


Fig. 1. Summaries of hybrid speciation in unisexual whiptail lizards. Primary hybridization (i.e., a hybridization event between two diploid sexual species that is associated with transition to unisexual reproduction) precedes genome addition (i.e., ploidy elevation) (left). Sexual species associated with each process are indicated. Colored bars in STRUcTURE plot (each representing data for a single lizard) show estimates of the genome-wide proportion of genetic ancestry derived from sexual parental species for each unisexual individual. Violin plots show 95% confidence intervals for formation time estimates of diploid unisexual lineages. kya, thousand years ago. Photo is of *Aspidoscelis deppii*, a sexual species that is both an ancestor of the parthenogenetic lineage *Aspidoscelis cozumelus* and has been involved in introgressive hybridization with *Aspidoscelis guttatus* (photo credit: A.J.B.).

of isolation and lineage divergence followed by secondary contact has occurred repeatedly across broad time scales in *Aspidoscelis*. Consequently, whiptails exhibit many replicated instances of both introgressive hybridization and hybrid speciation, which occurs through transitions in reproductive mode or ploidy elevation. This provides an opportunity to test competing hypotheses explaining the observed association between hybridization and the evolution of unisexuality in vertebrates. Three main hypotheses have been proposed. The phylogenetic constraint hypothesis proposes that particular sexual lineages are inherently predisposed to produce parthenogenetically competent individuals upon hybridization (6). The balance hypothesis predicts that as divergence among hybridizing sexual species increases, the probability that their hybrid offspring will produce unreduced oocytes also increases, allowing for the evolution of unisexual reproduction (14). This is balanced against the increasing likelihood that hybrids will experience impairments in fecundity and/or viability arising from genetic incompatibilities if divergence between their parents is too high. Therefore, this hypothesis predicts that hybridizing lineages that are at intermediate levels of evolutionary divergence produce unisexual offspring, whereas those that are at too low or too high levels lead to sexual offspring or a failed cross, respectively. A further refinement on both of these hypotheses has been called the rare formation hypothesis, in which complex genetic preconditions are necessary to initiate unisexual reproduction in hybrids (15).

These mechanisms are not mutually exclusive, but we see limited evidence for the phylogenetic constraint hypothesis in whiptails. Of the six specific primary hybridization events associated with transitions to parthenogenetic reproduction, only two of the 10 sexual parental species have been involved in more than one of these events (Fig. 2 and table S6). Further, if we model unisexual parental status as a discrete trait on the whiptail phylogeny, we see virtually no evidence of phylogenetic signal (Pagel's $\lambda = 6.6 \times 10^{-5}$; Fritz and Purvis's $D = 0.68$). By contrast, we see conspicuous support for the balance hypothesis. At low levels of divergence, hybridization among sexual whiptail species results only in introgression (Fig. 3A). Only those hybridization events spanning deep levels of evolutionary divergence result in the production of unisexual lineages, and we did not detect evidence of introgression between the parental species of the parthenogenetic lineages. Divergence time estimates between the pairs of sexual parental species that have generated parthenogenetic lineages span a broad range (~15 to 25 million years; fig. S11 and table S7), but these hybridization events have all occurred in the recent past, which suggests that these lineages are

Table 1. Summary of introgressive hybridization events and their support from alternative analyses.

Posterior probability (PP) for one species in a pair as under reticulation edge in Bayesian analysis, with alternative species as an ancestor. ΔL indicates improvement in likelihood score when focal hybrid node is present in the network. Support for focal hybridization event by *f* branch analysis is indicated by the corresponding Z score. The HyDe column indicates the largest Z score calculation for a test that includes both focal species. Support in best admixture graph model is indicated by an X. Dashes indicate analysis is not applicable (PhyloNet) or introgressive hybridization was not supported by analysis (others).

Hybridizing species	Admixture graph	PhyloNet (PP)	PhyloNetworks ΔL	<i>f</i> branch	HyDe
<i>tigris tigris-tigris aethiops</i>	-	1.0	-	10.1	-
<i>deppii deppii-guttatus</i>	X	-	-	12.3	80.2
<i>lineattissimus exoristus-guttatus</i>	X	1.0	3.5	10.1	73.0
<i>lineattissimus exoristus-cf. deppii infernalis</i>	X	1.0	-	10.5	73.0
<i>scalaris scalaris-scalaris colossus</i>	X	1.0	5.0	8.3	37.1
<i>costatus zweifeli-communis*</i>	X	1.0	6.3	11.4	31.5
<i>sackii-costatus costatus*</i>	X	1.0	51.2	22.8	65.4
<i>sackii-parvisocius*</i>	X	1.0	9.4	10.3	14.0
<i>costatus zweifeli-costatus costatus</i>	X	1.0	-	15.4	64.7
<i>mexicanus-motaguae*</i>	X	1.0	51.2	5.8	58.7
<i>calidipes-costatus zweifeli/communis</i>	X	1.0	-	11.1	20.1
<i>scalaris-costatus occidentalis</i>	X	1.0	-	9.6	40.9

*Phylogenetic discordance between mitochondrial DNA and nuclear DNA datasets in Barley *et al.* (13) also supports these introgressive hybridization events.

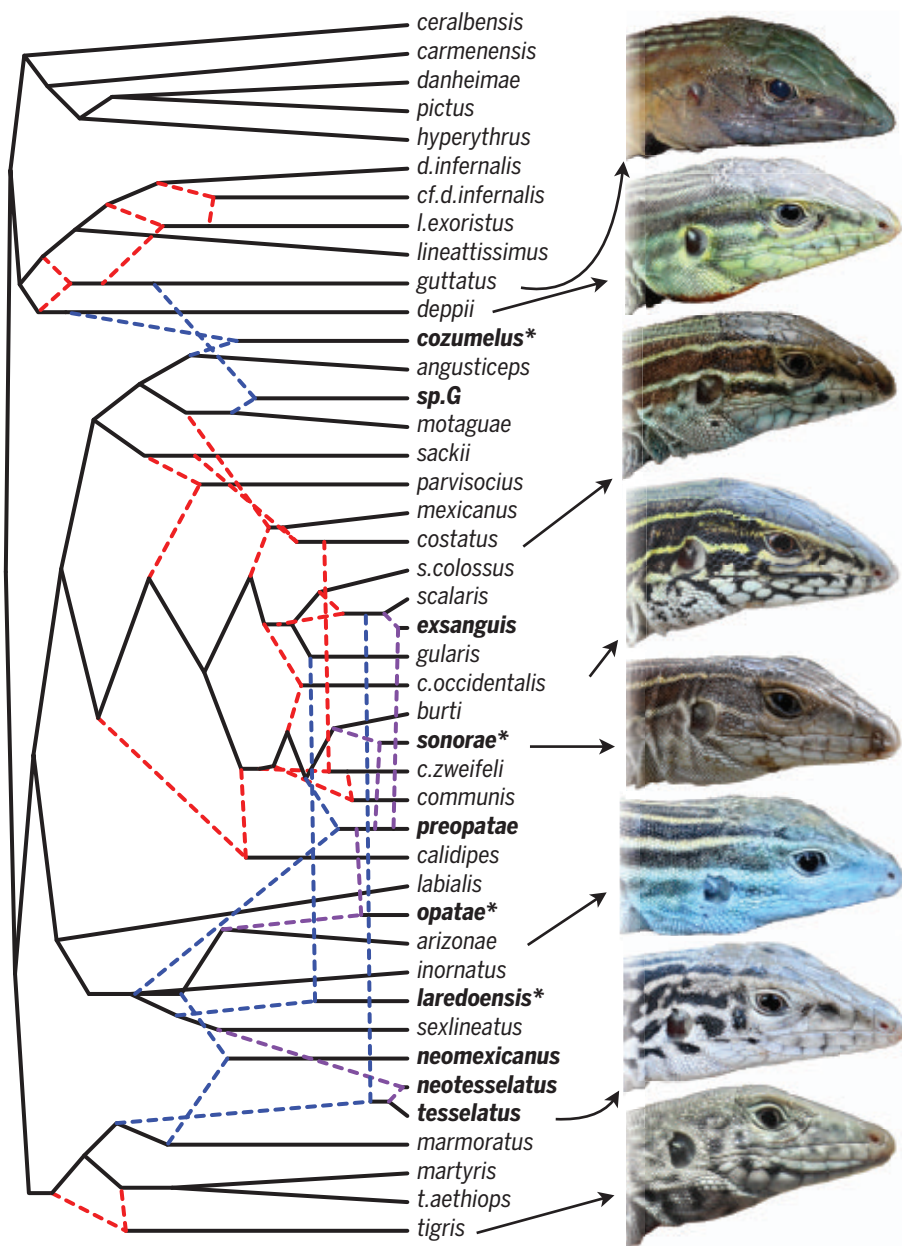


Fig. 2. Evolutionary network history of the whiptails. Red reticulation edges represent introgressive hybridization events between sexual species. Blue reticulation edges denote the evolution of diploid parthenogenetic lineages through the primary hybrid process (Fig. 1). Purple reticulation edges denote the evolution of triploid parthenogenetic lineages through the genome addition process. These edges illustrate instances of hybridization between diploid parthenogenetic lineages and sexual species. Parthenogenetic taxa are indicated in bold. Some of these, indicated by asterisks, have formed multiple times through independent hybridization events between the same sexual species (table S6) (photo credit: A.J.B. and A.N.-M.d.O.).

evolutionarily ephemeral owing to constraints imposed by a lack of genetic recombination (Fig. 1). Finally, we see some evidence in support of the rare formation hypothesis. A limited number of extant unisexual lineages are found in nature compared with the theoretical number of sexual parental combinations whose divergence time falls within this range and whose geographic distributions overlap (6 of 28 combinations). This suggests that specific

genotypic combinations may also be required to induce changes in reproductive mode. This is consistent with laboratory crossing experiments that have had only rare success in producing parthenogenetic lineages (16, 17).

Our results suggest that time since divergence plays a major role in determining the outcomes of hybridization in whiptails, with outcomes switching from introgressive hybridization to hybrid speciation when the diver-

gence time between species exceeds ~10 million years (Fig. 2, fig. S11, and table S7). There is a gap in evolutionary relatedness between pairs of whiptails involved in these two types of hybridization outcomes [between a genetic distance of ~0.0097 and 0.0126 based on the restriction site-associated DNA (RAD)-tag data]. Therefore, it remains an open question whether, within that divergence range, hybridization between pairs of whiptails could conceivably result in either outcome. Predictions of the balance hypothesis are similar to those made by theoretical models of homoploid hybrid speciation that are not accompanied by a transition in reproductive mode (18). This may reflect the importance of particular types of reproductive incompatibilities in both processes that become more likely to arise at intermediate levels of divergence. Incompatibilities in meiosis genes could be necessary to cause transitions from sexual to parthenogenetic reproduction in hybrid individuals (19). In homoploid hybrid speciation, incompatibilities that lead to assortative mating and discrimination against parental populations may be important for initiating reproductive isolation in hybrid populations.

Although patterns regarding the evolution of parthenogenesis and mechanisms of polyploidization appear to be largely conserved in *Aspidoscelis* (Fig. 1), patterns of introgression between species exhibit substantial variation (fig. S2). In their foundational comparative study of speciation, Coyne and Orr (20) have demonstrated that the amount of reproductive isolation between species of *Drosophila* is correlated with their time since divergence. If generalizable, this result predicts that in nature, the amount of introgression between species should be negatively correlated with their time since divergence (21, 22). We expect this relationship because, as species diverge, they should accumulate an increasing number of incompatibilities that both decrease the probability that hybrids can serve as a bridge for gene exchange and increase the probability that introgressed regions will be purged by negative selection over time. Consistent with this prediction, we see a strong negative correlation between the admixture proportion (i.e., the shared proportion of the genome) and the amount of evolutionary divergence that separates two species that have hybridized (Fig. 3B). Within species, allopatric whiptail subspecies show high variability in admixture proportions, with some values similar to those estimated from sister taxa and some much larger (table S8). These results are consistent with a polygenic threshold speciation model, which has garnered support from recent empirical studies (23, 24). Under this model, once two lineages reach a threshold of divergence, Dobzhansky-Muller incompatibilities are typically in sufficient number

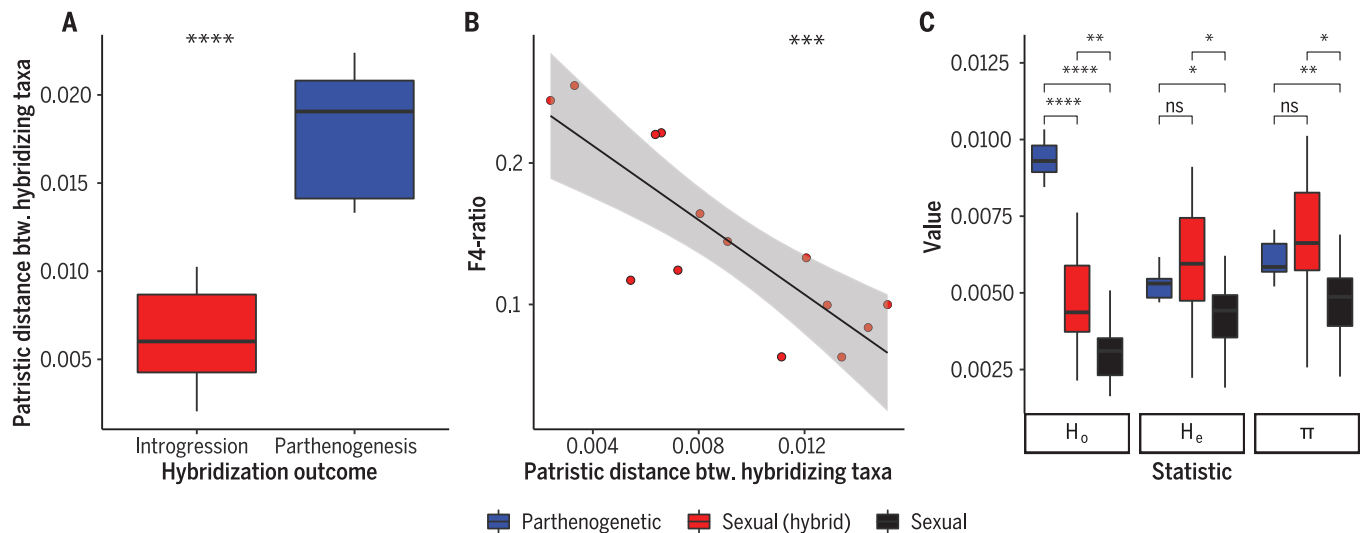


Fig. 3. Illustration of outcomes of hybridization between whiptail lizards.

(A) Hybridization between more closely related sexual species leads to introgression, whereas hybridization between more distantly related sexual species leads to the evolution of parthenogenetic lineages. Box plots are of patristic distances calculated from the posterior distribution of branch lengths separating hybridizing species. (B) For sexual species between which we find strong evidence of introgressive hybridization (Table 1), the amount of the genome that is shared (as estimated by the F4 ratio) is negatively correlated with their time since divergence (here quantified by the patristic distance that separates them). (C) Genome-wide measures of

observed heterozygosity (H_0), nucleotide diversity (H_e), and gene diversity (π) for diploid parthenogenetic lineages and sexual species (see table S9 for data). Because they are clones of F1 hybrids, parthenogenetic lineages have elevated observed heterozygosity compared with that in sexual species. Nucleotide and gene diversity tend to be higher in sexual species that have a hybrid genetic background compared with those that do not, illustrating the role of hybridization as a source of genetic diversity in whiptails. Asterisks indicate significance level for t tests [(A) and (C)] or linear regression with 95% confidence interval (B). **** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; ns, not significant.

to begin suppressing introgression. This effect may determine whether sufficient reproductive isolation has evolved to prevent lineage fusion upon secondary contact. Thus, multiple thresholds may be important to speciation and hybridization outcomes in whiptails at different time scales.

Models for speciation with gene flow make contrasting predictions about the effects of hybridization on diversification, including the formation of genomic islands of speciation within the genome, contributions to adaptation and the evolution of new phenotypes, and increasing rates of diversification (3, 25, 26). To examine the effects that hybridization has had on genomic diversity during speciation in whiptails, we compared estimates of observed heterozygosity, nucleotide diversity, and gene diversity between lineages that have been formed by different speciation processes (Fig. 3C and table S9). Parthenogenetic lineages have maintained high levels of observed heterozygosity because they are clones of F1 hybrid offspring. Sexual species with a history of introgressive hybridization frequently have higher estimates of nucleotide and gene diversity than those that do not. These same values in some sexual species even exceed those calculated for the parthenogenetic lineages, which likely reflects these species having had multiple hybrid partners during their evolutionary history.

Causes and consequences of hybridization in whiptails extend beyond the genome and

mode of reproduction. For example, one of the earliest hypotheses to explain the geographic distribution of parthenogenetic lineages proposed that climatic fluctuations during the Pleistocene may have contributed to shifting distributions of species, providing opportunities for secondary contact and hybridization in doing so (14, 27). This idea is consistent with results of an ancestral state biogeographic reconstruction for whiptails, which suggests that many of the parents of parthenogenetic lineages recently dispersed into the areas that they now occupy and where these lineages formed (fig. S12). Hybridization may also play a role in mediating phenotypic evolution. In *Aspidoscelis*, gigantic body size has evolved at least three different times (fig. S13). These giants are associated with some of the strongest (i.e., large D statistics indicating hybridization with multiple species) and most notable (i.e., with evolutionarily divergent species) signals of hybridization in the evolutionary history of the clade (table S10).

Further insights could be gained through comparative studies of other unisexual vertebrates. In *Darevskia* lizards, the specificity of the sexual species involved in the formation of the multiple parthenogenetic lineages and the apparent lack of correlation between gene flow and divergence have been interpreted as support for the phylogenetic constraint hypothesis (28). The evolutionary divergence separating parents of the unisexuals in these

two clades appears to be comparable (table S11). *Cobitis* fishes reproduce by gynogenesis, and crossing data in this group are also consistent with the balance hypothesis (29). The rare formation hypothesis has been a favored explanation in the gynogenetic Amazon molly (*Poecilia formosa*) (30). Age estimates for the unisexual whiptails based on the RAD data are much younger than those for *P. formosa* (median formation times ~160,000 to 250,000 versus ~500,000 generations ago). Resolving the mechanisms that underlie these similarities and differences will help advance our understanding of some of the longest-standing areas of inquiry in biology.

REFERENCES AND NOTES

- J. J. Kolbe, M. Leal, T. W. Schoener, D. A. Spiller, J. B. Losos, *Science* **335**, 1086–1089 (2012).
- S. Lamichhaney et al., *Nature* **518**, 371–375 (2015).
- M. D. McGee et al., *Nature* **586**, 75–79 (2020).
- S. A. Taylor, E. L. Larson, *Nat. Ecol. Evol.* **3**, 170–177 (2019).
- R. Abbott et al., *J. Evol. Biol.* **26**, 229–246 (2013).
- J. C. Avise, *Clonality: The Genetics, Ecology, and Evolution of Sexual Abstinence in Vertebrate Animals* (Oxford Univ. Press, 2008).
- N. H. Barton, *Mol. Ecol.* **10**, 551–568 (2001).
- K. L. Laskowski, C. Doran, D. Bierbach, J. Krause, M. Wolf, *Nat. Ecol. Evol.* **3**, 161–169 (2019).
- J. W. Wright, in *Herpetology of the North American Deserts: Proceedings of a Symposium*, P. R. Brown, J. W. Wright, Eds. (Southwestern Herpetologists Society, 1994), pp. 255–271.
- H. Gadow, *Proc. Zool. Soc. Lond.* **76**, 277–375 (1906).
- C. E. Burt, *Bull. U.S. Natl. Mus.* **154**, 1–286 (1931).
- Supplementary materials are available online.

13. A. J. Barley *et al.*, *Mol. Phylogenet. Evol.* **132**, 284–295 (2019).
14. C. Moritz *et al.*, in *Evolution and Ecology of Unisexual Vertebrates*, R. M. Dawley, J. P. Bogart, Eds. (New York State Museum, 1989), pp. 87–112.
15. M. Stöck, K. P. Lampert, D. Möller, I. Schlupp, M. Scharf, *Mol. Ecol.* **19**, 5204–5215 (2010).
16. A. A. Lutes, D. P. Baumann, W. B. Neaves, P. Baumann, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9910–9915 (2011).
17. C. J. Cole *et al.*, *Bull. Mus. Comp. Zool.* **161**, 285–321 (2017).
18. D. A. Marques, J. I. Meier, O. Seehausen, *Trends Ecol. Evol.* **34**, 531–544 (2019).
19. A. A. Newton *et al.*, *Development* **143**, 4486–4494 (2016).
20. J. A. Coyne, H. A. Orr, *Evolution* **43**, 362–381 (1989).
21. J. A. P. Hamlin, M. S. Hibbins, L. C. Moyle, *Evol. Lett.* **4**, 137–154 (2020).
22. D. R. Matute, B. S. Cooper, *Evolution* **75**, 764–778 (2021).
23. C. Roux *et al.*, *PLOS Biol.* **14**, e2000234 (2016).
24. M. Schumer *et al.*, *eLife* **3**, e02535 (2014).
25. J. B. W. Wolf, H. Ellegren, *Nat. Rev. Genet.* **18**, 87–100 (2017).
26. K. S. Pfennig, *Science* **318**, 965–967 (2007).
27. J. W. Wright, C. H. Lowe, *Copeia* **1968**, 128–138 (1968).
28. S. Freitas *et al.*, *Evolution* **76**, 899–914 (2022).
29. K. Janko *et al.*, *Mol. Ecol.* **27**, 248–263 (2018).
30. W. C. Warren *et al.*, *Nat. Ecol. Evol.* **2**, 669–679 (2018).
31. A. Barley *et al.*, Data for: Complex patterns of hybridization and introgression across evolutionary timescales in Mexican whiptail lizards (*Aspidoscelis*), Mendeley Data (2018).
32. A. Barley, A. Nieto-Montes de Oca, N. Manríquez-Morán, R. Thomson, The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization, Zenodo (2022).

ACKNOWLEDGMENTS

We thank G. Pauly and N. Camacho at the Natural History Museum of Los Angeles County, B. Hollingsworth at the San Diego Natural History Museum, J. Campbell, the Sam Noble Museum, the UTEP Biodiversity Collections, J. Cole and the American Museum of Natural History, the Museum of Vertebrate Zoology, R. Klabacka and the Auburn University Museum of Natural History, the California Academy of Sciences, the LSU Museum of Natural Science, and the Museum of Southwestern Biology for providing specimen and/or tissue loans in support of this research. We also thank U. García-Vázquez and A. Peralta for providing tissue samples for this study. We thank D. Lara-Tufiño and A. Wright for assistance with fieldwork. A.N.-M.d.O. thanks the Programa de Apoyos para la Superación del Personal Académico of the Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, for financial support during a sabbatical leave in 2018 to 2019. We thank the New Mexico Department of Game and Fish; the California Department of Fish and Wildlife; the State of Arizona Game and Fish Department; the Coronado National Forest; and the Secretaría de Medio Ambiente y Recursos Naturales, Dirección General de Fauna Silvestre for providing sampling permits for this research. We thank S. McDevitt and the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley and G. Wiley at the Oklahoma Medical Research Foundation for assistance with the Illumina sequencing. Some analyses involved use of the University of Hawai'i Information Technology Services high-performance computing cluster and the CIPRES Science Gateway. We thank T. Reeder for comments on the manuscript and valuable discussions about whiptail biology. **Funding:** This study received support from National Science Foundation grant DEB-1754350 (to R.C.T.); National Science Foundation grant DEB-190017 (to R.C.T. and A.J.B.); and Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México (Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica IN217818) (to A.N.-M.d.O.). **Author contributions:** Conceptualization: A.J.B., A.N.-M.d.O., N.L.M.-M., and R.C.T. Methodology: A.J.B. and R.C.T. Investigation: A.J.B., A.N.-M.d.O., N.L.M.-M., and R.C.T. Visualization: A.J.B., A.N.-M.d.O., N.L.M.-M., and R.C.T. Funding acquisition: A.J.B., A.N.-M.d.O., N.L.M.-M., and R.C.T. Writing – original draft: A.J.B. Writing – review & editing: A.J.B., A.N.-M.d.O., N.L.M.-M., and R.C.T. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Newly created genomic data are available at NCBI BioProject (SRA PRJNA 827355). Genetic data from previous studies are available at NCBI BioProject (SRA PRJNA707030 and PRJNA 762930) and Mendeley Data (31). Other data and code for running analyses are available from Zenodo (32). **License information:** Copyright © 2022 the authors, some rights reserved;

exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1593
Materials and Methods

Figs. S1 to S13
Tables S1 to S11
References (33–89)
MDAR Reproducibility Checklist
Data S1 to S3

Submitted 9 November 2021; accepted 28 June 2022
10.1126/science.abn1593

NANOPHOTONICS

Cavity-mediated electron-photon pairs

Armin Feist^{1,2†}, Guanhao Huang^{3,4†}, Germaine Arend^{1,2†}, Yujia Yang^{3,4†}, Jan-Wilke Henke^{1,2}, Arslan Sajid Raja^{3,4}, F. Jasmin Kappert^{1,2}, Rui Ning Wang^{3,4}, Hugo Lourenço-Martins^{1,2}, Zheru Qiu^{3,4}, Junqiu Liu^{3,4}, Ofer Kfir^{1,2}, Tobias J. Kippenberg^{3,4*}, Claus Ropers^{1,2*}

Quantum information, communication, and sensing rely on the generation and control of quantum correlations in complementary degrees of freedom. Free electrons coupled to photonics promise novel hybrid quantum technologies, although single-particle correlations and entanglement have yet to be shown. In this work, we demonstrate the preparation of electron-photon pair states using the phase-matched interaction of free electrons with the evanescent vacuum field of a photonic chip-based optical microresonator. Spontaneous inelastic scattering produces intracavity photons coincident with energy-shifted electrons, which we employ for noise-suppressed optical mode imaging. This parametric pair-state preparation will underpin the future development of free-electron quantum optics, providing a route to quantum-enhanced imaging, electron-photon entanglement, and heralded single-electron and Fock-state photon sources.

Optical parametric processes generate quantum correlations of photons without changing the state of the optical medium involved. Entangled photons from parametric down-conversion (1) are an essential resource for heralded single-photon sources, quantum communication, and quantum teleportation (2). Spatially separated entanglement afforded by such “twin beam” pairs was extended to various physical contexts, with examples in atomic beams (3), electron spins (4), and photon-phonon correlations in micromechanical systems (5). Free-electron beams are an emerging target for quantum manipulation and sensing, promising quantum-enhanced imaging (6, 7), spectroscopy (8–10), and excitation (11–14). A variety of technologies that bridge electron microscopy and photonics (15) are being established to join the most powerful probes in atomic-scale imaging and spectroscopy, respectively. Among these, stimulated near-field scattering offers mode-specific probing of optical properties, whereas spontaneous electron energy loss and cathodoluminescence access electronic transitions and the

total photonic density of states (15). Structural design has been shown to allow for a tailoring of the spectral and spatial properties of electron-driven radiation (16, 17). Harnessing quantum optics approaches, measurements of photon statistics are used to reveal single quantum emitters (18) or photon bunching (19, 20). Theoretical work predicted single-photon cathodoluminescence into a waveguide (21), and recent experiments studied the electron-induced excitation of whispering gallery modes (22, 23) and optical fibers (24). However, impeded by a lack of mode-specific and sufficiently strong coupling, correlations between single electrons and well-defined photonic states have remained elusive. In this study, we use spontaneous inelastic scattering via the evanescent field of a high-quality factor (Q) photonic chip-based optical microresonator to generate free-electron cavity-photon pair states. We characterize the dual-particle heralding efficiencies and demonstrate coincidence imaging of the cavity mode with a substantial background suppression.

The interaction of electron beams with cavities and resonant structures represents a universal scheme for generating electromagnetic radiation. In the quantum optical description, the inelastic scattering can be modeled as a coupling of free electrons to optical vacuum fields (8). Scattering with the evanescent field of the optical microresonator, an electron at energy E , generates intracavity photons at frequencies ω in an energy-conserving manner (Fig. 1A), described by the scattering matrix

¹Max Planck Institute for Multidisciplinary Sciences, D-37077 Göttingen, Germany. ²4th Physical Institute – Solids and Nanostructures, University of Göttingen, D-37077 Göttingen, Germany. ³Institute of Physics, Swiss Federal Institute of Technology Lausanne (EPFL), CH-1015 Lausanne, Switzerland. ⁴Center for Quantum Science and Engineering, Swiss Federal Institute of Technology Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

*Corresponding author. Email: claus.ropers@mpinat.mpg.de (C.R.); tobias.kippenberg@epfl.ch (T.J.K.)

†These authors contributed equally to this work.

$\hat{S} = \exp(g_{\text{qu}} \hat{a}^\dagger \hat{b} - h.c.)$. Here, \hat{a}^\dagger is the creation operator of the optical mode, g_{qu} is the vacuum coupling strength (25, 26), and $h.c.$ is the Hermitian conjugate. The ladder operator \hat{b} reduces the energy of an electron by that of one photon. The interaction induces entanglement between the electron energy and the cavity population and results in the state

$$|\psi_e, \psi_{\text{ph}}\rangle = \sum_{n=0}^{\infty} c_n |E - n\hbar\omega\rangle |n\rangle \quad (1)$$

(ψ_e , electron state; ψ_{ph} , photonic state; \hbar , reduced Planck's constant) with the coefficients $c_n = \exp\left(-\frac{|g_{\text{qu}}|^2}{2}\right) \frac{g_{\text{qu}}^n}{\sqrt{n!}}$ corresponding to Poissonian scattering probabilities (25). For a weak vacuum coupling strength $|g_{\text{qu}}| \ll 1$, the state is dominated by the zero- and one-photon contributions

$$|\psi_e, \psi_{\text{ph}}\rangle \propto |E\rangle |0\rangle + g_{\text{qu}} |E - \hbar\omega\rangle |1\rangle + \mathcal{O}(g_{\text{qu}}^2) \quad (2)$$

where \mathcal{O} is order of approximation. Our measurements are designed to probe this state by detecting single photons in coincidence with inelastically scattered electrons by the energy of one photon. In the experiment, the continuous electron beam of a transmission electron microscope [120-keV beam energy, 25-nm focal diameter; see (27) section S1.3] traverses a photonic chip-based microresonator (Fig. 1C) and interacts with the vacuum fields to populate empty cavity modes.

The Si_3N_4 microresonator (28) is designed for low optical loss, efficient fiber coupling, and free-space near-field access [for optical characteristics, see (27) S1.2]. Enhancing the interaction by electron-light phase matching (29–32), we engineered the resonator cross section (2.1 μm by 650 nm) (26) for maximum photon generation probability per electron, which is predicted up to 10% around 0.8-eV photon energy (a wavelength of 1.55 μm).

Cavity photons generated in the clockwise propagating mode (Fig. 1A) are coupled to a bus waveguide and are further guided by optical fibers to a single-photon avalanche diode (SPAD). The energy and arrival time of each electron is measured with an event-based detector behind a magnetic prism spectrometer (Fig. 1B), allowing for electron-photon coincidence experiments.

First, we spatially and spectrally map the electron-induced cavity excitation by placing the electron beam in proximity of the resonator (Fig. 2, A and B). Measured with an optical spectrometer, the overall spectral range of detected radiation spans from 1520 to 1620 nm, primarily limited by the bandwidth of electron-light phase matching and consistent with numerical simulations [Fig. 2C; see (27) S2.2]. The spectrum exhibits a comb-like structure (Fig. 2D) as a result of free-electron coupling

to the microresonator modes a_μ [μ , mode index; see (27) S2.1 for a multimode description of the interaction]. The 1.58-nm spacing of the emission peaks matches the optically characterized quasi-transverse magnetic mode free-spectral range and confirms the predominant photon extraction from this mode family [see (27) S1.2].

Spatial characterization of the intracavity excitation is obtained by raster scanning of the electron beam (Fig. 2, A and B) and detection of emitted light with a SPAD. Figure 2E displays the obtained single-mode (top) and spectrally integrated (bottom) photon rates, which decay exponentially with distance from the structure, tracing the near-field mode profile in this spectral range. The strongest photon generation is observed for electrons passing the ring resonator tangentially, as expected for phase-matched electron-light interaction. Figure 2F shows the photon generation rates along the chip surface. The single-mode count rate shows a pronounced oscillation in the radial distance to the center of the ring resonator. This is caused by the interference between sequential interactions (Fig. 2A) of an electron with the cavity vacuum field (33), proving the conservation of mutual coherence in the scattering process. Similar Ramsey-type interference is also observed for an externally pumped ring resonator (26). Integrating the full mode spectrum of the cathodoluminescence averages out the oscillatory pattern, in agreement with numerical simulations (Fig. 2G). The total scattering probability is modeled as the sum $\mathbb{P} = \sum_\mu |g_{\text{qu},\mu}|^2$ of the individual mode contributions $g_{\text{qu},\mu}$ [see (27) S1.1 for details].

The spontaneous creation of photons by single electrons satisfies energy-momentum conservation, lowering the energy of an electron by $\hbar\omega$ and transferring the corresponding momentum to the cavity photon (21). We probe this correlated electron-photon pair state by coincidence measurements of both particles (Fig. 3). The electron beam is held fixed in the near field of the cavity at a distance of ~ 160 nm from the surface. At this position, we detect photons with a probability of 4.6×10^{-5} per electron passing the structure. Considering coupling and detection losses [see (27) S1.4], this corresponds to an intrinsic generation probability of $\sim 2.5\%$.

The arrival time and kinetic energy of each electron is measured by event-based spectroscopy, using the stream of photon events recorded by the SPAD for time tagging [Fig. 1A and (27) S1.5 for details]. Figure 3A shows the energy- and time-dependent histogram of electron arrivals relative to the photon event closest in time. The two main features observed are a time-independent background of accidental coincidences around zero energy loss and an anticipated sharp correlation peak around 0.8-eV energy loss and 0-ns time delay.

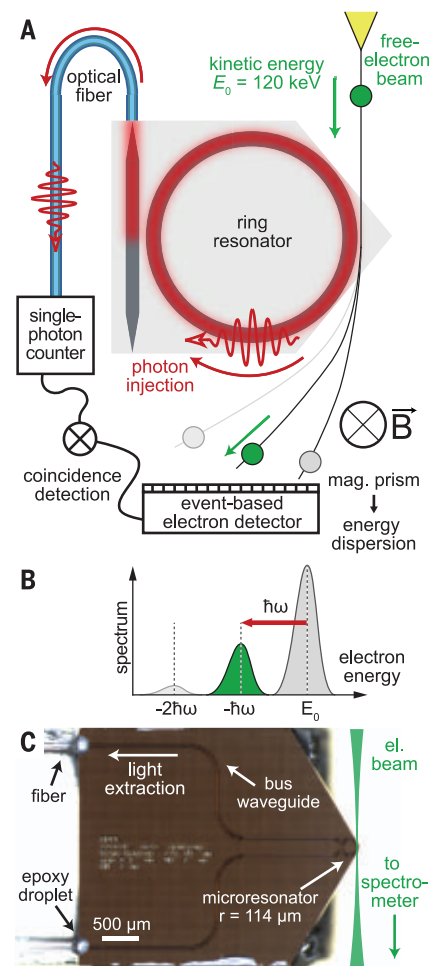


Fig. 1. Coincidence detection of electron-photon pairs generated at a photonic chip. (A) High-energy electrons traverse a fiber-coupled Si_3N_4 microring resonator, generating intracavity photons. Single photons and inelastically scattered electrons are detected in coincidence. \vec{B} , magnetic field. (B) Electron spectrum after interaction. (C) Optical microscope image of the photonic chip with a sketch of the passing electron beam. r , microresonator radius.

The spectrum of the correlated electrons (Fig. 3B, red) is downshifted by the energy of one photon with respect to the zero-loss peak (ZLP) (black). As a result of the narrow electron-light phase-matching bandwidth of ~ 50 meV, these spectral distributions are very similar in broadening and shape [~ 0.5 -eV width; see (27) S1.6 for details].

We next insert a mechanical slit in the electron spectrometer (Fig. 3C) to reject electrons that were not scattered, allowing for higher total electron flux, and thus photon rates, without detector saturation. The background-corrected time profile (inset) shows the precise temporal structure of the correlated electron-photon pairs with a width of 3.91 ns. In principle, the temporal correlation trace should reproduce

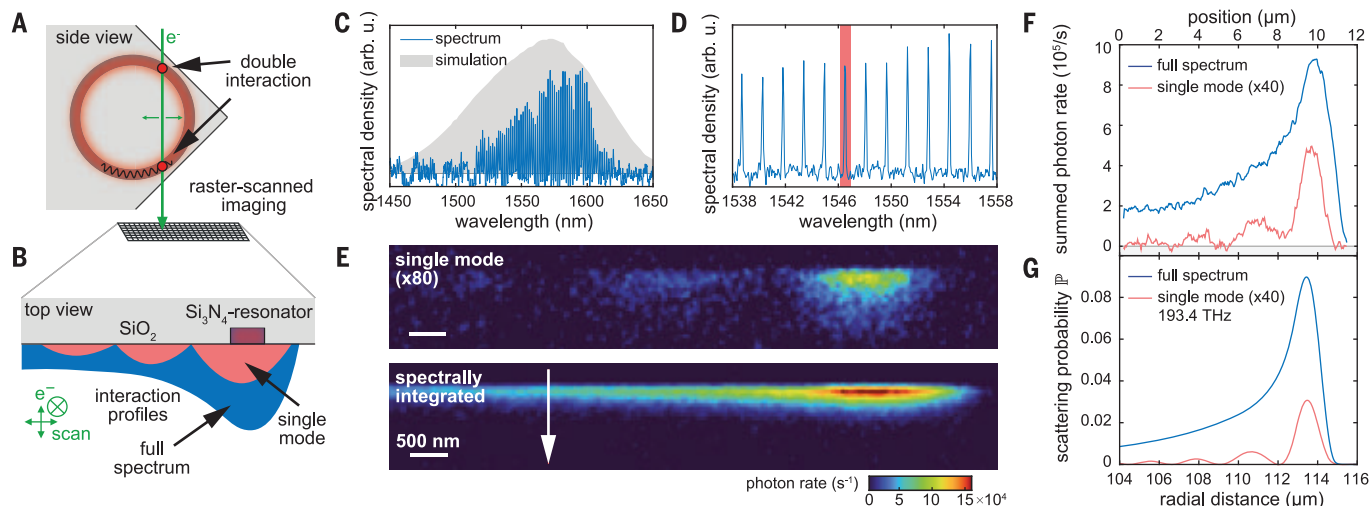


Fig. 2. Spatial and spectral mapping of intracavity photon generation. (A and B) Schematic of the measurement geometry. The electron beam (green) passes parallel to the chip surface (gray) and interacts twice (red dots) with the ring resonator (dark red). (C) Optical emission spectrum (blue; electron beam passes resonator tangentially) and simulated spectral envelope (gray). (D) Close-up view of the spectral mode comb and a selected single cavity mode (red). (E) Raster-scanned photon-generation maps [saturation corrected; see (27) S1.4] for the selected mode and the full spectrum [intensity patterns illustrated in (B) with red and blue]. (F) Signals from (E) integrated along the direction of the white arrow. (G) Simulation of the position-dependent electron scattering probability (50-nm distance).

the cavity decay time, but it is not resolved in the present experiments [current device: ~ 0.5 -ns lifetime for $Q \sim 5.5 \times 10^3$; see (27) S1.2].

By selecting loss-scattered electrons within a specific time window, we identify correlated events on a single-particle basis. The measurement of the electron energy for each event, in contrast to conventional optical spontaneous parametric down-conversion, presents a direct measure of the energy quanta n deposited in the optical cavity. The single-electron energy detection at the first energy loss sideband therefore represents an optical state projection from the entangled state (Eq. 1) onto the (non-classical) single-photon intracavity Fock state (13, 21, 25). This general principle, illustrated in the schematic of Fig. 3E, enables interparticle heralding schemes for either electrons or photons. Notably, interactions with multiple electrons—relevant for studying electron-electron correlations—can be excluded owing to the multihit capability of the detector.

Such heralded particle sources are quantified in terms of the measured rates of electrons (R_e), photons (R_p), and correlated events (R_{pe}) and the Klyshko heralding efficiencies $\eta_K^i = R_{pe}/R_j$ ($i, j = e, p, i \neq j$), which describe the conditional probability of experimentally detecting a heralded particle [see (27) S1.6]. For the data shown in Fig. 3C, we measure $\eta_K^e \sim 0.11\%$ and $\eta_K^p \sim 57\%$ for photons and electrons, respectively. The less-efficient heralding of photons follows from considerably higher losses in their output coupling and detection [see (27) S1.4]. Not being a fundamental physical limitation, we expect appreciable improvements on the photon collection efficiency with technical optimizations, including use of super-

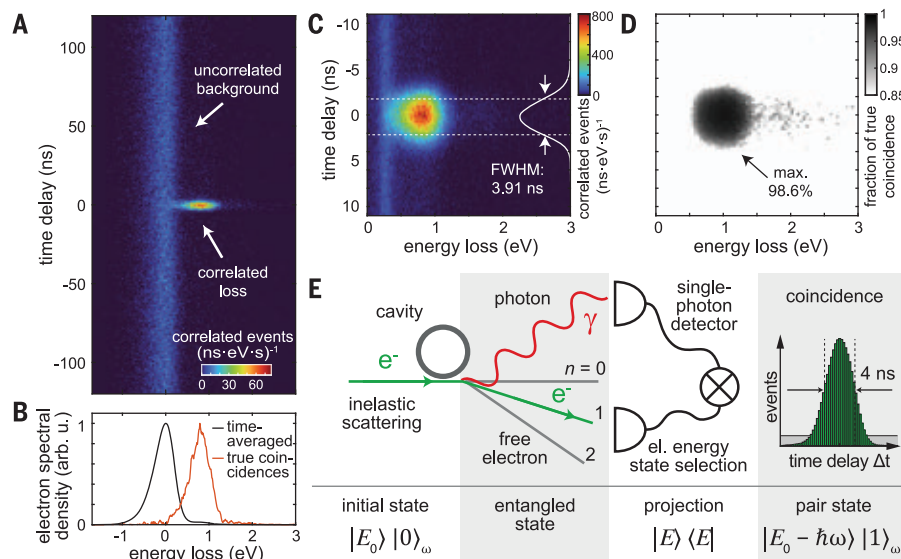


Fig. 3. Time- and energy-correlated electron-photon pairs. (A) Two-dimensional correlation histogram over the electron energy and relative photon timing (30-s integration, 0.65-pA electron current on detector) showing a strong coincidence peak. (B) Spectral distributions of the uncorrelated and correlated electrons (background subtracted). (C and D) Close-up view of correlation histogram (ZLP blocked, 60-s integration, 46-pA beam current at sample) with coincidence time trace (white) and fraction of true coincidences. FWHM, full width at half maximum. (E) Illustration of cavity-mediated inelastic electron-photon scattering and coincidence measurement.

conducting detectors and strongly overcoupled resonators. Taking into account particle losses in transmission and detection, we estimate intrinsic heralding efficiencies η_i^i of $\sim 50\%$. In principle, near-unity intrinsic heralding efficiencies can be achieved (21) [see detailed estimate of η_i^i in (27) S2.3].

Employing both particle channels facilitates correlation-enhanced measurements, isolating physical scattering events from uncorrelated noise such as detector dark counts. As a figure of merit for noise suppression, the fraction of true coincidences $R_{pe}/(R_{pe} - R_{acc})$, with the uncorrelated (accidental) background rate R_{acc}

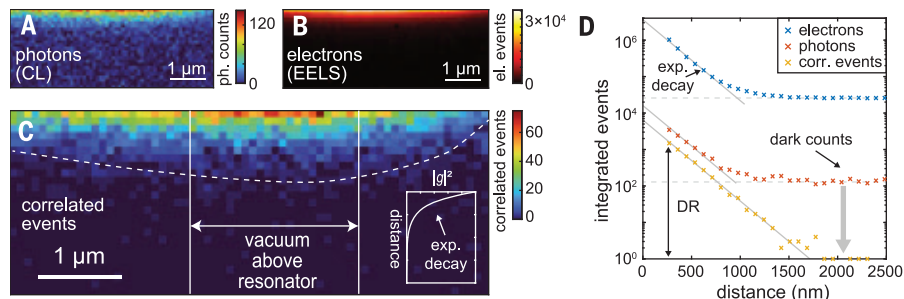


Fig. 4. Correlation-enhanced cavity mode imaging. Spatial distributions of (A) photons [cathodoluminescence (CL)], (B) loss-scattered electrons [electron energy-loss spectroscopy (EELS)], and (C) correlated events (30-ms integration per pixel, starting ~ 270 nm above the chip surface; see Fig. 2 for measurement geometry). The positions of the resonator and the mode envelope (solid and dashed lines, respectively, serve as a guide to the eye) are indicated. (Inset) Sketch of the cavity field decay into the vacuum (scattering probability $|g|^2$). The energy windows applied to (B) and (C) are optimized individually [see (27) S1.6]. (D) Logarithmic plot comparing the distance-dependent exponential mode decay from (A) to (C) (integrated in the area above the resonator, effectively 1-s acquisition per point). The lowest noise and highest dynamic range (DR) (more than three orders of magnitude) are observed for the time-correlated imaging.

is defined for selecting a specific time-delay and energy-loss window (Fig. 3D), reaching 98.6%—i.e., a coincidence-to-accidental ratio (CAR) of ~ 75 [see (27) S1.6]. This demonstrates the high-fidelity generation of correlated electron-photon pairs, promising quantum-enhanced imaging.

Here we provide a proof-of-concept demonstration of coincidence-gated raster mapping of the resonator mode. Specifically, Fig. 4, A and B, shows the time-integrated electron and photon signals, respectively, and Fig. 4C displays the correlated events only. To quantify the correlation-induced improvement in image contrast, Fig. 4D compares the respective count rates for the individual and correlated signals on a logarithmic scale. Both the photon and electron signals trace the exponential decay of the evanescent field away from the structure, leveling off at constant values for larger distances. These background offsets limit the dynamic range of the mode imaging and are determined by the uncorrelated noise rates—i.e., the photodetector dark counts (130 counts/s) and residual overlap of the ZLP with the energetic gate (1.5×10^4 counts/s), respectively. Rejecting the uncorrelated background, the correlated signal (Fig. 4D, yellow curve) resolves the evanescent decay of the cavity field over longer distances. Despite a slightly smaller signal ($\eta_K^e < 1$), we improve the dynamic range by two orders of magnitude [see (27) S2.5].

From a broader perspective, harnessing correlations of electrons with radiative emission (34) shows promise for enhancing contrast and resolution, as recently shown in the study of core-level (35) and valence electronic excitations in nanostructured materials (36). Using postselection in the electronic and photonic degrees of freedom (9, 13), such schemes can be generalized to trace state-specific scattering

cross sections and create heralded pair states as a function of linear or angular momentum, polarization or spin, or frequency or energy.

Our findings demonstrate and apply the parametric generation of electron-photon pair states. Besides showing the capability for contrast enhancement in correlation-gated imaging, we implement flexible on-chip electron-heralded photon and photon-heralded single-electron sources. The integrated photonics platform allows for flexible electron-light phase-matching and spectral tuning of the generated cavity state. Postselection of multiloss electron events will facilitate a unique approach to generate tailored and higher-order photon Fock states. By merging free-electron quantum optics with integrated photonics, we provide a pathway toward a new class of hybrid quantum technology that relies on entangled electrons and photons, with applications ranging from photonic quantum state synthesis to quantum-enhanced sensing and imaging.

REFERENCES AND NOTES

- P. G. Kwiat *et al.*, *Phys. Rev. Lett.* **75**, 4337–4341 (1995).
- N. Gisin, R. Thew, *Nat. Photonics* **1**, 165–171 (2007).
- F. Borselli *et al.*, *Phys. Rev. Lett.* **126**, 083603 (2021).
- B. Hensen *et al.*, *Nature* **526**, 682–686 (2015).
- R. Riedinger *et al.*, *Nature* **556**, 473–477 (2018).
- P. Kruit *et al.*, *Ultramicroscopy* **164**, 31–45 (2016).
- A. E. Turner, C. W. Johnson, P. Kruit, B. J. McMorran, *Phys. Rev. Lett.* **127**, 110401 (2021).
- V. Di Giulio, M. Kociak, F. J. G. de Abajo, *Optica* **6**, 1524 (2019).
- E. Rotunno *et al.*, arXiv:2106.08955 [quant-ph] (2021).
- M. Tsarev, A. Ryabov, P. Baum, *Phys. Rev. Lett.* **127**, 165501 (2021).
- Y. Pan, A. Gover, *Phys. Rev. A* **99**, 052107 (2019).
- Z. Zhao, X.-Q. Sun, S. Fan, *Phys. Rev. Lett.* **126**, 233402 (2021).
- A. Ben Hayoun *et al.*, *Sci. Adv.* **7**, eaabe4270 (2021).
- O. Kfir, V. Di Giulio, F. J. G. de Abajo, C. Ropers, *Sci. Adv.* **7**, eaaf6380 (2021).
- A. Polman, M. Kociak, F. J. García de Abajo, *Nat. Mater.* **18**, 1158–1171 (2019).
- J. Christopher *et al.*, *Nanophotonics* **9**, 4381–4406 (2020).
- C. Roques-Carmes *et al.*, *Science* **375**, ea9m9293 (2022).
- R. Bourrellier *et al.*, *Nano Lett.* **16**, 4317–4321 (2016).

- S. Meuret *et al.*, *Phys. Rev. Lett.* **114**, 197401 (2015).
- M. Solà-García *et al.*, *ACS Photonics* **8**, 916–925 (2021).
- X. Bendaña, A. Polman, F. J. García de Abajo, *Nano Lett.* **11**, 5099–5103 (2011).
- N. Müller *et al.*, *ACS Photonics* **8**, 1569–1575 (2021).
- Y. Auad *et al.*, *Nano Lett.* **22**, 319–327 (2022).
- M. Scheucher, T. Schachinger, T. Spielauer, M. Stöger-Pollach, P. Haslinger, arXiv:2110.05126 [physics.optics] (2021).
- O. Kfir, *Phys. Rev. Lett.* **123**, 103602 (2019).
- J.-W. Henke *et al.*, *Nature* **600**, 653–658 (2021).
- Materials and methods are available as supplementary materials.
- M. H. P. Pfeiffer *et al.*, *Optica* **3**, 20 (2016).
- M. Kozák *et al.*, *Opt. Express* **25**, 19195–19204 (2017).
- N. V. Sapra *et al.*, *Science* **367**, 79–83 (2020).
- R. Dahan *et al.*, *Science* **373**, ea7j128 (2021).
- O. Kfir *et al.*, *Nature* **582**, 46–49 (2020).
- N. Talebi, *New J. Phys.* **16**, 053021 (2014).
- R. J. Graham, J. Spence, H. Alexander, *Proc. MRS* **82**, 235 (1986).
- D. Jannis, K. Müller-Caspar, A. Béché, J. Verbeeck, *Appl. Sci.* **11**, 9058 (2021).
- N. Varkentina *et al.*, arXiv:2202.12520 [cond-mat.mes-hall] (2022).
- A. Feist *et al.*, Figure data for cavity-mediated electron-photon pairs, version 1, Zenodo (2022); <https://doi.org/10.5281/zenodo.6639005>.

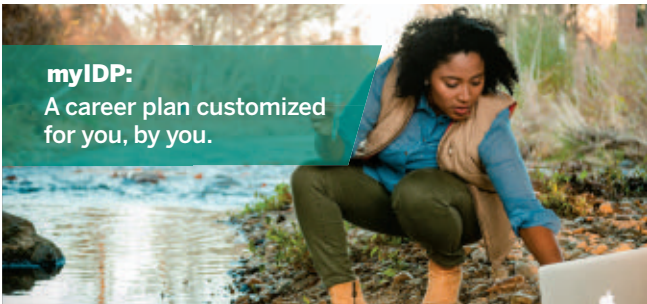
ACKNOWLEDGMENTS

We thank R. Haindl and E. Maddox for discussions on electron event clustering, M. Möller for supporting the scanning transmission electron microscope alignment, and J. Borchert and T. Weitz for providing us with the near-infrared spectrometer. We thank members of the eBEAM consortium—in particular A. Polman, F. J. García de Abajo, N. Talebi, J. Verbeeck, and M. Kociak—for useful discussions and feedback. **Funding:** All samples were fabricated in the Center of MicroNanoTechnology (CMI) at EPFL. This material is based on work supported by the Air Force Office of Scientific Research under award FA9550-19-1-0250. This work was further supported by the Swiss National Science Foundation under grant agreement 185870 (Ambizione). The experiments were conducted at the Göttingen UTEM Lab, funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) through 432680300/SFB 1456 (project C01) and the Gottfried Wilhelm Leibniz program, and the European Union's Horizon 2020 (EU H2020) research and innovation programme under grant agreement 101017720 (FET-Proactive EBEAM). Y.Y. acknowledges support from the EU H2020 research and innovation program under the Marie Skłodowska-Curie IF grant agreement 101033593 (SEPhIM). O.K. acknowledges the Max Planck Society for funding from the Manfred Eigen Fellowship for postdoctoral fellows from abroad. **Author contributions:** Conceptualization: C.R., O.K., and T.J.K. Methodology: A.F. and G.A. (event-based spectroscopy), A.S.R., O.K., and J.L. (sample design), and G.H. and C.R. (theory). Investigation: A.F., G.A., J.-W.H., and F.J.K. (transmission electron microscopy experiments and optical setup), Y.Y. and G.H. (numerical simulation), and A.S.R. and Y.Y. (sample packaging and characterization). Formal analysis: A.F., G.A., G.H., and H.L.-M. Resources: J.L., Z.Q., and R.N.W. (sample fabrication). Project administration, supervision, and funding acquisition: C.R. and T.J.K. Visualization: A.F., G.H., G.A., and Y.Y. Writing: A.F., G.H., G.A., Y.Y., T.J.K., and C.R., with comments and feedback from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The code and data used to produce the plots within this work are available in Zenodo (37). **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo5037
Materials and Methods
Supplementary Text
Figs. S1 to S9
Table S1
References (38–80)

Submitted 25 February 2022; accepted 7 July 2022
10.1126/science.abo5037



For your career in science, there's only one **Science**

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.



Visit the website and start planning today!
myIDP.sciencecareers.org

ScienceCareers In partnership with: MAAS



Director Brookhaven National Laboratory

BROOKHAVEN SCIENCE ASSOCIATES (BSA) is seeking an internationally recognized science leader with a demonstrated record of high achievement in leading-edge scientific research and management to serve as the Director of the U.S. Department of Energy's (DOE) Brookhaven National Laboratory.

One of 10 national labs primarily supported by DOE's Office of Science, Brookhaven Lab is a multidisciplinary laboratory with seven Nobel Prize-winning discoveries, 37 R&D 100 Awards, and 75 years of pioneering research.

You must be able to inspire and engage the Laboratory and its regional, national, and international partners with an exciting vision for the future, and lead the Brookhaven Lab community in formulating and implementing a strategic vision to deliver on the outcomes that are embedded in the Lab's strategic mission.

You will have a proven track record in attracting, developing, and retaining the highest-quality diverse scientific, managerial, and professional leadership to ensure a future talent pipeline, and be able to demonstrate leadership and commitment to organizational values vital to the Lab's culture.

Brookhaven Lab is an equal opportunity employer that values diversity, equity and inclusion. We are committed to ensuring that all qualified applicants receive consideration for employment and will not be discriminated against on the basis of race, color, religion, sex, sexual orientation, gender identity, national origin, age, status as a veteran, disability or any other federal, state or local protected class.



Learn more about the Laboratory Director role and apply – jobs.bnl.gov



YOUR NEXT BIG SCIENTIFIC DISCOVERY: A NEW JOB.

[Find your next job at ScienceCareers.org](http://ScienceCareers.org)

Whether you're looking to get ahead, get into, or just plain get advice about careers in science, there's no better or more trusted authority. Get the scoop, stay in the loop with *Science Careers*.

ScienceCareers

FROM THE JOURNAL SCIENCE MAAS

By Olivia Lanes

Finding a new calling

When I arrived at the mountaintop observatory, I was bursting with excitement. I was about to finish my junior year studying astrophysics, with plans to apply for a Ph.D. in the field, and I had the rare opportunity to go on an observing run. As I arrived at the tiny cabin that would be my home for the next few days, I finally felt like an astronomer. Unfortunately, the excitement didn't last. As I sat alone, staring at a computer screen instead of the night sky, pressing a combination of buttons over and over to control the telescope and retrieve data, I had to admit: This wasn't what I had imagined.

I had wanted to be an astronomer ever since my father gave me the DVD box set of *Cosmos: A Personal Voyage* when I was a teenager. Watching Carl Sagan talk about the heavens, I thought I had found my calling. But after my experience at the observatory, I had to face a new truth. If this was the astronomer's life, then the astronomer's life wasn't for me.

I returned to college on the verge of panic. I still wanted to pursue a Ph.D. to learn more about the natural world and conduct independent research one day. But where was my future, if not in astronomy?

I thought I might be better off doing research that was more interdisciplinary and collaborative. I recalled a quantum physics class I had taken and thought that field might fit the bill. My academic adviser was skeptical. I can't blame him; I was throwing away a ton of original astronomical data to start from scratch in a new field, and I only had 2 weeks left in the school year to design a new thesis project. But I embraced the challenge. As I read and thought about my potential new field, my brain buzzed with excitement—and anxiety about what my adviser would say. I put together a proposal I was eager to pursue, and to my joy and relief, my adviser approved it.

In some ways, my new project mirrored the observatory experience, as I sat for extended periods in a darkened lab, aligning lasers and shining them through a pair of slits. But being able to shape and control quantum phenomena inspired me in a way astronomy did not. I also recruited a friend to work with me, which made the research less lonely. As we talked through the project, explaining complex topics to each other and reflecting on the big picture of what we wanted to achieve, a revelation dawned: I had been drawn to *Cosmos* not because of the subject, but because of the finesse with which Sagan explained mind-blowing concepts.



“New possibilities can be far more rewarding than doggedly sticking to a years-old dream.”

That was the spark of inspiration I wanted to follow.

I began to try out my communication skills, working with high school physics students and tutoring in the college physics room. I got a real charge from these experiences. Even though they meant I had to work later in the lab to get my research done, I felt I had finally found my niche. I decided to pursue a Ph.D. in quantum—a field where massive breakthroughs are often barely visible to the wider world. This was my chance to be both a scientist and a communicator. During my Ph.D., I volunteered at local elementary schools and participated in other communication and outreach efforts. Sometimes it was hard to justify the amount of time they took away from my thesis work, but speaking to excited students gave me the

extra bump of motivation I needed to return to my research.

As I was completing my Ph.D., I knew I wanted to find work that let me both contribute to science and draw more people into the field. I interviewed for a wide range of jobs, including consulting work, industry positions, and even professorships at liberal arts colleges. It felt like serendipity when IBM reached out about two potential jobs, one focused on community outreach and the other on quantum engineering. I asked whether I could do both. Now, in addition to conducting research, I lecture at universities, run a summer school, and work with educators.

It's not what I imagined, sitting on my living room floor as a teenager. But as I've learned more about science and myself, I've realized that remaining open to new possibilities can be far more rewarding than doggedly sticking to a years-old dream. ■

Olivia Lanes is a researcher and team lead for education at IBM. Send your career story to SciCareerEditor@aaas.org.



Notice is hereby provided that a meeting of the members of the American Association for the Advancement of Science (the “Association”) will convene on September 16, 2022, to consider proposed amendments to the Association’s Charter that would collectively permit the Association’s Board of Directors to adopt, without further action, future amendments to the Association’s Constitution and Bylaws and permit meetings of the Association’s members to occur anywhere in the United States. The full text of the proposed amendments can be found below. The amendments were endorsed by the Board of Directors and approved by the Council on July 22, 2022.

The meeting will take place September 16, 2022, 11 a.m. EDT, at the Boston Marriott Newton Hotel (2345 Commonwealth Avenue, Newton, MA 02466). Members wishing to attend the meeting virtually will also have the opportunity to do so. Please email GovernanceModernization@aaas.org for virtual attendance instructions.

Votes will be accepted in-person or by proxy vote. Details of the proxy vote will be emailed to all members in good standing on or about September 2, 2022. Members who are not physically in attendance, including those attending the meeting virtually, must cast their vote via this proxy before the meeting.

Additionally, at the next annual election of Association officers, members will also be asked to consider amendments to the Association’s Constitution and Bylaws that would permit the Association’s Board of Directors to amend both such governing documents without further action. Notice of such election will be provided at a future date.

Proposed Amendments to the Association’s Charter

The Charter of American Association for the Advancement of Science, as amended, is hereby further amended by inserting the following after Section 3A:

Section 3B. The directors may make, amend or repeal the by-laws of said corporation in whole or in part.

Section 3C. Meetings of the members may be held anywhere in the United States.



PRIZE FOR IMMUNOLOGY

REWARDING HIGH-RISK RESEARCH.
SUPPORTING EARLY-CAREER SCIENTISTS.
HELPING TO FIND CURES FASTER.
APPLY TODAY

Now accepting applications for the Michelson Philanthropies & Science Prize for Immunology.

The Michelson Philanthropies and Science Prize for Immunology focuses on transformative research in human immunology, with trans-disease applications to accelerate vaccine and immunotherapeutic discovery. This international prize supports investigators 35 and younger, who apply their expertise to research that has a lasting impact on vaccine development and immunotherapy. It is open to researchers from a wide range of disciplines including computer science, artificial intelligence/machine learning, protein engineering, nanotechnology, genomics, parasitology and tropical medicine, neurodegenerative diseases, and gene editing.

Application deadline: Oct. 1, 2022.

For more information visit:
www.michelsonmedicalresearch.org

#MichelsonPrizes



"The Michelson Philanthropies & Science Prize for Immunology will greatly impact my future work. As I am just starting my scientific career, it will illuminate my work, spark interest and support me to continue my research in this field."

Paul Bastard, MD, PhD,
Laboratory of Human Genetics of Infectious Diseases, Imagine Institute (INSERM, University of Paris), Paris, France; and The Rockefeller University, New York.

Dr. Bastard received the inaugural Grand Prize for his essay: "Why do people die from COVID-19: Autoantibodies neutralizing type I interferons increase with age."

**GRAND PRIZE:
\$30,000**

**FINALIST PRIZE:
\$10,000**