

When surgery opens a way  
to study the brain p. 256

Saliva sharing indicates close  
relationships pp. 260 & 311

Creating quantum dots  
in glass p. 307

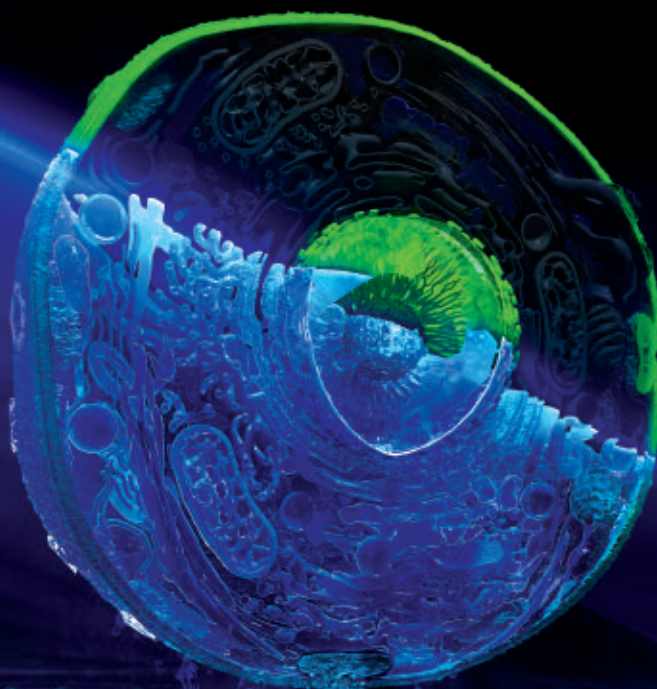
# Science

\$15  
21 JANUARY 2022  
science.org

AAAS

## QUICK PICKS

Sorting cells by  
high-throughput  
imaging p. 315



# Ignite Your Passion

Your passion for discovery transforms a precious sample into a flicker of hope.

It takes more than technology to fight disease. Your vision ignites the drive to discover, the freedom to optimize, and the confidence to impact outcomes.

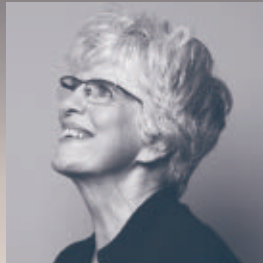
Ignite your passion at [explore.agilent.com/ignite-change](https://explore.agilent.com/ignite-change)





# 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.**

*"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](mailto:philanthropy@aaas.org). Additional details are also available online at [www.aaas.org/1848Society](http://www.aaas.org/1848Society).

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





# Forum launches new era in intelligent computing

The Zhejiang Lab and AAAS partnered to highlight the latest advances driving intelligent computing through a series of lectures and a new journal, *Intelligent Computing*.

On November 16, 2021, more than 10,000 people worldwide viewed the “Innovation Forum on Intelligent Computing” online, which took place at the Zhejiang Lab in Hangzhou, China. This research institute was established in 2017 by the Zhejiang Provincial Government, Zhejiang University, and Alibaba Group. Cosponsored by the Zhejiang Lab, *Science*, and *Science Robotics*, the forum consisted of presentations from computer science and artificial intelligence (AI) experts from across the globe. The presenters focused on this targeted theme: computing for and by intelligence.

In his opening remarks, Shiqiang Zhu, president of the Zhejiang Lab, stated, “Driven by burgeoning technological and industrial revolutions, human society is advancing toward an intelligent era where everything can be digitalized.” That transition depends on a cluster of recent advances in computing. “Artificial intelligence has become a leading technology in global tech innovation, with algorithms, data, and computational power as its three pillars,” Zhu noted. “In particular, intelligent computing has become a primary booster for the development of AI.”

But as Zhu explained, despite these advances, intelligent computation faces many difficulties, including a growing demand for computational power and its enormous energy requirements. “Our goal is to bring [scientists] together to confront the huge challenges we face as a society, and explore the limits of science, thereby setting and leading the trends of intelligent computing in the future,” Zhu said.

Keeping track of those trends requires the collaboration of experts around the world. “We believe in science without borders,” said Bill Moran, publisher, *Science* family of journals, *Science*/AAAS. “We believe in science

advancing society.” In considering this partnership between the Zhejiang Lab and AAAS, he asked: “How can intelligent computing serve science and advance society?”

Answering that question will be one of the goals of the Science Partner Journal (SPI) program’s *Intelligent Computing*—a new peer-reviewed journal that was launched during a signing ceremony at the forum; it will be published through a partnership between the Zhejiang Lab and *Science*/AAAS. Speaking of the new journal, Moran added: “Part of our goal is to work with other scientific societies and research centers to share our knowledge and our expertise in peer-reviewed publishing.”

## From the “miracle year” to now

In the forum, scientists discussed the current state of intelligent computing and its promise for the future. In the first keynote presentation, Zhu said, “Intelligent computing is neither a substitute for supercomputing or cloud computing, nor a simple integration of existing computing. Instead, it is [defined as] the best way to use computing resources and the most appropriate calculation method to solve practical problems according to the needs of [a given] task.” Zhu pointed out that intelligent computing “not only needs to make full use of existing calculations and algorithms, but also promotes the formation of new ones.” Consequently, he sees intelligent computing serving as the foundational technology of a “smart” society.

In the second keynote speech, Hai Jin, chair professor of computer science and engineering at Huazhong University of Science and Technology, described the range of applications for graph computing. Increasing the scale of data and working with increasingly diverse applications creates the need to explore approaches to “the architecture, memory, and algorithm levels,” he explained.



Some of the most exciting advances in intelligent computing over the past few decades have come from Jürgen Schmidhuber, director of the Artificial Intelligence Initiative at King Abdullah University of Science and Technology in Saudi Arabia. In his keynote address, he called 1990–1991 the “miracle year” when his team created the idea of deep learning, which included some automated methods of training a system.

As scientists learn more about biological intelligence, their discoveries can lead to new developments in computing technology, as demonstrated by Yulia Sandamirskaya, leader of the applications research team of Intel’s neuromorphic computing lab. In her talk, she observed that human intelligence spends more time controlling movements than, for example, playing chess; thus artificial neural networks can improve by emulating the processing behind human movements.

Elaborating on the evolutionary development of AI, Yaochu Jin, Alexander von Humboldt Professor for Artificial Intelligence at Bielefeld University in Germany, discussed biological evolution, development, and learning. Then, he described how these mechanisms might serve as complementary tools for automating the design of autonomous AI-based computing systems.

Looking further into automating software, Zhi Jin, a professor of software engineering at Peking University in China, reviewed the early top-down approaches to this process and the potential for a bottom-up approach. She concluded that a hybrid of the two approaches might work best, but that the preferred approach could vary by application.

Addressing numerical applications for high-performance computing, Guangwen Yang, director of the National Supercomputing Center in Wuxi and head of the Intelligent Supercomputing Center of the Zhejiang Lab, mentioned astrophysics and life sciences, advanced manufacturing and materials science, and more. Delving even deeper into such applications, he commented on ways to make intelligent supercomputers.

Wei D. Lu, professor of electrical engineering and computer science at the University of Michigan in Ann Arbor, described energetic bottlenecks in computing. He explained how to improve the efficiency of computation in various ways, including the use of memristors, which emulate the processes of synapses in the nervous system.

Jincang Zhang, distinguished professor of condensed matter and materials physics at Shanghai University in China, reviewed the development of methods in computational materials science, ranging from traditional to data-driven approaches. He showed how advances in

AI—such as predicting the mechanical properties of biopolymer gels—can improve computational materials science.

In the final presentation, Michael Lee, editor of *Science Robotics*, presented a workshop on the editorial process in the *Science* family of journals. This included an overview of many of the journals and the

types of articles they publish, plus some of the key features that an accepted paper will include, such as overcoming technological limits, answering longstanding questions, altering the way scientists look at a problem, and opening new opportunities for R&D.

### Staying the course

In closing the forum, Xinlong Zhao, vice president of Zhejiang Lab, asserted that “intelligent computing is an important cornerstone for human–cyber–physical integration.”

For that reason, he added, it is “critical to create a worldwide cooperative ecosystem for intelligent computing and promote collaborative innovation, in order to greatly boost the development of computing science and technology.”

Research in this area creates many challenges, however. One that Zhao noted is the need for “not only innovative computing architecture, but also a whole set of supporting theoretical systems, which, from a global perspective, have just begun to emerge.”

As intelligent computing develops, it will benefit society in many ways. Citing one example, Zhao said, “The intelligent computing digital reactor initiated by Zhejiang Lab will become a highly efficient innovation engine and a means to serve scientific and technological ingenuity and social development.”

The partnership between the Zhejiang Lab and *Science*/AAAS reflects Zhao’s last point: “The Innovation Forum on Intelligent Computing will become a significant catalyst for academic exchange.” He expressed keen interest in continuing to shape the forum into “an annual flagship event in computing science and technology, and a high-level, open international platform for academic exchange.”

“Our goal is to bring [scientists] together to confront the huge challenges we face as a society, and explore the limits of science, thereby setting and leading the trends of intelligent computing in the future.”

— Shiqiang Zhu

Sponsored by



之江实验室  
ZHEJIANG LAB



# You heard the message.

We've told you before that NEB® offers a broad portfolio of reagents for purification, quantitation, detection, synthesis and manipulation of RNA. But did you know that these products are available from bench-scale to commercial-scale to enable both academic and industrial needs? Further, we provide these products at quality levels that support vaccine and diagnostic manufacturing. Experience improved performance and increased yields, enabled by our expertise in enzymology.



**RNA purification:** Extract up to 100 µg of high quality, total RNA from a variety of sample types with the Monarch® Total RNA Miniprep Kit. Monarch RNA Cleanup Kits

can quickly and easily clean up and concentrate RNA in just minutes, with no carryover contamination.



**RNA detection:** Optimize your RT-qPCR across a variety of sample types with Luna®. High-concentration mixes and kits optimized for multiplexing enable sensitive detection of

SARS-CoV-2. Simple, one-step solutions for LAMP and RT-LAMP are also available.



**RNA-seq:** NEBNext® kits are available for RNA library preparation, rRNA depletion and poly(A) mRNA isolation. Save time with streamlined workflows, reduced hands-on

time and automation compatibility.



**RNA synthesis:** Synthesize high-quality RNA with reagents designed to simplify your workflow, including HiScribe™ IVT kits and capping reagents. GMP-grade\* reagents are

available for mRNA synthesis of therapeutics and vaccines.

Find more details on products available, request samples, and access helpful RNA-related resources at [www.neb.com/RNA2021](http://www.neb.com/RNA2021).

\*\*GMP-grade\* is a branding term NEB uses to describe reagents manufactured at our Rowley, MA facility, where we utilize procedures and process controls to manufacture reagents in compliance with ISO 9001 and ISO 13485 quality management system standards. NEB does not manufacture or sell products known as Active Pharmaceutical Ingredients (APIs), nor do we manufacture products in compliance with all of the Current Good Manufacturing Practice regulations.

One or more of these products are covered by patents, trademarks and/or copyrights owned or controlled by New England Biolabs, Inc. For more information, please email us at [gbd@neb.com](mailto:gbd@neb.com). The use of these products may require you to obtain additional third party intellectual property rights for certain applications.

© Copyright 2021, New England Biolabs, Inc.; all rights reserved.



be INSPIRED  
drive DISCOVERY  
stay GENUINE

# CONTENTS

21 JANUARY 2022

VOLUME 375

ISSUE 6578

260

## NEWS

### IN BRIEF

**248** News at a glance

### IN DEPTH

#### **250** A controversial train heads for the Maya forest

Critics fear Mexico's presidential megaproject could threaten ecology, archaeology  
By R. Pérez Ortega and I. Gutiérrez Jaber

#### **252** EU grants restrict U.K. and Swiss research

Rules on first grants from Horizon Europe highlight ongoing diplomatic disputes  
By G. Guglielmi

#### **253** Studies reveal dangers of SARS-CoV-2 infection in pregnancy

Vaccination helps prevent stillbirths, critical care  
By M. Wadman

#### **254** Mars rover detects carbon signature that hints at past life source

Dramatically "light" carbon could also be explained by atmospheric reactions or cosmic dust  
By P. Voosen

#### **255** China falls silent about its recruitment efforts

Information about "talent programs" that drew U.S. scrutiny is no longer available  
By D. Normile

### FEATURES

#### **256** Window of opportunity

When surgery to treat neurological conditions lets researchers peer into the brain, ethical questions abound  
By K. Servick

PODCAST

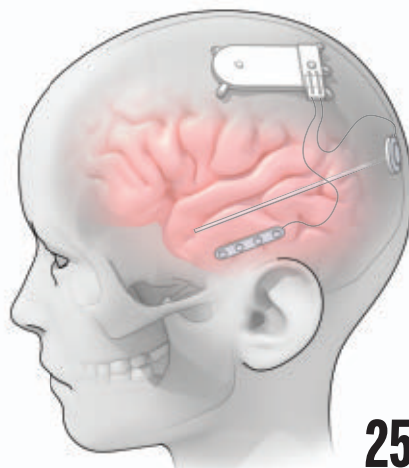
## INSIGHTS

### PERSPECTIVES

#### **260** Kids attend to saliva sharing to infer social relationships

Saliva sharing suggests "thick," intimate bonds  
By C. Fawcett

REPORT p. 311



256

#### **262** One cell, many fates

A synthetic gene circuit enables programming of many stable states in mammalian cells  
By C. Kunze and A. S. Khalil

RESEARCH ARTICLE p. 284

#### **263** Many-particle electron states in graphene

Scanning tunneling microscopy probes ground state competition in a magnetic field  
By M. Morgenstern and M. Goerbig

REPORT p. 321

#### **264** Epstein-Barr virus and multiple sclerosis

Infection with Epstein-Barr virus is the trigger for the development of multiple sclerosis

By W. H. Robinson and L. Steinman

REPORT p. 296

#### **266** The global nitrogen-phosphorus imbalance

The imbalance has grave consequences for natural ecosystems and global food security  
By J. Peñuelas and J. Sardans

#### **267** Nervous system consequences of COVID-19

Neurological symptoms highlight the need to understand pathophysiological mechanisms  
By S. Spudis and A. Nath

### POLICY FORUM

#### **270** How privacy's past may shape its future

An account of privacy's evolutionary roots may hold lessons for policies in the digital age  
By A. Acquisti et al.



## BOOKS ET AL.

**273 Biodiversity, food, and culture**

A loss to one is a loss to all  
By *L. Newman*

**274 Lost and found**

A lyrical meditation on wayfinding offers cultural context and hope for the navigationally challenged  
By *M. Bécu and C. F. Doeller*

## LETTERS

**275 Australia's biodiversity crisis and opportunity**

By *E. G. Ritchie*

**275 Protect the Amazon's Indigenous lands**

By *G. Mataveli and G. de Oliveira*

**276 Mining and Brazil's Indigenous peoples**

By *L. Ferrante and P. M. Fearnside*

**276 Technical Comment abstracts**

## RESEARCH

## IN BRIEF

**278** From *Science* and other journals

## REVIEW

**281 Sexual selection**

Sexual selection and the ascent of women: Mate choice research since Darwin  
*G. G. Rosenthal and M. J. Ryan*  
REVIEW SUMMARY; FOR FULL TEXT:  
DOI.ORG/10.1126/SCIENCE.AB16308

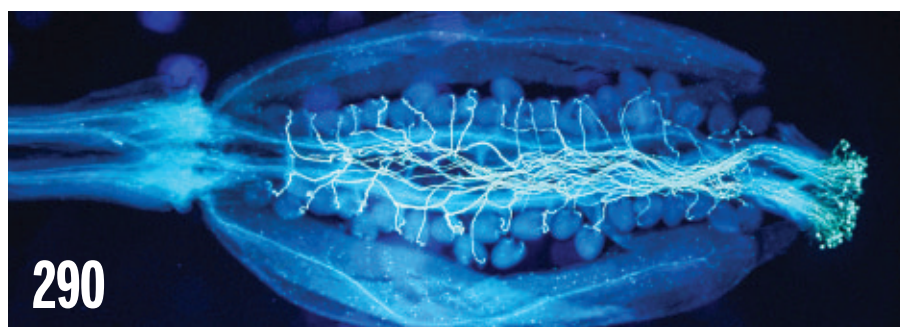
## RESEARCH ARTICLES

**282 Coronavirus**

Structural basis for continued antibody evasion by the SARS-CoV-2 receptor binding domain  
*K. G. Nabel et al.*  
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:  
DOI.ORG/10.1126/SCIENCE.ABL6251

**283 Protein design**

Reconfigurable asymmetric protein assemblies through implicit negative design  
*D. D. Sahtoe et al.*  
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:  
DOI.ORG/10.1126/SCIENCE.ABJ7662

**284 Synthetic biology**

Synthetic multistability in mammalian cells  
*R. Zhu et al.*  
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:  
DOI.ORG/10.1126/SCIENCE.ABG9765  
PERSPECTIVE p. 262

**285 Attosecond science**

Attosecond coherent electron motion in Auger-Meitner decay  
*S. Li et al.*

**290 Plant science**

RALF peptide signaling controls the polytubule block in *Arabidopsis*  
*S. Zhong et al.*

## REPORTS

**296 Multiple sclerosis**

Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis  
*K. Bjørnevik et al.*  
PERSPECTIVE p. 264

**302 Solar cells**

Conformal quantum dot-SnO<sub>2</sub> layers as electron transporters for efficient perovskite solar cells  
*M. Kim et al.*

**307 Perovskites**

Three-dimensional direct lithography of stable perovskite nanocrystals in glass  
*K. Sun et al.*

**311 Social psychology**

Early concepts of intimacy: Young humans use saliva sharing to infer close relationships  
*A. J. Thomas et al.*  
PERSPECTIVE p. 260; PODCAST

**315 Research technology**

High-speed fluorescence image-enabled cell sorting  
*D. Schraivogel et al.*

**321 2D materials**

Visualizing broken symmetry and topological defects in a quantum Hall ferromagnet  
*X. Liu et al.*  
PERSPECTIVE p. 263

**326 Molecular motors**

Structural and functional insight into regulation of kinesin-1 by microtubule-associated protein MAP7  
*L. S. Ferro et al.*

**331 Coronavirus**

SARS-CoV-2 vaccine protection and deaths among US veterans during 2021  
*B. A. Cohn et al.*

**336 Coral reefs**

Protecting connectivity promotes successful biodiversity and fisheries conservation  
*L. Fontoura et al.*

## DEPARTMENTS

**245 Editorial**

It's not too late  
By *Eric Topol*

**247 Editorial**

Strengthening scientific integrity  
By *Alondra Nelson and Jane Lubchenco*

**350 Working Life**

A mentor's journey  
By *Georgia Foustier*

## ON THE COVER

A target cell expressing a fluorescently tagged protein (green) is illuminated by a blue laser and selected from a swirling pool of cells. New cell sorting technology combines the throughput and quantitative power of flow cytometry with the spatial resolution of multicolor fluorescence microscopy to enable isolation of cells with complex phenotypes at speeds up to 15,000 cells per second. See page 315.



Illustration: Tobias Wüstefeld/Illustratoren.de

Science Staff .....	246
New Products .....	347
Science Careers .....	348

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.

# It's not too late

**A**s the Biden administration took office last January, with the pandemic peaking at more than 130,000 COVID-19 hospitalizations in the United States, there were high hopes for a new plan of “sticking to the science” and expectations that public health policies, communication, and trust would return to levels not seen for many years. That didn't happen. Why?

For one, managing a public health crisis requires a navigational system based as closely as possible on real-time capture, analytics, and sharing of essential and relevant information. But during the Omicron wave, we haven't a clue as to whether and to what extent hospitalizations are being prevented by vaccinations and boosters. The president and leading members of public health agencies have made numerous statements and public appearances throughout the pandemic, but not a single word has been uttered about fixing the profoundly deficient data systems. In May 2021, the Centers for Disease Control and Prevention (CDC) declared that there would be close tracking of all hospitalizations and deaths among the vaccinated. That never happened, though not because of the CDC.

The Department of Health and Human Services (HHS), led by Secretary Xavier Becerra, has chosen to avoid mandating that data be collected. Is this unwillingness a reaction to fear of political backlash? Amid an ongoing public health crisis, this inaction is unacceptable.

Mr. Becerra has also shirked other responsibilities. There has been obvious in-fighting between the Food and Drug Administration (FDA), the National Institutes of Health (NIH), the CDC, and the White House COVID-19 Response Team about issues ranging from booster shot recommendations to isolation guidelines that omit any testing for the virus. A leader of HHS should be mediating such conflicts, with the goal of a single, consistent, and unified public message. That, along with data that justify government decisions, would promote clarity and trust.

The CDC, as the primary agency that provides the public with pandemic guidance, is not exempt from responsibility. In August 2021, President Biden, with support of the White House team, NIH, CDC, and FDA, announced that all adults would get booster shots 6 months after their primary vaccination. Six

months became 8 months, then 5 months—or perhaps the plan would be limited to certain vaccines or to certain age groups, or was not necessary at all. Is it any wonder that only 23% of Americans have received a booster—in contrast to more than 50% of people in many other countries, including the United Kingdom, Israel, Ireland, and Denmark? Data from Israel's Ministry of Health, available in late July and August during the Delta wave, made it abundantly clear that boosters would be necessary for preventing symptomatic infections and severe disease, but it took until late November for the CDC to announce that all vaccinated adults should get a booster shot. This inevitably led to preventable hospitalizations and deaths. Now with the Omicron wave, for which booster shots have about 90% effectiveness against hospitalization, the mistakes and delays in communication have become magnified. Millions of Americans would have been far better protected against severe disease from both Delta and Omicron had the FDA and CDC followed the data.

The recent gaffe with the truncated isolation guidance—10 days abruptly cut to 5—was of lesser severity with respect to endangering people, but the lack of any science or evidence to back up the change added to public confusion.

Simply acknowledging that despite the uncertainty, it is necessary to take some action to maintain essential workers and the health care workforce in face of the onslaught, would have helped immensely. Furthermore, recognition of best practices for isolation from other countries, using testing for two consecutive days as guidance, would have been helpful. But the unwillingness to do so has been thematic, and particularly ironic given the void of American data or, in many cases, sound practices, like masking. The world was alerted by the South African biomedical community of Omicron's extreme infectiousness in November, but there has yet to be a recommendation for improved, medical-quality masking.

With the pandemic far from over, public health agencies must take corrective action: gather and share the data, avoid any further infighting, provide consistent, unified communication, and have the Secretary of HHS be an integral part of the pandemic response. It isn't too late to get all of this on track.

—Eric Topol



**Eric Topol** is a professor of Molecular Medicine and executive vice president of Scripps Research, and founder and director of the Scripps Research Translational Institute, La Jolla, CA, USA. etopol@scripps.edu

“...data that justify government decisions would promote clarity and trust.”

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

Executive Editor Monica M. Bradford

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

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

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis INTERNATIONAL EDITOR Martin Enserink DEPUTY NEWS EDITORS Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff SR. CORRESPONDENTS Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi ASSOCIATE EDITORS Jeffrey Brinnard, Kelly Servick, Catherine Maticic NEWS REPORTERS Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman INTERN Tess Joose 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), Elisabeth 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 PHOTOGRAPHY MANAGING EDITOR William Douthitt WEB STRATEGY MANAGER Kara Estelle-Powers MULTIMEDIA MANAGING PRODUCER Joel Goldberg DESIGN EDITOR Chrystal Smith DESIGNER Christina Aycock GRAPHICS EDITOR Nirja Desai INTERACTIVE GRAPHICS EDITOR Kelly Franklin SENIOR GRAPHICS SPECIALISTS Holly Bishop, Nathalie Cary SENIOR SCIENTIFIC ILLUSTRATOR Valerie Altounian SCIENTIFIC ILLUSTRATOR Ashley Mastin SENIOR PHOTO EDITOR Emily Petersen PHOTO EDITOR Kaitlyn Dolan SOCIAL MEDIA STRATEGIST Jessica Hubbard SOCIAL MEDIA PRODUCER Sabrina Jenkins WEB DESIGNER Jennie Pajewski SENIOR PODCAST PRODUCER Sarah Crespi VIDEO PRODUCER Meagan Cantwell

## 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 ANALYST Kurt Ennis FINANCIAL ANALYST Isacco Fusi ADVERTISING SYSTEM ADMINISTRATOR Tina Burks DIGITAL/PRINT STRATEGY MANAGER Jason Hillman SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS Marcus Spiegel ASSISTANT MANAGER DIGITAL/PRINT Rebecca Doshi SENIOR CONTENT AND PUBLISHING SYSTEMS SPECIALIST Jacob Hedrick SENIOR CONTENT SPECIALISTS Steve Forrester, Lori Murphy PRODUCTION SPECIALIST Kristin Wovk 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 Dhre

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

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

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

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 Gonçalves 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 Yoshimi Toda (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David RIGHTS AND PERMISSIONS ASSOCIATE Elizabeth Sandler LICENSING ASSOCIATE Virginia Warren

### 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  
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 Claire M. Fraser

PRESIDENT Susan G. Amara

PRESIDENT-ELECT Gilda A. Barabino

TREASURER Carolyn N. Ainslie

CHIEF EXECUTIVE OFFICER

Sudip Parikh

BOARD Cynthia M. Beall

Rosina M. Bierbaum

Ann Bostrom

Janine Austin Clayton

Laura H. Greene

Kaye Husbands Fealing

Maria M. Klawe

Robert B. Millard

William D. Provine

### 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.

Sebastian Amigorena, Inst. Curie

James Analytis, UC Berkeley

Trevor Archer, NIEHS, NIH

Paola Arlotto, Harvard U.

David Awschalom, U. of Chicago

Delia Baldassarri, NYU

Nenad Ban, ETH Zürich

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

Emily Brodsky, UC Santa Cruz

Ron Brookmeyer, UCLA (\$)

Christian Büchel, UKE Hamburg

Dennis Burton, Scripps Res.

Carter Tribble Butts, UC Irvine

György Buzsáki, NYU School of Med.

Mariana Byndloss, Vanderbilt U. Med. Ctr.

Anmarie Carlton, UC Irvine

Simon Cauchemez, Inst. Pasteur

Ling-Ling Chen, SIBCB, CAS

M. Keith Chen, UCLA

Zhijian Chen, UT Southwestern Med. Ctr.

Ib Chorkendorff, Denmark TU

Amander Clark, UCLA

James J. Collins, MIT

Robert Cook-Deegan, Arizona State U.

Virginia Cornish, Columbia U.

Carolyn Coyne, Duke U.

Roberta Croce, VU Amsterdam

Ismaila Dabo, Penn State U.

Jeff L. Dangi, UNC

Chiara Daraio, Caltech

Nicolas Dauphas, U. of Chicago

Christian Davenport, U. of Michigan

Franz de Waal, Emory U.

Claude Desplan, NYU

Sandra Díaz, U. Nacional de Córdoba

Samuel Díaz-Muñoz, UC Davis

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

Andrea Encalada, U. San Francisco de Quito

Nader Engheta, UPenn

Karen Ersche, U. of Cambridge

Beate Escher, UFZ & U. of Tübingen

Barry Everitt, U. of Cambridge

Vanessa Ezenwa, U. of Georgia

Michael Feuer, GWU

Toren Finkel, U. of Pitt. Med. Ctr.

Gwynn Flowers, Simon Fraser U.

Peter Fratzl, 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 Zürich

Hua Guo, U. of New Mexico

Taejoop Ha, Johns Hopkins U.

Sharon Hammes-Schiffer, Yale U.

Wolf-Dietrich Hardt, ETH Zürich

Louise Harra, U. Coll. London

Carl-Philipp Heisenberg, IST Austria

Ykä Helariutta, U. of Cambridge

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 Zürich

Kai-Uwe Hinrichs, U. of Bremen

Deirdre Hollingsworth, U. of Oxford

Randall Hult, Rice U.

Auke Ijspeert, EPFL

Darrell Irvine, MIT

Akiko Iwasaki, Yale U.

Stephen Jackson, USGS & U. of Arizona

Erich Jarvis, Rockefeller U.

Peter Jonas, IST Austria

Matt Kaeblerlein, U. of Wash.

William Kaelin Jr., Dana-Farber Cancer Inst.

Daniel Kammen, UC Berkeley

Kisuk Kang, Seoul Nat. U.

Sabine Kastner, Princeton U.

V. Narry Kim, Seoul Nat. U.

Robert Kingston, Harvard Med.

Nancy Knowlton, Smithsonian Institution

Etienne Koehlin, École Normale Supérieure

Alex L. Kolodkin, Johns Hopkins U.

Julia Krupic, U. of Cambridge

Paul Kubas, U. of Calgary

Gabriel Langer, Scripps Res. (\$)

Mitchell A. Lazar, UPenn

Wendell Lim, UCSF

Luis Liz-Marzán, CIC bioMaGUNE

Omar Lizardo, UCLA

Jonathan Losos, Wash. U. in St. Louis

Ke Lu, Inst. of Metal Res., CAS

Christian Lüscher, U. of Geneva

Jean Lynch-Stieglitz, Georgia Inst. of Tech.

David Lyons, U. of Edinburgh

Fabienne Mackay, QIMR Berghofer

Aime Magurran, U. of St. Andrews

Asifa Majid, U. of York

Oscar Marin, King's Coll. London

Charles Marshall, UC Berkeley

Christopher Marx, U. of Idaho

David Masopust, U. of Minnesota

Geraldine Masson, CNRS

Jason Matheny, Georgetown U.

Heidi McBride, McGill U.

C. Robertson McClung, Dartmouth

Rodrigo Medellin, U. Nacional Autónoma de México

C. Jessica Metcalf, Princeton U.

Baoxia Mi, UC Berkeley

Tom Misteli, NCI, NIH

Alison Moutsering-Reif, NIEHS, NIH (\$)

Suresh Naidu, Columbia U.

Danielle Navarro, U. of New South Wales

Daniel Nettle, Newcastle U.

Beate Escher, UFZ & U. of Tübingen

Beatriz Noheida, U. of Groningen

Helga 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, U. of British Columbia Med. Ctr.

Philip Phillips, UIUC

Matthieu Piel, Inst. Curie

Kathrin Plath, UCLA

Martin Plenio, Ulm U.

Katherine Pollard, UCSF

Elvira Poloczanska, Lawrence Berkeley Nat. Lab

Julia Pongratz, Ludwig Maximilians U.

Philippe Poulin, CNRS

Jonathan Pritchard, Stanford U.

Lei Stanley, U. of Cambridge

Trevor Robbins, U. of Cambridge

Joeri Rogelj, Imperial Coll. London



# Strengthening scientific integrity

**A** robust democracy requires a common well-spring of reliable information. During his first days in office, US President Biden affirmed that evidence-based decision-making—informed by vigorous science and unimpeded by political interference—would be a pillar of his administration. He directed ambitious actions to implement that goal, including the creation of an interagency Scientific Integrity Task Force, which has just released the first-ever, comprehensive assessment of scientific integrity policy and practices in the US government.

The task force included 48 scientists, statisticians, engineers, lawyers, and policy-makers with a diversity of experiences from 29 federal agencies, and it received input from hundreds of outside experts from academia, the nonprofit sector, industry, and the public. The group found that although federal agency science is generally sound—that is, reported violations of scientific integrity policies are small in number compared to the magnitude of the federal scientific enterprise—there have been lapses that could undermine public trust in science and jeopardize federal scientists' and technologists' morale and motivation to innovate.

For example, during Hurricane Dorian's approach in 2019, then US President Trump tweeted information contradicting the official forecast of scientists at the National Oceanic and Atmospheric Administration (NOAA). Senior political leadership then directed NOAA to issue a press release supporting the president's inaccurate forecast, in effect manipulating scientific information, jeopardizing public safety, and undermining public confidence in government. In another instance, the Trump administration added a citizenship question to the 2020 census, overruling Census Bureau scientists who cited clear evidence that doing so would depress the response rate and could jeopardize census quality.

These and other violations informed the task force's recommendations, including the importance of best practices that continually reinforce a culture of integrity across the government. The report recommends the creation of a permanent interagency Scientific Integrity Council to facilitate dissemination and uptake of best practices, and communication training for scientists so that they can be more effective in explaining results to their policy superiors, to the media, and to the public.

It also emphasizes the importance of meaningful and appropriate consequences for violations.

In 2009, under US President Obama, the Office of Science and Technology Policy (OSTP) identified six principles of scientific integrity: science and technology positions in the executive branch should be filled by candidates with appropriate experience; agencies should have rules to ensure the integrity of their scientific process; research that informs agency decisions should be subject to peer review; barring restrictions, scientific or technological findings that inform policy decisions should be available to the public; agencies should address instances in which the integrity of scientific and technological processes and information may be compromised; and agencies should adopt procedures that ensure the integrity of scientific and technological processes and information used to inform decision-making.

Drawing on the 2021 task force report, the OSTP now proposes five additional principles. Because science benefits from dissent within the scientific community to sharpen ideas and thinking, scientists' ability to freely voice legitimate disagreement should not be constrained. Another principle is that scientific integrity policies should apply to all federal agencies and departments engaged in the production, analysis, use, and communication of evidence, science, and technology. Moreover, these policies must apply to political appointees, career employees,

and contractors. A further principle is grounded in the knowledge that science needs to be understood and actively considered during decision-making. Therefore, scientists should routinely participate actively in policy-making. Also, to promote accountability to the American public, federal scientists should be able to speak freely about their unclassified research, including to the press. And, accountability must be upheld. Violations of scientific integrity policies should be taken seriously and considered comparable to violations of government ethics rules.

In phase two, OSTP will work to implement these best practices and make the new principles operational. Every day, federal scientists and technologists help to tackle the greatest challenges that society faces. Let's make sure that this crucial work is supported and protected.

—Alondra Nelson and Jane Lubchenco

**Alondra Nelson**

is the deputy director for science and society at the White House Office of Science and Technology Policy. [alondra.r.nelson@ostp.eop.gov](mailto:alondra.r.nelson@ostp.eop.gov)

**Jane Lubchenco**

is the deputy director for climate and environment at the White House Office of Science and Technology Policy. [jane.a.lubchenco@ostp.eop.gov](mailto:jane.a.lubchenco@ostp.eop.gov)

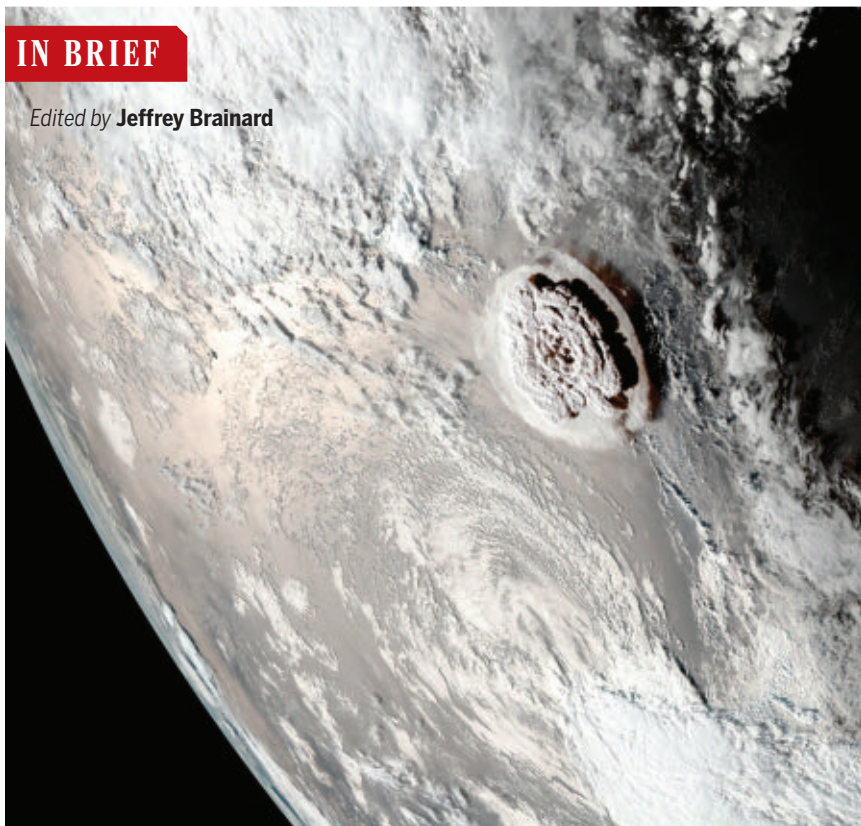
**“Violations of scientific integrity policies should be taken seriously...”**

## Bosom peril

Synonym for breast cancer, among the growing number of “tortured phrases” found by researchers in more than 3000 journal articles. The phrases may have been created by automatic paraphrasing software used to defeat plagiarism detection software—which can introduce technical errors. (*Bulletin of the Atomic Scientists*)

### IN BRIEF

Edited by Jeffrey Brainard



A volcano's massive plume spread across some 260 kilometers, as captured by the GOES-West satellite.

### VOLCANOLOGY

## Eruption in Tonga is biggest in 30 years

**A**s officials assessed the extent of death and damage caused by last week's rapid, violent volcanic eruption in Tonga, scientists began to piece together the global impact of the rare event. The explosion at Hunga Tonga-Hunga Ha'apai on 15 January caused a tsunami and likely created the highest cloud recorded in 3 decades, its ash and smoke rising 39 kilometers, according to an early estimate. The eruption, perhaps fueled by the lava's contact with cold shallow seawater, had the force of a hydrogen bomb, its rapidly rising gases creating atmospheric ripples that wrapped around the planet several times—the first time such a wave has been captured by modern satellites. Although the ash lofted into the stratosphere by the volcano could slightly cool local temperatures for the next few months, the volcano created little in the way of longer lived, light-blocking particles—such as those seen after the eruption of Mount Pinatubo in 1991, which lowered the global average temperature by about a half-degree for 2 years.

## U.S. global science lead ebbs

**INNOVATION** | The United States does not—and cannot—lead the world in science, the National Science Foundation's top advisory board told U.S. policymakers this week in a review of recent trends. This year's edition of the National Science Board's biennial *Science and Engineering Indicators*—a quantitative assessment of the global research enterprise—confirms the United States has lost its top ranking to China on several metrics in recent years, including the overall number of published papers and patents issued, and is no longer unrivaled in selected fields. “It would be the height of hubris to think we could lead in everything,” says the board's Julia Phillips, a retired physicist. “The important thing is for the United States to decide where it cannot be No. 2.” The board emphasized the need to expand the U.S. domestic scientific workforce and retain support for collaborative research across borders. “And if we continue to lead the world in basic research,” Phillips says, “then we're still in a really good position.”

## Gift funds academic software

**COMPUTING** | Schmidt Futures, a philanthropy co-founded by former Google CEO Eric Schmidt, said this week it will give universities \$40 million over the next 5 years to help their scientists obtain better software for use in their research. Academic researchers often rely on shaky, homemade computer software written by students and postdoctoral researchers, and some U.S. research grants that do pay for software engineering only cover a part-time position. The Virtual Institute for Scientific Software will fund centers at the Georgia Institute of Technology, Johns Hopkins University, the University of Cambridge, and the University of Washington to hire full-time software engineers, paying salaries high enough to compete with industry and government.

## Alzheimer's drug faces roadblock

**DRUG DEVELOPMENT** | In a highly unusual move, U.S. Medicare administrators proposed last week to sharply limit

## ANIMAL BEHAVIOR

### Arctic hares are epic hoppers

**R**abbits and their relatives tend to be homebodies, rarely straying more than a few kilometers. But at least one arctic hare traveled a record 388 kilometers. She averaged 8 kilometers per day, and one day hopped 31 kilometers. The animal was one of 25 arctic hares (*Lepus arcticus*) outfitted with satellite tracking tags on Ellesmere Island in the Canadian High Arctic, near northern Greenland. Sandra Lai, a wildlife biologist at the University of Quebec in Rimouski, and colleagues attached the transmitters in September 2019 and tracked them through the winter. Like the long-distance champion, most of the tagged hares traveled southwest for more than 100 kilometers, the researchers reported last month in *Ecology*. They suspect the hares were migrating to find food near Lake Hazen, one of the largest lakes north of the Arctic Circle, where the region's milder climate encourages plant growth.



when the government would pay for use of Biogen's controversial Alzheimer's drug, aducanumab. Marketed under the name Aduhelm, the antibody targets the abnormal buildup of the sticky brain protein amyloid. It won approval from the U.S. Food and Drug Administration (FDA) in June 2021 despite conflicting clinical trial results and skepticism among many scientists about its safety and effectiveness. On 11 January, President Joe Biden's administration proposed to cover Aduhelm's \$28,200-per-year price tag only for Medicare beneficiaries enrolled in a randomized controlled clinical trial—a restriction that would also apply to future FDA-approved anti-amyloid drugs. The policy is designed to enable researchers to collect more data on the risks and benefits of the drug. The agency expects to make a final decision in the spring.

### A boost for 'challenge' trials

**BIOMEDICINE** | Supporters of research in which people are intentionally infected with a disease to test vaccines have received a \$2 million boost to advocate for expanding the use of such "challenge trials." Open Philanthropy, an organization funded largely by Facebook co-founder Dustin Moskovitz and Cari Tuna, his spouse, this month committed the money

to the nonprofit 1Day Sooner. It was founded in April 2020 to recruit and represent volunteers for challenge trials to test vaccines against SARS-CoV-2, work that two U.K. research groups are conducting. The new funding will cover one-third of 1Day Sooner's budget for the next 2 years, as it explores challenge trials' potential in diseases such as those caused by Group A Strep, shigella, and tuberculosis bacteria, 1Day Sooner co-founder Josh Morrison says. Challenge trials can sound risky, but managers typically provide proven, conventional treatments to the volunteers if the experimental vaccine doesn't work and they become seriously ill.

### Decades on, EPA nixes pollutant

**TOXICOLOGY** | Highlighting just how slowly the wheels of science-based regulation can grind, the U.S. Environmental Protection Agency (EPA) will, for the first time in 3 decades, add a new chemical to the list of air pollutants it regulates. The decision to regulate 1-bromopropane, a solvent used by dry cleaners and auto shops, comes 20 years after researchers first raised concerns about the chemical's potential effects on human health and 12 years after groups formally petitioned the agency to act. In 1990, Congress gave EPA the power to expand a list of nearly 200 pollutants it

regulates under the Clean Air Act. But the agency had never exercised that power until this month. EPA says it will next propose how it will regulate the solvent—which could take a year to finalize.

### University of Michigan fires chief

**#METOO** | The University of Michigan (UM) last week terminated its president, immunologist and physician Mark Schlissel, for conducting a yearslong affair with a subordinate. UM's Board of Regents found he used his university email account from 2019 to 2021 to send suggestive and flirtatious messages to the co-worker—behavior the regents called "egregious" and "inconsistent with promoting the dignity and reputation" of the university. Schlissel himself had denounced such relationships in July 2021, when he announced a new university policy, according to *The Ann Arbor News*. He told the newspaper there would be zero tolerance for someone in a leadership position to "solicit a personal or romantic relationship with someone they have a supervisory authority or career influence over," adding, "That's exceptionally important because of the power dynamic. It makes it difficult sometimes for folks to effectively say no." The regents last week named former UM President Mary Sue Coleman, a chemist, as interim president.





## IN DEPTH

The Maya Train (artist's concept) will carry tourists to natural reserves and archaeological sites.

### ENVIRONMENT

# A controversial train heads for the Maya forest

Critics fear Mexico's presidential megaproject could threaten ecology, archaeology

By **Rodrigo Pérez Ortega** and  
**Inés Gutiérrez Jaber**, in Mexico City

Every day at sunset, a 3-million-bat whirlwind emerges from a cave and floods the night sky of the Calakmul Biosphere Reserve. The cave—*El Volcán de los Murciélagos* (the Bat Volcano)—hosts at least seven bat species and is a pillar of the region's ecosystem. Ecologist Rodrigo Medellín Legorreta of the National Autonomous University of Mexico (UNAM), University City, calls it “the most important bat colony of the neotropical region.” But he and other scientists say the bats and other native species could be in jeopardy as a new train makes its way through the Maya rainforest.

The Maya Train, named after the Indigenous people of the Yucatán Peninsula, is a controversial \$9.8 billion megaproject that aims to transport more than 40,000 passengers daily across 1500 kilometers of southeastern Mexico. It's backed by Mexican President Andrés Manuel López Obrador, whose administration says it will boost transportation in the Yucatán Peninsula and bring much-needed development and tourism. The United Nations predicts the train will create almost 1 million new jobs and double economic growth in the region. López Obrador has promised that, like other megaprojects, the train will be finished by the end of 2023.

But many scientists, along with some Maya activists, worry the train and accompanying development will have devastating, irreversible impacts, fragmenting the rainforest, reducing and polluting habitats, and disturbing ancient sites. “Wherever the train crosses, it will affect the archaeology and rainforest,” says Romel Rubén González Díaz, a Maya activist and coordinator at the Regional Indigenous and Popular Council of Xpujil.

An executive summary of the project's risks, released by Mexico's National Council on Science and Technology in 2019, warned of possible damage. It said the train would threaten at least 10 protected natural areas and nearly 1300 archaeological sites, and affect more than 143,000 Indigenous people living along the proposed route, with the rise in tourism potentially worsening pervasive human and drug trafficking.

The full report has still not been published, and track is already being laid. Critics hope there's still time to shape the train's second phase of 900 kilometers, which is still on the drawing board, and minimize damage.

In June 2020, amid heated debate, the National Fund for Tourism Development (Fonatur) began to build the train's first phase, which will transport locals and tourists 635 kilometers from Palenque in Chiapas state, near an ancient Maya city, to Izamal in Yucatán state, a colonial city founded over an ancient Maya settlement

(see map, p. 251). The route passes through savanna, mangroves, and rainforest that house thousands of species, many of them endangered, including jaguars, tapirs, scarlet macaws, Yucatán black howler monkeys, and many species of bats.

Fonatur's scientific advisers note that trains are more environmentally friendly than highways, and almost half the train will run along the route of existing tracks. Decisions about the train “are based on scientific evidence,” says neuroscientist Javier Velázquez Moctezuma, a Fonatur adviser.

But some researchers say the project was poorly planned and rushed. In November 2021, López Obrador issued a presidential decree categorizing the train and other megaprojects as matters of public interest and national security in a move to speed construction and overcome multiple injunctions filed to stop it. And 2 weeks ago, soil instability forced planners to shift the route of phase two. “It would've been wise for the planning to be ready many years before,” says Gerardo Ceballos González, a UNAM ecologist who has supervised some of the train's environmental mitigation plans. “But we have to work with what we have.”

Scientists and Indigenous leaders including González Díaz also contend the environmental analysis to date has been incomplete and superficial. In January 2021, more than 160 academics criticized the Environmental Impact Manifestation

(MIA), a report required by law and approved on 30 November 2020, 6 months after construction had begun. They posted a lengthy report claiming the MIA fails to acknowledge the impacts of ecosystem fracture, worsening air quality, noise, vibration, and pollution. The government has not responded to criticism, says Ana Esther Ceceña Martorella, an economist at UNAM's Institute of Economic Research. "It doesn't seem like they want to listen."

Fonatur officials say the plans include mitigation strategies, including many underpasses or overpasses allowing animals to cross tracks or highways. But Shelley Alexander, a landscape ecologist at the University of Calgary, says that although some animals adapt well to wildlife crossings, large carnivores such as jaguars have a harder time. The wildlife crossings planned for the Maya Train are "a Band-Aid on a very sensitive habitat." Still, she says: "It's important not to completely condemn the project based on the fact it might not be perfect, because that can then stop any [mitigation] from happening."

The most important mitigation strategy, Ceballos González says, is to expand and further fund already-protected areas near the train's route, such as the Calakmul Biosphere Reserve and Tulum National Park. Those measures could constrain future development and prevent further ecological damage.

As for the Bat Volcano, the train's route originally paralleled a highway only 10 meters from the cave. Bat researchers said this would put low-flying bats—already being struck by cars on the highway—at risk and disturb the cave with noise and vibration. Fonatur says the train will now run behind the cave, but official maps aren't yet clear on the exact route; scientists say the tracks should be at least 700 meters away.

Elsewhere the train runs near famous Maya archaeological sites such as Chichén Itzá and Tulum, as well as thousands of unexplored or undiscovered ruins. To document sites that may be disturbed by the train, 240 archaeologists and 1200 other workers and specialists have explored more than 800 kilometers of the route in 11 months, according to the National Institute of Anthropology and History (INAH) and Fonatur. "It has been a challenge," says Fonatur's lead archaeologist, Jesús Samuel Hernández Hernández. So far, they have recorded more than 19,000 Maya artifacts, including stelae, monuments, and house foundations, and more than 160 burials. This will yield "more than 600 papers and books for the next 10 to 15 years," Hernández Hernández predicts. He argues that by identifying unknown sites, the project will decrease looting.

But Francisco Estrada Belli, an archaeologist who studies Maya culture at Tulane University, says the train will bring "inevitable losses" to archaeology. He acknowledges the salvage work but worries that because the excavations are not led by a research question, most of what is found will remain in archives and not contribute to archaeological knowledge, as has happened in other salvage work around the world.

One archaeologist working on the excavations agrees. "It's just an immense amount of data," she says, asking to be anonymous because she is not authorized to talk to the press. "I don't think anyone is going to be able to really do true research." She says managers pressure archaeologists to work faster. When a burial is found, they log data

be accessible to the scientific community.

Archaeologists also worry about the new development the train will bring. The thousands of tourists who already visit archaeological sites in the Yucatán Peninsula have a detrimental impact on the ruins, Estrada Belli says. Their numbers are expected to skyrocket. The Maya Train is projected to attract about 3 million visitors each year to the ancient city of Calakmul, for example, compared with 40,000 each year today. It may also spur local population growth, Medellín Legorreta says. "[This] terrifies me." The Yucatán Peninsula, now home to some 5.1 million people, "cannot sustain an increase in population."

Pérez Rivas says INAH and Fonatur are working together to create a strategy to

## No stopping this train

Construction has begun on the first phase of the Maya Train, which is expected to bring a flood of tourists to the Yucatán Peninsula. Scientists worry a planned second segment will also slice across protected natural areas, fragmenting the rainforest and opening up remote archaeological sites.



quickly and seal them by the next day with little interpretation. "It's like a formality so that everything is checked off and the train construction can start."

Hernández Hernández and others insist the digs aren't rushed. "During salvage excavations there are always time pressures," says INAH archaeologist Manuel Eduardo Pérez Rivas, who leads the salvage project. "But we have the technical and academic elements to do things efficiently and carefully."

At the moment, data on artifacts and locations are kept confidential to avoid looting, Pérez Rivas says. But after sites are registered and protected, he says, the data will

manage tourism responsibly and protect the region's heritage, species, and ecosystems.

Ceballos González is hopeful the train can be built with little damage. "We have to take advantage of the opportunity the train brings to try and mitigate some of the impacts already happening," such as illegal logging and hunting, and intensive farming, he says.

But many remain skeptical, including Medellín Legorreta. "Academia and the world are waiting to see what will happen with the Maya Train." ■

Inés Gutiérrez Jaber is a journalist in Mexico City.

## EUROPE

# EU grants restrict U.K. and Swiss research

Rules on first grants from Horizon Europe highlight ongoing diplomatic disputes

By **Giorgia Guglielmi**

**F**or most of the 397 early-career scientists who learned last week that they had received €1.5 million grants, the first to be awarded under the European Union's mammoth new funding program, it was an ecstatic moment, a chance to launch a lab. But for 28 winners based in Switzerland and 46 in the United Kingdom—two countries whose attempts to join the €95.5 billion Horizon Europe funding scheme have been thwarted by broader diplomatic disputes—it has been bittersweet. To use the awards, the Swiss researchers need to leave the country, and the U.K. winners remain in limbo, uncertain whether the same condition applies to them.

"It's a loss for everybody, because sooner or later the outcomes of these projects are impacting on society," says Lidia Borrell-Damián, secretary general of Science Europe, a Brussels-based organization that represents European funding agencies.

In June 2021, the European Commission, the executive arm of the EU, excluded Switzerland from Horizon Europe after the country pulled out of talks about an overarching treaty with Europe that would replace more than 120 outdated bilateral deals on trade, immigration, and other issues. The Commission is using Swiss membership in Horizon Europe as a "bargaining chip" in the broader talks, says Thomas Jorgensen, senior policy coordinator at the European University Association. "The choices that have been made are not helpful to European science."

Switzerland-based researchers who won the early-career grants, which are awarded by the European Research Council (ERC), have been told they can only receive the funding if they move to an institution in the EU or a non-EU country that has already joined Horizon Europe, including Israel, Turkey, and Norway. "This would mean relocating myself and finding the most suitable institu-

tion in another country, and then starting from scratch," says climate historian Heli Huhtamaa at the University of Bern, who was awarded an ERC starting grant to study how past volcanic eruptions influenced climate and human societies.

The Swiss government launched a backup system for funding successful applicants who want to stay, so Huhtamaa, like most of the Switzerland-based awardees contacted by *Science*, says she will forgo the ERC award, remain in country, and collect an equivalent award from the national scheme. The national grants can't be transferred to other countries, however,

which may limit researchers' mobility.

What's more, national funding schemes aren't as competitive as their European counterpart, so the awards don't carry the same prestige, says economist Elliott Ash at ETH Zurich, who was awarded an ERC starting grant to develop artificial intelligence tools to analyze the behavior of judges. "An ERC grant opens doors; a national grant won't be able to do that," he says.

Alternate funding schemes must be an emergency measure rather than a permanent solution—or Switzerland-based researchers risk becoming isolated from their EU partners and losing valuable research ties, adds Marcel Tanner, president of the Swiss Academies of Arts and Sciences. An earlier ban from EU research funding in 2014—a response to Swiss restrictions on immigration—led to a sharp decrease in the number of international collaborations led by Swiss universities, according to a recent study. They recovered only after Switzerland regained full access to European research funding in 2017.

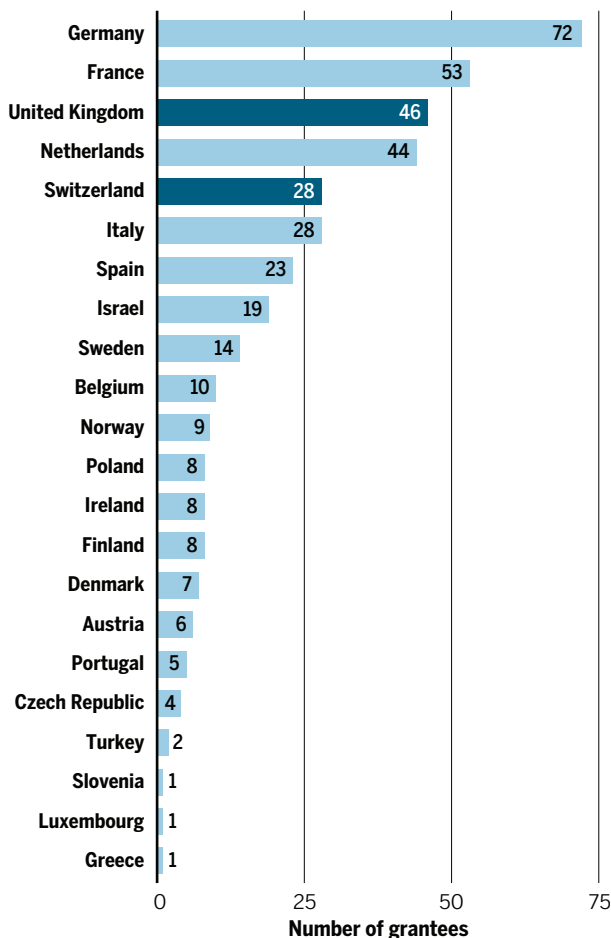
A spokesperson for the Swiss government says joining Horizon Europe remains Switzerland's "declared goal." But Stefanie Walter, an expert in political science and international relations at the University of Zurich, says the issues at stake in the treaty negotiations are thornier than before. She says the impasse could last for years, which will be "really bad" for science.

U.K.-led international projects also suffered a severe blow after the nation's 2016 vote to leave the EU, the same study found. Qualms about Brexit may have reduced the appeal of engaging in partnerships with U.K. researchers, says Benedetto Lepori, an expert in higher education at the University of Lugano, who led the research. "Uncertainty creates a lot of problems," he says.

Now, 2 years after Brexit was formally executed, ongoing diplomatic wrangles are creating a new barrier. Although the United Kingdom reached a deal in December 2020

## Left in limbo

On 10 January, the European Research Council announced €619 million in grants to 397 early-career researchers across 22 countries. But 74 grantees based in the United Kingdom and Switzerland are restricted from using the money because of broader diplomatic disputes.





to join Horizon Europe (*Science*, 8 January 2021, p. 110), the Commission has delayed approval over a separate dispute about border issues with Northern Ireland. ERC warned U.K.-based starting-grant winners that, like the Switzerland-based awardees, they may have to move elsewhere to receive their award if the Horizon Europe deal fails.

In that case, the government would provide backup funding for ERC starting-grant winners who stay in the country, as well as a wider alternative funding scheme—a “bold and ambitious offer that delivers many of the benefits of Horizon association,” U.K. science minister George Freeman announced in November 2021.

For now, U.K.-based awardees, who are expected to sign their grant contracts by April, don't know whether their funding will come from ERC or from the U.K. government. Seven U.K.-based winners contacted by *Science* said they don't plan to relocate their lab to the EU in the short term, but some said they may move in the future if the uncertainty drags on. “If an institution told me, ‘You can keep this funding if you come to Europe,’ I would very seriously start considering that possibility,” says Marcelo Lozada-Hidalgo, a physicist at the University of Manchester whose ERC starting grant is to develop nanoscale sieves to separate molecules.

Many U.K. researchers are frustrated: They must wait while other ERC winners can begin to collect their money and get started buying equipment and hiring graduate students. “It feels like our careers and our ability to do the kind of research that we're hoping to do is at the mercy of the post-Brexit climate, and it's not something that we can have any control over,” says David Doyle, a social psychologist at the University of Exeter who has an ERC grant to study the psychosocial outcomes of people undergoing hormone therapy to match their gender identity.

The national funding scheme that Freeman promised is small comfort, some say. The idea that it could match the collaborative opportunities of Horizon Europe is “pie-in-the-sky nonsense,” says James Wilsdon, a science policy researcher at the University of Sheffield. “You can't replicate those unilaterally as a single nation,” he says.

Kieron Flanagan, who studies science and technology policy at Manchester, is optimistic for a breakthrough on the U.K.-European impasse on research, but says it may take years. “The question is, how much damage gets done in the meantime?” ■

Giorgia Guglielmi is a journalist in Basel, Switzerland, who does part-time communications work for the Friedrich Miescher Institute for Biomedical Research.

## COVID-19

# Studies reveal dangers of SARS-CoV-2 infection in pregnancy

## Vaccination helps prevent stillbirths, critical care

By Meredith Wadman

**T**wo studies last week delivered sobering news about the dangers of COVID-19 to unvaccinated pregnant women and their babies. Perhaps the most disturbing data came from a first-of-its-kind analysis that tracked all the tens of thousands of recent pregnancies in Scotland. It found that unvaccinated women who became infected with the coronavirus during pregnancy were far more likely than the general pregnant population to have a stillborn infant or one that dies in the first month of life. All of the infected women who lost their fetuses or infants were unvaccinated.

The unvaccinated mothers themselves were also more endangered: Among 104 pregnant women with a SARS-CoV-2 infection who required critical care, 102 were unvaccinated. Unvaccinated women also had a far higher rate of hospitalization than their vaccinated counterparts in the study of nearly 88,000 pregnant women.

Yet the study found that in October 2021, months after COVID-19 vaccines became widely available in the United Kingdom, fewer than one-third of pregnant Scots delivering babies had been fully vaccinated. Given that more than 77% of adult women of childbearing age in the general Scottish population were vaccinated by then, the result highlights a large vaccine hesitancy among the pregnant that is mirrored in many places. “This should shake us up and really be a call to action,” says Yalda Afshar, a high-risk obstetrician at Ronald Reagan UCLA Medical Center. “Vaccination is the clear action item to improve health for pregnant people and their babies.”

Using data from an ongoing Scottish population study, researchers at the University of Edinburgh and colleagues elsewhere tracked pregnant women between December 2020 and October 2021. Although the risk of poor outcomes was generally elevated for unvaccinated pregnant women who got COVID-19 at any point during gestation, it was starkly higher for those infected late in pregnancy, the group reports in *Nature Medicine*. In the 620 mothers who contracted COVID-19 in the 28 days before they delivered their babies,

the study recorded 14 fetal or infant deaths, 10 of them stillbirths. That amounts to 22.5 deaths per 1000 births, compared with 5.6 among all Scottish pregnancies from March 2020 through October 2021.

The Scottish pregnancy study also highlighted risks to the unvaccinated women's own health: Ninety-eight percent of critical care admissions for study participants with SARS-CoV-2 infection and 91% of hospitalizations were in unvaccinated women. A pregnant woman critically ill with COVID-19 “should be an anomaly [and] not a daily occurrence,” says study first author Sarah Stock, a maternal and fetal medicine specialist at the University of Edinburgh.

Many pregnant women have been reluctant to get COVID-19 vaccines, with some wary of possible harm to their babies. Stock

**“This should shake us up ...”**

Yalda Afshar,  
Ronald Reagan UCLA  
Medical Center

and colleagues examined birth outcomes in the more than 18,000 Scots who were vaccinated during their pregnancies. Vaccination during pregnancy, including receiving a shot within 28 days of giving birth, did not increase preterm births or deaths of infants in the weeks before and after birth.

That finding “is really important” and confirms other recent studies, says Sarah Mulkey, who studies congenital viral infections at Children's National Hospital.

The second study, published in *The Lancet Digital Health*, examined electronic health records from more than 18,000 pregnant women in five U.S. states who were tested for COVID-19 between March 2020 and February 2021. The analysis matched 882 unvaccinated women who had a confirmed infection and mild to moderate symptoms with other pregnant women who tested negative, finding that infected women were significantly more likely to have preterm births or stillborn infants.

The timing of an infection, but not symptom severity, matters. The earlier in pregnancy a mother was infected with SARS-CoV-2, the earlier a baby was likely to be born, reported study first author Samantha Piekos of the Institute for Systems Biology and her colleagues. “Even mild COVID-19 infections put pregnant people at increased risk for preterm delivery,” Piekos says. ■



Martian rock samples drilled from the Vera Rubin ridge (tiny hole, lower left) were enriched in carbon-12.

## PLANETARY SCIENCE

# Mars rover detects carbon signature that hints at past life

Dramatically “light” carbon could also be explained by atmospheric reactions or cosmic dust

By **Paul Voosen**

**S**ince 2012, NASA’s Curiosity rover has trundled across Mars, drilling into rocks and running the grit through a sophisticated onboard chemistry lab, aiming to tease out evidence for life. This week, a team of rover scientists announced an intriguing signal, one that may or may not be evidence of past life, but is, at the very least, surprisingly weird. The team found that the carbon trapped in a handful of rocks probed by the rover is dramatically enriched in light isotopes of carbon. On Earth, the signal would be seen as strong evidence for ancient microbial life.

Given that this is Mars, however, the researchers avoid making any grand claims, and they have concocted nonbiological explanations involving ultraviolet (UV) light and stardust. But those alternatives are at least as far-fetched as a scenario in which microbes emitted the enriched carbon as methane gas. And so the study does “inch up the plausibility” that life once existed on the planet—and could still today, says Christopher House, a biogeochemist at Pennsylvania State University, University Park, who led the study, which was published this week in the *Proceedings of the National Academy of Sciences*.

Mark Harrison, a planetary scientist at the University of California, Los Angeles, who is not a part of the rover team, says the enriched carbon is a strong hint of ancient life. But he says the authors are “appropriately conservative,” and notes that such signatures are debated even on Earth.

The study relies on a time-honored insight: Life is lazy. Carbon exists in two stable isotopic forms: Most is “light” carbon-12, and a small fraction is carbon-13, weighed down by an extra neutron. Because of this extra neutron, carbon-13 tends to make molecules with slightly tougher bonds. As a result, life favors the easier to divide carbon-12, and most biological organic molecules are enriched in carbon-12. Methane from rice paddies, for example, is more enriched in light carbon than nonbiological methane from hydrothermal seafloor vents.

The team looked at rock samples drilled during Curiosity’s journey across Gale crater. An oven in the rover’s belly converted trace amounts of carbon in the rock into methane gas. A laser then probed the gas and revealed the methane’s isotopic makeup. For rocks from six sites, the ratio of carbon-12 to carbon-13 was much higher than an Earth-based reference standard. “These are dramatic signals,” House says.

Concentrating light carbon to such high levels might have taken multiple steps. The researchers envision methane emissions from deep subsurface microbes feeding on the slightly light carbon found in magma. (The martian atmosphere, depleted in light carbon, is unlikely to have provided the carbon.) Then, other microbes at the surface would feed on the methane, further enriching the light carbon and fixing it in the fossil record when they died.

Still, the rover has seen no physical traces of ancient microbes, so the researchers say it’s also possible deep microbes started the enrichment, but UV light drove it the rest of the way. The UV light might have broken apart the microbial methane, further enriching its light carbon while creating daughter products like formaldehyde that would eventually settle on the surface.

Or perhaps the young Solar System, including early Mars, passed through an interstellar cloud of gas and dust, which is believed to happen every 100 million years or so. The carbon in such dust is light, matching the levels seen by Curiosity, to judge by samples trapped in meteorites. The cloud might have blocked sunlight and plunged Mars into a deep freeze, causing widespread glaciation and preventing other carbon sources from diluting the light carbon in the rain of cosmic dust. House concedes that the scenario requires an incredible coincidence of events, and there’s no evidence of glaciation at Gale crater. But he says it can’t be ruled out.

More prosaically, a few studies suggest UV rays can generate the signal without help from biology at all, by breaking down atmospheric carbon dioxide to produce carbon monoxide that is enriched in carbon-12. Yuichiro Ueno, a planetary scientist at the Tokyo Institute of Technology, says he has recently confirmed the process can occur in unpublished lab results. “The reported carbon isotope ratios are exactly what I have expected,” he says. Ueno says early Mars may have had a different atmosphere, perhaps rich in hydrogen that would have reacted with the carbon monoxide to form a host of organic molecules. Those would eventually fall out of the air, depositing the signature Curiosity detected.

All these scenarios would play out in the ancient past. But Curiosity also sniffs for carbon in today’s martian air. It has detected methane, but at levels far too low to measure carbon isotopes. (Confoundingly, sensitive instruments in orbit see no methane.) Should light carbon ever be detected in a thicker plume of methane, it would open an even more exciting possibility, House says. “Even though we’re looking at a potentially ancient process, the methane today could be from the same biosphere sustained till now.” ■

## RESEARCH POLICY

# China falls silent about its recruitment efforts

Information about “talent programs” that drew U.S. scrutiny is no longer available

By **Dennis Normile**

**T**he criminal charges against Harvard University chemist Charles Lieber—and dozens of others ensnared in the U.S. Department of Justice’s China Initiative—have put a spotlight on the Thousand Talents Program (TTP), a Chinese government effort that brought Lieber and other scientists from overseas to China’s universities and research institutes. U.S. authorities have portrayed the program as an effort to pilfer know-how and innovation, a claim many scientists dispute. But as the scrutiny of the TTP grew, the program slipped out of sight.

Official mentions of the TTP have disappeared, and lists of TTP awardees once posted on government and university websites are no longer available. But experts say the TTP has simply been folded into other programs, and recruitment is continuing. More than ever, the effort focuses on scientists of Chinese origin, and part-time appointments of the type that Lieber had have become rare.

China launched the TTP in 2008, aiming to boost the country’s research output and quality. At the time, more than 90% of Chinese who earned Ph.D.s in the United States remained there for at least 5 years after completing their studies, according to a May 2020 report by David Zweig and Siqin Kang of the Hong Kong University of Science and Technology. The TTP offered returnees—and foreign researchers willing to relocate—competitive salaries and funding to establish labs. Although some half-time appointments were allowed, the program aimed for full-time researchers.

There were few takers. So in 2010 the part-time option was expanded, allowing recruits to maintain their jobs overseas if they spent at least part of the year in China. In 2011, close to 75% of 500 TTP scholars Zweig and Kang identified were on part-time agreements. (A 2019 U.S. Senate report claims the TTP had attracted more than 7000 “high-end professionals” by 2017 but didn’t specify how many were part time.)

The program has paid off for China. A 2020 study by Cong Cao, a China sci-

ence policy specialist at the University of Nottingham’s campus in Ningbo, China, showed scholars in China with overseas experience published more papers, and with higher impact, than stay-at-home peers. Universities also benefited from the association with star scientists. Lieber’s presence, for example, may have helped the little-known Wuhan University of Technology (WUT) attract prospective students, says Futao Huang, a higher education scholar at Hiroshima University.

But part-time options like Lieber’s also facilitated “double dipping,” Zweig says, where researchers with full-time posts abroad were also getting handsomely paid for time supposedly spent in China. Lieber’s contract, for example, called for him to

told *Science* in 2020. “The TTP recruited people to build up academic programs, not to steal ideas,” says Jay Siegel, a U.S. chemist who left the University of Zurich in 2013 to head a new pharmacy program at Tianjin University with TTP support. Of 23 academics targeted under the China Initiative, only two have been charged with intellectual property theft (*Science*, 10 December 2021, p. 1306). Lieber was found guilty of lying to federal authorities about his Chinese ties and failing to report the resulting income.

China has responded to the criticism as it often does: by becoming increasingly secretive. Information on the talents programs “seemed to start disappearing around the time that the China Initiative was launched” in 2018, says Emily Weinstein, an analyst at Georgetown University’s Center for Security and Emerging Technology (CSET). In 2019, the TTP and its spinoffs were absorbed into a High-End Foreign Expert Recruitment Plan, one of 27 currently active national plans, according to CSET, which gleans the information from fleeting mentions on Chinese websites. (Ministries and agencies have their own specialized programs.) “No relevant statistics” are publicly available about recruiting success, says Lu Miao, a policy analyst at the Center for China and Globalization, a Beijing think tank.

Still, the continued existence of the programs “indicates their usefulness to the country,” Cao says. Although most programs are open to non-Chinese, the number moving to China “is probably still insignificant,” he adds.

Siegel, now a Switzerland-based educational consultant, says China’s talent programs have gotten so much bad publicity that U.S. universities “have become reluctant to work with anyone who has any connection to TTP.” Doing so may become illegal as well: The U.S. Congress is considering legislation prohibiting federally funded researchers from participating in China’s talent programs. Siegel and many others think such a step would be misguided. Participation by Americans “brought a lot of U.S. influence into China and Chinese understanding back to the U.S.,” Siegel says. ■



work “at or for” WUT “not less than nine months a year,” according to the indictment against him, in return for a monthly fee of up to \$50,000 and \$1.7 million to set up a lab at WUT. Some Chinese academics complained that nonresident scientists got big salaries and research support for little in return. In 2017, the government clarified that part-timers were to be in China “for no less than 2 months a year,” Huang says.

U.S. authorities took a dim view of the deals for different reasons. “China pays scientists at American universities to secretly bring our knowledge and innovation back to China,” then-FBI Director Christopher Wray said in a July 2020 speech at the Hudson Institute in Washington, D.C.

Such claims are “simply wrong and false,” Yigong Shi, a molecular biologist who left Princeton University in 2008 to head the life sciences department at Tsinghua University,





An awake patient moves her fingers during surgery to remove a brain tumor.

# WINDOW OF OPPORTUNITY

When surgery to treat neurological conditions lets researchers peer into the brain, ethical questions abound

By **Kelly Servick**

In 2019, Kate Folladori spent a month sitting in a hospital room hoping she'd have a seizure. Since her diagnosis with epilepsy nearly 20 years earlier, a series of medications had failed to bring relief. Now, a team at Baylor St. Luke's Medical Center had placed wire electrodes into her brain to record neural activity. The doctors hoped to learn where her seizures originated—and whether she might be a candidate for tissue-removing surgery or a brain stimulation implant to suppress them.

As the weeks wore on, Folladori got restless. Time became warped by boredom, and her surroundings felt surreal. “One moment that I remember specifically was it was raining outside ... and it felt to me like I was watching a television show where it was raining.”

Breaking the monotony were visits from a group of neuroscientists who recorded activity in Folladori's brain while she did simple tasks. She might press a button when a cue appeared on a computer screen or watch short videos intended to evoke different moods. The studies weren't aimed at helping Folladori or even at treating epilepsy; they addressed more basic questions about vision and emotion in the brain. But for Folladori, they were a rare bright spot. “[Having] people from the outside to make you laugh, and to give you something to do, and to give you a goal—that was everything to me,” she says.

Folladori, in turn, offered something rare and valuable to the research team, led by her neurosurgeon, Sameer Sheth of Baylor College of Medicine. The intimate view of brain activity the scientists gleaned from those tests is impossible without invasive surgery, which would be unethical to perform solely for research's sake.

People who take part in these intracranial studies—often during epilepsy monitoring or brain surgery performed when the patient is awake—“are giving an invaluable gift,” says Khara Ramos, former director of the neuroethics program at the National Institutes of Health (NIH) who is now at the Dana Foundation. Noninvasive methods of studying brain function such as functional magnetic resonance imaging and electroencephalography can “give you

good spatial resolution or good temporal resolution, but not both,” she says. But a fine wire placed in contact with brain tissue can detect the activity of neurons with precision on the scale of millimeters and milliseconds. And researchers can relate that activity with a person's real-time report of the experience.

“We can essentially gain access to the very basic neural mechanism of the human condition,” says Itzhak Fried, a neurosurgeon at the University of California (UC), Los Angeles.

Thanks partly to the rise of invasive brain stimulation treatments for diseases such as

“There is a really unique situation of vulnerability that patients are in,” Chiong says.

He and others have raised questions about how to verify that patients' participation really is voluntary, how to make clear to participants that the research is separate from clinical care, and how to ensure that researchers' desire to collect useful data doesn't compromise or interfere with that care.

Those concerns have motivated one group of researchers to develop a set of ethical commitments to guide studies in the field, published this week in *Neuron*. “I've been heartened by the conscientiousness of the neurosurgical community that we have,” Chiong says, “but there's certainly opportunities for abuse.”



To try to pinpoint the source of her seizures, Kate Folladori spent weeks with electrodes in her brain, which allowed her to participate in research.

**PIGGYBACKING** on a surgery to explore basic brain function isn't new. Starting in the 1930s, Canadian neurosurgeon Wilder Penfield treated patients for epilepsy by removing small regions of the brain. During the operation, he also explored their exposed brains, stimulating the tissue with an electrical probe and asking the patients, who were awake, what they experienced. Such experiments led to the famous homunculus: a map of which brain regions represent various body parts.

In the past 20 years, researchers have benefited from the rise of other skull-penetrating medical treatments. Those include Folladori's seizure-monitoring electrodes and implanted devices that deliver electrical stimulation to stop seizures, treat severe obsessive-compulsive disorder, and control symptoms of movement disorders such as Parkinson's disease. Implanted stimulation devices are also being studied for other conditions, including post-traumatic stress disorder and depression.

Awake surgeries to insert such devices or resect tumors can sometimes be paused briefly for an unrelated experiment. Fried estimates roughly 30 groups in North America now do intracranial human neuroscience in epilepsy surgery patients—up from fewer than 10 when he started in the field, about 20 years ago.

Researchers can also tap into therapeutic devices that stay in the brain long-term, some of which both deliver electrical stimulation and read out neural activity. Such implants are still underused sources of neural data, says UC Los Angeles (UCLA)

Parkinson's and epilepsy and to a recent U.S. federal funding program, intracranial human neuroscience is burgeoning. “There has been a significant expansion of groups that are capable of doing this work,” says Winston Chiong, a neurologist and ethicist at UC San Francisco.

But the research opportunities that come with intimate access to people's brains also raise complex ethical issues. Basic science studies tacked onto medical procedures typically offer no clinical benefit to participants. People are often recruited into a study as they prepare for serious surgery, sometimes by an investigator who is also their surgeon.



neuroscientist Nanthia Suthana, who has used their recordings to study learning, memory, and spatial navigation. Another rare opportunity comes from people with paralysis or limb loss. Some of these patients agree to have neural recording devices implanted for research studies that may lead to new brain-computer interface approaches to restore lost movement or communication.

Intracranial research faces a unique set of constraints. For one, researchers typically can't record from any brain region they want. "We adjust our question to where the electrodes are," Fried says.

Because a brain region above the ears called the temporal lobe is among the most common sites of seizures, Fried and others have designed much of their research around its functions, which include memory and language processing. For example, recordings by Fried's team in epilepsy patients have revealed the underpinnings of the "memory moment"—when neurons encoding a memory activate, about 1 second before a person reports that memory coming to mind.

The precise locations of electrodes also vary between patients, making data hard to align across participants, notes Evelina Fedorenko, a neuroscientist at the Massachusetts Institute of Technology. Her team relies on intracranial recordings to study how the brain uses both general and language-specialized mechanisms to understand language. Another issue for the field, she says, is that because eligible participants are scarce, there's little incentive to conduct experiments that aim to replicate previous results rather than break new ground. "People just want to test whatever new cool hypothesis they have," Fedorenko says.

In a further challenge, many powerful research tools used in lab animals, including genetic manipulation of brain cells, are simply off limits in people. When grant applications to do human intracranial research receive review, says Jim Gnadt, a program director at the National Institute of Neurological Disorders and Stroke, "it's hard for them to compete with the critter studies because they're not as invasive, they're not as modern." So in 2017, the NIH neuroscience technology initiative, Brain Research Through Advancing Innovative Neurotechnologies, created a new program specifically to fund research opportunities offered by intracranial human recordings and to encourage interdisciplinary collaboration.

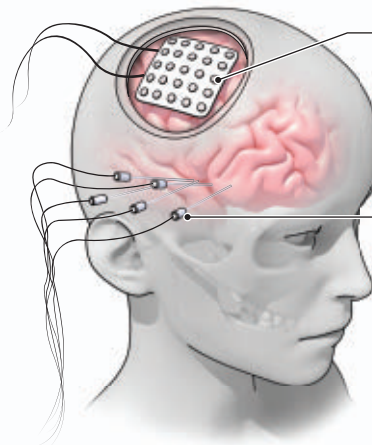
A consortium of investigators supported by the program has become a key forum for ethical discussions, and in the current *Neuron* paper, they lay out an ethical framework. Chiong, who was not involved in writing the paper, thinks other researchers will take it

## A view from beneath the skull

Scientists can run invasive studies of the human brain only in special cases. Medical devices implanted to assess or treat certain conditions offer the chance to gather additional data for research. Listening in on neurons at close range can yield basic insights into brain function.

### Epilepsy monitoring electrodes

Doctors use temporary probes to find the sources of seizures and determine whether diseased tissue can be removed or treated with a stimulation device. During a stay in an epilepsy monitoring unit, patients may participate in research studies.



### Subdural grid

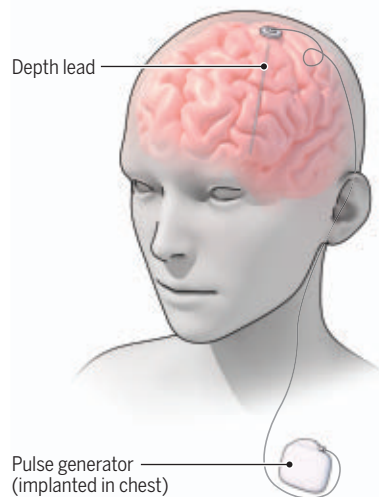
Surgeons sometimes place a thin plastic sheet of electrodes, typically slightly smaller than a credit card, on the brain's surface, under its protective membrane, the dura.

### Stereoelectroencephalography probes

A more common monitoring technology uses small, precise incisions in the skull to insert fine wires about 1 millimeter in diameter that can record from regions deep in the brain.

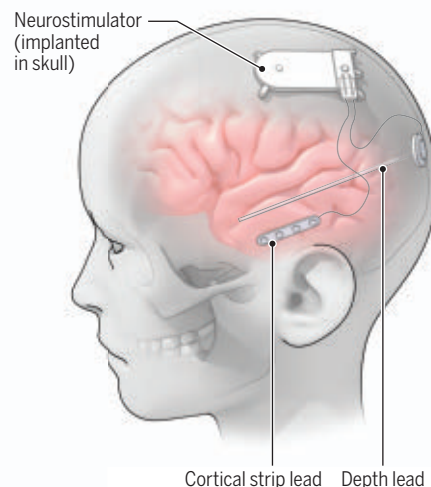
### Deep brain stimulation surgery

Stimulation deep in the brain can relieve symptoms of Parkinson's disease and other disorders. While inserting the treatment device, researchers can collect data from the device itself, fine microelectrodes that help guide its placement, or an electrode strip inserted for research.



### Responsive neurostimulation implant

A device for treating epilepsy can use different types of electrical leads to monitor brain activity and deliver stimulation that can prevent seizures. Researchers may use data downloaded from the device to study neural activity during lab experiments or daily activities.



seriously. "There's going to be a fair amount of pressure to make sure you're operating within that framework," he says. "Investigators are kind of looking around at what other people are doing and wanting to be sure that everybody's playing by the same rules."

One tenet of the new paper: Scientific considerations should not influence clinical decisions.

That guideline might sound straightforward. But for some procedures, including implanting epilepsy monitoring electrodes, multiple methods are acceptable, says Nader Pouratian, a member of the consortium and a neurosurgeon at the University of Texas Southwestern Medical Center. Surgeons use their discretion in clinical decisions that, in turn, influence

what research data can be collected.

For example, debate is ongoing in deep brain stimulation (DBS) surgery about whether patients should be under general anesthesia or awake for part of the procedure, Sheth notes. Many doctors have switched to asleep procedures for patient comfort and convenience, he says, whereas other clinicians assert that having patients responsive as surgeons determine where to place the implant can lead to better outcomes.

Unresponsive patients can't answer questions or do tasks for a research study. When asleep DBS surgery became standard at Sheth's center in 2019, he was faced with asking patients to agree to an awake surgery that was "still a very appropriate



way of doing it, but not how I usually do it.” Uncomfortable with posing that question, Sheth stopped doing research involving such patients.

**HOW TO RECRUIT** participants into those studies is itself fraught. Bioethicists have long discouraged “dual-role consent,” in which a physician who is also a study investigator invites a patient to participate. Patients may feel a sense of obligation or obedience to the physician in charge of their care, the thinking goes—and may misinterpret the study as having therapeutic benefit.

But Pouratian says some investigators in the NIH consortium asserted they were the best person to consult with patients and obtain consent because they understood both the study and the complexities of the brain surgery itself. He feels “a little conflicted” over the idea of leading the consent discussion. “They’re my patients—of course they’re going to want to consent for me,” he says. Pouratian and UCLA bioethicist Ashley Feinsinger have an NIH grant to study motivations of participants in nontherapeutic intracranial studies and their perceptions of risks and benefits. Feedback so far suggests trust in a physician or researcher plays an important role in how patients think about their participation.

Folladori can attest to that. “I really liked Dr. Sheth, and that was part of why I wanted to [participate in research],” she says. “If someone else had asked, I don’t think I would have said no, but I wonder if my feelings about it going in would have been different.”

Pouratian, Sheth, and others now use a hybrid consent process: A surgeon introduces the study and is available to answer questions, but another member of the study team not involved in the patient’s care walks through the consent documents and the signing process. The new *Neuron* paper says the consent process can vary across studies and institutions, “as long as the distinction between clinical care and research is explicit.”

Measuring and communicating risk is also challenging. Most researchers agree that asking a person to play a few computer games or answer questions in the epilepsy monitoring unit carries little risk beyond fatigue. Harder to quantify is the risk of experiments done midsurgery, which can extend a patient’s time in the operating room, typically by 20 to 30 minutes.

Very long surgeries are associated with higher rates of infection than short ones,

Sheth notes. But how much additional risk comes from extending a surgery from 3.5 hours to 4? “One could assume it’s very small,” he says, “but it may not be zero.”

Sometimes researchers temporarily place an extra strip of electrodes over the surface of the brain during a surgery to collect more data. “I’m very clear [with patients] that we’re doing something additionally that we normally would not do,”

**“I’ve been heartened by the conscientiousness of the neurosurgical community ... but there’s certainly opportunities for abuse.”**

**Winston Chiong,**  
University of California, San Francisco

Pouratian says of those situations. In a study published in April 2021 in *Neurosurgery*, he and colleagues analyzed 367 surgeries to implant DBS devices. Temporary placement of additional electrodes for research didn’t come with higher rates of complications such as bleeding in the brain, they found. Yet Pouratian still tells patients this step carries risks. “There’s a risk with everything we do, clinically speaking,” he says, even if “it doesn’t increase the overall adverse event rate.”

How well patients understand and remember what they are told about risks and other study details is another uncertainty. Neuroethicist Anna Wexler of the Perelman School of Medicine at the University of Pennsylvania and her team surveyed 22 people with Parkinson’s who had agreed to participate in research during surgery. (The study recorded brain activity during eye movements to explore how the brain makes rapid, flexible decisions.) Encouragingly, no participants had the erroneous impression that the study held direct benefits for them.

But by about 1 week after the informed-consent process, only about 23% could recall either of the two study risks communicated to them—an increased risk of infection and a potential loss of confidentiality associated with sharing their data.

Wexler notes that the patients might have better understood the risks at the time they were communicated. She adds

that little prior evidence is available for comparison, on how well patients with Parkinson’s recall information about either research or their treatment.

Still, the authors suggest future studies might explore ways to improve understanding and retention, such as a “teach-back” approach, in which participants explain details of the consent form to study staff.

Feinsinger and Pouratian are pursuing a different question: What do patients see as the value of joining these studies? At the annual meeting of the International Neuroethics Society in November 2021, Feinsinger presented feedback from interviews with 14 people between 2 months and 2 years after they took part in nontherapeutic research during implantation of DBS electrodes for a movement disorder. The conversations revealed a strong faith that basic science would pay off in future treatments for their or other brain diseases.

That was the case for Corey Westgate, who took part in studies by Suthana’s group at UCLA that relied on readouts from Westgate’s

implanted seizure-preventing device. After decades struggling with convulsive seizures, “I want this to stop,” she says, “and if me doing research can help that, then I would love to do it as much as I can.”

Suthana says the studies weren’t focused on treating epilepsy; they explored how the brain navigates through space and remembers landmarks. But the results could improve understanding of memory impairments common in epilepsy patients, she says.

Feinsinger notes that researchers need to make sure patients’ hopes are realistic. “Are we responsible for inferences patients will make about the translational likelihood of this research?” she asks. “I think to some extent, we are.”

For Folladori, a month in the epilepsy monitoring unit allowed her to participate in several studies, but she never had the seizure her doctors were waiting for. Fortunately, they used signs of abnormal activity from some of the implanted electrodes, among other clues, to find a target for a stimulation device that has kept her seizure-free for 2 years.

The experience has shaped her attitude toward research. “The reason I’m here is because of the scientific process that took place before me,” she says. “If I can participate in that in any way whatsoever, then I will absolutely do that.”

“Well,” she adds, reconsidering, “maybe not any way whatsoever.” ■

# INSIGHTS

## PERSPECTIVES

### DEVELOPMENTAL PSYCHOLOGY

## Kids attend to saliva sharing to infer social relationships

Saliva sharing suggests “thick,” intimate bonds

By **Christine Fawcett**<sup>1,2</sup>

One of the great tasks of child development is to make sense of other people's behavior by appealing to their internal thoughts, beliefs, and desires [“intuitive psychology” (1)], as well as the groups and relationships that they are a part of [“intuitive sociology” (2)]. There has long been debate over the extent to which these intuitions can be learned from social experience or whether they require some evolved capacity to interpret and categorize behavior. On page 311 of this issue, Thomas *et al.* (3) ex-

amine children's intuitions about so-called “thick” relationships—intimate bonds that people have with kin or romantic partners that are characterized by certain behaviors and obligations (4). The authors suggest that saliva sharing between individuals is a cue that young children use to infer thick relationships, and that these inferences are based on evolutionary processes that have shaped how young children interpret the social world.

Thick relationships have been described as those in which individuals' physical bodies can be seen as sharing a common “essence” (4). Thickly related individuals come

together in hugging, cuddling, and kissing, and they nurture through breastfeeding, food sharing, and other caretaking behaviors. These behaviors then become outward cues that can be used to infer the underlying thick relationship. Developmental psychology research has shown that infants use behavioral cues to infer different kinds of social relationships, rather than thick ones. For example, they expect that people who talk to and look at each other will later cooperate (5) and that people who move in synchrony (6) or have similar preferences (7) are friends. However, these previous studies have primarily examined friendship or other thin social relationships, rather than thick ones. In one notable exception, 12- to 16-month-old infants' expectations for a distressed individual to be approached by a caregiver were related to their attachment relationship with their own caregiver (8). This suggests that already in the first years of life, infants use their own experiences in relationships to make predictions about others' caregiving behaviors.

PHOTO: IGOR ALECSANDER/GETTY IMAGES



Saliva sharing, such as when kissing or sharing food, is a cue that young children use to infer "thick," intimate relationships with kin.



It is possible that learning from one's own experiences to inform expectations for others applies for the findings of Thomas *et al.* as well. In one experiment, 5- to 7-year-old children predicted that others who shared an eating utensil or licked the same ice cream cone were part of a nuclear family. Further experiments examined 8.5- to 10-month-old infants' and 16- to 18-month-old toddlers' intuitions about saliva sharing and social behaviors. When seeing someone in distress, these infants and toddlers tended to look toward a person who had previously been in a saliva-sharing interaction with the distressed individual (e.g., eating from the same orange slice), suggesting that they expected that person to respond and comfort, which would follow if they were in a thick relationship. Thomas *et al.* also collected data from a representative US sample of parents of in-

fants and toddlers that confirmed that infants are more likely to have experienced saliva-sharing interactions with those who are in thicker relationships with them, such as relatives, compared with other individuals they meet, such as friends or nannies. Finally, an ethnographic analysis suggested that around the world, saliva sharing is a cue for relationship thickness and can be used to both initiate and maintain close kin relationships.

In infants' everyday experiences, caregivers frequently share food with them, kiss them on the face, wipe away their drool with a bare hand, and so on, whereas these interactions are rare with noncaregivers or nonfamily members. But what is it about saliva-sharing behaviors that make them a reliable cue to relationship thickness? The thought of exchanging saliva with a stranger most likely elicits feelings of disgust; however, that is not the case if someone is a close intimate or associate—a romantic partner, a close friend, one's own child, or even a pet dog. It has been proposed that the emotion of disgust evolved to protect against contamination, such as can occur when an individual comes into contact with the bodily fluids of another person (9). Yet taking care of an infant, for example, requires such contact, so humans may have also evolved an exception to the rule: Those in a person's closest, thickest relationships do not elicit disgust, no matter the amount of drool or dirty diapers they produce. Indeed, research has shown that people find the smell of their own infant's diapers to be preferable to that of other infants (10).

Because dampened disgust for close kin provides an evolutionary explanation for saliva-sharing behavior occurring in thick relationships, that young children use such behavior to infer relationship thickness need not have direct evolutionary origins itself. There are numerous examples of infants relying on past experiences to inform their expectations for others' behavior, including goal-directed actions (11, 12), allocation of resources (13), and comforting (8). Thus, it is reasonable to assume that the connection between saliva sharing and thick relationships could have become part of human social structures through learning and cultural transmission. That is, by carrying out certain behaviors selectively with those closest to us, expectations and cultural traditions are established that these behaviors indicate closeness and can even be appropriated to initiate closeness, as Thomas *et al.* suggest from the results of their ethnography, citing examples of cultures where saliva sharing is used in wedding ceremonies or to welcome a new infant.

Thomas *et al.* bring questions and concepts from sociology, anthropology, and psychology together with methods from developmental and comparative psychology. Their findings reach across disciplinary boundaries and provide insight into how young children make sense of the complex social structures around them. Yet their work opens up further questions as well. For example, although their ethnography suggests that saliva sharing within thick relationships is culturally widespread, it is not known if the findings would generalize beyond the US population that was studied. It could be that variation in parenting practices across, or within, cultures leads to variation in children's expectations about thick relationships. Another critical question for future research is to clarify the mechanisms through which young children acquire the connection between saliva sharing and relationship thickness. Thomas *et al.* propose that this intuition could be a "cognitive primitive"—something that children are highly prepared to learn based on evolutionary pressures (14). However, as described above, it could rather be that evolutionary pressures are only needed to shape the behaviors that occur within thick relationships, whereas children's sensitivity to these connections is learned through their experiences.

These findings not only illuminate what young children understand about the social structures around them but also spark further questions regarding how children come to acquire these expectations and how universal they might be. ■

#### REFERENCES AND NOTES

1. H. M. Wellman, *The Child's Theory of Mind* (MIT Press, 1992).
2. L. A. Hirschfeld, *Race in the Making: Cognition, Culture, and the Child's Construction of Human Kinds* (MIT Press, 1998).
3. A. J. Thomas, B. Woo, D. Nettle, E. Spelke, R. Saxe, *Science* **375**, 311 (2022).
4. A. P. Fiske *et al.*, *Eur. J. Soc. Psychol.* **39**, 1294 (2009).
5. C. Fawcett, G. Gredebäck, *Dev. Sci.* **16**, 841 (2013).
6. C. Fawcett, B. Tunçgenç, *J. Exp. Child Psychol.* **160**, 127 (2017).
7. Z. Liberman *et al.*, *J. Exp. Psychol. Gen.* **143**, 966 (2014).
8. S. C. Johnson *et al.*, *Psychol. Sci.* **18**, 501 (2007).
9. V. Curtis, A. Biran, *Perspect. Biol. Med.* **44**, 17 (2001).
10. T. I. Case, B. M. Repacholi, R. J. Stevenson, *Evol. Hum. Behav.* **27**, 357 (2006).
11. G. Gredebäck, A. Melinder, *Cognition* **114**, 197 (2010).
12. L. van den Berg, G. Gredebäck, *Dev. Sci.* **24**, e13036 (2021).
13. T. Ziv, J. A. Sommerville, *Child Dev.* **88**, 1930 (2017).
14. L. Thomsen, S. Carey, in *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*, M. R. Banaji, S. A. Gelman, Eds. (Oxford Series in Social Cognition and Social Neuroscience, Oxford Univ. Press, 2013), pp. 17–22.

#### ACKNOWLEDGMENTS

I am grateful to T. Falck-Ytter, L. Forssman, and G. Gredebäck for their helpful comments and feedback.

<sup>1</sup>Uppsala Child and Baby Lab, Department of Psychology, Uppsala University, Uppsala, Sweden. <sup>2</sup>Department of Psychology, Stockholm University, Stockholm, Sweden. Email: christine.fawcett@psyk.uu.se



## SYNTHETIC BIOLOGY

# One cell, many fates

A synthetic gene circuit enables programming of many stable states in mammalian cells

By **Colin Kunze**<sup>1,2</sup> and **Ahmad S. Khalil**<sup>1,2,3</sup>

**T**he ability of genetically identical cells to take on diverse and specialized roles, which are maintained over long time scales, underlies critical biological processes, including multicellular development. More than 60 years ago, Waddington invoked the concept of multistability, a property of dynamical systems, to rationalize how a cell progresses from an undifferentiated state to various distinct cell fates during development (1). It has since been revealed that even within a single tissue, there is extraordinary diversity of cell states; yet how they are generated and maintained remains unclear. On page 284 of this issue, Zhu *et al.* (2) describe MultiFate, a genetic circuit design that unlocks controllable and scalable multistability in mammalian cells. They generate seven stable cell states and transition cells between states or completely destabilize a once-stable state with exquisite control. This will enable the engineering of a range of multicellular behaviors in mammalian cells.

Bistability, the simplest form of multistability, has been studied extensively in natural contexts and synthetic systems. Bistability has been shown to underlie a host of biological processes, including cell fate decisions in frog oocyte maturation (3). The study of natu-

ral bistable systems has revealed important features such as the requirement for positive feedback. However, unraveling the main ingredients of multistable systems is challenging because of the complexity of regulatory networks and deep interconnectedness with auxiliary pathways. This difficulty can be overcome by using a synthetic biology approach, in which a minimal set of non-native genetic components are introduced into cells to recapitulate complex biological functions (4). In a pioneering study, a synthetic bistable “toggle switch” was constructed in bacteria that allowed cells to flip between two stable states (5). This study and others established a blueprint for how to investigate biological functions, such as bistability, from the bottom up and inspired potential applications of these synthetic circuits (6). Recently, the synthetic circuit toolbox has grown to include tristable and quadrastable systems, promising increased functionality of engineered cells (7, 8). However, a clear procedure to expand multistability—and perhaps even exceeding the phenotypic capacity of natural systems—has remained elusive.

The design framework for MultiFate is inspired by natural transcription networks that regulate stem cell differentiation and development. These networks commonly feature positive feedback loops that involve master regulatory transcription factors (TFs) and promiscuous binding among the TFs, which have been implicated in generating multistability (9). Promiscuously dimerizing TFs are peculiar because they can have opposing cellular functions; for example, the pluripo-

tency TF octamer-binding protein 4 (OCT4) can drive either pluripotency or endodermal differentiation depending on its dimerizing partner (10). MultiFate exploits this peculiarity, using autoactivating synthetic TFs with promiscuous interactivity to ultimately generate multistability (see the figure).

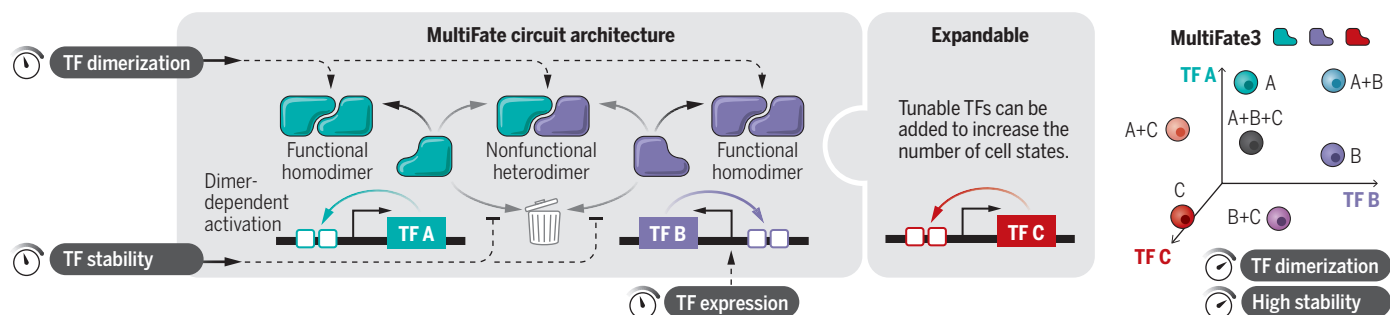
Zhu *et al.* first built MultiFate2, a circuit composed of two TFs. When activated by inducing dimerization, cells dispersed into one of three stable states, akin to inducing differentiation in a pluripotent cell. This tristable landscape is predicted by a mathematical model to transform into a bistable one with changes in TF stability (i.e., degradation rate). In reshaping the landscape, reduced TF stability generates irreversible transitions. Subsequent restoration of TF stability does not cause cells to spontaneously return to the previously destabilized state. This progression evokes comparisons to cellular differentiation and suggests how to maintain engineered multipotency in cells.

A major roadblock to implementing multistability thus far has been the increasing degree of design complexity required as more states are added. In MultiFate, increasing the number of states is straightforward: Expansion of MultiFate2 (tristability) to MultiFate3 (septastability) is achieved by adding another TF with the same design principles. With MultiFate3, Zhu *et al.* demonstrated several discrete changes in the potential landscape, reshaping septastability to hexastability to tristability with progressively diminishing levels of TF stability. Higher levels of synthetic multistability open the door to exploring these stepwise irreversible transitions. Through theoretical analysis of their system, Zhu *et al.* find that the number of stable steady states available increases rapidly with the number of TFs used, potentially attaining 256 distinct stable states with only nine TFs. With MultiFate, expandability is not hampered by complexity of circuit design but is instead restrained by the physical lim-

<sup>1</sup>Biological Design Center, Boston University, Boston, MA, USA. <sup>2</sup>Department of Biomedical Engineering, Boston University, Boston, MA, USA. <sup>3</sup>Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. Email: khalil@bu.edu

## Synthetic multistability in mammalian cells

The MultiFate circuit uses synthetic transcription factors (TFs), each with a distinct zinc finger DNA binding domain but identical dimerization domains. This allows TFs to homodimerize (active) or heterodimerize (inactive). Small molecules control dimerization, independent activation of TFs, and TF stability, allowing controlled generation of multistable states. In MultiFate3, a three-TF circuit, up to seven stable cell states can be generated under high TF stability and dimerizing conditions.



its of introducing additional highly expressed TFs. Exploration of these limits may offer clues about how organisms can generate morphological and phenotypic complexity with a relatively small number of conserved regulatory components (11).

The clever use of modular TF components with tunable molecular interactions (12, 13) allows us to ask questions about the design and control of multistable landscapes. For example, Zhu *et al.* observed an asymmetrical population distribution of cell states upon initial activation of MultiFate due to differences in TF binding and transcriptional efficiencies. This could be advantageous for an organism whose development requires different abundances of cell types. MultiFate provides a foundational tool for exploring what system parameters control those ratios and how narrow or broad a regime can be to produce the desired distribution of fates. Furthermore, MultiFate adds to a growing set of new tools to investigate the roles various cells play in developmental biology. These bottom-up approaches to cell fate circuits complement top-down, single-cell sequencing techniques that provide high-resolution maps of developmental trajectories. Comparison of synthetic and natural trajectories could clarify the stability and functional relevance of the many cell states that have been identified across tissues and organisms.

MultiFate also provides a platform for exploring how transcriptional differentiation circuits interface with other controllers of cell state, such as cell signaling. MultiFate coupled with emerging synthetic cell-cell signaling systems, such as the SynNotch receptor (14), could produce more sophisticated developmental trajectories that provide insight into natural development (15). Finally, MultiFate may enable the engineering of a general cell therapy tool that encodes many potential therapeutic states and can be guided to individually tailored fates. ■

#### REFERENCES AND NOTES

1. C. H. Waddington, *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology* (Allen and Unwin, 1957).
2. R. Zhu *et al.*, *Science* **375**, eabg9765 (2022).
3. W. Xiong, J. E. Ferrell Jr., *Nature* **426**, 460 (2003).
4. M. Elowitz, W. A. Lim, *Nature* **468**, 889 (2010).
5. T. S. Gardner *et al.*, *Nature* **403**, 339 (2000).
6. D. R. Burrill *et al.*, *Genes Dev.* **26**, 1486 (2012).
7. J. Santos-Moreno, E. Tasiudi, J. Stelling, Y. Schaerli, *Nat. Commun.* **11**, 2746 (2020).
8. F. Wu, R.-Q. Su, Y.-C. Lai, X. Wang, *eLife* **6**, e23702 (2017).
9. P. C. Facon *et al.*, *PLoS ONE* **9**, e102873 (2014).
10. S. Stefanovic *et al.*, *J. Cell Biol.* **186**, 665 (2009).
11. M. Levine, R. Tjian, *Nature* **424**, 147 (2003).
12. A. S. Khalil *et al.*, *Cell* **150**, 647 (2012).
13. C. J. Bashor *et al.*, *Science* **364**, 593 (2019).
14. L. Morsut *et al.*, *Cell* **164**, 780 (2016).
15. S. Toda, L. R. Blauch, S. K. Y. Tang, L. Morsut, W. A. Lim, *Science* **361**, 156 (2018).

#### ACKNOWLEDGMENTS

A.S.K. is a co-founder of K2 Biotechnologies and a scientific advisor for Senti Biosciences and Chroma Medicine.

10.1126/science.abn6548

#### MATERIALS SCIENCE

# Many-particle electron states in graphene

## Scanning tunneling microscopy probes ground state competition in a magnetic field

By Markus Morgenstern<sup>1</sup> and Mark Goerbig<sup>2</sup>

**T**wo-dimensional electron systems in a magnetic field provide a paradigm for unraveling the complexity of electron-electron interactions. In particular, because of the large number of possible electron states at the same energy level, such a two-dimensional system provides a rich playground for studying the different ways electrons can be arranged. Consequently, there are a plethora of many-particle ground states with similar energies to be explored. However, these ground states are difficult to distinguish from each other without direct observation of the electron arrangements (1). On page 321 of this issue, Liu *et al.* (2) report that they have deciphered the secret of one of these ground states by imaging the electron distribution with atomic resolution using scanning tunneling microscopy.

Ever since the 2004 landmark paper by Novoselov and Geim that described how to probe the quantum properties of graphene (3), the material has been widely studied for the fundamental insights it offers into electronic systems (4). Graphene, defined as a single layer of carbon atoms arranged in a two-dimensional honeycomb lattice structure, is ideal for studying the intriguing ways that electrons interact with each other in a quantum-mechanical and relativistic manner.

Since early research into the material, one point of interest has been the arrangement of electrons in graphene under an external magnetic field (5). For neutral graphene under a magnetic field, the number of possible electron states with the same energy level is exactly twice the number of electrons. Thus, the electrons in this partially filled level arrange among themselves by each choosing between the possible states. The possible states are primarily distinguished by one label for each of two properties—for the spin (up or down) and for the so-called valley (*K*

or *K'*) that determines the possible positions of the electrons. For the *K* valley, the electrons of neutral graphene are located on carbon atoms with two carbon neighbors to the right and one neighbor to the left, whereas for the *K'* valley, the electrons occupy carbon atoms with one neighbor to the right and two neighbors to the left.

The exact combination of labels for any particular electron is determined by the exchange energy, which is a quantum-mechanical effect that occurs between identical particles—in this case, the electrons. The exchange energy plays a part in the repulsive electron-electron interaction by favoring a collective state where all electrons share the same labels to minimize their repulsion.

However, for graphene with an overall neutral charge in a magnetic field, it is impossible for all electrons to occupy states with the same label combination. Here, one-half of the states have to be occupied with electrons, whereas only one-quarter of the electron states have spin up and *K* as labels, one-quarter have spin up and *K'*, one-quarter have spin down and *K*, and one-quarter have spin down and *K'*. Hence, the electrons must partially choose states with different labels. For example, if all electrons have the same valley label, half of their spins will be up and half of them down. In reality, the exact combinations of labels for the electrons are more complex and are very difficult to predict (5). One reason is that none of the labels is preselected by the mutual electron repulsion. One must also consider the quantum-mechanical superposition of label choices, meaning that each electron can simultaneously have the up and down spin label and also the *K* alongside the *K'* valley label. Moreover, as is usual for such superposition states, the different label choices are related to something known as the quantum-mechanical phase factor. These phase factors are necessary to describe the electrons as waves and as particles at the same time. In the superposition state, the wave of the *K* state and the wave of the *K'* state are overlapped. The phase factor of the superposition describes how the peaks of the two waves are positioned with respect to each other.

<sup>1</sup>II. Institute of Physics B and JARA-FIT, RWTH-Aachen University, 52074 Aachen, Germany. <sup>2</sup>Laboratoire de Physique des Solides, CNRS, Université Paris Saclay, Bât. 510, 91405 Orsay cedex, France.  
Email: mmorgenstern@physik.rwth-aachen.de

Liu *et al.* used scanning tunneling microscopy to determine the valley labels of the electrons. The percentage of the valley superposition is directly visible as the relative occupation of the two types of atoms in the mapped electron density (see the figure). The authors tracked this percentage and the phase factor of the superposition, when increasing the magnetic field, and observed that the percentage changes continuously from a 100% *K* valley occupation to a 50/50 mixed superposition state of *K* and *K'*. Moreover, the magnetic field where the transition between these two types of valley ferromagnets takes place depends on the substrate on which the graphene is deposited. The reason is the substrate's different influence on the two types of carbon sites in graphene.

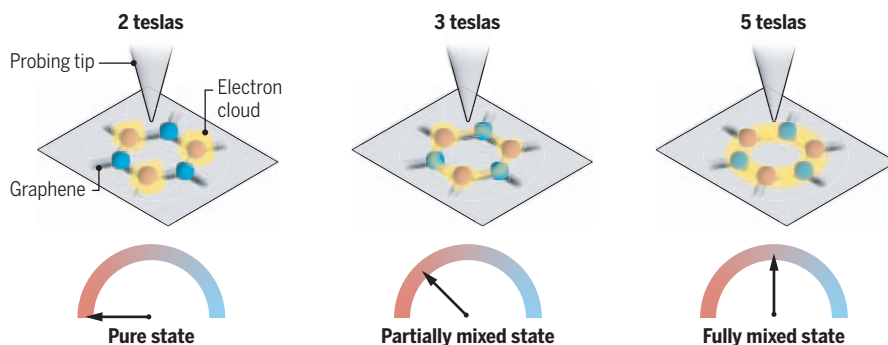
Besides revealing unprecedented details of the change in electron arrangement when changing the magnetic field, Liu *et al.* probed the electron system around positions where more electrons are located than

tribution of the constituting electrons. The emergence of such new particles from the ensemble of electrons is a central beauty of many-particle electron systems.

The experiments by Liu *et al.* mark a milestone for probing real-space patterns of electron arrangements occurring as a result of strong interactions. This can be extended by mapping the spin degree of freedom using spin-polarized scanning tunneling microscopy (10). Hence, details of all four degrees of freedom in the ground state arrangement could be disclosed as crucial to decipher other ground states where the spin labels also exhibit a superposition (11). Moreover, the authors found well-known indications of even more-complex electron arrangements in their experiments. In these arrangements, some of the emergent particles carry only one-third of the charge of a single electron (12). Using the approach of Liu *et al.*, one could begin to tackle the mysteries of these and many other emergent particles by direct imaging (13). ■

## How electrons rearrange themselves in graphene

The tip of a scanning tunneling microscope maps the electron distribution (shown as a yellow haze) on the two different types of carbon sites (shown as red and blue balls) at different magnetic fields. The distribution changes continuously from being only on the red sites to being equally distributed on both types of sites, marking a quantum-mechanical transition of the electron arrangement.



on average. It has long been conjectured that such areas are surrounded by a well-defined distribution of the percentage and of the quantum-mechanical phase factor of the valley label (6, 7, 8). On each circle around the charge center, percentage and phase factor change in a way that eventually leads to a new particle—the charged skyrmion (6), which is a localized structure made of many interacting electrons. Although there have been indirect experimental observations for charged skyrmions (9), they have never been observed in real space. Liu *et al.* mapped out the percentage and phase factor of the valley label and found the whirlpool-type texture around a charged position. The valley skyrmion has a size of 7 nm and shows excellent agreement with model calculations (7, 8). As particles, such skyrmions are robust, although they are made of a complex dis-

## REFERENCES AND NOTES

1. G. Murthy, R. Shankar, *Rev. Mod. Phys.* **75**, 1101 (2003).
2. X. Liu *et al.*, *Science* **375**, 321 (2022).
3. K. S. Novoselov *et al.*, *Nature* **438**, 197 (2005).
4. L. Balents, C. R. Dean, D. K. Efetov, A. F. Young, *Nat. Phys.* **16**, 725 (2020).
5. M. O. Goerbig, *Rev. Mod. Phys.* **83**, 1193 (2011).
6. S. L. Sondhi, A. Karlhede, S. A. Kivelson, E. H. Rezayi, *Phys. Rev. B* **47**, 16419 (1993).
7. Y. Lian, M. O. Goerbig, *Phys. Rev. B* **95**, 245428 (2017).
8. J. Atteia *et al.*, *Phys. Rev. B* **103**, 035403 (2021).
9. S. E. Barrett, G. Dabbagh, L. N. Pfeiffer, K. W. West, R. Tycko, *Phys. Rev. Lett.* **74**, 5112 (1995).
10. M. Bode, *Rep. Prog. Phys.* **66**, 523 (2003).
11. A. F. Young *et al.*, *Nature* **505**, 528 (2014).
12. M. Reznikov, R. de Picciotto, T. G. Griffiths, M. Heiblum, V. Umansky, *Nature* **399**, 238 (1999).
13. Z. Papić *et al.*, *Phys. Rev. X* **8**, 011037 (2018).

## ACKNOWLEDGMENTS

The authors acknowledge M. Pratzer for the figure outline and the German Research Foundation (Mo 858/15-1) and the Agence Nationale de la Recherche (ANR-17-CE30-0029) for financial support.

## DISEASE

# Epstein-Barr virus and multiple sclerosis

Infection with Epstein-Barr virus is the trigger for the development of multiple sclerosis

By William H. Robinson<sup>1,2</sup> and Lawrence Steinman<sup>3</sup>

Infection with the Epstein-Barr virus (EBV) has long been postulated to trigger multiple sclerosis (MS) (1). Prior analyses demonstrated increased serum antibodies to EBV in ~99.5% of MS patients compared with ~94% of healthy individuals (2). On page 296 of this issue, Bjornevik *et al.* (3) analyzed EBV antibodies in serum from 801 individuals who developed MS among a cohort of >10 million people active in the US military over a 20-year period (1993–2013). Thirty-five of the 801 MS cases were initially EBV seronegative, and 34 became infected with EBV before the onset of MS. EBV seropositivity was nearly ubiquitous at the time of MS development, with only one of 801 MS cases being EBV seronegative at the time of MS onset. These findings provide compelling data that implicate EBV as the trigger for the development of MS.

How does a virus with tropism for B cells develop into a disease of the central nervous system (CNS)? In MS, there is an inflammatory attack against the myelin sheath and the axons that it insulates. Ultimately, neurons themselves are injured. In MS, B cells and their activated progeny, plasmablasts, express integrin  $\alpha 4$ , which has adhesive properties that allow these antibody-producing cells to move from the bone marrow to the peripheral circulation and then across the blood-brain barrier (BBB), where they take residence inside the brain and its internal lining (4). A distinct feature of MS is the synthesis of

<sup>1</sup>Division of Immunology and Rheumatology, Department of Medicine, Stanford University, Stanford, CA, USA. <sup>2</sup>VA Palo Alto Health Care System, Palo Alto, CA, USA. <sup>3</sup>Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, USA. Email: w.robinson@stanford.edu; steinman@stanford.edu



immunoglobulins by clonal expansions of plasmablasts within the brain. When these immunoglobulins, found in cerebrospinal fluid (CSF) from patients with MS, are applied to an electrophoretic gel, they form bands of restricted mobility, called oligoclonal immunoglobulin bands, representing clonal expansions of plasmablasts. These antibodies target myelin-producing glial cells, thereby damaging them (4).

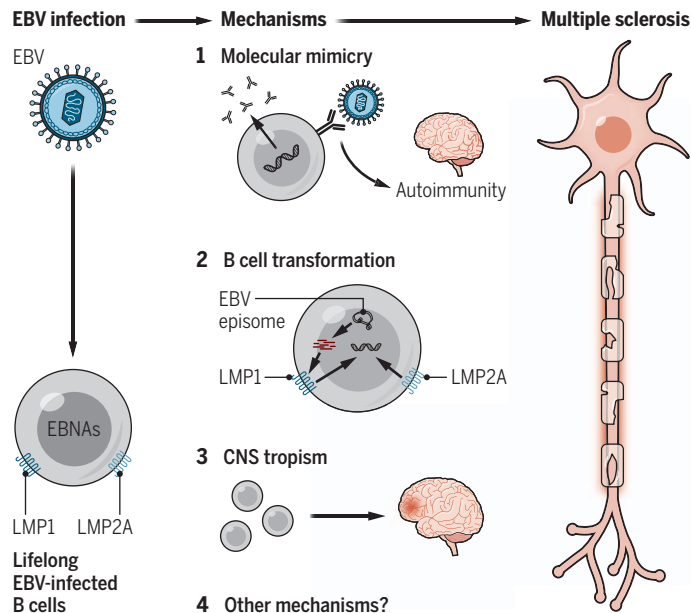
Multiple studies have identified EBV-infected B cells in the brains of MS patients (5, 6). Understanding how infection of B cells with EBV initiates the pathology seen in MS is now ripe for a deeper understanding of the roles of these clonally expanded B cells and plasmablasts. Depletion of B cells with monoclonal antibodies targeting CD20 has emerged as one of the most efficacious therapies for MS (7). However, because of the BBB, CD20 monoclonal antibody therapies do not reach the CNS in sufficient amounts, and moreover, antibodies to CD20 do not deplete their progeny, antibody-producing plasmablasts and plasma cells, which lack CD20.

The mechanism (or mechanisms) of EBV-mediated MS development remains elusive. Possibilities include molecular mimicry, through which EBV viral protein sequences mimic human myelin proteins and other CNS proteins and thereby induce autoimmunity against myelin and CNS antigens (8). EBV transformation of B cells could also lead to clonal expansion of pathogenic plasmablasts. EBV may transform B cells through disruption of several pathways: EBV latent membrane protein 2A (LMP2A) mimics B cell receptor (BCR) signaling. LMP1 mimics CD40 receptor signaling, a costimulatory pathway that is important for B cell-T cell interaction. Additionally, EBV encodes an interleukin-10-like protein, which activates B cells (9). EBV might also mediate bystander damage to the axon and its surrounding sheath, or defective clearance of infected B cells. CD8<sup>+</sup> T cells specific for EBV lytic proteins are present in MS brain lesions, and a persistent EBV infection in the CNS might stimulate CD8<sup>+</sup> T cell responses that mediate CNS injury (4–8) (see the figure).

There are multiple reports suggesting that molecular mimicry might induce MS. Serum antibodies from MS patients to the EBV small capsid protein BFRF3 cross-react

## Model for multiple sclerosis development

In at-risk individuals, Epstein-Barr virus (EBV) infection of B cells promotes the development of multiple sclerosis through several possible mechanisms. These include molecular mimicry (1) by EBV nuclear antigen 1 (EBNA-1), B cell transformation (2) through latent membrane protein 1 (LMP1) and LMP2A, induction of B cell trafficking (3) to the central nervous system (CNS), and/or other unknown mechanisms (4).



with the cytoplasmic protein septin-9 and are associated with demyelination (10). Another study showed serum antibodies from MS patients are cross-reactive between amino acids 411–440 of the viral protein EBV nuclear antigen 1 (EBNA-1) and the human chloride-channel protein, anoctamin 2 (ANO2), which is associated with electrical conduction in axons (11). MS serum antibodies targeting EBNA-1 residues 411–426 that cross-react with myelin basic protein have also been identified (12). Clonally expanded antibodies in the CSF of MS patients targeting EBNA-1 residues 386–405 that cross-react with the CNS cell adhesion molecule, glialCAM, have also been described (4). It is intriguing that three contiguous regions of mimicry have been reported in a small region of the EBNA-1 protein; this may arise through immune surveillance in a process called epitope spreading.

Increased incidence of EBV infection is associated with other autoimmune diseases, including systemic lupus erythematosus (SLE). Serologic reactivation of EBV (production of EBV serum antibodies after resolution of acute infection) is associated with transition to clinical SLE. EBNA-1, through amino acid regions distinct from those implicated in MS, has been shown to mediate molecular mimicry of nuclear antigens associated with SLE pathogenesis (13, 14). Mice engineered to express a CD40-LMP1 fusion protein exhibited

increased EBNA-1-mediated molecular mimicry and lupus-like clinical features (15). Whether EBV infection activates other inflammatory mechanisms common to MS and other autoimmune diseases, including SLE, is under investigation.

Nearly everyone is infected with EBV, but only a small fraction develop MS. Thus, other factors, such as genetic susceptibility, are important in MS pathogenesis. Certain genes, such as those encoding the antigen-presenting human leukocyte antigen (HLA) proteins, determine the portion of a protein that is presented to the immune system. Other genes control modifications in EBV-associated proteins, including phosphorylation. Such genes are critical for modulating molecular mimicry (4, 11). Thus, given these additional gating factors in MS pathogenesis, infection with EBV is likely to be necessary, but not sufficient, to trigger development of MS.

Infection with EBV is the initial pathogenic step in MS, but additional fuses must be ignited for the full pathophysiology.

There may be new opportunities for therapy: Would a vaccine against EBV protect against MS? Can the B cells that dwell in the CSF be killed or inactivated with therapeutics? Would antivirals that target EBV provide effective therapy, especially when given early in the course of disease? Now that the initial trigger for MS has been identified, perhaps MS could be eradicated. ■

## REFERENCES AND NOTES

1. A. Bar-Or et al., *Trends Mol. Med.* **26**, 296 (2020).
2. A. Ascherio, K. L. Munger, *Ann. Neurol.* **61**, 288 (2007).
3. K. Bjornevik et al., *Science* **375**, 296 (2022).
4. T. V. Lanz et al., *Research Square* 10.21203/rs.3.rs-1239863/v1 (2022).
5. M. A. Moreno et al., *Neurol. Neuroimmunol. Neuroinflamm.* **5**, e466 (2018).
6. B. Serafini et al., *J. Virol.* **93**, e00980 (2019).
7. S. L. Hauser et al., *N. Engl. J. Med.* **376**, 221 (2017).
8. R. S. Fujinami, M. B. Oldstone, *Science* **230**, 1043 (1985).
9. M. S. Kang, E. Kieff, *Exp. Mol. Med.* **47**, e131 (2015).
10. J. W. Lindsey, *J. Neuroimmunol.* **310**, 131 (2017).
11. K. Tengvall et al., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16955 (2019).
12. N. R. Jog et al., *J. Autoimmun.* **106**, 102332 (2020).
13. M. T. McClain et al., *Nat. Med.* **11**, 85 (2005).
14. A. Sabbatini et al., *Eur. J. Immunol.* **23**, 1146 (1993).
15. M. E. Munroe et al., *Front. Immunol.* **11**, 606936 (2021).

## ACKNOWLEDGMENTS

W.H.R. is a coinventor on a patent application filed by Stanford University that includes antibodies to EBV.

Published online 13 January 2022

10.1126/science.abm7930





## POLLUTION

# The global nitrogen-phosphorus imbalance

The imbalance has grave consequences for natural ecosystems and global food security

By **Josep Peñuelas**<sup>1,2</sup> and **Jordi Sardans**<sup>1,2</sup>

**E**xponential increases in the human population and its activities are accelerating global changes, from the climate to land use to loss of species. The rise in atmospheric concentrations of greenhouse gasses, mainly CO<sub>2</sub> from the combustion of fossil fuels, is the most well-known driver of global change (1). Emission of greenhouse gases, which also include methane (CH<sub>4</sub>) and nitrous oxide (N<sub>2</sub>O), are stoking global warming as well as more frequent and extreme weather events, such as droughts and floods. Land use and pollution also have major impacts on Earth's future (1). Among these ongoing anthropogenic changes, the biospheric nutrient imbalance between nitrogen (N) and phosphorus (P) is less known and deserves more attention.

In 2003, a pioneering study by Sterner and Elser reported that the aquatic N/P ratio determined the community structure and function of plankton in lakes (2). When the concentrations of both N and P are not limited, the rate of protein synthesis by plankton depends mostly on the amount of P-rich RNA that the organisms produce

and is therefore negatively correlated with the cellular N/P ratio. Thus, lower N/P ratios are associated with faster protein synthesis and higher growth rates of plankton (2). This negative correlation has multiple ecological consequences for the structure and functioning of ecosystems, as has been reported in all types of ecosystems.

Human activities have substantially altered this N/P ratio in water, soils, and organisms over the past five decades (3). The much faster increase of anthropogenic inputs of reactive N to the biosphere than inputs of P has led to a global increase in the N/P ratio. The main anthropogenic sources of reactive N include the many kinds of nitrogen oxides from burning fossil fuels, the planting of N-fixing crops, and the use of N-rich fertilizers as well as their runoff into waterways. Although there are also human activities that have increased the amount of P in soils and waters—for example, from the application of P-rich fertilizers and detergents—the overall increase in the input of P is dwarfed by that of N. This increase in both N and P has led to eutrophication (excess of nutrients) of waters and soil. Some countries have implemented water-treatment strategies to decrease N and P concentrations. However, the technology used by these water-treatment plants retains more P than N and therefore increases the N/P ratio as an unintended consequence (4).

The global N/P ratio of anthropogenic inputs has increased from about 19:1 in molar basis in the 1980s to 30:1 in 2020 (3). The interactions of these N/P ratios in water, soil, and organisms with drivers of global change, such as warming and increasing atmospheric CO<sub>2</sub> concentrations, further increase the N/P ratio in some biomes (5).

The current global N/P ratio of anthropogenic inputs is thus larger than the averages of the main ecosystem compartments such as soil (up to 22:1), humus (up to 17:1), ocean water (up to 16:1), terrestrial plants (up to 30:1), and plankton (up to 16:1). This imbalance at the global scale may be even greater at the local-regional scale because the inputs of N and P are not evenly spread around the world, and because N and P have very different mobilities in the environment. P tends to have low water solubility and volatilization, is often adsorbed and precipitates in soil as salt minerals, and is buried in sediments, thus tending to remain near the emission sources. By contrast, N is much more water-soluble and volatile and thus tends to spread out over a larger radius from its sources (3).

The biological impacts of the increasing N/P imbalance have already been observed and reported in the trophic communities of several continental water bodies (6), in the structure and function of soil communities (7), and in the species composition of plant

<sup>1</sup>Consejo Superior de Investigaciones Científicas (CSIC), Global Ecology Unit CREAL-CSIC-UAB (Universitat Autònoma de Barcelona), Bellaterra 08193, Barcelona, Catalonia, Spain.

<sup>2</sup>CREAF, Cerdanyola del Vallès 08193, Barcelona, Catalonia, Spain. Email: josep.penuelas@uab.cat



The use of nitrogen-rich fertilizers and their runoff into waterways are major contributors to the growing imbalance of nitrogen and phosphorus.

communities (8). The impact of the imbalance will continue to increase as the imbalance continues to tip in the same direction.

This increasing N/P imbalance can have severe consequences not only for natural ecosystems but also for human societies because crop production and food security will be affected. The resource gap in remedying this imbalance at the regional level may also broaden the economic gap between rich and poor countries (9). N-containing fertilizers have an unlimited source—the atmosphere—from which N can be extracted through the Haber-Bösch reaction. This innovation has allowed a continuous increase in the production and use of these N fertilizers since the 1950s (10). By comparison, P sources have largely been limited to mines and are concentrated in very few countries, such as Morocco (9), so P might eventually become economically inaccessible to low-income and food-deficient countries as these sources become depleted. In the future, P-producing nations are likely to manage their reserves to maximize profit for both their domestic mining and farming industries, making P-based fertilizers increasingly unaffordable for farmers in the poorest countries (9) and worsening the N/P imbalance in regions where the problem is the most prominent.

Imbalanced soil N/P ratios can also affect the chemical composition of crops, which can have implications on public health. For example, in some regions P accumulates in soils and water bodies, and the N/P ratio decreases, as a result of an excessive use of inorganic and organic P fertilizers (3). Food produced in these environments may lead to an overconsumption of P in the local population, which can have negative implications for their health (11). The implications of this global imbalance between N and P could also have impacts on several other human infectious and noninfectious illnesses that are strongly associated with diet, such as coeliac disease (12).

Besides the imbalance observed in the N/P ratio, human activities also generate imbalances among other elements. For example, changes in the ratio of carbon (C) and N relative to iron, zinc, calcium, and potassium, among others, have been observed in plant tissues (12). The increasing atmospheric concentrations of CO<sub>2</sub> are likely a driver of the increase of C in plants, which in turn have developed more compounds that reduce the concentrations of these other elements (12). This indirectly leads to the displacement of the elemental composition (elementome) (13) of organisms, com-

munities, and entire ecosystems owing to imbalances of the anthropogenic biospheric inputs of C and N relative to P and other elements in recent decades and is likely to exacerbate in the coming decades.

The time has come for national and international environmental agencies and policy-makers to recognize the risks of unbalanced N/P ratios and other parallel imbalances in elemental stoichiometry to the biosphere and humanity. The international environmental agencies and policy-makers should address the problem through a coordinated international policy. Observations, experimentation, theory, and modeling at different temporal and spatial scales are warranted to evaluate, predict, and provide possible solutions to these anthropogenic nutritional imbalances and their effects on nature and humans. Among these possible solutions, increasing the efficiency of use and cycling of N and P—for example, through precision agriculture to avoid misuse of fertilizers, methods to increase plant accessibility to P sources, use of innovative management techniques and biotechnologies to improve nutrient-use efficiency, stimuli and subsidies for recycling P through legislative regulations and instruments at the national or regional administrative level, or reduction of livestock production—has been suggested as the most effective approach to prevent imbalanced N/P ratios for food production and reduce environmental problems that involve N and P. This research will determine whether these nutritional imbalances should be added to the planetary boundaries instead of only considering N and P separately (14). ■

#### REFERENCES AND NOTES

1. P. Ekins, J. Gupta, *Glob. Sustain.* **2**, e20 (2019).
2. W. Sterner, J. Elser, *Ecological Stoichiometry: The Biology of Elements from Molecules to Biosphere* (Princeton Univ. Press, 2003).
3. J. Peñuelas, I. A. Janssens, P. Ciais, M. Obersteiner, J. Sardans, *Glob. Change Biol.* **26**, 1962 (2020).
4. Y. Tong et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11566 (2020).
5. K. Yue et al., *Glob. Change Biol.* **23**, 2450 (2017).
6. J. J. Elser et al., *Science* **326**, 835 (2009).
7. C. Ren et al., *Biol. Fertil. Soils* **53**, 457 (2017).
8. Y. Fujita et al., *Nature* **505**, 82 (2014).
9. M. Obersteiner, J. Peñuelas, P. Ciais, M. van der Velde, I. A. Janssens, *Nat. Geosci.* **6**, 897 (2013).
10. V. Smil, *Nature* **400**, 415 (1999).
11. M. S. Razzaque, *Clin. Sci. (Lond.)* **120**, 91 (2011).
12. R. E. Paseka, A. R. Bratt, K. L. MacNeill, A. Burian, C. R. See, *Front. Ecol. Evol.* **7**, 378 (2019).
13. J. Peñuelas et al., *Ecology* **100**, e02652 (2019).
14. W. Steffen et al., *Science* **347**, 12598553 (2015).

#### ACKNOWLEDGMENTS

Our research is supported by the Spanish Government (grants PID2019-110521GB-I00 and PID2020115770RB-I), Fundación Ramon Areces (grant ELEMENTAL-CLIMATE), Catalan Government (grants SGR 2017-1005 and AGAUR-2020PANDE00117), and European Research Council (Synergy grant ERC-SyG-2013-610028, IMBALANCE-P).

#### VIEWPOINT: COVID-19

## Nervous system consequences of COVID-19

Neurological symptoms highlight the need to understand pathophysiologic mechanisms

By Serena Spudich<sup>1</sup> and Avindra Nath<sup>2</sup>

**A**lthough severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is considered a respiratory pathogen, myriad neurologic complications—including confusion, stroke, and neuromuscular disorders—manifest during acute COVID-19. Furthermore, maladies such as impaired concentration, headache, sensory disturbances, depression, and even psychosis may persist for months after infection, as part of a constellation of symptoms now called Long Covid. Even young people with mild initial disease can develop acute COVID-19 and Long Covid neuropsychiatric syndromes. The pathophysiological mechanisms are not well understood, although evidence primarily implicates immune dysfunction, including nonspecific neuroinflammation and antineural autoimmune dysregulation. It is uncertain whether unforeseen neurological consequences may develop years after initial infection. With millions of individuals affected, nervous system complications pose public health challenges for rehabilitation and recovery and for disruptions in the workforce due to loss of functional capacity. There is an urgent need to understand the pathophysiology of these disorders and develop disease-modifying therapies.

Initial reports of neurologic syndromes accompanying COVID-19 described changes in level of consciousness or cognitive dysfunction, weakness, and headache in hospitalized patients that might be attributable to any severe acute illness with respiratory and metabolic disturbances. Subsequently, reports of strokes and acute inflammation

<sup>1</sup>Yale School of Medicine, New Haven, CT, USA.

<sup>2</sup>National Institute of Neurological Diseases and Stroke, National Institutes of Health, Bethesda, MD, USA.  
Email: serena.spudich@yale.edu



or demyelination of the central or peripheral nervous system highlighted specific cerebrovascular and neural tissue involvement. As the number of cases increased globally, it was recognized that SARS-CoV-2 not only induces respiratory symptoms but also can affect multiple organ systems, including the kidneys, gastrointestinal tract, heart, and brain.

Clinical neurological and psychiatric syndromes in patients with acute COVID-19 have been delineated by surveillance studies of hospitalized patients. A UK-wide study of hospitalized patients identified the most common neurologic conditions as anosmia (loss of smell), stroke, delirium, brain inflammation, encephalopathy, primary psychiatric syndromes, and peripheral nerve syndromes (1). Varied timing of onset suggests that these conditions have diverse pathophysiological mechanisms. For example, cerebrovascular complications co-occur with or even predate the onset of respiratory symptoms, whereas central inflammatory and peripheral nerve conditions manifest on average 2 weeks later, suggesting that they may result from peri- or postinfectious processes (2).

Despite early speculation that SARS-CoV-2 may enter the central nervous system (CNS) via migration through the nasal cavity and the olfactory pathway or trafficking across the blood-brain barrier, analysis of cerebrospinal fluid (CSF) from living patients with neuropsychiatric manifestations has almost uniformly failed to detect viral RNA by reverse transcription polymerase chain reaction. Instead, the preponderance of evidence from CSF and brain tissue suggests that immune activation and inflammation within the CNS is the primary driver of neurologic disease in acute COVID-19. Indeed, histopathological studies of brain tissue from patients who died with acute COVID-19 reveal only limited detection of SARS-CoV-2 nucleic acid or viral protein in the brain (3, 4), consistent with findings in CSF from live patients. Direct examination of autopsy brain tissue has caveats—those who died with acute COVID-19 had severe disease that may not be representative of the majority of those infected with SARS-CoV-2. Many had systemic or metabolic derangements prior to death that may contribute to pathology in a nonspecific way. However, when infrequently detected, infected cells in human brain lack surrounding clusters of inflammatory cells, suggesting that SARS-CoV-2 presence in the CNS does not incite classic viral encephalitis.

Examination of CSF samples from living patients reveals neuroinflammation and

aberrant neuroimmune responses during acute COVID-19. CSF shows up-regulation in the expression of interferon-regulated genes in dendritic cells, along with activated T cells and natural killer (NK) cells. This is accompanied with an increase in interleukin-1 (IL-1) and IL-12, which is not seen in blood plasma (5). Additionally, CSF-specific clonal expansion of T cells and antibodies that recognize epitopes of SARS-CoV-2 spike protein that cross-react with neural antigens suggest compartmentalization of the immune response (5, 6), although the possibility of a persistent infection with restricted viral replication cannot be entirely excluded. During this acute phase, other markers of monocyte activation and neuronal injury can also be detected in CSF (7). In the following sub-acute phase, patients with severe manifestations show diminished interferon re-

**“Many people who experience neurologic symptoms that linger after acute COVID-19 are less than 50 years old and were healthy and active prior to infection.”**

sponses and markers of T cell exhaustion in CSF (8).

Autopsy studies of patients with acute COVID-19 show infiltration of macrophages, CD8<sup>+</sup> T lymphocytes in perivascular regions, and widespread microglial activation throughout the brain (3). Single-cell analysis of brain tissue has also confirmed CD8<sup>+</sup> T lymphocyte infiltration and microglial activation without evidence of SARS-CoV-2 RNA detection in cells of the brain parenchyma (9). The robust, generalized, and SARS-CoV-2-specific immune responses observed in the CNS are puzzling in the absence of readily detectable virus and may suggest transient infection of the brain very early in infection or low concentrations of viral antigen in the CNS. Systemic activation of immune cells may additionally lead to up-regulated expression of cell surface markers that facilitate amplified trafficking into the nervous system, even in the absence of targeted CNS antigens (see the figure).

Does widespread vascular dysfunction contribute to nervous system complications of COVID-19? Acute COVID-19 is associated with heightened risk of stroke compared with influenza illness of similar severity, even after correcting for stroke risk factors (10). Overt cerebrovascular events during acute COVID-19 often occur in those with vulnerabilities to vascular

disease (such as advanced age and cardiac disease). Increases in blood markers of vascular inflammation as well as thrombosis and infarction in other tissues can also be found in patients with COVID-19 and stroke, suggesting that endothelial inflammation and coagulopathy contribute to these events (11). Indeed, system-wide vascular dysfunction can characterize severe acute COVID-19 and has the potential to contribute to manifestations of organ system failure and systemic inflammation in those most severely ill (12). It is plausible that subtle forms of generalized vascular dysfunction, including thrombotic microangiopathy (microscopic blood clots) in the brain, may lead to neurological symptoms even in the absence of clinically apparent stroke. Additionally, high-field magnetic resonance examination of brain tissue demonstrates microvascular damage in structures plausibly related to neurologic manifestations of COVID-19, consistent with endothelial activation and widespread vascular injury observed in other organs (4).

Since early in the COVID-19 pandemic, patients have described lingering syndromes following acute infection, now called Long Covid.

These syndromes often include predominant neurologic and psychiatric symptoms, such as difficulty with memory, concentration, and ability to accomplish everyday tasks, frequent headaches, alterations in skin sensation, autonomic dysfunction, intractable fatigue, and in severe cases, delusions and paranoia. Many people who experience neurologic symptoms that linger after acute COVID-19 are less than 50 years old and were healthy and active prior to infection. Notably, the majority were never hospitalized during their acute COVID-19 illness, reflecting mild initial disease. Many of the symptoms experienced by individuals with Long Covid are similar to those of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), which is also considered to be a postinfectious syndrome caused by a variety of infectious agents. Because the pathophysiology of ME/CFS is poorly understood and there are no effective disease-modifying therapies available, it is likely that the study of Long Covid may benefit ME/CFS patients as well. There is also overlap in symptoms of post-Lyme disease, suggesting that there may be common host susceptibility factors that underlie these illnesses.

The heterogeneity of symptoms affecting individuals with Long Covid and the difficulties in ascertaining which symptoms may be a consequence of SARS-CoV-2 in-

fection versus aggravation of preexisting or coincidental conditions pose enormous challenges for mechanistic understanding and approaches to treatment. Few studies have systematically categorized or examined the natural history of Long Covid symptoms, let alone studied their biology. Of 3762 respondents in an online study of people with persistent symptoms after documented or suspected COVID-19, many had ongoing symptoms up to 7 months after initial infection, including prominent neuropsychiatric syndromes (13). Serial imaging routinely captured in the UK Biobank cohort has revealed focal areas of brain atrophy in individuals after documented COVID-19 compared with a parallel group without COVID-19, suggesting a potential biomarker for brain effects of SARS-CoV-2 (14). Studies of positron emission tomography (PET) imaging

also show decreased metabolic activity in the brain in people with Long Covid (15). However, the pathophysiology leading to these symptoms and cerebral changes is unknown. Potential etiologies are mainly extrapolated from current understanding of nervous system pathogenesis during acute COVID-19. These include residual immune activation or persistent autoimmune disturbance, ongoing endothelial activation or vascular dysfunction, or residual of injury accrued during acute disease. Systematic neurologic studies of carefully phenotyped individuals with neurological Long Covid symptoms are essential. These patients often also experience stigma, employment difficulties, and mental health challenges. Thus, diagnostic certainty and therapeutic interventions are needed to address this major public health concern.

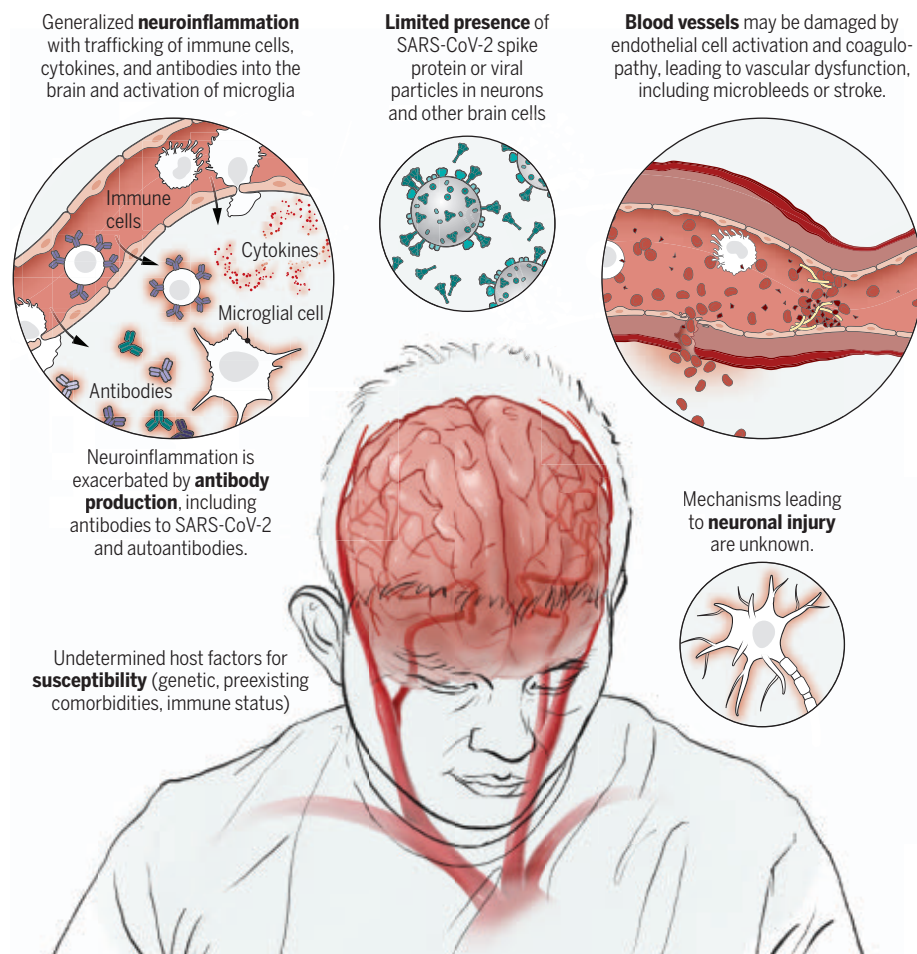
The full extent of the long-term neuro-

logical complications of COVID-19 has not been realized. Observations of neuroinflammation and neuronal injury in acute COVID-19 have raised the possibility that infection may accelerate or trigger future development of neurodegenerative diseases such as Alzheimer's or Parkinson's diseases. No information is yet available regarding neurodevelopmental trajectories in children, who usually experience mild COVID-19 and manifest few neurologic or psychiatric symptoms during or after acute illness. Those who experience the rare multisystem-inflammatory syndrome in children (MIS-C) may be at particular risk for neurological sequelae owing to widespread endothelial activation, often involving the brain.

What are the host factors that account for the wide variability in clinical manifestations such that some patients develop acute neurologic illness, and others develop persistent postinfectious complications? It will be critical to characterize the pattern(s) of immune dysregulation in Long Covid patients. Is it possible that persistent immune dysregulation underlies ongoing symptoms? If so, this may be driven by host antigens with autoimmune responses, or a persistent viral infection with restricted viral replication in tissue reservoirs. Whether antiviral or immune-targeted interventions early in the disease course or prophylactic vaccination against COVID-19 will alter the trajectory of neurologic complications of COVID-19 is also unknown. Investigations that include longitudinal studies with neurological and psychiatric assessments and rigorous host-pathogen studies of systemic and nervous system interactions have the potential to answer these questions. Ultimately, interventional trials based on these discoveries are needed to determine approaches to curtail or reverse nervous system effects of COVID-19 that are experienced by huge numbers of people globally. ■

## Putative neuropathogenic effects of SARS-CoV-2

Infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can lead to neuropsychiatric effects during acute COVID-19, including confusion, stroke, and neuromuscular disorders. These may arise from neuroinflammation, coagulopathy, neuronal injury, and possibly viral infection in the central nervous system. Causes of Long Covid symptoms affecting the nervous system may result from the emergence and persistence of these mechanisms.



## REFERENCES AND NOTES

1. A. Varatharaj et al., *Lancet Psychiatry* **7**, 875 (2020).
2. A. L. Ross Russell et al., *Brain Commun.* **3**, fcab168 (2021).
3. J. Matschke et al., *Lancet Neurol.* **19**, 919 (2020).
4. M. H. Lee et al., *N. Engl. J. Med.* **384**, 481 (2021).
5. E. Song et al., *Cell Rep. Med.* **2**, 100288 (2021).
6. C. Franke et al., *Brain Behav. Immun.* **93**, 415 (2021).
7. A. Edén et al., *Neurology* **96**, e294 (2021).
8. M. Heming et al., *Immunity* **54**, 164 (2021).
9. J. F. Fullard et al., *Genome Med.* **13**, 118 (2021).
10. A. E. Merkler et al., *JAMA Neurol.* **77**, 1366 (2020).
11. L. S. McAlpine et al., *Stroke* **52**, e233 (2021).
12. L. A. Teuwen et al., *Nat. Rev. Immunol.* **20**, 389 (2020).
13. H. E. Davis et al., *EClinicalMedicine* **38**, 101019 (2021).
14. G. Douaud et al., *medRxiv* 10.1101/2021.06.11.21258690 (2021).
15. J. A. Hospet et al., *Brain* **144**, 1263 (2021).

10.1126/science.abm2052

## POLICY FORUM

## TECHNOLOGY POLICY

# How privacy's past may shape its future

An account of privacy's evolutionary roots may hold lessons for policies in the digital age

By **Alessandro Acquisti**<sup>1</sup>, **Laura Brandimarte**<sup>2</sup>,  
**Jeff Hancock**<sup>3</sup>

**C**ontinued expansion of human activities into digital realms gives rise to concerns about digital privacy and its invasions, often expressed in terms of data rights and internet surveillance. It may thus be tempting to construe privacy as a modern phenomenon—something our ancestors lacked and technological innovation and urban growth made possible. Research from history, anthropology, and ethnography suggests otherwise. The evidence for peoples seeking to manage the boundaries of private and public spans time and space, social class, and degree of technological sophistication. Privacy—not merely hiding of data, but the selective opening and closing of the self to others—appears to be both culturally specific and culturally universal (1). But what could explain the simultaneous universality and diversity of a human drive for privacy? An account of the evolutionary roots of privacy may offer an answer and teach us about privacy's digital future and how to manage it (2).

## THE EVOLUTIONARY ROOTS OF PRIVACY

Norms about, opportunities for, and even definitions of privacy fluctuate over time and, markedly, across cultures. Yet, privacy-seeking behaviors emerge across peoples throughout history and across geography: from the Greek philosophers of the fourth century BCE “discuss[ing] the concept of private life” [(3), p. 120], to Chinese thinkers developing “a sharp distinction between the concepts of public and private” by the third century BCE (p. 221); from lovers in ancient Rome who would need to “borrow the house of an indulgent friend” to escape the prying eyes of their servants [(4), p. 72], to 1950s Javanese culture, where little physical seclusion was available and people manifested their privacy needs through stiff social interactions (5). Evidence of privacy mores even

arises from within the ancient holy books of monotheistic religions—from the Quran to the Talmud to the Bible (6).

The notion that a concern for privacy may thus have evolutionary roots—and examples supporting it—can be found in the writings of leading privacy scholars from half a century ago (5, 7). Under this account, what we refer to today as privacy evolved from physical needs for security and self-interest. By using signals to assess threats and opportunities in their physical proximity, humans and other species enhance their chance of survival and evolutionary success. In our past, the ability to detect and react appropriately to the presence of others would have offered an evolutionary advantage (for instance, distinguishing kin from stranger and adapting behavior accordingly—from openness to withdrawal, from cooperative to protective). In this view, the evolutionary advantage of being able to process and appropriately react to sensorial cues that suggest the presence of others is clear.

Over time, in humans, evolutionary adaptation developed more sophisticated forms of social cognition. In the case of privacy, those developments may have followed various possible vectors, balancing the tensions between individual freedom and collective welfare. In one of them, our ancestral ability to detect and react to the physical presence of others adapted into reputation and impression management. Barkow (8), for example, suggests that the primary evolutionary function of the self could have been impression management. Selecting information about oneself to make favorable impressions on others and using reticence and withdrawal to “reduce the risk of saying or doing something that others might regard negatively” [(9), p. 520] are instruments, then and now, for such management.

There may, therefore, be a line connecting the seemingly diverse manifestations of privacy across human history: privacy as the selective, self-interested opening and closing of the self to others (1). This dialectic is a universal characteristic of the human species. And the boundaries between self and others take multiple forms—from physical to informational to regulatory. This

explains the diversity in dimensions, definitions, and norms regarding privacy over time and across the literature.

The ability to leverage sensorial cues to detect others and modulate appropriate responses is advantageous not just in terms of protection from risk, but also extraction of benefits. Supporting the evolutionary account, both dynamics have been observed in animal species (5), such as cats seeking seclusion under duress (protection against risk), and chimps lower in the social hierarchy concealing mating activities or a coveted source of food from higher-status males (extraction of benefits). Both dynamics are also surprisingly consistent with modern economic accounts of informational privacy as a self-interested process of selective sharing: An individual may rationally want to share with a marketer their interests and preferences, so as to get beneficial targeted offers (extraction of benefits), but may not want to share with the marketer their willingness to pay for those interests, so as to avoid price discrimination (protection against risk) (10).

Sensorial cues that indicate that other humans are present still influence how we define boundaries between public and private. Consider the familiar feeling alerting us that someone is staring at us as we were lost in thought: a “sense” of privacy, so to say. In our own research, we found that participants were less likely to divulge sensitive personal information in an online survey when they visually detected the presence of another person, either in the same room or—notably—in an adjacent room from which that person could not have seen what the participant was typing (11).

## A PRIVACY GAP

And herein lies the problem. If people do in fact rely, in part, on sensorial cues to navigate privacy choices, the more we transition from physical to digital interactions, the less equipped we may be for informed digital privacy decisions. At the extreme, offline sensorial cues that we depend upon may, online, be absent, subdued, or even intentionally manipulated through so-called dark patterns. Thus, privacy (and security) responses common in the physical world

<sup>1</sup>Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA. <sup>2</sup>Eller College of Management, The University of Arizona, Tucson, AZ, USA. <sup>3</sup>Department of Communications, Stanford University, Stanford, CA, USA.  
Email: [acquisti@andrew.cmu.edu](mailto:acquisti@andrew.cmu.edu)



may be muted or entirely unavailable online. This may explain the specific hurdles that we face in protecting privacy online, and the surprising observation of seemingly careless online behaviors by individuals who claim to care about their privacy (5). In our nondigital life, Altman (7) noted, behaviors that help us manage privacy boundaries are common and ubiquitous. We often engage in them with little conscious awareness. We lower our voice during intimate conversations and raise it when we want a large audience to hear us; we cover a document that we are reading to protect it from prying eyes, or raise it up and emphatically show it when we want to make a point. Instead, online, the evolutionary mechanisms that Westin (4) and Klopfer and Rubenstein (6) identified cannot help us: We do not see Google leaning over our shoulders to track our sensitive searches; we do not hear

as has the value that individuals and societies can harness from personal data.

This mismatch does not imply that online disclosures are inherently damaging. Nor should we fall into a naturalistic fallacy of elevating senses to be sole arbiters of privacy decision-making, even online. Rather, a privacy mismatch implies that, if the ability to regulate privacy boundaries is in an individual's best interest, once that evolutionarily rooted ability is impaired, individuals become vulnerable, online, to new threats.

We refer, in particular, to ubiquitous instances of microlevel influence on decision-making which, in the aggregate, shape portentous macrolevel dynamics. Although offline and online surveillance risks do differ along numerous dimensions (including the likelihood of material harm), the privacy mismatch is, in fact, happening at the same time as ramifications of in-

popular with industry, overreliance on notice and consent mechanisms, disjointed from baseline privacy safeguards, is ineffectual and can backfire—because we are wired to react to privacy invasions viscerally, not just deliberately. Those mechanisms burden individuals with “responsibilization” for problems they did not create and cannot truly control. Responsibilization also creates unequal burdens, as it disadvantages certain groups more than others: To get the most out of notice and consent, for example, people need resources—time, education, or the economic leverage to not consent to unfair policies—which are not equally distributed across the population. Even considerable regulatory interventions such as the General Data Protection Regulation (GDPR) in the European Union resulted in a proliferation of consent mechanisms that burden users with manipulative interfaces and implied consent (15). Notice and consent approaches fail to recognize that much of our privacy behavior is intuitive and visceral, rendering them archaic solutions for a modern problem.

## USING POLICY TO FOSTER AND EMBED PRIVACY TECHNOLOGIES

So, what should be done? If, as a society, we determine that privacy is still valuable to us, then to maintain it we should embed privacy by default into the fabric of our digital systems. We cannot demand that people overcome an evolved sense of privacy strongly reliant on sensorial cues unavailable online. Any approach that places not just the ability to choose, but ultimately the responsibility to protect, on individuals themselves, will—according to this account—fail. Privacy mismatches will keep rising in frequency and importance with the growth of the information economy, well past the point where any approach relying upon individual responsibility alone could contain them (12). Individuals' personal information will keep being collected across too many instances and vectors for our bounded cognitive resources to manage efficiently or effectively. And humans will keep tuning out even visible signs of electronic surveillance.

Attempts to reproduce online the visceral cues of the physical world are therefore unlikely, alone, to solve the problem: An evolutionary mechanism of explanation does not necessarily demand an evolutionary mechanism of change. For those same reasons, even well-meaning proposals [such as data propertization from economics, nudges from behavioral research, and simplified notices from usability research (13)], though appealing, may hardly make a difference absent a combination of technology and policy inter-



the US National Security Agency stepping closer to listen to our videoconferences. Yet, humans may be wired to rely in part on those very sensorial cues to assess the privacy implications of their behaviors. This discrepancy may create a gap, or mismatch, in our ability to manage privacy-sensitive scenarios in digital realms relative to the physical world—a hypothesis we have developed (11) and that Sharif, Green, and Jettinghoff have also proposed (12).

This privacy mismatch has distinctive features vis à vis other evolutionary mismatches of modernity. For example, where the mismatch between human-evolved physiology and the modern diet arises entirely within physical systems, the very boundaries of privacy have evolved with the digital transition—from predominantly physical to predominantly informational. Through that transition, privacy's role and relevance in society have evolved as well—

formational privacy issues become consequential not merely at the individual but at the societal level. Privacy externalities (13)—the collective ramifications of individual disclosures—are becoming evident, as algorithmic personalization spawns filter bubbles, amplification of disinformation, and political polarization, with implications ranging from public health to the safety of democratic elections. And though the privacy mismatch arises from our transition to digital interactions, the consequences are no longer merely digital, but physical as well: Consider the material threats associated with “doxxing,” or even episodes of genocidal violence fostered by data-driven personalization algorithms (14).

An evolutionary account of privacy provides the underlying reason why a dominant approach to privacy management in the United States—notice and consent—has failed to address these problems. Though

ventions that embed privacy safeguards into the core of our information systems.

Consider an analogy. Thanks to continuous technological improvements, the speed and acceleration of car production kept growing over time. Once cars reached velocities that rendered drivers' reaction times unreliable tools for avoiding collisions, the solution was not to teach drivers to develop faster reaction abilities, but rather to develop policy interventions (e.g., mandatory safety standards on cars for accident avoidance and damage reduction) and technological fixes (e.g., anti-lock braking systems, airbags) that countered the challenges arising from other technological progress. Better safety in cars was the result of deliberate policy intervention (driving investment in technical and infrastructural changes), not merely of driver education or market forces. In the case of privacy, policy intervention can both instill baseline safeguards (such as those embedded in the Organization for Economic Cooperation and Development's Guidelines on the Protection of Privacy and Transborder Flows of Personal Data) and foster the deployment of technologies that make those safeguards possible without hurting modern society's reliance on data.

Recent decades have, in fact, not only produced a burst of innovation around data as a critical asset for economic and societal development; they have also generated innovations in statistics, cryptography, and computer science that may address the challenges of creeping surveillance. Protocols from differential privacy to homomorphic encryption to federated learning point at the possibility of protecting individuals' privacy while allowing beneficial analytics to advance. Tools from artificial intelligence and machine learning deployed in privacy assistants suggest a future in which computerized agents may represent users' data interests when they interact with services, help them evaluate privacy risks, and identify mismatches between users' preferences and systems settings. These developments portend a world where privacy by design is possible without undermining the value of data. Indeed, economic research suggests that data protection is not inherently welfare-decreasing (13), and the use of differentially private algorithms to achieve the dual goal of producing accurate statistics while protecting privacy is being investigated.

Yet, we believe that those technologies are unlikely to fulfill their promise unless they are embedded at the core of our information infrastructures. To achieve that, we first need to resolve, through policy, an inception problem: Without a policy intervention to support their deployment, the incen-

tives for the vast array of industry players who control individuals' data may be insufficient to reach a critical mass of adoption, and only isolated individual market agents would act (13). Returning to the car-safety analogy, the externalities of privacy call for regulatory intervention, similarly to how externalities arising from unsafe driving led to policy responses.

Privacy legal scholarship has also evolved with the times, expanding the notion of privacy protection from mere control over data flows to encompass issues of autonomy and protection from bias, and proposing new approaches, such as construing data holders as data "fiduciaries" who hold legal obligations to act in the best interests of their customers. Those efforts are promising. In our view, any regulatory effort genuinely intent on addressing the challenges of privacy will have to be deliberate in avoiding the pitfall of spawning new iterations of ineffectual notice and consent mechanisms—which are easily gamed (15) and at best provide necessary but insufficient conditions for privacy management. The problems with recent legislations like GDPR make the limitations of this approach abundantly clear, with consumers fatigued by the constant clicking of buttons to waive privacy rights.

Instead, we argue for regulation that accounts for the richer understanding of privacy that scholarship has produced (one that goes beyond mere control over data or user consent) and that concentrates on fostering mass-scale deployment of privacy technology. Such efforts may include mandating products and services compliance with user-centered privacy technologies (including intelligent agents representing user interests, and not just preferences); incentivizing the usage of privacy-preserving, analytics-retaining algorithms among data holders; and fostering corporate practices that minimize user burden and the likelihood of coercion and manipulation. Several alternative paths could achieve those objectives—from standards setting to coordinated R&D efforts; from leveraging incentives to relying on penalties and fines for noncompliance. Regulatory initiatives should, therefore, be preceded and accompanied by a concerted policy effort to promote the development of these tools—by which we refer to efforts aimed both at improving technical solutions and analyzing their downstream impacts at the individual, organizational, and societal levels.

The history of privacy tells us that the drive to maintain a private space may be as universal as the drive to commune (and that the two drives are in fact intimately related)—and that humans invariably attempt to carve out those spaces even when the

odds are stacked against them by surveillance, whether digital or physical (13). The reason why concerns over privacy endure, despite privacy being repeatedly declared dead, may be in part cultural and in part related to visceral, evolutionary roots. The current state of privacy also tells us, however, that those spaces have become unquestionably harder for individuals to manage. Solutions that predominantly rely on notice and consent mechanisms are inadequate—because they fail to take into account the visceral roots of our sense of privacy and thus can be easily gamed by platforms and service providers. Understanding and then accounting for those ancestral roots of privacy may be critical to secure its future. ■

## REFERENCES AND NOTES

1. I. Altman, *J. Soc. Issues* **33**, 66 (1977).
2. In addition to the references listed at the end of this article, an annotated bibliography of further scholarly works related to arguments presented in this manuscript can be found at <https://www.heinz.cmu.edu/~acquisti/companion-science-privacy-past-future-evolution.html>.
3. B. Moore Jr., *Privacy: Studies in Social and Cultural History: Studies in Social and Cultural History* (Sharpe, 1984).
4. P. Aries, G. Duby, *A History of Private Life: From Pagan Rome to Byzantium* (Belknap, 1987), vol. 1.
5. A. Westin, *Privacy and Freedom* (Simon & Schuster, 1967), chapter 1.
6. A. Acquisti, L. Brandimarte, G. Loewenstein, *Science* **347**, 509 (2015).
7. P. H. Klopfer, D. I. Rubenstein, *J. Soc. Issues* **33**, 52 (1977).
8. J. H. Barkow, *Darwin, Sex, and Status: Biological Approaches to Mind and Culture* (Univ. of Toronto Press, 1989).
9. R. F. Baumeister, M. R. Leary, *Psychol. Bull.* **117**, 497 (1995).
10. H. R. Varian, in *Privacy and Self-Regulation in the Information Age* (National Telecommunications and Information Administration, US Department of Commerce, 1996), chapter 1.
11. A. Acquisti, L. Brandimarte, J. T. Hancock, "Are There Evolutionary Roots To Privacy Concerns?" *Privacy Law Scholars Conference* (Berkeley, CA, 2013).
12. A. Shariff, J. Green, W. Jettinghoff, *Curr. Dir. Psychol. Sci.* **30**, 159 (2021).
13. A. Acquisti, L. Brandimarte, G. Loewenstein, *J. Consum. Psychol.* **30**, 736 (2020).
14. A. Warofka, "An Independent Assessment of the Human Rights Impact of Facebook in Myanmar," *Facebook* (2018; revised 2020); <https://about.fb.com/news/2018/11/myanmar-hria/>.
15. C. Utz, M. Degeling, S. Fahl, F. Schaub, T. Holz, (Un)informed consent: Studying GDPR consent notices in the field. *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, London, UK, November 2019.

## ACKNOWLEDGMENTS

The authors acknowledge support from the National Science Foundation through Awards 1228857 (Evolutionary Approaches to Privacy and Information Security, 2012) and 1514192 (Understanding and Exploiting Visceral Roots of Privacy and Security Concerns, 2015). A.A. acknowledges support from the Alfred P. Sloan Foundation through grant G-2015-14111 and from the Carnegie Corporation of New York via an Andrew Carnegie Fellowship. The authors are thankful for comments provided by the reviewers, as well as J. Bailenson, E. Carbone, D. Chang, J. Flagg, C. Hoofnagle, L. Jiang, L. John, G. Loewenstein, J. Spiegel, R. Steed, and by participants at several workshops (including PLSC 2013, SHB 2014, and WEIS 2015) and seminars.



## FOOD SYSTEMS

# Biodiversity, food, and culture

A loss to one is a loss to all

By **Lenore Newman**

In northern Tanzania near the shore of Lake Eyasi, a bird known as the honey-guide has struck a fruitful bargain with the Hadza people. The unassuming olive-green birds find honey hidden in the massive baobab trees that grow nearby, but they are no match for the bees. Their human collaborators, once led there by the birds, smoke the bees into submission. Working together, bird and human can share the spoils. But will this millennia-old cooperation come to an end as cattle farming encroaches on Hadza land?

Through beautiful stories of nature and culture such as this, veteran BBC reporter Dan Saladino asks critical questions about biodiversity loss and the future of food. *Eating to Extinction* is a sprawling wander through the world's food system with an urgent message. The modern food system rests on a persistent paradox: Although we depend on fewer and fewer breeds and cultivars for sustenance—we source 50% of our calories from just three crops: wheat, rice, and corn—we require the variety once found in our food species for both resilience and joy. *Eating to Extinction* firmly links biodiversity loss to cultural loss.

The book is a series of variations on a theme. After a brief (perhaps too brief) overview of the history of food, it unfolds in short acts in which the author elegantly explores a food and the people who sustain it. Key messages emerge as this narrative unfolds: We are rapidly depleting the wild and historical variation in our foodstuffs, critical elements of the food system persist only thanks to the efforts of a few dedicated individuals, and initiatives to bank biological diversity are fraught and tenuous.

Some territory has been covered widely before, such as wild foods and domesticated



Two Hadza men eat honey from combs in the Gideru Ridge region of Tanzania.

fruits, but Saladino distinguishes himself by exposing readers to lesser-known corners of these categories. Standouts include the Australian murnong root, which once underpinned aboriginal foodways, and rare local fruits lost in the sweep of globalization, including the Sicily vanilla orange and the Ugandan banana. The author then turns to the overlooked backbone of the food system: carbohydrates. Here, the reader is treated to an interplay of history and geography through crops such as Kavalita wheat, Orkney barley, and Oaxaca corn, complete with stories of seed vaults and landraces.

The breadth of *Eating to Extinction* expands with the animal section, which is less commonly explored territory. Saladino reminds us that we have only domesticated 14 animal species, and the big five—cattle, goats, pigs, sheep, and chickens—provide most terrestrial animal protein. I had never thought about why we do not eat zebra, but now I know they are too aggressive to domesticate.

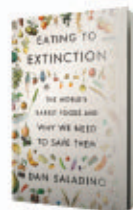
This chapter glossed over the ethical issues of the animal system. At this point in history, to uncritically discuss eating a sentient creature such as the whale falls flat. The pig and bison sections also felt a little rushed, but the book comes back to life in the fish section, where Saladino poses the paradox of the Atlantic salmon (wild fish are extremely rare yet farmed salmon are ubiquitous) and explores the cultural knowledge being threatened by overfishing of the Mauritania mullet.

The book's exploration of vegetable crops is a welcome counterpoint to the better-known

fruit species. Saladino's discussion of the Geechee red pea is the best exploration of the interplay between culture and vegetable crops since Joel Denker's *The Carrot Purple*. As he moves on to the Okinawa soybean, we are reminded how important soy is in the world food system, yet once again, a critical rare variety is tended by a single farmer.

The book's pivot to cheese, alcohol, and stimulants could appear tacked on, but this section includes some of Saladino's strongest cultural messages. Readers learn about the devastating impact of war on crops in a discussion of criollo cacao and wild Ethiopian coffee and how the great leap forward in China's pu-erh-producing region of southern Yunnan destroyed some of the oldest wild tea trees on the planet. Here, I was reminded that I can enjoy a nice Stilton from Neal's Yard in Borough Market with a sparkling glass of perry only because a handful of producers took the time and expended the energy to resurrect vanishing traditions.

*Eating to Extinction* covers such a vast swath of time, space, and biology that the epilogue is a bit of a letdown with its call to engage with the world of wild local food. There was room here to talk about policy and organized action to protect our culinary heritage, particularly at the governmental level. This is a minor point, however. Ultimately, Saladino's latest work is an important addition to the urgent study of the foods we know and love and to the threats that imperil them. ■



**Eating to Extinction**  
Dan Saladino  
Farrar, Straus and Giroux,  
2022. 464 pp.

The reviewer is at the Food and Agriculture Institute, University of the Fraser Valley, Abbotsford, BC, Canada, and the author of *Lost Feast: Culinary Extinction and the Future of Food* (ECW Press, 2019). Email: lenore.newman@ufv.ca

10.1126/science.abn1066



## NEUROSCIENCE

# Lost and found

A lyrical meditation on wayfinding offers cultural context and hope for the navigationally challenged

By **Marcia Bécu<sup>1,2</sup>** and **Christian F. Doeller<sup>1,2,3</sup>**

**A**n acquaintance comes to mind who gets lost after a few turns. Another can maintain, as if by magic, a good sense of orientation in any setting, whether a crowded metropolis or deep in the woods. Have you ever wondered why people differ so much when it comes to wayfinding? Or which brain mechanisms allow us to know where we are and the way to reach distant locations?

In *Dark and Magical Places*, Christopher Kemp reviews the major discoveries that have built this fascinating research field, from the seminal works of psychologist Edward C. Tolman, who worshipped mazes as the paragon of methods to reveal the brain's navigational secrets, to descriptions of the elegant cellular machinery, scattered across the brain, that renders navigation possible. Throughout the book, he tells the stories of the technological advances that accompanied relevant discoveries, from a simple bag of flour with holes in it tied to subjects while they navigated a maze to the most advanced virtual reality and neuroimaging techniques.

Through a series of interviews with key researchers, we learn that the hippocampus, a small area hidden deep in the brain, is empowered by place cells, which will always fire when we occupy a given location in space. By alternating between representations of the past (the well-known route), the present (where we are now), and the possible futures (paths that we have not yet traveled), place cells are thought to enable memory consolidation and mental

projection through time. Alongside the hippocampus, Kemp describes the “usual suspects” that form our navigation hub: the entorhinal cortex, the metric system that contains grid cells, the retrosplenial cortex, and much more. He also provides examples of clinical cases that help to emphasize the subtlety of the brain's spatial machinery, as well as what happens when one piece of the network fails.

Kemp's quest is not limited to contemporary neuroscientific insights; he also embarks on a journey through evolutionary time. He explains that symbolic thinking,



Practice can help those with a poor sense of direction improve their spatial abilities.

long-distance trading, and social connections helped to distinguish *Homo sapiens* from its early contemporaries. Through rich and amusing interviews with paleoneurologists and paleopsychologists, we learn that the spreading of modern humans across the planet is likely a reflection of our powerful spatial abilities—abilities made possible by the expansion of the parietal lobe.

Guided by the need to understand his own navigational shortcomings, Kemp charts evidence supporting the influence of both nature and nurture on our wayfinding skills. We were not born genetically equal in spatial aspects, he reveals. Genetic studies on twins have found that spatial skills—both poor and strong—are highly heritable and tend to cluster in individuals and in single families.

## Dark and Magical Places: The Neuroscience of Navigation

Christopher Kemp  
Norton, 2022. 256 pp.



Kemp concludes with an international examination of how cultures, environments, and languages affect navigation. Here, he describes the navigational superpowers of the Tsimane who navigate the Amazon rainforest and the linguistic singularities of the Tenejapan Mayans. The latter, we learn, do not rely on “left” and “right” as directional indicators but rather use “uphill” or “downhill,” for example, “Please pass the cup that is uphill.”

If you, like the author, feel as though you are among the “constantly lost,” this book will alleviate any feeling of social guilt. It also suggests ways to improve. First: Navigate. Spatial memory obeys the “use it or lose it” rule, and those of us who are too dependent on GPS technologies may run the risk of losing it. Second: Focus and look around. Paying attention to your environment is critical to successful navigation. Turn around, if you can, and have a look at your surroundings from different perspectives. Finally, identify a stable landmark and orient yourself relative to it at all times.

This tome is no textbook. It is designed to be accessible to a large audience: Nonscientists will benefit from Kemp's capacity to render vivid representations of the complexity of living organisms, but scientists, too, are likely to gain something from reading this book. Kemp offers context,

both historical and societal, for the scientific discoveries he presents, as well as some behind-the-scenes stories, as told to him by scientists themselves.

Spatial cognition is a rich and productive research domain, earning John O'Keefe, May-Britt Moser, and Edvard I. Moser the Nobel Prize in Physiology or Medicine in 2014 and garnering more than 1000 publications in 2020, but it is a topic that is too prolific to be considered in its entirety. Kemp wisely chooses to concentrate instead on the big picture, offering a poetical overview of the field that highlights how much of the magic the brain uses to represent space remains to be understood. ■

The reviewers are at the <sup>1</sup>Centre for Neural Computation, Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, and K.G. Jebsen Centre for Alzheimer's Disease, Kavli Institute for Systems Neuroscience, Norwegian University of Science and Technology, Trondheim, Norway; <sup>2</sup>Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; and <sup>3</sup>Wilhelm Wundt Institute of Psychology, Leipzig University, Leipzig, Germany. Email: marcia.becu@ntnu.no

## LETTERS



The Critically Endangered spotted tree frog is not included in Australia's priority threatened species list.

Edited by **Jennifer Sills**

## Australia's biodiversity crisis and opportunity

Australia is failing to meet its international obligations to conserve its unique native biodiversity and ecosystems. Most of Australia's plants and animals are found nowhere else on Earth, but since colonization about 230 years ago, at least 100 endemic species have been driven to extinction (1), and 17 ecosystems spanning the continent are now showing signs of collapse (2). Many more species face the same grim fate, with more than 1900 species and ecological communities currently listed as of conservation concern (3) under Australia's centerpiece environmental legislation, the Environment Protection and Biodiversity Conservation Act 1999. Numerous reports (4) demonstrate that Australia is simply not doing enough to address key threats to biodiversity, including land clearing and urbanization, invasive species, altered fire regimes, pollution, disease, and climate change. Despite being a member of the G20, Australian federal and state government environmental spending is well short of what's required (5) to reverse the nation's biodiversity extinction trajectory.

A stark example of this failure is the newly announced priority threatened species list. Just 100 threatened species (6)—fewer than 6% of the country's listed threatened species—are earmarked for conservation attention and AUS \$10 million of new funding, equating to about \$100,000 per species. Of Australia's Critically Endangered or Endangered species, only 2 of 25 frog species (8%), 7 of 53 invertebrate species (13.2%), and 28 of 776 plant species (3.6%) make the priority list (3, 6).

Stronger environmental laws, combined with a substantial increase in investment in environmental and conservation spending, will not only benefit Australia's biodiversity but also undoubtedly deliver substantial social, cultural, and economic benefits (7). The international community is moving to implement a new post-2020 global biodiversity framework, and heads of state recently met at the United Nations Climate Change Conference to chart a course to avert the climate change crisis deepening. Australia must be a leader of change, not a laggard.

**Euan G. Ritchie**

School of Life and Environmental Sciences,  
Centre for Integrative Ecology, Deakin  
University, Burwood, VIC 3125, Australia.  
Email: e.ritchie@deakin.edu.au

## REFERENCES AND NOTES

1. J. C. Z. Woinarski *et al.*, *Biol. Conserv.* **239**, 108261 (2019).
2. D. M. Bergstrom, *Glob. Change. Biol.* **27**, 1692 (2021).
3. Department of Agriculture, Water, and the Environment, Australian Government, "Species Profile and Threats Database" (1999); <http://environment.gov.au/cgi-bin/sprat/public/sprat.pl>. To find the number of threatened species and communities: Under "Other EPBC Act Lists," click "Threatened Fauna" (545 species in total, minus 76 Extinct, Extinct in the Wild, and Conservation Dependent species), "Threatened Flora" (1399 species in total, minus 37 Extinct species), and "Ecological Communities" (95 total), for a total of 1926 species and ecological communities. To find total threatened frogs, invertebrates, and plants: Under "Other EPBC Act Lists," click "Threatened Fauna" (25 frogs and 53 "other animals" categorized as Critically Endangered or Endangered) and "Threatened Flora" (776 species categorized as Critically Endangered or Endangered).
4. Australia State of the Environment 2016 (2018); <https://soe.environment.gov.au/>.
5. B. A. Wintle *et al.*, *Conserv. Lett.* **12**, e12682 (2019).
6. Department of Agriculture, Water, and the Environment, Australian Government, "100 Priority Species" (2021); [www.awe.gov.au/environment/biodiversity/threatened/publications/100-priority-species](http://www.awe.gov.au/environment/biodiversity/threatened/publications/100-priority-species).
7. "Nature as a climate solution: Country, culture and nature-based solutions for mitigating climate change" (Conservation Futures, 2021); <http://conservationfutures.org.au/publications/>.

10.1126/science.abn5705

## Protect the Amazon's Indigenous lands

The dramatic increase in the Brazilian Legal Amazon (BLA) deforestation rates since 2019 (1, 2) contrasts with Brazil's goal to achieve zero illegal deforestation in 2028 (3). The 2021 BLA official deforestation rate was the highest in 15 years. The total deforested area in 2021 (13,235 km<sup>2</sup>) is 69% above the average annual deforestation rate from 2012 to 2021 (4). This increase in deforestation affects not only private lands but also protected areas, including Indigenous lands.

Indigenous lands act as "shields" against deforestation (5). Alarming, deforestation within these lands is increasing. The annual average deforestation rate within Indigenous lands in the past 3 years (419 km<sup>2</sup>) is 80.9% above the average annual rate from 2012 to 2021 (4).

One example of this noticeable increase is the Apyterewa Indigenous land, located in the state of Pará. A deforestation incursion started in 2018 (6); however, no effective law enforcement actions were taken to stop land grabbers. Consequently, deforestation in Apyterewa spiked in 2021, accounting for 20.7% of the deforestation within all Indigenous lands (4). The increase in mining and smoke pollution that follows deforestation puts at risk Indigenous peoples (7, 8), especially the isolated ones. It also violates International Labour Organization Convention 169, of



which Brazil is a signatory. This agreement guarantees the participation of Indigenous peoples in the management and conservation of their territories (9).

The conservation of Indigenous lands is paramount for honoring Brazil's legal commitments, maintaining Amazonian environmental stability, fighting climate change, and guaranteeing traditional peoples' well-being. The existence of laws for preserving Amazonian remaining forests and traditional peoples' rights is not enough. Effective law enforcement actions are required to protect the last intact and preserved Amazon frontiers (10, 11).

Guilherme Mataveli\* and Gabriel de Oliveira<sup>2</sup>

<sup>1</sup>Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), São José dos Campos, SP, 12227-010, Brazil.

<sup>2</sup>Department of Earth Sciences, University of South Alabama, Mobile, AL 36688, USA.

\*Corresponding author.

Email: mataveli@alumni.usp.br

#### REFERENCES AND NOTES

1. C. H. L. Silva Junior et al., *Nat. Ecol. Evol.* **5**, 144 (2021).
2. H. Escobar, *Science* **10.1126/science.aba3238** (2019).
3. Brazilian Ministry of Environment (MMA), "Confira a íntegra do discurso do ministro Joaquim Leite na plenária da Cúpula do Clima" (2021); [www.gov.br/mma/pt-br/noticias/confira-a-integra-do-discurso-do-ministro-joaquim-leite-na-plenaria-da-cupula-do-clima](http://www.gov.br/mma/pt-br/noticias/confira-a-integra-do-discurso-do-ministro-joaquim-leite-na-plenaria-da-cupula-do-clima) [in Portuguese].
4. TerraBrasilis, PRODES (Deforestation), Analyses—Legal Amazon (2021); [http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/biomes/legal\\_amazon/rates](http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/biomes/legal_amazon/rates). Annual deforestation rates in the BLA: "Deforestation rates—Legal Amazon—States" bar graph. Total deforestation by year in Indigenous lands: Select "Legal Amazon" under "Deforestation increments" on the left. Then select "Indigenous Areas" at the top and see the bar graph labeled "Deforestation increments—Legal Amazon—Indigenous Areas."
5. P. F. R. Paiva et al., *Biodivers. Conserv.* **29**, 19 (2019).
6. G. de Oliveira et al., *Forests* **11**, 829 (2020).
7. A. C. Rorato et al., *Environ. Res. Lett.* **15**, 1040a3 (2020).
8. G. de Oliveira et al., *Science* **369**, 634 (2020).
9. International Labour Organization (ILO), "Indigenous and Tribal Peoples Convention (No. 169)" (1989); [www.ilo.org/dyn/normlex/en/?p=NORMLEX\\_PUB:12100:0:NO:12100:P12100\\_ILO\\_CODE:C169](http://www.ilo.org/dyn/normlex/en/?p=NORMLEX_PUB:12100:0:NO:12100:P12100_ILO_CODE:C169).
10. R. Trancoso, *Environ. Res. Lett.* **16**, 041004 (2021).
11. K. V. Conceição et al., *Land Use Pol.* **108**, 105663 (2021).

10.1126/science.abn4936

## Mining and Brazil's Indigenous peoples

Mining companies have 3481 requests pending with Brazil's National Mining Agency for permission to prospect in the country's Indigenous lands, even though it is currently illegal (1). The companies submitting the requests will gain priority for approval if mining in these areas is made legal in the future. The Bolsonaro presidential administration has curtailed inspections and greatly weakened protection of Indigenous lands against invasion (2). The frequency and scale of assaults on Indigenous land, especially from mining,

have increased greatly (3). Urgent judicial action, or credible threats of judicial action, could stop this trend.

On 5 December 2021, investigative journalists discovered that Brazil's federal government had authorized seven gold-mining projects in the "Dog's Head" (cabeça do cachorro), an area inhabited by 23 Indigenous peoples in the northwestern corner of the country (4). Brazil's federal Public Ministry initiated an investigation into the authorizations, and on 27 December the authorizations were canceled (5). The prosecutors in the federal Public Ministry believe that the mining authorizations in the Dog's Head were preparation for the opening of Indigenous lands throughout Amazonia to mining (6), a change that would become legal once the National Congress approves a bill that would open Indigenous land to mining, dams, and agribusiness (PL 191/2020) (7). President Bolsonaro, who submitted the bill, has requested special priority for it from the coalition of political parties that supports him and now controls both houses of Congress (8).

More than 20,000 "wildcat" gold miners (*garimpeiros*) have invaded the Yanomami Indigenous land, located just northeast of the Dog's Head (9). Both the Yanomami Indigenous land and the Dog's Head are also under the threat of corporate mining based on the many pending requests for mining licenses. Both of these areas contain isolated Indigenous peoples who are not in contact with the majority of Brazilian society (7). Isolated peoples are in particular danger given the insufficiently protective policies of Brazil's agency for Indigenous affairs (10). On 8 December 2021, Brazil's Ministry of Women, Family, and Human Rights began a process to classify cattle ranchers and wildcat gold miners as "traditional peoples," which would allow them to legally remain in conservation units (protected areas for biodiversity) and, potentially, in Indigenous lands (11).

Brazil's Public Ministry must act quickly to obtain judicial orders revoking the mining clearances. Judicial orders will also be needed to induce the presidential administration to remove illegal miners from the Yanomami and many other Indigenous lands. The countless actions of the Bolsonaro presidential administration in violation of Indigenous rights can no longer be ignored by the International Criminal Court in The Hague, where multiple cases remain pending (12).

Lucas Ferrante<sup>1\*</sup> and Philip M. Fearnside<sup>2</sup>

<sup>1</sup>National Institute for Research in the Amazon (INPA), Ecology Graduate Program, 69060-001, Manaus, AM, Brazil. <sup>2</sup>National Institute for

Research in the Amazon (INPA), 69067-375

Manaus, AM, Brazil.

\*Corresponding author.

Email: lucasferrante@hotmail.com

#### REFERENCES AND NOTES

1. F. Amato, "Pedidos de pesquisa mineral em terra indígena são quase 3,5 mil, embora atividade seja proibida." *G1* (2020) [in Portuguese].
2. M. M. Vale et al., *Biol. Conserv.* **255**, 108994 (2021).
3. L. Ferrante, P. M. Fearnside, *Die Erde* **152**, 200 (2021).
4. V. Sassini, "General Helene autoriza avanço de garimpo em áreas preservadas na Amazônia," *Folha de São Paulo* (2021) [in Portuguese].
5. "Governo recua e cancela autorizações de garimpo na Amazônia," *Folha de São Paulo* (2021) [in Portuguese].
6. V. Sassini, "MPF suspeita que atos de Helene buscam preparar terreno para mineração em terra indígena," *Folha de São Paulo* (2021) [in Portuguese].
7. S. Villén-Pérez, L. Anaya-Valenzuela, D. Conrado da Cruz, P. M. Fearnside, *Glob. Environ. Change* **72**, 102398 (2022).
8. M. Angelo, "Jair Bolsonaro pede a Arthur Lira prioridade na aprovação do PL que libera mineração em terras indígenas," *Observatório da Mineração* (2021) [in Portuguese].
9. A. Athila, C. Zacchini, "Yanomamis revivem ameaça de extermínio com garimpo e omissão governamental," *Folha de São Paulo* (2021) [in Portuguese].
10. E. S. Bigio, L. L. Santos, T. Moreira, "Índigenas isolados podem ser exterminados com o avanço da Funai," *Folha de São Paulo* (2021) [in Portuguese].
11. C. Prizibiszki, "Governo tenta incluir garimpeiros e pecuaristas na lista de comunidades tradicionais," *O Eco* (2021) [in Portuguese].
12. Articulação dos Povos Indígenas do Brasil, International "Complaints Dossier of Brazil's Indigenous Peoples" (2021); [https://apiboficial.org/files/2021/08/DOSSIE\\_en\\_v3web.pdf](https://apiboficial.org/files/2021/08/DOSSIE_en_v3web.pdf).

10.1126/science.abn6753

#### TECHNICAL COMMENT ABSTRACTS

##### Comment on "The influence of juvenile dinosaurs on community structure and diversity"

Roger B. J. Benson, Caleb M. Brown, Nicolás E. Campione, Thomas M. Cullen, David C. Evans, Lindsay E. Zanno

Schroeder et al. (Reports, 26 February 2021, p. 941) reported a size gap among predatory dinosaur species. We argue that the supporting dataset is skewed toward Late Cretaceous North America and that the gap was likely absent during other intervals in most geographic regions. We urge broader consideration of this hypothesis, with quantitative evaluation of preservational and dataset biases.

Full text: [dx.doi.org/10.1126/science.abj5976](https://doi.org/10.1126/science.abj5976)

##### Response to Comment on "The influence of juvenile dinosaurs on community structure and diversity"

Katlin M. Schroeder, S. Kathleen Lyons, Felisa A. Smith

The analysis of dinosaur ecology hinges on the appropriate reconstruction and analysis of dinosaur biodiversity. Benson et al. question the data used in our analysis and our subsequent interpretation of the results. We address these concerns and show that their reanalysis is flawed. Indeed, when occurrences are filtered to include only valid taxa, their revised dataset strengthens our earlier conclusions.

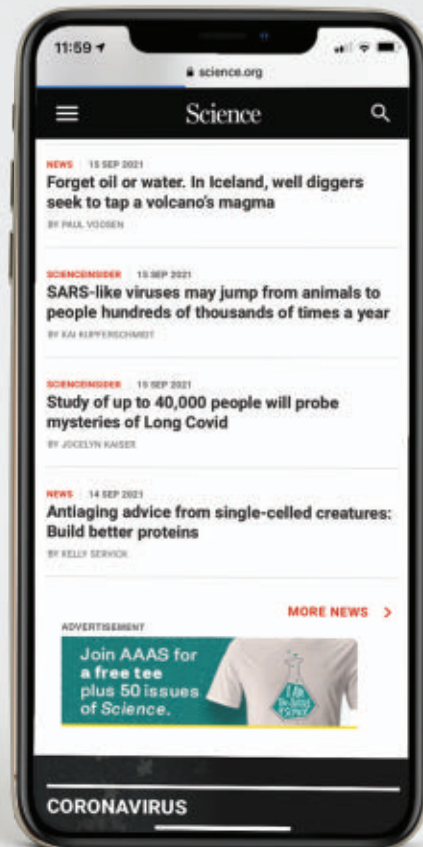
Full text: [dx.doi.org/10.1126/science.abj7383](https://doi.org/10.1126/science.abj7383)



# NEWS FROM Science

**Up-to-the-minute research and policy news you won't find in print**

Visit us online to read all the news coverage that there just wasn't enough room to print in this issue.



Science.org/news

# RESEARCH


## IN SCIENCE JOURNALS

Edited by Michael Funk

### CORAL REEFS

#### Protecting connectivity

**C**oral reefs are highly threatened by climate change and human pressure. Recent research has focused extensively on how to protect reefs from these impacts, but a fact that is often missed is that reefs are not isolated systems. Fish and coral larvae are actively exchanged across regions, and some reefs supply more whereas others receive more. Fontoura *et al.* looked at the source/sink dynamics of reefs globally and found that maintaining these networks, including dispersal corridors, is essential for biodiversity conservation and sustainable fisheries. Furthermore, they found that the majority of key source reefs and corridors remain unprotected. —SNV *Science*, abj8432, this issue p. 336



Protecting connectivity between regions is important for maintaining biodiversity and ecosystem services in coral reefs, such as this one in Indonesia.

### MOLECULAR MOTORS

#### Biphasic regulation of kinesin-1 by MAP7

Motor proteins are differentially regulated by microtubule-associated proteins (MAPs) to deliver intracellular cargos to their destinations. MAP7 recruits the molecular motor kinesin-1 to microtubules and activates subsequent motility, but the underlying mechanism is unclear. Using cryo-electron microscopy, Ferro *et al.* determined a near-atomic structure of MAP7 on the microtubule. They found that the microtubule-binding site of MAP7 overlaps with kinesin. However, by tethering kinesin-1 to the microtubule, the MAP7 projection domain

enabled the motor to diffuse to available sites on microtubules partially decorated by MAP7. These results reveal the mechanistic basis of MAP7-mediated activation of kinesin motility despite their competing binding to tubulin. —SMH

*Science*, abf6154, this issue p. 326

### SOLAR CELLS

#### Tailoring tin oxide layers

Mesoporous titanium dioxide is commonly used as the electron transport layer in perovskite solar cells, but electron transport layers based on tin(IV) oxide quantum dots could be more efficient, with a better-aligned conduction band and a higher carrier mobility. Kim *et al.* show that such quantum

dots could conformally coat a textured fluorine-doped tin oxide electrode when stabilized with polyacrylic acid. Improved light trapping and reduced nonradiative recombination resulted in a certified power-conversion efficiency of 25.4% and high operational stability. In larger-area minimodules, active areas as high as 64 square centimeters maintained certified power conversion efficiencies of more than 20%. —PDS

*Science*, abh1885, this issue p. 302

### STRUCTURAL BIOLOGY

#### Identifying ideal monoclonal antibodies

The emergence of new severe acute respiratory syndrome

coronavirus 2 (SARS-CoV-2) strains highlights the importance of identifying antigen-binding monoclonal antibodies rapidly and accurately. Existing methodologies are time-consuming and laborious. Antanasijevic *et al.* used detailed cryo-electron microscopy structural analysis of serum containing polyclonal antibodies binding to an HIV antigen in combination with next-generation sequencing of B cell receptors to identify HIV-specific monoclonal antibodies. The novel sequences identified from the polyclonal serum were validated using antigen binding and structural assays. This approach requires that B cell receptor sequences are available at the same time that serum is collected

for structural analysis and will prove valuable in situations in which there is rapid development of antibody-escape variants to known pathogens such as HIV or SARS-CoV-2. —PAM

*Sci. Adv.* 10.1126/sciadv.abk2039 (2022).

## CORONAVIRUS

### Veterans' virus affairs

The US Department of Veterans Affairs (VA) provides life-long health care to military personnel, and VA data offer a unique resource for public health surveillance in the United States. Cohn *et al.* examined data from over 780,000 individuals to track the effects of vaccination rollout as the Delta variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in the United States between February and September of 2021. Although messenger RNA (mRNA) and viral vector vaccines have effectively prevented clinically significant disease, SARS-CoV-2 Delta transmission has surged. Vaccine breakthrough infections have predominated in those given the viral vector vaccine but also occur among those given mRNA vaccines. By September 2021, the protection offered by the viral vector vaccine during the study declined to about 13% against infection and to about 50% against death. Unvaccinated individuals remained at the highest risk of infection, severe disease, and death. —CA

*Science*, abm0620, this issue p. 331

## PLANT SCIENCE

### Block to polyspermy

The next generation needs enough, but not too many, nuclear genomes. Zhong *et al.* show how the small mustard plant

*Arabidopsis* both blocks polyspermy and adds second-chance insurance when the first fertilization effort goes awry. Signals from the synergid cells in the female gametophyte invite a nearby pollen tube, which secretes peptides to block other pollen tubes from tagging along. The blockade persists as the pollen tube grows to its target. If the pollen tube successfully releases its pair of sperm (plants have a dual fertilization system), the pollen tube's signaling system fades to silence. If the female gametophyte successfully receives the sperm nuclei, invitational signals from the synergid cells also fade to silence. If, however, fertilization fails, the persistent signal from the synergid cells continues to attract pollen tubes and, because pollen tube rupture silenced the first pollen tube's block to polytuby, secondary pollen tubes are able to give another try at fertilization. —PJH

*Science*, abl4683, this issue p. 290

## CANCER

### A Yes to new therapies?

Patients with hepatocellular carcinoma (HCC) have few therapeutic options. The transcriptional coactivators YAP/TAZ are implicated in HCC development, although mutations in their genes are not associated with disease. Guégan *et al.* show that the activity of the tyrosine kinase Yes upstream of YAP/TAZ promotes HCC cell proliferation and tumor development in a manner dependent on the phosphorylation and nuclear translocation of YAP/TAZ. Data from HCC patients showed that increased Yes activity, rather than mutations in Yes or increased Yes abundance, was correlated with poor survival. —JFF

*Sci. Signal.* 15, eabj4743 (2022).

## IN OTHER JOURNALS

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



## HOST DEFENSE

### Biting into antiviral defense?

**T**he oral bacterium *Porphyromonas gingivalis* has been implicated in periodontal disease, and Rodriguez-Hernandez *et al.* report that it can also interfere with host antiviral responses in mice. The bacterium disrupts interferon (IFN)- $\lambda$  production by oral epithelial cells by the downregulation of IFN-activating transcription factors. Moreover, *P. gingivalis* promotes insensitivity to IFN by enhancing the proteolytic degradation of IFN receptors. Finally, *P. gingivalis* levels correlate with suppressed IFN responses in human oral tissues. Future studies will be needed to tease out whether *P. gingivalis* overabundance has clinically relevant effects on viral infection and transmission. —STS *Proc. Natl. Acad. Sci. U.S.A.* 118, e2105170118 (2021).

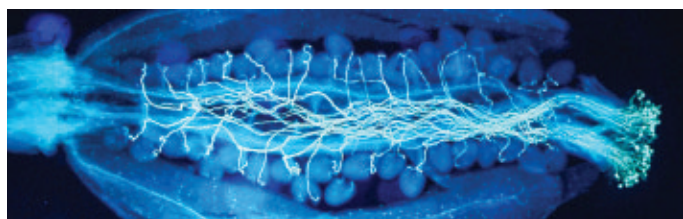
*Porphyromonas gingivalis*, shown here in a computer-created illustration, associates with periodontal disease and disrupts antiviral defenses.

## PHASE TRANSITION

### On the nucleation of quasicrystals

Quasicrystal structures are solids with symmetries of their atomic arrangements that are forbidden for periodic crystals. Despite decades of active research on quasicrystals and

their discovery in nature, the mechanism of quasicrystal nucleation remains unclear. Using high-index saddle dynamics applied to a Landau free-energy functional, Yin *et al.* obtained an interesting picture of critical nuclei and minimum energy paths that could shed light on how quasicrystals



Microscopy image of an *Arabidopsis* pistil filled with pollen tubes stretching from pollen grains (far right) toward individual ovules.



## SOCIAL MEDIA

## Does Twitter bend to the Right?

Concerns are increasing about political polarization and its potential for causing widespread social disruption in a world also facing environmental disruption. Political discussions form the bulk of exchanges on the social media platform Twitter. In 2016, Twitter introduced machine learning for ranking content tuned to individual preferences. Huszár *et al.* investigated whether this new mode of content provision selectively amplifies political flavors. This long-term experiment includes a control group free of algorithmic personalization. In tweets made by elected politicians from six of the seven countries studied, it appears the political Right had a louder voice than any other persuasion. Meanwhile, Levin *et al.* introduce recent interdisciplinary research, drawing inspiration from evolutionary theory and systems science, to explore the influences reshaping global politics in potentially dangerous directions. —CA

*Proc. Natl. Acad. Sci. U.S.A.* **119**, e2025334119 (2022); **118**, e2116950118 (2021).



emerge from the crystalline phase. The proposed methodology is general and applicable to other physical systems undergoing first-order phase transitions. —YS

*Proc. Natl. Acad. Sci. U.S.A.* **118**, 2106230118 (2021).

## CELL METABOLISM

## Probing the TCA cycle

The tricarboxylic acid (TCA) cycle is key to cellular metabolism, being responsible for generating ATP, the energy source for cellular processes. In addition, the TCA cycle produces precursors for synthesizing other biological molecules, including nonessential amino acids. Despite its central role, the consequences of TCA dysfunction are poorly understood. Ryan *et al.* investigated the consequences of inhibiting the TCA cycle in

murine kidney epithelial cells by either inhibiting or ablating two key enzymes. The authors monitored the effects of disruption using metabolomics, transcriptomics, and proteomics. An impaired TCA cycle alters amino acid and redox metabolism and activates the integrated stress response of the cell to rewire transcription and translation and compensate for the amino acid and redox stress. —VV

*eLife* **11**, e72593 (2021).

## STEM WORKFORCE

## Retention over recruitment

Retention in academia remains a persistent barrier to proportional representation. Shaw *et al.* used 25 years of National Science Foundation data on the racial and ethnic composition of

STEM academics, from undergraduates to tenured professors, to construct a null model of ethnic and racial representation in the US science community. Comparisons between the null model and current representation allowed them to measure the effects of retention while controlling for recruitment at different academic stages. Their results show that failed retention, occurring at different stages depending on race and ethnicity, negatively affected Black, Indigenous, and Hispanic scholars the most. A substantial drop in racial/ethnic representation between students (graduate and undergraduate) and researchers (postdocs and faculty) was found, suggesting that recruiting diverse students is not enough, and an increased

focus on inclusion/retention is needed. —MMc

*PLoS ONE* **16**, e0259710 (2021).

## ORGANIC CHEMISTRY

## Copper lights up acid coupling

Carboxylic acids are relatively abundant feedstocks for the preparation of more elaborate compounds in fine chemical, pharmaceutical, and agrochemical synthesis. Recently, transformation of these acids into redox-active esters has enabled efficient coupling chemistry using light-activated catalysts. Li *et al.* report that a simple copper salt can take the place of both the ester group and the catalyst, single-handedly oxidizing the acid for decarboxylative coupling under blue light. A variety of sulfonamides, amides, and alcohols served as effective coupling partners. —JSY

*Nat. Chem.* 10.1038/s41557-021-00834-8 (2022).

## CELL TRACKING

## Let ELEPHANT do the work

Tracking individual cells and their progeny through development is a labor-intensive and technically challenging task, and it often takes months to analyze a single dataset. Sugawara *et al.* describe a setup they call efficient learning using sparse human annotations for nuclear tracking (ELEPHANT), which performs cell tracking in three dimensions (3D) with minimal user input. ELEPHANT optimizes incremental deep learning using sparse annotations to detect nuclei in 3D and then links these nuclei through time in 4D image datasets. ELEPHANT successfully tracked cell lineages during embryonic development in nematodes, during limb regeneration in a crustacean, as well as in human intestinal organoids and breast carcinoma cells. All in all, data analysis time was reduced from months to weeks. —SMH

*eLife* **11**, e69380 (2022).

ALSO IN *SCIENCE* JOURNALS

Edited by Michael Funk

## CORONAVIRUS

## Evasive maneuvers by the spike protein

Throughout the course of the COVID-19 pandemic, variants have arisen in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus that increase infectivity or reduce its susceptibility to existing antibodies. Nabel *et al.* focus on mutations in the spike protein, which is found on the viral surface and is responsible for binding and entering host cells, and show that the structure exhibits plasticity in resisting neutralizing antibodies. Working with pseudoviruses, the authors identify combinatorial mutations that might lead to resistance. They identify an antibody that neutralizes a pseudotype with a highly mutated spike, but also show that SARS-CoV-2 can acquire a glycan to escape this neutralization. Understanding the consequences of changes in the antigenic landscape of the spike protein is important if we are to rapidly respond to new variants of concern. —VV

*Science*, abj6251, this issue p. 282

## PROTEIN DESIGN

## Designing dynamic protein complexes

Protein complexes play important roles in biological processes, and many complexes are dynamic, with subunits exchanging to facilitate different functions. It has been challenging to design stable and soluble monomeric proteins that reversibly associate into hetero-oligomers. Sahtoe *et al.* used a strategy called implicit negative design to construct proteins with interaction interfaces that drive association with a selected partner but not self-association. The resulting designs are stably folded in solution and provide the modules for assembly into a wide variety of complexes. They can be functionalized, allowing

target proteins to be displayed in defined geometries, and complex subunits can be exchanged by varying the available concentrations of components. —VV

*Science*, abj7662, this issue p. 283

## SYNTHETIC BIOLOGY

## Building synthetic cell fate selection

An important goal for synthetic biology is to establish control systems that allow the direction of cells into multiple stable states, much like biological signaling systems do during organismal development. Zhu *et al.* devised a system that allows such control through designed zinc finger transcription factors that interact with one another through homo- and heterodimerization and can be regulated by small molecules that control dimerization and stability of the transcription factors (see the Perspective by Kunze and Khalil). Mathematical modeling allowed computational prediction of the system's behavior, and the introduction of three designed transcription factors in cultured mammalian cells allowed direction of the cells into seven distinct, stable states. Understanding such multistability is useful in synthetic biology and can help to determine its roles in development and disease processes. —LBR

*Science*, abg9765, this issue p. 284;

see also abn6548, p. 262

## ATTOSECOND SCIENCE

## Coherent electron motion in real time

Charge transfer plays a fundamental role in many chemical and biological processes, yet many important questions about its mechanics at the electronic level remain unanswered. Recent development of attosecond x-ray free-electron laser sources have enabled site-specific valence excitations in molecular systems

on a time scale shorter than the natural charge dynamics. Li *et al.* use this technology to produce coherent superpositions of core-excited states in the model molecule nitric oxide, which is composed of biorelevant atoms. The authors map the time-dependent current of the Auger-Meitner emission by means of an angular streaking measurement. The present work demonstrates the ability to use x-ray free-electron lasers for exploring electronic coherences in the charge transfer on the attosecond time scale. —YS

*Science*, abj2096, this issue p. 285

## MULTIPLE SCLEROSIS

## Stronger evidence for viral connection

Multiple sclerosis is a chronic demyelinating disease of the central nervous system. The underlying cause of this disease is not known, but Epstein-Barr virus is thought to be a possible culprit. However, most people infected with this common virus do not develop multiple sclerosis, and it is not feasible to directly demonstrate causation of this disease in humans. Using data from millions of US military recruits monitored over a 20-year period, Bjornevik *et al.* determined that Epstein-Barr virus infection greatly increased the risk of subsequent multiple sclerosis and that it preceded the development of disease, supporting its potential role in the pathogenesis of multiple sclerosis (see the Perspective by Robinson and Steinman). —YN

*Science*, abj8222, this issue p. 296;

see also abm7930, p. 264

## PEROVSKITES

## Perovskite nanocrystals under glass

Perovskite nanocrystals (PNCs) such as cesium lead triiodide ( $\text{CsPbI}_3$ ) can display bright photoemission with narrow

linewidths for display applications, but their long-term stability requires passivation and encapsulation steps after synthesis in solution. Sun *et al.* created three-dimensional arrays of PNCs in doped metal oxide glasses using ultrafast laser pulses that caused local melting and subsequent crystallization. They tuned the bandgap of PNCs and their photoluminescence between 480- and 700-nanometer wavelengths by transforming the composition from  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x)_3$  to  $\text{CsPbI}_3$ . These encapsulated PNCs exhibited long-term stability after prolonged heating or organic solvent and ultraviolet light exposure. —PDS

*Science*, abj2691, this issue p. 307

## SOCIAL PSYCHOLOGY

## A kiss tells the tale

Young humans are remarkably helpless, relying entirely on the adult humans around them for survival. However, not all adults are as invested in the care of a particular child, and there is benefit in being able to determine from a very young age which relationships are close. Thomas *et al.* tested young children and infants to determine whether they were able to identify close, or “thick,” relationships based on whether individuals participated in activities that involve sharing saliva, such as eating, kissing, or sharing utensils (see the Perspective by Fawcett). The children expected relationships like these to be closer than other relationships, indicating that they can distinguish closeness very early in life. —SNV

*Science*, abh1054, this issue p. 311;

see also abn5157, p. 260

## RESEARCH TECHNOLOGY

## Sorting cells by intracellular features

Fluorescence-activated cell sorting, reported in *Science* 52 years ago, has revolutionized

biomedical research, giving us the ability to isolate cells according to the expression of labeled proteins. So far, however, flow cytometric cell sorting has been blind to spatial processes such as intracellular protein localization, which is traditionally measured using microscopy. Schraivogel *et al.* combined ultrafast microscopy and image analysis with a flow cytometric cell sorter to unlock spatial phenotypes for high-throughput sorting applications. The authors show how this technology can be used to rapidly isolate cells with complex cellular phenotypes and how it can accelerate genome-wide microscopy-based CRISPR screening. —DJ

*Science*, abj3013, this issue p. 315

## 2D MATERIALS

### Following a crossover

Superfluidity in fermionic systems occurs through the pairing of fermions into bosons, which can undergo condensation. Depending on the strength of interactions between fermions, the pairs range from large and overlapping to tightly bound. The crossover between these two limits has been explored in ultracold Fermi gases. Liu *et al.* observed the crossover in an electronic system consisting of two layers of graphene separated by an insulating barrier and placed in a magnetic field (see the Perspective by Morgenstern and Goerbig). In this two-dimensional (2D) system, the pairs were excitons formed from an electron in one layer and a hole in the other. The researchers used magnetic field and layer separation to tune the interactions and detected the signatures of superfluidity through transport measurements. —JS

*Science*, abm3770, this issue p. 321;  
see also abn2049, p. 263

## T CELLS

### Looping in FoxP3

The identity and function of T regulatory cells ( $T_{reg}$ ) relies on the activity of the transcription factor FoxP3, but its precise

mechanism of action in controlling  $T_{reg}$ -specific gene expression is not well understood. Using chromatin conformation capture with immunoprecipitation, Ramirez *et al.* mapped the enhancer-promoter architecture of conventional CD4<sup>+</sup> T cells and  $T_{reg}$ s and then identified FoxP3-interacting enhancer-promoter loops (EPLs). FoxP3 interacted with EPLs at core  $T_{reg}$  signature genes and was associated with increased enhancer-promoter connectivity, whereas genetic inactivation of FoxP3 resulted in decreased H3K27Ac looping at the same loci. These results provide insight into FoxP3's interactions with EPLs and support a model in which FoxP3 directly regulates the expression of many of its target genes through enhancer-promoter connections. —CO

*Sci. Immunol.* **6**, eabj9836 (2022).

## MUSCLE DISEASE

### T cell-independent muscle loss

Although recent data suggest the involvement of autoimmune mechanisms in the pathophysiology of sporadic inclusion body myositis (IBM), the role of T cell-mediated autoimmunity remains to be elucidated. Britson *et al.* developed a humanized xenograft rodent model of IBM that recapitulates the main hallmarks of the disease. The authors showed that T cell depletion was able to reduce major histocompatibility complex class I (MHC-I) upregulation within myofibers but did not reduce pathology of the transcriptional protein TDP-43 or rimmed vacuole formation. The results indicate that in this xenograft model, loss of TDP-43 function and muscle degeneration in IBM occur independently of T cell infiltration. —MM

*Sci. Transl. Med.* **14**, eabi9196 (2022).

## CORONAVIRUS

### Neurologic effects of COVID-19

Although predominantly a respiratory virus, it has become

clear that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can affect numerous organ systems, including the brain. Those with acute COVID-19 report various neuropsychiatric syndromes that can persist, and new symptoms can develop in those with so-called Long Covid. In a Perspective, Spudich and Nath discuss the neurological symptoms that seem to come with SARS-CoV-2 infection and the possible mechanisms, including neuroinflammation, autoimmunity, and vascular dysfunction, that are also associated with increased risk of stroke. It is important to understand the underlying pathology of neurological dysfunction caused by SARS-CoV-2 infection to prevent possible long-term sequelae. —GKA

*Science*, abm2052, this issue p. 267

## SEXUAL SELECTION

### The power of choice

In his book *The Descent of Man, and Selection in Relation to Sex*, Darwin proposed the concept of sexual selection, specifically that the vagaries of choice during reproduction could also shape patterns of adaptation—and thus evolution. Despite this groundbreaking thinking, his Victorian ideas about women affected his ability to see just how influential mate choice, especially from the female side, could be. Rosenthal and Ryan review progress that has been made in this area in the 150 years since Darwin first proposed sexual selection, paying special attention to how more equitable thinking has led to significant scientific advances in our understanding. —SNV

*Science*, abi6308, this issue p. 281



## REVIEW SUMMARY

## SEXUAL SELECTION

Sexual selection and the ascent of women:  
Mate choice research since Darwin

Gil G. Rosenthal\* and Michael J. Ryan\*

**BACKGROUND:** Charles Darwin's *Descent of Man, and Selection in Relation to Sex* tackled the two main controversies arising from the *Origin of Species*: the evolution of humans from animal ancestors and the evolution of sexual ornaments. Most of the book focuses on the latter, Darwin's theory of sexual selection. Research since supports his conjecture that songs, perfumes, and intricate dances evolve because they help secure mating partners. Evidence is overwhelming for a primary role of both male and female mate choice in sexual selection—not only through premating courtship but also through intimate interactions during and long after mating.

But what makes one prospective mate more enticing than another? Darwin, shaped by misogyny and sexual prudery, invoked a “taste for the beautiful” without speculating on the origin of the “taste.” How to explain when the “final marriage ceremony” is between two rams? What of oral sex in bats, cloacal rubbing in bonobos, or the sexual spectrum in humans, all observable in Darwin's time? By explaining desire through the lens of those male traits that caught his eyes and those of his gender and culture, Darwin elided these data in his theory of sexual evolution.

Work since Darwin has focused on how traits and preferences coevolve. Preferences can evolve even if attractive signals only predict offspring attractiveness, but most attention has gone to the intuitive but tenuous premise that mating with gorgeous partners yields vigorous offspring.

By focusing on those aspects of mating preferences that coevolve with male traits, many of Darwin's influential followers have followed the same narrow path. The sexual selection debate in the 1980s was framed as “good genes versus runaway”: Do preferences coevolve with traits because traits predict genetic benefits, or simply because they are beautiful? To the broader world this is still the conversation.

**ADVANCES:** Even as they evolve toward ever-more-beautiful signals and healthier offspring, mate-choice mechanisms and courter traits are locked in an arms race of coercion and resistance, persuasion and skepticism. Traits favored by sexual selection often do so at the expense of chooser fitness, creating sexual con-

flict. Choosers then evolve preferences in response to the costs imposed by courtiers.

Often, though, the current traits of courtiers tell us little about how preferences arise. Sensory systems are often tuned to nonsexual cues like food, favoring mating signals resembling those cues. And preferences can emerge simply from selection on choosing conspecifics. Sexual selection can therefore arise from chooser biases that have nothing to do with ornaments.

Choice may occur before mating, as Darwin emphasized, but individuals mate multiple

times and bias fertilization and offspring care toward favored partners. Mate choice can thus occur in myriad ways after mating, through behavioral, morphological, and physiological mechanisms.

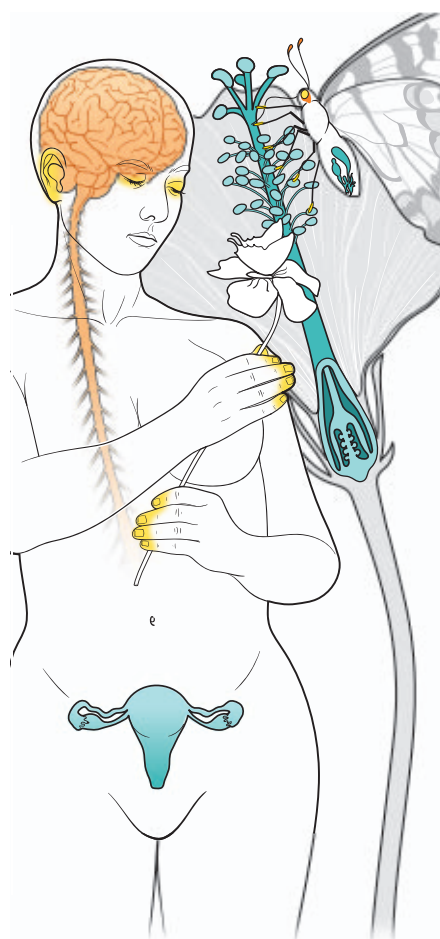
Like other biological traits, mating preferences vary among individuals and species along multiple dimensions. Some of this is likely adaptive, as different individuals will have different optimal mates. Indeed, mate choice may be more about choosing compatible partners than picking the “best” mate in the absolute sense. Compatibility-based choice can drive or reinforce genetic divergence and lead to speciation.

The mechanisms underlying the “taste for the beautiful” determine whether mate choice accelerates or inhibits reproductive isolation. If preferences are learned from parents, or covary with ecological differences like the sensory environment, then choice can promote genetic divergence. If everyone shares preferences for attractive ornaments, then choice promotes gene flow between lineages.

**OUTLOOK:** Two major trends continue to shift the emphasis away from male “beauty” and toward how and why individuals make sexual choices. The first integrates neuroscience, genomics, and physiology. We need not limit ourselves to the feathers and dances that dazzled Darwin, which gives us a vastly richer picture of mate choice. The second is that despite persistent structural inequities in academia, a broader range of people study a broader range of questions.

This new focus confirms Darwin's insight that mate choice makes a primary contribution to sexual selection, but suggests that sexual selection is often tangential to mate choice. This conclusion challenges a persistent belief with sinister roots, whereby mate choice is all about male ornaments. Under this view, females evolve to prefer handsome males who provide healthy offspring, or alternatively, to express flighty whims for arbitrary traits. But mate-choice mechanisms also evolve for a host of other reasons

Understanding mate choice mechanisms is key to understanding how sexual decisions underlie speciation and adaptation to environmental change. New theory and technology allow us to explicitly connect decision-making mechanisms with their evolutionary consequences. A century and a half after Darwin, we can shift our focus to females and males as choosers, rather than the gaudy by-products of mate choice. ■

**Mate choice mechanisms across domains of life.**

Sensory periphery for stimulus detection (yellow), brain for perceptual integration and evaluation (orange), and reproductive structures for postmating choice among pollen or sperm (teal).

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

\*Corresponding author. Email: gil.rosenthal@unipd.it (G.G.R.); mryan@utexas.edu (M.J.R.)

Cite this article as G. G. Rosenthal, M. J. Ryan, *Science* 375, eabi6308 (2022). DOI: 10.1126/science.abi6308

**S READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abi6308>

## REVIEW

## SEXUAL SELECTION

# Sexual selection and the ascent of women: Mate choice research since Darwin

Gil G. Rosenthal<sup>1,2\*</sup> and Michael J. Ryan<sup>3,4\*</sup>

Darwin's theory of sexual selection fundamentally changed how we think about sex and evolution. The struggle over mating and fertilization is a powerful driver of diversification within and among species. Contemporaries dismissed Darwin's conjecture of a "taste for the beautiful" as favoring particular mates over others, but there is now overwhelming evidence for a primary role of both male and female mate choice in sexual selection. Darwin's misogyny precluded much analysis of the "taste"; an increasing focus on mate choice mechanisms before, during, and after mating reveals that these often evolve in response to selection pressures that have little to do with sexual selection on chosen traits. Where traits and preferences do coevolve, they can do so whether fitness effects on choosers are positive, neutral, or negative. The spectrum of selection on traits and preferences, and how traits and preferences respond to social effects, determine how sexual selection and mate choice influence broader-scale processes like reproductive isolation and population responses to environmental change.

One hundred and fifty years ago, Charles Darwin published his second great book, *The Descent of Man, and Selection in Relation to Sex* (1). Darwin's book on sexual selection is traditionally viewed as a supplement to *On the Origin of Species*, tackling two distinct topics from his theory that were steeped in controversy. The societal influence of *The Descent of Man, and Selection in Relation to Sex* has perhaps been as far-reaching as Darwin's natural selection book. Although our descent from a nonhuman ancestor is settled science, sexual selection—specifically through mate choice—continues to fascinate (2) and frustrate (3).

## Mate choice: The fuzzy center of Darwin's theory

This sesquicentennial offers an appropriate time to reevaluate "Darwin's really dangerous idea" (4). Darwin's evolutionary science and abolitionist politics (5) were animated by the "contingent fact" of human equality (6). Published on the heels of the US war to abolish slavery, *The Descent of Man, and Selection in Relation to Sex* (*The Descent*) is perhaps the best anyone could have done on this score given that Darwin was steeped in the structural racism of his time, tribe, and place. Darwin believed in the manifest superiority of white Anglo-Saxon Protestants and the inferiority of "savages" like the Fuegians. Critically, however, Darwin emphasizes that this superiority stems from factors that are (largely) not immuta-

ble, but rather dependent on culture and environment. He cites his personal experiences with Orundellico ("Jemmy Button") and John Edmonstone—"civilized" men from savages and slaves—to reinforce his point that heritable differences existed between ethnic groups but were largely due to the caprices of sexual selection acting on human populations.

As detailed in Richards (7), Darwin's sexual selection was a revolutionary tool to take on both the theological racism of polygenist creationists like Louis Agassiz, who posited separate-and-unequal Creations on each continent, along with the pseudoscientific racism of his white scientific contemporaries, who argued for the essential inferiority of Black and Brown people. In the first third of *The Descent*, Darwin argues that we all share a recent common ancestor and that our mental and moral differences are largely the product of culture and environment. In two chapters near the end, he argues that our physical differences—skin color, hair, and so forth—are just superficial by-products of sexual selection. These two parts of *The Descent* add up to a scientific case for, if not equality among humans, enough brotherhood among men to put chattel slavery beyond the moral pale even in a forthrightly white-supremacist society (7).

There is a lively debate over the nature and extent of Darwin's racism (8). By contrast, Darwin's essentialist misogyny—his belief that women were immutably inferior to men—is unambiguous. Here, Darwin argues for sexual selection not as a force driving superficial differences among groups, but rather as the driving force for the biological supremacy of males: "hence man has ultimately become superior to woman" (1), p. 565]. As we detail below, Darwin's dismissal of female agency

(9) and promiscuity (10) continues to shadow the field of sexual selection. Yet the broad logic of Darwin's theory stands today:

Sexual selection depends on the success of certain individuals over others of the same sex, in relation to the propagation of the species, whereas natural selection depends on the success of both sexes, at all ages, in relation to the general conditions of life. The sexual struggle is of two kinds; it is between the individuals of the same sex, generally the males, to drive away or kill their rivals, the females remaining passive—while it is also the struggle between individuals of the same sex, to excite or charm those of the opposite sex, generally the females, which no longer remain passive, but select the more agreeable partners. [(1), p. 630]

In the quote above, Darwin makes it clear that through mate choice, females become active agents of sexual selection rather than mere passive participants. As we note below, however, this empowerment of females was not extended to his own species. As Richards (7) explains, Darwin naturalized female choice among animals and normalized male choice among humans.

Sexual selection was revealed to Darwin in the form of sexual dimorphism, and it is sexual dimorphism that continues to fuel our inquiries. Just as Darwin invokes sexual selection to explain human differences, he spends 10 chapters—three on birds alone—on an encyclopedic analysis of differences between the sexes across the animal kingdom. Throughout, his focus is on the color, song, and morphology of males. How could these sexually dimorphic traits evolve when they seemed to decrease the survivorship of the sex that bore them? The answer was simple: These traits were favored because they increased the likelihood of mating, either by prevailing over rivals for access to reproductive opportunities or by wooing prospective partners.

Much of what we know about sexual selection is well-documented and relatively intuitive—traits that make males more successful are favored by selection. If selection takes the form of intrasexual competition, males evolve weapons that make it easier to fight, as well as ornaments and behaviors used in aggressive signaling (11). Conventional signaling theory, whereby communication systems evolve to minimize costs for both signalers and receivers, provides a well-supported framework for the evolution of sexually dimorphic traits through male-male competition. These traits are typically "honest"; that is, variation in sexually dimorphic signals predicts variation in signalers' fighting ability (12).

We also know that if traits "excite or charm" potential mates, they will give a reproductive

<sup>1</sup>Department of Biology, University of Padova, Padova, Italy.

<sup>2</sup>Centro de Investigaciones Científicas de las Huastecas

"Aguazarca," Calnali, Hidalgo, Mexico. <sup>3</sup>Department of Integrative Biology, University of Texas, Austin, TX, USA. <sup>4</sup>Smithsonian Tropical Research Institute, Balboa, Republic of Panama.

\*Corresponding author. Email: gil.rosenthal@unipd.it (G.G.R.); mryan@utexas.edu (M.J.R.)

advantage to courtiers with those traits, which is in line with the broad strokes of what Darwin proposed. It is this advantage that favored the evolution of the elaborate, gaudy, and conspicuous traits that are often used in mate attraction—traits that are not necessarily always “honest” because they do not always convey information about inherent survival attributes of males (2, 3, 9). Although Darwin’s interest in sexual selection was piqued by a certain kind of dimorphism, greater expression of both armaments and ornaments in males, we now know that these Darwinian sex roles are often reversed, with females battling it out for access to males and also being subject to the vicissitudes of male mate choice (13, 14).

### Ears, brain, clitoris: The unwhisperable substrates of Darwin’s theory

Darwin’s radical hypothesis assigned females a pivotal role in mate choice. His contemporaries scoffed at his assertion that females actively decided what is sexually attractive and thus which males got to mate (7). Tellingly, Darwin did not suggest the same agency in female humans (15) and generally expressed deeply misogynistic views on women’s roles in the mating endeavor. Throughout *The Descent*, he views women’s intellectual inferiority as a manifest fact of nature [(1), p. 565].

Darwin’s reflexive misogyny makes for an uncharacteristically euphemistic, muddled view of mate-choice mechanisms. Whether human, bird, or butterfly, Darwin’s females are allowed neither strategy nor lust: They are charmed or dazzled by beauty, ingénues entranced by cameos of eligible Lords on display at the Victoria and Albert. Darwin observes above the waistline and at a distance, with a chaste focus on “beautiful” plumage and song; he is squeamish about sex and particularly about female sexual desire. Even in butterflies, courtship culminates not in copulation but in a “final marriage ceremony.” What little there is about the “messier” aspects of sex, such as rubbing and licking and sniffing and secreting, is cloaked in euphemism or in Latin, and some things are actively avoided, unwhisperable in any language: homosexuality, ejaculate mixing, female orgasm. Darwin ignored multiple mating and therefore all of postmating sexual selection, including sperm competition and cryptic female choice.

Darwin may have had little choice in adopting his prudishness in such a socially charged environment: Even 26 years later, when Havelock Ellis tackled such topics in his *Psychology of Sex*, a bookseller was charged with distributing obscene material (16). Such public outrage also greeted Kinsey (17, 18) in the USA a half century later.

Darwin set up another major hurdle that had little to do with prurience or misogyny and indeed ironically contradicted his dismiss-

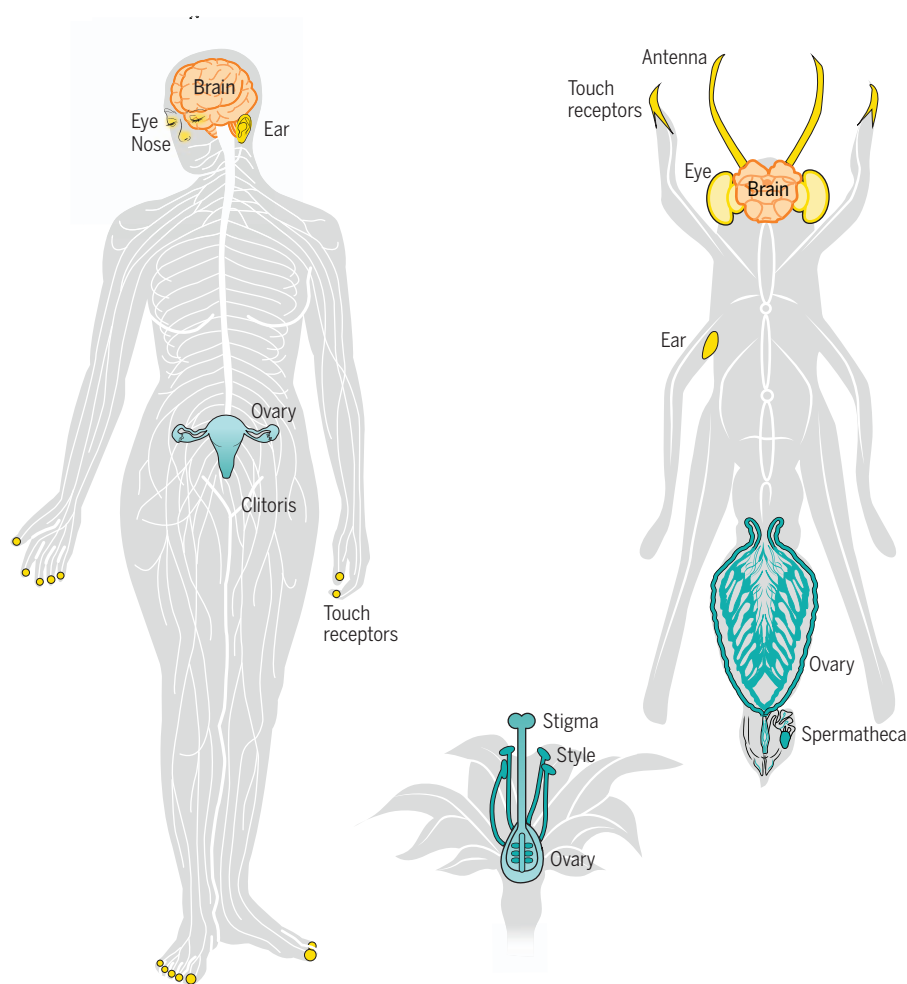
sal of the females of his own species: the requirement for sophisticated “mental faculties” to exercise choice. Certainly, the brain shapes mating outcomes in species that have one, but a brain is not required for choice. Indeed, the oldest and most universal forms of sexual selection take the form of interactions among gametes and within the reproductive tract (Fig. 1, areas in blue).

### What Darwin couldn’t see: Postmating sexual selection, variation among females, and male choice

Postmating sexual selection—sperm competition among males, fertilization bias by females, and allocation bias by both sexes—was beyond the pale for Darwin, yet such mate choice after mating is often a stronger agent of sexual selection than the premating “taste for the beautiful” (19). Yet postmating sexual selection is only one of several avenues where Darwin’s denial of women’s intelligence, agency, or promiscuity led him astray on mate choice in humans and nonhumans alike: not only ig-

norning multiple mating, but ignoring variation among choosers across scales, and the mutual nature of mate choice and its consequences.

Although Darwin was quite interested in variation among males, he treated females and their preferences as uniform, rather than as distinct individuals. Kinsey, whose earlier career as an entomologist endowed him with an appreciation for individual variation (6), was an early pioneer in quantifying variation in sexual proclivities across genders. But much research on mate choice continues to follow Darwin’s lead in ignoring preference variation. Preferences can be diametrically opposed in sister species, choosers from the same population with different experiences, and even the same chooser over the course of ontogeny. For example, older satin bowerbird females respond positively to high-intensity male courtship displays, whereas their younger sisters are startled away (20). A host of variables, from the developmental trajectory of sensory systems (21) to diet (22), predation risk (23), pathogen infection (24), and maternal hormone



**Fig. 1. Analogous mate-choice mechanisms in a flowering plant, an insect, and a mammal.** Sensory periphery for stimulus detection (yellow), brain for perceptual integration and evaluation (orange), and reproductive structures for postmating choice among pollen or sperm (teal).



transfer (25), can have marked effects on how preferences vary [reviewed in (3), chapters 9 to 12]. Female spadefoot toads take preference modulation to the logical extreme, flipping their mate preferences for conspecifics versus heterospecifics depending on climatic conditions (26). Despite such notable examples, a recent meta-analysis (27) suggests that such adaptive plasticity in mate choice may be an exception: Environmental context effects on mate choice are small and inconsistent.

Darwin's treatment of females as uniform might have contributed to his failure to appreciate the importance of male mate choice in nonhumans: If females are all the same, then there is no variation to choose upon. Male mate choice is most pronounced in species where males are a limited resource for females, like male-pregnant pipefishes; in these cases, females compete aggressively with each other and for the attentions of males. Even in highly polygynous systems where males compete for females, males devote more courtship effort to more-fecund females and females they haven't mated previously. A recent surge of interest in male mate choice also expands the importance of females in sexual selection theory (14).

Without variation in females, there is no male mate choice; and without male mate choice, it's a one-way process. Darwin recognized, in passing, that mate choice can be a reciprocal process, with incompatible mates exhibiting "mutual repugnance to pairing" [(1), p. 182]. Yet throughout the renaissance of sexual selection a century later, studies emphasized mate choice interactions as unilateral, with information and stimulation flowing in only one direction, usually between the male courter and the female chooser. This was despite the fact that contemporaneous studies of reproductive physiology emphasized the importance of interactions between individuals during courtship (28).

Perhaps Darwin would have arrived at a broader understanding of mate choice if he had diverted his fascination with birds to a wider array of systems. Despite *The Descent's* extensive survey of sexual dimorphism in display traits, Darwin swept behavior under the rug, focusing on animals that seemed to conform to traditional Victorian sex roles. He did not give much thought to sexual behavior outside the context of heterosexual reproduction, nor did he address hermaphroditism. Whereas terrestrial vertebrates and arthropods have only one functional sex over a lifetime, sequential and simultaneous hermaphroditism are widespread in fishes, flowering plants, and mollusks, among others. With simultaneous hermaphrodites, there is of course no sexual dimorphism, and postmating interactions play a primary role in sexual selection (29). With sequential hermaphroditism, sexual selection

theory has provided the framework to understand when individuals should change sex and in what direction (30), and how social dynamics shift the timing of sex change from individual optima (31).

The fields of animal behavior, neuroscience, and genetics have given us the tools we need to elucidate mechanisms of sexual selection, measure their consequences, and build quantitative theory. Perhaps more notably, much of the latter-day progress on mate choice and sexual selection occurred because we are less afraid to talk about mate-choice mechanisms and mating outcomes. Using value-neutral, gender-neutral terms—multiple mating rather than promiscuity, arousal thresholds rather than coyness and eagerness—helps us avoid some of Darwin's pitfalls.

Sexual selection research is also no longer a Victorian patriarch's club. The resurgence of the study of sexual selection in the 1970s saw queer and women scientists at the center of the field (32–36), and studies increasingly engage their entanglement with implicit bias and public views on sexuality (37). Although this piece comes from yet two more straight cisgender men, our intellectual world is nevertheless shaped by perspectives that were missing in Darwin's time.

Feminism and changing sexual mores have eased the misogyny and prudery that clouded Darwin's thinking on the "taste for the beautiful." Yet like Darwin, we mostly remain charmed by beautiful males, rather than by the biological mechanisms that made them beautiful (38) (Fig. 1). Nevertheless, we are starting to learn much by focusing on the basics of choice and preference (38). Below, we focus on three key aspects of mate choice that Darwin approached but that have taken 150 years to begin unraveling. We concentrate first on the "taste for the beautiful": the mechanisms that are the evolving agents of sexual selection. We then discuss how "tastes" are shaped by the social environment, which in turn constrains how preferences are realized into choices. Finally, we address how preferences come to be and return to Darwin's original impetus for *The Descent*—namely, sexual selection as an agent of diversification among populations and species. Research since Darwin reveals that whereas mate choice is a fundamental agent of sexual selection, sexual selection may not be that important to mate choice.

### What have we learned about sexual selection since Darwin?

Mate choice can occur before, during, and after mating. Almost any aspect of a chooser's biology can potentially bias matings and therefore serve as an agent of sexual selection. Even a comprehensive account of more than 600 pages (3) could not encompass everything from membrane-bound chemoreceptors to the

muscles surrounding the vaginal wall to the cortical circuits integrating third-party social information with a hedonically labeled representation of a potential mate. The fact that nearly anything can be a mate-choice mechanism is important first because each of these mate-choice mechanisms have different histories of selection in sexual and nonsexual contexts, and different ways of covarying—or not—with traits under selection. Second, so-called "mental faculties" are simply the most complex of countless factors that influence mate choice. Yet third, mental faculties—the brain—are the key to understanding mating outcomes in most animals (39).

The 20th-century rise of animal communication as a field of study (40) provided a profitable framework for studying mating decisions and their consequences. Although communication theory provides a framework for incorporating mutual feedback and two way-interactions (28), the simplest case is where a courter emits a signal and the signal is transmitted through the environment. The chooser detects and analyzes that signal, compares it to other signals, and makes a mating decision. We begin by offering an overview of how mate-choice mechanisms vary among species and individuals at each of these stages—sensory transduction, perceptual integration, and evaluation.

### The sensory periphery

The first steps in responding to a stimulus constitute important mate-choice mechanisms. Any communication signal must be detected in its environment, and choosers are more likely to favor signals that result in easier detectability and greater sensory stimulation (41, 42). Signals tend to match the tuning of peripheral end organs, be these correlations between signal colors and photopigments, sounds and inner-ear tuning, or odors and odorant receptors. How does this match come to be? Phylogenetic analyses across the animal kingdom suggest that whereas signal-receiver properties (i.e., the properties of the signal and the preferences for them) sometimes coevolve tightly, courter traits often evolve in response to preexisting biases, a process known as sensory exploitation (43, 44). Some of these biases may be ancient and tightly constrained. For example, habituation and release from habituation, starting at the sensory periphery and cascading through the brain, are universal properties of neural networks (45). Choosers often have preferences for signals of greater magnitude and/or more complex signals. These preferences might have nothing to do with what these signals mean or do not mean or how costly they are or are not to signalers, but rather how they grasp and hold the receiver's attention. Alerting signals are a good example, as they seem to function in attracting the receiver's attention and little else (46).

Of course, the cost of the signal to a signaler will shape the evolution of that signal even if it means little to the receiver.

Biases that are involved in mate choice need not have evolved for mate choice. Often they arise in response to ecological selection on sensory tuning and other preference mechanisms—not because of mating decisions but because these mechanisms function in another context (47, 48). For example, in the haplochromine cichlids of Lake Victoria, the ambient light environment influences evolution of photopigment tuning, which in turn influences female preference for male color and, consequently, the evolution of the color itself (49). Nevertheless, there are sense organs that have evolved purely in the service of sex (2, 3). Membrane-bound odorant receptor molecules, for example, often respond narrowly to conspecific pheromones. In fact, replacing a *Drosophila* antennal receptor gene with the moth gene that recognizes a moth sex pheromone makes female flies attracted to moth pheromones (50).

### Perceptual integration

Detection is only the first step in responding to a stimulus. Choosers must both pay attention to courters and integrate complex streams of sensory information within or across modalities. It is pertinent to remind ourselves that just as there are colors, sounds, and smells that are important to other animals but imperceptible to us, so too is our perception of complex signals shaped by our own biases regarding what's salient and what's attractive. A male's "beauty" to humans tells us little about what his target audience cares about. During courtship, for example, female peahens spend more time gazing at a male's comparatively drab lower feathers rather than the colorful tail feathers prized by humans (51). The same study suggested that these ocelli serve a rather prosaic function, to elicit a female's attention at a distance rather than influencing the female's assessment of the male in close encounters. All else being equal, a signal that elicits more initial attention is favored by sexual selection (46).

But all else is not equal. Different signals often interact with each other in ways not predicted by the conspicuousness or attractiveness of individual components. Perfumes and curries are enticing because of emergent olfactory percepts that are different from their individual components. Túngara frogs are a well-studied example of perceptual integration along multiple axes. First, females go beyond just attending to a particular acoustic frequency; they are consistently choosy about the temporal structure of the male's "whine" call, imposing strong stabilizing sexual selection. Neurons in the female's auditory mid-brain exhibit greater activity in response to the mat-

ing call's downward frequency sweep compared to other stimuli (52). As choosy as they are about the whine, they are permissive about acoustic accoutrements to the whine. Given that a male makes the appropriate whine, additional novel call components that stimulate the female's auditory system make him more attractive (53). Finally, these complex calls come with a complex visual stimulus: inflation of a conspicuous pigmented vocal sac. A silently inflating vocal sac is ignored by females, but adding it to a call makes the multimodal display more attractive and perhaps easier to find (54). How receivers integrate different components of signals is key to how mate choice operates as an agent of selection.

Just as signal detection is constrained by ecological selection, so too is perception. The olfactory system in fruit flies is involved in two important functions: feeding and mating. The olfactory receptor neurons that are involved in feeding project to the mushroom body, and those involved in mating project to the lateral horn. Males are more likely to court in the presence of rotting fruit, which is where females lay eggs. A specialized olfactory receptor neuron that responds to volatiles of rotting fruit enhances the male's motivation to court when stimulated. This neuron, unlike others that detect rotting fruit, projects to the fly's mating brain, the lateral horn, rather than the mushroom body (55). Thus, not only is the function of a food detection neuron co-opted for mating decisions, so is its neuroanatomical position in the brain. Understanding how and where in the brain mating decisions are made is a major area of investigation in sexual selection and mate choice.

Just as sensory biases can favor certain signals, so too can perceptual biases that arise from integrating sensory information within or across modalities (44). The complex courtship structures of the greater bowerbird provide a spectacular example. Males arrange objects around the bower to create a forced perspective to make themselves appear larger to females (Fig. 2) (56). Such illusions occur often in courtship (57).

### Evaluation and hedonic marking

Sensory and perceptual biases of receivers can have important influences on the evolution of mating signals, but they are certainly not the whole story. A courter's signal may be processed similarly by two choosers who come to opposite conclusions about its attractiveness (58, 59). Both social (see below) and environmental effects can easily cause choosers to lose, intensify, or reverse their preferences—even as most aspects of stimulus processing are conserved. Recent studies of guppies (60) and spadefoot toads (61), for example, have shown that neural responses in brain areas that process sensory stimuli need not predict

responses in areas of the brain involved in decision-making. In some cases, gatekeepers in the brain differentially influence how sensory information is transferred to areas where behavioral responses are generated (62, 63).

Not only do neural mechanisms act to filter out unappealing stimuli, they can readily label them "good" or "bad" (58, 59). In insects, structural changes to a single gene are sufficient to cause a "flip" in axon targeting responsible for attraction versus avoidance of intersexual pheromone cues (64). Small genetic or environmental changes can thus potentially have disjunct effects on the strength and direction of sexual selection.

### Mate choice after the brain

Observations of nature would have shown Darwin that there is no "final marriage ceremony"; he was probably too mortified to write it down. Mate choice is mutual partly because not only males but also females mate multiply in many species. Genetic paternity assignments have uncovered multiple mating throughout the animal kingdom, and this is arguably the greatest contribution of the molecular revolution to animal behavior. Across over 500 studies in socially monogamous birds, for example, fully a third of broods contain at least one offspring from outside the pair bond (65). Getting to mate is, thus, simply the end of the beginning. In species with internal fertilization, sexual selection continues to occur during and long after mating. Parker's insight that sperm from different males compete with each other (66) was followed by Eberhard's work on cryptic female choice (67). Mechanical, chemosensory, and somatosensory structures within the reproductive tract serve as mate-choice mechanisms as much as brains and head-bound sense organs do (10). For example, the ancestral state in mammals is that stimulation of the clitoris—and the resultant somatosensory input to the brain—is required to trigger ovulation (68).

Biassing fertilization is only one example of how mate choice can occur during and after mating. Australian redback spiders provide a stark example, where the male's courtship finale positions him between his mate's jaws; males eaten by a female fertilize a greater share of her eggs (69). In birds, females will often spend more time feeding the offspring of attractive males than unattractive males. This effect persists when attractive males are experimentally rendered unattractive (70). Any time that an individual mates with multiple partners over the course of its lifetime, there is the potential for postmating sexual selection.

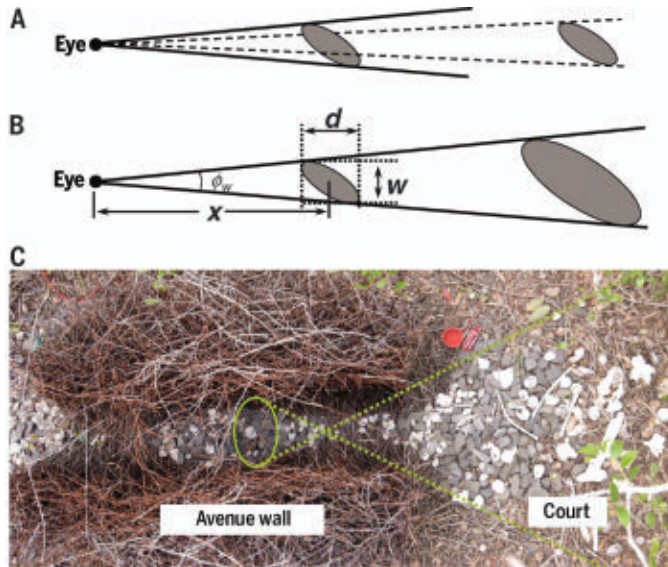
### Mate choice is a social process

A chooser's "taste for the beautiful" does not develop in a vacuum, and the way that taste is expressed depends on a chooser's internal



**Fig. 2. Forced perspective and mate choice in bowerbirds.**

(A) When two objects are the same size but different distances from the viewer, the more distant object subtends a smaller angle on the eye. (B) In forced perspective, as long as the width ( $w$ ) and depth ( $d$ ) of the larger object increase with distance from the viewer ( $x$ ), the larger object will subtend the same visual angle as the smaller one ( $\phi_w$  is the visible angle). (C) This top view of a bower shows where the female is positioned in the avenue (oval) when she observes the courting male in his court. The female moves her head between the walls, leading to a predetermined field of view (dotted line). [Reprinted from (56) with permission]



state and external inputs. The physical environment and ecological community can have big effects on how preferences are organized and expressed, and the same is true of social interactions.

The social environment—conspecific individuals and groups of individuals interacting with a focal actor—shapes preferences and choices at all scales, from maternal effects when the brain starts to develop to audience effects at the time of mating. Third parties can influence mating outcomes long before choosers reach maturity and can act to reverse mating decisions in the moment. These social effects play a determining role in how preferences are shaped and, in turn, how those preferences are realized in the individual choices that together generate sexual selection. We can divide these influences broadly into two general categories: those that shape preferences before they are expressed and those that shape how preferences map onto choices when choosers are comparing potential mates.

### Imprinting and social exposure

We know little about how cellular epigenetics shape mate choice (71), but epigenetic modification in the broader sense is ubiquitous. Parents and siblings frequently influence adult preferences when choosers are still in a nest or with a parent. Cultural imprinting on a genetic parent acts to powerfully couple variation in traits and variation in preferences (72). These effects can be irreversibly specified during early development, or they can be the flexible consequences of short-term exposure (73). Choosers often favor mere familiarity or, conversely, mere novelty in courtiers; these biases act to make sexual selection positively or nega-

tively frequency-dependent, respectively. A special and extreme case of preference for familiarity is pair bonding in socially monogamous vertebrates (74).

Sexual imprinting and pair bonding are domain-specific and circumscribed to species with particular life histories. A more universal if underappreciated driver of preference variation, and therefore of sexual selection, may be associative learning. Pairing sexual reward with an arbitrary object elicits fetishes in rats and quail, whereby a piece of cloth, for example, becomes required for sexual arousal (75, 76). Just as positive interactions with a mate can positively reinforce associations with an arbitrary stimulus, it is reasonable to suppose that negative interactions with a phenotype should lead to sexual avoidance. Notably, different receivers may label the same courter trait arousing or aversive (59). As previously mentioned, young female bowerbirds flee from the high-intensity courtship displays that entice their more experienced elders (77).

Genotype-by-environment interactions are as pervasive for social effects as for everything else and are yet another factor that biases stimulus preferences (78); learning biases are no exception. Marler's "instinct to learn" (79) means that structural differences among species, sexes, and/or individuals act to bias the effects of experience on mate choice. For example, female sheephead swordtails prefer familiar male phenotypes, whereas their sister species, highland swordtails, are repelled by familiarity (80).

### Mate sampling and decision-making

Numerous studies suggest adaptive flexibility in preferences, whereby choosers become more

permissive as mates become scarce, risks increase (81, 82), and time becomes short. Human standards of beauty broaden over the course of a social evening as closing time at bars approaches (83), just as do preferences in some female fishes in the days before death (84).

Mate assessment involves comparing potential mates, and the strength of sexual selection depends critically on how many mates an individual samples and how these individuals are remembered and compared. Sounds are ephemeral, and in most cases an auditory memory is required to compare them. This is especially true in choruses of insects and frogs in which bouts of singing are interrupted by periods of silence. Auditory memory can be flexible; studies of two frog species show that addition of call syllables (85) or a visual cue (86) instantiates sufficient memory to exceed intercall bout intervals.

Mating outcomes thus depend on whether a chooser is remembering individual courtiers or simply mating with anyone above a threshold. Often, however, experience with multiple courtiers can affect both the stringency of the threshold and the attractiveness of courtiers relative to one another. This can produce nonlinearities that complicate sexual selection. Rational choice theory posits that the relative attraction between A and B should be independent of the presence of an alternative, C. We know that this assumption is violated quite often in humans and other animals, and some recent studies show an effect of such competitive decoys on mate choice (87, 88).

Finally, the decisions of individual choosers depend on other choosers. For example, males in internally fertilizing species tend to avoid females with previous mates because of increased sperm competition. Male Atlantic mollies are more likely to court unattractive females when other males are nearby (89). In females, by contrast, mate choice copying is ubiquitous and powerful, at least in experimental settings. Dozens of studies show that a preference for a typically more attractive versus unattractive male can be reversed if a female sees the unattractive male being courted by another female (90, 91). Mate choice copying can also be influenced by the phenotype of the model female (92).

In females, copying may often be favored as long as popular males have abundant sperm and few sexually transmitted diseases. But there has been little success in uncovering its fitness benefits (90). This might be because mate choice copying need not have evolved for mate choice, but instead might be a special case of domain-general social facilitation that evolved in another social context (93). For example, a recent study in humans shows that copying the decision of others has the same magnitude of effect when they evaluate facial attraction or attraction of hands, as well as

abstract art (94). Similar across-domain comparisons in nonhuman animals would be worthwhile.

### Mate choice has consequences for choosers, not just courters

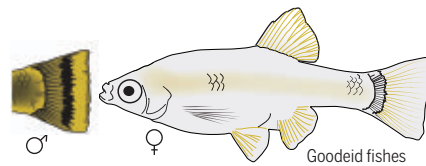
We saw above that preferences frequently arise outside the context of mating decisions. These effects of pleiotropy (95), sensory drive (47), sensory exploitation (96), and sensory traps (34) have notably changed the way we think of mate choice. The brain is an important sex organ but, for females at least, mainly has other things on its mind (Fig. 3A).

It is only once there are choices to be made among partners that preference mechanisms become subject to selection in the context of mate choice (Fig. 3B). We can ask three primary questions about the fitness consequences of mating decisions. First, are they good, neutral, or bad for a given measure of fitness (Fig. 3, B and C)? Second, are those fitness costs (like sexually transmitted diseases) or benefits (like protein-rich spermatophores) directly experienced by a chooser or their offspring (Fig. 3B), or are preference mechanisms under indirect selection because they are genetically correlated with courtship traits that are under direct selection, and thus preferences evolve through genetic hitchhiking (Fig. 3C)? Finally, are fitness consequences additive, meaning that costs and benefits from a courter are independent of who the chooser is, or are they complementary, depending on the interaction between courter and chooser (Fig. 3D)?

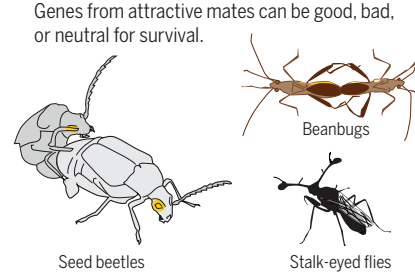
Direct selection as a result of mating outcomes plays a major role in the evolution of sexual and social phenotypes. There are some systems where direct benefits are evident as tangible resources, such as food, protection, and parental care, but every sexual interaction carries costs and benefits. Even broadcast spawners and wind pollinators synchronize the timing of gamete release and express mate choice through sperm-egg interactions (97). And there is always a downside risk to making mistakes, such as initiating the acrosome reaction in response to a nonviable sperm or other foreign object.

Direct selection therefore has a powerful influence on mate-choice mechanisms at the gamete level, but direct benefits and costs apply to mating decisions even when gametes are not involved. It is useful to think about sexual selection as a special case of social selection (34, 35). In complex organisms, sexual behavior is about more than exchanging gametes, and it can have fitness consequences that go beyond reproducing with a partner. Consider same-sex interactions between gull parents, or as a mediator of social hierarchies in bonobos (3). Nonreproductive sexual behavior carries similar downside risks of trauma and pathogen transmissions as reproductive sex.

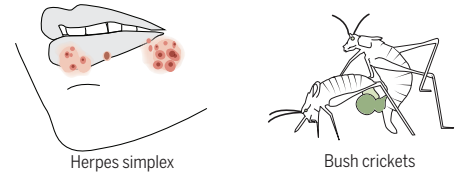
### A Mate-choice mechanisms often come from non-sexual functions. Sensory and perceptual biases shape preferences.



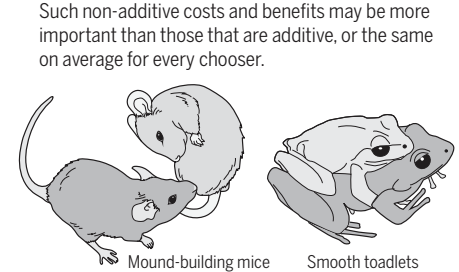
### C Mating decisions are under indirect selection because they co-evolve with courters. Genes from attractive mates can be good, bad, or neutral for survival.



### B Mating decisions are under direct selection. Phenotypic benefits and costs abound.



### D Fitness consequences are dependent on the interaction between chooser and courter. Such non-additive costs and benefits may be more important than those that are additive, or the same on average for every chooser.



**Fig. 3. Origin and evolution of mating preferences.** (A) Mating preferences can emerge from nonsexual biases like foraging (goodeid fish) (107). (B) Once preferences are expressed as such, they are under direct selection (herpes sore in human, spermatophore transfer in bush crickets) (142). (C) Traits and preferences coevolve; attractive courters can have good genes (bean bugs) (143), bad genes (seed beetles) (144), or neutral genes (stalk-eyed flies) (145) with respect to offspring viability. (D) Consequences of mate choice are often complementary; *Uperoleia laevis* frogs drown if their mate is too big (115); mound-building mouse pairs start breeding faster if they have matching personalities (146).

An overlooked cost of mate choice is that associated with cognitive processing. Making decisions takes time, energy, and neural hardware. Animals can suffer cognitive overload, especially when there is information being delivered in multiple sensory modalities (98). There may be hedonic rewards to decreasing neural computational costs—pleasure from fluency of processing (99). In a percoid fish, females prefer male visual signals that are easier to parse (100).

Like any other trait, mate-choice mechanisms evolve in response to direct selection on both their sexual and nonsexual function. Those mechanisms in turn drive the evolution of sexual signals. The coevolutionary dynamics between mating signals and mating preferences depend not only on the direct costs and benefits associated with mating decisions but also on the indirect genetic benefits of mating outcomes. Attention has centered on the genetic benefits of mating decisions (Fig. 3C), specifically the additive genetic benefits (Fig. 3D). These benefits (and costs) are usually partitioned into viability—offspring vigor and health—and offspring attractiveness.

A preference that successfully predicts offspring viability is favored by a so-called “good genes” process. Preferences are favored because attractive courters sire offspring that thrive. But preferences are also favored simply because attractive mates produce attractive offspring: the Fisher-Lande-Kirkpatrick pro-

cess of genetic hitchhiking (101). In both cases, a preference evolves not because it is under direct selection but instead because preference genes become statistically associated with the good genes for survival or with the genes for attractive traits (102). These mechanisms are likely episodic but ubiquitous and may play an important role in hindering or facilitating adaptation to new environments. Hitchhiking happens automatically if there is genetic variance in both sexually dimorphic traits and preferences, whereas good genes require that traits predict offspring viability. Although there are few studies on preferences, meta-analyses suggest ample genetic variance in traits (103), whereas genetic effects on offspring viability (“good genes”) are generally small and likely to have only a small effect on preference evolution (101, 104). Sex chromosomes provide interesting dynamics in sexual selection: Y chromosomes that make males more attractive, for example, can accumulate “bad genes” that delay maturity or reduce viability, whereas selfish W chromosomes can select for traits that harm males but favor daughters [(105); see also (106)].

### Sexual cooperation and sexual conflict

The coevolutionary scenarios discussed so far return us to Darwin’s emphasis on mate-choice mechanisms as agents of selection. In the simplest evolutionary scenario for sexual selection through mate choice, only one thing need evolve:

a display trait that elicits a latent “taste for the beautiful” in choosers, whether or not that “taste” has any heritable variation. This is the case for the terminal yellow band of male goodeid fishes, which evolved to elicit female attention to a similar-looking insect larva.

The terminal yellow band was advantageous for males, but costly for females, who lost weight chasing after male tails rather than food (Fig. 3A). In goodeid lineages with yellow bands, females evolved greater sophistication, decoupling the feeding response to insects from the sexual response to males (107, 108). The coevolution of the preference and trait here is an example of sexual conflict. Males gain fitness by increasing their mating success at the expense of female foraging success. Females, in turn, increase their foraging success while making it more difficult to mate.

Sexual conflict always occurs when partners have divergent interests in a mating interaction. This is the flip side of sexual reproduction as cooperation, where producing and/or rearing offspring is to each partner’s mutual benefit. In the extreme case of lifetime genetic monogamy, there is no sexual conflict; your partner’s lifetime reproductive success is your own. If both males and females mate multiply, sexual conflict is extreme. In *Drosophila*, experimentally enforcing monogamy, and therefore arresting sexual selection and sexual conflict, eliminates genetic load by relaxing selection on male seminal fluid proteins, which are toxic to females, and on female resistance to toxicity, which reduces male fertilization success (109).

Sexual conflict can take two forms. The first is intralocus sexual conflict. Here, an allele that is favored in males is disadvantaged in females, or vice versa. A single locus, RXFP2, is associated with much of the variation in horn size in male Soay sheep. One allele makes horns larger, giving males an advantage in male-male competition. An alternative allele makes smaller horns but is associated with increased offspring viability. Similarly, male seed beetles that succeed in sperm competition sire daughters that are less likely to survive (110).

In both cases, females are saddled with “bad genes” when they mate with competitively successful males; their offspring are less likely to thrive, even if those sons that survive to maturity have a competitive advantage. Preferences for traits that confer “bad genes” can only persist if those “bad genes” are limited to courtiers. In two species of poeciliid fish, some of the variance in male attractiveness comes from genes on the Y chromosome. Haplotypes that make males more attractive accumulate mutations that reduce survival, such that females face a trade-off: Either mate with unattractive males who produce sons more likely to survive but less likely to reproduce if they do, or attractive sons who will have higher fit-

ness if they make it to maturity (111–113). Also, if males but not females suffer reduced survivorship, and females are the heterogametic sex, then preference for more attractive males can be favored by selection even if they have breeding value for lower variability (105, 106).

In contrast to intralocus sexual conflict, interlocus sexual conflict follows more conventional dynamics analogous to predator-prey coevolution. Sexual cannibalism provides perhaps more than an analogy. Where mates are scarce, males and females cooperate; if a male and a female mate once, sacrifice increases the fitness of both partners. In systems where mates are dense and males encounter multiple females, they may endeavor to mate with a female but deprive her of a meal as they search for their next partner. Finally, if females encounter multiple males, they may choose to eat a male but eschew his sperm. When male and female interests over mating diverge, the stage is set for antagonistic coevolution.

Most of the evolutionary dynamics discussed so far are additive (Fig. 3): A courter’s effect on a chooser or her offspring is independent of that chooser. But one individual’s ideal mate is often another’s nightmare. Such nonindependent or complementary processes constitute major, underappreciated sources of selection on mate choice, if not sexual selection (114).

Phenotypic compatibility, assortative mating by body size (115) or personality (116), or synchronization of reproductive state (7, 8) are all major outcomes of mate choice. In any sexually reproducing system, compatibility is a fundamental, yet underappreciated, force in mate-choice evolution. Choosers benefit by choosing partners with compatible genes (117) such as conspecifics or individuals with complementary immune genes (118, 119).

Conspecific pollen precedence and conspecific sperm precedence are examples of selection favoring a preference for a compatible genotype—in this case, a conspecific. Self-incompatibility in plants (120) is, at the other extreme, also a preference for a compatible genotype.

### Sexual selection and gene flow

A clearly complementary outcome is when choosers decide between conspecifics and heterospecifics. Where do conspecific mate preferences come from, and why do choosers prefer heterospecifics? When populations stop exchanging genes, they can develop incompatibilities in sexual communication, as a special case of genetic incompatibilities among divergent regions in the genome (121). Traits and preferences coevolve along different trajectories in different populations. This can happen because of stochastic processes like mutation-order effects (122) and genetic hitchhiking (123) and can be accelerated by ecological divergence (121) or antagonistic coevolution due

to sexual conflict (124). Sexual selection can thus accelerate divergence in allopatry, when external barriers prevent gene flow. Darwin’s intuition for differences among human groups—that sexual selection could lead to morphological diversification—is theoretically sound and has received ample support in animals.

Darwin sought to explain recent divergence among isolated populations within the same species—humans. But what happens when two different but closely related species come into secondary contact? Here, selection for compatible genes favors reinforcement, i.e., divergence of traits and preferences in sympatry. Selection against hybridization in different locations can lead to signal-receiver divergence among conspecific populations, also known as cascade reinforcement (125).

When there is gene flow among populations, sexual selection plays an even more interesting role. Environmental and social effects on individual mating decisions, described above, can modulate hybridization between species (126). If traits evolve purely because they are attractive, theory suggests that sexual selection homogenizes populations, because choosers will mate with genetically divergent courtiers bearing a preferred trait (127). By contrast, if display traits and/or mating biases are subject to divergent ecological selection, theory suggests they can reinforce reproductive isolation and divergence in sympatry (128). “Magic traits” (129), which are defined as traits that are involved in both reproductive isolation and ecological divergence such as visual sensitivity in Lake Victoria cichlids (49, 130), link divergent ecological selection for different color sensitivity to assortative mating by color. Perhaps counterintuitively, preference-trait combinations can contribute most easily to reproductive isolation if preferences are learned from genetic parents (131). In poison frogs, imprinting on maternal phenotypes could maintain coexistence of distinct color morphs in sympatry through the congruent actions of the two main mechanisms of sexual selection. Females mate with males with their mother’s color pattern, promoting assortative mating by color morph. Males attack rival males with the same pattern, giving rare color morphs an advantage (132).

Finally, sexual selection is just beginning to be reconciled with an emerging view of macroevolution as a reticulate process heavily influenced by gene flow among divergent lineages. Hybridization was accepted in Darwin’s time as an important force in the evolution of plants and microorganisms but was largely dismissed as an aberration in animals, the “grossest blunder in sexual preference,” according to Fisher [(133), p. 150]. A love of racist typology perhaps hindered the field’s appreciation of the importance of hybridization, which we now understand to play an important role



in adaptation (134) and speciation (135). Environmental modulation of individual mating decisions may play a key role in macroevolutionary processes (126), including the evolution of *Homo*; for example, how did mate choice, sexual coercion, and sexual combat modulate gene flow between Neanderthals and humans?

### The ascent of choice: Sexual decisions as a linchpin of evolution

In the broadest strokes, Darwin's theory of sexual selection is one of the best documented and most important contributions in evolutionary biology (7). Darwin's specific idea of sexual selection by mate choice redefined females as not merely passive actors over which males raise combat, but instead as the arbiters who wield the sword of selection separating those males who do and do not reproduce. Yet Darwin's idea of female empowerment evaporated when applied to his own species. Darwin's misogyny infiltrated his understanding of human mating patterns and polluted his general notion of the value of women in society. His elisions on female desire and agency contributed to a caricature of Darwinian evolution that has compromised and fragmented the way we study sexuality. Variation among humans, the explicandum for Darwin's long argument about sexual selection and human equality, has become the almost exclusive province of the social sciences, whereas evolutionary psychology searches for human universals (3).

### Mate choice mechanisms matter

Knowing the nature of the "taste for the beautiful" lets us make predictions about how sexual selection should facilitate or hinder gene flow between species. For example, an early model of sympatric speciation through mate choice (136) was discredited because it posited discrete flips in preference or antipathy for courter traits. As argued above, however, empirical studies of evaluative mechanisms suggest that such flips may occur. Similarly, recent theoretical models have shown that learning from genetic parents facilitates assortative mating and therefore genetic divergence (137, 138). The mechanisms underlying signal-receiver congruence, and the fitness consequences of mate choice, are crucial to predicting whether sexual selection acts to promote or inhibit genetic exchange.

### Mate-choice evolution is not chained to sexual selection

The obsession over "good genes" and the "evolution of beauty" hinges on genetic variation in preferences. In a chooser-centric view, this means heritable variation in preference mechanisms. Such variation remains elusive, notably in female vertebrates. This is foremost limited by the low repeatability of mating preferences, with a modal repeatability of zero in

vertebrate studies (3). At the other end of the spectrum, preference mechanisms under strong stabilizing selection may be stable over decades (139), perhaps showing little opportunity for selection. Even if we could measure stable preference phenotypes in individuals, individual social experience and life history add an incredible amount of variation to genotypic effects, some of which may be adaptive. The effect of this nongenetic variation will generally be to reduce the heritability of preference and therefore the opportunity for genetic coevolution between traits and preferences. In people as in fruit flies, it is impossible to understand sexual selection without considering how genes interact with their environment, especially their sociocultural environment.

Darwin's observation that mutual attraction and mutual repugnance shape mating outcomes deserves renewed focus now that we understand the importance of multiple mating and mating decisions by both males and females. Understanding mate choice as a dynamic process involving choices by at least two actors has the potential to overturn intuitive predictions about the evolution of choices and sex roles. An instructive example is lekking topi antelope females, who aggressively compete for access to a preferred male. The preferred male, sperm limited and in demand, prefers unmated females. Being the most preferred male topi antelope is unambiguously good for male fitness, but being the most preferred female fruit fly is not. Being chosen can be costly as well. Attractive female fruit flies suffer reduced fecundity as a result of constant harassment by males (77).

Understanding multiway interactions in their social and ecological context is needed to tackle Darwin's original focus with *The Descent*, the nature of diversity among humans. Human diversity has become close to a taboo subject in the biological study of sexual selection, with the notable exception of a growing body of work on the neuroendocrinology and genetics of same-sex attraction (37). The broader sexual spectrum deserves study, as does Darwin's primary conjecture that sexual selection generated differences in appearance among human populations.

We need to better understand conflict and cooperation dynamics in sexual interactions before we can generalize about whether sexual selection is on the whole good or bad for populations. Does it reinforce purifying and ecological selection for "good genes," those that increase viability, or do "bad genes," those that decrease viability, hitchhike along with attractiveness (129)? A recent model (140) suggests that sexual conflict over reproductive investment can stabilize into cooperation over a broad range of conditions.

Indeed, removing the opportunity for mate choice often results in reduced population fit-

ness (141), but studies seldom disentangle the effects of choosing compatible mates from those of selecting the best-adapted mates. We need more empirical data to disambiguate the consequences of mate choice from those of sexual competition. In particular, we need to distinguish consequences that are additive from those that are complementary. Numerous studies show a benefit of premating choice to chooser fitness, but these conflate additive "good genes"—choosers all picking only the best mates—with complementary outcomes where choosers get to pick their preferred individual partner. Complementarity of genotypes—whether whole conspecific genomes or at discrete loci like the major histocompatibility complex (118, 119)—compatibility of reproductive physiological states (28), and compatibility of behavior (129) are of primary importance to the evolution of mate-choice mechanisms.

So where are we 150 years after Darwin suggested a sex-based theory to explain natural flamboyance and human diversity? Natural selection theory had its predecessors, such as Darwin's grandfather Erasmus, as well as other scientists who focused on the same explanation, as did Alfred Russel Wallace. But sexual selection theory seems to have been birthed purely by Darwin without historical precedents and also without contemporary support (7). Sexual selection theory is almost purely Darwinian, and the controversy it generated 150 years ago might be different in the details but certainly not in the societal discontent it continues to breed, especially in some corners of the social sciences. The Victorian prejudices that infiltrated Darwin's science have evolved somewhat but not gone extinct. Our focus continues to be shaped by the flamboyant ornaments and displays that caught Darwin's eye. Directional sexual selection on courter traits may well emerge as a "spandrel" from perceptual biases, avoidance of downside risk, and selection of the right partner.

### REFERENCES AND NOTES

1. C. Darwin, *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
2. M. J. Ryan, *A Taste for the Beautiful: The Evolution of Attraction* (Princeton Univ. Press, 2018).
3. G. G. Rosenthal, *Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans* (Princeton Univ. Press, 2017).
4. R. O. Prum, Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2253–2265 (2012). doi:10.1098/rstb.2011.0285; pmid:22777014
5. A. Desmond, J. Moore, *Darwin's Sacred Cause: How a Hatred of Slavery Shaped Darwin's Views on Human Evolution* (Houghton Mifflin Harcourt, 2014).
6. S. J. Gould, *The Flamingo's Smile: Reflections in Natural History* (Penguin Books 1985).
7. E. Richards, *Darwin and the Making of Sexual Selection* (Univ. of Chicago Press, 2017).
8. A. Fuentes, in *A Most Interesting Problem*, J. M. DeSilva, Ed. (Princeton Univ. Press, 2021), pp. 144–161.
9. R. O. Prum, *The Evolution of Beauty, How Darwin's Forgotten Theory of Mate Choice Shapes the Animal World—and Us* (Anchor, 2017).

10. R. C. Firman, C. Gasparini, M. K. Manier, T. Pizzari, Postmating female control: 20 years of cryptic female choice. *Trends Ecol. Evol.* **32**, 368–382 (2017). doi: [10.1016/j.tree.2017.02.010](https://doi.org/10.1016/j.tree.2017.02.010); pmid: [28318651](https://pubmed.ncbi.nlm.nih.gov/28318651/)
11. A. Berglund, A. Bisazza, A. Pilastro, Armaments and ornaments: An evolutionary explanation of traits of dual utility. *Biol. J. Linn. Soc. Lond.* **58**, 385–399 (1996). doi: [10.1111/j.1095-8312.1996.tb01442.x](https://doi.org/10.1111/j.1095-8312.1996.tb01442.x)
12. J. Maynard Smith, D. Harper, *Animal Signals* (Oxford Univ. Press, Oxford, 2003).
13. T. Janicke, I. K. Häderer, M. J. Lajeunesse, N. Anthes, Darwinian sex roles confirmed across the animal kingdom. *Sci. Adv.* **2**, e1500983 (2016). doi: [10.1126/sciadv.1500983](https://doi.org/10.1126/sciadv.1500983); pmid: [26933680](https://pubmed.ncbi.nlm.nih.gov/26933680/)
14. I. Schlupp, *Male Choice, Female Competition, and Female Ornaments in Sexual Selection* (Oxford Univ. Press, Oxford, 2021).
15. H. Dunsworth, in *A Most Interesting Problem, What Darwin's Decent of Man Got Right and Wrong About Human Evolution*, J. DeSilva, Ed. (Princeton Univ. Press, 2021), pp. 183–203.
16. E. M. DuPont, Henry Havelock Ellis (1859-1939). *Embryo Project Encyclopedia* (2008).
17. A. C. Kinsey, *Sexual Behavior in the Human Male* (Saunders, 1948).
18. A. C. Kinsey, W. B. Pomeroy, C. E. Martin, P. H. Gebhard, *Sexual Behavior in the Human Female* (Indiana Univ. Press, 1953).
19. L. Marie-Orleach, N. Vellnow, L. Schärer, The repeatable opportunity for selection differs between pre- and postcopulatory fitness components. *Evol. Lett.* **5**, 101–114 (2020). doi: [10.1002/evl3.210](https://doi.org/10.1002/evl3.210); pmid: [33552539](https://pubmed.ncbi.nlm.nih.gov/33552539/)
20. S. W. Coleman, G. L. Patricelli, G. Borgia, Variable female preferences drive complex male displays. *Nature* **428**, 742–745 (2004). doi: [10.1038/nature02419](https://doi.org/10.1038/nature02419); pmid: [15085130](https://pubmed.ncbi.nlm.nih.gov/15085130/)
21. M. J. Ryan, S. A. Perrill, W. Wilczynski, Auditory tuning and call frequency predict population-based mating preferences in the cricket frog *Acris crepitans*. *Am. Nat.* **139**, 1370–1383 (1992). doi: [10.1086/285391](https://doi.org/10.1086/285391)
22. H. K. Kindsvater, S. E. Simpson, G. G. Rosenthal, S. H. Alonzo, Male diet, female experience, and female size influence maternal investment in swordtails. *Behav. Ecol.* **24**, 691–697 (2013). doi: [10.1093/beheco/ars213](https://doi.org/10.1093/beheco/ars213)
23. P. M. Willis, G. G. Rosenthal, M. J. Ryan, An indirect cue of predation risk counteracts female preference for conspecifics in a naturally hybridizing fish *Xiphophorus birchmanni*. *PLOS ONE* **7**, e34802 (2012). doi: [10.1371/journal.pone.0034802](https://doi.org/10.1371/journal.pone.0034802); pmid: [22529936](https://pubmed.ncbi.nlm.nih.gov/22529936/)
24. S. Cotton, J. Small, A. Pomiankowski, Sexual selection and condition-dependent mate preferences. *Curr. Biol.* **16**, R755–R765 (2006). doi: [10.1016/j.cub.2006.08.022](https://doi.org/10.1016/j.cub.2006.08.022); pmid: [16950102](https://pubmed.ncbi.nlm.nih.gov/16950102/)
25. A. Bonisoli-Alquati *et al.*, Effects of egg testosterone on female mate choice and male sexual behavior in the pheasant. *Horm. Behav.* **59**, 75–82 (2011). doi: [10.1016/j.jybeh.2010.10.013](https://doi.org/10.1016/j.jybeh.2010.10.013); pmid: [21029735](https://pubmed.ncbi.nlm.nih.gov/21029735/)
26. K. S. Pfennig, Facultative mate choice drives adaptive hybridization. *Science* **318**, 965–967 (2007). doi: [10.1126/science.1146035](https://doi.org/10.1126/science.1146035); pmid: [17991861](https://pubmed.ncbi.nlm.nih.gov/17991861/)
27. L. R. Dougherty, Meta-analysis shows the evidence for context-dependent mating behaviour is inconsistent or weak across animals. *Ecol. Lett.* **24**, 862–875 (2021). doi: [10.1111/ele.13679](https://doi.org/10.1111/ele.13679); pmid: [33471386](https://pubmed.ncbi.nlm.nih.gov/33471386/)
28. D. S. Lehrman, in *Sex and Behavior*, F. A. Beach, Ed. (Wiley, 1965), pp. 355–380.
29. M. Beekman, B. Nieuwenhuis, D. Ortiz-Barrientos, J. P. Evans, Sexual selection in hermaphrodites, sperm and broadcast spawners, plants and fungi. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150541 (2016). doi: [10.1098/rstb.2015.0541](https://doi.org/10.1098/rstb.2015.0541); pmid: [27619704](https://pubmed.ncbi.nlm.nih.gov/27619704/)
30. R. R. Warner, D. R. Robertson, E. G. Leigh Jr., Sex change and sexual selection. *Science* **190**, 633–638 (1975). doi: [10.1126/science.1188360](https://doi.org/10.1126/science.1188360); pmid: [1188360](https://pubmed.ncbi.nlm.nih.gov/1188360/)
31. B. H. Roberts *et al.*, Faster juvenile growth promotes earlier sex change in a protandrous hermaphrodite (*barramundi Lates calcarifer*). *Sci. Rep.* **11**, 2276 (2021). doi: [10.1038/s41598-021-81727-1](https://doi.org/10.1038/s41598-021-81727-1); pmid: [33500452](https://pubmed.ncbi.nlm.nih.gov/33500452/)
32. Z. Tang-Martinez, Rethinking Bateman's principles: Challenging persistent myths of sexually reluctant females and promiscuous males. *J. Sex Res.* **53**, 532–559 (2016). doi: [10.1080/00224499.2016.1150938](https://doi.org/10.1080/00224499.2016.1150938); pmid: [27074147](https://pubmed.ncbi.nlm.nih.gov/27074147/)
33. P. A. Gowaty, Y.-K. Kim, W. W. Anderson, No evidence of sexual selection in a repetition of Bateman's classic study of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11740–11745 (2012). doi: [10.1073/pnas.1207851109](https://doi.org/10.1073/pnas.1207851109); pmid: [22689966](https://pubmed.ncbi.nlm.nih.gov/22689966/)
34. M. J. West-Eberhard, Sexual selection, social competition, and evolution. *Proc. Am. Philos. Soc.* **123**, 222–234 (1979).
35. M. J. West-Eberhard, Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**, 155–183 (1983). doi: [10.1086/413215](https://doi.org/10.1086/413215)
36. S. B. Hrdy, *The Woman That Never Evolved* (Harvard Univ. Press, 2009).
37. A. Ganna *et al.*, Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science* **365**, eaat7693 (2019). doi: [10.1126/science.aat7693](https://doi.org/10.1126/science.aat7693); pmid: [31467194](https://pubmed.ncbi.nlm.nih.gov/31467194/)
38. A. M. Achorn, G. G. Rosenthal, It's not about him: Mismeasuring 'good genes' in sexual selection. *Trends Ecol. Evol.* **35**, 206–219 (2020). doi: [10.1016/j.tree.2019.11.007](https://doi.org/10.1016/j.tree.2019.11.007); pmid: [31858995](https://pubmed.ncbi.nlm.nih.gov/31858995/)
39. M. J. Ryan, Darwin, sexual selection, and the brain. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2008194118 (2021). doi: [10.1073/pnas.2008194118](https://doi.org/10.1073/pnas.2008194118); pmid: [33593899](https://pubmed.ncbi.nlm.nih.gov/33593899/)
40. C. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948). doi: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x)
41. M. J. Ryan, A. Keddy-Hector, Directional patterns of female mate choice and the role of sensory biases. *Am. Nat.* **139**, S4–S35 (1992). doi: [10.1086/285303](https://doi.org/10.1086/285303)
42. M. Andersson, *Sexual Selection*. (Princeton Univ. Press, 1994).
43. M. J. Ryan, Sensory systems, sexual selection, and sensory exploitation. *Oxf. Surv. Evol. Biol.* **7**, 157–195 (1990).
44. M. J. Ryan, M. E. Cummings, Perceptual biases and mate choice. *Annu. Rev. Ecol. Evol. Syst.* **44**, 437–459 (2013). doi: [10.1146/annurev-ecolsys-110512-135901](https://doi.org/10.1146/annurev-ecolsys-110512-135901)
45. S. Dong, D. F. Clayton, Habituation in songbirds. *Neurobiol. Learn. Mem.* **92**, 183–188 (2009). doi: [10.1016/j.nlm.2008.09.009](https://doi.org/10.1016/j.nlm.2008.09.009); pmid: [18845267](https://pubmed.ncbi.nlm.nih.gov/18845267/)
46. R. H. Wiley, *Noise Matters: The Evolution of Communication* (Harvard Univ. Press, 2015).
47. J. A. Endler, Signals, signal conditions and the direction of evolution. *Am. Nat.* **139**, S125–S153 (1992). doi: [10.1086/285308](https://doi.org/10.1086/285308)
48. M. E. Cummings, J. A. Endler; Handling editor: Rebecca C. Fuller, 25 Years of sensory drive: The evidence and its watery bias. *Curr. Zool.* **64**, 471–484 (2018). doi: [10.1093/cz/zoy043](https://doi.org/10.1093/cz/zoy043); pmid: [30108628](https://pubmed.ncbi.nlm.nih.gov/30108628/)
49. O. Seehausen *et al.*, Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008). doi: [10.1038/nature07285](https://doi.org/10.1038/nature07285); pmid: [18833272](https://pubmed.ncbi.nlm.nih.gov/18833272/)
50. A. Kurtovic, A. Widmer, B. J. Dickson, A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nature* **446**, 542–546 (2007). doi: [10.1038/nature05672](https://doi.org/10.1038/nature05672); pmid: [17392786](https://pubmed.ncbi.nlm.nih.gov/17392786/)
51. J. L. Yorzinski, G. L. Patricelli, J. S. Babcock, J. M. Pearson, M. L. Platt, Through their eyes: Selective attention in peahens during courtship. *J. Exp. Biol.* **216**, 3035–3046 (2013). doi: [10.1242/jeb.087338](https://doi.org/10.1242/jeb.087338); pmid: [23885088](https://pubmed.ncbi.nlm.nih.gov/23885088/)
52. K. L. Hoke, M. J. Ryan, W. Wilczynski, Integration of sensory and motor processing underlying sexual behaviour in túngara frogs. *Proc. Biol. Sci.* **274**, 641–649 (2007). doi: [10.1098/rspb.2006.0038](https://doi.org/10.1098/rspb.2006.0038); pmid: [17254988](https://pubmed.ncbi.nlm.nih.gov/17254988/)
53. M. J. Ryan, X. E. Bernal, A. S. Rand, Female mate choice and the potential for ornament evolution in túngara frogs *Physalaemus pustulosus*. *Curr. Zool.* **56**, 343–357 (2010). doi: [10.1093/czoolo/56.3.343](https://doi.org/10.1093/czoolo/56.3.343)
54. R. C. Taylor, B. A. Klein, J. Stein, M. J. Ryan, Faux frogs: Multicomponent signalling and the value of robotics in animal behaviour. *Anim. Behav.* **76**, 1089–1097 (2008). doi: [10.1016/j.janbeh.2008.01.031](https://doi.org/10.1016/j.janbeh.2008.01.031)
55. Y. Grosjean *et al.*, An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. *Nature* **478**, 236–240 (2011). doi: [10.1038/nature10428](https://doi.org/10.1038/nature10428); pmid: [21964331](https://pubmed.ncbi.nlm.nih.gov/21964331/)
56. L. A. Kelley, J. A. Endler, Illusions promote mating success in great bowerbirds. *Science* **335**, 335–338 (2012). doi: [10.1126/science.1212443](https://doi.org/10.1126/science.1212443); pmid: [22267812](https://pubmed.ncbi.nlm.nih.gov/22267812/)
57. L. A. Kelley, J. L. Kelley, Animal visual illusion and confusion: The importance of a perceptual perspective. *Behav. Ecol.* **25**, 450–463 (2014). doi: [10.1093/beheco/arti118](https://doi.org/10.1093/beheco/arti118)
58. R. L. Rodríguez, Back to the basics of mate choice: The evolutionary importance of Darwin's sense of beauty. *Q. Rev. Biol.* **95**, 289–309 (2020). doi: [10.1086/711781](https://doi.org/10.1086/711781)
59. G. G. Rosenthal; Handling editor: Becky Fuller, Evaluation and hedonic value in mate choice. *Curr. Zool.* **64**, 485–492 (2018). doi: [10.1093/cz/zoy054](https://doi.org/10.1093/cz/zoy054); pmid: [30108629](https://pubmed.ncbi.nlm.nih.gov/30108629/)
60. N. I. Bloch *et al.*, Early neurogenomic response associated with variation in guppy female mate preference. *Nat. Ecol. Evol.* **2**, 1772–1781 (2018). doi: [10.1038/s41559-018-0682-4](https://doi.org/10.1038/s41559-018-0682-4); pmid: [30297748](https://pubmed.ncbi.nlm.nih.gov/30297748/)
61. S. S. Burmeister, V. G. Rodriguez Moncalvo, K. S. Pfennig, Differential encoding of signals and preferences by noradrenaline in the anuran brain. *J. Exp. Biol.* **223**, jeb.214148 (2020). doi: [10.1242/jeb.214148](https://doi.org/10.1242/jeb.214148); pmid: [32647019](https://pubmed.ncbi.nlm.nih.gov/32647019/)
62. K. L. Hoke, M. J. Ryan, W. Wilczynski, Candidate neural locus for sex differences in reproductive decisions. *Biol. Lett.* **4**, 518–521 (2008). doi: [10.1098/rsbl.2008.0192](https://doi.org/10.1098/rsbl.2008.0192); pmid: [18611839](https://pubmed.ncbi.nlm.nih.gov/18611839/)
63. K. L. Hoke, M. J. Ryan, W. Wilczynski, Sexually dimorphic sensory gating drives behavioral differences in tungara frogs. *J. Exp. Biol.* **213**, 3463–3472 (2010). doi: [10.1242/jeb.043992](https://doi.org/10.1242/jeb.043992); pmid: [20889827](https://pubmed.ncbi.nlm.nih.gov/20889827/)
64. Z. Kárpáti, T. Dekker, B. S. Hansson, Reversed functional topology in the antennal lobe of the male European corn borer. *J. Exp. Biol.* **211**, 2841–2848 (2008). doi: [10.1242/jeb.017319](https://doi.org/10.1242/jeb.017319); pmid: [18723543](https://pubmed.ncbi.nlm.nih.gov/18723543/)
65. L. Brouwer, S. C. Griffith, Extra-pair paternity in birds. *Mol. Ecol.* **28**, 4864–4882 (2019). doi: [10.1111/mec.15259](https://doi.org/10.1111/mec.15259); pmid: [31587397](https://pubmed.ncbi.nlm.nih.gov/31587397/)
66. L. W. Simmons, N. Wedell, Fifty years of sperm competition: The structure of a scientific revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20200060 (2020). doi: [10.1098/rstb.2020.0060](https://doi.org/10.1098/rstb.2020.0060); pmid: [33070719](https://pubmed.ncbi.nlm.nih.gov/33070719/)
67. W. G. Eberhard, *Female Control: Sexual Selection by Cryptic Female Choice* (Princeton Univ. Press, 1996).
68. M. Pavličev, G. Wagner, The evolutionary origin of female orgasm. *J. Exp. Zool. B Mol. Dev. Evol.* **326**, 326–337 (2016). doi: [10.1002/jez.b.22690](https://doi.org/10.1002/jez.b.22690); pmid: [27478160](https://pubmed.ncbi.nlm.nih.gov/27478160/)
69. M. Andrade, Sexual selection for male sacrifice in redback spiders. *Science* **271**, 70–72 (1996). doi: [10.1126/science.271.5245.70](https://doi.org/10.1126/science.271.5245.70)
70. T. Limbourg, A. C. Mateman, C. M. Lessells, Opposite differential allocation by males and females of the same species. *Biol. Lett.* **9**, 20120835 (2013). doi: [10.1098/rsbl.2012.0835](https://doi.org/10.1098/rsbl.2012.0835); pmid: [23193045](https://pubmed.ncbi.nlm.nih.gov/23193045/)
71. A. M. Makowicz *et al.*, Foundations and Frontiers in Mate Choice Review of: Rosenthal, G. 2017. *Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans*. Princeton Univ. Press, 648 pp. ISBN: 978-0-691-15067-3; \$US55.00 HB. *Evolution* **74**, 1575–1583 (2020). doi: [10.1111/evo.14018](https://doi.org/10.1111/evo.14018)
72. D. J. Yeh, J. W. Boughman, G.-P. Saetre, M. R. Servedio, The evolution of sexual imprinting through reinforcement. *Evolution* **72**, 1336–1349 (2018). doi: [10.1111/evo.13500](https://doi.org/10.1111/evo.13500); pmid: [29741268](https://pubmed.ncbi.nlm.nih.gov/29741268/)
73. M. N. Verzijden *et al.*, The impact of learning on sexual selection and speciation. *Trends Ecol. Evol.* **27**, 511–519 (2012). doi: [10.1016/j.tree.2012.05.007](https://doi.org/10.1016/j.tree.2012.05.007); pmid: [22705159](https://pubmed.ncbi.nlm.nih.gov/22705159/)
74. H. Walum, L. J. Young, The neural mechanisms and circuitry of the pair bond. *Nat. Rev. Neurosci.* **19**, 643–654 (2018). doi: [10.1038/s41583-018-0072-6](https://doi.org/10.1038/s41583-018-0072-6); pmid: [30301953](https://pubmed.ncbi.nlm.nih.gov/30301953/)
75. J. G. Pfau *et al.*, Who, what, where, when (and maybe even why)? How the experience of sexual reward connects sexual desire, preference, and performance. *Arch. Sex. Behav.* **41**, 31–62 (2012). doi: [10.1007/s10508-012-9935-5](https://doi.org/10.1007/s10508-012-9935-5); pmid: [22402996](https://pubmed.ncbi.nlm.nih.gov/22402996/)
76. H. Çetinkaya, M. Domjan, Sexual fetishism in a quail (*Coturnix japonica*) model system: Test of reproductive success. *J. Comp. Psychol.* **120**, 427–432 (2006). doi: [10.1037/0735-7036.120.4.427](https://doi.org/10.1037/0735-7036.120.4.427); pmid: [17115864](https://pubmed.ncbi.nlm.nih.gov/17115864/)
77. G. L. Patricelli, J. A. C. Uy, G. Walsh, G. Borgia, Male displays adjusted to female's response. *Nature* **415**, 279–280 (2002). doi: [10.1038/415279a](https://doi.org/10.1038/415279a); pmid: [11796996](https://pubmed.ncbi.nlm.nih.gov/11796996/)
78. N. W. Bailey, M. Zuk, Socially flexible female choice differs among populations of the Pacific field cricket: Geographical variation in the interaction coefficient psi ( $\Psi$ ). *Proc. Biol. Sci.* **279**, 3589–3596 (2012). doi: [10.1098/rspb.2012.0631](https://doi.org/10.1098/rspb.2012.0631); pmid: [22648156](https://pubmed.ncbi.nlm.nih.gov/22648156/)
79. P. Marler, in *The Epigenesis of Mind: Essays on Biology and Cognition*, S. Carey, R. Gelman, Eds. (Lawrence Erlbaum Assoc., Hillsdale, New York, 1991), pp. 37–66.
80. M. N. Verzijden, Z. W. Culumber, G. G. Rosenthal, Opposite effects of learning cause asymmetric mate preferences in hybridizing species. *Behav. Ecol.* **23**, 1133–1139 (2012). doi: [10.1093/beheco/ars086](https://doi.org/10.1093/beheco/ars086)
81. A. V. Hedrick, L. M. Dill, Mate choice by female crickets is influenced by predation risk. *Anim. Behav.* **46**, 193–196 (1993). doi: [10.1006/anbe.1993.1176](https://doi.org/10.1006/anbe.1993.1176)
82. A. S. Rand, M. E. Bridaroli, L. Dries, M. J. Ryan, Light levels influence female choice in túngara frogs: Predation risk

- assessment? *Copeia* **1997**, 447–450 (1997). doi: [10.2307/1447770](#)
83. J. Pennebaker *et al.*, Don't the girls get prettier at closing time: A country and western application to psychology. *Pers. Soc. Psychol. Bull.* **5**, 122–125 (1979). doi: [10.1177/014616727900500127](#)
  84. C. Passos, B. Tassinio, F. Reyes, G. G. Rosenthal, Seasonal variation in female mate choice and operational sex ratio in wild populations of an annual fish, *Austrolebias reicherti*. *PLOS ONE* **9**, e01649 (2014). doi: [10.1371/journal.pone.0101649](#); pmid: [25029019](#)
  85. K. L. Akre, M. J. Ryan, Complexity increases working memory for mating signals. *Curr. Biol.* **20**, 502–505 (2010). doi: [10.1016/j.cub.2010.01.021](#); pmid: [20206525](#)
  86. B. Zhu *et al.*, Multisensory modalities increase working memory for mating signals in a treefrog. *J. Anim. Ecol.* **90**, 1455–1465 (2021). doi: [10.1111/1365-2656.13465](#); pmid: [33666233](#)
  87. L. Locatello, F. Poli, M. B. Rasotto, Context-dependent evaluation of prospective mates in a fish. *Behav. Ecol. Sociobiol.* **69**, 1119–1126 (2015). doi: [10.1007/s00265-015-1924-y](#); pmid: [26097281](#)
  88. A. M. Lea, M. J. Ryan, Irrationality in mate choice revealed by túngara frogs. *Science* **349**, 964–966 (2015). doi: [10.1126/science.1240122](#); pmid: [26315434](#)
  89. M. Plath, D. Blum, I. Schlupp, R. Tiedemann, Audience effect alters mating preferences in a livebearing fish, the Atlantic molly, *Poecilia mexicana*. *Anim. Behav.* **75**, 21–29 (2008). doi: [10.1016/j.anbehav.2007.05.013](#)
  90. A. D. Davies, Z. Lewis, L. R. Doughterty, A meta-analysis of factors influencing the strength of mate-choice copying in animals. *Behav. Ecol.* **31**, 1279–1290 (2020). doi: [10.1093/beheco/araa064](#)
  91. B. C. Jones, E. H. DuVal, Mechanisms of social influence: A meta-analysis of the effects of social information on female mate choice decisions. *Front. Ecol. Evol.* **7**, 390 (2019). doi: [10.3389/fevo.2019.000390](#)
  92. S. E. Hill, M. J. Ryan, The role of model female quality in the mate choice copying behaviour of sailfin mollies. *Biol. Lett.* **2**, 203–205 (2006). doi: [10.1098/rsbl.2005.0423](#); pmid: [17148362](#)
  93. B. G. Galef Jr, Imitation in animals: History, definition, and interpretation of data from the psychological laboratory, in *Social Learning: Psychological and Biological Perspectives* (Erlbaum, 1988), pp. 3–28.
  94. S. E. Street *et al.*, Human mate-choice copying is domain-general social learning. *Sci. Rep.* **8**, 1715 (2018). doi: [10.1038/s41598-018-19770-8](#); pmid: [29379046](#)
  95. M. Kirkpatrick, Sexual selection by female choice in polygynous animals. *Annu. Rev. Ecol. Syst.* **18**, 43–70 (1987). doi: [10.1146/annurev.es.18.110187.000355](#)
  96. M. J. Ryan, J. H. Fox, W. Wilczynski, A. S. Rand, Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. *Nature* **343**, 66–67 (1990). doi: [10.1038/343066a0](#); pmid: [2296291](#)
  97. D. R. Levitan, Effects of gamete trait on fertilization in the sea and the evolution of sexual dimorphism. *Nature* **382**, 153–155 (1996). doi: [10.1038/382153a0](#)
  98. M. Del Giudice, B. J. Crespi, Basic functional trade-offs in cognition: An integrative framework. *Cognition* **179**, 56–70 (2018). doi: [10.1016/j.cognition.2018.06.008](#); pmid: [29909281](#)
  99. R. Reber, N. Schwarz, P. Winkielman, Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* **8**, 364–382 (2004). doi: [10.1207/s15327957pspr0804\\_3](#); pmid: [15582859](#)
  100. S. V. Hulse, J. P. Renoult, T. C. Mendelson, Sexual signaling pattern correlates with habitat pattern in visually ornamented fishes. *Nat. Commun.* **11**, 2561 (2020). doi: [10.1038/s41467-020-16389-0](#); pmid: [32444815](#)
  101. L. S. Mead, S. J. Arnold, Quantitative genetic models of sexual selection. *Trends Ecol. Evol.* **19**, 264–271 (2004). doi: [10.1016/j.tree.2004.03.003](#); pmid: [16701266](#)
  102. M. Kirkpatrick, M. J. Ryan, The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33–38 (1991). doi: [10.1038/350033a0](#)
  103. Z. M. Prokop, Ł. Michalczyk, S. M. Drobnik, M. Herdegen, J. Radwan, Meta-analysis suggests choosy females get sexy sons more than "good genes." *Evolution* **66**, 2665–2673 (2012). doi: [10.1111/j.1558-5646.2012.01654.x](#); pmid: [22946794](#)
  104. M. Kirkpatrick, N. H. Barton, The strength of indirect selection on female mating preferences. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1282–1286 (1997). doi: [10.1073/pnas.94.4.1282](#); pmid: [9037044](#)
  105. P. Muralidhar, Mating preferences of selfish sex chromosomes. *Nature* **570**, 376–379 (2019). doi: [10.1038/s41586-019-1271-7](#); pmid: [31168095](#)
  106. J. Seger, R. Trivers, Asymmetry in the evolution of female mating preferences. *Nature* **319**, 771–773 (1986). doi: [10.1038/319771a0](#)
  107. C. M. Garcia, E. Ramirez, Evidence that sensory traps can evolve into honest signals. *Nature* **434**, 501–505 (2005). doi: [10.1038/nature03363](#); pmid: [15791255](#)
  108. E. D. Broder *et al.*, Evolutionary novelty in communication between the sexes. *Biol. Lett.* **17**, 20200733 (2021). doi: [10.1098/rsbl.2020.0733](#); pmid: [33529546](#)
  109. B. Hollis *et al.*, Sexual conflict drives male manipulation of female postmating responses in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8437–8444 (2019). doi: [10.1073/pnas.1821386116](#); pmid: [30962372](#)
  110. T. Bilde, A. Foged, N. Schilling, G. Arnqvist, Postmating sexual selection favors males that sire offspring with low fitness. *Science* **324**, 1705–1706 (2009). doi: [10.1126/science.1171675](#); pmid: [19556506](#)
  111. K. P. Lampert *et al.*, Determination of onset of sexual maturation and mating behavior by melanocortin receptor 4 polymorphisms. *Curr. Biol.* **20**, 1729–1734 (2010). doi: [10.1016/j.cub.2010.08.029](#); pmid: [20869245](#)
  112. R. Brooks, Negative genetic correlation between male sexual attractiveness and survival. *Nature* **406**, 67–70 (2000). doi: [10.1038/35017552](#); pmid: [10894542](#)
  113. M. J. Ryan, C. M. Pease, M. R. Morris, A genetic polymorphism in the swordtail *Xiphophorus nigrensis*: Testing the prediction of equal fitnesses. *Am. Nat.* **139**, 21–31 (1992). doi: [10.1086/285311](#)
  114. B. D. Neff, T. E. Pitcher, Mate choice for nonadditive genetic benefits and the maintenance of genetic diversity in song sparrows. *J. Evol. Biol.* **22**, 424–429 (2009). doi: [10.1111/j.1420-9101.2008.01646.x](#); pmid: [19032502](#)
  115. J. G. Robertson, Female choice increases fertilization success in the Australian frog, *Uperoleia laevigata*. *Anim. Behav.* **39**, 639–645 (1990). doi: [10.1016/S0003-3472\(05\)80374-4](#)
  116. W. Schuett, T. Tregenza, S. R. X. Dall, Sexual selection and animal personality. *Biol. Rev. Camb. Philos. Soc.* **85**, 217–246 (2010). doi: [10.1111/j.1469-185X.2009.00101.x](#); pmid: [19922534](#)
  117. P. G. Byrne, J. S. Keogh, D. M. O'Brien, J. D. Gaitan-Espitia, A. J. Silla, Evidence that genetic compatibility underpins female mate choice in a monandrous amphibian. *Evolution* **75**, 529–541 (2021). doi: [10.1111/evo.14160](#); pmid: [33389749](#)
  118. D. Andreou, C. Elizaguirre, T. Boehm, M. Milinski, Mate choice in sticklebacks reveals that immunogenes can drive ecological speciation. *Behav. Ecol.* **28**, 953–961 (2017). doi: [10.1093/beheco/axx074](#); pmid: [29622924](#)
  119. B. D. Neff, T. E. Pitcher, Genetic quality and sexual selection: An integrated framework for good genes and compatible genes. *Mol. Ecol.* **14**, 19–38 (2005). doi: [10.1111/j.1365-294X.2004.02395.x](#); pmid: [15643948](#)
  120. S. Takayama, A. Isogai, Self-incompatibility in plants. *Annu. Rev. Plant Biol.* **56**, 467–489 (2005). doi: [10.1146/annurev.arplant.56.032604.144249](#); pmid: [15862104](#)
  121. E. Mayr, *Animal Species and Evolution*. (Harvard Univ. Press 1963).
  122. T. C. Mendelson, M. D. Martin, S. M. Flaxman, Mutation-order divergence by sexual selection: Diversification of sexual signals in similar environments as a first step in speciation. *Ecol. Lett.* **17**, 1053–1066 (2014). doi: [10.1111/ele.12313](#); pmid: [24943881](#)
  123. J. C. Uyeda, S. J. Arnold, P. A. Hohenlohe, L. S. Mead, Drift promotes speciation by sexual selection. *Evolution* **63**, 583–594 (2009). doi: [10.1111/j.1558-5646.2008.00589.x](#); pmid: [19087180](#)
  124. S. Gavrillets, Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* **403**, 886–889 (2000). doi: [10.1038/35002564](#); pmid: [10706284](#)
  125. K. A. Dyer, E. R. Bewick, B. E. White, M. J. Bray, D. P. Humphreys, Fine-scale geographic patterns of gene flow and reproductive character displacement in *Drosophila subquinaria* and *Drosophila recens*. *Mol. Ecol.* **27**, 3655–3670 (2018). doi: [10.1111/mec.14825](#); pmid: [30074656](#)
  126. G. G. Rosenthal, Individual mating decisions and hybridization. *J. Evol. Biol.* **26**, 252–255 (2013). doi: [10.1111/jeb.12004](#); pmid: [23323999](#)
  127. M. R. Servedio, R. Bürger, The counterintuitive role of sexual selection in species maintenance and speciation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8113–8118 (2014). doi: [10.1073/pnas.1316484111](#); pmid: [24821767](#)
  128. M. R. Servedio, G. S. Van Doorn, M. Kopp, A. M. Frame, P. Nosil, Magic traits in speciation: 'magic' but not rare? *Trends Ecol. Evol.* **26**, 389–397 (2011). doi: [10.1016/j.tree.2011.04.005](#); pmid: [21592615](#)
  129. M. R. Servedio, J. W. Boughman, The role of sexual selection in local adaptation and speciation. *Annu. Rev. Ecol. Syst.* **48**, 85–109 (2017). doi: [10.1146/annurev-ecolsys-110316-022905](#)
  130. O. Seehausen *et al.*, Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008). doi: [10.1038/nature07285](#); pmid: [18833272](#)
  131. D. J. Yeh, M. R. Servedio, Reproductive isolation with a learned trait in a structured population. *Evolution* **69**, 1938–1947 (2015). doi: [10.1111/evo.12688](#); pmid: [26031568](#)
  132. Y. Yang, M. R. Servedio, C. L. Richards-Zawacki, Imprinting sets the stage for speciation. *Nature* **574**, 99–102 (2019). doi: [10.1038/s41586-019-1599-z](#); pmid: [31578486](#)
  133. R. A. Fisher, *The Genetical Theory of Natural Selection*. (Oxford Univ. Press, Oxford, 1930).
  134. C. Chen, K. S. Plennig, Female toads engaging in adaptive hybridization prefer high-quality heterospecifics as mates. *Science* **367**, 1377–1379 (2020). doi: [10.1126/science.aaz5109](#); pmid: [32193328](#)
  135. R. Abbott *et al.*, Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013). doi: [10.1111/j.1420-9101.2012.02599.x](#); pmid: [23323997](#)
  136. G. F. Turner, M. T. Burrows, A model of sympatric speciation by sexual selection. *Proc. Biol. Sci.* **260**, 287–292 (1995). doi: [10.1098/rspb.1995.0093](#)
  137. M. N. Verzijden, R. F. Lachlan, M. R. Servedio, Female mate-choice behavior and sympatric speciation. *Evolution* **59**, 2097–2108 (2005). doi: [10.1111/j.0014-3820.2005.tb00920.x](#); pmid: [16405155](#)
  138. D. J. Yeh, Assortative mating by an obliquely transmitted local cultural trait promotes genetic divergence: A model. *Am. Nat.* **193**, 81–92 (2019). doi: [10.1086/7009583](#); pmid: [30624103](#)
  139. M. J. Ryan *et al.*, Nineteen years of consistently positive and strong female mate preferences despite individual variation. *Am. Nat.* **194**, 125–134 (2019). doi: [10.1086/704103](#); pmid: [31318282](#)
  140. M. R. Servedio, J. M. Powers, R. Lande, T. D. Price, Evolution of sexual cooperation from sexual conflict. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23225–23231 (2019). doi: [10.1073/pnas.1904138116](#); pmid: [31611370](#)
  141. J. G. Cally, D. Stuart-Fox, L. Holman, Meta-analytic evidence that sexual selection improves population fitness. *Nat. Commun.* **10**, 1017 (2019). doi: [10.1038/s41467-019-10074-7](#); pmid: [31043615](#)
  142. G. U. C. Lehmann, A. W. Lehmann, Material benefit of mating: The bushcricket spermatophylax as a fast uptake nuptial gift. *Anim. Behav.* **112**, 267–271 (2016). doi: [10.1016/j.anbehav.2015.12.022](#)
  143. K. Suzuki, N. Juni, D. Yamamoto, Enhanced mate refusal in female *Drosophila* induced by a mutation in the spinster locus. *Appl. Entomol. Zool.* **32**, 235–243 (1997). doi: [10.1303/aaz.32.235](#)
  144. T. Bilde, A. A. Maklakov, K. Meisner, L. la Guardia, U. Friberg, Sex differences in the genetic architecture of lifespan in a seed beetle: Extreme inbreeding extends male lifespan. *BMC Evol. Biol.* **9**, 33 (2009). doi: [10.1186/1471-2148-9-33](#); pmid: [19200350](#)
  145. G. S. Wilkinson, P. R. Reillo, Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly. *Proc. Biol. Sci.* **255**, 1–6 (1994). doi: [10.1098/rspb.1994.0001](#)
  146. M. Rangassamy, M. Dalmaz, C. Féron, P. Gouat, H. G. Rödel, Similarity of personalities speeds up reproduction in pairs of a monogamous rodent. *Anim. Behav.* **103**, 7–15 (2015). doi: [10.1016/j.anbehav.2015.02.007](#)

## ACKNOWLEDGMENTS

We thank M. Kirkpatrick, C. Fitzpatrick, O. Dorsey, and especially an anonymous reviewer and T. Mendelson for their helpful criticisms and C. M. Rosenthal Struminger for the original Fig. 1 illustration and line art in Fig. 3.

10.1126/science.abi6308



## RESEARCH ARTICLE SUMMARY

## CORONAVIRUS

## Structural basis for continued antibody evasion by the SARS-CoV-2 receptor binding domain

Katherine G. Nabel†, Sarah A. Clark†, Sundaresh Shankar†, Junhua Pan†, Lars E. Clark, Pan Yang, Adrian Coscia, Lindsay G. A. McKay, Haley H. Varnum, Vesna Brusic, Nicole V. Tolan, Guohai Zhou, Michaël Desjardins, Sarah E. Turbett, Sanjat Kanjilal, Amy C. Sherman, Anand Dighe, Regina C. LaRocque, Edward T. Ryan, Casey Tylek, Joel F. Cohen-Solal, Anhdao T. Darcy, Davide Tavella, Anca Clabbers, Yao Fan, Anthony Griffiths, Ivan R. Correia, Jane Seagal, Lindsey R. Baden, Richelle C. Charles, Jonathan Abraham\*

**INTRODUCTION:** As severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replicates under selective pressure from natural and vaccine-induced immunity, variants of concern (VOCs) continue to emerge. Through adaptative evolution, these variants acquire mutations in the spike protein receptor binding domain (RBD) that binds the cellular receptor angiotensin-converting enzyme 2 (ACE2). The effects of spike protein mutations on immune responses make it important to monitor viral variants. While previously studied VOCs contain one to three RBD mutations that at times overlap, the potential for composite variants that contain larger numbers of mutations is being closely monitored.

**RATIONALE:** As parts of the world continue to face waves of infection and mitigation strategies are relaxed, viral replication in human hosts under antibody selective pressure continues to shape the antigenic landscape of the SARS-CoV-2 spike protein. As variants containing composite mutations begin to emerge, proactive approaches

examining the impact of variants before they become dominant strains are critical.

**RESULTS:** We determined the x-ray crystal structure of human ACE2 in complex with a SARS-CoV-2 RBD that contains six substitutions that arose during persistent infection of an immunocompromised individual. We found that structural plasticity at the RBD-ACE2 interface allowed the RBD to tolerate a large number of mutations while retaining ACE2 affinity. We generated a panel of pseudotypes bearing composite RBD mutations (up to seven) from immunocompromised host-derived sequences and VOCs. Composite variants more adeptly evaded therapeutic antibody neutralization than did previously studied VOCs. After first immunization but before the second dose of an mRNA vaccine, we observed a loss in vaccine recipient serum neutralizing activity for all variants tested, although the severity differed depending on the variant. However, sampling after the second immunization revealed detectable neutralizing activity against all variants in the serum of

vaccine recipients, including against a pseudotype that contains seven composite RBD mutations [denoted receptor binding mutant-2 (RBM-2)]. To identify evolutionary barriers that restrict neutralization breadth, we used the SARS-CoV spike protein to isolate a neutralizing antibody from a COVID-19 convalescent donor. Through structural analysis and functional assays, we show that N-linked glycan acquisition by the SARS-CoV-2 RBD confers pseudotype resistance to neutralization by the isolated cross-reactive antibody and at least one other antibody that binds a similar, otherwise highly conserved epitope. Therefore, acquisition of an N-linked glycan on the SARS-CoV-2 RBD is an additional means through which the virus could continue to evade immune responses.

**CONCLUSION:** We find that accumulation of large numbers of RBD mutations is facilitated by structural plasticity at the RBD-ACE2 interface and further erodes the activity of therapeutic antibodies and serum from vaccine recipients. Furthermore, acquisition of an N-linked glycan on the SARS-CoV-2 RBD is an additional neutralization escape pathway that should be closely monitored during viral antigenic drift. ■

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

\*Corresponding author. Email: jonathan\_abraham@hms.harvard.edu

†These authors contributed equally to this work.

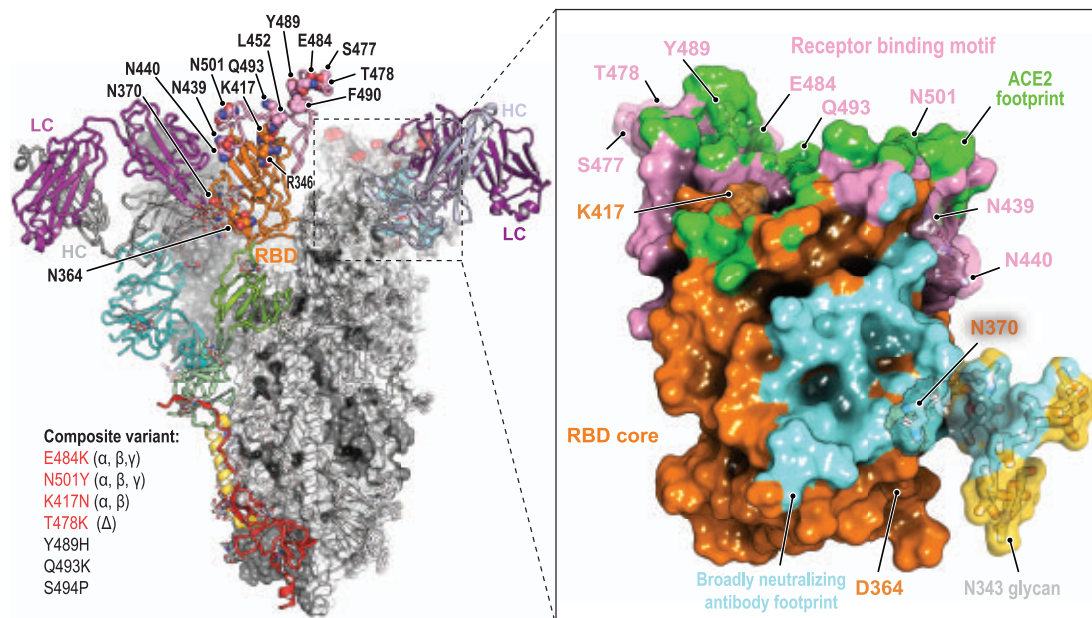
This is an open-access article distributed under the terms of the Creative Commons Attribution license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cite this article as K. G. Nabel et al., *Science* 375, eabl6251 (2022). DOI: 10.1126/science.abl6251

**READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abl6251>

### Immune escape at the SARS-CoV-2 spike protein RBD.

Structural plasticity accommodates the accumulation of composite substitutions in the RBD ACE2 binding site and allows the RBD to adeptly escape therapeutic antibodies. Cross-neutralizing antibodies bind the RBD core, but acquisition of an N-linked glycan at RBD residue Asn<sup>370</sup> (N370) drives further neutralization escape. Single-letter abbreviations for the amino acid residues are as follows: D, Asp; E, Glu; F, Phe; H, His; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; and Y, Tyr. LC, light chain; HC, heavy chain.



## RESEARCH ARTICLE

## CORONAVIRUS

## Structural basis for continued antibody evasion by the SARS-CoV-2 receptor binding domain

Katherine G. Nabel<sup>1†</sup>, Sarah A. Clark<sup>1†</sup>, Sundaresh Shankar<sup>1†</sup>, Junhua Pan<sup>1†</sup>, Lars E. Clark<sup>1</sup>, Pan Yang<sup>1</sup>, Adrian Coscia<sup>1</sup>, Lindsay G. A. McKay<sup>2</sup>, Haley H. Varnum<sup>1</sup>, Vesna Brusic<sup>1</sup>, Nicole V. Tolan<sup>3</sup>, Guohai Zhou<sup>4</sup>, Michaël Desjardins<sup>5,6</sup>, Sarah E. Turbett<sup>7,8</sup>, Sanjat Kanjilal<sup>5,9</sup>, Amy C. Sherman<sup>5</sup>, Anand Dighe<sup>8</sup>, Regina C. LaRocque<sup>7</sup>, Edward T. Ryan<sup>7,10</sup>, Casey Tylek<sup>11</sup>, Joel F. Cohen-Solal<sup>11</sup>, Anhdao T. Darcy<sup>11</sup>, Davide Tavella<sup>11</sup>, Anca Clabbers<sup>11</sup>, Yao Fan<sup>11</sup>, Anthony Griffiths<sup>2</sup>, Ivan R. Correia<sup>11</sup>, Jane Seagal<sup>11</sup>, Lindsey R. Baden<sup>4,5,12</sup>, Richelle C. Charles<sup>7</sup>, Jonathan Abraham<sup>1,5,12,13\*</sup>

Many studies have examined the impact of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants on neutralizing antibody activity after they have become dominant strains. Here, we evaluate the consequences of further viral evolution. We demonstrate mechanisms through which the SARS-CoV-2 receptor binding domain (RBD) can tolerate large numbers of simultaneous antibody escape mutations and show that pseudotypes containing up to seven mutations, as opposed to the one to three found in previously studied variants of concern, are more resistant to neutralization by therapeutic antibodies and serum from vaccine recipients. We identify an antibody that binds the RBD core to neutralize pseudotypes for all tested variants but show that the RBD can acquire an N-linked glycan to escape neutralization. Our findings portend continued emergence of escape variants as SARS-CoV-2 adapts to humans.

As severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continues to replicate in humans under selective pressure from natural and vaccine-induced immunity, variants of concern (VOCs) with increased transmissibility or virulence continue to emerge (1). Through adaptive evolution, these variants acquire mutations in the spike (S) protein receptor binding domain (RBD) that binds the cellular receptor human angiotensin-converting enzyme 2 (ACE2) (1–3). Many of these mutations are within the RBD receptor binding motif (RBM), a hypervariable loop that mediates most of the ACE2 contacts (2, 3). The RBD is the primary target of

neutralizing antibodies in naturally acquired or vaccine-elicited humoral immunity (4, 5). The spike protein N-terminal domain (NTD) is also a target of neutralizing antibodies, and VOCs have NTD mutations that include deletions at an antigenic supersite for neutralizing antibody binding (6, 7). The effects of spike protein mutations on immune responses (8–13) make it important to monitor viral variants.

While previously studied VOCs contain one to three RBD mutations that at times overlap (1), the potential for composite variants is being closely monitored. For example, the B.1.617.2 (Delta) variant can acquire the K417N<sub>RBD</sub> mutation (Lys<sup>417</sup>→Asn) found in the B.1.351 (Beta) variant, generating the Delta AY.2 variant, for a total of three RBD mutations (Fig. 1A). Similarly, as shown in recently deposited sequences from samples collected in Angola, the Beta variant can acquire the L452R<sub>RBD</sub> mutation found in the Delta and B.1.429/427 (Epsilon) variants, for a total of four RBD mutations (Fig. 1A and table S1). Further complicating matters, variant monitoring efforts are still undersampling viral evolution. For example, a virus recently sequenced from travelers returning from Tanzania contained a previously undocumented combination of RBD mutations (E484K<sub>RBD</sub>, T478R<sub>RBD</sub>, and R346K<sub>RBD</sub>) with NTD deletions that would likely alter the spike protein antigenic surface and result in antibody escape (table S1).

Here, we investigate the structural plasticity of the SARS-CoV-2 spike protein RBD and its capacity to evade neutralizing antibodies.

## Results

## Structure of an evolved receptor binding domain–ACE2 complex

We previously generated two SARS-CoV-2 spike proteins that each contain six RBD changes that were detected during persistent infection of an immunocompromised individual infected with a SARS-CoV-2 strain containing the D614G<sub>S</sub> mutation (14–16). This individual received treatment with REGN-COV2 (17, 18), but several of the RBD substitutions had occurred even before administration of this therapeutic antibody cocktail (14–16). Lentivirus pseudotypes bearing these spike proteins, denoted day 146\* and day 152\* (Fig. 1A and table S2), were refractory to neutralization by V<sub>H</sub>3-53 heavy chain gene–derived neutralizing antibodies, a potent class of neutralizing antibodies that have been repeatedly isolated from convalescent donors (19–25). These pseudotypes were also resistant to neutralization by components of REGN-COV2 (17, 18) and by polyclonal immunoglobulin G (IgG) purified from the serum of COVID-19 convalescent donors (14). Substitutions in the day 146\* and day 152\* spike proteins, noted in samples sequenced from this individual in the spring and summer of 2020, foreshadowed those in currently circulating VOCs at three positions: N501<sub>RBD</sub>, E484<sub>RBD</sub>, and T478<sub>RBD</sub> (Fig. 1A). The day 146\* and day 152\* spike proteins also contain substitutions that are not in current dominant strains but could have serious effects if acquired. For example, the S494P<sub>RBD</sub> substitution is a therapeutic antibody (LY-CoV555) escape mutation (26) that, as of 27 September 2021, was present in more than 12,000 human-derived SARS-CoV-2 sequences on public research databases (GISAID) (27). Additionally, the Q493K<sub>RBD</sub> mutation, which is found in more than 100 human-derived SARS-CoV-2 sequences on GISAID as of 27 September 2021, confers resistance to multiple therapeutic antibodies [REGN10933, CB6 (LY-CoV016), and LY-CoV555] and V<sub>H</sub>3-53 gene–derived antibodies (14, 16, 17, 28).

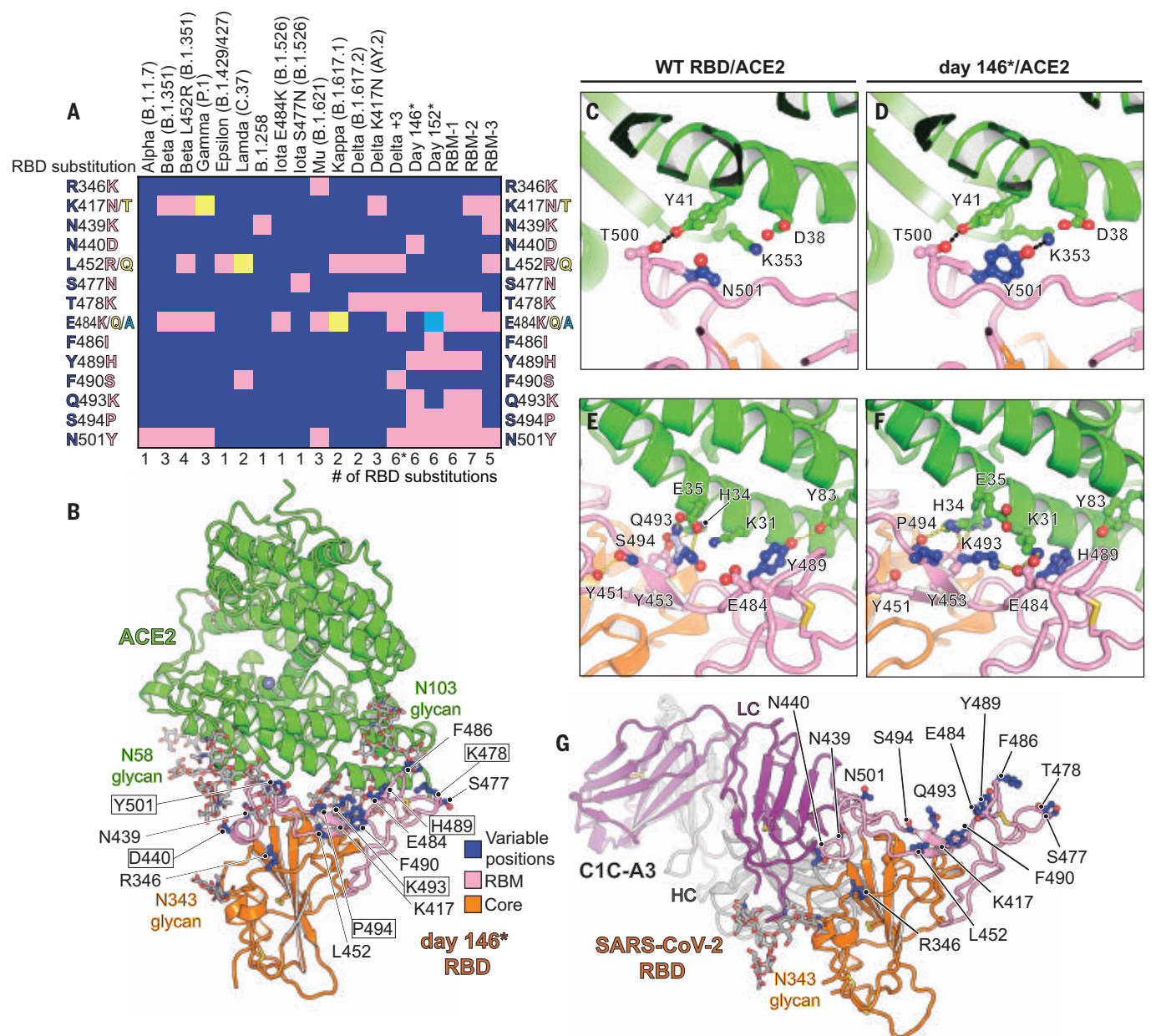
To determine the impact of their combined mutations on human ACE2 binding, we generated recombinant RBDs for the day 146\* and day 152\* spike protein mutants. The affinity of the day 152\* mutant monomeric RBD for the monomeric ACE2 ectodomain was substantially lower (binding affinity,  $K_D$ , of 2.4  $\mu$ M) than that of wild-type (Wuhan-Hu-1) RBD (54 nM, consistent with other reports) (3, 29), suggesting that its mutations compromise ACE2 binding (fig. S1 and table S3). For comparison, the affinity we measured of the SARS-CoV RBD for human ACE2 was 0.26  $\mu$ M, about ninefold higher than the affinity for the day 152\* RBD (fig. S1 and table S3). The affinity of the day 152\* RBD for ACE2 is comparable to that of the RBDs of some bat coronaviruses that are closely related to SARS-CoV-2 and

<sup>1</sup>Department of Microbiology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. <sup>2</sup>Department of Microbiology and National Emerging Infectious Diseases Laboratories, Boston University School of Medicine, Boston, MA 02118, USA. <sup>3</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA. <sup>4</sup>Center for Clinical Investigation, Brigham and Women's Hospital, Boston, MA 02115, USA. <sup>5</sup>Division of Infectious Diseases, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. <sup>6</sup>Division of Infectious Diseases, Department of Medicine, Centre Hospitalier de l'Université de Montréal, Montreal QC H2X 0C1, Canada. <sup>7</sup>Division of Infectious Diseases, Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA. <sup>8</sup>Department of Pathology, Massachusetts General Hospital, Boston, MA 02114, USA. <sup>9</sup>Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA 02215, USA. <sup>10</sup>Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. <sup>11</sup>AbbVie Bioresearch Center, Worcester, MA 01605, USA. <sup>12</sup>Massachusetts Consortium on Pathogen Readiness, Boston, MA, USA. <sup>13</sup>Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA.

\*Corresponding author. Email: jonathan\_abraham@hms.harvard.edu

†These authors contributed equally to this work.





**Fig. 1. Structure of intrahost evolved RBD bound to human ACE2.** (A) Key RBD substitutions discussed in the text and the SARS-CoV-2 variants that contain them. 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. (B) Day 146\* RBD-ACE2 ectodomain x-ray crystal structure. RBD residues that are mutated in variants discussed in the text are shown. Boxed residues are mutated in the day 146\* RBD as compared with the Wuhan-Hu-1 (wild-type)

SARS-CoV-2 RBD. The Delta +3 variant contains an additional RBD mutation that is not shown in the schematic diagram (see table S2). (C) Wild-type RBD-ACE2 contacts near N501<sub>RBD</sub> [Protein Data Bank (PDB) ID 6MOJ] (2). (D) Day 146\* RBD contacts near Y501<sub>RBD</sub>. (E) Wild-type SARS-CoV-2 RBD-ACE2 interactions near Q493<sub>RBD</sub>. (F) Day 146\* RBD interactions near K493<sub>RBD</sub>. (G) Cryo-EM structure of the SARS-CoV-2 RBD bound to the C1C-A3 antibody Fab. RBD residues discussed in the text are labeled. LC, light chain; HC, heavy chain.

bind human ACE2 (e.g., RaTG13 virus RBD affinity of 3.9  $\mu$ M) (30). The day 146\* RBD, however, had a similar affinity ( $K_d$  of 46 nM) for ACE2 as that of the Wuhan-Hu-1 SARS-CoV-2 RBD (fig. S1 and table S3).

We determined the x-ray crystal structure of the day 146\* RBD bound to the human ACE2 ectodomain (Fig. 1B, fig. S2, and table S4). This structure is similar to previously determined

structures of ACE2-SARS-CoV-2 RBD complexes (2, 3), except we observed contacts between two N-linked glycans on ACE2 (attached to N53<sub>ACE2</sub> and N90<sub>ACE2</sub>) and the RBD (fig. S3). Removing the N90<sub>ACE2</sub> glycan, which interacts with the RBD in both copies of the crystal asymmetric unit (fig. S3), increased Wuhan-Hu-1 SARS-CoV-2 and day 146\* RBD affinity for ACE2, although the effect was modest (fig.

S1 and table S3). This finding is consistent with prior work implicating the N90<sub>ACE2</sub> glycan, which is removed in a human polymorphism (T92I<sub>ACE2</sub>), as a barrier to SARS-CoV-2 RBD binding to ACE2 (31, 32).

The N501Y<sub>RBD</sub> substitution is found in multiple VOCs (Fig. 1A); once it surfaced in the immunocompromised individual, it was retained at later time points (14–16). As also



shown in a cryo-electron microscopy (cryo-EM) structure of the SARS-CoV-2 spike protein containing the N501Y<sub>RBD</sub> substitution bound to ACE2 (33), the side chain of Y501<sub>RBD</sub> interacts with Y41<sub>ACE2</sub> and K353<sub>ACE2</sub> with no notable structural change (Fig. 1, C and D). E484<sub>RBD</sub> is a critical target of antibodies against SARS-CoV-2 and is mutated in several variants (12, 34, 35). In structures of Wuhan-Hu-1 SARS-CoV-2 RBD bound to ACE2, E484<sub>RBD</sub> is near but does not directly contact the receptor (Fig. 1E). In the day 146\* RBD–ACE2 complex structure, the K493<sub>RBD</sub> side chain reaches over the RBD surface to recruit the E484<sub>RBD</sub> side chain to form a new salt bridge with K31<sub>ACE2</sub> (Fig. 1F). The nearby Y489<sub>RBD</sub> mutation, which removes a polar contact with ACE2, better accommodates repositioning of E484<sub>RBD</sub> because the histidine is smaller than the tyrosine side chain and would avoid potential steric clashes with E484<sub>RBD</sub> in this binding mode (Fig. 1, E and F). A second rotamer for residue H34<sub>ACE2</sub> forms additional RBD contacts to fill a gap created by the reorganization of local interactions (Fig. 1, E and F). This structural plasticity may explain how the RBD tolerates an unexpectedly large number of mutations during intrahost evolution yet retains the ability to bind ACE2 tightly. It is also consistent with the large sequence divergence in the RBD residues that contact ACE2 among SARS-related coronaviruses that share this cellular receptor.

#### Neutralization escape of therapeutic antibodies

RBD-targeting antibodies can be categorized into classes on the basis of whether they bind an overlapping footprint with ACE2 and recognize only an open or both an open and a closed RBD on the spike protein trimer (36). CB6 (equivalent to LY-CoV016 or etesevimab) is a class 1, V<sub>H</sub>3-66-derived antibody that blocks ACE2 binding and can only bind the RBD when it is open, and LY-CoV555 (bamlanivimab) is a class 2 antibody that blocks ACE2 binding but can bind both open and closed RBDs (21, 37). LY-CoV016 and LY-CoV555 are used together as a cocktail and bind epitopes that partially overlap on the RBM such that both cannot bind simultaneously (21, 37). REGN10933 is a class 1 antibody, and REGN10987 is a class 3 antibody that sterically blocks ACE2 binding but binds the RBM outside the main ACE2 binding site; both are used as a cocktail (REGN-COV2) (17, 18).

Structural plasticity at the RBD–ACE2 interface suggests that the RBD could tolerate many more mutations than found in current VOCs. We next generated pseudotypes for spike protein variants that contain composite mutations. The Delta variant, which contains the L452<sub>RBD</sub> and T478<sub>RBD</sub> substitutions, has become a dominant strain across the globe (38). We generated a pseudotype for the Delta AY.2 variant, which contains the K417N<sub>RBD</sub>

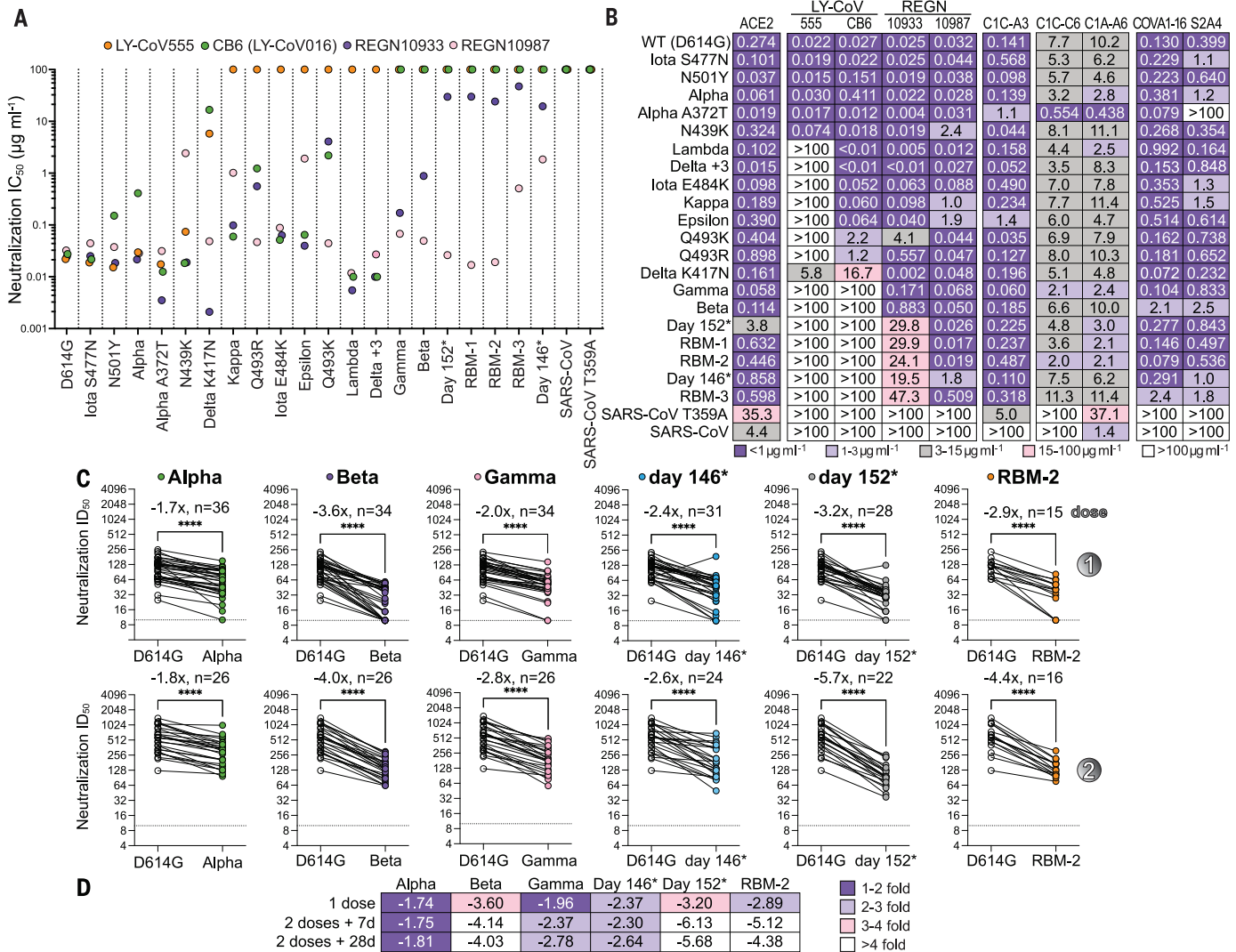
mutation that is usually found in the Beta variant, and a Delta variant containing the N501Y<sub>RBD</sub>, E484K<sub>RBD</sub>, and F490S<sub>RBD</sub> mutations usually found in the Beta, P.1 (Gamma), and C.37 (Lambda) variants (referred to here as “Delta +3”) (Fig. 1A and tables S1 and S2). The set of RBD mutations for the latter strain occurred in deposited sequences from samples collected in Turkey (table S1). We also generated pseudotypes in which we combined spike protein substitutions detected in the immunocompromised host with mutations found in the Beta variant, which we chose because this VOC is highly resistant to antibody neutralization (10, 12, 39). Starting with a day 146\* spike protein sequence, which contains an NTD deletion, we incorporated either one (E484K<sub>RBD</sub>) or two (E484K<sub>RBD</sub> and K417N<sub>RBD</sub>) additional substitutions; these are referred to as receptor binding mutant-1 (RBM-1) and RBM-2, respectively (Fig. 1A and table S2). Additionally, starting with the Beta variant spike protein sequence, we generated a variant pseudotype that contains two additional mutations associated with immune evasion (L452R<sub>RBD</sub> and N439K<sub>RBD</sub>) (40, 41). This pseudotype is referred to as RBM-3 (Fig. 1A and table S2). An ACE2-Fc fusion protein neutralized RBM-1, RBM-2, and RBM-3 pseudotypes, suggesting that all entered cells by binding ACE2 (Fig. 2B and fig. S4A).

We tested the activity of therapeutic antibodies against Delta AY.2, Delta +3, RBM-1, RBM-2, RBM-3, and additional variant pseudotypes with known resistance profiles to serve as comparators in the same assay (Fig. 2, A and B, and fig. S4A). LY-CoV555 was the most affected by escape mutations, followed by CB6 (from which LY-CoV016 is derived) (Fig. 2, A and B, and fig. S4A). The Q493K<sub>RBD</sub> mutation conferred absolute resistance to LY-CoV555, generated 80-fold resistance to CB6, and also compromised REGN10933 activity, consistent with previous reports (Fig. 2, A and B, and fig. S4A) (14, 16, 17, 26). In addition to the expected loss of activity of LY-CoV555 and CB6 against Beta and Gamma variants (9, 11, 12, 42), LY-CoV555 and CB6 lost all activity against day 146\*, day 152\*, RBM-1, RBM-2, and RBM-3 pseudotypes (Fig. 2, A and B, and fig. S4A). Whereas the Delta variant is known to resist neutralization by LY-CoV555 but retain sensitivity to neutralization by CB6/LY-CoV016 (38), the Delta AY.2 pseudotype was resistant to both agents (Fig. 2, A and B, and fig. S4A). This is expected because CB6/LY-CoV016 is derived from a V<sub>H</sub>3-66 antibody (21), and the additional mutation the Delta AY.2 variant contains with respect to Delta (K417N<sub>RBD</sub>) confers resistance to CB6/LY-CoV016 and other members of the V<sub>H</sub>3-53 and V<sub>H</sub>3-66 class of neutralizing antibodies (9, 14, 16, 26, 43). The Delta +3 pseudotype, which despite containing six RBD mutations does not contain the

K417N<sub>RBD</sub> substitution, only escaped neutralization by LY-CoV555 (Fig. 2, A and B; fig. S4A; and table S2). Although the distribution of LY-CoV016 and LY-CoV555 was paused in the United States in the summer of 2021 as the prevalence of Gamma and Beta VOCs increased, the distribution of this antibody cocktail has since been resumed with the rise of Delta as the predominant strain. However, our findings emphasize the importance of close monitoring of Delta AY.2 and of other Delta variants for acquisition of the K417N<sub>RBD</sub> mutation.

Although REGN10933 lost substantial activity against the Beta variant, which is consistent with other reports (9, 12, 42), it still had a median inhibitory concentration (IC<sub>50</sub>) value of <1 µg ml<sup>−1</sup> in our assays (Fig. 2, A and B, and fig. S4A). However, resistance markedly worsened with the day 146\*, day 152\*, RBM-1, RBM-2, and RBM-3 pseudotypes, with 800- to 1900-fold loss of neutralizing activity (IC<sub>50</sub> values ranging from 20 to 47 µg ml<sup>−1</sup>). REGN10987 potentially neutralized many of the variant pseudotypes we examined. While we observed the expected resistance to REGN10987 neutralization by variants containing the N439K<sub>RBD</sub> or the adjacent N440D<sub>RBD</sub> substitutions (14, 16), we also observed some loss of activity against Epsilon and B.1.617.1 (Kappa), which was not expected because none of their substitutions fall within the REGN10987 RBD footprint (Fig. 1A and Movie 1). Nonetheless, other reports have also noted varying degrees of modest in vitro resistance of Epsilon and Kappa variants to REGN10987 neutralization (39, 42). Notably, the day 146\* and RBM-3 pseudotypes were the only ones to gain resistance to both antibodies in REGN-COV2, because they contain substitutions in the REGN10933 (e.g., Q493K<sub>RBD</sub>, or E484K<sub>RBD</sub> and K417N<sub>RBD</sub>) and the REGN10987 binding sites (N439K<sub>RBD</sub> or N440D<sub>RBD</sub>) (Fig. 2, A and B; fig. S4A; and Movie 1) (14). We observed on GISAID instances of “day 146\*-like” viruses that would be expected to resist neutralization by LY-CoV555, LY-CoV016, REGN10933, and REGN10987, because they contain the Q493K<sub>RBD</sub> and N439K<sub>RBD</sub> substitutions. One strain contains the N501Y<sub>RBD</sub>, Q493K<sub>RBD</sub>, and N439K<sub>RBD</sub> mutations (sequenced once in South Africa), and the other contains the N501Y<sub>RBD</sub>, Q493K<sub>RBD</sub>, L452R<sub>RBD</sub>, N439K<sub>RBD</sub>, and N440F<sub>RBD</sub> mutations (sequenced once in the United Kingdom) (table S1).

The broadly neutralizing antibody S309 (44), a class 3 antibody that binds the RBD but does not interfere with ACE2 binding and from which the therapeutic antibody sotrovimab is derived, was active against all variants we tested (fig. S4A). However, we could not calculate reliable neutralization IC<sub>50</sub> values because of variable non-neutralizable pseudotype fractions (fig. S4A). The presence of a non-neutralizable fraction is unexplained but has



**Fig. 2. Neutralization escape from therapeutic antibodies and mRNA vaccine-elicited serum.** (A) Summary of neutralization IC<sub>50</sub> values for lentivirus pseudotypes with the indicated monoclonal antibodies. (B) Tabulated IC<sub>50</sub> values for lentivirus pseudotypes with the indicated monoclonal antibodies and an ACE2-Fc fusion protein (ACE2). (C) Mean ID<sub>50</sub> neutralization titers for the indicated variant pseudotypes at the time of the second immunization but before

vaccination ("dose 1"), or 28 days after second immunization ("dose 2") with mRNA-1273 or BNT162b2. The fold change of the mean ID<sub>50</sub> neutralization titer with respect to D614G<sub>S</sub> pseudotype is shown in each panel. Each experiment was performed twice independently in triplicate (*n* = 6). Wilcoxon matched-pairs signed rank test; \*\*\*\**P* < 0.0001. (D) Tabulated fold change of mean ID<sub>50</sub> neutralization titers for the indicated pseudotypes as compared with D614G<sub>S</sub> pseudotype.

been noted in other reports when human cells overexpressing ACE2, as opposed to African green monkey (Vero) target cells, were used to examine S309 neutralizing activity (45, 46).

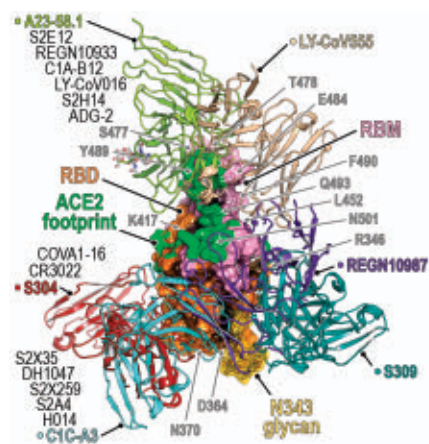
**Neutralization escape of mRNA vaccine recipient sera**

Messenger RNA (mRNA)-based vaccines encoding the SARS-CoV-2 spike protein elicit robust neutralizing antibody responses (47–49). We directly compared immune evasion of the day 146\*, day 152\*, and RBM-2 pseudotypes to the B.1.1.7 (Alpha), Beta, and Gamma pseudotypes in sera obtained from individuals who had received a two-dose series of an mRNA vaccine (BNT162b2 or mRNA-1273) (Fig. 2, C

and D, and fig. S5). In addition to RBD substitutions, day 146\*, day 152\*, RBM-1, and RBM-2 spike proteins all contain NTD deletions spanning residues 141 to 144, which are positioned near NTD mutations in Alpha, Beta, and Delta in a key antigenic supersite (table S2) (6, 7). As similar NTD deletions found in Alpha, Beta, and Delta prevent binding of some neutralizing antibodies (6, 7, 46), they would be expected to escape neutralization by some NTD-targeting antibodies in addition to escaping neutralization by RBD-targeting antibodies. After initial immunization but before the second dose, we observed a loss in neutralizing activity for all variants, although the severity of this loss varied. Variants that

contain any substitution at E484<sub>RBD</sub> combined with an NTD deletion (Beta, day 152\*, and RBM-2) were more effective at evading antibody responses than variants that had an E484<sub>RBD</sub> substitution without an NTD deletion (Gamma) or an NTD deletion but no E484<sub>RBD</sub> substitution (day 146\*) (Fig. 2, C and D; fig. S5; and table S2). These findings are consistent with the role of E484<sub>RBD</sub> as a major driver in neutralization escape of polyclonal antibody responses to SARS-CoV-2 (35) and observations that Beta more robustly escapes antibody neutralization than Gamma (9, 13). They further suggest that variants that have an NTD supersite deletion and an E484<sub>RBD</sub> substitution are the most concerning





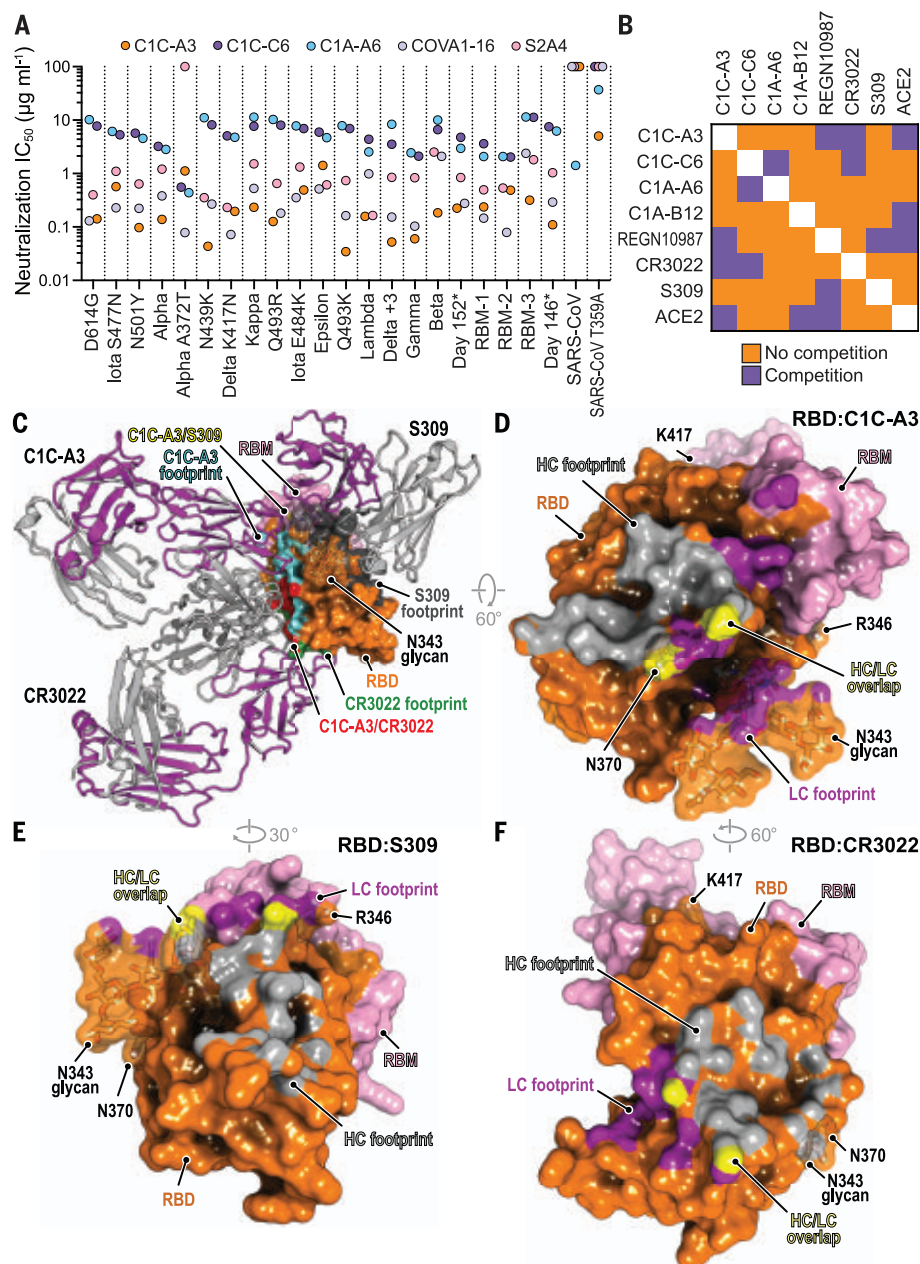
**Movie 1. Antibody footprints on an evolving SARS-CoV-2 RBD.** Antibodies are classified according to Barnes *et al.* (36). PDB IDs are listed in parentheses. Key RBD residues discussed in the main text are highlighted.

when it comes to resistance to polyclonal antibodies.

One-quarter of sampled individuals had no detectable activity against the Beta and RBM-2 pseudotypes after a single immunization (Fig. 2, C and D). However, sampling at 7 and 28 days after the second immunization revealed detectable neutralizing activity against all variants in all vaccine recipients, including against the RBM-2 pseudotype, which contains seven RBD mutations (Fig. 2, C and D, and fig. S5). Thus, repeated administration of an mRNA vaccine encoding constructs of the SARS-CoV-2 spike protein used in current formulations may provide sufficient neutralizing antibody breadth and potency to yield baseline serum neutralizing activity against variants that are more extensively mutated than the current dominant strains.

#### Identification of SARS-CoV cross-reactive antibodies

The RBD is also the major target of neutralizing antibodies against SARS-CoV, which caused a small outbreak of viral pneumonia from 2003 to 2004, although with a much higher case fatality rate (50, 51). Polyclonal antibody responses against SARS-CoV-2 poorly cross-neutralize SARS-CoV (52, 53). To identify barriers that restrict neutralization breadth, we performed single memory B cell sorting with the SARS-CoV spike protein to mine the memory B cell repertoire of a COVID-19 convalescent individual (“C1”). Polyclonal IgG from C1 plasma neutralized SARS-CoV-2 pseudotype but had weak activity against SARS-CoV pseudotype (fig. S6A). From C1 peripheral blood mononuclear cells, using a prefusion stabilized SARS-CoV spike protein (S2P) (54), we cloned 17 cross-reactive antibodies. Of these, 11 antibodies bound both the SARS-CoV and



**Fig. 3. Neutralization of SARS-CoV-2 variants by an RBD core-targeting antibody.** (A) Summary of neutralization  $IC_{50}$  values for pseudotypes and the indicated antibodies. (B) Summary of the results of BLI-based competition assays. (C) Superposition of the CR3022 (PDB ID 6W41) (55) and S309 (PDB ID 6WPS) (44) structures onto the C1C-A3-bound RBD structure. Antibody Fabs are shown as ribbon diagrams, and the RBD is shown in surface representation. Antibody footprints are shown on the RBD surface. (D) RBD footprint of C1C-A3. (E) RBD footprint of S309 (PDB ID 6WPS) (44). (F) RBD footprint of CR3022 (PDB ID 6W41) (55). In panels (D) to (F), key RBD residues discussed in the main text are highlighted.

the SARS-CoV-2 spike protein in an enzyme-linked immunosorbent assay (ELISA) (fig. S6C and table S5). Only two RBD-binding antibodies, C1C-A3 (“A3”) and C1C-C6 (“C6”), neutralized SARS-CoV-2 pseudotypes in our assays (Figs. 2B and 3A and fig. S6F). Despite binding to the SARS-CoV spike protein and RBD by ELISA, A3 and C6 did not neutralize SARS-CoV pseudotype (fig. S6, F and G). We also included

C1A-A6 (“A6”) in these assays, a SARS-CoV-2 neutralizing antibody we previously isolated from the C1 donor using prefusion stabilized SARS-CoV-2 S2P in single B cell sorting experiments (14). Unlike A3 and C6, A6 neutralized SARS-CoV pseudotypes (Figs. 2B and 3A and fig. S6F). We determined Fab RBD binding affinities using biolayer interferometry (BLI) (fig. S7 and table S3) and confirmed A3 and



A6 activity against infectious SARS-CoV-2 in a plaque reduction neutralization assay (fig. S4B).

A3 neutralized almost all SARS-CoV-2 variant pseudotypes with a neutralization  $IC_{50}$  value of  $<1 \mu g \text{ ml}^{-1}$ , including Beta, Gamma, Delta AY.2, Delta +3, RBM-1, RBM-2, and RBM-3 pseudotypes; the Epsilon variant was the only exception, with an  $IC_{50}$  value of  $1.9 \mu g \text{ ml}^{-1}$  (Figs. 2B and 3A and fig. S4A). C6 and A6 also broadly neutralized variants, but with higher baseline  $IC_{50}$  values, even against D614G<sub>S</sub> pseudotypes (ranging from 2.0 to  $11.4 \mu g \text{ ml}^{-1}$ ) (Figs. 2B and 3A and fig. S4A).

To determine where on the RBD A3, C6, and A6 bind, we carried out competition studies with C1A-B12 (14), a class 1 antibody; REGN10987 (17, 18) and S309 (44), two class 3 antibodies; and CR3022 (55), a class 4 antibody (Fig. 3B, fig. S8, and Movie 1). A3 competed with CR3022 and REGN10987 for RBD binding, C6 competed with CR3022, and C6 and A6 competed with each other (Fig. 3B and fig. S8). A6 did not compete with any of the other antibodies tested. Among A3, C6, and A6, only A3 competed with binding of an ACE2-Fc fusion protein, suggesting that A3 blocks cellular attachment.

#### Antibody C1C-A3 binds the conserved RBD core

We determined the 3.1-Å cryo-EM structure of the A3 Fab bound to the SARS-CoV-2 spike

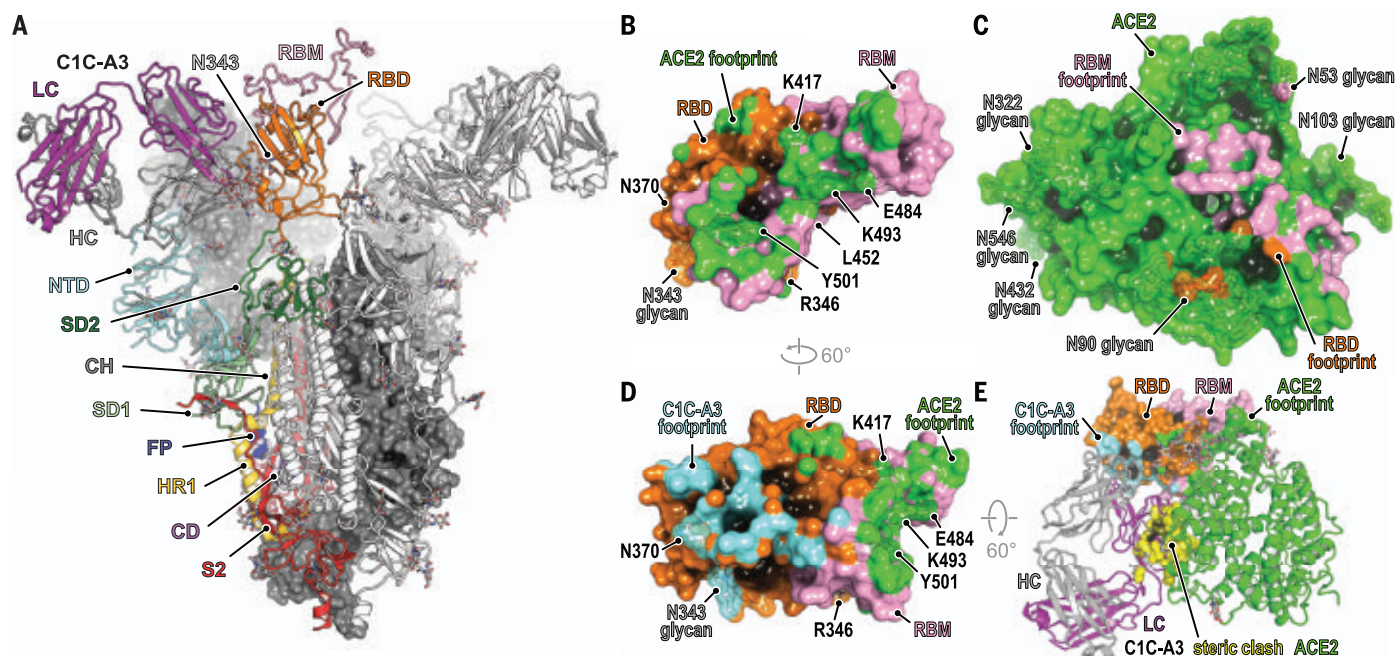
protein ectodomain (Fig. 4A, figs. S9 and S10, and table S6). A3 binds the RBD core with the spike protein trapped in the three open RBD conformation (Fig. 4A). In agreement with competition assays (Fig. 3B), A3 interacts with RBD residues that overlap significantly with those of CR3022 (Fig. 3, C, D, and F, and Movie 1). A3 is therefore a class 4 antibody, a class that includes CR3022, S2A4, S304, S2X35, H014, COVA1-16, S2X259, and DH1047 (4, 56–59) (Movie 1 and fig. S11). Although the A3 and S309 footprints on the RBD do not overlap, and S309 (a class 3 antibody) can bind the closed spike protein trimer (44), both antibodies contact the N-linked glycan attached to N343<sub>RBD</sub> but approach it from different faces (Fig. 3C and Movie 1).

The A3 Fab avoids the RBD–ACE2 interface, which contains the majority of key antibody escape mutations, but, like other class 4 antibodies, nonetheless binds the RBD in a manner that would sterically interfere with ACE2 binding (Fig. 4, B to E, and fig. S11). On the basis of its epitope, in addition to retaining activity against all variants we tested, A3 would also have activity against emergent and preemergent SARS-CoV-2 variants; these include a variant sequenced from travelers from Tanzania that contains the E484K<sub>RBD</sub>, T478R<sub>RBD</sub>, and R346K<sub>RBD</sub> mutations, and

B.1.621 (Mu), a variant detected early in 2021 in Colombia that has since spread internationally and contains the E484K<sub>RBD</sub>, N501Y<sub>RBD</sub>, and R346K<sub>RBD</sub> mutations (Figs. 1G and 4D and table S1). The R346K<sub>RBD</sub> mutation falls within the RBD core and is in the S309 binding site but is not within A3's footprint (Fig. 3, D and E, and Movie 1). However, S309 would likely retain activity against SARS-CoV-2 variants that contain the R346K<sub>RBD</sub> mutation, as the residue that is at the position analogous to SARS-CoV R346<sub>RBD</sub> is a lysine in SARS-CoV, and S309 neutralizes both SARS-CoV and SARS-CoV-2 (44, 60).

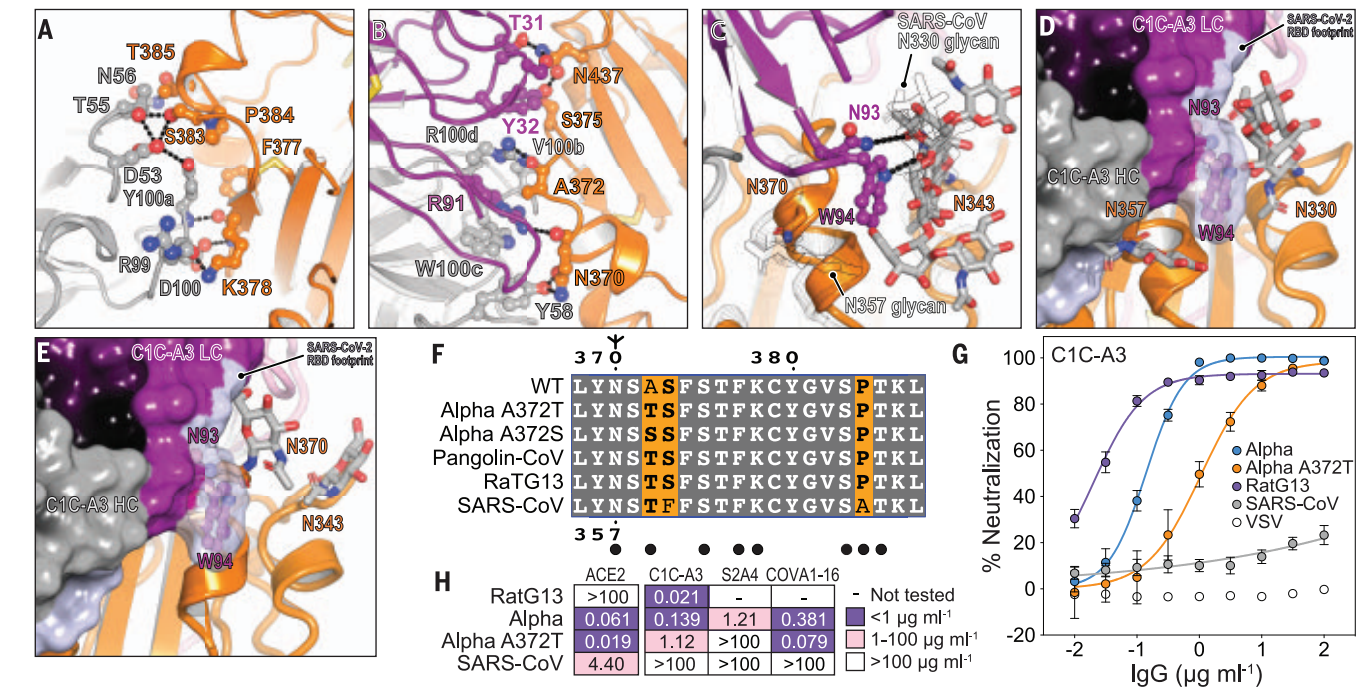
#### RBD core glycan addition drives neutralization escape

Despite A3's breadth against SARS-CoV-2 variant pseudotypes (Figs. 2B and 3A), A3 does not neutralize SARS-CoV pseudotype (fig. S6, F and G). The A3 epitope is highly conserved between SARS-CoV-2 and SARS-CoV; however, N370<sub>RBD</sub> is a site of N-linked glycosylation in SARS-CoV (N357<sub>RBD</sub> in SARS-CoV numbering) and in animal coronaviruses but not in SARS-CoV-2 (Fig. 5, A to C and F) (61). An N-linked glycan attached to SARS-CoV-2 N370<sub>RBD</sub> would introduce steric clashes with the A3 antibody heavy and light chains (Fig. 5D). Furthermore, calculations of antibody-accessible



**Fig. 4. Structural basis for C1C-A3 neutralization.** (A) Cryo-EM structure of the C1C-A3–Fab SARS-CoV-2 spike protein complex. Two of the three spike protein protomers are shown in surface representation. One protomer is shown as a ribbon diagram with labeled subdomains. The trimer model shown was generated by superposition of an RBD–C1C-A3 Fab model generated by subparticle classification of the RBD region onto the coordinates of the trimeric spike protein–C1C-A3 Fab complex (see materials and methods). SD1, subdomain 1; SD2, subdomain 2; FP, fusion peptide; HR1, heptad repeat 1; CD, connector domain; S2, additional portions

of S2 subunit. (B) Surface representation of the SARS-CoV-2 day 146\* RBD showing the ACE2 footprint, including surfaces contacted by ACE2 N-linked glycans. Key RBD positions discussed in the text are labeled. (C) Surface representation of ACE2, showing the day 146\* RBD and RBM footprints. (D) Surface representation of the RBD highlighting C1C-A3 Fab and ACE2 footprints. (E) Overlay of the C1C-A3 Fab–RBD complex with the day 146\* RBD–ACE2 complex. Atoms within 1.54 Å of each other are shown in yellow surface representation to highlight steric clashes. Key RBD residues discussed in the text are labeled in (B) and (D).



**Fig. 5. Structural basis for immune evasion of a RBD core-targeting antibody.** (A and B) C1C-A3 antibody contacts with the SARS-CoV-2 RBD core. (C) C1C-A3 contacts with the N343<sub>RBD</sub> glycan with structural superposition of the SARS-CoV RBD (PDB ID 6NB6) (78). N-linked glycans found at N330<sub>RBD</sub> and N357<sub>RBD</sub> in SARS-CoV and the analogous N343<sub>RBD</sub> and N370<sub>RBD</sub> positions in SARS-CoV-2 are highlighted. (D) Superposition of the C1C-A3 Fab-SARS-CoV-2 RBD structure with the SARS-CoV RBD (PDB ID 6NB6) (78) showing that a glycan attached at SARS-CoV N357<sub>RBD</sub> may interfere with antibody binding. The SARS-CoV-2 RBD is not shown for clarity. (E) Superposition of the C1C-A3-SARS-CoV-2 RBD with the RaTG13 virus RBD (PDB ID 7CN4) (79) showing that

a glycan attached at RaTG13 virus N370<sub>RBD</sub> would be more readily accommodated because the helix that contains it would be rotated away from the Fab. The SARS-CoV-2 RBD is omitted for clarity. (F) Sequence alignment of the RBD core region contacted by C1C-A3. SARS-CoV-2 numbering is shown at the top of the alignment, and SARS-CoV numbering is shown at the bottom. Circles indicate antibody contacts. (G) C1C-A3 neutralization curves for the indicated lentivirus pseudotypes. Data are plotted as the mean ± standard deviation of the mean. The experiment was performed twice in triplicate (n = 6). For some data points, error bars are smaller than symbols. (H) Tabulated neutralization IC<sub>50</sub> values for the indicated pseudotypes.

surface areas using molecular dynamic trajectories of a fully glycosylated SARS-CoV-2 spike protein with a modeled N370<sub>RBD</sub> glycan suggest that its addition would restrict A3 epitope access and could also affect binding of other class 4 antibodies (fig. S12) (61, 62).

Partial occupancy of the glycan attached to SARS-CoV N357<sub>RBD</sub> in recombinant protein preparations may explain why we observed spike protein and RBD binding but a lack of SARS-CoV pseudotype neutralization (fig. S6, F and G). In surface plasmon resonance binding assays, A3 IgG bound tightly to the SARS-CoV RBD only when the RBD was enzymatically deglycosylated (fig. S13 and table S7). Consistent with the role of the SARS-CoV N357<sub>RBD</sub> N-linked glycan as a barrier to A3 neutralization, introducing a substitution that would prevent its addition (T359A<sub>RBD</sub>) sensitized SARS-CoV pseudotypes to A3 neutralization (IC<sub>50</sub> value of 5 µg ml<sup>-1</sup>) (Fig. 2B and fig. S4A).

The A372S/T<sub>RBD</sub> mutations, which would introduce an N-linked glycosylation motif and allow for modification of N370<sub>RBD</sub> in the SARS-CoV-2 spike protein, are found in human-derived SARS-CoV-2 sequences (GISAID) (27),

including on sequences for VOCs Alpha and Delta, without apparent geographic restriction (48 sequence counts as of 10 October 2021 and detected in at least 14 countries) (Fig. 5F and table S8). Although the mutations are currently rare, their presence in sequence databases suggests that SARS-CoV-2 strains containing these mutations can replicate in humans. To confirm that an N-linked glycan could be added to N370<sub>RBD</sub>, we conducted glycan analysis on recombinant SARS-CoV-2 RBD containing the A372S<sub>RBD</sub> substitution and observed 90% occupancy of an N-linked glycan at position N370<sub>RBD</sub> (fig. S13B).

Because acquisition of a putative N-linked glycan at N370<sub>RBD</sub> was the most frequent on the Alpha variant at the time of our initial analysis, we generated an Alpha pseudotype that contains the A372T<sub>RBD</sub> substitution (Alpha A372T). We tested the effect of this substitution on three class 4 antibodies: A3, the antibody we isolated here; S2A4, an antibody that does not cross-react with the SARS-CoV RBD (4); and COVA1-16, an antibody that has weak cross-neutralizing activity against SARS-CoV (57). The mutation resulted in eightfold greater

resistance to A3 neutralization (IC<sub>50</sub> value of 1.1 µg ml<sup>-1</sup>, as compared with 0.14 µg ml<sup>-1</sup> with the Alpha pseudotype) and complete resistance to S2A4 neutralization (Figs. 2B, 3A, and 5, G and H, and fig. S4A). S2A4 and COVA1-16 neutralized variants with potency that was overall comparable to A3 in most cases (Figs. 2B and 3A and fig. S4A). COVA1-16, probably because it has some activity against SARS-CoV [above the limit of detection in our assays, but 29 µg ml<sup>-1</sup> as reported by Liu *et al.* (57)], retained activity against Alpha A372T pseudotype (Figs. 2B, 3A, and 5H, and fig. S4A). The Fab binding pose of certain class 4 antibodies, therefore, may allow them to avoid steric hindrance from an N-linked glycan attached to N370<sub>RBD</sub> (S2X259 is one such antibody) (Movie 1) (56).

**Antibody C1C-A3 neutralizes a related coronavirus**

Coronaviruses that circulate in animals and have spike protein RBDs that can bind human ACE2 are a continued threat. RaTG13 virus, which is closely related to SARS-CoV-2 phylogenetically, is one example (63). The RaTG13



virus spike protein contains a threonine at RBD position 372, which would allow for N370<sub>RBD</sub> glycosylation (Fig. 5F). Despite the presence of the N-linked glycan, A3 potently neutralized RaTG13 virus pseudotype (neutralization IC<sub>50</sub> value of 21 ng ml<sup>-1</sup>), suggesting that A3 neutralization breadth extends to preemergent coronaviruses that are closely related to SARS-CoV-2 (Fig. 5, G and H). Structural superposition reveals that the N370<sub>RBD</sub> glycan on the RaTG13 RBD is positioned in a manner that may not block A3 epitope access but could interfere with binding of other antibodies that bind nearby epitopes on the RBD core (Fig. 5E).

## Discussion

As variants containing composite mutations begin to emerge, continued SARS-CoV-2 immune evasion will remain a big concern. We found that accumulation of large numbers of RBD mutations, which mimics accelerated spike protein evolution occurring in a persistently infected immunocompromised host (14–16), is facilitated by structural plasticity at the ACE2–RBD interface (Fig. 1, B to F). The severity of the phenotypes we observed in vitro suggests that further evolved variants will more adeptly escape therapeutic antibody neutralization than currently circulating VOCs, with potential resistance to two-component antibody cocktails (Fig. 2, A and B).

After two mRNA vaccine immunizations and as early as 7 days after the second dose, all mRNA vaccine recipients had detectable neutralizing activity against pseudotypes containing an NTD supersite deletion and RBDs with six to seven mutations (e.g., day 146\*, day 152\*, and RBM-2), with mean neutralization ID<sub>50</sub> values decreased by 2.3- to 6.1-fold (Fig. 2, C and D, and fig. S5). While the precise epitopes targeted by this residual vaccine-elicited serum neutralizing activity remain to be determined, we surmise that antibodies targeting the RBD core (e.g., those that bind away from the RBM) at least in part account for some of this activity. As the RBD is a major target of vaccine-elicited and naturally acquired humoral immunity to SARS-CoV-2 (4, 5), and the RBM is a critical site of potent neutralizing antibody binding (19, 21–25, 64) that is the most antibody-accessible and the least masked by glycan and conformational shielding (fig. S12), continued RBM evolution may guide antibody responses toward more conserved neutralizing epitopes on the RBD core.

We mined genome sequences in the GISAID database for substitutions that would introduce additional N-linked glycans onto the RBD. This analysis identified D364N<sub>RBD</sub> as an additional mutation that would introduce a putative N-linked glycosylation site in a surface-exposed loop in the footprint of some class 4 antibodies (Movie 1). The independent acquisition

of N-linked glycosylation sites (through the A372S/T<sub>RBD</sub> and D364N<sub>RBD</sub> substitutions) on the same surface of the RBD core, but not on other RBD sites, suggests that this region may be a target of immune selective pressure.

While glycan addition may allow neutralization escape, this change could come at a cost to viral fitness and infectivity. Indeed, the A1114G:T372A mutation that removed the glycan in the SARS-CoV-2 RBD appeared under selective pressure, and addition of the glycan decreases viral replication in human lung epithelial (Calu-3) cells by more than 60-fold (65). A recent molecular dynamics study suggests that introducing the glycan at N370<sub>RBD</sub> in SARS-CoV-2 would favor the closed conformation with the N370<sub>RBD</sub> glycan stabilizing the closed RBD structure on the trimeric spike protein (66). A lack of a glycan at position N370<sub>RBD</sub>, therefore, may increase SARS-CoV-2 ACE2-binding and infectivity by favoring the open state but may also make SARS-CoV-2 more vulnerable to neutralizing antibodies that can only bind the RBD in the open conformation.

Although addition of the N370<sub>RBD</sub> glycan may be associated with a cost to viral fitness, should the selective immune pressure be considerable at this site over a long enough time scale, this may also afford the virus an opportunity to acquire permissive secondary mutations during evolution that restore viral fitness, as is observed in influenza virus drug resistance (67). Such compensatory mutations would be ones that promote ACE2 binding and RBD opening; for example, the D614G<sub>S</sub> mutation (68), which favors the open conformation, and the N501Y<sub>RBD</sub> mutation, which introduces more favorable interactions with ACE2 (Fig. 1D).

As parts of the world continue to face waves of infection and mitigation strategies are relaxed, viral replication in human hosts under antibody selective pressure will continue to shape the antigenic landscape of the SARS-CoV-2 spike protein. With vigorous variant monitoring efforts underway to help design next-generation antibody-based therapeutics, and with mRNA- or DNA-based vaccines that can be updated to rapidly adapt to new variants, proactively examining the consequences of further viral evolution before the next highly antibody resistant strain emerges is of utmost importance.

## Materials and methods summary

We isolated monoclonal antibodies from the blood of a COVID-19 convalescent individual using single B cell sorting with prefusion-stabilized SARS-CoV spike protein ectodomain as bait and using established protocols (14, 54). We obtained venous blood samples from healthy mRNA-1273 and BNT162b2 vaccine recipients. We produced recombinant glycoproteins and antibodies or Fabs in tran-

siently transfected mammalian cells grown in suspension culture and purified these proteins using affinity-based methods. We used ELISAs to measure antibody binding and BLI or surface plasmon resonance to determine kinetic parameters of binding. We packaged lentivirus pseudotypes by transient transfection of HEK293T cells, as previously described (14). We used HEK293T cells expressing human ACE2 in pseudotype neutralization assays or Vero E6 cells in plaque reduction neutralization tests, as previously described (14). We collected x-ray diffraction data on crystals of a day 146\*-SARS-CoV-2 RBD complex at the Advanced Photon Source (APS, Argonne, IL) NE-CAT beamline and used established procedures for data processing, molecular replacement, atomic model building, and refinement (69–73). We used mass spectrometry to perform glycopeptide analysis. After data collection on a Titan Krios cryo-electron microscope equipped with a Gatan K3 camera, we used single-particle cryo-EM to determine the structure of a prefusion-stabilized SARS-CoV-2 spike protein ectodomain (7) complexed with C1C-A3 Fab complex using established procedures for image processing, atomic model building, and refinement (72–77).

## REFERENCES AND NOTES

- W. T. Harvey et al., SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021). doi: 10.1038/s41579-021-00573-0; pmid: 34075212
- J. Lan et al., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020). doi: 10.1038/s41586-020-2180-5; pmid: 32225176
- J. Shang et al., Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221–224 (2020). doi: 10.1038/s41586-020-2179-y; pmid: 32225175
- L. Piccoli et al., Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042.e21 (2020). doi: 10.1016/j.cell.2020.09.037; pmid: 32991844
- A. J. Greaney et al., Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. *Sci. Transl. Med.* **13**, eabi9915 (2021). doi: 10.1126/scitranslmed.abi9915; pmid: 34103407
- K. R. McCarthy et al., Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science* **371**, 1139–1142 (2021). doi: 10.1126/science.abf6950; pmid: 33536258
- M. McCallum et al., N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332–2347.e16 (2021). doi: 10.1016/j.cell.2021.03.028; pmid: 33761326
- P. Supasa et al., Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell* **184**, 2201–2211.e7 (2021). doi: 10.1016/j.cell.2021.02.033; pmid: 33743891
- W. Dejnirattai et al., Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **184**, 2939–2954.e9 (2021). doi: 10.1016/j.cell.2021.03.055; pmid: 33852911
- D. Planas et al., Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **27**, 917–924 (2021). doi: 10.1038/s41591-021-01318-5; pmid: 33772244
- P. Wang et al., Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747–751.e4 (2021). doi: 10.1016/j.chom.2021.04.007; pmid: 33887205
- P. Wang et al., Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021). doi: 10.1038/s41586-021-03398-2; pmid: 33684923



- ## ACKNOWLEDGMENTS

This work is based on research conducted at the Northeastern Collaborative Access Team (NE-CAT) beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Pilatus 6M detector on 24-ID-C beamline is funded by a NIH-ORIP HEI grant

(S10 RR029205). 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. We thank the staff at NE-CAT for assistance with x-ray data collection. Cryo-EM data were collected at the Harvard Cryo-Electron Microscopy Center for Structural Biology. We thank Harvard Cryo-EM Center staff members for assistance with cryo-EM data collection. We acknowledge support for COVID-19-related structural biology research at Harvard from the Nancy Lurie Marks Family Foundation. We thank B. Chen for providing a human ACE2 ectodomain-expressing cell line, and H. Choe for providing cDNA encoding the SARS-CoV spike protein.

**Funding:** This work was funded in part by the Massachusetts Consortium on Pathogen Readiness (MassCPR) (J.A.). J.A. was also supported by the Star-Friedman Award Challenge for Promising Scientific Research. L.R.B. is supported by the Harvard Clinical and Translational Science Center, from the National Center for Advancing Translational Science (1UL1TR002541-01); by Barbara and Amos Hostetter; and by the Chleck Family Foundation. S.E.T., R.C.L., E.T.R., and R.C.C. receive grant support from the Centers for Disease Control and Prevention (U01CK000490). K.G.N., A.C., and H.H.V. are supported by award T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. **Author contributions:** Conceptualization: K.G.N., S.A.C., S.S., J.P., L.E.C., P.Y., A.Co., L.G.A.M., N.V.T., M.D.,

S.E.T., S.K., A.C.S., A.D., C.T., J.F.C.-S., A.T.D., D.T., A.Cl., Y.F., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. Methodology: K.G.N., S.A.C., S.S., J.P., L.E.C., P.Y., A.Co., L.G.A.M., H.H.V., N.V.T., G.Z., M.D., S.E.T., S.K., A.C.S., A.D., R.C.L., E.T.R., C.T., J.F.C.-S., A.T.D., D.T., A.Cl., Y.F., A.G., I.R.C., J.S., L.R.B., R.C., and J.A. Investigation: K.G.N., S.A.C., S.S., J.P., L.E.C., P.Y., A.Co., L.G.A.M., H.H.V., V.B., N.V.T., S.E.T., S.K., M.J., A.C.S., C.T., J.F.C.-S., A.T.D., D.T., A.C., Y.F., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. Visualization: K.G.N., S.A.C., S.S., J.P., L.E.C., P.Y., A.Co., L.G.A.M., H.H.V., N.V.T., G.Z., M.D., S.E.T., S.K., A.C.S., C.T., J.F.C.-S., A.T.D., D.T., A.Cl., Y.F., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. Funding acquisition: R.C.L., E.T.R., R.C.C., and J.A. Project administration: K.G.N., S.A.C., S.S., J.P., N.V.T., M.D., S.E.T., S.K., A.C.S., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. Supervision: A.D., J.F.C.-S., Y.F., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. Writing—original draft: J.A. Writing—review & editing: K.G.N., S.A.C., S.S., J.P., L.E.C., A.C., L.G.A.M., H.H.V., G.Z., M.D., S.E.T., S.K., A.C.S., R.C.L., E.T.R., C.T., J.F.C., A.T.D., D.T., A.C., Y.F., A.G., I.R.C., J.S., L.R.B., R.C.C., and J.A. **Competing interests:** S.A.C., L.E.C., and J.A. are inventors on a provisional patent application that covers antibodies described in this publication filed by the president and fellows of Harvard University. S.E.T. receives monetary compensation from UpToDate, which provides clinical decision support. **Data and materials availability:** C1C-A3 antibody heavy chain and light chain variable region gene sequences have been deposited in GenBank under accession codes OL621229 and OL621230, respectively. Protein Data Bank (PDB) IDs for the day 146<sup>\*</sup> RBD human ACE2 ectodomain complex, the unliganded C1C-A3 Fab, the C1C-A3 Fab SARS-CoV-2 spike protein ectodomain complex, and focused

refinement of the C1C-A3 Fab/RBD region are 7SN0, 7SN1, 7SN3, and 7SN2, respectively. Electron Microscopy Data Bank (EMDB) accession numbers for the C1C-A3 Fab SARS-CoV-2 spike protein ectodomain complex and focused refinement of the C1C-A3 Fab/RBD region are EMD-25210 and EMD-2509, respectively. All materials are available upon request. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abl6251](https://science.org/doi/10.1126/science.abl6251)

Materials and Methods

Figs. S1 to S13

Tables S1 to S9

References (80–100)

MDAR Reproducibility Checklist

28 July 2021; accepted 29 November 2021

Published online 2 December 2021

10.1126/science.abl6251

## RESEARCH ARTICLE SUMMARY

## PROTEIN DESIGN

## Reconfigurable asymmetric protein assemblies through implicit negative design

Danny D. Sahtoe<sup>†</sup>, Florian Praetorius<sup>†</sup>, Alexis Courbet, Yang Hsia, Basile I. M. Wicky, Natasha I. Edman, Lauren M. Miller, Bart J. R. Timmermans, Justin Decarreau, Hana M. Morris, Alex Kang, Asim K. Bera, David Baker\*

**INTRODUCTION:** For many current challenges in synthetic biology, it would be desirable to have bio-orthogonal and modular sets of interacting proteins that are folded and soluble when alone but rapidly and specifically associate when mixed. Although pairs of proteins with these properties are found in nature, with the exception of single-helix coiled-coil peptides that are not folded in isolation, it has been very challenging to generate new bio-orthogonal pairs by design. This is because designed, largely nonpolar interfaces that drive association between two different chains can also mediate self-association of individual chains into large oligomers or aggregates that disassociate very slowly. For example, designed heterodimeric

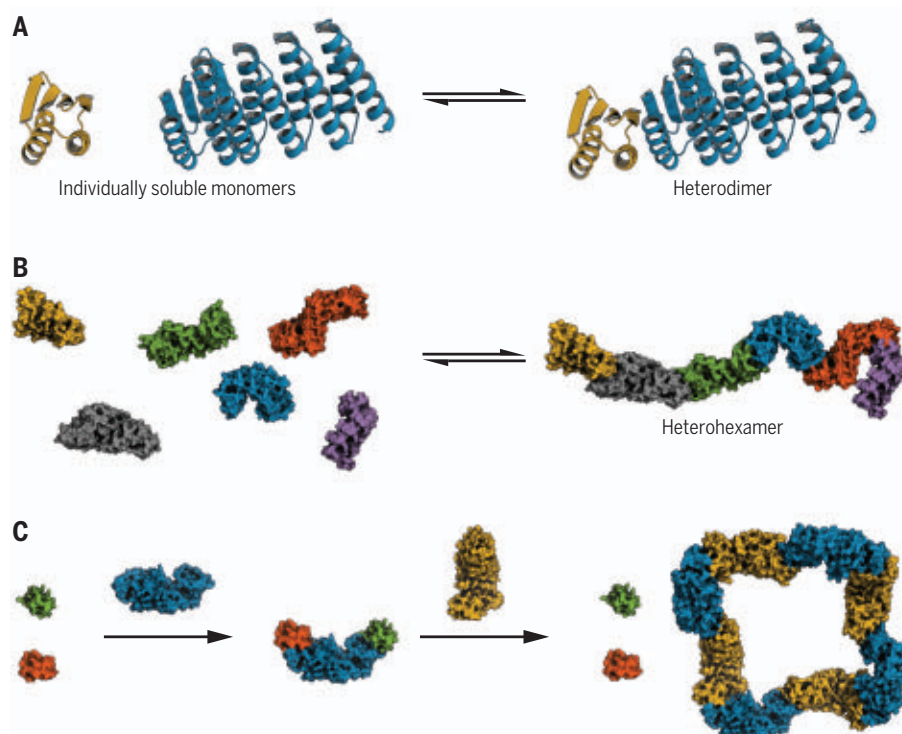
helical bundles with specific and orthogonal interfaces do not readily assemble from individually purified monomers but require coexpression, thermal or chemical denaturation, or incubation for more than a week for assembly from individually prepared components.

**RATIONALE:** We sought to design heterodimers that (i) spontaneously assemble upon mixing of stable and soluble individual components, (ii) allow dynamic exchange of components, and (iii) are amenable to rigid fusion to enable the assembly of higher-order hetero-oligomeric complexes with defined structures. We reasoned that these properties could be achieved with rigid designed protomers with exposed  $\beta$  sheet

edges that associate to form a continuous  $\beta$  sheet in the heterodimer (see the figure): Any off-target interaction that does not allow for strand-pair formation should be highly unfavorable because of the high thermodynamic cost of the burial of backbone polar groups away from water in the absence of compensating hydrogen-bonding interactions, thus reducing the probability of undesired homo-oligomerization.

**RESULTS:** We designed 12  $\alpha$ - $\beta$  heterodimers that readily assemble from individually expressed and purified monomers. By rigidly fusing two different protomers to designed helical repeat proteins, we generated bivalent connector proteins that can bind two different partners. Using one or more of these bivalent connectors, we successfully assembled linear heterotrimers, heterotetramers, heteropentamers, and a heterohexamer with distinct shapes (see the figure). We further assembled branched heterotetramers using trivalent connectors that can bind three different partners in defined orientations. By rigidly fusing our protomers to previously designed homooligomers, we created hubs that can bind three or four copies of the same binding partner. Using symmetry-aware helical fusion, we designed a closed C4-symmetric ring. All of these hetero-oligomeric complexes readily assemble from individually prepared components. The components function as designed in living cells, mediating the assembly of liquid-liquid condensates or more static aggregates depending on the interaction affinities, and designed assemblies can be reconfigured by addition of components providing access to lower free-energy states (see the figure).

**CONCLUSION:** Our reversible heterodimeric assemblies open the door to many exciting new synthetic biology and nanomaterial applications. Bivalent connectors can be used to induce interaction between otherwise monomeric proteins to modulate biological function, and symmetric hubs can present multiple copies of ligands or antigens to cluster cell-surface receptors. Bio-orthogonal signaling systems can be constructed by using the heterodimer components in synthetic receptors and ligands. Reconfigurable higher-order nanomaterials—one-dimensional (1D) fibers, 2D lattices, and 3D nanocages and crystals—can be created using our designed components to drive geometrically precise association between the material components. ■



**Schematic representation of reconfigurable protein assemblies.** (A and B) Formation of a heterodimer (A) or heterohexamer (B) from monomers that are stable and soluble in isolation. (C) Two components A (orange) and C (green) are monomeric and do not interact. Addition of a bivalent connector B (blue) brings them in close proximity. Subsequent addition of component B' (yellow) leads to the formation of a symmetric B<sub>4</sub>B'<sub>4</sub> ring and release of free A and C.

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

\*Corresponding author. Email: [dabaker@uw.edu](mailto:dabaker@uw.edu)

<sup>†</sup>These authors contributed equally to this work.

Cite this article as D. D. Sahtoe *et al.*, *Science* **375**, eabj7662 (2022). DOI: [10.1126/science.abj7662](https://doi.org/10.1126/science.abj7662)

**S READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abj7662>



## RESEARCH ARTICLE

## PROTEIN DESIGN

## Reconfigurable asymmetric protein assemblies through implicit negative design

Danny D. Sahtoe<sup>1,2,3†</sup>, Florian Praetorius<sup>1,2†</sup>, Alexis Courbet<sup>1,2,3</sup>, Yang Hsia<sup>1,2</sup>, Basile I. M. Wicky<sup>1,2</sup>, Natasha I. Edman<sup>1,2,4,5</sup>, Lauren M. Miller<sup>1,2</sup>, Bart J. R. Timmermans<sup>1,2</sup>, Justin Decarreau<sup>1,2</sup>, Hana M. Morris<sup>1,2</sup>, Alex Kang<sup>1,2</sup>, Asim K. Bera<sup>1,2</sup>, David Baker<sup>1,2,3\*</sup>

Asymmetric multiprotein complexes that undergo subunit exchange play central roles in biology but present a challenge for design because the components must not only contain interfaces that enable reversible association but also be stable and well behaved in isolation. We use implicit negative design to generate  $\beta$  sheet-mediated heterodimers that can be assembled into a wide variety of complexes. The designs are stable, folded, and soluble in isolation and rapidly assemble upon mixing, and crystal structures are close to the computational models. We construct linearly arranged hetero-oligomers with up to six different components, branched hetero-oligomers, closed C4-symmetric two-component rings, and hetero-oligomers assembled on a cyclic homo-oligomeric central hub and demonstrate that such complexes can readily reconfigure through subunit exchange. Our approach provides a general route to designing asymmetric reconfigurable protein systems.

**D**ynamic reconfigurable multiprotein complexes play key roles in central biological processes (1). The subunits are generally monomeric in isolation, allowing the assemblies to reconfigure by successive addition or removal of one or more components. Such modulation is essential to their function; for example, subunit loss and addition underlie the molecular mechanisms of protein complexes that drive DNA replication and transcription (2, 3). The ability to de novo design such multicomponent reconfigurable protein assemblies would enable the realization of sophisticated new functions. Previous design efforts have generated cyclic oligomeric and higher-order symmetric nanostructures such as icosahedral nanocages with as many as 120 subunits and two-dimensional (2D)-layers with many thousands of regularly arrayed components (4–8). Essential to this is the symmetry and cooperativity of assembly, which strongly favors just one of a large number of possible states. Once formed, these assemblies are therefore typically quite static and exchange subunits only on long time scales, which is advantageous for applications such as nanoparticle vaccine design and multivalent receptor engagement (9).

The design of reconfigurable asymmetric assemblies is more challenging, because there

is no symmetry “bonus” favoring the target structure (as is attained, for example, in the closing of an icosahedral cage) and because the individual subunits must be stable and soluble in isolation in order to reversibly associate. Reconfigurable asymmetric protein assemblies could, in principle, be constructed using a modular set of protein-protein interaction pairs (heterodimers), provided that first, the individual subunits are stable and monomeric in isolation so that they can be added and removed; second, the interacting pairs are specific; and third, they can be rigidly fused through structured connectors to other components. Rigid fusion, as opposed to fusion by flexible linkers, is important to program the assembly of structurally well-defined complexes; most higher-order natural protein complexes have, despite their reconfigurability, distinct overall shapes that are critical for their function. Although there are designed orthogonal sets of interacting proteins that have one of these properties, designed proteins that have all of these properties are lacking. The components of designed helical-hairpin heterodimers (10, 11) on their own form homodimers or other higher-order homomeric aggregates that disassemble on very long time scales (10, 12), making them unsuitable for use in constructing reconfigurable higher-order assemblies. Heterodimeric coiled coils assemble from peptides that are soluble and monomeric, but the monomers are unfolded before binding their partners (13, 14), complicating their use in structurally defined rigid fusions.

We set out to design sets of interacting protein pairs for constructing reconfigurable assemblies (Fig. 1A). The first challenge is the

systematic design of proteins with interaction surfaces that drive association with cognate partners but not self-association. Hydrophobic interactions drive protein complex assembly, but these same hydrophobic interactions can also promote homomerization. Previously designed heterodimeric helical bundles featured, in addition to hydrophobic interactions, explicit hydrogen-bond networks that contribute to binding specificity and make the interface more polar. However, the individual protomers, either helical hairpins or individual helices, lack a hydrophobic core and are thus flexible and unstable as monomers, allowing a wide range of potential off-target homo-oligomers to form (Fig. 1B). Explicit negative-design methods favor one state by considering the effect of amino acid substitutions on the free energies of both states (15–17). However, such methods cannot be readily applied to disfavor self-association, because there are, in general, a large number of possible self-associated states that cannot be systematically enumerated.

We instead sought to use implicit negative design (18) by introducing three properties that collectively make self-associated states unlikely to have low free energy: First, in contrast to the flexible coiled coils and helical hairpins used in previous designs, we aimed for well-folded individual protomers stabilized by substantial hydrophobic cores; this property limits the formation of slowly exchanging homo-oligomers (Fig. 1B). Second, we constructed interfaces in which each protomer has a mixed  $\alpha$ - $\beta$  topology and contributes one exposed  $\beta$  strand to the interface, giving rise to a continuous  $\beta$  sheet across the heterodimer interface (19–21) (Fig. 1C). The exposed polar backbone atoms of this “edge strand” limit self-association to arrangements that pair the  $\beta$  edge strands; most other homomeric arrangements are unlikely because they result in the energetically unfavorable burial of the polar backbone atoms on the  $\beta$  edge strand (Fig. 1C). Third, taking advantage of the restrictions in possible undesired states resulting from the two properties noted above, we explicitly modeled the limited number of homomeric states and designed in additional elements that were likely to sterically occlude such states (Fig. 1D).

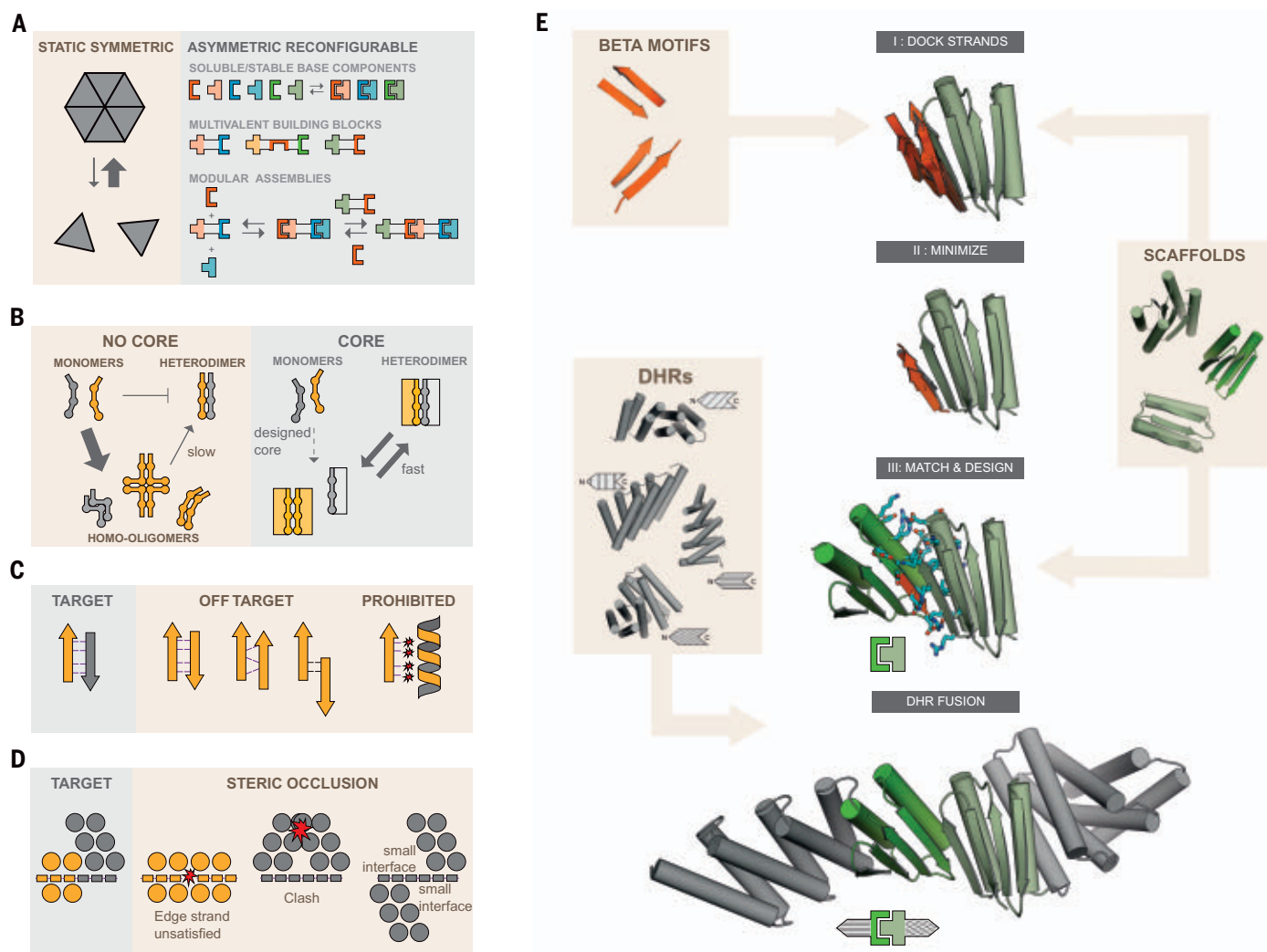
## Results

To implement these properties, we chose to start with a set of mixed  $\alpha$ - $\beta$  scaffolds that were designed by Foldit players (22). The selected designs contain sizable hydrophobic cores, exposed edge strands required for  $\beta$  sheet extension (19), and one terminal helix (either N or C) available for rigid helical fusion (Fig. 1E) (23). Using blueprint-based backbone building (24, 25), we designed additional helices at the other terminus for a subset of the scaffolds to

<sup>1</sup>Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. <sup>2</sup>Institute for Protein Design, University of Washington, Seattle, WA 98195, USA. <sup>3</sup>Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA. <sup>4</sup>Molecular and Cellular Biology Graduate Program, University of Washington, Seattle, WA 98195, USA. <sup>5</sup>Medical Scientist Training Program, University of Washington, Seattle, WA 98195, USA.

\*Corresponding author. Email: dabaker@uw.edu

†These authors contributed equally to this work.



**Fig. 1. Strategies for the design of asymmetric hetero-oligomeric complexes.**

(A) Many design efforts have focused on cooperatively assembling symmetric complexes (left) with little subunit exchange. In this study, we sought to create asymmetric hetero-oligomers from stable heterodimeric building blocks that can modularly exchange subunits (right). Design strategies for preventing subunit self-association are illustrated in the following panels. (B) Protomers that have a substantial hydrophobic core (right rectangles) are less likely to form stable homo-oligomers than protomers of previously designed heterodimers that lack hydrophobic monomer cores. (C) In  $\beta$  sheet extended interfaces, most homodimer states that bury non-hydrogen-bonding polar edge-strand atoms are energetically inaccessible. Potential homodimers are more likely to form by  $\beta$  sheet extension. These are restricted to only two orientations (parallel and antiparallel) and a limited

number of offset registers. Arrows and ribbons represent strands and helices, respectively; thin lines indicate hydrogen bonds; and red stars indicate unsatisfied polar groups. (D) By modeling the limited number of  $\beta$  sheet homodimers across the  $\beta$  edge strand, structural elements may be designed that specifically block homodimer formation or make it unlikely because of small interfaces but still allow heterodimer formation. Circles indicate helices, rectangles indicate  $\beta$  strands, and stars indicate steric clashes. (E) To design reversible heterodimers,  $\beta$  strands are docked to the edge strands of hydrophobic core-containing protein scaffolds [in this paper, from Foldit (22)], a second scaffold is superimposed on the docked strand creating a protein-protein complex, the amino acids at the protein-protein interface are optimized for high affinity and specificity binding, and finally DHRs are fused to the terminal helices.

enable rigid fusion at both the N and C termini (fig. S1). Heterodimers with paired  $\beta$  strands across the interface were generated by superimposing one of the two strands from each of a series of paired  $\beta$  strand templates onto an edge  $\beta$  strand of each scaffold (Fig. 1E, top) and then optimizing the rigid body orientation and the internal geometry of the partner  $\beta$  strand of the template to maximize hydrogen-bonding interactions across the interface

(Fig. 1E, second row). This generates a series of disembodied  $\beta$  strands that form an extended  $\beta$  sheet for each scaffold; for each of these, an edge  $\beta$  strand from a second scaffold was superimposed on the disembodied  $\beta$  strand to form an extended  $\beta$  sheet (Fig. 1E, third row). The interface side chain-side chain interactions in the resulting protein-protein docks were optimized using Rosetta combinatorial sequence design (26). To limit excessive hydrophobic inter-

actions, we generated explicit hydrogen-bond networks across the heterodimer interface (17) or constrained the amino acid composition to favor polar residues while penalizing buried unsatisfied polar groups (27). This resulted in interfaces that, outside of the polar hydrogen bonding of the  $\beta$  strands, contained both hydrophobic interactions and polar networks. To further disfavor unwanted homodimeric interactions (Fig. 1D, right), we rigidly fused

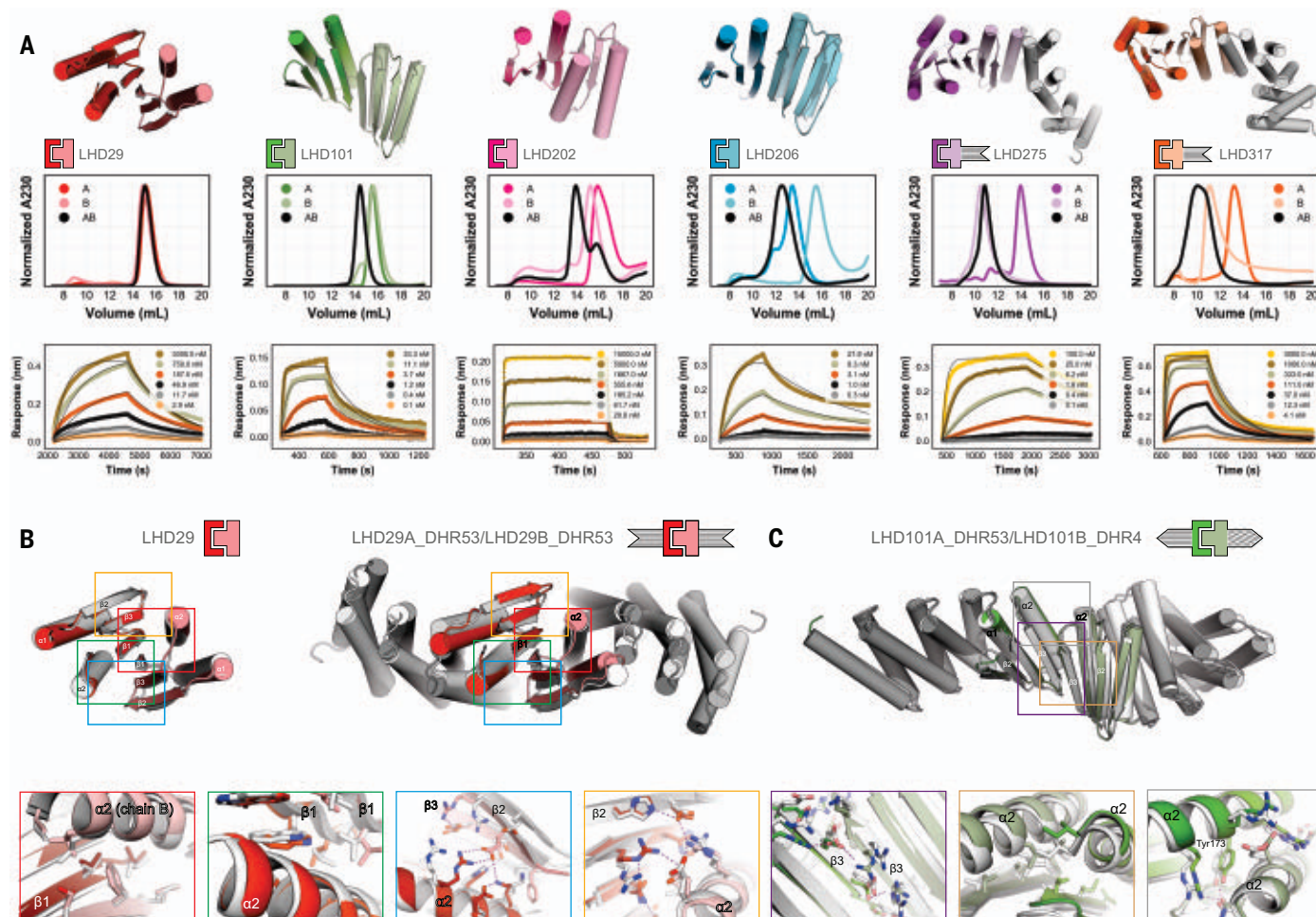
designed helical repeat proteins (DHRs) to terminal helices (23, 28). Because these DHRs have different shapes, they also serve to diversify building-block shapes for subsequent higher-order assembly design. Designed heterodimers were selected for experimental characterization on the basis of binding energy, the number of buried unsatisfied polar groups, buried surface area, and shape complementarity (see Materials and methods).

We coexpressed the selected heterodimers in *Escherichia coli* using a bicistronic expression system that encodes one of the two protomers with a C-terminal polyhistidine tag and the other protomer either with no tag or with a green fluorescent protein (GFP) tag at the N terminus. Complex formation was initially assessed using nickel affinity chromatography; designs for which both protomers were present in SDS-polyacrylamide gel electrophoresis (PAGE) after nickel pulldown were

subjected to size exclusion chromatography (SEC) and liquid chromatography–mass spectrometry (LC-MS). Of the 238 tested designs, 71 passed the bicistronic screen and were selected for individual expression of protomers. Of these, 32 formed heterodimers from individually purified monomers as confirmed by SEC, native MS, or both (Fig. 2A and figs. S2 and S3A). In SEC titration experiments, some protomers were monomeric at all injection concentrations, whereas others self-associated at higher concentrations (fig. S4). Both LHD101 protomers and their fusions were monomeric even at injection concentrations greater than 100  $\mu$ M (fig. S4). LHD275A, LHD278A, LHD317A, and a redesigned version of LHD29 with a more polar interface (LHD274) were also predominantly monomeric (figs. S4 and S5). Designs for which isolated protomers were poorly expressed, polydispersed in SEC, or did not yield stable, soluble, and functional

rigid DHR fusions were discarded along with designs that were very similar to other designs but otherwise stable and soluble. The remaining 11 heterodimers span three main structural classes [Fig. 2A, fig. S2, and data S1 (LHD components)]. In class one, the central extended  $\beta$  sheet is buttressed on opposite sides by helices that contribute additional interface interactions (LHD29 and LHD202 in Fig. 2A); in class two, the helices that provide additional interactions are on the same side of the extended central sheet (LHD101 and LHD206 in Fig. 2A); and in class three, both sides of the central  $\beta$  sheet extension are flanked by helices (LHD275 and LHD317 in Fig. 2A).

We monitored the kinetics of heterodimer formation and dissociation through biolayer interferometry (BLI) (Fig. 2A, fig. S2, and table S1) by immobilizing individual biotinylated protomers onto streptavidin-coated sensors and adding the designed binding partner.



**Fig. 2. Designed heterodimer characterization.** (A) Characterization of six designed heterodimers. Design models are shown in the top row; the color scheme for the different designs is maintained throughout the paper. Normalized SEC traces of individual protomers (A and B) and complexes (AB) are shown in the middle row and kinetic binding traces with global kinetic fits of in vitro biolayer interferometry binding assays in the bottom row. (B and C) Crystal structures (in colors) of the designs LHD29, LHD29A53/B53, and LHD101A53/B4 overlaid on design models (light gray). Colored rectangles in the full models (top row) match the corresponding detailed views (bottom row). Sequences and models for all proteins are provided in data S1.



Unlike previously designed heterodimers, binding reactions equilibrated rapidly, with affinities ranging from micromolar to low nanomolar (fig. S3C and table S1). Association rates were quite fast and ranged from  $10^6 \text{ M}^{-1} \text{ s}^{-1}$  for the fastest heterodimer to  $10^2 \text{ M}^{-1} \text{ s}^{-1}$  for the slowest heterodimer LHD29, which is still an order of magnitude faster than the fastest associating designed helical hairpin heterodimer DHD37 (10) (Fig. 2A, fig. S6A, and tables S1 and S2). For LHD101 and LHD206, we independently determined the equilibrium dissociation constant ( $K_d$ ) with a split luciferase-based binding assay in *E. coli* lysates and obtained very similar values, indicating that heterodimer association is not affected by high concentrations of noncognate proteins (fig. S6, D and E, and table S3).

We determined the crystal structures of two class-one designs, LHD29 (2.2 Å) and LHD29A53/B53 (2.6 Å) in which both protomers are fused to DHR53 (Fig. 2B and table S4). In the central extended  $\beta$  sheet, the LHD29 design closely matches the crystal structure (red and green boxes in Fig. 2B and table S5). Aside from backbone  $\beta$  sheet hydrogen bonds, this part of the interface is supported by primarily hydrophobic packing interactions between the side chains of each interface  $\beta$  edge strand. The two flanking helices on opposite sides of the central  $\beta$  sheet (blue and orange boxes in Fig. 2B) contribute predominantly polar contacts to the interface and are also similar in the crystal structure and design model. Apart from crystal contact-induced subtle backbone rearrangements in strand two of LHD29B that promote the formation of a polar interaction network (blue box in Fig. 2B), most interface side chain-side chain interactions agree with the design model. As for unfused LHD29, the interface of LHD29A53/B53 resembles the designed model; at the fusion junction and repeat protein regions, deviations are slightly larger (table S5).

We also determined the structure of a class-two design, LHD101A53/B4 (2.2 Å), in which protomer A is fused to DHR53 and B to DHR4 (Fig. 2C and tables S4 and S5). The crystal structure agrees well with the design model at both the interface and fusion junctions, as well as the repeat protein regions. In class-two designs, the interface  $\beta$  strand pair is reinforced by flanking helices that, unlike in class-one designs, are in direct contact with both each other and the interface  $\beta$  sheet. The solvent-exposed side of the  $\beta$  interface consists primarily of electrostatic interactions (purple box in Fig. 2C), whereas the buried side consists exclusively of hydrophobic side chains. Together with apolar side chains on the flanking helices of both protomers, these residues form a closely packed core interface (brown box in Fig. 2C) that is further stabilized by solvent-exposed polar interactions between

the flanking helices. Notably, the designed semiburied polar interaction network centered on Tyr<sup>173</sup> is recapitulated in the crystal structure (gray box in Fig. 2C).

As described above, the third of our implicit negative-design principles was to incorporate structural elements incompatible with  $\beta$  sheet extension in homodimeric species (Fig. 1D). To assess the utility of this principle, we took advantage of the limited number of possible off-target edge-strand interactions that can form (Fig. 1C); we docked all protomers against themselves on the edge strand that participates in the heterodimer interface and calculated the Rosetta binding energy after relaxation of the resulting homodimeric dock (fig. S7). Homodimer docks of the protomers that chromatographed as monomers in SEC had unfavorable energies compared with those that showed evidence of self-association in agreement with our initial hypothesis (Fig. 1D), and visual inspection of these docks suggested that homodimerization was likely prevented by the presence of sterically blocking secondary-structure elements (fig. S7).

Twenty-eight additional rigid fusion proteins that were generated using the 11 base heterodimers and LHD274 (Fig. 3A) retained both the oligomeric state and binding activity of the unfused counterparts, indicating that the designed heterodimers are quite robust to fusion (figs. S3D, S6E, and S8). There are 74 different possible heterodimeric complexes that can be assembled from these fusions, each with different shapes. Most of the fusions involve protomers of LHD274 and LHD101; fusions to LHD101 protomers alone enable the formation of 30 distinct heterodimeric complexes (fig. S9).

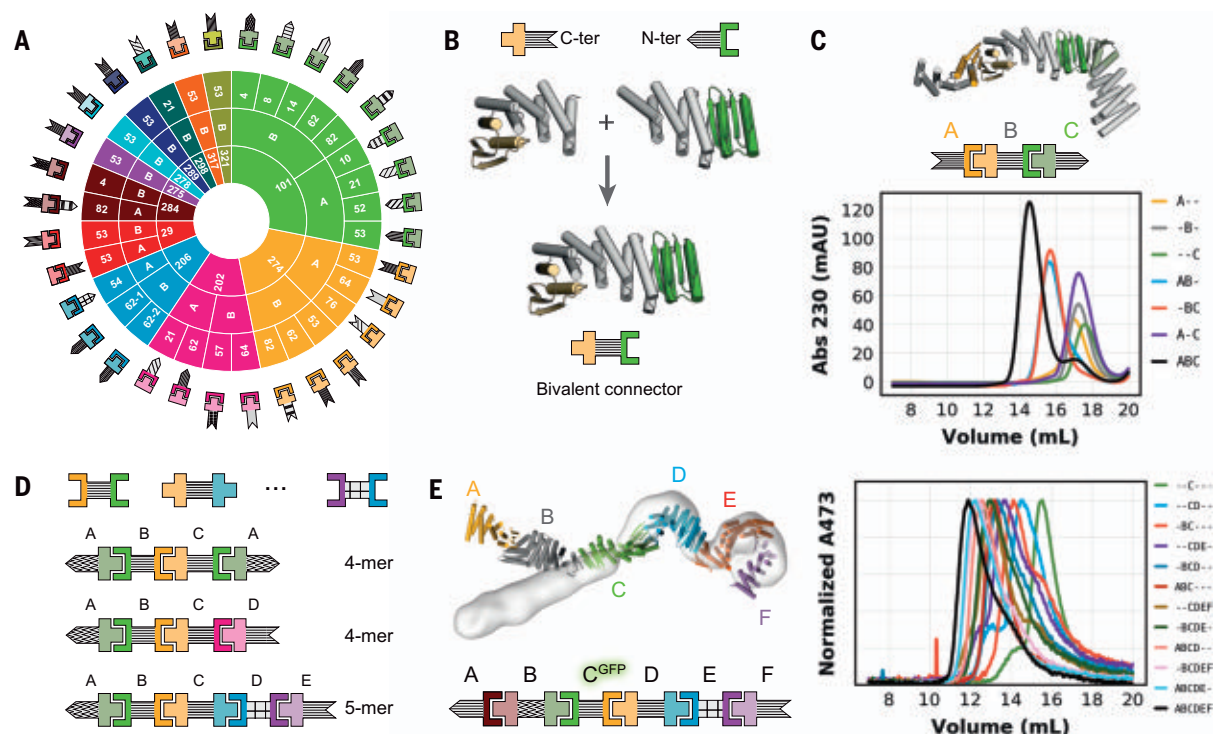
Larger multicomponent hetero-oligomeric protein assemblies require subunits that can interact with more than one binding partner at the same time. To this end, we generated single-chain bivalent connector proteins. Designed protomers that share the same DHR as the fusion partner and have compatible termini can be simply spliced together into a single protein chain on overlapping DHR repeats (Fig. 3B). Mixing a linear connector ("B") with its two cognate binding partners ("A" and "C") yields a linearly arranged heterotrimer ("ABC") in which the two terminal capping components A and C are connected through component B but otherwise are not in direct contact with each other (Fig. 3C). We analyzed the assembly of this heterotrimer and controls by SEC (Fig. 3C) and observed stepwise assembly of the ABC heterotrimer with clear baseline separation from AB and BC heterodimers, as well as from monomeric components (Fig. 3C). Using experimentally validated linear connectors created using the above-described modular splicing approach (Fig. 3D, fig. S10A, and data S1), we assembled 20 heterotrimers in total,

including one verified by negative-stain electron microscopy (nsEM) (figs. S10B and S11). The absence of off-target complexes in these assemblies corroborates the orthogonality of the heterodimer interfaces (fig. S12).

By using more than one connector subunit, larger linear hetero-oligomers can be generated. We constructed and confirmed assembly of ABCA and ABCD heterotetramers, each containing two different linear connectors (B and C) and either one or two terminal caps (two A or A and D), an ABBA heterotetramer using a homodimeric central connector (two B) and one terminal cap (two A), and a nsEM-verified heteropentamer (ABCDE) containing three different linear connectors and two caps (Fig. 3D and figs. S13 and S14). We followed the assembly of an ABCDEF heterohexamer in SEC by GFP-tagging one of the components and monitoring GFP absorbance. The full assembly, as well as subassemblies generated as controls, eluted as monodisperse peaks, with elution volumes agreeing well with expected assembly sizes (Fig. 3E). nsEM reconstruction of the hexamer confirmed that all components were present (Fig. 3E and fig. S15A). Deviation of the experimentally observed shape from the design model likely arises from small deviations from the model in one of the components that cause a lever-arm effect (Fig. 2B).

In total, by combining the bivalent connectors with each other and with monovalent terminal caps, we constructed 36 hetero-oligomers with up to six different chains and confirmed their assembly by SEC and EM [Fig. 3, C and E; figs. S10, S11, S13, and S15; and data S1 (experimentally validated assemblies)]. This number can be readily increased to 489 by including all available components [Fig. 3A, fig. S10A, and data S1 (all\_theoretical\_assemblies)]. Because all fusions have structured helical linkers, the overall molecular shapes of the complexes and the spatial arrangement of individual components are well defined, which should be useful for scaffolding and other applications. Our linear assemblies resemble elongated modular multiprotein complexes found in nature (fig. S15B), like the Cullin RING E3 ligases (29) that mediate ubiquitin transfer by geometrically orienting the target protein and catalytic domain.

We next sought to go beyond linear assemblies and build branched and closed assemblies. Trivalent connectors can be generated from heterodimers in which one protomer has both N- and C-terminal helices (LHD275A, LHD278A, LHD289A, and LHD317A). Such protomers can be fused to two helical repeat proteins and spliced together with different halves of other heterodimer protomers via a common DHR repeat (Figs. 3, A and B, and 4A). The resulting branched trivalent connectors ("A") are capable of binding the three cognate binding partners ("B," "C," and "D")



**Fig. 3. Design of higher-order assemblies.** (A) Schematic overview of experimentally validated heterodimer-DHR fusions. In the colored circle at the center, the inner ring represents the heterodimer, the middle ring the protomer chain that is fused, and the outer ring the DHR (28) fusion partner. In the design model cartoons outside the colored circle, the patterning of the DHRs (in gray) is consistent throughout the paper. (B) Schematic representation of the design-free alignment method used to generate bivalent connectors from heterodimer-DHR fusions. Shown are LHD274B fused to the N terminus of DHR53 (274B53) (top left), LHD101A fused to the C terminus of DHR53 (101A53) (top right), and bivalent connector DFBO (bottom). (C) Representations of a heterotrimer (top) comprising the bivalent connector in (B) (“B”) and two of the rigid fusions shown in (A) (“A” is 274A53 and “C” is 101B62) and SEC traces for all possible

combinations of the trimer components (bottom). Abs 230, absorbance at 230 nm; mAU, milli-arbitrary units. (D) Schematic representations of three examples of bivalent connectors (see fig. S10A for a full list) and experimentally validated higher-order assemblies (see figs. S10 and S11). (E) Shown on the left is an overlay of the heterohexameric design model (in colors) and nsEM density (light gray). Shown on the right are SEC traces of partial and full mixtures of the hexamer components (“A” is 284A82, “B” is DF284, “C” is DFA-GFP, “D” is DF206, “E” is DF275A, and “F” is 275B). Absorbance was monitored at 473 nm to follow the GFP-tagged component C. Sequences, models, and chain-to-construct mapping are provided in data S1, affinities of individual interactions in tables S1 and S3, and the mapping of schemes to names for individual components in fig. S25.

simultaneously and conceptually resemble Ste5 and related scaffolding proteins that organize mitogen-activated protein (MAP) kinase signal transduction pathways in eukaryotes (30). Through SEC analyses, we verified the assembly of two different tetrameric branched ABCD complexes, each containing one trivalent branched connector bound to three terminal caps (Fig. 4A and fig. S16). For one of these, the complex was confirmed by nsEM class averages and 3D reconstructions, which indicate not only that all binding partners are present but also that the shape closely matches the designed model (Fig. 4A and fig. S16A).

A different type of branched assemblies are “star shaped” oligomers with cyclic symmetries, akin to natural assemblies formed by immunoglobulin M (IgM) and the inflammasome (31, 32). Using the alignment approach described above (Fig. 3B), we fused our building blocks (Fig. 3A) to previously designed homo-oligomers (23, 33) that terminate in helical

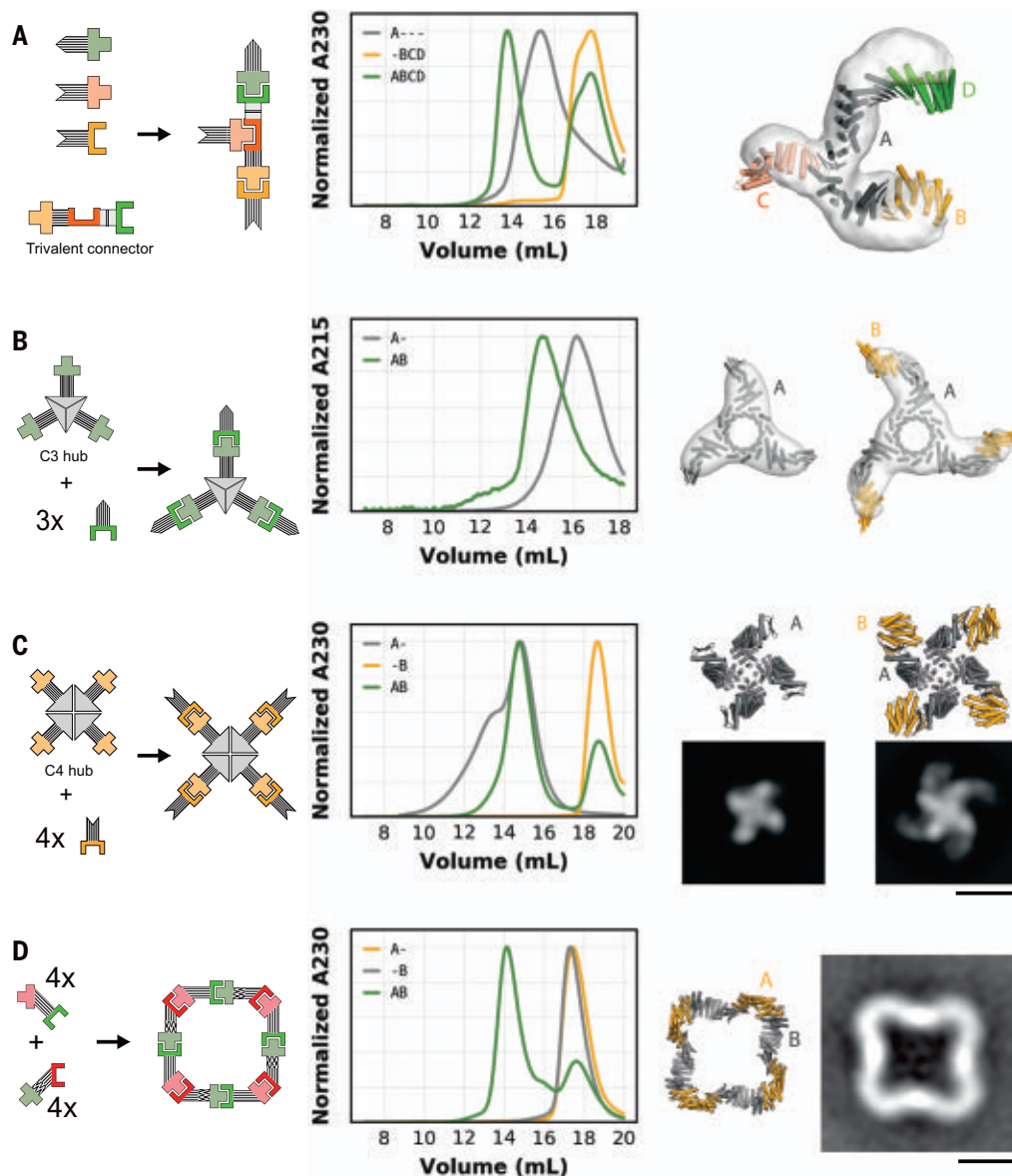
repeat proteins (Fig. 4, B and C). Such fusions yield central homo-oligomeric hubs (“A<sub>n</sub>”) that can bind multiple copies of the same binding partner (“n\*B”). We generated C3- and C4-symmetric “hubs” that can bind three or four copies of their binding partners, respectively (Fig. 4, B and C). In both cases, the oligomeric hubs are stable and soluble in isolation and readily form the target complexes when mixed with their binding partners, as confirmed by SEC, nsEM class averages, and 3D reconstructions (Fig. 4, B and C, and figs. S17 to S19). For the C4-symmetric hub, in the absence of its binding partner, we observed an additional concentration-dependent peak on SEC (Fig. 4C and figs. S18A and S19A), indicating formation of a higher-order complex. This is likely a dimer of C4 hubs, because the C4 hub contains the redesigned protomer LHD274B that, despite its reduced homodimerization propensity compared with parent design LHD29B, still weakly homodimerizes (fig. S5). Addition of

the binding partner drives reconfiguration of this higher-order assembly into the on-target octameric (A<sub>4</sub>B<sub>4</sub>) complex (Fig. 4C).

In addition to linear and branched assemblies, we designed closed symmetric two-component assemblies. Designing these presents a more complex geometric challenge, because the interaction geometry of all pairs of subunits must be compatible with a single closed 3D structure of the entire assembly. We used architecture-aware rigid helical fusion (7, 34) to generate two bivalent connector proteins from the crystal-verified fusions of LHD29 and LD101 (Fig. 2B) that allow assembly of a perfectly closed C4-symmetric hetero-oligomeric two-component ring (Fig. 4D). Individually expressed and purified components are stable and soluble monomers in isolation, as confirmed by SEC, multiangle light scattering (MALS) and native MS (Fig. 4D and fig. S20). Upon mixing, the components form a higher-order complex that, by native MS and MALS, comprises four copies of each component.

**Fig. 4. Design of branched and closed hetero-oligomeric assemblies.**

Schematic depictions of designs are shown in the first column, SEC binding data in the second column, and designed models in colors overlayed on a nsEM reconstruction in the third column. **(A)** A trivalent connector ("A" is TF10) binds three different binding partners ("B" is 274A53, "C" is 317B, and "D" is 101B62). **(B)** The C3-symmetric hub presenting three copies of LHD101B (chain A) binds three copies of its binding partner ("B" is 101A53). **(C)** The C4-symmetric hub presenting four copies of LHD274B binds its cognate binding partner (274A53). Representative nsEM class averages are shown on the right. **(D)** C4-symmetric closed ring comprising two components (A and B) assemblies from constituent components. A representative nsEM class average is shown on the right. Scale bars, 10 nm.



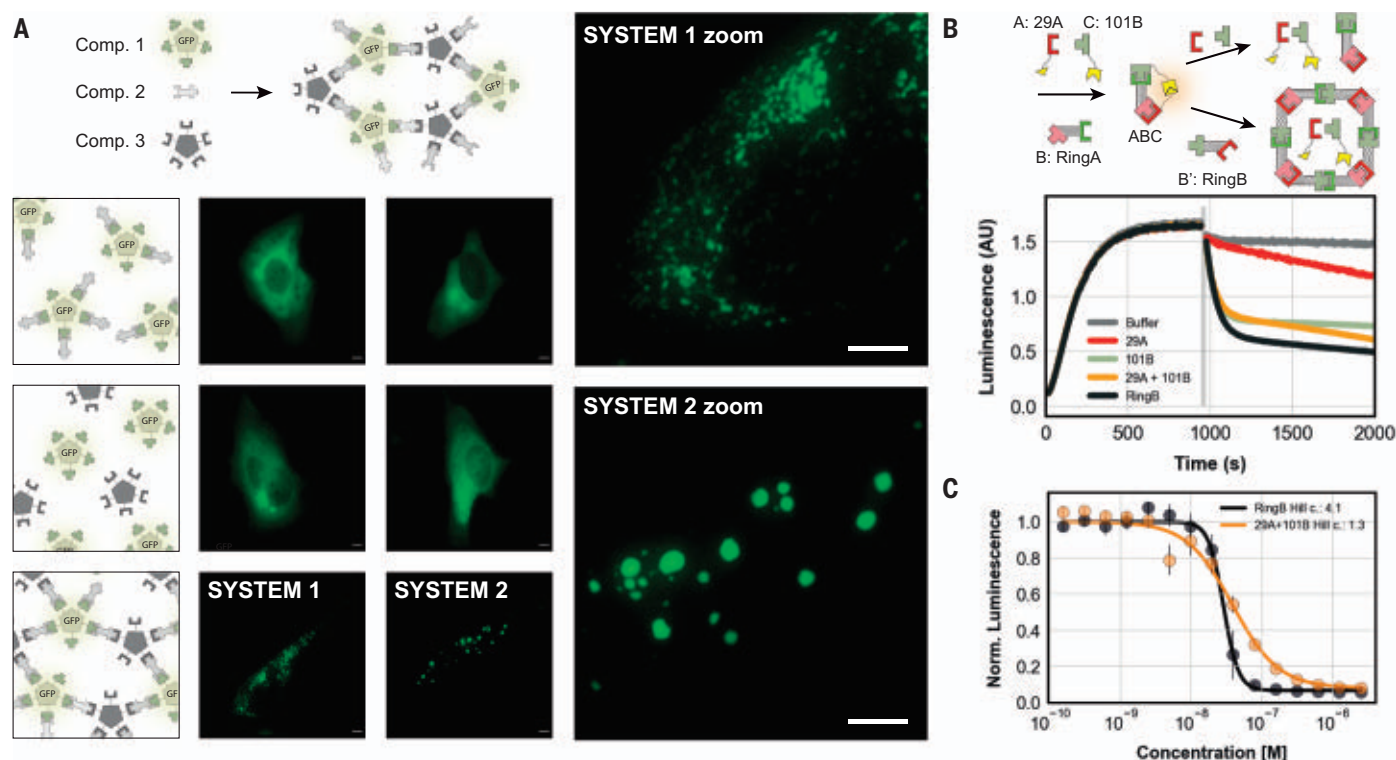
nsEM confirmed that this higher-order complex is similar to the designed C4-symmetric ring (Fig. 4D and fig. S21).

To determine whether our components function as designed in living cells, and to evaluate their use in constructing conditional assemblies, we fused one heterodimer protomer to a previously designed GFP-tagged C5 homo-oligomer (7) and a second protomer of a different heterodimer to an untagged C5 homo-oligomer. Transient expression of the two constructs in HeLa cells led to a distributed and diffuse GFP signal throughout the cell (Fig. 5A and fig. S22), suggesting that the components do not interact with each other or self-associate. However, when a bivalent connector (Fig. 3B) designed to link the two homo-oligomers was also expressed, the GFP signal redistributed into discrete puncta con-

sistent with the expected three-component extended meshwork (Fig. 5A and fig. S22). Notably, changing just one of the two heterodimer interfaces in the assembly from a high- to a low-affinity interface had a notable effect on the morphology of the puncta. When both interfaces had nanomolar affinity (system 1 in Fig. 5A), there were many small puncta, whereas substitution with a micromolar affinity heterodimer with a more rapid dissociation rate led to large droplet-like puncta (system 2 in Fig. 5A). These results show that, as designed, the components of the heterodimers are well behaved in isolation and assemble when combined in cells. The morphology differences further suggest that the ability to modulate dissociation rates and affinities of designed components could be advantageous for probing phase transitions in cells.

Because our designed building blocks are stable in solution and not kinetically trapped in off-target homo-oligomeric states, the assemblies they form can, in principle, reconfigure, as outlined in Fig. 1A and observed for the C4-symmetric hub shown in Fig. 4C. To examine reconfiguration dynamics, we constructed an ABC linear heterotrimer in which the B connector component is one of the two components of the ring shown in Fig. 4D and the A and C capping components are tagged with split luciferase fragments. In the absence of B, components A and C do not interact, and luciferase activity is not reconstituted (Fig. 5B). Upon addition of B, the heterotrimer forms, resulting in luciferase activity (Fig. 5B). Addition of the other ring component (B') to the preformed ABC trimer leads to a rapid decrease in luciferase activity, consistent with





**Fig. 5. Inducible and reconfigurable assemblies.** (A) Cross-linking of homopentamers by bivalent connectors in cells. Schematic representations of the components are at the top and in the first column, and fluorescence microscopy images of cells expressing different combinations of the components are in the second and third columns. High-affinity system 1 (second column) uses LHD101 and LHD275; low-affinity system 2 (third column) uses LHD101 and LHD321. See fig. S22 for additional control images. Scale bars, 5  $\mu$ m. (B) Schematic representation of an ABC heterotrimer (top) with split luciferase activity (yellow

shapes) undergoing subunit exchange through addition of non-luciferase-tagged components. Real-time luminescence measurements (bottom) of samples containing the mixture ABC shown at the top left. The gray bar indicates the addition of either buffer (gray trace), component RingB, or non-luciferase-tagged components LHD29A and LHD101B. AU, arbitrary units. (C) Titration of either component RingB or non-luciferase-tagged components LHD29A and LHD101B to the preformed ABC heterotrimer. Data are fitted to the Hill equation. Error bars represent SD.

disassembly of the trimer and formation of the ring (Fig. 5B and fig. S23, A to C). Because ring formation is cooperative owing to the additional interactions made upon ring closure, we reasoned that the concentration dependence of ABC trimer dissociation would be steeper upon addition of B' than with untagged A and C. To investigate this, we titrated B' and non-luciferase-tagged variants of A and C into the preformed trimer. There was a steep concentration dependence to the loss in luciferase signal upon addition of B' with a Hill coefficient of 4.1 (Fig. 5C and fig. S23D), consistent with the cooperative formation of a symmetrically closed ring ( $B_4B'_4$ ). By contrast, the loss of luciferase signal upon addition of nontagged A and C had a Hill coefficient close to 1, as expected for formation of a non-cooperative linear assembly (Fig. 5C and fig. S23D). In both cases, reconfiguration occurred on the several-minute time scale (fig. S23, B and C). We also observed reconfiguration of heterotrimers using SEC and BLI (fig. S24). This behavior, although common in naturally

evolved protein complexes, has been difficult to achieve by design, because it requires that the individual components not self-associate on their own. Our design principles pave the way for the design of functions requiring reconfigurable multiprotein complexes.

### Discussion

Our implicit negative-design principles enable the de novo design of heterodimer pairs for which the individual protomers are stable in solution and readily form their target heterodimeric complexes upon mixing, unlike previously designed assemblies. Rigid fusion of components through structured helical linkers enables the design of higher-order asymmetric multiprotein complexes in which individual subunits have well-defined positions relative to each other. Although rigidly fused building blocks may still exhibit flexibility (molecular breathing), fusion with structured connectors allows more control of subunit orientation than can be achieved by flexible linker fusion and enables fine-tuning of protein complex

geometries. Because of the small sizes of our unfused protomers (between 7 and 15 kDa without DHR or tags), complexes can readily be functionalized through genetic fusion of subunits with proteins of interest. Our bivalent or trivalent connectors can then be used to colocalize and geometrically position two or three such target protein fusions, respectively, and our symmetric hubs can be used to colocalize and position multiple copies of the same target fusion. Because of the modularity of our system, the same set of target fusions can be arranged in multiple different arrangements with adjustable distances, angles, and copy numbers by simply using different components (fig. S25). Because of the solubility and stability of the designs in isolation, complexes can be assembled stepwise (see, for example, Fig. 5A). The asymmetric complexes generated with our components will, in general, have low assembly cooperativity, so the fraction of fully assembled complex will be sensitive to the concentrations of the individual components over a broad range, enabling

subunit exchange and complex reconfiguration in response to signal inputs for synthetic biology and other applications. Because the thermodynamics and kinetics of our designed interfaces are not altered by fusion, the fraction of full assemblies and subassemblies, as well as assembly dynamics, can, in principle, be predicted based on the properties of the individual interfaces (fig. S23A). We expect that the design approach and components presented here will lead to a new generation of reconfigurable protein assemblies for a wide range of applications, including intracellular control for synthetic biology, design of protein logic gates, reprogramming cells from the outside by arraying receptor binding modules with specific geometries, processive multienzyme complexes, and designed molecular machines.

## Materials and methods

### Protein design

#### Docking procedure

As scaffolds for generating edge-strand heterodimers, we used mixed  $\alpha$ - $\beta$  proteins designed by citizen scientists (22) and variants of the Foldit scaffolds that were either expanded with additional helices (see backbone generation methods) and/or fused to designed helical repeat (DHR) proteins (28). Edge-strand docking was performed as described previously (19). Exposed edge strands suitable for docking were identified by calculating the solvent-accessible surface area of the  $\beta$  sheet backbone atoms in all the scaffolds used in the docking procedure. Next, the  $C\alpha$  atoms of each strand of short two-stranded parallel and antiparallel  $\beta$  sheet motifs were aligned to the exposed edge strand, yielding an aligned clashing strand and free docked strand. After removal of the aligned clashing strand, the docked strand was trimmed at the N and/or C terminus to remove potential clashes and subsequently minimized using Rosetta FastRelax (35) to optimize backbone-to-backbone hydrogen bonds. Docks that failed a specified threshold value (typically  $-4$  using ref2015) for the backbone hydrogen-bond score term in Rosetta (hbond\_lr\_bb) were discarded. The minimized docked strands were then geometrically matched to the scaffold library using the MotifGraftMover to create a docked protein-protein complex (36).

#### Interface design

The interface residues of the docked heterodimer complexes were optimized using Rosetta combinatorial sequence (37–40) design using “ref2015,” “beta\_nov16,” or “beta\_genpot” as score functions (41). The interface polarity of the docked heterodimer complexes were fine-tuned in several ways (see supplementary materials for a description of the design xml’s).

First, the HBNetMover (11) was used to install explicit hydrogen-bond networks that contained at least three hydrogen bonds across the interface. Later design rounds consisted of two separate interface sequence-optimization steps. First, interface residues were optimized without compositional constraints, yielding a substantial number of hydrophobic interactions in the interface. The best designs were subsequently selected, and hydrophobic residue pairs with the lowest Rosetta energy interactions across the interface were stored as a seed hydrophobic interaction hotspot (42). In a second round, a polar-interaction network was designed around the fixed hydrophobic hotspot interaction using compositional constraints that favor polar interactions (27). Designs were filtered on interface properties such as binding energy, buried surface area, shape complementarity, degree of packing, and presence of unsatisfied buried polar atoms. A final selection was made by visual inspection of models.

#### Homodimer self-docking

In later design rounds, the propensity for homodimerization was explicitly assessed in silico. Each individual chain of a heterodimer was docked onto itself through edge-strand docking (19) (see also the Docking procedure section). This creates a set of disembodied strands that pair with the scaffold edge strand that also participates in the heterodimeric complex. Homodimer docks were generated by aligning the heterodimerizing edge strand of a second copy of the scaffold back onto the disembodied docked strand (see fig. S7A). Docks with different  $\beta$  register offsets and orientations (parallel and antiparallel) were created. Docks were next converted to polypeptide and clash-checked. Docks where the repulsive Rosetta scoreterm (fa\_rep) was higher than 250 (scorefunction ref2015) were discarded (i.e., no homodimer possible). Surviving docks were converted to full atom models and minimized using FastRelax (35) followed by scoring and assessing of homodimer interface metrics such as binding energy, buried surface area, shape complementarity, degree of packing, and presence of unsatisfied buried polar atoms.

#### Backbone generation and scaffold design

De novo designed protein scaffolds created by Foldit players (22) were expanded with C-terminal polyvaline helices using blueprint-based backbone generation (24, 25). The amino acid identities of the newly built helices and their surrounding region were optimized using Rosetta combinatorial sequence design using a flexible backbone. The resulting models were folded in silico using Rosetta folding simulations, and trajectories that converged to the designed model

structure without off-target minima were selected for rigid fusion and heterodimer design.

#### Design of rigid fusions

To generate rigid fusions of scaffolds or heterodimers to DHRs, we adapted the HFuse pipeline (7, 23): Fusion junctions were designed using the Fastdesign mover to allow backbone movement, and additional filters were included to ensure sufficient contact between the DHR and the fusion partner. When fusing to heterodimers, an additional filter was used to prevent additional contacts between the DHR and the other protomer of the dimer. Bivalent connectors were generated by aligning two proteins that share the same DHR along their shared helical repeats and subsequently splicing together the sequences. To build the C3-symmetric hub, we used a previously published crystal structure of a 12-repeat toroid ring (33). The starting structure was relaxed, its  $z$  axis was aligned, and it was cut into three C3 symmetric chains. Then the HFuse software (7, 23) was used to sample DHR fusions to the exposed helical C termini, and the newly created interfaces were redesigned using RosettaScripts. For the C4-symmetric hub, we used a previously published C4 symmetric homo-oligomer that already contained a N-terminal DHR. Both DHR-containing hubs were fused to LHD protomers in the same way as described above for the bivalent connectors.

#### Design of C4 rings

Using the relaxed crystal structures of LHD29 and LHD101 fused to their respective DHRs, the WORMS software (7, 9, 34) was used to fuse the two heterodimers into cyclic symmetrical rings. Because one construct has exposed N termini and the other has exposed C termini, they were able to be fused head to tail without introduction of further building blocks. Briefly, the first three repeats of each repeat protein were allowed to be sampled as fusion points to ensure that the heterodimer interface was not altered. After fusion into cyclic structures, fixed backbone junction design was applied to the new fusion point using RosettaScripts (39), optimizing for shape complementarity (43). One design from each symmetry—C3, C4, C5, and C6—was selected for experimental testing.

#### Protein expression and purification

Synthetic genes encoding designed proteins and their variants were purchased from Genscript or Integrated DNA technologies (IDT). Bicistronic genes were ordered in pET29b, with the first cistron being either without tag or with an N-terminal sfGFP tag followed by the intercistronic sequence TAAAGAAGGAGATATCA-TATG. The second cistron was tagged with a polyhistidine His6 $\times$  tag at the C terminus.

Plasmids encoding the individual protomers were ordered in pET29b with an N-terminal polyhistidine His6× tag followed by a TEV cleavage site, N-terminal polyhistidine His6× tag followed by a snac cleavage site or C-terminal polyhistidine His6× tag preceded by a snac tag. For enzymatic biotinylation reactions, an Avi-Tag was included at either the N or C terminus (see data S1 for detailed construct information). Proteins were expressed in BL21 LEMO *E. coli* cells by autoinduction using TBII media (Mpbio) supplemented with 50x5052, 20 mM MgSO<sub>4</sub>, and trace metal mix or in almost TB media containing 12 g of peptone and 24 g of yeast extract per liter supplemented with 50x5052, 20 mM MgSO<sub>4</sub>, trace metal mix, and 10× phosphate buffer. Proteins were expressed under antibiotic selection at 37°C overnight or at 18°C for 24 hours after initial growth for 6 to 8 hours at 37°C. Cells were harvested by centrifugation at 4000g and lysed by sonication after resuspension of the cells in lysis buffer (100 mM Tris pH 8.0, 200 mM NaCl, 50 mM imidazole pH 8.0) containing protease inhibitors (Thermo Scientific) and Bovine pancreas DNaseI (Sigma-Aldrich). Proteins were purified by immobilized metal affinity chromatography (IMAC). Cleared lysates were incubated with 2 to 4 ml of nickel-nitriloacetic acid (NTA) beads (Qiagen) for 20 to 40 min before washing beads with 5 to 10 column volumes of lysis buffer, 5 to 10 column volumes of high-salt buffer (10 mM Tris pH 8.0, 1 M NaCl), and 5 to 10 column volumes of lysis buffer. Proteins were eluted with 10 ml of elution buffer (20 mM Tris pH 8.0, 100 mM NaCl, 500 mM imidazole pH 8.0).

Designs were finally polished using SEC on either Superdex 200 Increase 10/300GL or Superdex 75 Increase 10/300GL columns (GE Healthcare) using 20 mM Tris pH 8.0, 100 mM NaCl or 20 mM Tris pH 8.0, 300 mM NaCl. Cyclic assemblies of C3 and C4 symmetries were purified using a Superose 6 increase 10/300GL (GE Healthcare). The two component C4 rings were purified by SEC in 25 mM Tris pH 8.0, 300 mM NaCl. Peak fractions were verified by SDS-PAGE and LC/MS and stored at concentrations between 0.5 and 10 mg/ml at 4°C or flash frozen in liquid nitrogen for storage at -80°C. Designs that precipitated at low concentration upon storage at 4°C could, in general, be salvaged by increasing the salt concentration to 300 to 500 mM NaCl.

For structural studies, designs with a polyhistidine tag and TEV recognition site were cleaved using TEV protease (his6-TEV). TEV cleavage was performed in a buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, and 1 mM TCEP using 1% (w/w) his6-TEV and allowed to proceed overnight at room temperature. Uncleaved protein and his6-TEV were separated

from cleaved protein using IMAC followed by SEC. Designs carrying a C-terminal SNAC-polyhistidine tag [GGSHHWGS(...)HHHHHH] were cleaved chemically by on-bead nickel-assisted cleavage (44): Nickel-bound designs were washed with 10 column volumes of lysis buffer followed by 5 column volumes of 20 mM Tris pH 8.0, 100 mM NaCl. Proteins were subsequently washed with 5 column volumes of SNAC buffer (100 mM CHES, 100 mM acetone oxime, 100 mM NaCl, pH 8.6). Beads were next incubated with 5 column volumes of SNAC buffer with 2 mM NiCl<sub>2</sub> for more than 12 hours at room temperature on a shaking platform to allow cleavage to take place. Next, the flow-through that contained cleaved protein was collected. The flow-throughs of two additional washes (SNAC buffer/SNAC buffer with 50 mM Imidazole) of 3 to 5 column volumes were also collected to harvest any remaining weakly bound protein. Cleaved proteins were finally purified by SEC.

For mammalian cell expression, synthetic genes encoding designed proteins were purchased from Genscript and cloned into mammalian expression vectors. LHD101B-C5 was cloned into the KpnI-XbaI site of pCDNA3.1+N-eGFP in frame with enhanced GFP (eGFP). Both LHD275B\_53\_0\_LHD101A and LHD321B\_53\_LHD101A were cloned into the NheI-XbaI site of pCDNA3.1+C-HA. LHD275A-C5 and LHD321A-C5 were cloned into the KpnI-XbaI site of pCDNA3.1+N-HA.

#### Cell culture and transient transfections

HeLa cells (ATCC CCL-2) were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco) that was supplemented with 1 mM L-glutamine (Gibco), 4.5 g/liter D-glucose (Gibco), 10% fetal bovine serum (FBS), and (1×) nonessential amino acids (Gibco). Cells were cultured at 37°C and 5% CO<sub>2</sub> and passaged twice per week. To passage, cells were dissociated using 0.05% trypsin EDTA (Gibco) and split 1:5 or 1:10 into a new tissue culture (TC)-treated T75 flask (Thermo Scientific ref 156499).

HeLa cells were plated at 20,000 cells per well in Cellview cell culture slides (Greiner Bio-One ref 543079). Twenty-four hours later, cells were transiently transfected at a concentration of 1875 ng total DNA per well and 1 µg/µl PEI-MAX (Polysciences) mixed with Opti-MEM medium (Gibco). Transfected cells were incubated at 37°C and 5% CO<sub>2</sub> for 24 to 36 hours before being imaged.

#### Fluorescence microscopy and image processing

3D images were acquired with a commercial OMX-SR system (GE Healthcare). A 488-nm Topica diode laser was used for excitation. Emission was collected on a PCO.edge scientific complementary metal-oxide semiconductor (sCMOS) camera using an Olympus 60× 1.42NA PlanApochromat oil immersion lens. Images

were acquired in a 1024 by 1024 field of view (pixel size 6.5 µm) with no binning. Acquisition was controlled with AcquireSR acquisition control software. Z-stacks were collected with a step size of 500 nm and 15 slices per image. Images were deconvolved with an enhanced ratio using SoftWoRx 7.0.0 (GE Healthcare). Cell images were sum-projected using Fiji v2.1.0. Scale bars equal 5 µm.

#### Enzymatic protein biotinylation

Avi-tagged (GLNDIFEAAQKIEWHE; see supplementary materials) proteins were purified as described above. The BirA500 (Avidity, LLC) biotinylation kit was used to biotinylate 840 µl of protein from the IMAC elution in a 1200 µl (final volume) reaction according to the manufacturer's protocol. Reactions were incubated at 4°C overnight and purified using SEC on a Superdex 200 10/300 Increase GL (GE Healthcare) or S75 10/300 Increase GL (GE Healthcare) in SEC buffer (20 mM Tris pH 8.0, 100 mM NaCl).

#### Biolayer interferometry

Biolayer interferometry experiments were performed on an OctetRED96 BLI system (ForteBio, Menlo Park, CA). Streptavidin-coated biosensors were first equilibrated for at least 10 min in Octet buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P20) supplemented with 1 mg/ml bovine serum albumin (SigmaAldrich). Enzymatically biotinylated designs were immobilized onto the biosensors by dipping the biosensors into a solution with 10 to 50 nM protein for 30 to 120 s. This was followed by dipping in fresh octet buffer to establish a baseline for 120 s. Titration experiments were performed at 25°C while rotating at 1000 rpm. Association of designs was allowed by dipping biosensors in solutions containing designed protein diluted in octet buffer until equilibrium was approached followed by dissociation by dipping the biosensors into fresh buffer solution to monitor the dissociation kinetics. Steady-state and global kinetic fits were performed using the manufacturer's software (Data Analysis 9.1) assuming a 1:1 binding model.

#### SEC binding assays

Complexes and individual components were diluted in 20 mM Tris pH 8.0, 100 mM NaCl. After overnight equilibration of the mixtures at room temperature or 4°C, 500 µl of sample was injected onto a Superdex 200 10/300 increase GL (dimers, linear assemblies) or Superose 6 increase 10/300 GL (symmetric assemblies) (all columns from GE healthcare) using the absorbance at 230 or 473 nm (for GFP-tagged components) as readout. Dimers were mixed at monomer concentrations of 5 µM or higher. Trimer and ABCD tetramer mixtures contained 5 µM of the bivalent



connector and 7.5  $\mu\text{M}$  of each terminal cap. (Lower absolute concentrations with the same ratios were used for some trimers.) ABCA tetramer mixtures contained 5  $\mu\text{M}$  per bivalent connector and 15  $\mu\text{M}$  terminal cap. The hexamer mixture contained 3  $\mu\text{M}$  of components C and D, 3.6  $\mu\text{M}$  of B and E, and 4.4  $\mu\text{M}$  of A and F. The branched assembly shown in Fig. 4A contained 2.8  $\mu\text{M}$  of the trivalent connector and 4  $\mu\text{M}$  of each cap. For the exchange experiment shown in Fig. S24A, the ABC trimer was preincubated at concentrations of 6  $\mu\text{M}$  B and 9  $\mu\text{M}$  each of A and C. C' was then added to reach a final concentration of 2  $\mu\text{M}$  B, 3  $\mu\text{M}$  each of A and C, and 6  $\mu\text{M}$  C'.

### Native mass spectrometry

Sample purity, integrity, and oligomeric state were analyzed by on-line buffer exchange MS in 200 mM ammonium acetate using a Vanquish ultra-high performance LC system coupled to a Q Exactive ultra-high mass range Orbitrap mass spectrometer (Thermo Fisher Scientific). A self-packed buffer exchange column was used (P6 polyacrylamide gel, BioRad) (45). The recorded mass spectra were deconvolved with UniDec version 4.2+ (46).

### Crystal structure determination

For all structures, starting phases were obtained by molecular replacement using Phaser (47). Diffraction images were integrated using XDS (48) or HKL2000 (49) and merged and scaled using Aimless (50). Structures were refined in Phenix (51) using phenix.autobuild and phenix.refine or Refmac (52). Model building was performed using COOT (53).

Proteins were crystallized using the vapor diffusion method at room temperature. LHD29 crystals grew in 0.2 M sodium iodide, 20% PEG3350; LHD29A53/B53 crystals in 3.2 M ammonium sulfate, 0.1 M citric acid pH 4.0; and LHD101A53/B4 crystals in 2.4 M sodium malonate pH 7.0. Crystals were harvested and cryoprotected using 20% PEG200 for LHD29, 20% PEG400 for LHD29A53/B53, and 20% glycerol for LHD101A53/B4 before data was collected at the Advanced Light Source (Berkeley, USA). The structures were solved by molecular replacement using either computationally designed models of individual chains A or B or the full heterodimer complex as search models.

The root mean square deviation (RMSD), TMscore, and local distance difference test (LDDT) metrics between the designed models and corresponding crystal structures were calculated as described previously (54, 55). Protein structure graphics were prepared using PyMOL (Schrödinger).

### Electron microscopy

SEC peak fractions were concentrated before nsEM screening. Samples were then immediately diluted 5 to 150 times in Tris-buffered

saline (TBS) buffer (25 mM Tris pH 8.0, 25 mM NaCl) depending on sample concentration. A final volume of 5  $\mu\text{l}$  was applied to negatively glow discharged, carbon-coated 400-mesh copper grids (01844-F, TedPella, Inc.) and then washed with Milli-Q Water and stained using 0.75% uranyl formate as previously described (56). Air-dried grids were imaged on a FEI Talos L120C TEM (FEI Thermo Scientific, Hillsboro, OR) equipped with a 4K  $\times$  4K Gatan OneView camera at a magnification of 57,000 $\times$  and pixel size of 2.51. Micrographs were imported into CislTEM software or cryoSPARC software, and a circular blob picker was used to select particles that were then subjected to 2D classification. Ab initio reconstruction and homogeneous refinement in Cn symmetry were used to generate 3D electron density maps (57, 58).

### Constructs for luciferase assays

Split luciferase reporter constructs were ordered as synthetic genes from Genscript. Each design was N-terminally fused to a superfolder GFP (sfGFP) (for protein quantification in lysate) and C-terminally fused to either smBIT or lgBiT of the split luciferase components. A Strep-tag was included at the N terminus for purification, and a glycine-serine (GS)-linker was inserted between the design and the split luciferase component.

### Expression for multiplexed luciferase assay

Plasmids were transformed into Lemo21(DE3) cells (New England Biolabs) and grown in 96-deep-well plates overnight at 37°C in 1 ml of LB containing 50  $\mu\text{g}/\text{ml}$  of kanamycin sulfate. The next day, 100  $\mu\text{l}$  of overnight cultures were used to inoculate 96-deep-well plates containing 900  $\mu\text{l}$  of TBII medium (MP Biomedicals) with 50  $\mu\text{g}/\text{ml}$  of kanamycin sulfate, and the cultures were grown for 2 hours at 37°C before induction with 0.1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Protein expression was carried out at 37°C for 4 hours before the cells were harvested by centrifugation (4000g, 5 min). Cell pellets were resuspended in 100  $\mu\text{l}$  of lysis buffer (10 mM sodium phosphate, 150 mM NaCl, pH 7.4, 1 mg/ml lysozyme, 0.1 mg/ml DNase I, 5 mM  $\text{MgCl}_2$ , 1 tablet per 50 ml of cOmplete protease inhibitor (Roche), 0.05% v/v Tween 20), and cells were lysed by performing three freeze-thaw cycles (1 hour incubations at 37°C followed by freezing at  $-80^\circ\text{C}$ ). The lysate was cleared by centrifugation (4000g, 20 min), and the soluble fraction was transferred to a 96-well assay plate (Corning, cat. no. 3991). Concentrations of the constructs in soluble lysate were determined by sfGFP fluorescence using a calibration curve.

### Lysate production for multiplexed luciferase assay

Neutral lysate for preparing serial dilutions was prepared by transforming Lemo21(DE3)

with the pUC19 plasmid. Transformations were used to inoculate small overnight cultures, which were used to inoculate 0.5-liter TBII cultures (all cultures contained 50  $\mu\text{g}/\text{ml}$  of carbenicillin). Cells were grown for 24 hours at 37°C before being harvested. Pellets were resuspended in the same lysis buffer, followed by sonication. The lysate density was adjusted with lysis buffer to have its OD280 (optical density at 280 nm) match that of pUC19 control wells from the 96-well expression plate.

### Expression and purification of luciferase constructs

Plasmids were transformed into Lemo21 (DE3) cells and used directly to inoculate 50 ml of autoinduction media (TBII supplemented with 0.5% w/v glucose, 0.05% w/v glycerol, 0.2% w/v lactose monohydrate, and 2 mM  $\text{MgSO}_4$ , 50  $\mu\text{g}/\text{ml}$  kanamycin sulfate). The cultures were incubated at 37°C for 20 to 24 hours before harvesting the cells by centrifugation (4000g, 5 min). Cells were resuspended in 10 ml of lysis buffer [100 mM Tris, 150 mM NaCl, pH 8, 0.1 mg/ml lysozyme, 0.01 mg/ml DNase I, 1 mM phenylmethylsulfonyl fluoride (PMSF)] and lysed by sonication. The insoluble fraction was cleared by centrifugation (16,000g, 45 min), and the proteins were purified from the soluble fraction by affinity chromatography using Strep-Tactin XT Superflow High-Capacity resin (IBA Lifesciences). Elutions were performed with 100 mM Tris, 150 mM NaCl, 50 mM biotin, pH 8, and the proteins were further purified by SEC using a Superdex 200 10/300 increase column equilibrated with 20 mM sodium phosphate, 100 mM NaCl, pH 7.4, 0.05% v/v Tween 20.

### Luciferase binding assays

All assays were performed in 20 mM sodium phosphate, 100 mM NaCl, pH 7.4, 0.05% v/v Tween 20. Depending on the source of the protein used in the assay (purified components or lysate), soluble lysate components were also present. Reactions were assembled in 96-well plates (Corning, cat. no. 3686) in the presence of Nano-Glo substrate (Promega, cat. no. N1130) and diluted 100 $\times$  or 500 $\times$  for kinetics and endpoint measurements, respectively, and the luminescence signal was recorded on a Synergy Neo2 plate reader (BioTek).

Kinetic binding assays were performed under pseudo first-order conditions, with the final concentration of one protein at 1 nM and the other at 10 nM. Stock solutions were mixed in a 1:1 volume ratio in the presence of substrate, and the dead-time between mixing and starting the measurement (typically 15 to 30 s) was added during data processing. For long kinetic measurements (Fig. S6A), the proteins were premixed and kept in a sealed tube at room temperature over the course of the experiment. Aliquots were taken at regular intervals, mixed

with substrate, and immediately recorded. All kinetic measurements were fitted to a single exponential decay function:

$$S = A \cdot \exp(-k_{\text{obs}} \cdot t) + B$$

where  $t$  is time (the independent variable) and  $S$  is the observed luminescence signal (the dependent variable) and the fitted parameters are the amplitude  $A$ , the observed rate constant  $k_{\text{obs}}$ , and the endpoint luminescence  $B$ .

Equilibrium binding assays were performed with one component kept constant at 1 nM while titrating the other protein. Serial dilutions curves were prepared over 12 points, with a one-quarter dilution factor between each step. The concentration of protein in the soluble lysate provided the highest concentration point of the curve. To avoid serial dilution of the other lysate components, all stocks were prepared with neutral lysate. The assembled plates were incubated overnight at room temperature before adding substrate and immediately measuring luminescence. The data was fitted to the following equation to obtain  $K_d$  values:

$$S = S_0 + S_1 \cdot f_{AB} + a_2 \cdot B_T \cdot S_2$$

$$f_{AB} = [A_T + B_T + K_d - (A_T + B_T + K_d)^2 - 4A_TB_T] / 2A_T$$

where  $A_T$  and  $B_T$  are the total concentrations of each species (the independent variables,  $A_T = 1$  nM,  $B_T$  is the titrated species) and  $S$  is the observed signal (the dependent variable). The fitted parameters are the pre-saturation baseline  $S_0$ , the postsaturation baseline  $S_1$ , and the correction terms  $a_2$  and  $S_2$ .

Ternary complex equilibrium binding experiments were performed with pure protein, using the concentration indicated in fig. S23 for the constant components, and titrating B. After assembly, the plates were incubated overnight before adding substrate and immediately measuring luminescence.

Ternary complex reconfiguration kinetics (Fig. 5B and fig. S23) were measured with pure proteins. Components A (1 nM) and C (100 nM) were briefly preincubated in the presence of substrate (1/500 dilution) before adding component B (50 nM) to start the reaction. Once the association reactions were complete, the assay plate was briefly taken out of the plate reader; out-competing protein(s) (100 nM each in Fig. 5B and fig. S23B and 1000 nM each in fig. S23C) were added to the reactions; and data acquisition was resumed.

Ternary complex thermodynamic out-competitions (Fig. 5C and fig. S23D) were measured with purified proteins. Final concentrations of components A-smBiT, B, and C-IgBiT were 1, 50, and 100 nM final, respec-

tively. The out-competitor(s) (B' or untagged A+C) were titrated from 10 uM down to about 1 pM over 24 points, with a one-half dilution factor between each step. Reactions were incubated at room temperature for 2 to 5 hours before adding substrate (1/500 dilution) and measuring luminescence. The averages of four experiments were fitted to the Hill equation:

$$S = S_0 + (S_1 - S_0) / [1 + (K/L)^n]$$

where  $L$  is the total concentration of the out-competitor(s) (the independent variable) and  $S$  is the observed signal (the dependent variable). The fitted parameters are the presaturation baseline  $S_0$ , the postsaturation baseline  $S_1$ , the transition midpoint  $K$ , and the Hill coefficient  $n$ .

### Simulation of ternary complex

Systems of ordinary differential equations describing the kinetics of interactions between the species involved in the formation of the ternary complex (fig. S23A) were numerically integrated using `scipy.integrate.odeint` as implemented in `Scipy` (version 1.6.3). Steady-state values were used to determine the distribution of species at thermodynamic equilibrium.

The ternary system is composed of the following species: A, B, C, AB, BC, and ABC. The following set of equations was used to describe the system:

$$\begin{aligned} d[A]/dt &= -k_1[A][B] + k_{-1}[AB] - \\ &\quad k_1[A][BC] + k_{-1}[ABC] \end{aligned}$$

$$\begin{aligned} d[B]/dt &= -k_1[A][B] + k_{-1}[AB] - \\ &\quad k_2[B][C] + k_{-2}[BC] \end{aligned}$$

$$\begin{aligned} d[C]/dt &= -k_2[B][C] + k_{-2}[BC] - \\ &\quad k_2[AB][C] + k_{-2}[ABC] \end{aligned}$$

$$\begin{aligned} d[AB]/dt &= k_1[A][B] - k_{-1}[AB] + \\ &\quad k_{-2}[ABC] - k_2[AB][C] \end{aligned}$$

$$\begin{aligned} d[BC]/dt &= k_2[B][C] - k_{-2}[BC] + \\ &\quad k_{-1}[ABC] - k_1[A][BC] \end{aligned}$$

$$\begin{aligned} d[ABC]/dt &= k_1[A][BC] - k_1[ABC] + \\ &\quad k_2[AB][C] - k_{-2}[ABC] \end{aligned}$$

where  $k_i$  describes bimolecular association rate constants and  $k_{-i}$  represents unimolecular dissociation rate constants.  $K_1 = k_{-1}/k_1$ , and  $K_2 = k_{-2}/k_2$  describe the affinity of the A:B and B:C interfaces, respectively.

### REFERENCES AND NOTES

1. S. E. Tusk, N. J. Delalez, R. M. Berry, Subunit exchange in protein complexes. *J. Mol. Biol.* **430**, 4557–4579 (2018). doi: [10.1016/j.jmb.2018.06.039](https://doi.org/10.1016/j.jmb.2018.06.039); pmid: [29959924](https://pubmed.ncbi.nlm.nih.gov/29959924/)

2. C. Engel, S. Neyer, P. Cramer, Distinct mechanisms of transcription initiation by RNA polymerases I and II. *Annu. Rev. Biophys.* **47**, 425–446 (2018). doi: [10.1146/annurev-biophys-070317-033058](https://doi.org/10.1146/annurev-biophys-070317-033058); pmid: [29792819](https://pubmed.ncbi.nlm.nih.gov/29792819/)
3. P. M. J. Burgers, T. A. Kunkel, Eukaryotic DNA replication fork. *Annu. Rev. Biochem.* **86**, 417–438 (2017). doi: [10.1146/annurev-biochem-061516-044709](https://doi.org/10.1146/annurev-biochem-061516-044709); pmid: [28301743](https://pubmed.ncbi.nlm.nih.gov/28301743/)
4. S. Gonen, F. DiMaio, T. Gonen, D. Baker, Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **348**, 1365–1368 (2015). doi: [10.1126/science.1258897](https://doi.org/10.1126/science.1258897); pmid: [26089516](https://pubmed.ncbi.nlm.nih.gov/26089516/)
5. Y. Hsia et al., Design of a hyperstable 60-subunit protein dodecahedron. [corrected]. *Nature* **535**, 136–139 (2016). doi: [10.1038/nature18010](https://doi.org/10.1038/nature18010); pmid: [27309817](https://pubmed.ncbi.nlm.nih.gov/27309817/)
6. N. P. King et al., Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108 (2014). doi: [10.1038/nature13404](https://doi.org/10.1038/nature13404); pmid: [24870237](https://pubmed.ncbi.nlm.nih.gov/24870237/)
7. Y. Hsia et al., Design of multi-scale protein complexes by hierarchical building block fusion. *Nat. Commun.* **12**, 2294 (2021). doi: [10.1038/s41467-021-22276-z](https://doi.org/10.1038/s41467-021-22276-z); pmid: [33863889](https://pubmed.ncbi.nlm.nih.gov/33863889/)
8. A. J. Ben-Sasson et al., Design of biologically active binary protein 2D materials. *Nature* **589**, 468–473 (2021). doi: [10.1038/s41586-020-03120-8](https://doi.org/10.1038/s41586-020-03120-8); pmid: [33408408](https://pubmed.ncbi.nlm.nih.gov/33408408/)
9. R. Divine et al., Designed proteins assemble antibodies into modular nanocages. *Science* **372**, eabd9994 (2021). doi: [10.1126/science.abd9994](https://doi.org/10.1126/science.abd9994); pmid: [33795432](https://pubmed.ncbi.nlm.nih.gov/33795432/)
10. Z. Chen et al., Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111 (2019). doi: [10.1038/s41586-018-0802-y](https://doi.org/10.1038/s41586-018-0802-y); pmid: [30568301](https://pubmed.ncbi.nlm.nih.gov/30568301/)
11. S. E. Boyken et al., De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352**, 680–687 (2016). doi: [10.1126/science.1248865](https://doi.org/10.1126/science.1248865); pmid: [27151862](https://pubmed.ncbi.nlm.nih.gov/27151862/)
12. Z. Chen et al., De novo design of protein logic gates. *Science* **368**, 78–84 (2020). doi: [10.1126/science.aby2790](https://doi.org/10.1126/science.aby2790); pmid: [32241946](https://pubmed.ncbi.nlm.nih.gov/32241946/)
13. H. Gradišar, R. Jerala, De novo design of orthogonal peptide pairs forming parallel coiled-coil heterodimers. *J. Pept. Sci.* **17**, 100–106 (2011). doi: [10.1002/psc.1331](https://doi.org/10.1002/psc.1331); pmid: [21234981](https://pubmed.ncbi.nlm.nih.gov/21234981/)
14. C. L. Edgell, A. J. Smith, J. L. Beesley, N. J. Savery, D. N. Woolfson, De novo designed protein-interaction modules for in-cell applications. *ACS Synth. Biol.* **9**, 427–436 (2020). doi: [10.1021/acssynbio.9b00453](https://doi.org/10.1021/acssynbio.9b00453); pmid: [31977192](https://pubmed.ncbi.nlm.nih.gov/31977192/)
15. A. Leaver-Fay, R. Jacak, P. B. Stranges, B. Kuhlman, A generic program for multistate protein design. *PLOS ONE* **6**, e20937 (2011). doi: [10.1371/journal.pone.0020937](https://doi.org/10.1371/journal.pone.0020937); pmid: [21754981](https://pubmed.ncbi.nlm.nih.gov/21754981/)
16. A. Leaver-Fay et al., Computationally designed bispecific antibodies using negative state repertoires. *Structure* **24**, 641–651 (2016). doi: [10.1016/j.str.2016.02.013](https://doi.org/10.1016/j.str.2016.02.013); pmid: [26996964](https://pubmed.ncbi.nlm.nih.gov/26996964/)
17. J. J. Havranek, P. B. Harbury, Automated design of specificity in molecular recognition. *Nat. Struct. Biol.* **10**, 45–52 (2003). doi: [10.1038/nsb877](https://doi.org/10.1038/nsb877); pmid: [12459719](https://pubmed.ncbi.nlm.nih.gov/12459719/)
18. S. J. Fleishman, D. Baker, Role of the biomolecular energy gap in protein design, structure, and evolution. *Cell* **149**, 262–273 (2012). doi: [10.1016/j.cell.2012.03.016](https://doi.org/10.1016/j.cell.2012.03.016); pmid: [22500796](https://pubmed.ncbi.nlm.nih.gov/22500796/)
19. D. D. Sahtoe et al., Transferrin receptor targeting by de novo sheet extension. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021569118 (2021). doi: [10.1073/pnas.2021569118](https://doi.org/10.1073/pnas.2021569118); pmid: [33879614](https://pubmed.ncbi.nlm.nih.gov/33879614/)
20. P. B. Stranges, M. Machiusi, M. J. Miley, A. Tripathy, B. Kuhlman, Computational design of a symmetric homodimer using  $\beta$ -strand assembly. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20562–20567 (2011). doi: [10.1073/pnas.1115124108](https://doi.org/10.1073/pnas.1115124108); pmid: [22143762](https://pubmed.ncbi.nlm.nih.gov/22143762/)
21. H. Remaut, G. Waksman, Protein-protein interaction through  $\beta$ -strand addition. *Trends Biochem. Sci.* **31**, 436–444 (2006). doi: [10.1016/j.tibs.2006.06.007](https://doi.org/10.1016/j.tibs.2006.06.007); pmid: [16828554](https://pubmed.ncbi.nlm.nih.gov/16828554/)
22. B. Koepnick et al., De novo protein design by citizen scientists. *Nature* **570**, 390–394 (2019). doi: [10.1038/s41586-019-1274-4](https://doi.org/10.1038/s41586-019-1274-4); pmid: [31168091](https://pubmed.ncbi.nlm.nih.gov/31168091/)
23. T. J. Brunette et al., Modular repeat protein sculpting using rigid helical junctions. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 8870–8875 (2020). doi: [10.1073/pnas.1908768117](https://doi.org/10.1073/pnas.1908768117); pmid: [32245816](https://pubmed.ncbi.nlm.nih.gov/32245816/)
24. Y.-R. Lin et al., Control over overall shape and size in de novo designed proteins. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5478–E5485 (2015). doi: [10.1073/pnas.1509508112](https://doi.org/10.1073/pnas.1509508112); pmid: [26396255](https://pubmed.ncbi.nlm.nih.gov/26396255/)
25. N. Koga et al., Principles for designing ideal protein structures. *Nature* **491**, 222–227 (2012). doi: [10.1038/nature11600](https://doi.org/10.1038/nature11600); pmid: [23135467](https://pubmed.ncbi.nlm.nih.gov/23135467/)
26. J. K. Leman et al., Macromolecular modeling and design in Rosetta: Recent methods and frameworks. *Nat. Methods* **17**,

- 665–680 (2020). doi: [10.1038/s41592-020-0848-2](https://doi.org/10.1038/s41592-020-0848-2); pmid: [32483333](https://pubmed.ncbi.nlm.nih.gov/32483333/)
27. B. Coventry, D. Baker, Protein sequence optimization with a pairwise decomposable penalty for buried unsatisfied hydrogen bonds. *PLOS Comput. Biol.* **17**, e1008061 (2021). doi: [10.1371/journal.pcbi.1008061](https://doi.org/10.1371/journal.pcbi.1008061); pmid: [33684097](https://pubmed.ncbi.nlm.nih.gov/33684097/)
  28. T. J. Brunette *et al.*, Exploring the repeat protein universe through computational protein design. *Nature* **528**, 580–584 (2015). doi: [10.1038/nature16162](https://doi.org/10.1038/nature16162); pmid: [26675729](https://pubmed.ncbi.nlm.nih.gov/26675729/)
  29. J. R. Lydeard, B. A. Schulman, J. W. Harper, Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Rep.* **14**, 1050–1061 (2013). doi: [10.1038/embor.2013.173](https://doi.org/10.1038/embor.2013.173); pmid: [24232186](https://pubmed.ncbi.nlm.nih.gov/24232186/)
  30. L. K. Langeberg, J. D. Scott, Signalling scaffolds and local organization of cellular behaviour. *Nat. Rev. Mol. Cell Biol.* **16**, 232–244 (2015). doi: [10.1038/nrm3966](https://doi.org/10.1038/nrm3966); pmid: [25785716](https://pubmed.ncbi.nlm.nih.gov/25785716/)
  31. H. W. Schroeder Jr., L. Cavacini, Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52 (2010). doi: [10.1016/j.jaci.2009.09.046](https://doi.org/10.1016/j.jaci.2009.09.046); pmid: [20176268](https://pubmed.ncbi.nlm.nih.gov/20176268/)
  32. P. Broz, V. M. Dixit, Inflammasomes: Mechanism of assembly, regulation and signalling. *Nat. Rev. Immunol.* **16**, 407–420 (2016). doi: [10.1038/nri.2016.58](https://doi.org/10.1038/nri.2016.58); pmid: [27291964](https://pubmed.ncbi.nlm.nih.gov/27291964/)
  33. L. Doyle *et al.*, Rational design of  $\alpha$ -helical tandem repeat proteins with closed architectures. *Nature* **528**, 585–588 (2015). doi: [10.1038/nature16191](https://doi.org/10.1038/nature16191); pmid: [26675735](https://pubmed.ncbi.nlm.nih.gov/26675735/)
  34. I. Vulovic *et al.*, Generation of ordered protein assemblies using rigid three-body fusion. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015037118 (2021). doi: [10.1073/pnas.2015037118](https://doi.org/10.1073/pnas.2015037118); pmid: [34074752](https://pubmed.ncbi.nlm.nih.gov/34074752/)
  35. M. D. Tyka *et al.*, Alternate states of proteins revealed by detailed energy landscape mapping. *J. Mol. Biol.* **405**, 607–618 (2011). doi: [10.1016/j.jmb.2010.11.008](https://doi.org/10.1016/j.jmb.2010.11.008); pmid: [21073878](https://pubmed.ncbi.nlm.nih.gov/21073878/)
  36. A. Chevalier *et al.*, Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017). doi: [10.1038/nature23912](https://doi.org/10.1038/nature23912); pmid: [28953867](https://pubmed.ncbi.nlm.nih.gov/28953867/)
  37. P. Hosseinzadeh *et al.*, Comprehensive computational design of ordered peptide macrocycles. *Science* **358**, 1461–1466 (2017). doi: [10.1126/science.aap7577](https://doi.org/10.1126/science.aap7577); pmid: [29242347](https://pubmed.ncbi.nlm.nih.gov/29242347/)
  38. B. Dang *et al.*, De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10852–10857 (2017). doi: [10.1073/pnas.1710695114](https://doi.org/10.1073/pnas.1710695114); pmid: [28973862](https://pubmed.ncbi.nlm.nih.gov/28973862/)
  39. S. J. Fleishman *et al.*, RosettaScripts: A scripting language interface to the Rosetta macromolecular modeling suite. *PLOS ONE* **6**, e20161 (2011). doi: [10.1371/journal.pone.0020161](https://doi.org/10.1371/journal.pone.0020161); pmid: [210731610](https://pubmed.ncbi.nlm.nih.gov/210731610/)
  40. G. Bhardwaj *et al.*, Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329–335 (2016). doi: [10.1038/nature19791](https://doi.org/10.1038/nature19791); pmid: [27626386](https://pubmed.ncbi.nlm.nih.gov/27626386/)
  41. R. F. Alford *et al.*, The Rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theory Comput.* **13**, 3031–3048 (2017). doi: [10.1021/acs.jctc.7b00125](https://doi.org/10.1021/acs.jctc.7b00125); pmid: [28430426](https://pubmed.ncbi.nlm.nih.gov/28430426/)
  42. S. Chaudhury, S. Lyskov, J. J. Gray, PyRosetta: A script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **26**, 689–691 (2010). doi: [10.1006/jmbi.1993.1648](https://doi.org/10.1006/jmbi.1993.1648); pmid: [8263940](https://pubmed.ncbi.nlm.nih.gov/8263940/)
  43. M. C. Lawrence, P. M. Colman, Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950 (1993). doi: [10.1006/jmbi.1993.1648](https://doi.org/10.1006/jmbi.1993.1648); pmid: [8263940](https://pubmed.ncbi.nlm.nih.gov/8263940/)
  44. B. Dang *et al.*, SNAC-tag for sequence-specific chemical protein cleavage. *Nat. Methods* **16**, 319–322 (2019). doi: [10.1038/s41592-019-0357-3](https://doi.org/10.1038/s41592-019-0357-3); pmid: [30923372](https://pubmed.ncbi.nlm.nih.gov/30923372/)
  45. Z. L. VanAernum *et al.*, Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nat. Protoc.* **15**, 1132–1157 (2020). doi: [10.1038/s41596-019-0281-0](https://doi.org/10.1038/s41596-019-0281-0); pmid: [32005983](https://pubmed.ncbi.nlm.nih.gov/32005983/)
  46. M. T. Marty *et al.*, Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015). doi: [10.1021/acs.analchem.5b00140](https://doi.org/10.1021/acs.analchem.5b00140); pmid: [25799115](https://pubmed.ncbi.nlm.nih.gov/25799115/)
  47. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007). doi: [10.1107/S0021889807021206](https://doi.org/10.1107/S0021889807021206); pmid: [19461840](https://pubmed.ncbi.nlm.nih.gov/19461840/)
  48. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010). doi: [10.1107/S0907444909047337](https://doi.org/10.1107/S0907444909047337); pmid: [20124692](https://pubmed.ncbi.nlm.nih.gov/20124692/)
  49. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997). doi: [10.1016/S0076-6879\(97\)76066-X](https://doi.org/10.1016/S0076-6879(97)76066-X)
  50. M. J. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011). doi: [10.1107/S0907444910045749](https://doi.org/10.1107/S0907444910045749); pmid: [21460441](https://pubmed.ncbi.nlm.nih.gov/21460441/)
  51. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi: [10.1107/S0907444909052925](https://doi.org/10.1107/S0907444909052925); pmid: [20124702](https://pubmed.ncbi.nlm.nih.gov/20124702/)
  52. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997). doi: [10.1107/S0907444996012255](https://doi.org/10.1107/S0907444996012255); pmid: [15299926](https://pubmed.ncbi.nlm.nih.gov/15299926/)
  53. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158); pmid: [15572765](https://pubmed.ncbi.nlm.nih.gov/15572765/)
  54. Y. Zhang, J. Skolnick, TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005). doi: [10.1093/nar/gki524](https://doi.org/10.1093/nar/gki524); pmid: [15849316](https://pubmed.ncbi.nlm.nih.gov/15849316/)
  55. V. Mariani, M. Biasini, A. Barbato, T. Schwede, ICDT: A local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728 (2013). doi: [10.1093/bioinformatics/btt473](https://doi.org/10.1093/bioinformatics/btt473); pmid: [23986568](https://pubmed.ncbi.nlm.nih.gov/23986568/)
  56. B. L. Nannenga, M. G. Iadanza, B. S. Vollmar, T. Gonen, Overview of electron crystallography of membrane proteins: crystallization and screening strategies using negative stain electron microscopy. *Curr. Protoc. Protein Sci.* **72**, 17.15.1–17.15.11 (2013). doi: [10.1093/bioinformatics/btt473](https://doi.org/10.1093/bioinformatics/btt473); pmid: [23986568](https://pubmed.ncbi.nlm.nih.gov/23986568/)
  57. T. Grant, A. Rohou, N. Grigorieff, cisTEM, user-friendly software for single-particle image processing. *eLife* **7**, e35383 (2018). doi: [10.7554/eLife.35383](https://doi.org/10.7554/eLife.35383); pmid: [29513216](https://pubmed.ncbi.nlm.nih.gov/29513216/)
  58. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi: [10.1038/nmeth.4169](https://doi.org/10.1038/nmeth.4169); pmid: [28165473](https://pubmed.ncbi.nlm.nih.gov/28165473/)
  59. D. D. Sahtoe, F. Praetorius, A. Courbet, Y. Hsia, N. I. Edman, D. Baker, Data for: Reconfigurable asymmetric protein assemblies through implicit negative design, Version 1, Zenodo (2021). doi: [10.5281/zenodo.5717329](https://doi.org/10.5281/zenodo.5717329)

## ACKNOWLEDGMENTS

We acknowledge Baker lab members for discussion; B. Sankaran and beamline scientists at the Advanced Light Source for crystallographic data collection and support; F. Busch, A. Norris, and the Wysocki lab for native mass spectrometry measurements; R. Mout for C5-LHD101B construct sequences; L. Carter for SEC-MALS analysis; M. Ahlrichs, C. Ogohara, and M. Murphy for mammalian cell transfections; C. Miller for discussions on mammalian cell assays; and T. Sixma and H. Dietz for critical reading of the manuscript. **Funding:** This work was funded by EMBO long term fellowship ALTF 1295-2015 and ALTF 139-2018 (D.D.S. and B.I.M.W.); Washington Research Foundation Innovation Fellowship (D.D.S.); Human Frontiers Science Program long term fellowship (F.P. and A.C.); DARPA Biostasis HR00118S0034 (Y.H. and D.B.); Open Philanthropy Project Improving Protein Design Fund (D.B., F.P., A.K.B., and A.K.); The Audacious Project at the Institute for Protein Design (L.M.M., H.M.M., B.J.R.T., and N.I.E.); Eric and Wendy Schmidt by recommendation of the Schmidt Futures (D.B. and Y.H.); The Howard Hughes Medical Research Institute (D.B. and A.C.); and NIH Resource for Native Mass Spectrometry Guided Structural Biology P41 GM128577 (V. Wysocki, Ohio State University). **Author contributions:** Conceptualization: D.D.S., F.P., D.B.; Methodology: D.D.S., F.P.; Investigation: D.D.S., F.P., A.C., Y.H., B.I.M.W., N.I.E., L.M.M., H.M.M., B.J.R.T., J.D., A.K.B.; Visualization: D.D.S., F.P.; Funding acquisition: D.D.S., F.P., A.C., B.I.M.W., D.B.; Supervision: D.B.; Writing – original draft: D.D.S., F.P., D.B.; Writing – review and editing: D.D.S., F.P., D.B. D.D.S. and F.P. developed the hetero-oligomer design pipeline, performed design calculations and experiments, and analyzed all data. A.C. designed and characterized the homo-oligomeric C3 hub. Y.H. designed and characterized the two component C4 ring. B.I.M.W. performed and analyzed split luciferase binding assays. N.I.E. designed and characterized the homo-oligomeric C4 hub. A.C., Y.H., and N.I.E. performed nsEM and 3D reconstructions. B.J.R.T. designed scaffolds. L.M.M. and H.M.M. purified designs. D.D.S. and J.D. analyzed mammalian cell-based assays. D.D.S., A.B. and A.K. determined crystal structures. D.B. supervised research. **Competing interests:** D.D.S., F.P., A.C., N.I.E., Y.H., B.J.R.T., and D.B. are inventors on a provisional patent application submitted by the University of Washington for the design, composition, and function of the proteins created in this study. **Data and materials availability:** Crystallographic models have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (accession codes 6wmk, 7mwq, and 7mwr). All data are available in the main text or the supplementary materials. Design scripts, protein sequences, design models, and models of assemblies are also available through Zenodo (59).

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abj7662](https://science.org/doi/10.1126/science.abj7662)

Figs. S1 to S25

Tables S1 to S5

Reference (60)

MDAR Reproducibility Checklist

Data S1 and S2

31 May 2021; accepted 13 December 2021  
10.1126/science.abj7662



## RESEARCH ARTICLE SUMMARY

## SYNTHETIC BIOLOGY

## Synthetic multistability in mammalian cells

Ronghui Zhu, Jesus M. del Rio-Salgado, Jordi Garcia-Ojalvo, Michael B. Elowitz\*

**INTRODUCTION:** Multistability allows genetically identical cells to exist in thousands of molecularly distinct and mitotically stable states. Building synthetic multistable circuits could provide insight into the minimal circuitry sufficient for multistability and establish a foundation for exploiting multicellularity in engineered cell therapies. However, efforts in mammalian cells have been limited to two-state systems or have used architectures that cannot be easily expanded. Beyond generating long-term multistability, an ideal synthetic architecture would also recapitulate key properties of natural cell fate control systems, including the ability to switch cells among states with transient external inputs, control the stability of particular states, and generate irreversible state transitions. Despite much work on identifying key genes and regulatory interactions in many natural cell fate control systems, it has remained unclear what circuit architectures could provide these capabilities.

**RATIONALE:** Natural cell fate control systems exhibit two prevalent features: positive auto-regulation and combinatorial protein-protein interactions. We designed a minimal circuit architecture based on similar principles, called MultiFate, in which a set of transcription factors competitively homo- and heterodimerize, with only the homodimers activating the expression of their own gene. Mathematical modeling showed that MultiFate can produce

diverse types of multistability, support controlled state switching, and enable irreversible state transitions. Critically, the use of heterodimerization to implement cross-inhibition allows the expansion of MultiFate to larger numbers of states simply by adding new transcription factors, without the need to reengineer existing components. These properties suggest that MultiFate could provide an ideal synthetic architecture for multistability.

**RESULTS:** To create MultiFate circuits, we first engineered a set of zinc finger transcription factors that enable homodimer-dependent self-activation and heterodimer-dependent inhibition. We then constructed a minimal circuit termed MultiFate-2 (comprising two of these factors), stably integrated it into CHO-K1 cells, and obtained several monoclonal MultiFate-2 cell lines. Flow cytometry and time-lapse imaging showed that MultiFate-2 cells could exist in three distinct expression states, expressing predominantly one factor, the other, or both. Each of these states was stable for extended time scales of weeks or more. Using external inducers, we were able to switch cells among states. Finally, consistent with model predictions, reducing protein stability resulted in a tristable-to-bistable bifurcation, selectively destabilizing the state expressing both factors while preserving states expressing single factors. Cells exiting the destabilized state did not return even when protein stability was restored, recapitulating

irreversible state transitions observed in many natural fate control systems.

To test the expandability of the MultiFate design, we integrated a third transcription factor into a MultiFate-2 cell line. As predicted by the model, the resulting MultiFate-3 cells could stably exist in seven distinct states for more than 18 days. Progressively reducing protein stability repeatedly bifurcated the system from septastability through hexastability to tristability, further recapitulating the progressive loss of cell fate potential in natural cell differentiation systems. Modeling indicates that the MultiFate system should be expandable beyond three transcription factors to generate hundreds of robust stable states.

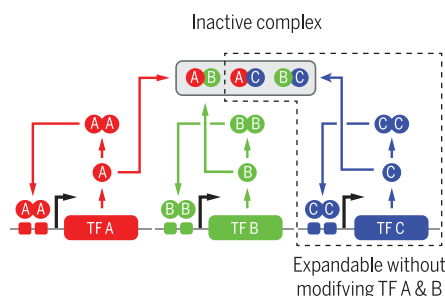
**CONCLUSION:** Recently, single-cell transcriptomic approaches have revealed a stunning diversity of natural cellular states, making the question of how such multistability is generated and controlled more urgent than ever. MultiFate demonstrates how a relatively simple, naturally inspired architecture can produce several hallmarks of natural multistability: They generate long-term multistability through combinations of transcription factors; they allow controlled state switching using external inducers; and they permit modulation of state stability, which allows hierarchical and irreversible cellular transitions. Because MultiFate can be readily expanded to generate more states by adding new transcription factors, it provides a scalable foundation for exploring circuit-level principles of multistability and enables multicellular applications in synthetic biology. ■

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

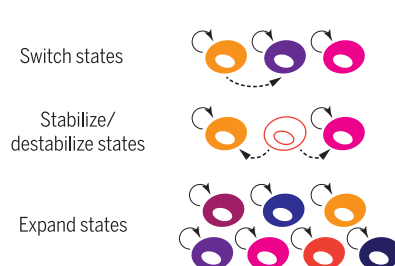
\*Corresponding author. Email: melowitz@caltech.edu  
Cite this article as R. Zhu *et al.*, *Science* 375, eabg9765 (2022). DOI: 10.1126/science.abg9765

**READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abg9765>

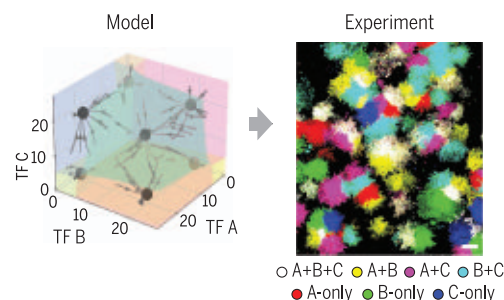
## MultiFate: a synthetic multistable circuit architecture



## MultiFate supports long-term, controllable, and expandable multistability



## Septastability: model and experiment



**MultiFate supports long-term, controllable, and expandable multistability.** Left: In MultiFate, transcription factors (TFs) homodimerize to self-activate and mutually inhibit one another through heterodimerization. Cross-inhibition through heterodimerization allows circuit expansion by adding additional transcription factors without modifying existing components. Center: This circuit design supports state switching, modulation of state stability, and expansion of states. Right: A MultiFate-3 circuit with three transcription factors generates seven stable states (attractors in phase diagram). Experiments show that these seven states, indicated by distinct transcription factor combinations (colors), are stably maintained as cells grow into colonies.

## RESEARCH ARTICLE

## SYNTHETIC BIOLOGY

## Synthetic multistability in mammalian cells

Ronghui Zhu<sup>1</sup>, Jesus M. del Rio-Salgado<sup>1</sup>, Jordi Garcia-Ojalvo<sup>2</sup>, Michael B. Elowitz<sup>1,3\*</sup>

In multicellular organisms, gene regulatory circuits generate thousands of molecularly distinct, mitotically heritable states through the property of multistability. Designing synthetic multistable circuits would provide insight into natural cell fate control circuit architectures and would allow engineering of multicellular programs that require interactions among distinct cell types. We created MultiFate, a naturally inspired, synthetic circuit that supports long-term, controllable, and expandable multistability in mammalian cells. MultiFate uses engineered zinc finger transcription factors that transcriptionally self-activate as homodimers and mutually inhibit one another through heterodimerization. Using a model-based design, we engineered MultiFate circuits that generate as many as seven states, each stable for at least 18 days. MultiFate permits controlled state switching and modulation of state stability through external inputs and can be expanded with additional transcription factors. These results provide a foundation for engineering multicellular behaviors in mammalian cells.

**M**ultistability allows genetically identical cells to exist in thousands of molecularly distinct and mitotically stable cell types or states (1, 2). Understanding natural multistable circuits and engineering synthetic ones have been long-standing challenges in developmental and synthetic biology (3–14). Building synthetic multistable circuits could provide insight into the minimal circuitry sufficient for multistability and would establish a foundation for exploiting multicellularity in engineered cell therapies. However, efforts in mammalian cells have been limited to two-state systems or have used architectures that cannot be easily expanded to larger numbers of states (5–7). An ideal synthetic multistable system would allow cells to remain in any of a set of distinct expression states over many cell cycles, despite biological noise. In addition, it would provide three key capabilities exhibited by its natural counterparts (Fig. 1A): (i) It would permit transient external inputs to switch cells between states, similar to the way in which signaling pathways direct fate decisions (15, 16). (ii) It would support control over the stability of different states and would enable irreversible transitions, similar to those that occur during natural differentiation (13, 14). (iii) It would be expandable by introducing additional components without reengineering an existing functional circuit, analogous to the expansion of cell types during evolution (17).

Natural mammalian multistable circuits provide inspiration for such a synthetic architecture.

In many natural fate control systems, transcription factors positively autoregulate their own expression and competitively interact with one another to form a variety of homodimers, heterodimers, and higher-order multimeric forms (Fig. 1B) (18–24). For example, during myogenesis, muscle regulatory factors such as MyoD heterodimerize with E proteins to activate their own expression and the broader myogenesis program, while Id family proteins disrupt this process through competitive dimerization (23, 24). Similarly, during embryogenesis, Sox2 and Sox17 competitively interact with Oct4 to control fate decisions between pluripotency and endodermal differentiation (21, 22). Related combinations of positive autoregulation and cross-inhibition could extend multistability behaviors beyond bistability and generate bifurcation dynamics that explain the partial irreversibility of cell differentiation (9, 12). Nonetheless, it remains unclear whether these natural architectures could be adapted to enable synthetic multistability. Here, we show how a synthetic multistable system based on principles derived from natural cell fate control systems can generate robust, controllable, expandable multistability in mammalian cells.

#### MultiFate generates diverse types of multistability through a set of promiscuously dimerizing, autoregulatory transcription factors

Inspired by natural fate control circuits, we designed a new synthetic multistable system called MultiFate. In MultiFate, transcription factors share a common dimerization domain, allowing them to competitively form both homodimers and heterodimers. The promoter of each transcription factor gene contains binding sites that can be strongly bound only by their own homodimers, allowing homodimer-dependent self-activation. By

contrast, heterodimers do not efficiently bind to any promoter in this design. Heterodimerization thus acts to mutually inhibit the activity of both constituent transcription factors.

Mathematical modeling shows how the MultiFate architecture provides each of the desired capabilities described above (Fig. 1A) in physiologically reasonable parameter regimes (Box 1 and table S1) (25). A MultiFate circuit with just two transcription factors, designated MultiFate-2, can produce diverse types of multistability containing two, three, or four stable fixed points, depending on protein stability and other parameter values (Fig. 1C and fig. S1A). In particular, a regime designated type II tristability is analogous to multilineage priming in uncommitted progenitor cells, with the double positive state playing the role of a multipotent progenitor (26–28). Transient expression of one transcription factor can switch cells between states (fig. S3 and movie S1). Reducing the protein stability of transcription factors can cause bifurcations that selectively destabilize certain states (Fig. 1C and fig. S1A). Finally, the model is expandable: Addition of a new transcription factor to the MultiFate-2 model generates a MultiFate-3 circuit that supports additional stable states with the same parameter values (Fig. 1D and fig. S2A). Together, these modeling results suggest that the MultiFate architecture can support a rich array of multistable behaviors.

#### Engineered zinc finger transcription factors enable homodimer-dependent self-activation and heterodimer-dependent inhibition

Synthetic zinc finger (ZF) transcription factors provide an ideal platform to implement the MultiFate circuit. They can recognize and activate a promoter containing target DNA binding sites with high specificity (29, 30). Further, engineered ZF DNA binding domains containing three fingers bind weakly as monomers to 9–base pair (bp) target sites but can bind much more strongly as homodimers to 18-bp tandem binding-site pairs (31, 32). This property allows homodimer-dependent transcriptional activity and potentially allows inhibition through heterodimerization.

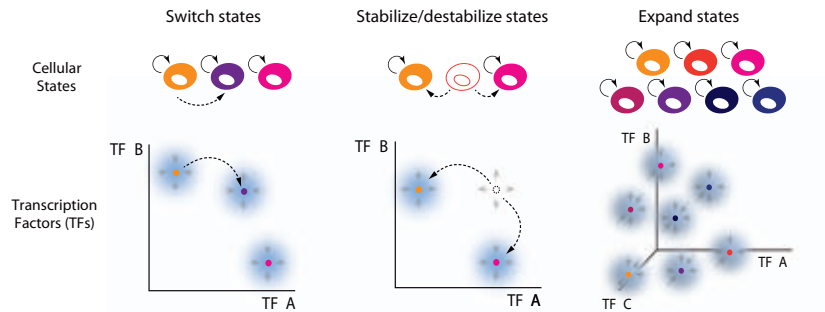
To engineer ZF transcription factors, we fused the Erbb2 ZF DNA binding domain to a GCN4 homodimerization domain and a VP48 transcriptional activation domain to create the synthetic transcription factor, termed ZF-GCN4-AD (Fig. 2A) (31). A transcription factor (ZF-AD) lacking GCN4 was used as a monomeric control. To assay their transcriptional activity, we constructed a reporter containing 18-bp homodimer binding sites driving the expression of Citrine (31). We then cotransfected each transcription factor, together with the reporter and an mTagBFP2 (33) cotransfection

<sup>1</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

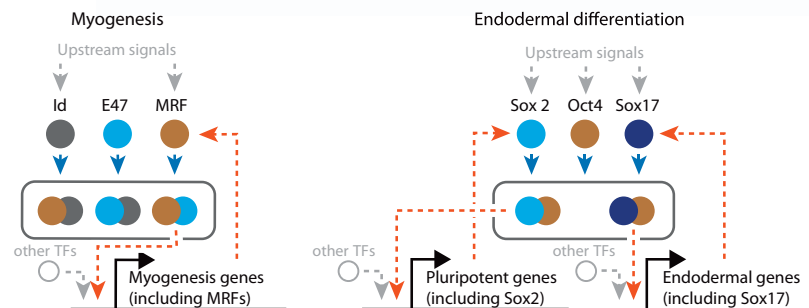
<sup>2</sup>Department of Experimental and Health Sciences, Universitat Pompeu Fabra, 08003 Barcelona, Spain. <sup>3</sup>Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA.

\*Corresponding author. Email: melowitz@caltech.edu

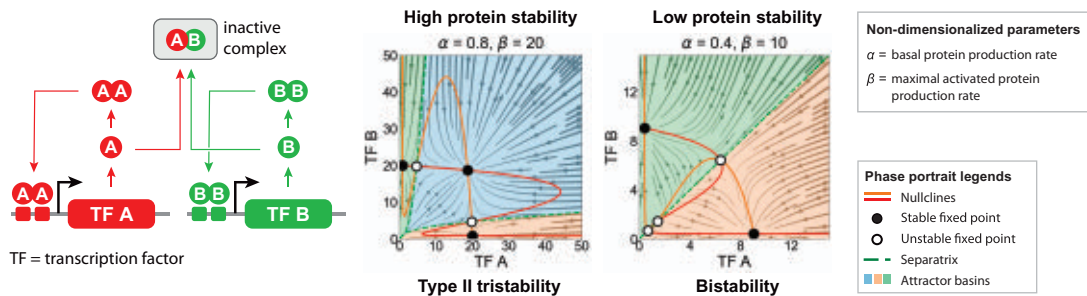
### A Ideal capabilities for a synthetic multistable circuit



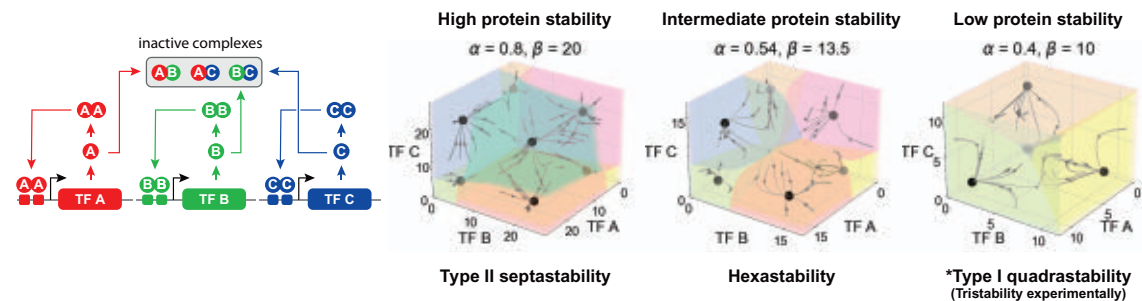
### B Natural multistable circuits use dimerization and autoregulation



### C MultiFate-2 circuit



### D MultiFate-3 circuit

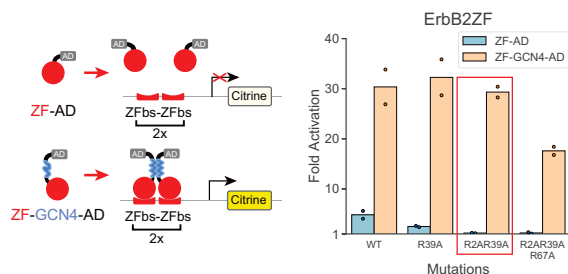
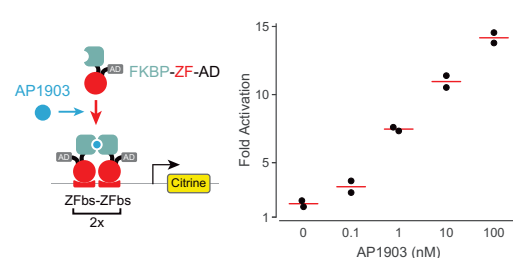
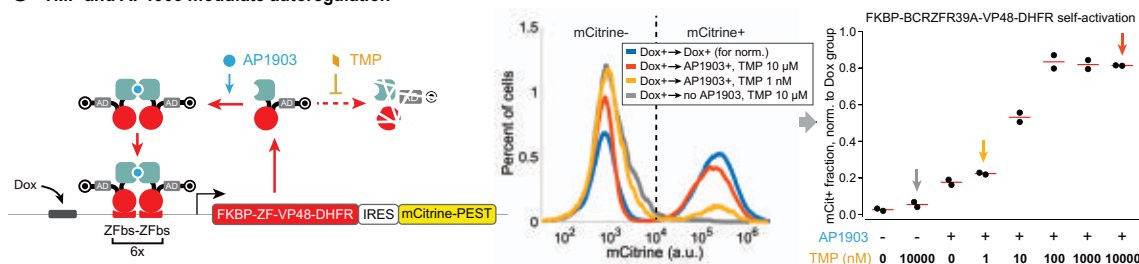
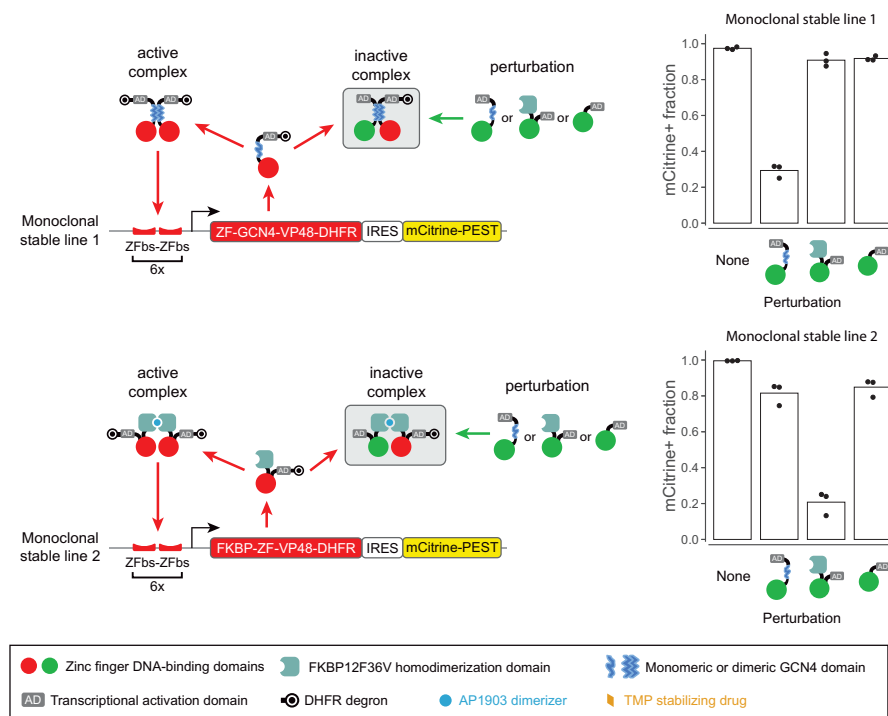


**Fig. 1. The naturally inspired MultiFate architecture generates diverse types of multistability in the model. (A)** A hypothetical synthetic multistable circuit is represented by colored cell cartoons (top) and attractors in a transcription factor phase space (bottom; axes represent transcription factor concentrations of TF A, TF B, and TF C). An ideal synthetic multistable circuit should generate multiple stable states, support control of state switching (left) and state stability (center), and allow easy expansion of states by addition of more transcription factors (right).

**(B)** Competitive protein-protein interactions and autoregulatory feedback are prevalent in natural multistable circuits that control myogenesis (left) and endodermal differentiation (right), as shown by these simplified and abridged diagrams. Blue arrows indicate competitive protein-protein interactions, which can involve higher-order multimerization. Orange dashed arrows indicate direct or

indirect positive transcriptional feedback. **(C and D)** Models of the MultiFate-2 circuit and MultiFate-3 circuit (Box 1) (25) generate diverse types of multistability in different parameter regimes (indicated above plots). In the model of the MultiFate-3 circuit, low protein stability generates four stable states (type I quadrastability), but the state in which all transcription factors are minimally expressed is unstable in the presence of biological noise (fig. S22), consistent with experimental results in Fig. 5B, low-TMP columns. See figs. S1 and S2 for complete lists of multistability regimes. All models used here are symmetric and nondimensionalized, with rescaled dimerization dissociation constant  $K_d = 1$  and Hill coefficient  $n = 1.5$  (Box 1). In both (C) and (D), each axis represents the dimensionless total concentration of each transcription factor. Note that in the nondimensionalized model, changing protein stability is equivalent to multiplying  $\alpha$  and  $\beta$  with the same factor (Box 1).

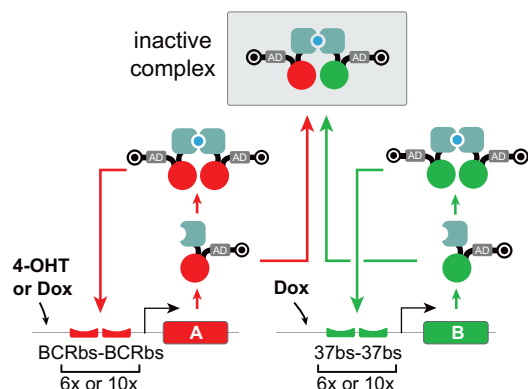


**A Engineered TFs exhibit dimer-dependent activity****B FKBP12F36V allows inducible dimerization****C TMP and AP1903 modulate autoregulation****D Competitive dimerization allows inhibition of positive autoregulation**

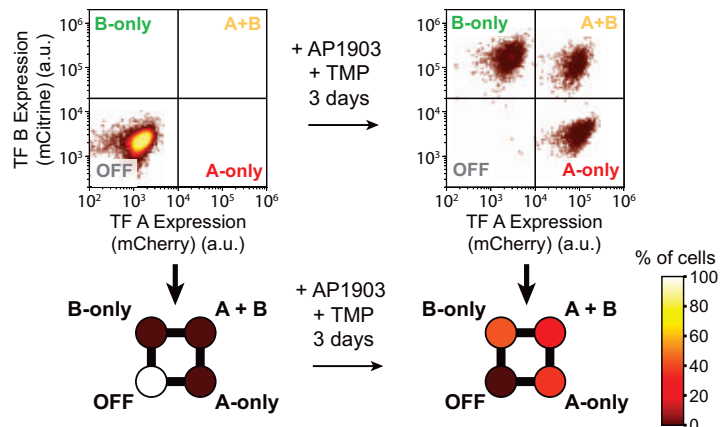
**Fig. 2. Engineered transcription factors enable homodimer-dependent auto-regulation and heterodimerization-based inhibition.** (A) ZF transcription factors enable homodimer-dependent activation. Left: Design of test constructs in which ErbB2ZF (31) (red circle), fused to VP48 (AD) and in some cases GCN4 (blue squiggle) domains, binds to target sites (red pads) to activate Citrine expression. Activators were expressed from a constitutive CAG promoter (52). Right: R-to-A mutations in ZF-modulated reporter activation by ZF-GCN4-AD and ZF-AD. The R2AR39A variant was selected because of high ZF-GCN4-AD activation and minimal ZF-AD activation. Fold activation is defined in fig. S4A. WT, wild-type variant. (B) Left: Design of test constructs in which FKBP12F36V (FKBP, light cyan partial box), fused to BCRZFR39A (red circle) and VP48 (AD), binds to target sites (red pads) to activate Citrine expression in the presence of AP1903 (cyan circle). Right: This design allows dose-dependent control of activation by AP1903. (C) Transcription factor self-activation can be controlled by TMP and AP1903.

Left: Design of the controllable self-activation circuit. IRES, internal ribosome entry site; PEST, constitutive degradation tag (53). Center: Stable polyclonal cells showed bimodal mCitrine distribution upon circuit activation. An empirical threshold at mCitrine =  $10^4$  separates the distribution into two fractions, and the normalized mCitrine<sup>+</sup> fraction was used to quantify the self-activation strength (25). Right: Colored arrows indicate data from the center panel. AP1903<sup>+</sup> samples had 100 nM AP1903. (D) Self-activation was inhibited by proteins with a different ZF and matching dimerization domains. Two monoclonal stable lines could spontaneously self-activate in media containing AP1903 and TMP (fig. S5B). Each perturbation construct is introduced by stable integration (25). The integrated construct in the “None” group did not express any perturbation protein. Red circle denotes 42ZFR2AR39AR67A; green circle denotes BCRZFR39A. In all panels, each dot represents one biological replicate, and each red line or bar indicates the mean of replicates. Lists of constructs and cell lines are in tables S2 and S3.

## A Experimental MultiFate-2 design

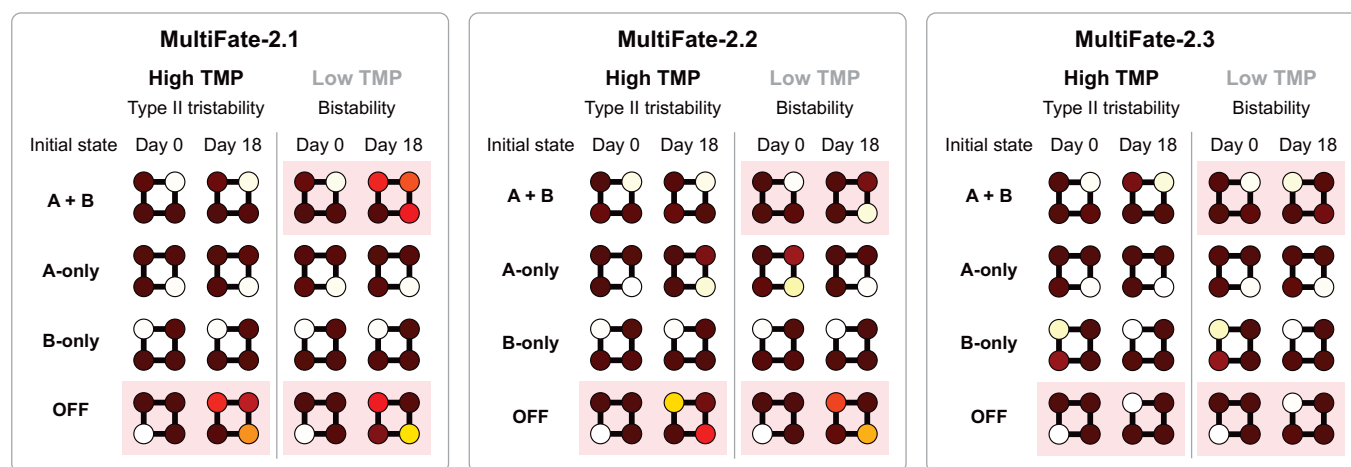


## B

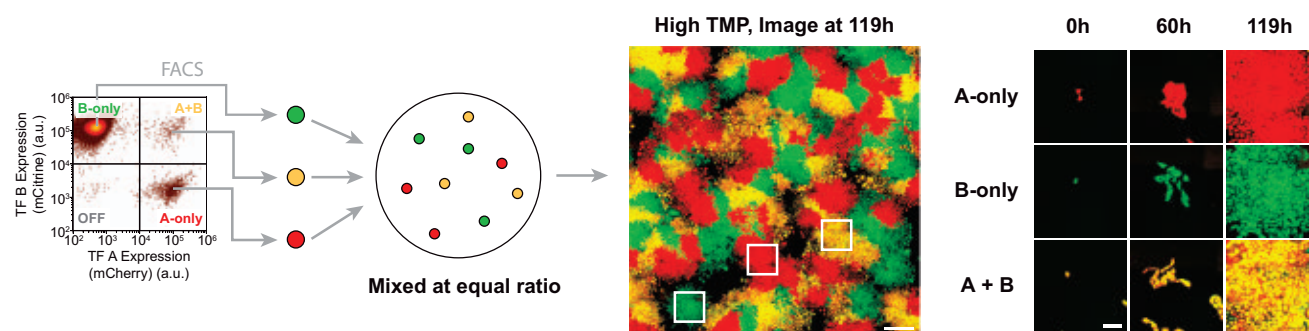


## C MultiFate-2 cell lines exhibit bistability and tristability

Unstable states - At least 10% cells escaped initial states



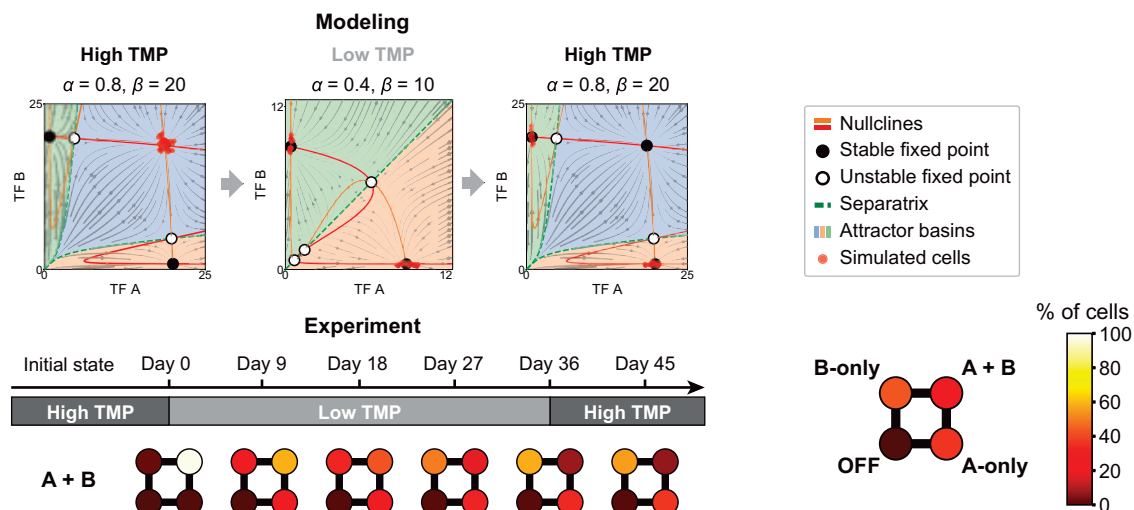
## D Time-lapse imaging reveals MultiFate-2.3 tristability



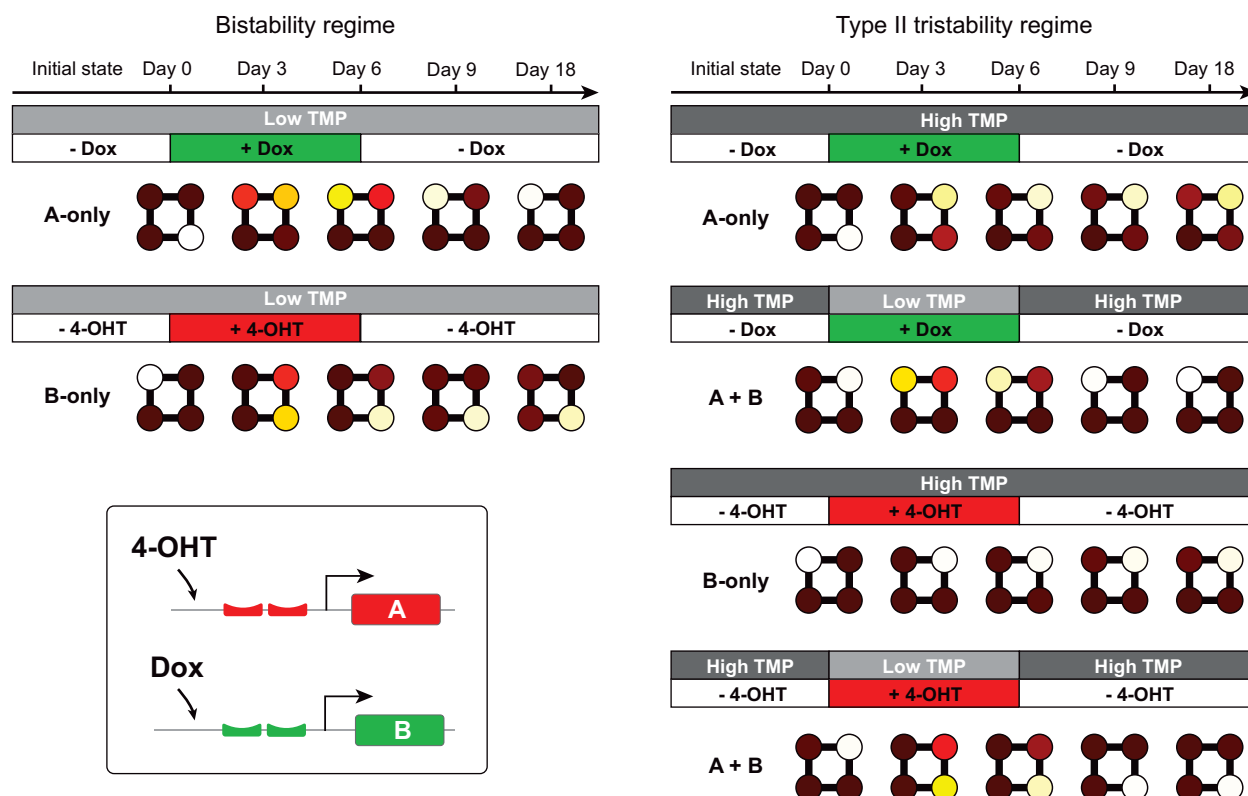
**Fig. 3. MultiFate-2 generates multiple stable states.** (A) The experimental MultiFate-2 design uses two self-activation cassettes differing only in their fluorescent proteins and their ZF DNA binding domains and binding sites. Each cassette expresses FKBP-ZF-VP16-DHFR-IRES-FP-PEST, where ZF represents either BCRZFR39A or 37ZFR2AR11AR39AR67A and FP represents either mCherry or mCitrine, for A and B, respectively. Detailed construct maps and differences among MultiFate-2 lines are shown in tables S2 and S3. (B) MultiFate-2.1 cells spontaneously activate A, B, or both cassettes upon addition of 100 nM AP1903 and 10  $\mu$ M TMP. Cell percentages in OFF, A-only, B-only, and A+B states were quantified and plotted as a square with four colored circles (31); a.u., arbitrary units. (C) Three MultiFate-2 lines all exhibited type II tristability in the high-TMP condition and bistability in the low-TMP condition. In all conditions, we added

100 nM AP1903. Exact concentrations of TMP are shown in figs. S9 to S11. Unstable states, defined by states having more than 10% cells escaping their initial states after 18 days, were marked in pink rectangles. Each square represents the mean fractions of three biological replicates. Initial A-only, B-only, and A+B cells were sorted from a population of cells in different states, while initial OFF cells came from cells in regular CHO media without any inducers. (D) A-only, B-only, and A+B states were each stable during growth from single MultiFate-2.3 cells into colonies over 5 days (119 hours) under a time-lapse microscope. Left: We first sorted mixed MultiFate-2.3 cell populations to separate cells in three different states. Then we seeded cells in these three states at equal ratio in the same well and performed time-lapse imaging (25). Scale bars, 500  $\mu$ m for the wide-field image, 100  $\mu$ m for zoomed-in images. "High TMP" = 100 nM AP1903 + 10  $\mu$ M TMP.

## A Modulating TF stability destabilizes specific states and allows irreversible transitions



## B External inducers switch cells between states



**Fig. 4. MultiFate-2 supports modulation of state stability and allows state switching. (A)** Escape from the destabilized A+B state was irreversible, as shown by both modeling and experiments that used MultiFate-2.1 cells. Top: The model used here is symmetric and nondimensionalized, with rescaled dimerization dissociation constant  $K_d = 1$  and Hill coefficient  $n = 1.5$  (Box 1). The  $x$  and  $y$  axes are total dimensionless concentrations of TF A and TF B, respectively. Simulated cells on phase portraits were calculated using the Gillespie algorithm (25, 54). Note that in the nondimensionalized model, changing protein

stability is equivalent to multiplying  $\alpha$  and  $\beta$  with the same factor (Box 1). Bottom: Throughout the experiment, we added 100 nM AP1903. Exact concentrations of TMP are shown in fig. S9. **(B)** MultiFate-2.3 cells can be switched between states by transient 4-OHT (25 nM) or Dox (500 ng/ml) treatment. In all conditions, we added 100 nM AP1903. Exact concentrations of TMP are shown in fig. S16. In all panels, initial A-only, B-only, and A+B cells were sorted from a population of cells in different states. Each square represents the mean fractions of three biological replicates.



### Box 1. Design of the MultiFate circuit.

Here we introduce the mathematical model of the MultiFate circuit and show how it can be used to design the experimental system and predict its behavior. For simplicity, we focus on a symmetric MultiFate-2 circuit whose two transcription factors share identical biochemical parameters and differ only in their DNA binding site specificity. A similar analysis of systems with more transcription factors and asymmetric parameters is presented in (25).

We represent the dynamics of protein production and degradation using ordinary differential equations (ODEs) for the total concentrations of the transcription factors A and B, denoted  $[A_{\text{tot}}]$  and  $[B_{\text{tot}}]$ , respectively. We assume that the rate of production of each protein follows a Hill function of the corresponding homodimer concentration,  $[A_2]$  or  $[B_2]$ , with maximal rate  $\beta$ , Hill coefficient  $n$ , and half-maximal activation at a homodimer concentration of  $K_M$ . A low basal protein production rate, denoted  $\alpha$ , is included to allow self-activation from low initial expression states. Finally, each protein can degrade and be diluted (as a result of cell division) at a total rate  $\delta$ , regardless of its dimerization state. To simplify analysis, we nondimensionalize the model by rescaling time in units of  $\delta^{-1}$  and rescaling concentrations in units of  $K_M$  (25), and obtain

$$\begin{aligned}\frac{d[A_{\text{tot}}]}{dt} &= \alpha + \frac{\beta[A_2]^n}{1+[A_2]^n} - [A_{\text{tot}}] \\ \frac{d[B_{\text{tot}}]}{dt} &= \alpha + \frac{\beta[B_2]^n}{1+[B_2]^n} - [B_{\text{tot}}]\end{aligned}$$

Here, Hill coefficient  $n$  only represents ultrasensitivity introduced by transcriptional activation. See (25) for a more detailed discussion on additional ultrasensitivity provided by homodimerization and molecular titration.

Because dimerization dynamics occur on a faster time scale than protein production and degradation (49), we assume that the distribution of monomer and dimer states remains close to their equilibrium values. This generates the following relationships between the concentrations of monomers,  $[A]$  and  $[B]$ , and dimers,  $[A_2]$ ,  $[B_2]$ , and  $[AB]$ :

$$\begin{aligned}[A]^2 &= K_d[A_2] \\ [B]^2 &= K_d[B_2] \\ 2[A][B] &= K_d[AB]\end{aligned}$$

Because the two transcription factors share the same dimerization domain, homo- and heterodimerization are assumed to occur with equal dissociation constants  $K_d$ . Additionally, conservation of mass implies that  $[A_{\text{tot}}] = [A] + [AB] + 2[A_2]$ , with a similar relationship for B. Introducing the equilibrium equations given above into this conservation law produces expressions for the concentrations of the activating homodimers in terms of the total concentrations of A and B:

$$\begin{aligned}[A_2] &= \frac{2[A_{\text{tot}}]^2}{K_d + 4([A_{\text{tot}}] + [B_{\text{tot}}]) + \sqrt{K_d^2 + 8([A_{\text{tot}}] + [B_{\text{tot}}])K_d}} \\ [B_2] &= \frac{2[B_{\text{tot}}]^2}{K_d + 4([A_{\text{tot}}] + [B_{\text{tot}}]) + \sqrt{K_d^2 + 8([A_{\text{tot}}] + [B_{\text{tot}}])K_d}}\end{aligned}$$

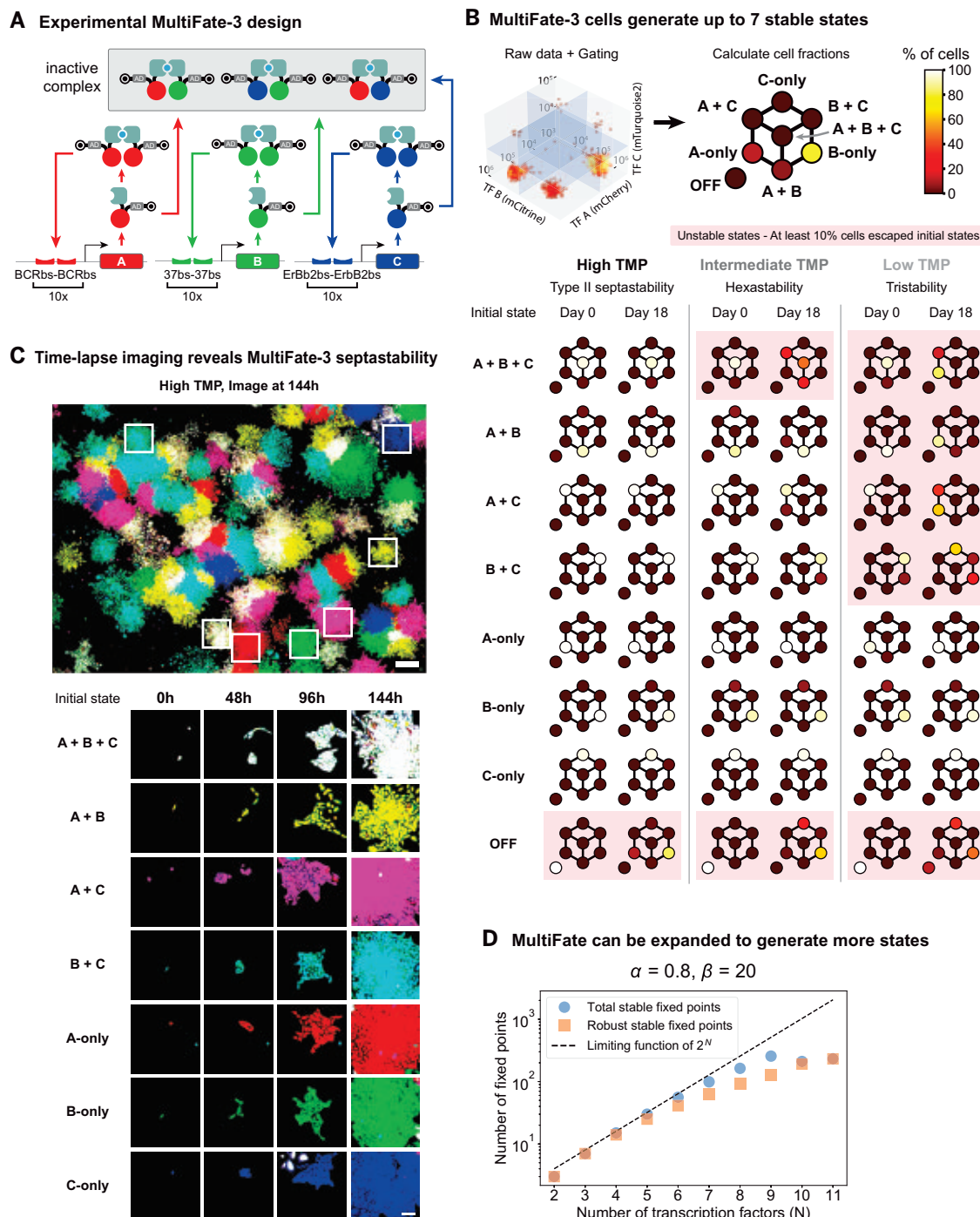
Inserting these expressions into the differential equations for  $[A_{\text{tot}}]$  and  $[B_{\text{tot}}]$  above, we obtain a pair of coupled ODEs with only  $[A_{\text{tot}}]$  and  $[B_{\text{tot}}]$  as variables.

To understand the behavior of this system in physiologically reasonable parameter regimes (table S1), we used standard approaches from dynamical systems analysis (25, 50). We first generated a phase portrait of variables  $[A_{\text{tot}}]$  and  $[B_{\text{tot}}]$  based on ODEs (labeled “TF A” and “TF B,” which are dimensionless total TF A or B concentrations), where the linewidth of a vector (Fig. 1C, gray arrows) at any point is proportional to the speed of that point. On the phase portrait, we plotted the nullclines (Fig. 1C, solid lines), defined by setting each of the ODEs above to zero. We then identified fixed points at nullcline intersections and determined their linear stability (Fig. 1C, black and white dots) (50). Finally, we delineated the basins of attraction for each stable fixed point (Fig. 1C, shaded regions).

Using this analysis, we identified parameter values that support type II tristability, a regime that minimally embodies the developmental concept of multilineage priming (26–28) (Fig. 1C and fig. S1B). Stronger self-activation (higher values of  $\beta$ ) was more likely to produce type II tristability (fig. S1B,  $\beta$  row and column). Too much leaky production (high  $\alpha$ ) allowed both transcription factors to self-activate, reducing the degree of multistability, whereas too little (low  $\alpha$ ) stabilized the undesired OFF state (fig. S1B,  $\alpha$  column). Strong dimerization (low  $K_d$ ) was essential for type II tristability (fig. S1B,  $K_d$  row and column). Finally, a broad range of Hill coefficients  $n \geq 1$  were compatible with type II tristability. Although higher values of  $n$  led to a reduced sensitivity to other parameters and allowed the system to tolerate higher values of  $\alpha$ , they also stabilized the OFF state (fig. S1B,  $n$  row and column). Together, these results suggested that an ideal design would maximize  $\beta$ , minimize  $K_d$ , and use intermediate values of  $\alpha$  and  $n$ .

On the basis of these conclusions, we incorporated multiple repeats of the homodimeric binding sites to maximize  $\beta$ , used strongly associating FKBP12F36V homodimerization domains (36) to minimize  $K_d$ , and modified the promoter sequences to allow some leaky expression to optimize  $\alpha$  (fig. S24) (25). Finally, although we did not directly control  $n$ , we expected that the repeated homodimeric binding sites should lead to some ultrasensitivity (51). These design choices produced the selected type II tristability in the experimental system (Fig. 3C).

A key feature of the MultiFate design is its ability to qualitatively change its multistability properties through bifurcations in response to parameter changes. In particular, the mathematical model predicts that protein stability can control the number of stable fixed points in phase space. In the nondimensionalized model, the protein degradation rate,  $\delta$ , does not appear explicitly but enters through the rescaling of  $\alpha$  and  $\beta$  by  $(\delta K_M)^{-1}$  (25). Thus, tuning protein stability is equivalent to multiplying both  $\alpha$  and  $\beta$  by a common factor, which we term the “protein stability factor.” Reducing protein stability shifts the nullclines closer to the origin, causing the two unstable fixed points to collide with the stable A+B fixed point in a subcritical pitchfork bifurcation (Fig. 1C) (50). The result is a bistable system with A-only and B-only stable fixed points at somewhat lower concentrations (Fig. 1C). To experimentally realize this bifurcation, we designed the circuit to allow external control of transcription factor protein stability using the drug-inducible DHFR degenon (Fig. 2C) (37). As predicted, reducing protein stability destabilized the A+B state but preserved the A-only and B-only stable states (Fig. 3C). In this way, model-based design enabled us to rationally engineer tristability as well as externally controllable transitions to bistability in the experimental system.



**Fig. 5. MultiFate architecture is expandable to include three and potentially even more transcription factors.** (A) The experimental MultiFate-3 design uses three self-activation cassettes differing only in their fluorescent proteins and their ZF DNA binding domains and binding sites. Each cassette expresses FKBP-ZF-VP16-DHFR-IRES-FP-PEST, where ZF represents either BCRZFR39A, 37ZFR2AR11-AR39AR67A, or ErbB2ZFR2AR39A and FP represents either mCherry, mCitrine, or mTurquoise2 for A, B, and C, respectively. See table S2 for detailed construct maps. (B) The MultiFate-3 line exhibited type II septastability, hexastability, and tristability in three different TMP conditions. Top: State percentages in each octant were quantified and plotted as eight colored circles (25). Bottom: High-TMP condition = 100 nM AP1903 + 100 nM TMP; intermediate-TMP condition = 100 nM AP1903 + 40 nM TMP; low-TMP condition = 100 nM AP1903 + 10 nM TMP. Except for OFF-state cells, cells in different initial states were sorted from a mixed

population of cells in the high-TMP condition. Initial OFF cells came from cells in regular CHO media without any inducers. Each plot represents the mean percentage of three biological replicates. (C) Cells in each of the seven states were stable during growth from single cells into colonies over 6 days under a time-lapse microscope. We sorted cells and seeded an equal ratio of cells in seven states using the same method as in Fig. 3D. Scale bars, 500  $\mu$ m for the wide-field image, 100  $\mu$ m for zoomed-in images. (D) MultiFate is expandable. The number of robust stable fixed points grows monotonically with the number of transcription factor species ( $N$ ) in the model. A robust stable fixed point is defined as a stable fixed point that has fewer than 10% of cells escaping at the end of stochastic simulations (25). The parameter set provided above the plot (with  $K_d = 1$  and  $n = 1.5$ ) is the same nondimensionalized parameter set used in MultiFate-2 and MultiFate-3 models under high protein stability.

marker, into Chinese hamster ovary K1 (CHO-K1) cells and analyzed Citrine expression by flow cytometry 36 hours later (Fig. 2A and fig. S4A) (25). The wild-type (WT) ZF-GCN4-AD factors strongly activated the reporter, as desired, whereas ZF-AD exhibited weaker, but still undesirable, basal activity (Fig. 2A and fig. S4B). Following previous work (29, 34, 35), we incorporated arginine-to-alanine mutations at key positions in the ZF known to weaken DNA binding, which decreased monomeric activity without reducing homodimer activity (Fig. 2A, red square). Replacing the GCN4 with the FKBP12F36V (FKBP) homodimerization domain (36) allowed us to achieve dose-dependent control of dimerization with the small molecule AP1903 (Fig. 2B). Finally, we repeated this general design to engineer a set of additional homodimer-dependent ZF transcription factors with orthogonal DNA binding specificities (fig. S4, B and C).

The MultiFate circuit design requires that each transcription factor positively autoregulates its own expression in a homodimer-dependent manner. To validate this capability, we designed a self-activation construct (Fig. 2C, left) in which a transcription factor with an FKBP dimerization domain is expressed from a promoter containing its own 18-bp homodimer binding sites (table S2). This construct allowed independent doxycycline (Dox)-inducible activation through upstream Tet3G (Takara Bio) binding sites. It also incorporated a dihydrofolate reductase (DHFR) degron (37), which can be inhibited by trimethoprim (TMP), permitting control of protein stability. Finally, we incorporated a destabilized mCitrine for dynamic readout of construct expression. We integrated this construct into Tet3G-expressing CHO-K1 cells, generating a stable polyclonal population for further analysis (table S3) (25).

To test for self-activation, we transiently induced transcription factor expression for 24 hours with Dox, and then withdrew Dox and checked whether cells could sustain circuit activation when dimerization strength and protein stability were varied by AP1903 and TMP, respectively. In the presence, but not the absence, of AP1903, cells exhibited a bimodal distribution of mCitrine fluorescence, with well-separated peaks (Fig. 2C, center), consistent with homodimer-dependent self-activation in a subset of cells. TMP, by stabilizing transcription factors, also promoted self-activation in a dose-dependent manner (Fig. 2C and fig. S5A). Thus, a single dimer-dependent transcription factor can self-activate and sustain its own expression in a controllable manner.

MultiFate's final requirement is the ability of one transcription factor to effectively inhibit another through heterodimerization. To test this, we selected monoclonal cell lines with the self-activating circuits, and then stably

integrated constructs expressing proteins with a different ZF DNA binding domain and a matching or mismatching dimerization domain to generate a polyclonal cell population for each perturbation construct (tables S2 and S3) (25). Consistent with inhibition through heterodimerization, the proteins with matching dimerization domains strongly inhibited the self-activating transcription factor, whereas similar proteins with nonmatching dimerization domains exhibited much weaker inhibition, possibly through nonspecific mechanisms (Fig. 2D and fig. S5B). Taken together, these results provided a set of engineered ZF transcription factors that exhibited controllable homodimer-dependent activation and heterodimer-dependent inhibition.

### The MultiFate-2 circuit generates tristability

To construct a complete MultiFate circuit, we selected two dimer-dependent transcription factors, designated A and B, with distinct DNA binding specificities but the same FKBP homodimerization domain. Their expressions were driven by promoters containing multiple repeats of their corresponding 18-bp homodimer binding sites (Fig. 3A and table S2). The promoters also incorporated Tet3G or ERT2-Gal4 response elements (38) to allow independent external activation of transcription. Factors A and B were transcriptionally coexpressed with destabilized mCherry or mCitrine fluorescent proteins, respectively, each placed after an internal ribosome entry site (IRES), allowing fluorescent readout of transcription rates in individual cells (fig. S6). We stably integrated both genes simultaneously in CHO-K1 cells expressing Tet3G and ERT2-Gal4 proteins, and then selected and further characterized three stable monoclonal cell lines, designated MultiFate-2.1, MultiFate-2.2, and MultiFate-2.3, with different promoter configurations (fig. S7A and table S3) (25).

To test whether MultiFate circuits support multistability, we activated the circuit by transferring MultiFate-2.1 cells to media containing AP1903 and TMP to allow dimerization and stabilizing the transcription factors. As expected in the regime of type II tristability (Fig. 1C), cells went from low expression of both transcription factors (OFF state) to one of three distinct states, with either A, B, or both transcription factors highly expressed (Fig. 3B). We designated these states A-only, B-only, and A+B, respectively. The three states were well separated by differences in either mCherry or mCitrine expression by a factor of ~25 to 50, and cells grew at similar rates among states (fig. S8). To assess their stability, we sorted cells from each of these states and cultured them continuously for 18 days (25). Strikingly, nearly all cells remained in the sorted state for this extended period (Fig. 3C, MultiFate-2.1 high-TMP columns, and fig. S9),

despite gene expression noise (observable from the spread of cellular fluorescence on flow cytometry plots). This showed that cells were attracted to these states. Stability required positive autoregulation, as withdrawal of AP1903 and TMP collapsed the expression of both factors within 2 days (fig. S9). Similar overall behavior was also observed in MultiFate-2.2 and MultiFate-2.3 (Fig. 3C and figs. S10 and S11). All three MultiFate-2 cell lines thus exhibited dynamics consistent with type II tristability (Fig. 1C).

Time-lapse imaging provided a more direct view of multistability. We cultured an equal ratio of single cells sorted from three different initial states in the same well and imaged them as they developed into colonies (Fig. 3D) (25). In almost all colonies (132 of 134), all cells maintained their initial states for the full duration of the movie, at least 5 days or seven to eight cell cycles (Fig. 3D, figs. S12A and S13A, and movie S2). Together with the flow cytometry analysis, these results demonstrate that all three MultiFate-2 lines can sustain long-term tristability.

### MultiFate-2 supports modulation of state stability and allows controlled state switching

The ability of a transient stimulus to destabilize multipotent states and trigger an irreversible fate change is a hallmark of many cell fate control systems (12–14). In the model, reducing protein stability can eliminate the A+B state while preserving the A-only and B-only states (Fig. 1C). As a result, cells initially occupying the A+B state transition to A-only or B-only states (Fig. 4A, top). When protein stability is restored to its initial value, the A+B attractor reappears. However, for the parameter sets analyzed here, cells remain within the attractor basins of A-only and B-only states and therefore do not return to the A+B state (Fig. 4A, top). Stochastic simulations of single-cell dynamics confirmed this irreversible (hysteretic) behavior (Fig. 4A, top, and movie S3).

To test whether similar bifurcation and hysteretic dynamics occur in the experimental system, we transferred A-only, B-only, and A+B cells from media containing high TMP concentrations ("high TMP") to similar media with reduced TMP concentrations ("low TMP"), which decreased protein stability by permitting degron function. As predicted, reducing protein stability selectively destabilized the A+B state, but not the A-only and B-only states, shifting cells from the A+B state to the A-only or B-only states (Fig. 3C, low-TMP columns, and Fig. 4A, bottom). Different MultiFate-2 cell lines exhibited different transition biases, reflecting clone-specific asymmetries in the experimental MultiFate-2 systems (Fig. 3C and figs. S9 to S11), in a manner consistent with an asymmetric MultiFate model (movie S3 and figs. S14 and S15) (25). Escape from the destabilized A+B state was irreversible, as cells remained in the A-only or B-only state even after they were



transferred back to the high-TMP media (Fig. 4A, bottom, and fig. S9). Thus, MultiFate's ability to support irreversible transitions allows it to produce behaviors resembling stem cell differentiation.

Finally, we asked to what extent we could deliberately switch cells from one state to another through transient perturbations. We used MultiFate-2.3, in which the A and B genes can be independently activated by 4-hydroxy-tamoxifen (4-OHT) and Dox, respectively, to address this question. In this line, the response elements for the inducers are adjacent to the homodimer binding sites. Therefore, the addition of inducers increases A or B expression up to, but not substantially beyond, the level produced by self-activation (Fig. 2C and fig. S16). In the bistable regime, transient induction of either transcription factor switched cells into the corresponding state, where they remained in the absence of further induction (Fig. 4B, left, and figs. S3A and S16A). In the tristable regime, the model predicted, and experiments confirmed, that transient induction of B by Dox could switch A-only cells to the A+B state, but not beyond it to the B-only state (Fig. 4B, top right; fig. S3B, first row; and fig. S16B). Combining transient Dox addition to induce B expression with TMP reduction to destabilize the A+B state successfully transitioned cells from the A+B to the B-only state (Fig. 4B, right second row, and fig. S3B, second row). The reciprocal experiments, in which we induced A expression with 4-OHT with or without reduced TMP, produced equivalent results (Fig. 4B, right column, lower two rows). Taken together, these results demonstrate that MultiFate-2 circuits allow modulation of state stability, irreversible cell state transitions, and direct control of state switching with transient external inducers.

### MultiFate is expandable

Because the MultiFate system implements mutual inhibition among transcription factors through heterodimerization, it can be expanded by adding additional transcription factors without reengineering existing components. In the model, adding a third transcription factor to a MultiFate-2 circuit produces a range of new stability regimes containing three, four, six, seven, or eight stable fixed points, depending on parameter values (Fig. 1D, fig. S2, and movie S4) (25). To test whether experimental MultiFate-2 circuits can be similarly expanded, we stably integrated a third ZF transcription factor, denoted C, containing the same FKBP dimerization domain as A and B, coexpressed with a third fluorescent protein, mTurquoise2, into the MultiFate-2.2 cell line to obtain the MultiFate-3 cell line (Fig. 5A, fig. S7B, and table S3) (25).

After the addition of AP1903 and TMP, MultiFate-3 cells went from low expression of all genes (OFF state) to one of seven distinct

expression states, termed A-only, B-only, C-only, A+B, A+C, B+C, and A+B+C states (Fig. 5B), consistent with a type II septastability regime (Fig. 1D and fig. S2A). Most cells occupied the B-only state ( $79.5 \pm 0.3\%$ ), reflecting asymmetries within the circuit (figs. S14 and S15). To assess the stability of these states, we sorted cells from each of the seven states and continuously cultured them in media containing AP1903 and TMP, analyzing the culture every 3 days by flow cytometry (25). Remarkably, each of the seven states was stable for the full 18-day duration of the experiment (Fig. 5B, high-TMP columns, and fig. S17). Long-term stability required AP1903 and TMP, as expected (fig. S18). Finally, cells from each state could be reset by withdrawal of AP1903 and TMP and then redifferentiated into all seven states when AP1903 and TMP were added back (fig. S18). This indicates that the observed stability is not the result of a mixture of clones permanently locked into distinct expression states.

To directly visualize the septastable dynamics of MultiFate-3, we cocultured single cells sorted from each of the seven states and performed live imaging as they grew into colonies (25). Consistent with the flow cytometry results, cells retained their initial states for the full 6-day duration of the experiment in almost every colony (153 of 157) (Fig. 5C, figs. S12B and S13B, and movie S5).

Like MultiFate-2, the number and stability of different states in MultiFate-3 can be modulated. In the model, reducing protein stability repeatedly bifurcates the system from type II septastability (seven stable states) through hexastability (six stable states) to tristability (three stable states) (Fig. 1D). This process resembles the progressive loss of cell fate potential during stem cell differentiation (39). To experimentally test this prediction, we transferred cells in each of the seven states cultured under the high-TMP (100 nM) condition (high protein stability) to similar media with intermediate-TMP (40 nM) or low-TMP (10 nM) conditions. As predicted by the model, the intermediate-TMP condition destabilized only the A+B+C state, but not the other six states (Fig. 5B, intermediate-TMP columns, and fig. S19), whereas the low-TMP condition destabilized all multiprotein states, preserving only the A-only, B-only, and C-only states (Fig. 5B, low-TMP columns, and fig. S20). Consistent with the model, these transitions were also irreversible: Restoring high TMP concentrations did not cause cells to repopulate previously destabilized states (fig. S21 and movie S6). Taken together, these results demonstrate that the MultiFate-3 circuit supports septastability and allows controlled bifurcations to produce irreversible cell state transitions.

Can the MultiFate architecture be expanded beyond three transcription factors? To un-

derstand higher-order systems, we modeled MultiFate circuits containing up to  $N = 11$  transcription factors (25). Using the same parameter values established for MultiFate-2 and MultiFate-3, the number of attractors reached a maximum of 256 at  $N = 9$ . Analysis of attractor escape rates in stochastic simulations revealed that most of these attractors were robust to gene expression noise (Fig. 5D and fig. S22) (25, 40). The number of attractors grew more slowly than the theoretical limit of  $\sim 2^N$  because stable attractors could only sustain high levels of up to four transcription factors at a time (fig. S23, middle row). This limitation reflects the diminishing share of the active homodimers relative to all dimers. Similarly, the combined basal expression of all transcription factors suppressed homodimer formation, resulting in a decline in the number of attractors for systems containing more than nine transcription factors (Fig. 5D and fig. S23, middle row). Finally, we note that the precise values of the maximum number of stable attractors can be modulated up or down by parameters that affect overall gene expression (fig. S23). Together, these results indicate that the MultiFate architecture can be expanded to generate large numbers of robust stable states.

### Discussion

The astonishing diversity of cell types in our own bodies underscores the critical importance of multistable circuits and provokes the fundamental question of how to engineer a robust, controllable, and expandable synthetic multistable system. We took inspiration from two ubiquitous features of natural multistable systems, namely competitive protein-protein interactions and transcriptional autoregulation, to design a synthetic multistable architecture that operates in mammalian cells. The MultiFate circuits exhibit many of the hallmarks of natural cell fate control systems. They generate as many as seven molecularly distinct, mitotically heritable cell states (Figs. 3 and 5). They allow controlled switching of cells between states with transient transcription factor expression (Fig. 4B), similar to fate reprogramming (16). They support modulation of state stability (Figs. 3 and 5) and permit irreversible cellular transitions through externally controllable parameters such as protein stability (Fig. 4A and fig. S21), similar to the irreversible loss of cell fate potential during stem cell differentiation (12). Finally, implementing cross-inhibition at the protein level makes MultiFate expandable by “plugging in” additional transcription factors without reengineering the existing circuit, a useful feature for synthetic biology. The same design principle may play a related role in natural systems, allowing the emergence of new cell states through transcription factor duplication and subfunctionalization

in a manner analogous to the stepwise expansion of MultiFate circuits demonstrated here (21, 22, 41, 42).

A remarkable feature of this circuit is its close agreement with predictions from a dynamical systems model (Box 1). Despite a lack of precise quantitative parameter values for many molecular interactions, the qualitative behaviors possible with this circuit design can be enumerated and explained from simple properties of the components and their interactions. More precise measurements of effective biochemical parameters and stochastic fluctuations could help to explain, eliminate, or exploit asymmetries and provide a better understanding of the time scales of state transitions.

MultiFate has a relatively simple structure, requiring a small number of genes, all of the same type, yet exhibits robust memory behaviors, scalability, and predictive design. Future work should extend MultiFate into a full-fledged synthetic cell fate control system. Coupling MultiFate to synthetic cell-cell communication systems such as synNotch (43, 44), MESA (45), syntheKines (46), engineered GFP (47), and auxin (48) should enable navigation of cells through a series of fate choices, recapitulating cell behaviors associated with normal development. MultiFate could also allow engineering of multicellular cell therapeutic programs. For example, one could engineer a stemlike state that can either self-renew or “differentiate” into other states that recognize and remember different input signals and communicate with one another to coordinate complex response programs. Such strategies will benefit from the ability of MultiFate to allow probabilistic differentiation into multiple different states in the same condition (fig. S14). In this way, we anticipate that the MultiFate architecture will provide a scalable foundation for exploring the circuit-level principles of cell fate control and will enable new multicellular applications in synthetic biology.

## Methods summary

We performed all tissue culture experiments with CHO-K1 cells (ATCC). For flow cytometry experiments characterizing ZF transcription factors (Fig. 2, A and B, and fig. S4), we cotransfected CHO-K1 cells with mTagBFP2 (as cotransfection marker), reporter, and ZF transcription factor (table S2). Cells were harvested after 36 hours and cell fluorescence was measured by flow cytometry. For experiments characterizing ZF transcription factor self-activation (Fig. 2C and fig. S5A), we stably integrated each self-activation construct (table S2) into polyclonal Tet3G-expressing CHO-K1 cells via PiggyBac (Systems Biosciences) to make a polyclonal cell line (table S3). We transiently activated the inte-

grated self-activation cassettes in each polyclonal line by adding Dox (Sigma-Aldrich) for 24 hours, then washed out Dox and transferred cells into different combinations of AP1903 and/or TMP (Sigma-Aldrich). After another 72 hours, cells were harvested and analyzed by flow cytometry. To test inhibition through competitive dimerization (Fig. 2D and fig. S5B), we selected two monoclonal self-activation lines with 42ZFR2AR39AR67A DNA binding domain and either GCN4 or FKBP dimerization domain. We stably integrated plasmids constitutively expressing different perturbation transcription factors in each monoclonal line, then transferred cells in media containing AP1903 and TMP to permit self-activation. The inhibition strength was quantified as the reduction of self-activation cell fractions.

We constructed MultiFate-2 lines by stably integrating corresponding constructs into polyclonal ERT2-Gal4-P2A-Tet3G-expressing CHO-K1 cells (table S3). We then used fluorescence-activated cell sorting to sort stable A+B cells in media containing AP1903 and TMP as single cells into 384-well plates to obtain monoclonal MultiFate-2 lines. We constructed MultiFate-3 cells by stably integrating the TF C self-activation cassette into MultiFate-2.2 cells, then used a similar sorting method to obtain the MultiFate-3 monoclonal cells (fig. S7).

For flow cytometry experiments characterizing state stability (Figs. 3C and 5B) and state-switching dynamics (Fig. 4), we sorted cells from each state into media containing the corresponding inducers. We continuously cultured these cells by trypsinizing cells and transferred 4% of cells into fresh media containing corresponding inducers every 3 days. The remaining 96% of cells were suspended in the flow cytometry buffer and analyzed by flow cytometry. For time-lapse imaging (Fig. 3D, Fig. 5C, and movies S2 and S5), we sorted cells from each state, mixed them with equal ratio, and sparsely plated cell mixture in the same well with media containing AP1903 and TMP. After 6 to 12 hours, we changed media and began imaging.

Mathematical models of MultiFate circuits are summarized in Box 1 and supplementary text (25). All data, computational and analysis codes, and sequence files are available at [data.caltech.edu/records/1882](http://data.caltech.edu/records/1882). Full materials and methods are available in (25).

## REFERENCES AND NOTES

1. S. Ramón y Cajal, *Histology of the Nervous System of Man and Vertebrates* (Oxford Univ. Press, 1995).
2. S. Huang, Multistability and Multicellularity: Cell Fates as High-Dimensional Attractors of Gene Regulatory Networks. *Comput. Syst. Biol.* **2006**, 293–326 (2006). doi: [10.1016/B978-012088786-6/50033-2](https://doi.org/10.1016/B978-012088786-6/50033-2)
3. T. S. Gardner, C. R. Cantor, J. J. Collins, Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000). doi: [10.1038/35002131](https://doi.org/10.1038/35002131); pmid: [10659857](https://pubmed.ncbi.nlm.nih.gov/10659857/)

4. C. M. Ajo-Franklin et al., Rational design of memory in eukaryotic cells. *Genes Dev.* **21**, 2271–2276 (2007). doi: [10.1101/gad.1586107](https://doi.org/10.1101/gad.1586107); pmid: [17875664](https://pubmed.ncbi.nlm.nih.gov/17875664/)
5. B. P. Kramer et al., An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870 (2004). doi: [10.1038/nbt980](https://doi.org/10.1038/nbt980); pmid: [15184906](https://pubmed.ncbi.nlm.nih.gov/15184906/)
6. T. Lebar et al., A bistable genetic switch based on designable DNA-binding domains. *Nat. Commun.* **5**, 5007 (2014). doi: [10.1038/ncomms6007](https://doi.org/10.1038/ncomms6007); pmid: [25264186](https://pubmed.ncbi.nlm.nih.gov/25264186/)
7. D. R. Burrill, M. C. Inniss, P. M. Boyle, P. A. Silver, Synthetic memory circuits for tracking human cell fate. *Genes Dev.* **26**, 1486–1497 (2012). doi: [10.1101/gad.189035.112](https://doi.org/10.1101/gad.189035.112); pmid: [22751502](https://pubmed.ncbi.nlm.nih.gov/22751502/)
8. J. Santos-Moreno, E. Tasiudi, J. Stelling, Y. Schaerli, Multistable and dynamic CRISPRi-based synthetic circuits. *Nat. Commun.* **11**, 2746 (2020). doi: [10.1038/s41467-020-16574-1](https://doi.org/10.1038/s41467-020-16574-1); pmid: [32488086](https://pubmed.ncbi.nlm.nih.gov/32488086/)
9. F. Wu, R.-Q. Su, Y.-C. Lai, X. Wang, Engineering of a synthetic quadrastable gene network to approach Waddington landscape and cell fate determination. *eLife* **6**, e23702 (2017). doi: [10.7554/eLife.23702](https://doi.org/10.7554/eLife.23702); pmid: [28397688](https://pubmed.ncbi.nlm.nih.gov/28397688/)
10. H. Y. Kueh, A. Champhekar, S. L. Nutt, M. B. Elowitz, E. V. Rothenberg, Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, 670–673 (2013). doi: [10.1126/science.1240831](https://doi.org/10.1126/science.1240831); pmid: [23868921](https://pubmed.ncbi.nlm.nih.gov/23868921/)
11. G. Yao, T. J. Lee, S. Mori, J. R. Nevins, L. You, A bistable Rb-E2F switch underlies the restriction point. *Nat. Cell Biol.* **10**, 476–482 (2008). doi: [10.1038/ncb1711](https://doi.org/10.1038/ncb1711); pmid: [18364697](https://pubmed.ncbi.nlm.nih.gov/18364697/)
12. S. Huang, Y.-P. Guo, G. May, T. Enver, Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007). doi: [10.1016/j.ydbio.2007.02.036](https://doi.org/10.1016/j.ydbio.2007.02.036); pmid: [17412320](https://pubmed.ncbi.nlm.nih.gov/17412320/)
13. W. Xiong, J. E. Ferrell Jr., A positive-feedback-based bistable ‘memory module’ that governs a cell fate decision. *Nature* **426**, 460–465 (2003). doi: [10.1038/nature02089](https://doi.org/10.1038/nature02089); pmid: [14647386](https://pubmed.ncbi.nlm.nih.gov/14647386/)
14. A. Gunne-Braden et al., GATA3 Mediates a Fast, Irreversible Commitment to BMP4-Driven Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell* **26**, 693–706.e9 (2020). doi: [10.1016/j.stem.2020.03.005](https://doi.org/10.1016/j.stem.2020.03.005); pmid: [32302522](https://pubmed.ncbi.nlm.nih.gov/32302522/)
15. K. M. Loh et al., Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. *Cell* **166**, 451–467 (2016). doi: [10.1016/j.cell.2016.06.011](https://doi.org/10.1016/j.cell.2016.06.011); pmid: [27419872](https://pubmed.ncbi.nlm.nih.gov/27419872/)
16. K. Takahashi, S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006). doi: [10.1016/j.cell.2006.07.024](https://doi.org/10.1016/j.cell.2006.07.024); pmid: [16904174](https://pubmed.ncbi.nlm.nih.gov/16904174/)
17. D. Arendt et al., The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 (2016). doi: [10.1038/nrg.2016.127](https://doi.org/10.1038/nrg.2016.127); pmid: [27818507](https://pubmed.ncbi.nlm.nih.gov/27818507/)
18. C. Sokolik et al., Transcription factor competition allows embryonic stem cells to distinguish authentic signals from noise. *Cell Syst.* **1**, 117–129 (2015). doi: [10.1016/j.cels.2015.08.001](https://doi.org/10.1016/j.cels.2015.08.001); pmid: [26405695](https://pubmed.ncbi.nlm.nih.gov/26405695/)
19. A. Bhattacharya, N. E. Baker, A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell* **147**, 881–892 (2011). doi: [10.1016/j.cell.2011.08.055](https://doi.org/10.1016/j.cell.2011.08.055); pmid: [22078884](https://pubmed.ncbi.nlm.nih.gov/22078884/)
20. H. Hosokawa et al., Transcription factor PU.1 Represses and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription Factor Binding. *Immunity* **48**, 1119–1134.e7 (2018). doi: [10.1016/j.immuni.2018.04.024](https://doi.org/10.1016/j.immuni.2018.04.024); pmid: [29924977](https://pubmed.ncbi.nlm.nih.gov/29924977/)
21. S. Stefanovic et al., Interplay of Oct4 with Sox2 and Sox17: A molecular switch from stem cell pluripotency to specifying a cardiac fate. *J. Cell Biol.* **186**, 665–673 (2009). doi: [10.1083/jcb.200901040](https://doi.org/10.1083/jcb.200901040); pmid: [19736317](https://pubmed.ncbi.nlm.nih.gov/19736317/)
22. I. Aksoy et al., Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J.* **32**, 938–953 (2013). doi: [10.1038/emboj.2013.31](https://doi.org/10.1038/emboj.2013.31); pmid: [23474895](https://pubmed.ncbi.nlm.nih.gov/23474895/)
23. M. J. Thayer et al., Positive autoregulation of the myogenic determination gene MyoD1. *Cell* **58**, 241–248 (1989). doi: [10.1016/0092-8674\(89\)90838-6](https://doi.org/10.1016/0092-8674(89)90838-6); pmid: [2546677](https://pubmed.ncbi.nlm.nih.gov/2546677/)
24. L. A. Neuhold, B. Wold, HLH forced dimers: Tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**, 1033–1042 (1993). doi: [10.1016/0092-8674\(93\)90725-6](https://doi.org/10.1016/0092-8674(93)90725-6); pmid: [7691411](https://pubmed.ncbi.nlm.nih.gov/7691411/)
25. See supplementary materials.
26. S. Bessonard et al., Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development* **141**, 3637–3648 (2014). doi: [10.1242/dev.109678](https://doi.org/10.1242/dev.109678); pmid: [25209243](https://pubmed.ncbi.nlm.nih.gov/25209243/)
27. P. Laslo et al., Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* **126**,

- 755–766 (2006). doi: [10.1016/j.cell.2006.06.052](https://doi.org/10.1016/j.cell.2006.06.052); pmid: 16923394
28. T. Miyamoto *et al.*, Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell* **3**, 137–147 (2002). doi: [10.1016/S1534-5807\(02\)00201-0](https://doi.org/10.1016/S1534-5807(02)00201-0); pmid: 12110174
  29. A. S. Khalil *et al.*, A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647–658 (2012). doi: [10.1016/j.cell.2012.05.045](https://doi.org/10.1016/j.cell.2012.05.045); pmid: 22863014
  30. R. R. Beerli, D. J. Segal, B. Dreier, C. F. Barbas 3rd, Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14628–14633 (1998). doi: [10.1073/pnas.95.25.14628](https://doi.org/10.1073/pnas.95.25.14628); pmid: 9843940
  31. J. J. Lohmueller, T. Z. Armel, P. A. Silver, A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res.* **40**, 5180–5187 (2012). doi: [10.1093/nar/gks142](https://doi.org/10.1093/nar/gks142); pmid: 22323524
  32. C. J. Bashor *et al.*, Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. *Science* **364**, 593–597 (2019). doi: [10.1126/science.aau8287](https://doi.org/10.1126/science.aau8287); pmid: 31000590
  33. O. M. Subach, P. J. Cranfill, M. W. Davidson, V. V. Verkhusha, An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLOS ONE* **6**, e28674 (2011). doi: [10.1371/journal.pone.0028674](https://doi.org/10.1371/journal.pone.0028674); pmid: 22174863
  34. P. S. Donahue *et al.*, The COMET toolkit for composing customizable genetic programs in mammalian cells. *Nat. Commun.* **11**, 779 (2020). doi: [10.1038/s41467-019-14147-5](https://doi.org/10.1038/s41467-019-14147-5); pmid: 32034124
  35. M. Elrod-Erickson, T. E. Benson, C. O. Pabo, High-resolution structures of variant Zif268-DNA complexes: Implications for understanding zinc finger-DNA recognition. *Structure* **6**, 451–464 (1998). doi: [10.1016/S0969-2126\(98\)00047-1](https://doi.org/10.1016/S0969-2126(98)00047-1); pmid: 9562555
  36. T. Clackson *et al.*, Redesigning an FKBP-ligand interface to generate chemical dimers with novel specificity. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10437–10442 (1998). doi: [10.1073/pnas.95.18.10437](https://doi.org/10.1073/pnas.95.18.10437); pmid: 9724721
  37. M. Iwamoto, T. Björklund, C. Lundberg, D. Kirik, T. J. Wandless, A general chemical method to regulate protein stability in the mammalian central nervous system. *Chem. Biol.* **17**, 981–988 (2010). doi: [10.1016/j.chembiol.2010.07.009](https://doi.org/10.1016/j.chembiol.2010.07.009); pmid: 20851347
  38. S. S. Gerety *et al.*, An inducible transgene expression system for zebrafish and chick. *Development* **140**, 2235–2243 (2013). doi: [10.1242/dev.091520](https://doi.org/10.1242/dev.091520); pmid: 23633515
  39. F. Notta *et al.*, Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016). doi: [10.1126/science.aab2116](https://doi.org/10.1126/science.aab2116); pmid: 26541609
  40. P. S. Swain, M. B. Elowitz, E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12795–12800 (2002). doi: [10.1073/pnas.162041399](https://doi.org/10.1073/pnas.162041399); pmid: 12237400
  41. K. Pougach *et al.*, Duplication of a promiscuous transcription factor drives the emergence of a new regulatory network. *Nat. Commun.* **5**, 4868 (2014). doi: [10.1038/ncomms5868](https://doi.org/10.1038/ncomms5868); pmid: 25204769
  42. J. González *et al.*, Diversification of Transcriptional Regulation Determines Subfunctionalization of Paralogous Branched Chain Amino transferases in the Yeast *Saccharomyces cerevisiae*. *Genetics* **207**, 975–991 (2017). doi: [10.1534/genetics.117.300290](https://doi.org/10.1534/genetics.117.300290); pmid: 28912343
  43. L. Morsut *et al.*, Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164**, 780–791 (2016). doi: [10.1016/j.cell.2016.01.012](https://doi.org/10.1016/j.cell.2016.01.012); pmid: 26830878
  44. S. Toda *et al.*, Engineering synthetic morphogen systems that can program multicellular patterning. *Science* **370**, 327–331 (2020). doi: [10.1126/science.abc0033](https://doi.org/10.1126/science.abc0033); pmid: 33060357
  45. K. A. Schwarz, N. M. Daringer, T. B. Dolberg, J. N. Leonard, Rewiring human cellular input-output using modular extracellular sensors. *Nat. Chem. Biol.* **13**, 202–209 (2017). doi: [10.1038/nchembio.2253](https://doi.org/10.1038/nchembio.2253); pmid: 27941759
  46. I. Moraga *et al.*, Synthesins are surrogate cytokine and growth factor agonists that compel signaling through non-natural receptor dimers. *eLife* **6**, e22882 (2017). doi: [10.7554/eLife.22882](https://doi.org/10.7554/eLife.22882); pmid: 28498099
  47. K. S. Staporowongkul, M. de Gennes, L. Coconci, G. Salbreux, J.-P. Vincent, Patterning and growth control in vivo by an engineered GFP gradient. *Science* **370**, 321–327 (2020). doi: [10.1126/science.abb8205](https://doi.org/10.1126/science.abb8205); pmid: 33060356
  48. Y. Ma, M. W. Budde, M. N. Mayalu, J. Zhu, R. M. Murray, M. B. Elowitz, Synthetic mammalian signaling circuits for robust cell population control. bioRxiv 278564 [preprint] (2020). doi: [10.1101/2020.09.02.278564](https://doi.org/10.1101/2020.09.02.278564)
  49. M. R. Bennett, D. Volfson, L. Tsimring, J. Hasty, Transient dynamics of genetic regulatory networks. *Biophys. J.* **92**, 3501–3512 (2007). doi: [10.1529/biophysj.106.095638](https://doi.org/10.1529/biophysj.106.095638); pmid: 17350994
  50. S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Hachette, 2014).
  51. K. J. Polach, J. Widom, A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. *J. Mol. Biol.* **258**, 800–812 (1996). doi: [10.1006/jmbi.1996.0288](https://doi.org/10.1006/jmbi.1996.0288); pmid: 8637011
  52. J. Miyazaki *et al.*, Expression vector system based on the chicken  $\beta$ -actin promoter directs efficient production of interleukin-5. *Gene* **79**, 269–277 (1989). doi: [10.1016/0378-1119\(89\)90209-6](https://doi.org/10.1016/0378-1119(89)90209-6); pmid: 2551778
  53. H. Chassin *et al.*, A modular degen library for synthetic circuits in mammalian cells. *Nat. Commun.* **10**, 2013 (2019). doi: [10.1038/s41467-019-09974-5](https://doi.org/10.1038/s41467-019-09974-5); pmid: 31043592
  54. D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434 (1976). doi: [10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
  55. X. Pan, C. Dalm, R. H. Wijffels, D. E. Martens, Metabolic characterization of a CHO cell size increase phase in fed-batch cultures. *Appl. Microbiol. Biotechnol.* **101**, 8101–8113 (2017). doi: [10.1007/s00253-017-8531-y](https://doi.org/10.1007/s00253-017-8531-y); pmid: 28951949
  56. Q. Liu, D. J. Segal, J. B. Ghiara, C. F. Barbas 3rd, Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5525–5530 (1997). doi: [10.1073/pnas.94.11.5525](https://doi.org/10.1073/pnas.94.11.5525); pmid: 9159105
  57. L. Bintu *et al.*, Transcriptional regulation by the numbers: Applications. *Curr. Opin. Genet. Dev.* **15**, 125–135 (2005). doi: [10.1016/j.gde.2005.02.006](https://doi.org/10.1016/j.gde.2005.02.006); pmid: 15797195
  58. E. Eden *et al.*, Proteome half-life dynamics in living human cells. *Science* **331**, 764–768 (2011). doi: [10.1126/science.1199784](https://doi.org/10.1126/science.1199784); pmid: 21233346
  59. B. Schwanhäusser *et al.*, Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011). doi: [10.1038/nature10098](https://doi.org/10.1038/nature10098); pmid: 21593866
  60. K. L. Frieda *et al.*, Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017). doi: [10.1038/nature20777](https://doi.org/10.1038/nature20777); pmid: 27869821
  61. B. Schwalb *et al.*, TT-seq maps the human transient transcriptome. *Science* **352**, 1225–1228 (2016). doi: [10.1126/science.aad9841](https://doi.org/10.1126/science.aad9841); pmid: 27257258
  62. J. J. Muldoon *et al.*, Model-guided design of mammalian genetic programs. *Sci. Adv.* **7**, eaabe9375 (2021). doi: [10.1126/sciadv.aabe9375](https://doi.org/10.1126/sciadv.aabe9375)
  63. J. A. Zitzewitz, O. Bilsel, J. Luo, B. E. Jones, C. R. Matthews, Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy. *Biochemistry* **34**, 12812–12819 (1995). doi: [10.1021/bi00039a042](https://doi.org/10.1021/bi00039a042); pmid: 7548036
  64. M. Schlosshauer, D. Baker, Realistic protein-protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. *Protein Sci.* **13**, 1660–1669 (2004). doi: [10.1110/ps.03517304](https://doi.org/10.1110/ps.03517304); pmid: 15133165
  65. S. Paulous, C. E. Malnou, Y. M. Michel, K. M. Kean, A. M. Borman, Comparison of the capacity of different viral internal ribosome entry segments to direct translation initiation in poly(A)-dependent reticulocyte lysates. *Nucleic Acids Res.* **31**, 722–733 (2003). doi: [10.1093/nar/gkg695](https://doi.org/10.1093/nar/gkg695); pmid: 12527782
  66. E. Balleza, J. M. Kim, P. Cluzel, Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods* **15**, 47–51 (2018). doi: [10.1038/nmeth.4509](https://doi.org/10.1038/nmeth.4509); pmid: 29320486
  67. L. He, R. Binari, J. Huang, J. Faló-Sanjuan, N. Perrimon, In vivo study of gene expression with an enhanced dual-color fluorescent transcriptional timer. *eLife* **8**, e46181 (2019). doi: [10.7554/eLife.46181](https://doi.org/10.7554/eLife.46181); pmid: 31140975
  68. J. E. Ferrell Jr., S. H. Ha, Ultrasensitivity part II: Multisite phosphorylation, stoichiometric inhibitors, and positive feedback. *Trends Biochem. Sci.* **39**, 556–569 (2014). doi: [10.1016/j.tibs.2014.09.003](https://doi.org/10.1016/j.tibs.2014.09.003); pmid: 25440716
  69. J. E. Ferrell Jr., S. H. Ha, Ultrasensitivity part III: Cascades, bistable switches, and oscillators. *Trends Biochem. Sci.* **39**, 612–618 (2014). doi: [10.1016/j.tibs.2014.10.002](https://doi.org/10.1016/j.tibs.2014.10.002); pmid: 25456048
  70. J. A. Miller, J. Widom, Collaborative competition mechanism for gene activation in vivo. *Mol. Cell. Biol.* **23**, 1623–1632 (2003). doi: [10.1128/MCB.23.5.1623-1632.2003](https://doi.org/10.1128/MCB.23.5.1623-1632.2003); pmid: 12588982
  71. L. A. Mirny, Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22534–22539 (2010). doi: [10.1073/pnas.0913805107](https://doi.org/10.1073/pnas.0913805107); pmid: 21149679
  72. N. E. Buchler, M. Louis, Molecular titration and ultrasensitivity in regulatory networks. *J. Mol. Biol.* **384**, 1106–1119 (2008). doi: [10.1016/j.jmb.2008.09.079](https://doi.org/10.1016/j.jmb.2008.09.079); pmid: 18938177
  73. C. Hsu, V. Jaquet, M. Gencoglu, A. Becskei, Protein Dimerization Generates Bistability in Positive Feedback Loops. *Cell Rep.* **16**, 1204–1210 (2016). doi: [10.1016/j.celrep.2016.06.072](https://doi.org/10.1016/j.celrep.2016.06.072); pmid: 27425609
  74. D. V. Israni, H.-S. Li, K. A. Gagnon, J. D. Sander, K. T. Roybal, J. Keith Jung, W. W. Wong, A. S. Khalil, Clinically-driven design of synthetic gene regulatory programs in human cells. bioRxiv 432371 [preprint] (2021). doi: [10.1101/2021.02.22.432371](https://doi.org/10.1101/2021.02.22.432371)
  75. T. Buder, A. Deutsch, M. Seifert, A. Voss-Böhme, CellTrans: An R Package to Quantify Stochastic Cell State Transitions. *Bioinform. Biol. Insights* **11**, 1–14 (2017). doi: [10.1177/1177932217712241](https://doi.org/10.1177/1177932217712241); pmid: 28659714

## ACKNOWLEDGMENTS

We thank M. Budde for suggestions on MultiFate circuit design; J. Tijerina at Caltech Flow Cytometry Facility for help with cell sorting; X. Wang and F. Horns for timely help with experiments during COVID and lab move; S. Xie for help with MultiFate-2 monoclonal screening; S. Xie and S. Satia for advice on coding; J. Bois for teaching and sharing Caltech BE150 course materials for mathematical modeling; A. Khalil for suggestions on the choice of zinc fingers; R. Kuintzle, F. Horns, L. Chong, Z. Chen, M. Flynn, H. Klumpe, M. Budde, B. Gu, J. Gregorowicz, and E. Mun for critical feedback; and other members of the Elowitz lab for scientific input and support. **Funding:** Supported by DARPA (HR0011-17-2-0008, M.B.E.); the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation (UWSC10142, M.B.E.); the Spanish Ministry of Science and Innovation and FEDER (PGC2018-101251-B-I00, J.G.-O.); “Maria de Maetzu” Programme for Units of Excellence in R&D (CEX2018-000792-M, J.G.-O.); and the Generalitat de Catalunya (ICREA Academia program, J.G.-O.). M.B.E. is a Howard Hughes Medical Institute investigator. **Author contributions:** R.Z. and M.B.E. conceived of the project. R.Z. and M.B.E. designed experiments. R.Z. performed experiments. R.Z. and M.B.E. analyzed data. R.Z., J.M.d.R.-S., J.G.-O., and M.B.E. did mathematical modeling. R.Z. and M.B.E. wrote the manuscript with input from all authors. **Competing interests:** R.Z. and M.B.E. are inventors on a US provisional patent application related to this work. **Data and materials availability:** All DNA constructs (table S2) and cell lines (table S3) are available from M.B.E. or through the Addgene repository under a material agreement with California Institute of Technology. All data generated and all the computational and data analysis and modeling code used in the current study are available at [data.caltech.edu/records/1882](https://data.caltech.edu/records/1882).

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abg9765](https://science.org/doi/10.1126/science.abg9765)

Materials and Methods

Supplementary Text

Figs. S1 to S24

Tables S1 to S4

References (55–75)

MDAR Reproducibility Checklist

Movies S1 to S6

10 February 2021; accepted 29 November 2021  
10.1126/science.abg9765



## 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: **[ScienceAdvances.org](https://www.scienceadvances.org)**

Science  
Advances  
AAAS

---

**GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS**

---

## RESEARCH ARTICLES

## ATTOSECOND SCIENCE

## Attosecond coherent electron motion in Auger-Meitner decay

Siqi Li<sup>1,2,†</sup>, Taran Driver<sup>1,3,4,†</sup>, Philipp Rosenberger<sup>1,3,5,6</sup>, Elio G. Champenois<sup>3</sup>, Joseph Duris<sup>1</sup>, Andre Al-Haddad<sup>7</sup>, Vitali Averbukh<sup>4</sup>, Jonathan C. T. Barnard<sup>4</sup>, Nora Berrah<sup>8</sup>, Christoph Bostedt<sup>7,9</sup>, Philip H. Bucksbaum<sup>2,3,10</sup>, Ryan N. Coffee<sup>1,3</sup>, Louis F. DiMauro<sup>11</sup>, Li Fang<sup>11,12</sup>, Douglas Garratt<sup>4</sup>, Averell Gattton<sup>1</sup>, Zhaoheng Guo<sup>1,10</sup>, Gregor Hartmann<sup>13</sup>, Daniel Haxton<sup>14</sup>, Wolfram Helml<sup>15</sup>, Zhirong Huang<sup>1,2</sup>, Aaron C. LaForge<sup>8</sup>, Andrei Kamalov<sup>1,2,3</sup>, Jonas Knurr<sup>3</sup>, Ming-Fu Lin<sup>1</sup>, Alberto A. Lutman<sup>1</sup>, James P. MacArthur<sup>1,2</sup>, Jon P. Marangos<sup>4</sup>, Megan Nantel<sup>1,2</sup>, Adi Natan<sup>3</sup>, Razib Obaid<sup>1,8</sup>, Jordan T. O'Neal<sup>2,3</sup>, Niranjana H. Shivaram<sup>1,16</sup>, Avi Schor<sup>1</sup>, Peter Walter<sup>1</sup>, Anna Li Wang<sup>3,10</sup>, Thomas J. A. Wolf<sup>1,3</sup>, Zhen Zhang<sup>1</sup>, Matthias F. Kling<sup>1,3,5,6</sup>, Agostino Marinelli<sup>1,3,\*</sup>, James P. Cryan<sup>1,3,\*</sup>

In quantum systems, coherent superpositions of electronic states evolve on ultrafast time scales (few femtoseconds to attoseconds; 1 attosecond = 0.001 femtoseconds =  $10^{-18}$  seconds), leading to a time-dependent charge density. Here we performed time-resolved measurements using attosecond soft x-ray pulses produced by a free-electron laser, to track the evolution of a coherent core-hole excitation in nitric oxide. Using an additional circularly polarized infrared laser pulse, we created a clock to time-resolve the electron dynamics and demonstrated control of the coherent electron motion by tuning the photon energy of the x-ray pulse. Core-excited states offer a fundamental test bed for studying coherent electron dynamics in highly excited and strongly correlated matter.

Interference is a pillar of quantum physics and a manifestation of one of its most profound consequences: the wavelike nature of matter. A quantum system can exist in a superposition of energy states whose relative quantum phases progress in time. This behavior can cause the states to interfere constructively or destructively as the system evolves, causing physical observables (e.g., charge density) to oscillate in time. Such oscillations are known as quantum beats and have a period of  $T_{QB} = \hbar/\Delta E$ , where  $\hbar$  is Planck's constant and  $\Delta E$  is the energetic separation between the states (1–5). To display a quantum beat, two conditions must be satisfied: First, the quantum system must be prepared in a superposition of two or more different energy states that have a well-defined (or coherent) relationship between their individual quantum phases, which remains stable over the beat period between the relevant phases. Second, the physical observable must be sensitive to the difference between the quantum phases of the energy states forming the coherent superposition.

In this work, we demonstrated the creation and observation of coherent superpositions of

core-excited states in molecules using attosecond x-ray pulses. These molecules decayed nonradiatively via the Auger-Meitner (AM) mechanism—a multielectron process in which the core vacancy created by an x-ray pulse is filled by one electron from a valence orbital, and another valence electron is emitted to conserve energy. The AM process is the dominant mechanism for relaxation following x-ray absorption in most biologically relevant molecules, and in any molecules composed of light atoms such as carbon, oxygen, and nitrogen.

We demonstrated how coherence in short x-ray pulses is imprinted on excited electronic states in x-ray-matter interaction and how this coherence affects the attosecond evolution of the excited electronic wave packet. To this end, we measured the time-dependent AM yield and found that it was sensitive to the quantum coherence of the electronic wave packet, as well as the differences in the excited state populations. The coherence of the wave packet was manifested as femtosecond modulations (or quantum beats) in the time-dependent electron yield. The effect of the wave packet coherence on the relaxation process could have

implications for a broad class of other ultrafast experiments in which the need for high temporal resolution necessitates the use of broadbandwidth x-ray pulses.

Time-resolved measurements of any correlated electron interaction (including AM decay) are challenging because of the extreme time scale (few to subfemtosecond) on which electron-electron interactions occur. Previous time-resolved measurements have extracted a single parameter ( $\Gamma$ ) to characterize the decay of a core-excited system (6–9). In the case of short excitation or ionization pulses,  $\Gamma$  corresponds to the lifetime of the core-excited state, but for long pulses the extracted decay constant is altered by interferences with the excitation process (9–11). Our distinct combination of short excitation pulses and a sufficiently long observation window allowed for a direct time-resolved measurement of the AM emission process. We measured a quantum beat, demonstrating the creation and observation of electronic coherence in a core-excited molecular system. Our technique of mapping coherent electronic motion to the AM decay profile offered a distinctive test-bed for studies of electronic coherence in highly excited and strongly correlated systems.

## Measurement

Our experimental setup is shown in Fig. 1A. Isolated soft x-ray attosecond pulses from a free-electron laser (12), tuned near the oxygen 1s  $\rightarrow \pi$  resonance in nitric oxide (NO) (~530 to 540 eV), irradiated a gas target in the presence of a circularly polarized,  $2.3\text{ }\mu\text{m}, 5 \times 10^{12}\text{ W/cm}^2$  laser field. The momentum distribution of the resultant photoelectrons was recorded by a coaxial velocity map imaging spectrometer (c-VMI) (13). Interaction with the x-ray pulse produced electrons from several different photoionization channels: direct ionization of nitrogen K-shell electrons, KLL AM emission resulting from the nitrogen K-shell vacancy, and resonant oxygen AM emission following O 1s  $\rightarrow \pi$  excitation. These channels are labeled in Fig. 1B, which shows the electron momentum distribution recorded without the  $2.3\text{-}\mu\text{m}$  laser field. The 1s  $\rightarrow \pi$  excitation in nitric oxide corresponds to the promotion of an oxygen 1s electron to the degenerate  $2\pi$  molecular orbital, which is already partially occupied by an unpaired valence electron. The resonant AM emission following this excitation has a dominant feature corresponding to channels where one

<sup>1</sup>SLAC National Accelerator Laboratory, Menlo Park, CA, USA. <sup>2</sup>Department of Physics, Stanford University, Stanford, CA, USA. <sup>3</sup>Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. <sup>4</sup>The Blackett Laboratory, Department of Physics, Imperial College London, London, UK. <sup>5</sup>Max Planck Institute of Quantum Optics, Garching, Germany. <sup>6</sup>Physics Department, Ludwig-Maximilians-Universität Munich, Garching, Germany. <sup>7</sup>Paul Scherrer Institute, Villigen, Switzerland. <sup>8</sup>Physics Department, University of Connecticut, Storrs, CT, USA. <sup>9</sup>LUXS Laboratory for Ultrafast X-ray Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. <sup>10</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA. <sup>11</sup>Department of Physics, The Ohio State University, Columbus, OH, USA. <sup>12</sup>Department of Physics, University of Central Florida, Orlando, FL, USA. <sup>13</sup>Institut für Physik und CINSaT, Universität Kassel, Kassel, Germany. <sup>14</sup>KLA Corporation, Milpitas, CA, USA. <sup>15</sup>Department of Physics, TU Dortmund University, Dortmund, Germany. <sup>16</sup>Department of Physics and Astronomy and Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, IN, USA.

\*Corresponding author: Email: marinelli@slac.stanford.edu (A.M.); jcryan@slac.stanford.edu (J.P.C.)

†These authors contributed equally to this work.



of the degenerate  $2\pi$  electrons participates in the decay, leading to excited cationic states. There is a small contribution from the channel where both  $2\pi$  electrons participate, resulting in a  $2\pi^0$  ground configuration of the cation (14).

The circularly polarized laser field mapped the temporal profile of the electron emission on to the momentum measured at the detector. When electrons were released from the molecule following interaction with the x-ray pulse, their trajectory was altered by the presence of the infrared (IR) laser field, similar to the principle of a time-resolving streak camera (15, 16).

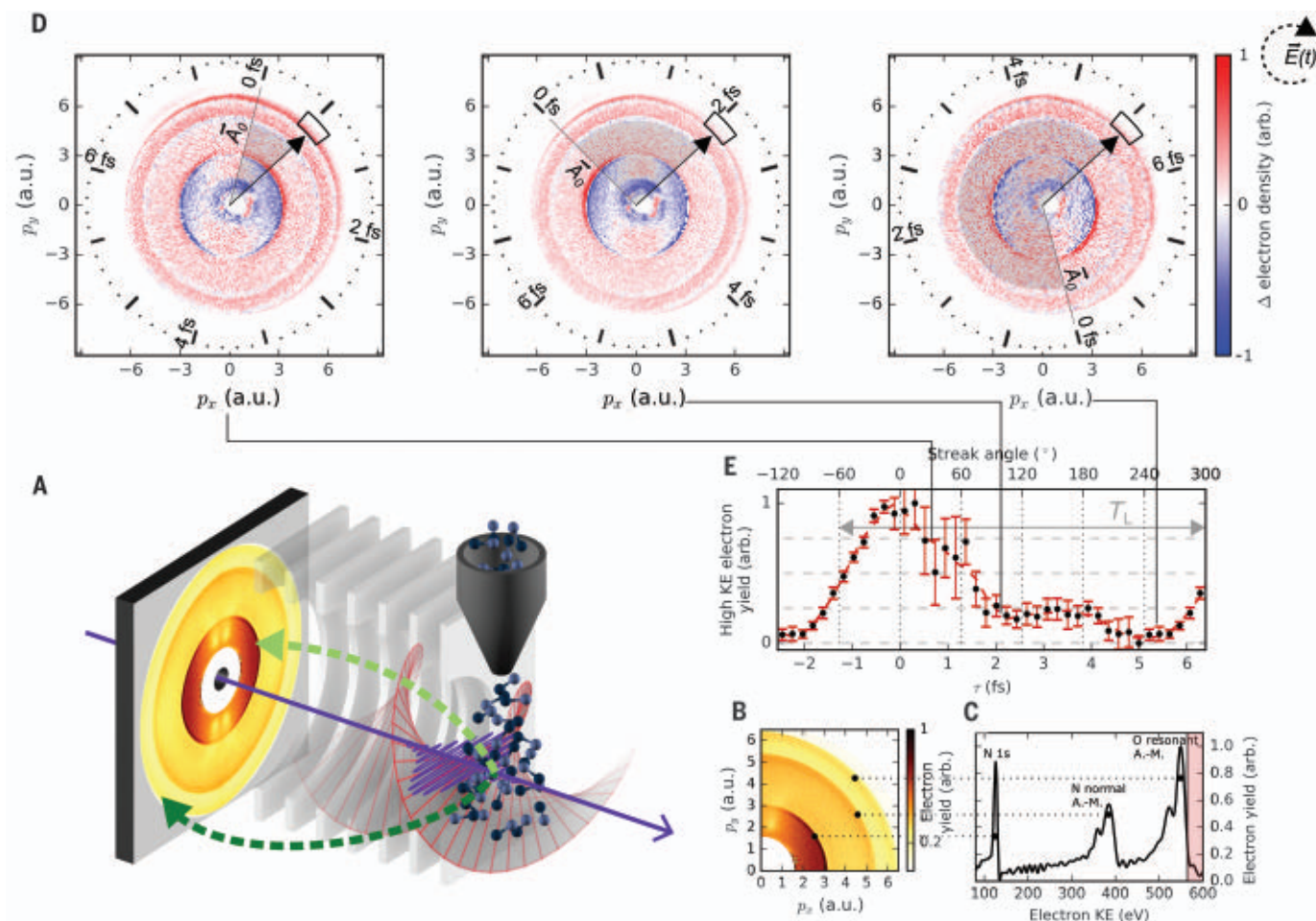
This interaction altered (or “streaked”) the final electron momentum, which was measured at the detector. In a semi-classical approximation, the final momentum of an ionized electron is given by

$$\vec{p}(t \rightarrow \infty) = \vec{p}_0 + e\vec{A}(t_0) \quad (1)$$

where  $\vec{A}(t_0) = -\int_{-\infty}^{t_0} \vec{\mathcal{E}}_L(t') dt'$  is the vector potential of the circularly polarized laser field,  $\vec{\mathcal{E}}_L(t)$ , at the time of ionization  $t_0$ ,  $e$  is the charge of an electron (1 atomic unit), and  $\vec{p}_0$  is the momentum of the electron in the absence of

the IR laser field. All the quantities are expressed in atomic units.

In our measurement, the temporal duration of the circularly polarized “streaking” laser field ( $\sim 100$  fs) was much longer than the laser period ( $T_L = 7.7$  fs). This fact implies that over the time scale of a single laser period, the vector potential had nearly constant amplitude ( $|\vec{A}|$ ) but a direction that rotated with constant angular velocity  $2\pi/T_L$ . Thus, Eq. 1 describes how the streaking technique encodes the temporal evolution of the electron emission rate onto the electron momentum spectrum: An



**Fig. 1. Experimental observation of Auger-Meitner emission.** (A) NO gas is ionized by an attosecond x-ray free-electron laser (XFEL) pulse ( $\sim 530$  to  $540$  eV, central photon energy) in the presence of a  $2.3\text{-}\mu\text{m}$  circularly polarized streaking field. The resultant photoelectron momentum distribution is measured by a coaxial velocity map imaging spectrometer (c-VMI) (13). The streaking field maps the instantaneous ionization rate onto the measured photoelectron momentum distribution. (B) Single-color electron momentum spectrum projected along the axis of the c-VMI in the absence of the streaking field. Atomic units are denoted here and throughout as “a.u.” We define  $p_x$  to lie along the x-ray polarization axis. (C) Applying an inverse Abel transform to this image, we retrieve the electron kinetic energy distribution (“arb.” denotes arbitrary units). (D) Change in the projected momentum distribution as the direction of the streaking laser vector potential ( $\vec{A}_0$ ,

light-gray line pointing along 0 fs of the stopwatch face) is varied. The projected momentum distribution is presented as a difference image where the electron momentum shown in (B) is used as a background. To observe the temporal evolution of AM emission, we monitor the AM yield in a small ( $15^\circ$ ) region of the detector [black box shown in the panels of (D); energetic position also shown in pale red in (C)]. The time dependence of this yield is shown in black dots in (E) (dashed red line shows trace with high-frequency noise filter applied; see supplementary materials for further details). The AM yield in (E) is plotted as a function of angle between the streaking laser vector potential at the time of ionization and the angle of the detection box, which is shown as a gray shaded area in (D).  $\vec{E}$  shows the direction of rotation of the electric field. The red error bars have a total length of four times the SEM of the measured electron yield,  $\pm 2\sigma_x$ .



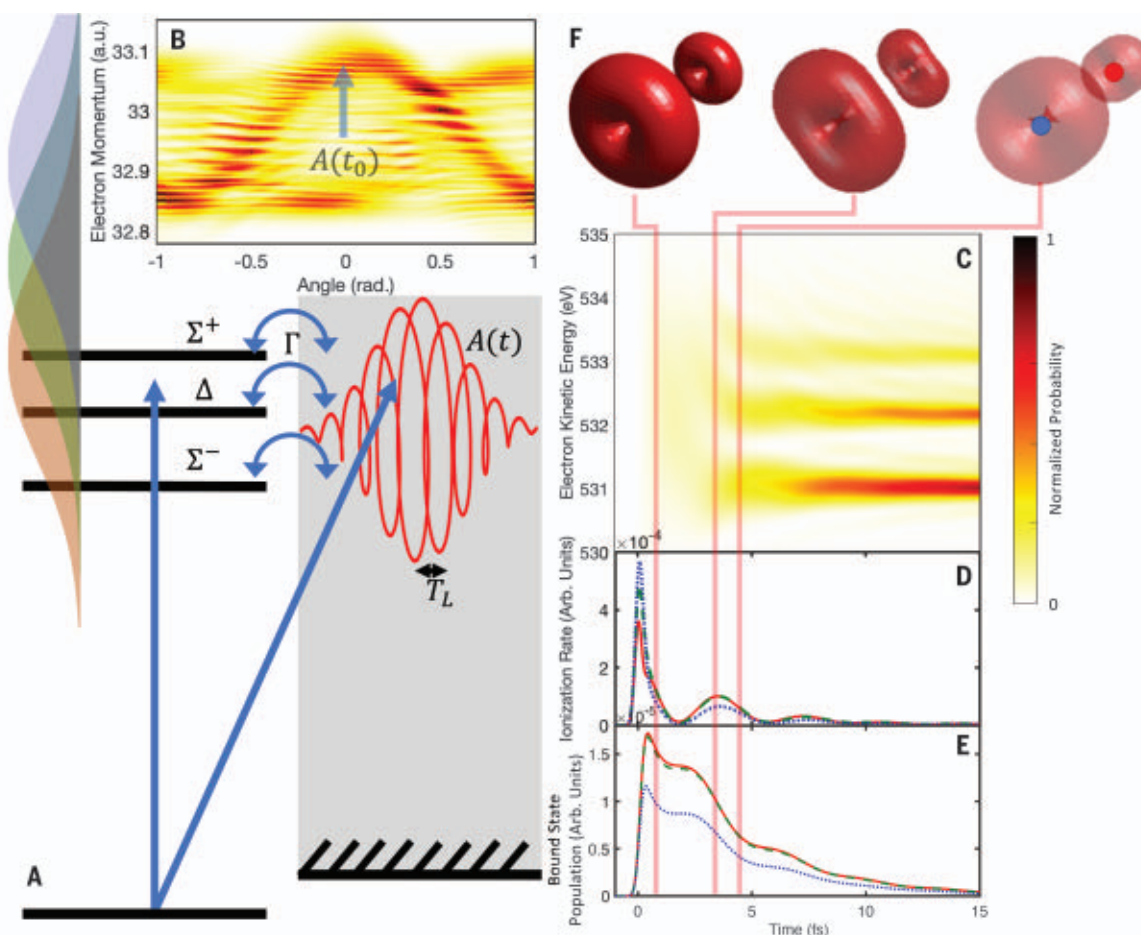
electron emitted at  $t_i$  will experience a momentum shift in the direction of  $\vec{A}(t_i)$ . Because the period  $T_L$  of the circularly polarized laser is well known, if two photoemission features are found to have momentum shifts that differ by an amount  $\Delta\theta$ , this difference implies that the photoemission events were separated by a time  $\Delta\tau$ :

$$\Delta\tau = \frac{\Delta\theta}{2\pi} \times T_L \quad (2)$$

This mapping of angle-to-time resembles the face of a clock, which has led to the term “attoclock” being used to describe this type of time-resolved measurement (17–19).

Our method for extracting the temporal profile of the AM electron yield is illustrated in Fig. 1, D and E. Figure 1D shows the variation in the differential electron yield for measurements with three different x-ray arrival times, or directions of the streaking laser vector potential,  $\vec{A}_0$ . The differential images show the difference between the averaged electron image when the vector potential of the IR laser was chosen to lie along the line labeled  $\vec{A}_0$  on the figure (and labeled as “0 fs” on the clock face) and the averaged electron image where the IR laser was intentionally mistimed with the x-rays to ensure that there was no effect from the streaking field. To

extract the time-dependent emission rate of resonant AM electrons, we monitored a small angular region on the detector (black wedge in Fig. 1D) and plot this yield as a function of streak angle, or the angle between the observation bin and the streaking laser vector potential, in Fig. 1E. The observation region was chosen to be slightly higher in momentum than the center of the field-free resonant emission spectrum shown in Fig. 1B. The electron yield in this radial bin therefore mapped to the number of electrons released into the continuum at the time the vector potential  $\vec{A}_0(t)$  is pointed in the angular direction of the observation region. The lower momentum



**Fig. 2. Model for Auger-Meitner emission.** (A) Schematic representation of the model used for AM emission. Subfemtosecond x-ray pulses coherently excite four resonances (labeled  $^2\Sigma^+$ ,  $^2\Sigma^-$ , and a doubly degenerate  $^2\Delta$ ). In addition to the resonant pathway, electrons can be directly ionized by the x-ray pulse, leading to interfering paths from the ground state to the field-dressed continuum [although the direct ionization pathway is a minor channel (14)]. (B) Calculated photoelectron momentum spectrum for 0.5-fs x-ray pulses centered at 533-eV photon energy in the presence of a 2.3- $\mu\text{m}$  laser field. The blue arrow shows the direction of IR laser vector potential,  $A(t)$ , at the x-ray arrival time,  $t_0$ . (C) Kinetic energy distribution of the continuum

electron as a function of time in the absence of the streaking laser field. (D) Time-dependent ionization rate for this wave function, summed over electron kinetic energy, for a central photon energy of 533 eV (red), 534.5 eV (green), and 536 eV (blue). (E) Total population of the core-excited states as a function of time delay for the same photon energies as in (D). (F) Time evolution of the electron density of the bound electronic states. The three-dimensional contour is drawn at 20% of the maximum electron density, and its transparency represents the overall bound-state population, which decays via AM emission. The blue and red dots in the rightmost panel show the positions of the nitrogen and oxygen atoms, respectively.

limit of this region is  $\sim 6.4$  atomic units, and the upper limit extends to include all electrons at higher momenta. Equation 2 can be used to convert the streak angle into a time delay, and this value is used to label the clock face in Fig. 1D and the lower horizontal axis in Fig. 1E.

At the Linac Coherent Light Source (LCLS), the synchronization of the streaking laser and x-ray pulse has a jitter of roughly  $\sim 500$  fs (20), which is orders of magnitude below the required precision for directly timing the AM process. Thus, to produce the images shown in Fig. 1D, we must use a single-shot diagnostic of the relative arrival time between the x-rays and laser pulse. As described above, in addition to driving resonant excitation near the oxygen K-edge, the attosecond x-ray pulse ionized electrons from the nitrogen K-shell of the NO molecule (see Fig. 1, B and C). This direct photoionization process produced high energy ( $\sim 120$  eV) electrons. The photoionization delay between the arrival of the x-ray pulse and the appearance of these fast photoelectrons in the continuum was negligibly small ( $\lesssim 5$  as)

compared to the streaking laser period  $T_L$  of 7.7 fs (21–23). Therefore, the momentum shift observed for the nitrogen K-shell photoemission feature provided an accurate, single-shot measurement of the direction of the streaking laser vector potential  $\vec{A}_0$  at the time of arrival of the x-ray pulse.

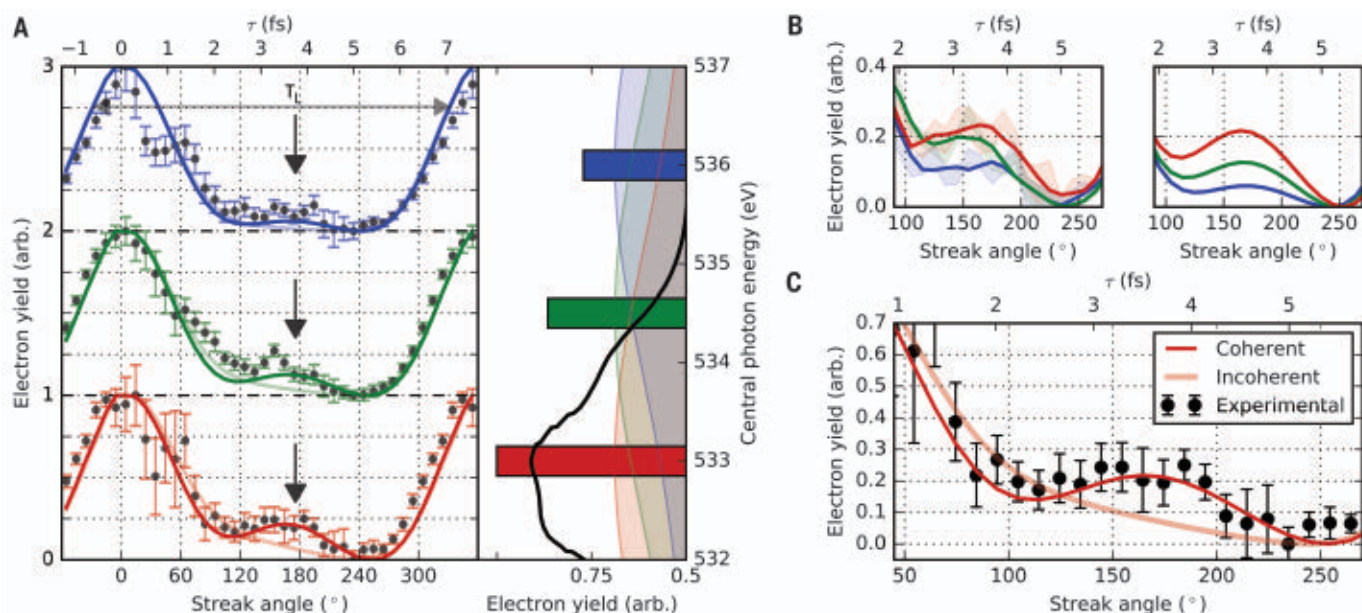
We monitor the AM yield in a small angular region of the detector to avoid introducing artifacts in the extracted time-dependent trace due to angular anisotropy in the AM emission (24). The period of the streaking field was chosen to be longer than the dominant time scale of the AM process. This fact simplifies interpretation of the streaking measurement by limiting the effect of “wrapping,” where electrons released into the continuum at time  $\tau$  and  $\tau + T_L$  experience a similar momentum kick from the streaking field.

The time-dependent electron yield shown in Fig. 1E shows a maximum at  $\tau = 0$ , when  $\vec{A}_0$  was directed along the detection direction and the the core-excited population (and AM emission rate) was at a maximum. In addition

to an exponentially decaying electron emission rate, we observed a revival in the time-dependent emission rate at  $\tau = 3.5$  fs.

## Model

We modeled our measurement according to the theory of attosecond streaking of multiple Fano resonances described by Wickenhauser *et al.* (25, 26). Our model, illustrated in Fig. 2A, included a ground state that is doubly degenerate and was resonantly coupled to four bound states, one of which ( $^2\Delta$ ) is also doubly degenerate, and thus is labeled as a single state in the figure. These bound states were also coupled to a single, structure-less continuum, which was dressed by the circularly polarized,  $2.3\text{-}\mu\text{m}$  streaking laser field. The coupling between the bound and continuum states was the result of electron correlation interactions and drove the AM decay process. The bound states had excitation energies of 531.5 eV ( $^2\Sigma^-$ ), 532.6 eV ( $^2\Delta$ ), and 533.5 eV ( $^2\Sigma^+$ ), which represented the core-excitation spectrum of nitric oxide (27). The continuum coupling constant



**Fig. 3. Comparison between model and experimental results for resonant Auger-Meitner emission.** (A) Measurement of time-resolved Auger-Meitner emission from core-excited NO. The left panel shows the experimentally measured time-dependent AM yield for the various central XFEL photon energies (black dots). Colored error bars have a total length of four times the SEM of the measured electron yield,  $\pm 2\sigma_x$ . This measurement is compared with the results of the model shown in Fig. 2 (solid colored lines). The right panel shows total electron yield, which decreases as the central photon energy moves away from the center of the  $1s \rightarrow \pi$  resonance (bars). The time-dependent yields change by a factor of 2 between the minimum (normalized to 0) and maximum (normalized to 1) values. The coherent bandwidth of the attosecond XFEL pulse spans  $\sim 5$  eV, as illustrated by a Gaussian curve of equivalent full width at half maximum at each central photon energy. The black line shows the  $O1s \rightarrow \pi$  feature reported in (27), comprising the  $^2\Sigma^-$ ,  $^2\Delta$ , and  $^2\Sigma^+$  electronic states. The revival at  $\tau \sim 3.5$  fs, marked by the black vertical arrow, is due to the rephasing

(constructive interference) of the AM emission from the core-excited states ( $^2\Sigma$  and  $^2\Delta$ ) populated by the x-ray pulse. The coherent revival is suppressed as the photon energy moves above the  $1s \rightarrow \pi$  resonance and the contribution from the direct photoionization channel increases. The photon energy-dependence of the quantum beat is shown in the magnified image in (B), for experiment (left) and simulation (right). The shaded area represents the streak-angle-dependent yield with corresponding error bar  $\pm 2\sigma_x$ , and the solid line shows this electron yield after application of a high-frequency filter along the time axis (see supplementary materials for further details). The color of the curves corresponds to the central photon energies shown on the left side of (A). (C) Comparison between two different models where core-excited states are populated coherently (deep red) and incoherently (pale red) at 533-eV central photon energy. The experimental measurement is shown in black dots with error bars  $\pm 2\sigma_x$ . Coherent interaction between the core-excited states is required to account for the measured data.

( $\Gamma = 170$  meV) was consistent with previous x-ray absorption measurements (27). The relative amplitude between transitions to the bound, core-excited states and direct photoionization of valence electrons to the continuum was represented by the Fano parameter,  $q_i$  (see supplementary materials) (28). We choose the value for  $q_i$  according to the measured absorption spectrum of NO (27).

The coherent bandwidth of the exciting x-ray source was  $\sim 5$  eV (12), which was sufficient to span all core-excited bound states in the model. Symmetry constraints did not allow for the coherent population of the  $^2\Sigma^-$  and  $^2\Sigma^+$  states because the  $^2\Sigma^-$  and  $^2\Sigma^+$  states each coupled to a different component of the doubly degenerate ground state (14). Moreover, each component of the ground state coupled to a different component of the degenerate  $^2\Delta$  state (14). Thus, the model only included coherence between the  $^2\Sigma^-$  and one of the  $^2\Delta$  states and the  $^2\Sigma^+$  and the other  $^2\Delta$  state, but not the  $^2\Sigma^-$  and  $^2\Sigma^+$  states. In both simulation and experiment, we tune the central wavelength of the x-ray source across the  $1s \rightarrow 2p\pi$  resonance (red, green, and blue shaded curves in Fig. 3, bandwidth drawn to scale with energetic separation of core-excited states).

In the simulation, we could calculate the energy-resolved continuum wave function in the absence of the streaking-field, shown in Fig. 2D, demonstrating the build-up of resonant features. The rate of electron emission (integrated over electron kinetic energy) is shown in Fig. 2E, and we clearly observe an oscillatory emission rate. Finally, in Fig. 2F we show the population of each core-excited state as a function of time, which again shows oscillatory behavior. The periodic modulation of the electron emission rate resulted from the coherent population of the two pairs of excited states  $^2\Sigma^-$  and  $^2\Delta$ , and  $^2\Delta$  and  $^2\Sigma^+$ . Electronic coherence between the pairs of excited states resulted in consecutive minima or maxima in the time-dependent ionization rate, owing to destructive or constructive interference between emission from the core-excited states. Because the core-excited wave packet consisted of states with different angular momentum projections along the molecular axis, the excited state wave packet produced an excited electron density that rotated around the molecular axis, as shown in Fig. 2C.

## Results

We directly modeled our experimental observable by computing the asymptotic ( $t \rightarrow \infty$ ) momentum distribution of ionized electrons within the strong-field approximation (SFA) (26) (Fig. 2B) and performing the same analysis routine as the one we applied to the experimental data. The asymmetry parameters describing emission from the oxygen  $1s \rightarrow ^2\Sigma^-$ ,  $O 1s \rightarrow ^2\Delta$ ,

and  $O 1s \rightarrow ^2\Sigma^+$  excitations were expected to be different for each of the electronic states (24) and have not previously been measured, meaning the contribution of each channel to emission in the direction of our observation window was not well defined. We fit the simulation to the experimental data using the lower kinetic energy limits of the small detector region defined in Fig. 1E, and the relative contribution from each decay channel at the precise region on the detector, as free parameters. With a separate measurement taken concurrent with the presented data, we determined an error distribution of  $\sigma = 30^\circ$  for single-shot vector potential determination. We accounted for this experimental error by convolution of the time-dependent electron yield with a Gaussian kernel of  $\sigma = 30^\circ$ . Further details are provided in the supplementary materials. We also accounted for the possibility of a small systematic error in  $t_0$  determination between experiment and theory, resulting from the finite temporal profile of the reference nitrogen K-shell photoline produced by the attosecond x-ray pulse (12). We identified an offset of  $\sim 1.7\%$  of the full detector angle.

Figure 3A shows the vector-potential direction-dependent electron yield measured at different x-ray excitation energies (black dots) compared with the simulated yield (solid line). The transient revival at  $\tau \sim 3.5$  fs resulting from electronic coherence in the core-excited state is indicated by the black arrow and is observed in both experiment and simulation. This feature is a quantum beat, occurring at the moment when the quantum phases of the coherently excited  $^2\Sigma^-$  and  $^2\Delta$  excitations realigned. This alignment caused constructive interference between the two core-excited states, and an increase in AM emission rate. The feature at  $\sim 1.3$  fs measured at central photon energy 536 eV was possibly due to the temporal build-up of the Fano interference between the resonant and direct excitation channels and has been qualitatively reproduced in further simulation. Analysis of the energetic positions of the Rydberg series converging to the oxygen K-edge (27) was not consistent with the interpretation that this modulation was due to further coherent excitation involving Rydberg states.

Figure 3B shows a magnified image of the revival feature. By tuning the central x-ray photon energy away from the center of the  $1s \rightarrow 2p\pi$  resonance, we could suppress the quantum beat in both experiment (left) and simulation (right), demonstrating control over the coherent evolution of the core-excited states. The beat was suppressed at higher photon energy because of an increased relative contribution from the direct channel versus the coherently excited resonant decay pathways. In Fig. 3C, we compare our measurement, for a central x-ray excitation energy of 533 eV, to our simulation, including (deep red) and ex-

cluding (pale red) the coherence between the core excited states. As expected, the revival feature could be reproduced only by including the coherence between the different core-excited electronic states. The result from incoherent summation of the AM emission from the different core-excited states failed to reproduce the feature at  $150^\circ$  streaking angle.

## Conclusion

This work reports the real-time measurement of electronic coherence in the temporal evolution of a core-excited molecule. Electronic coherence imparted a modulation in the time-dependent emission rate of AM electrons, driven by an isolated attosecond soft x-ray pulse from a free-electron laser. The AM emission occurred on a few-femtosecond time scale, and we time-resolved it using angular streaking. Our measurement provides a testbed for exploring the effect of electronic coherence in the photoexcitation dynamics and subsequent photochemical behavior of molecular systems. The existence of this electronic coherence provides the opportunity to explore interatomic site electronic wave packet coupling, which can reveal interactions between different parts of an extended system (29–31). Measuring this coupling can reveal important information on the system's fundamental physical properties (32, 33). For example, the spectral makeup of the observed modulations provides rich information on the composition of the excited superposition state. This information opens the possibility for resolving in time the evolution and decay of coherent electronic states, as they evolve and couple to subsequent nuclear motion in the first stages of a photochemical reaction (34–37).

## REFERENCES AND NOTES

1. A. T. Forrester, R. A. Gudmundsen, P. O. Johnson, *Phys. Rev.* **99**, 1691–1700 (1955).
2. E. Alexandrov, *Optika i Spektrosk.* **17**, 957 (1964).
3. T. Hadeishi, W. A. Nierenberg, *Phys. Rev. Lett.* **14**, 891–892 (1965).
4. J. Mauritsson et al., *Phys. Rev. Lett.* **105**, 053001 (2010).
5. E. Goulielmakis et al., *Nature* **466**, 739–743 (2010).
6. M. Drescher et al., *Nature* **419**, 803–807 (2002).
7. M. Uiberacker et al., *Nature* **446**, 627–632 (2007).
8. A. J. Verhoeef et al., *Laser Phys.* **21**, 1270–1274 (2011).
9. D. Haynes et al., *Nat. Phys.* **17**, 512–518 (2021).
10. O. Smirnova, V. S. Yakovlev, A. Scrinzi, *Phys. Rev. Lett.* **91**, 253001 (2003).
11. A. K. Kazansky, I. P. Sazhina, N. M. Kabachnik, *J. Phys. At. Mol. Opt. Phys.* **42**, 245601 (2009) Publisher: IOP Publishing.
12. J. Duris et al., *Nat. Photonics* **14**, 30–36 (2020).
13. S. Li et al., *AIP Adv.* **8**, 115308 (2018) Publisher: American Institute of Physics.
14. H. Wang et al., *Chem. Phys.* **289**, 31–44 (2003).
15. D. Bradley, B. Liddy, W. Sleat, *Opt. Commun.* **2**, 391–395 (1971).
16. Y. Tsuchiya, *IEEE J. Quantum Electron.* **20**, 1516–1528 (1984) Conference Name: IEEE Journal of Quantum Electronics.
17. J. Itatani et al., *Phys. Rev. Lett.* **88**, 173903 (2002).
18. N. Hartmann et al., *Nat. Photonics* **12**, 215–220 (2018).
19. P. Eckle et al., *Science* **322**, 1525–1529 (2008).
20. J. M. Glowia et al., *Opt. Express* **18**, 17620–17630 (2010).
21. J. M. Dahlström, A. L'Huillier, A. Maquet, *J. Phys. At. Mol. Opt. Phys.* **45**, 183001 (2012).
22. J. M. Dahlström et al., *Chem. Phys.* **414**, 53–64 (2013).



23. V. V. Serov, V. L. Derbov, T. A. Sergeeva, *Phys. Rev. A* **87**, 063414 (2013).
24. P. V. Demekhin *et al.*, *J. Phys. At. Mol. Opt. Phys.* **43**, 165103 (2010).
25. M. Wickenhauser, J. Burgdörfer, F. Krausz, M. Drescher, *J. Mod. Opt.* **53**, 247–257 (2006).
26. M. Wickenhauser, J. Burgdörfer, F. Krausz, M. Drescher, *Phys. Rev. Lett.* **94**, 023002 (2005).
27. R. Püttner *et al.*, *Phys. Rev. A* **59**, 3415–3423 (1999).
28. U. Fano, *Phys. Rev.* **124**, 1866–1878 (1961).
29. L. S. Cederbaum, J. Zobeley, *Chem. Phys. Lett.* **307**, 205–210 (1999).
30. F. Calegari *et al.*, *Science* **346**, 336–339 (2014).
31. P. M. Kraus *et al.*, *Science* **350**, 790–795 (2015).
32. L. S. Cederbaum, W. Domcke, J. Schirmer, W. V. Niessen, in *Advances in Chemical Physics*, I. Prigogine, S. A. Rice, Eds. (Wiley, 1986), pp. 115–159; <https://doi.org/10.1002/9780470142899.ch3>.
33. A. I. Kuleff, L. S. Cederbaum, *J. Phys. At. Mol. Opt. Phys.* **47**, 124002 (2014).
34. F. Lépine, M. Y. Ivanov, M. J. J. Vrakking, *Nat. Photonics* **8**, 195–204 (2014).
35. M. Vacher, M. J. Bearpark, M. A. Robb, J. P. Malhado, *Phys. Rev. Lett.* **118**, 083001 (2017).
36. V. Despré, N. V. Golubev, A. I. Kuleff, *Phys. Rev. Lett.* **121**, 203002 (2018).
37. A. Marciniak *et al.*, *Nat. Commun.* **10**, 337 (2019).
38. T. Driver, S. Li, A. Marinelli, J. Cryan, Attosecond coherent electron motion in Auger–Meitner decay, Zenodo (2022); <https://doi.org/10.5281/zenodo.5711217>.

## ACKNOWLEDGMENTS

**Funding:** S.L., Z. Z., and A.M. acknowledge support from US Department of Energy (DOE), BES Scientific User Facilities Division Field Work Proposal 100317; J.D. and A. M. were supported by the Laboratory Directed Research and Development Program in support of the Panofsky fellowship. The contributions from T.D., P.H.B., A.K., A.N., J.T.O., T.J.A.W., A.L.W., and J.P.C. were supported by the US DOE, Office of Science, Office of Basic Energy Sciences (BES), Chemical Sciences, Geosciences, and Biosciences Division (CSGB); E.G.C. was supported by the DOE Laboratory Directed Research and Development program at SLAC National Accelerator Laboratory, under contract DE-AC02-76SF00515. P.R. and M.F.K. acknowledge support by the German Research Foundation via KL-1439/10, and the Fellow program of the Max Planck Society, V.A. J.C.T.B., D.G., and J.P.M. gratefully acknowledge funding support from UK EPSRC grants EP/R019509/1, EP/T006943/1, and EP/I032517/1. N.B., R.O., and A.C.L. acknowledge the Chemical Sciences, Geosciences and Biosciences Division, US DOE, Office of Science, BES, grant DE-SC0012376. C.B. acknowledges the Swiss National Science Foundation and the National Center of Competence in Research—Molecular Ultrafast Science and Technology NCCR—MUST. L.F.D. and L.F. acknowledge support from NSF grant 1605042 and DOE DE-FG02-04ER15614. W.H. thanks the German BMBF for funding of the project “SpeAR XFEL” under contract 05K19PE1. Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the US DOE, Office of Science, BES, under Contract DE-AC02-76SF00515. **Author contributions:** S.L., A.M. and J.P.C. devised the experimental scheme. S.L., A.M., and J.P.C. developed the experimental apparatus. S.L., J.D., J.P.M., Z.Z., and A.M. prepared the attosecond x-ray pulses. M.-F.L., N.H.S., and P.W. prepared the experimental beam-line. All Authors participated in the collection and interpretation of the experimental data. T.D. led the data analysis. S.L., T.D., P.R., and E.G.C. worked on the single-shot “streaking” diagnostic. S.L., T.D., A.M., and J.P.C. prepared an initial version of the manuscript. All authors provided critical feedback in preparing the submitted manuscript. **Competing interests:** None declared. **Data and materials availability:** The partially analyzed raw data and the raw data from the calculations is available on the Zenodo repository (38). All (other) data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abj2096](https://science.org/doi/10.1126/science.abj2096)

Materials and Methods

Supplementary Text

Figs. S1 to S8

References (39–42)

28 April 2021; accepted 29 November 2021

Published online 6 January 2022

10.1126/science.abj2096

## PLANT SCIENCE

# RALF peptide signaling controls the polytubey block in *Arabidopsis*

Sheng Zhong<sup>1†</sup>, Ling Li<sup>1†</sup>, Zhijuan Wang<sup>1†</sup>, Zengxiang Ge<sup>1†‡</sup>, Qiyun Li<sup>1†</sup>, Andrea Bleckmann<sup>2</sup>, Jizong Wang<sup>1</sup>, Zihan Song<sup>1</sup>, Yihao Shi<sup>1</sup>, Tianxu Liu<sup>1</sup>, Luhan Li<sup>1</sup>, Huabin Zhou<sup>3</sup>, Yanyan Wang<sup>3</sup>, Li Zhang<sup>1</sup>, Hen-Ming Wu<sup>4</sup>, Luhua Lai<sup>3</sup>, Hongya Gu<sup>1,5</sup>, Juan Dong<sup>6</sup>, Alice Y. Cheung<sup>4</sup>, Thomas Dresselhaus<sup>2</sup>, Li-Jia Qu<sup>1,5\*</sup>

Fertilization of an egg by multiple sperm (polyspermy) leads to lethal genome imbalance and chromosome segregation defects. In *Arabidopsis thaliana*, the block to polyspermy is facilitated by a mechanism that prevents polytubey (the arrival of multiple pollen tubes to one ovule). We show here that FERONIA, ANJEA, and HERCULES RECEPTOR KINASE 1 receptor-like kinases located at the septum interact with pollen tube–specific RALF6, 7, 16, 36, and 37 peptide ligands to establish this polytubey block. The same combination of RALF (rapid alkalization factor) peptides and receptor complexes controls pollen tube reception and rupture inside the targeted ovule. Pollen tube rupture releases the polytubey block at the septum, which allows the emergence of secondary pollen tubes upon fertilization failure. Thus, orchestrated steps in the fertilization process in *Arabidopsis* are coordinated by the same signaling components to guarantee and optimize reproductive success.

Seed plants rely on tightly regulated fertilization mechanisms to secure fertility and reproductive success. Like in animals, the entrance of supernumerary sperm into a single egg—i.e., polyspermy—is restricted to ensure chromosomal balance and progeny health (1, 2). Fertilization in angiosperms is more complex because two sperm cells are carried by one pollen tube that grows in the maternal pistil tissues and ultimately releases its sperm cell cargo inside the ovule (3). Although hundreds of pollen tubes may grow into the transmitting tract of a pistil, usually only a single tube, in response to attractants, emerges from the septum in the vicinity of each ovule to target the ovule (Fig. 1A) (4). The block to polytubey (i.e., the emergence of multiple pollen tubes targeting an ovule) prevents the occurrence of polyspermy. This polytubey block is likely further reinforced by successful gamete fusion that triggers programmed cell death (PCD) of the persistent synergid cell, which leads to the elimination of

pollen tube attractants. Thus, the first pollen tube that emerges from the septum will have the privilege of fertilizing the female gametes, providing a precondition for conspecific pollen precedence (i.e., the preferential use of pollen from the same species for fertilization) (5). If the first pollen tube fails, however, fertilization success will be ensured by fertilization recovery (6, 7), in which the polytubey block is suspended to allow the emergence of secondary pollen tubes for another chance of fertilization. Therefore, plants can (i) restrict polyspermy by enforcing the polytubey block at the septum under normal circumstances and (ii) salvage fertility by removing the polytubey block when fertilization fails. Here, we report the molecular mechanisms by which the polytubey block is implemented or suspended, when needed.

## FERONIA, ANJEA, and HERCULES RECEPTOR KINASE 1 receptor kinases are required to establish the polytubey block

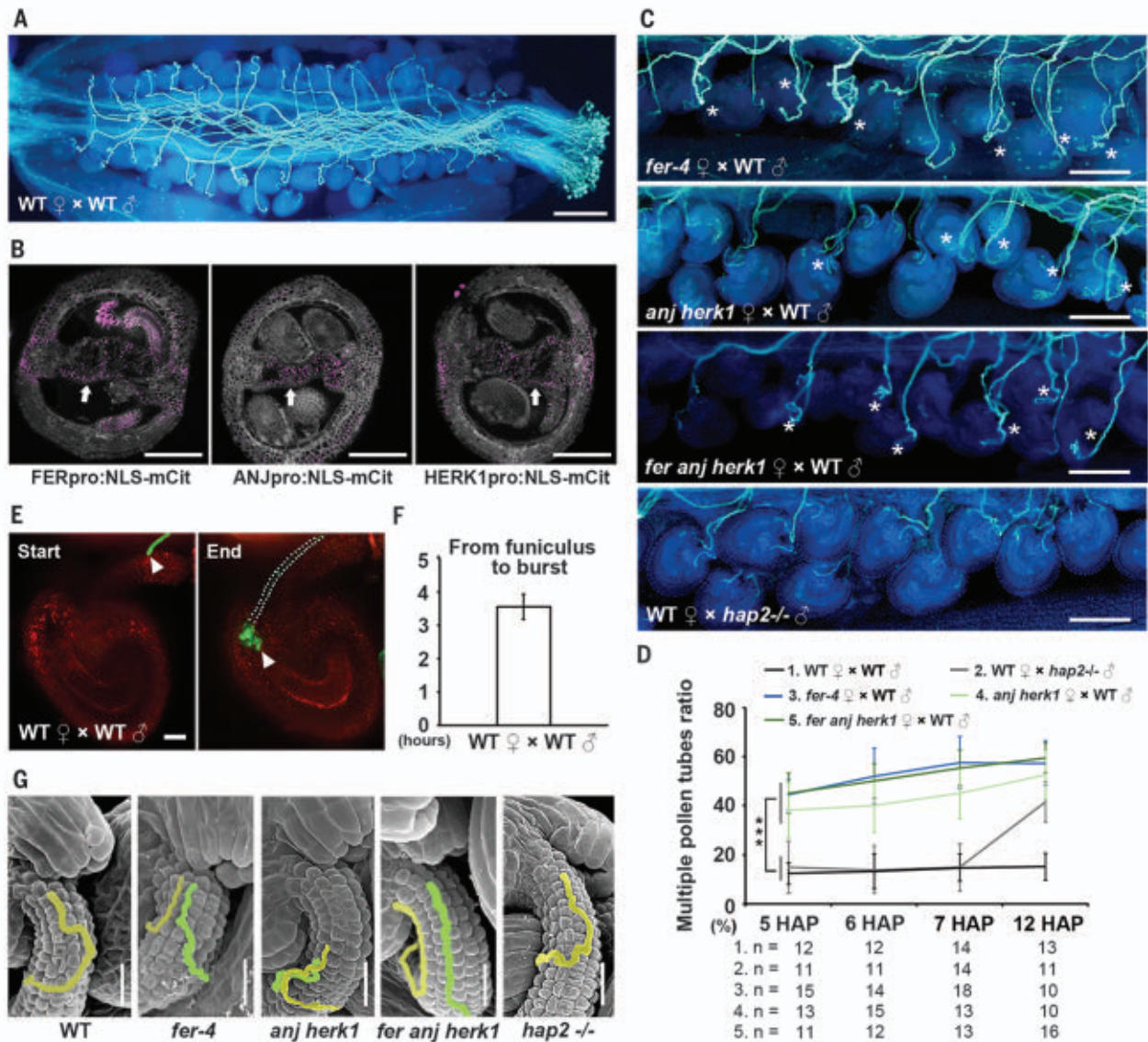
To identify factors that may establish the polytubey block, we conducted RNA sequencing (RNA-seq) analysis using transmitting tract and septum tissues and searched for candidate receptors that may perceive signals from the pollen tube. We found that seven malectin-like domain-containing receptor-like kinase (MLD-RLK) (also known as *Catharanthus roseus* RLK1-LIKE or CrRLK1L) genes were highly expressed [verified by real-time quantitative polymerase chain reaction (qPCR); fig. S1, A and B]. Transcriptional and translational markers showed that three of these genes—FERONIA (FER), ANJEA (ANJ), and HERCULES RECEPTOR KINASE 1 (HERK1)—were expressed in the ovule, transmitting tract, and septum epidermis (Fig. 1B and fig. S2). We further

<sup>1</sup>State Key Laboratory for Protein and Plant Gene Research, Peking-Tsinghua Center for Life Sciences at College of Life Sciences, Peking University, Beijing 100871, People's Republic of China. <sup>2</sup>Cell Biology and Plant Biochemistry, University of Regensburg, 93053 Regensburg, Germany. <sup>3</sup>College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, People's Republic of China. <sup>4</sup>Department of Biochemistry and Molecular Biology, Molecular and Cell Biology Program, Plant Biology Program, University of Massachusetts, Amherst, MA 01003, USA. <sup>5</sup>The National Plant Gene Research Center (Beijing), Beijing 100101, People's Republic of China. <sup>6</sup>The Waksman Institute of Microbiology, Rutgers the State University of New Jersey, Piscataway, NJ 08854, USA.

\*Corresponding author. Email: [quj@pku.edu.cn](mailto:quj@pku.edu.cn)

†These authors contributed equally to this work.

‡Present address: Institute of Science and Technology Austria, Klosterneuburg 3400, Austria.



**Fig. 1. FER, ANJ, and HERK1 control the polytubey block.** (A) A WT pistil pollinated by WT pollen showing that each ovule is targeted by a single pollen tube at 12 HAP. Scale bar, 200  $\mu$ m. (B) Cross sections of pistils showing promoter:reporter expression pattern of *FER*, *ANJ*, and *HERK1*. Arrows indicate the septum epidermal layer. Scale bars, 100  $\mu$ m. NLS-mCit, mCitrine with nuclear localization sequence. (C) Aniline blue staining showing multiple pollen tubes emerging at the septum at 5 HAP in *fer-4*, *anj herk1*, and *fer anj herk1* pistils pollinated with WT pollen, but rarely in WT pistils pollinated with WT pollen and *hap2*<sup>-/-</sup> pollen. Asterisks indicate multiple pollen tubes. The analysis was repeated at least three times. Scale bars, 100  $\mu$ m.

(D) Statistical analysis of multiple pollen tube ratios, as indicated. “n” refers to the number of pistils. Data are mean values  $\pm$  SDs; \*\*\**P* < 0.01 (Student’s *t* test).

(E) Semi-in vivo assay to determine the duration time of pollen tube growth along the funiculus until rupture in the embryo sac (white arrow heads indicate pollen tubes labeled with LAT52pro:GFP). Dashed dots show position of the pollen tube. Scale bar, 20  $\mu$ m. GFP, green fluorescent protein. (F) Statistics of (E). Data are mean values  $\pm$  SDs. (G) Scanning electron microscopy images of ovules of WT pistils pollinated with WT pollen and *hap2*<sup>-/-</sup> pollen and *fer-4*, *anj herk1*, and *fer anj herk1* pistils pollinated with WT pollen at 5 HAP. Scale bars, 20  $\mu$ m.

realized that a polytubey phenotype can be (i) caused by the failure of establishing the block at the septum or (ii) triggered by fertilization recovery at the later stage. To distinguish between these polytubey phenotypes, we exploited the *hap2/gcs1* mutant that is defective in gamete fusion and triggers fertilization recovery (8, 9). By CRISPR-Cas9, we obtained the loss-of-function mutant *hap2*<sup>-/-</sup> (fig. S3) and determined that, when depositing *hap2*<sup>-/-</sup>

pollen on wild-type (WT) pistils, the polytubey phenotype occurs at ~7 hours after pollination (HAP) (Fig. 1, C and D). To determine when the polytubey block occurs at the septum, we observed the growth behavior of WT pollen tubes in semi-in vivo assays (10). It takes a pollen tube  $3.5 \pm 0.4$  hours (*n* = 3 repeats of three to five pollen tubes each) to grow along the funiculus into the ovule and burst (Fig. 1, E and F). Therefore, pollen tubes would emerge

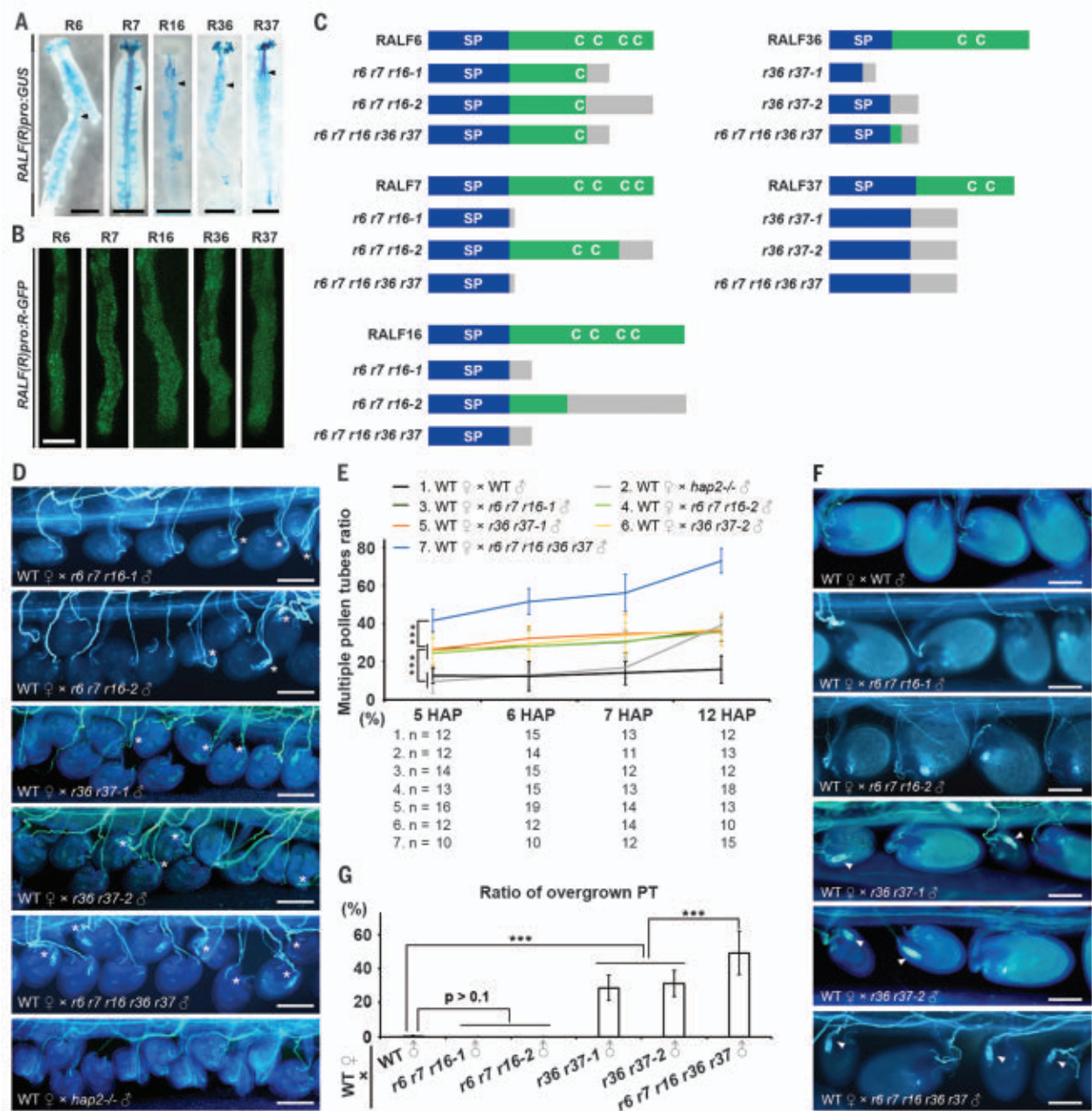
in vivo from the septum at ~4 HAP. Accordingly, a polytubey phenotype at the septum should be detectable at 4 HAP in vivo. Thus, we examined the emergence of polytubey at 5, 6, and 7 HAP, contrasting with 12 HAP, that has previously been used for characterizing *fer-4* (11, 12) and/or *anj herk1* mutants (13). Aniline blue staining showed that at 5 to 7 HAP, multiple pollen tubes emerged with comparable growth rates in the receptor mutants *fer-4* (11),



*anj herk1*, and *fer anj herk1* (13). This was much more frequent compared with the *hap2*<sup>-/-</sup> mutant (Fig. 1, C and D, and figs. S4 and S5A). Scanning electron microscopy demonstrated that at 5 HAP, multiple pollen tubes

grew at the funiculus of *fer-4*, *anj herk1*, or *fer anj herk1* ovules. This was rarely observed in WT or *hap2*<sup>-/-</sup> (Fig. 1G). The comparable polytubey ratios of *fer anj herk1* and *fer-4* indicated that FER and/or ANJ and HERK1

receptors may form a receptor complex required for establishing the polytubey block at the septum, which occurs much earlier than the initiation of fertilization recovery in the ovule (6, 7).



**Fig. 2. Pollen-specific RALF6, 7, 16, 36, and 37 peptides control the polytubey block and pollen tube reception.** (A) Promoter:GUS plants of five RALF genes show GUS signals in the pollen tubes (black arrowheads). Scale bars, 500 μm. (B) GFP signals in pollen tubes of plants containing RALF-GFP fusion proteins expressed by their native promoter. Scale bar, 10 μm. (C) Schematic diagram of RALF6, 7, 16, 36, and 37 peptide structures showing the positions of conserved cysteine residues and their CRISPR-Cas9-edited mutant structures in *ralf6 ralf7 ralf16* triple (abbreviated as *r6 r7 r16*), *ralf36 ralf37* double (*r36 r37*), and *ralf6 ralf7 ralf16 ralf36 ralf37* quintuple (*r6 r7 r16 r36 r37*) mutants. SP, signal peptide; C, conserved cysteine residue. Gray boxes indicate missense sequences resulting from frame shift mutations. (D) Aniline blue staining

showing multiple pollen tubes emerging at the septum at 5 HAP in WT pistils pollinated with *ralf* triple, double, and quintuple mutant and *hap2*<sup>-/-</sup> mutant pollen. Asterisks indicate multiple pollen tubes. The analysis was repeated at least three times. Scale bars, 100 μm. (E) Statistical analysis of multiple pollen tube emergence as observed in WT and (D). "n" refers to the number of pistils. (F) Aniline blue staining showing pollen tube overgrowth at the micropyle (white arrowheads) in WT pistils pollinated with WT pollen and *ralf* triple, double, and quintuple mutant pollen at 48 HAP. Arrowheads indicate overgrown pollen tubes. The analysis was repeated at least three times. Scale bars, 100 μm. (G) Statistical analysis of the pollen tube (PT) overgrowth phenotype in pistils shown in (F). Data are mean values ± SDs; \*\*\**P* < 0.01 (Student's *t* test).



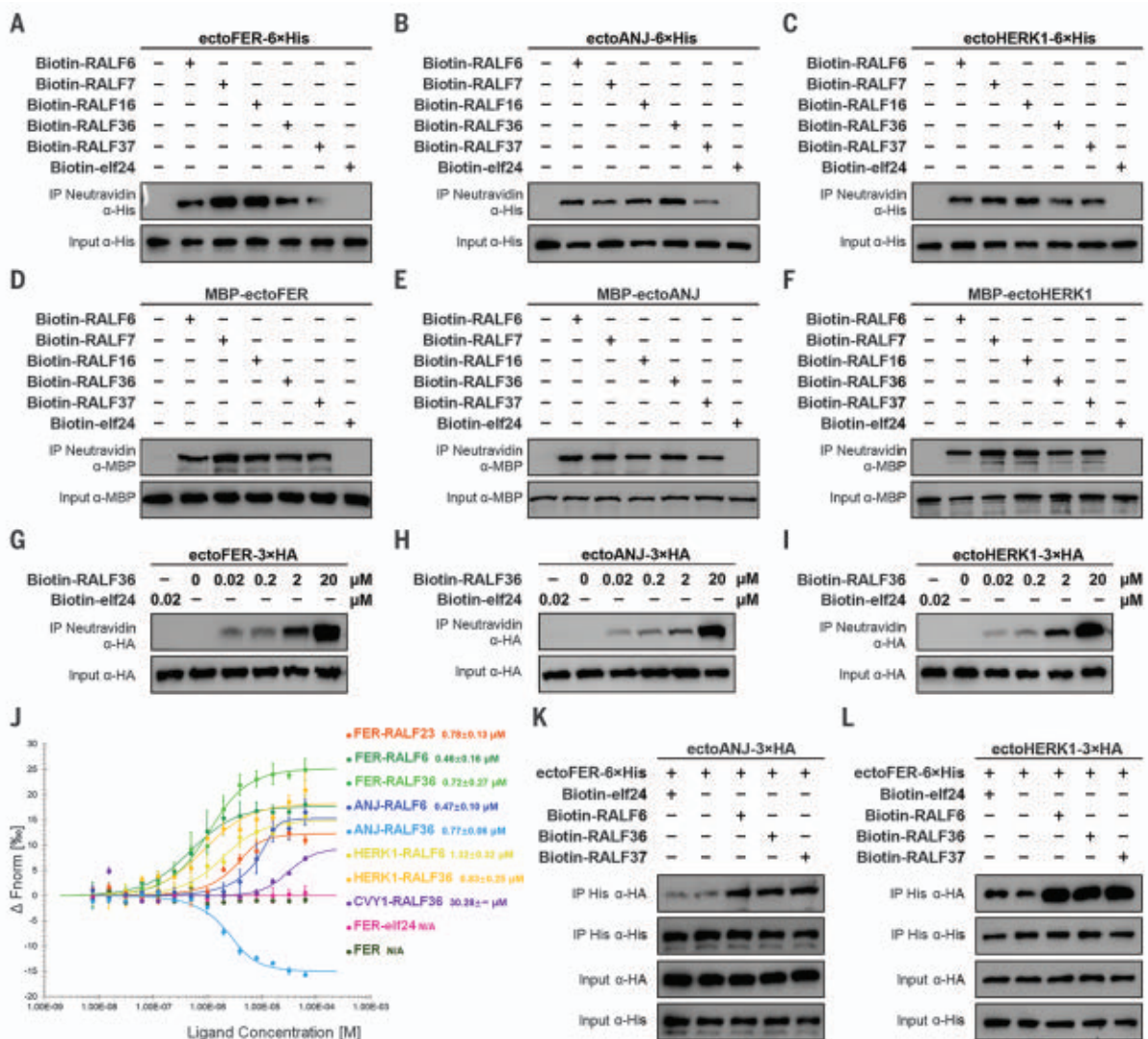
### Five RALF peptides from pollen tubes trigger the polytubey block

Peptides of the rapid alkalization factor (RALF) family function upstream of MLD-RLK receptors in plant development, immunity response, pollen tube perception, and rupture (12–17). To identify candidate RALF peptides involved in establishing the polytubey block, we examined the pollen-specific MYB transcription factor mutant, *myb97 myb101 myb120*, which showed similar pollen tube reception defects as those observed in *fer-4* and *anj herk1* mutants (18, 19). We found that the emergence of multiple pollen tubes can be observed at 5 to 7 HAP in WT pistils when pollinated by *myb* triple mutant pollen (fig. S6),

which indicates the compromised polytubey block at the septum. We thus conducted RNA-seq analysis of *myb* mutant pollen tubes and identified five RALF genes with expression levels that were lowered or absent in the *myb* triple mutant (fig. S7). They are *RALF6* (At1g60625), *RALF7* (At1g60815), *RALF16* (At2g32835), and two noncanonical *RALF36* (At2g32785) and *RALF37* (At2g32788) (20), which cluster into two subclades (fig. S7). Their expression in pollen and pollen tubes was further confirmed (Fig. 2, A and B).

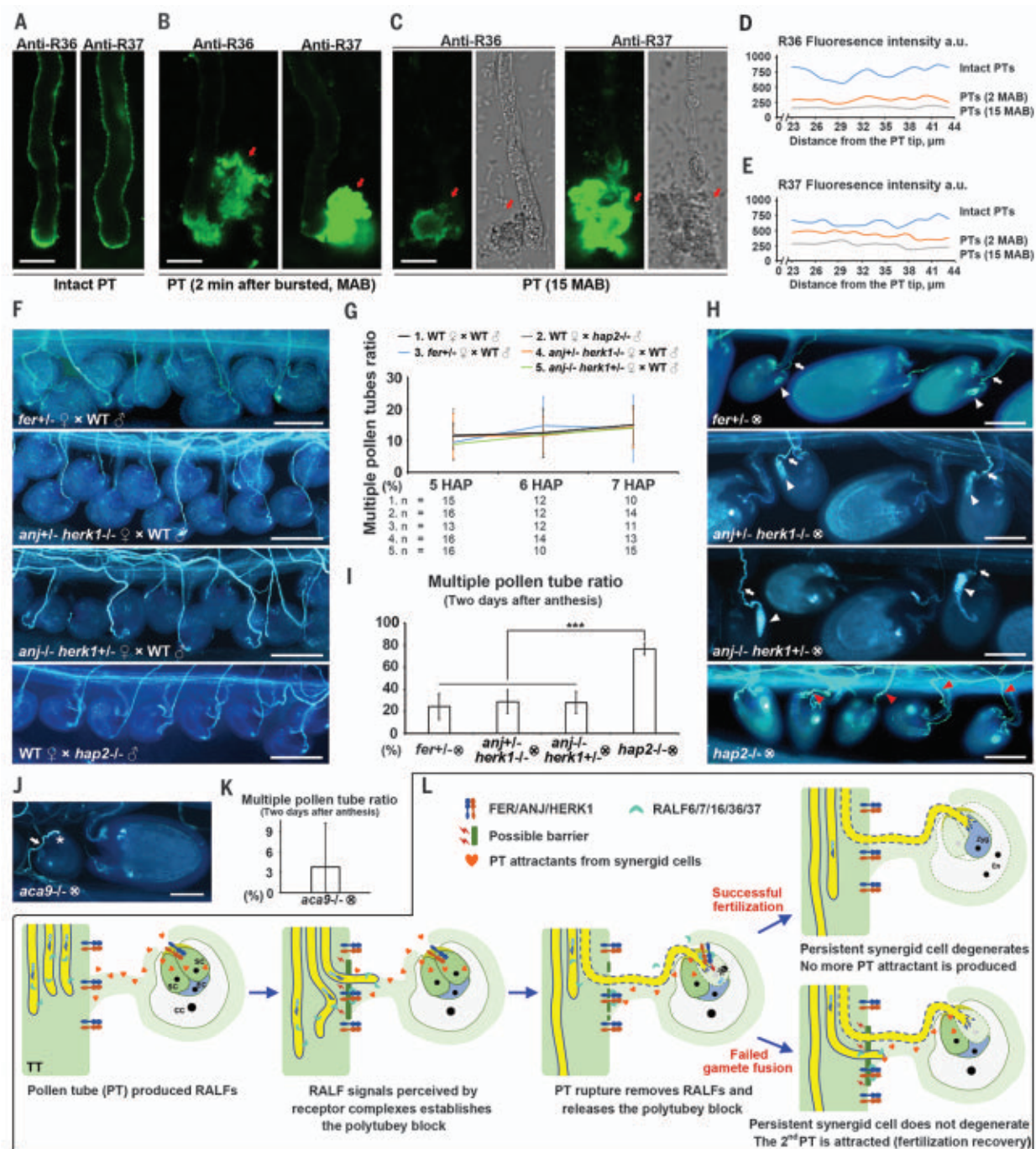
Next, we used CRISPR-Cas9 to generate *ralf36 ralf37* double, *ralf6 ralf7 ralf16* triple, and *ralf6 ralf7 ralf16 ralf36 ralf37* quintuple mutants. They all showed normal vegetative growth be-

havior (Fig. 2C and figs. S8 and S9). When WT pistils were pollinated with mutant pollen, all three *ralf* mutants showed the polytubey phenotype at 5 to 7 HAP (Fig. 2, D and E, and fig. S5B), resembling what was observed in *fer-4*, *anj herk1*, or *fer anj herk1* mutant pistils pollinated by WT pollen. The higher polytubey ratio caused by *ralf* quintuple than by *ralf* double or triple mutations suggests that the five RALF genes function collectively to establish a polytubey block. Therefore, we hypothesized that the five RALF peptides are pollen tube-produced signal molecules that are likely perceived by the FER-ANJ-HERK1 receptor complex at the septum to establish a polytubey block.



**Fig. 3. FER, ANJ, and HERK1 receptors interact with RALF6, 7, 16, 36, and 37 peptide ligands.** (A to F) Pull-down assays between 6×His-tagged [(A) to (C)] and MBP-tagged [(D) to (F)] ectodomains of FER, ANJ, and HERK1 and biotinylated RALF6, 7, 16, 36, and 37 and elf24. (G to I) Pull-down assay between HA-tagged FER, ANJ, and HERK1 ectodomains and biotinylated RALF36

obtained by elevating the concentration of RALF36. (J) Binding affinity as indicated by MST analysis. CVY1 and RALF36, FER and elf24, and FER only were used as controls. (K and L) Interaction between 6×His-tagged ectodomains of FER and HA-tagged ectodomains of ANJ and HERK1 by pull-down assays in the presence or absence of RALF6, 36, and 37 peptides.



**Fig. 4. Pollen tube rupture removes the polytubey block and coordinates fertilization recovery.** (A) Immunofluorescence imaging of RALF36 and 37 on the surface of intact pollen tubes. (B and C) Shortly after pollen tube burst (2 min), RALF36- and RALF37-derived fluorescence signal on the pollen tube surface decreased substantially (B) and could scarcely be detected at 15 min (C). Cytosolic RALF36 and 37 (red arrows) were released after pollen tube burst at the tip. The analysis was repeated at least three times. Scale bars, 10  $\mu$ m. (D and E) Quantification of immunofluorescence intensity along the shank of the pollen tube surface at stages shown in (A) to (C). Intact pollen tubes and pollen tubes 2 and 15 min after burst (MAB) were measured ( $n = 17, 18$ , and 23 for anti-R36, and  $n = 22, 20$ , and 18 for anti-R37). a.u., arbitrary units. (F) Aniline blue staining of pollen tube emerging status in *fer+/-*, *anj+/- herk1-/-*, and *anj-/- herk1+/-* pollinated with WT pollen and WT pistil pollinated with *hap2-/-* pollen at 5 HAP. The analysis was repeated at least three times. Scale bars, 100  $\mu$ m. (G) Statistical analysis of the multiple pollen tube ratio at

stages indicated in pistils shown in WT and (F). "n" refers to the number of pistils. (H) Aniline blue staining of self-crossed *fer+/-* and *hap2-/-* mutants 2 days after anthesis. White arrowheads indicate overgrown pollen tubes, white arrows mark single pollen tubes in the self-crossed *fer+/-* mutant, and red arrowheads point toward multiple pollen tubes in the self-crossed *hap2-/-* mutant. The analysis was repeated at least three times. Scale bars, 100  $\mu$ m. (I) Statistical analysis of multiple pollen tubes ratio in pistils shown in (H). (J) Aniline blue staining of self-crossed *aca9-/-* mutant 2 days after anthesis. White arrow indicates the single pollen tube, and white asterisk indicates blocked pollen tube. Scale bar, 100  $\mu$ m. (K) Statistical analysis of multiple pollen tubes ratio in pistils shown in (J). (L) Model of how pollen-specific RALFs act on female tissue located at FER, ANJ, and HERK1 receptor complexes to jointly coordinate establishment, maintenance, and removal of the polytubey block during fertilization progression. TT, transmitting tract; SC, synergid cell; EC, egg cell; CC, central cell; Zyg, zygote; En, endosperm. Data are mean values  $\pm$  SDs; \*\*\* $P < 0.01$  (Student's  $t$  test).



Because the FER-ANJ-HERK1 receptor complex also controls pollen tube reception at the micropyle (12–14, 21), we next investigated whether the five RALF peptides are also required for pollen tube reception. We observed that in WT pistils pollinated by *ralf* double and quintuple mutant pollen grains, higher ratios of pollen tube overgrowth (failures in reception) occurred at the micropyle [ $28.7 \pm 7.4\%$  ( $n = 30$ ),  $31.1 \pm 7.5\%$  ( $n = 30$ ), and  $49.2 \pm 12.8\%$  ( $n = 22$ ) for *ralf36 ralf37-1*, *ralf36 ralf37-2*, and the *ralf* quintuple mutant, respectively] (Fig. 2, F and G). Consistently, fertility analysis showed that male defects led to obviously reduced fertility in the *ralf* double mutants ( $70.8 \pm 7.8\%$ ,  $n = 30$ , and  $68.5 \pm 7.7\%$ ,  $n = 30$ , respectively) and more severely in the *ralf* quintuple mutant ( $57.1 \pm 10.9\%$ ,  $n = 22$ ,  $P < 0.01$  versus *ralf36 ralf37* mutants) (fig. S10). These defects were similar to those observed in *fer-4*, *anj herk1*, or *fer anj herk1* mutants (fig. S11) (12–14). These findings suggest that RALF6, 7, 16, 36, and 37 are likely the long-pursued ligands of FER and ANJ-HERK1 receptors required for pollen tube reception (12, 13). As both the establishment of a polytubey block at the septum and pollen tube reception in the ovule appear to require the same signaling components, an in-depth mechanistic study on their precisely controlled interaction was necessary.

#### FER, ANJ, and HERK1 physically interact with RALF6, 7, 16, 36, and 37 peptide ligands

In vitro pull-down assays showed that biotinylated RALF6, 7, 16, 36, and 37 bind to 6×His-tagged ectodomains of FER, ANJ, and HERK1 purified from insect cells or corresponding maltose-binding protein (MBP)-tagged ectodomains purified from *Escherichia coli* (Fig. 3, A to F). These interactions were strengthened in a peptide dose-dependent manner (Fig. 3, G to I). Microscale thermophoresis (MST) analysis further revealed that both canonical RALF6 and noncanonical RALF36 interact with FER, ANJ, and HERK1 with low equilibrium dissociation constants ( $K_d$ ) (Fig. 3J), which demonstrates that these RALFs are ligands of FER, ANJ, and HERK1. Moreover, in vitro pull-down assays showed that the addition of RALF6, 36, and 37 promoted the interactions between FER and ANJ-HERK1 receptors (Fig. 3, K and L), which suggests that the RALF6, 36, and 37 peptides may facilitate the formation of larger FER-ANJ-HERK1 heteromeric receptor complexes.

#### Pollen tube rupture releases the polytubey block and coordinates fertilization recovery

Because the emergence of secondary pollen tubes ensures reproductive success in cases of fertilization failure, the mechanism regulating the polytubey block has to be adjustable. We therefore investigated when and how the polytubey block is removed. After perception by the receptive synergid, the pollen tube bursts to

release two sperm cells for double fertilization within ~20 min (4, 22–26). FER-mediated pollen tube reception and production of reactive oxygen species (ROS) in the filiform apparatus region of the ovule have been shown to be required for pollen tube rupture (21). We hypothesized that the presence of RALF peptides is required not only for establishing but also for maintaining the polytubey block at the septum. Thus, pollen tube rupture that naturally terminates the production of RALF peptides would weaken or remove the polytubey block. To test this hypothesis, by using antibodies against RALF36 and 37, we detected immunofluorescence signals in the cell wall of the whole shank region in intact WT pollen tubes but not in the *ralf36 ralf37* mutant (Fig. 4, A, D, and E, and fig. S12). This indicates that the polytubey block can be maintained as long as the pollen tube grows inside the ovule. However, the intensity of RALF36 and 37 signals in the cell wall declined rapidly 2 min after pollen tube rupture (Fig. 4, B, D, and E) and was no longer detectable after 15 min (Fig. 4, C, D, and E), whereas released pollen tube content showed strong immunofluorescence (Fig. 4, B and C). Thus, pollen tube rupture results in the loss of RALF peptides that are required to maintain the polytubey block at the septum and allows secondary pollen tubes to exit the septum.

To genetically test this hypothesis, we examined the receptor mutants *fer+/-* (11), *anj-/- herk1+/-*, and *anj+/- herk1-/-* (13). Like WT, these heterozygous mutants should be able to establish and maintain the polytubey block at the septum. Compared with the *hap2-/-* control, none of the three heterozygous mutant pistils exhibited significant emergence of multiple WT pollen tubes at 5 to 7 HAP (Fig. 4, F and G, and fig. S5C). Absence of FER or ANJ-HERK1 in synergids was previously reported to result in failure of pollen tube rupture and impaired fertility, and pollen tube overgrowth can be easily visualized in the mutant ovule (12, 13, 21). We therefore investigated whether selfed heterozygous receptor mutants would show a polytubey phenotype that is triggered by fertilization recovery. Two days after anthesis, in those ovules with failed events of pollen tube rupture (i.e., pollen tube overgrowth), only low levels of polytubey were observed [*fer+/-*,  $24.3 \pm 11.8\%$  ( $n = 28$ ); *anj-/- herk1+/-*,  $28.6 \pm 10.5\%$  ( $n = 16$ ); and *anj+/- herk1-/-*,  $27.9 \pm 9.8\%$  ( $n = 21$ )], which were much lower than that of the *hap2-/-* mutant ( $76.4 \pm 5.2\%$ ,  $n = 12$ ) (Fig. 4, H and I), indicating that the removal of the polytubey block required for fertilization recovery was compromised. We next investigated another mutant, *aca9* [autoinhibited calcium adenosine triphosphatase (ATPase) 9], which showed defects in pollen tube rupture in the ovule (27) (Fig. 4J). Like the receptor mutants, *aca9* did not produce an increased level of polytubey 2 days

after anthesis [ $3.8 \pm 6.4\%$  ( $n = 20$ )] (Fig. 4, J and K). This further confirms that the removal of the polytubey block at the septum depends on pollen tube rupture in the embryo sac. Taken together, the FER, ANJ, and HERK1 receptor complexes not only mediate pollen tube-synergid recognition and subsequent pollen tube rupture in the embryo sac (12, 13), but also function to trigger the removal of the polytubey block at the septum for fertilization recovery.

#### Discussion

Here, we have elucidated a molecular mechanism of how FER, ANJ, and HERK1 receptor complexes located at the septum interact with pollen tube-produced RALF6, 7, 16, 36, and 37 peptide ligands to coordinately establish, maintain, and terminate the polytubey block and thus regulate the emergence of pollen tubes to ultimately prevent polyspermy and to ensure reproductive success. On the basis of this and previous studies, we suggest the following model (Fig. 4L): (i) Pollen tube attractants secreted from the synergid cells trigger the nearest pollen tube to exit the transmitting tract. (ii) RALFs secreted from this pollen tube activate FER, ANJ, and HERK1 signaling in septum epidermal cells to establish the polytubey block that prevents the emergence of additional pollen tubes. This male-female jointly established polytubey block remains activated during pollen tube growth into the ovule as a result of the continuous production of RALF peptides by the first-emerged pollen tube. (iii) After successful recognition by the same receptor complex (FER-ANJ-HERK1) in synergid cells, the pollen tube ruptures to release two sperm cells, fertilization will be completed within 20 min (21–26), and the polytubey block is then removed as RALFs quickly disappear from the ruptured pollen tube. (iv) Once fertilization is successful, the persistent synergid cell undergoes fertilization-dependent PCD and fuses with the fertilized central cell. Pollen tube attractants are dispersed, modified, and degraded (11, 28–30), which reduces the attraction of further pollen tubes from the septum despite the release of the polytubey block. Polyspermy is thus prevented. (v) When gamete fusion fails, the persistent synergid cell remains alive and continues to produce pollen tube attractants. Because the polytubey block is removed shortly after pollen tube rupture, secondary pollen tubes are able to emerge from the septum to salvage fertilization. The secondary pollen tube reestablishes the polytubey block, which also explains the low rate of tertiary pollen tubes (6, 23, 30).

This study demonstrates how *Arabidopsis* regulates the emergence of a single pollen tube at the septum to target each ovule and how fertilization success and recovery are interconnected with the activity of the same receptor



complexes. It will now be essential to elucidate the downstream mechanism or mechanisms that block the emergence of secondary pollen tubes at the septum. Further components of this polytubey block may include nitric oxide (11); secretion of cell wall components; and ROS that mediate FER-controlled pollen tube rupture (21), pollen hydration (31), and self-incompatibility (32).

## REFERENCES AND NOTES

1. D. P. Wolf, *Dev. Biol.* **64**, 1–10 (1978).
2. J. P. Evans, *Mol. Reprod. Dev.* **87**, 341–349 (2020).
3. J. Zhang *et al.*, *Nat. Plants* **3**, 17079 (2017).
4. S. Zhong, L.-J. Qu, *Curr. Opin. Plant Biol.* **51**, 7–14 (2019).
5. S. Zhong *et al.*, *Science* **364**, eaau9564 (2019).
6. K. M. Beale, A. R. Leydon, M. A. Johnson, *Curr. Biol.* **22**, 1090–1094 (2012).
7. R. D. Kasahara *et al.*, *Curr. Biol.* **22**, 1084–1089 (2012).
8. T. Mori, H. Kuroiwa, T. Higashiyama, T. Kuroiwa, *Nat. Cell Biol.* **8**, 64–71 (2006).
9. K. von Besser, A. C. Frank, M. A. Johnson, D. Preuss, *Development* **133**, 4761–4769 (2006).
10. R. Palanivelu, D. Preuss, *BMC Plant Biol.* **6**, 7 (2006).
11. Q. Duan *et al.*, *Nature* **579**, 561–566 (2020).
12. J. M. Escobar-Restrepo *et al.*, *Science* **317**, 656–660 (2007).
13. S. Galindo-Trigo *et al.*, *EMBO Rep.* **21**, e48466 (2020).
14. Z. Ge, T. Dresselhaus, L.-J. Qu, *Trends Plant Sci.* **24**, 978–981 (2019).
15. M. Haruta, G. Sabat, K. Stecker, B. B. Minkoff, M. R. Sussman, *Science* **343**, 408–411 (2014).
16. C. Li *et al.*, *eLife* **4**, e06587 (2015).
17. M. Stegmann *et al.*, *Science* **355**, 287–289 (2017).
18. A. R. Leydon *et al.*, *Curr. Biol.* **23**, 1209–1214 (2013).
19. Y. Liang *et al.*, *PLoS Genet.* **9**, e1003933 (2013).
20. A. Abarca, C. M. Franck, C. Zipfel, *Plant Physiol.* **187**, 996–1010 (2021).
21. Q. Duan *et al.*, *Nat. Commun.* **5**, 3129 (2014).
22. Y. Hamamura *et al.*, *Curr. Biol.* **21**, 497–502 (2011).
23. S. Sprunck *et al.*, *Science* **338**, 1093–1097 (2012).
24. P. Denninger *et al.*, *Nat. Commun.* **5**, 4645 (2014).
25. Y. Hamamura *et al.*, *Nat. Commun.* **5**, 4722 (2014).
26. T. Dresselhaus, S. Sprunck, G. M. Wessel, *Curr. Biol.* **26**, R125–R139 (2016).
27. M. Schiøtt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9502–9507 (2004).
28. R. Völz, J. Heydlauff, D. Ripper, L. von Lyncker, R. Groß-Hardt, *Dev. Cell* **25**, 310–316 (2013).
29. D. Maruyama *et al.*, *Cell* **161**, 907–918 (2015).
30. X. Yu *et al.*, *Nature* **592**, 433–437 (2021).
31. C. Liu *et al.*, *Science* **372**, 171–175 (2021).
32. L. Zhang *et al.*, *Curr. Biol.* **31**, 3004–3016.e4 (2021).

## ACKNOWLEDGMENTS

We thank D. Ye for providing *fer-4* and *myb97 myb101 myb120* mutant seeds; L. Smith for sharing *anj*, *herk1*, *anj herk1*, and *fer anj herk1* mutant seeds; J. F. Harper for providing *aca9* mutant seeds; and C. Li and Q. Duan for sharing *fer+/-* mutant seeds. **Funding:** L.-J.Q. was funded by the National Natural Science Foundation of China (grant nos. 31991202, 31830004, 31620103903, and 31621001), S.Z. was supported by the Young Elite Scientists Sponsorship Program by the China Association of Science and Technology (2019QNRC001), Z.G. was supported by a NSFC Young Scientists Fund (31900161), A.Y.C. was funded by the US Natural Science Foundation (IOS-1645854, MCB-1715764, and MCB-0955910), J.D. was funded by the National Institute of Health (RO1GM109080), and T.D. was supported by the German Research Foundation DFG (SFB924). **Author contributions:** S.Z. and L.-J.Q. conceived the project, and L.-J.Q. and H.G. supervised the project. S.Z., LiLi, Z.W., Z.G., and T.L. performed molecular cloning and CRISPR-Cas9-mediated mutant generation. S.Z., LiLi, and Z.W. performed phenotype observation and statistical analysis. LiLi and Z.W. analyzed the GUS activity with the help of Z.S., LuLi, and L.Z. S.Z., Z.G., and Y.S. performed RNA-seq analysis. Z.G., Q.L., and J.W. performed protein expression, protein purification, and all the protein-protein interaction assays with the help of H.Z., Y.W., and LuLa. A.B. and T.D. conducted receptor reporter localization assays. S.Z., J.D., H.-M.W., A.Y.C., T.D., and L.-J.Q. drafted the manuscript. All authors contributed to data analysis

and manuscript preparation. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl4683  
Materials and Methods  
Figs. S1 to S12

Tables S1 to S3  
References (33–47)  
MDAR Reproducibility Checklist

15 July 2021; accepted 30 November 2021  
10.1126/science.abl4683

## REPORTS

## MULTIPLE SCLEROSIS

# Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis

Kjetil Bjornevik<sup>1†</sup>, Marianna Cortese<sup>1†</sup>, Brian C. Healy<sup>2,3,4</sup>, Jens Kuhle<sup>5</sup>, Michael J. Mina<sup>6,7,8</sup>, Yumei Leng<sup>6</sup>, Stephen J. Elledge<sup>6</sup>, David W. Niebuhr<sup>9</sup>, Ann I. Scher<sup>9</sup>, Cassandra L. Munger<sup>1†</sup>, Alberto Ascherio<sup>1,10,11\*†</sup>

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system of unknown etiology. We tested the hypothesis that MS is caused by Epstein-Barr virus (EBV) in a cohort comprising more than 10 million young adults on active duty in the US military, 955 of whom were diagnosed with MS during their period of service. Risk of MS increased 32-fold after infection with EBV but was not increased after infection with other viruses, including the similarly transmitted cytomegalovirus. Serum levels of neurofilament light chain, a biomarker of neuroaxonal degeneration, increased only after EBV seroconversion. These findings cannot be explained by any known risk factor for MS and suggest EBV as the leading cause of MS.

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system of unknown etiology. The demyelination in the brain and spinal cord is an immune-mediated process (1) possibly triggered by a viral infection (2). Among the putative causal agents, the top candidate is Epstein-Barr virus (EBV) (3). EBV is a human herpesvirus that after infection persists in latent form in B lymphocytes throughout the life of the host (3). A causal role of EBV is supported by the increased MS risk after infectious mononucleosis (4), elevated serum antibody titers against EBV nuclear antigens (EBNAs) (5), and by the presence of EBV in MS demyelinated lesions reported in some (6–8), but not all (9), pathological studies. Evidence of causality, however, remains inconclusive.

Causality implies that some individuals who developed MS after EBV infection would not have developed MS if they had not been infected with EBV. Ruling out a randomized trial, the gold standard to study this counterfactual occurrence is an “experiment of nature,” a longitudinal investigation of MS incidence in a cohort of EBV-negative individuals, some of whom will be infected with EBV during the follow-up and some who will not. The ubiquitous nature of EBV, which infects ~95% of adults, and the fact that MS is a relatively rare disease, has until now impeded such an investigation. Over the course of a 20-year collaboration with the US military, we have identified cases of MS in a cohort composed of active-duty US military personnel between 1993 and 2013, a racially diverse population of >10 million individuals.

<sup>1</sup>Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, MA, USA. <sup>2</sup>Partners Multiple Sclerosis Center, Brigham and Women's Hospital, Boston, MA, USA. <sup>3</sup>Department of Neurology, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Biostatistics Center, Massachusetts General Hospital, Boston, MA, USA. <sup>5</sup>Neurologic Clinic and Policlinic, MS Center and Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel, University of Basel, Basel, Switzerland. <sup>6</sup>Division of Genetics, Brigham and Women's Hospital, Howard Hughes Medical Institute, Department of Genetics, and Program in Virology, Harvard Medical School, Boston, MA, USA. <sup>7</sup>Center for Communicable Disease Dynamics, Department of Epidemiology, and Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA, USA. <sup>8</sup>Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>9</sup>Department of Preventive Medicine and Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA. <sup>10</sup>Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA. <sup>11</sup>Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA.

\*Corresponding author. Email: aascheri@hsph.harvard.edu

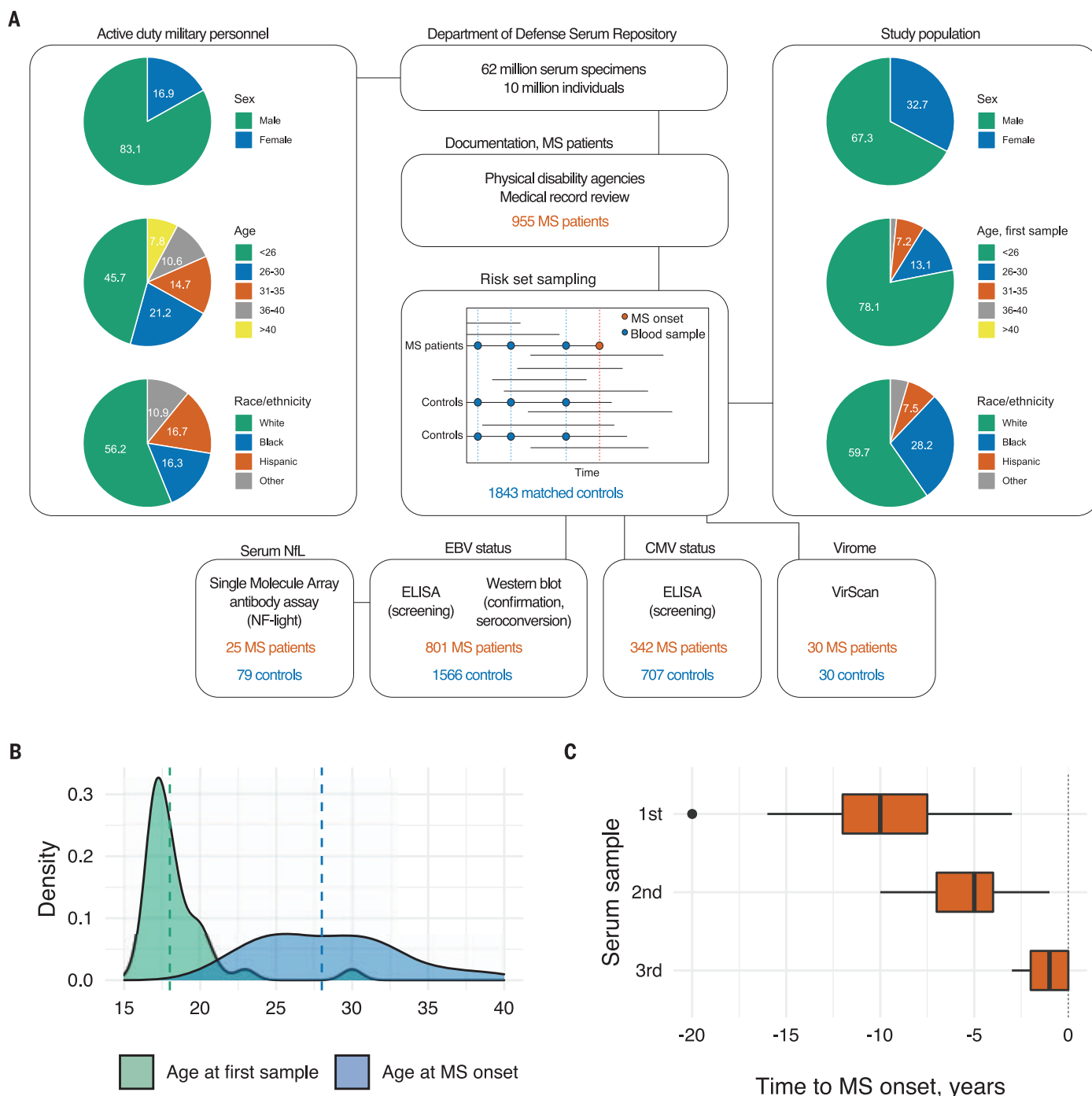
†These authors contributed equally to this work.

‡These authors contributed equally to this work.

Active-duty members are screened for HIV at the start of military service and biennially thereafter, and residual serum from these tests (>62 million serum samples) is archived in the Department of Defense Serum Repository

tory (DoDSR) (10). We used samples stored in the DoDSR to determine EBV status at time of first sample and the relation between EBV infection and MS onset during the period of active duty. In a preliminary study,

we found that 5.3% of individuals were EBV-negative at the time of first sample (11), corresponding to hundreds of thousands of EBV-negative young adults at risk of EBV infection and MS.

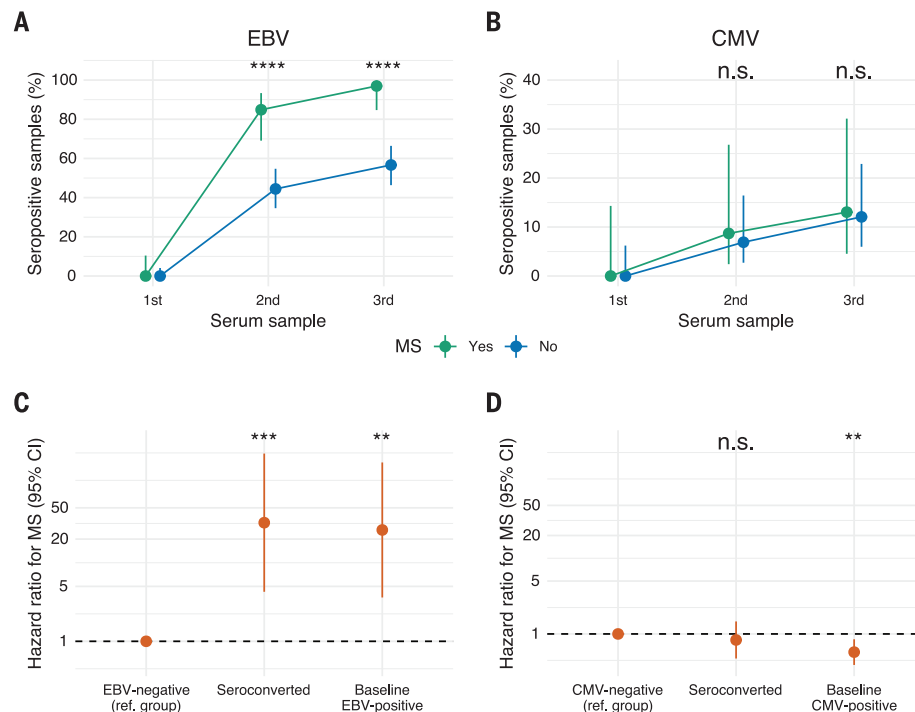


**Fig. 1. Study design.** (A) Residual serum samples from the DoDSR were obtained from 810 MS patients and 1577 matched controls. We assessed whether individuals were seropositive for EBV and CMV in up to three serum samples per person. We measured sNFL in those who were EBV-negative in the first serum sample. VirScan was used to profile the virome in a subset of MS

cases with serum samples collected shortly before and after symptom onset. (B) Density plot of age at onset among MS patients who were EBV-negative at time of first serum sample. The dashed line marks median age at onset. (C) Box plots of the time to first MS symptoms according to the serum sample.

**Fig. 2. EBV infection precedes MS onset and is associated with markedly higher disease risk.**

(A) Proportion of individuals who were EBV-positive at the time of the first, second, and third sample. The figure is restricted to those who were EBV-negative at baseline and with EBV measurement in three samples (33 of 35 MS patients and 90 of 107 controls). A significantly higher proportion of individuals who later developed MS were EBV-positive in the second (28 of 33 MS patients) and third (32 of 33 MS patients) sample compared with individuals who did not develop MS (second sample: 40 of 90 controls; third sample: 51 of 90 controls). \*\*\*\* $P < 0.0001$ , two-sided Fisher's exact test. (B) Proportion of individuals who were CMV-positive at the time of the first, second, and third sample collected in the study. The figure is restricted to those who were CMV- and EBV-negative at baseline. The proportion who were CMV-positive was similar in the second (two of 23 MS patients versus four of 60 controls) and third sample (three of 23 MS patients versus seven of 60 controls). All  $P > 0.05$ , two-sided Fisher's exact test. (C) Risk ratio for MS according to EBV status. EBV seroconversion by the time of the third sample and EBV seropositivity at the time of the first sample were associated with a 32-fold and 26-fold increased risk of developing MS, respectively, in matched analyses. \*\* $P < 0.01$  and \*\*\* $P < 0.001$ , two-sided univariable conditional logistic regression model. (D) Risk ratio for MS according to CMV status. \*\* $P < 0.01$ , two-sided univariable conditional logistic regression model.



We documented 955 incident MS cases among active-duty military personnel [including 315 cases from our preliminary study (11)]. For each MS case, we identified up to three serum samples collected before the date of MS onset (the first available, the last collected before disease onset, and one in between). Cases were matched to two randomly selected individuals without MS of the same age, sex, race/ethnicity, branch of military service, and dates of collection of blood samples who were on active military duty when the case was diagnosed (Fig. 1A and fig. S1). There were 801 MS cases and 1566 controls with samples available to assess EBV infection status. Most of the individuals in our study were <20 years of age at the time of their first blood collection (Fig. 1B), and those who developed MS had symptom onset a median of 10 years after time of first sample (Fig. 1C).

Only one of the 801 MS cases occurred in an individual who was EBV-negative in the last sample, which was collected at a median of 1 year before MS onset [hazard ratio (HR) for MS comparing EBV-positive versus EBV-negative = 26.5; 95% confidence interval (CI): 3.7 to 191.6;  $P = 0.001$ , conditional logistic regression]. At baseline, 35 MS cases and 107 controls were EBV-negative. All but one of these 35 EBV-negative MS cases became infected with EBV during the follow-up, and all seroconverted before the onset of MS (fig. S3). The median time from the first EBV-positive

sample to MS onset was 5 years (range: 0 to 10 years), and the median time from estimated EBV seroconversion, defined as the midpoint between the last seronegative sample and the first seropositive sample, to MS onset was 7.5 years (range: 2 to 15 years). The high seroconversion rate among individuals who developed MS during follow-up (97%) contrasts with the 57% rate of seroconversion observed among individuals who did not develop MS (Fig. 2A), a rate consistent with previous reports among EBV-negative young adults (12). The HR for MS comparing EBV seroconversion versus persistent EBV seronegativity was 32.4 (95% CI: 4.3 to 245.3,  $P < 0.001$ ) (Fig. 2C).

Behavioral, environmental, or personal characteristics may correlate with a predisposition to both infection and MS. To assess this possibility, we measured antibodies against cytomegalovirus (CMV), a herpesvirus that, like EBV, is transmitted through the saliva. CMV displays socioeconomic and racial/ethnic disparities in age at infection in the US population (13) similar to those of EBV (14), thus constituting an ideal negative control (15). Among those who were CMV-negative at baseline, seroconversion for CMV occurred at a similar rate in those who later developed MS and those who did not (Fig. 2B). MS risk was lower among CMV-positive than among CMV-negative individuals (Fig. 2D), consistent with a previous report and

with suggestions that the immune response to CMV attenuates the adverse effects of EBV (16).

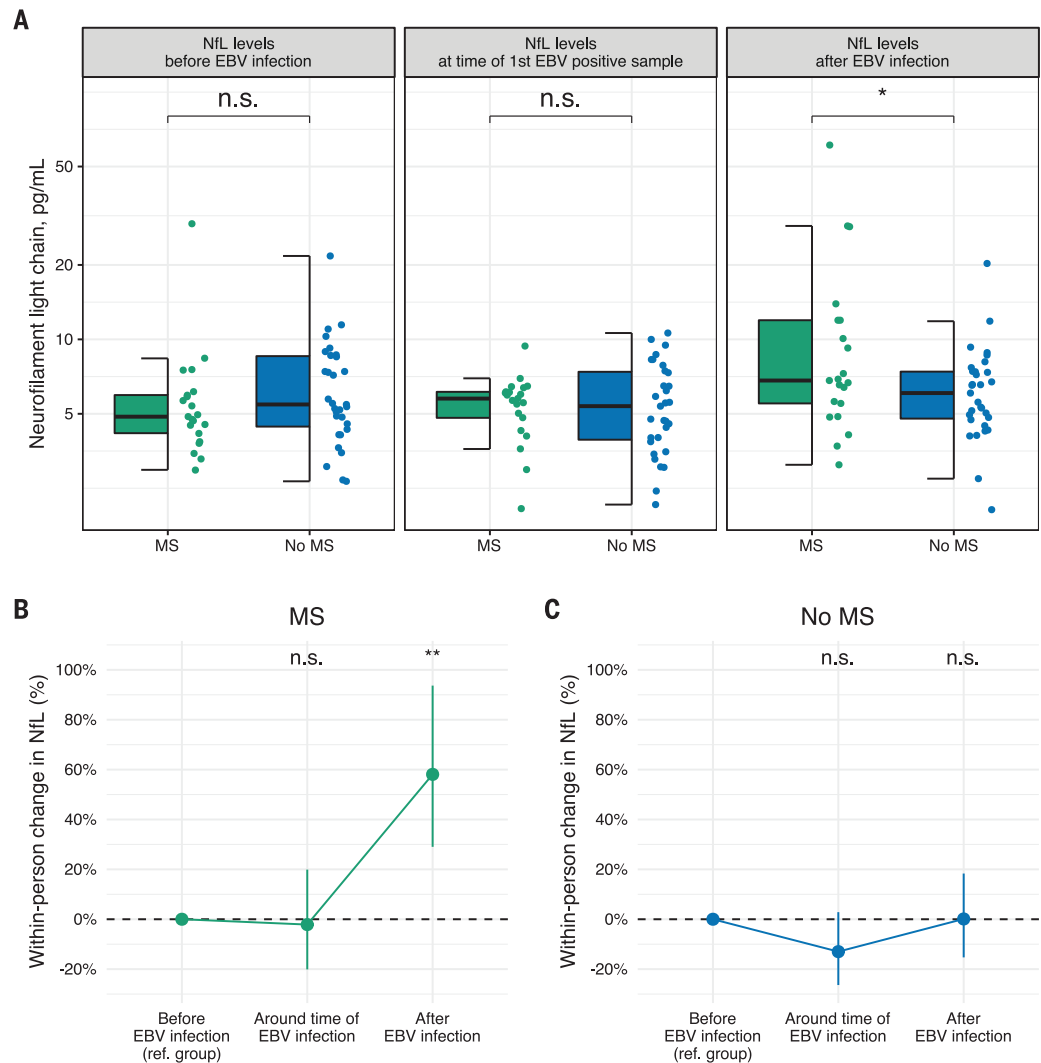
Similar to other neurological diseases, the pathological mechanisms underlying MS likely start several years before the first symptoms appear (17). To further elucidate the temporal relation between EBV infection and MS, we measured serum concentrations of neurofilament light chain (sNfL), a sensitive, albeit not disease-specific, biomarker of ongoing neuroaxonal degeneration (18), using an ultra-sensitive single-molecule assay (19) in the samples from those who were EBV-negative at baseline. We have previously reported that sNfL levels increase as early as 6 years before clinical MS onset and may be a more accurate marker of the time of initiation of the disease process (20). sNfL levels in individuals who were EBV-negative at baseline and went on to develop MS were similar to those of non-MS controls before and around the time of EBV infection but increased after EBV infection (Fig. 3, A to C, and fig. S4). Thus, there were no signs of neuroaxonal degeneration before EBV seroconversion in individuals who later developed MS, indicating that EBV infection preceded not only symptom onset but also the time of the first detectable pathological mechanisms underlying MS.

To further explore whether immune dysregulation during the preclinical phase could increase susceptibility to viral infections more



**Fig. 3. EBV infection precedes elevation of sNFL before the onset of MS.** (A) Box plots of sNFL levels before, around, and after the time of EBV infection. \* $P < 0.05$ , two-sided multivariable linear regression model adjusted for age and sex. (B) Within-person increase in sNFL levels in MS cases around and after time of EBV infection compared with before EBV infection. \*\* $P < 0.01$ , two-sided linear mixed-effects regression model.

(C) Within-person increase in sNFL levels in controls around and after time of EBV infection compared with before EBV infection. Error bars in (B) and (C) are 95% CIs. sNFL levels increased significantly more in MS cases than in controls in the sample collected after time of EBV infection compared with before EBV infection ( $P < 0.001$ , two-sided linear mixed-effects regression model).



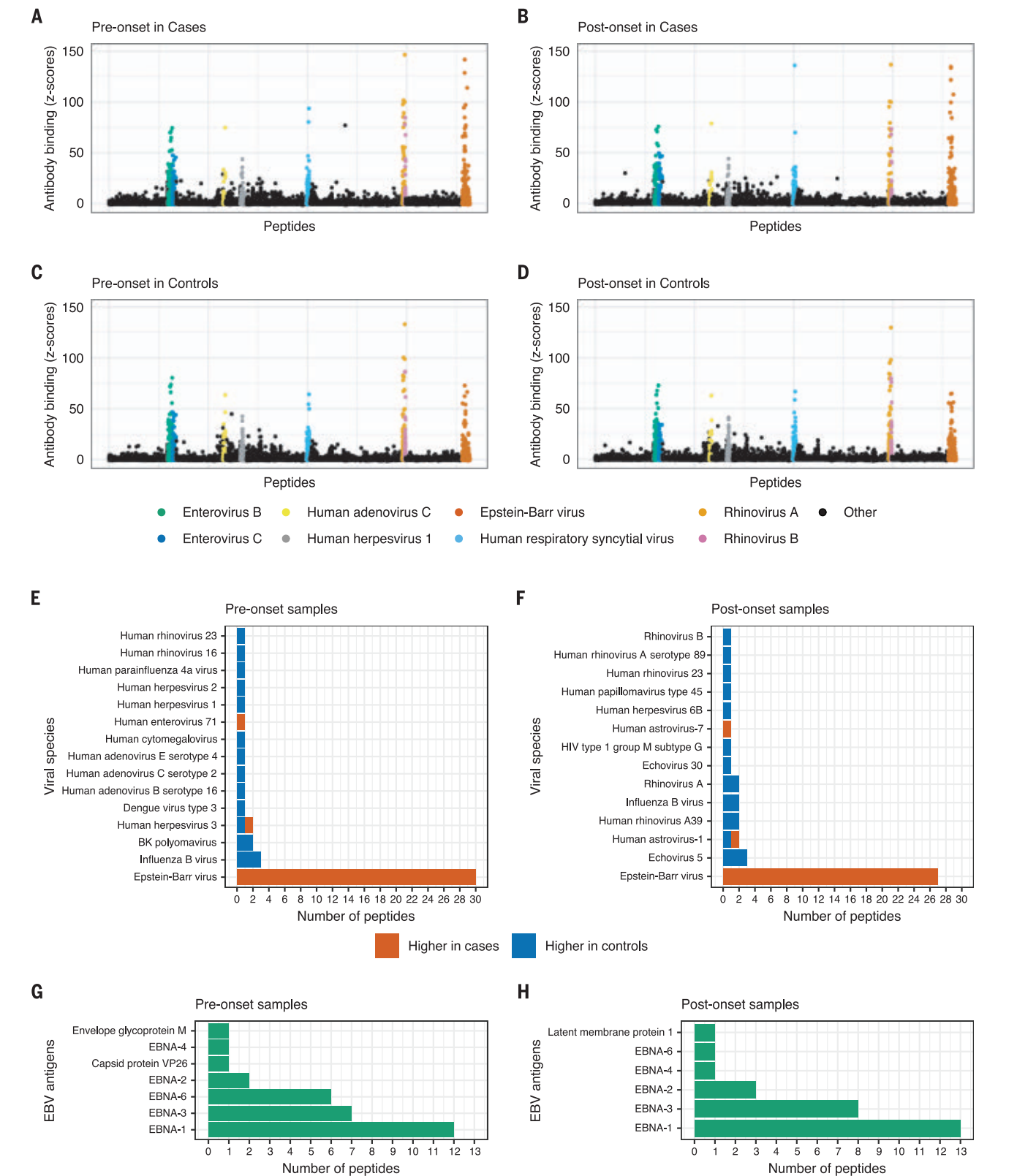
generally, we randomly selected 30 MS cases and 30 matched controls with serum samples collected shortly before (median: 1 year, range: 0 to 3) and soon after symptom onset (median: 1 year, range: 0 to 2) and conducted a comprehensive agnostic search of the anti-virome antibody response using VirScan, an assay based on a T7 phage-display immunoprecipitation and sequencing technology (21), which encodes the full proteome of most known human pathogenic viruses (~200 species, ~110,000 peptides) in 56-amino acid peptides with 28-amino acid overlap between adjacent peptides. VirScan thus enables comprehensive unbiased detection of antibodies raised against all linear peptides encoded in the genomes of viruses known to infect humans. The overall antibody response to viral peptides was similar in cases and controls at both time points, except for EBV (Fig. 4, A to D), arguing that the preclinical and early clinical phases in MS are not associated with immune dysregulation affecting general sus-

ceptibility to infections. Using a Z score of  $>3.5$  as an enrichment cutoff for identifying the presence of a peptide-specific antibody (22), the number of antibody-recognized EBV peptides with a nominally significant difference between MS cases and controls clearly stood out, both in the pre-onset (Fig. 4, E and G, and table S1) and post-onset (Fig. 4, F and H, and table S2) samples, which supports the specificity of the association between EBV and MS and argues against a second hit from another virus playing a major role in MS etiology.

A causal interpretation of our results requires ruling out the possibility that systematic differences between individuals who seroconverted and those who remained EBV-negative explain the results. These differences can be grouped into two categories: (i) confounding by known or unknown factors and (ii) reverse causation.

Confounding by known factors is virtually ruled out by the strength of the association.

To explain a 32-fold increase in MS risk, any confounder would have to confer a  $>60$ -fold increase in risk of EBV seroconversion and a  $>60$ -fold risk of MS (23). None of the known or suspected risk factors for MS has such strong associations. The next strongest known risk factor for MS, homozygosity for the HLA-DR15 allele, which confers a threefold increase in MS risk (24), is not associated with EBV positivity (25) and thus cannot explain the EBV-MS association. Rather, there is epidemiological (26) and experimental (27) evidence that EBV infection and HLA-DR15 may act synergistically in causing MS. Environmental factors are also far too weak to materially confound the EBV-MS association (28). The existence of a still unknown factor that increases the risk of both EBV infection and MS by  $>60$ -fold is rather implausible and there are no good candidates, even hypothetical ones. This conclusion would be robust even in the very unlikely case that EBV seroconversion in one of the MS cases



**Fig. 4. Antibodies against EBV peptides in MS cases detected using virome-wide screening.** (A to D) Scatter plots showing the mean Z scores of serum antibody enrichment against the entire viral peptidome in cases and controls in the pre- and post-onset samples. Each point represents one peptide. (E and F) Number of viral peptides with nominally significantly different antibody

binding between cases and controls in the pre- and post-onset samples (using a Z score >3.5 as an enrichment cut-off; compared using two-sided Fisher's exact test) and their mapping to viral species. (G and H) Mapping of the statistically significant EBV peptides to EBV antigens (for details, see tables S1 and S2).

was a false-positive result, in which case EBV infection would confer a 16-fold increase in MS risk.

Reverse causation could occur if the immune dysregulation during the preclinical phase of MS increases the susceptibility to EBV infection. In our agnostic search of the entire human virome during the preclinical phase of MS, we did not find other systematic differences in the antibody response to any pathogen except EBV that was related to previous infections in MS cases and controls, which makes it unlikely that immune dysregulation during this phase increases susceptibility to infections. This is consistent with previous studies reporting no difference in the frequency of infections in the 5 years preceding MS onset (29) or in individuals with untreated MS (30). Although in one study, hospitalizations for bacterial infections in adolescence were associated with MS risk, this association was modest and therefore cannot explain our study results (31). Additional arguments against reverse causality are that EBV seroconversion occurs before elevation of sNfL levels, which is an early marker of preclinical MS, and the long lag time (median: 7.5 years) between EBV infection and MS clinical onset. The increased MS risk 15 years or longer after infectious mononucleosis (32) and the observation that anti-EBNA antibodies are a strong and consistent predictor of MS risk in EBV-positive individuals up to 15 to 20 years later (33) provide further and independent evidence against reverse causation. Collectively, these findings strongly suggest that the occurrence of EBV infection, detectable by the elicited immune response, is a cause and not a consequence of MS.

One MS case was EBV-negative in the last sample, obtained 3 months before MS onset, which could suggest that EBV was not the cause of disease in this patient. This individual could have been infected with EBV after the last blood collection, could have failed to seroconvert in response to infection (an uncommon but nevertheless regular phenomenon seen after infections and vaccines), or could have been misdiagnosed. Another explanation is related to etiological diversity, which is common for any clinically defined disease. For example, all cases of paralytic poliomyelitis are by definition caused by poliovirus, but rare cases of acute flaccid paralysis, clinically indistinguishable from poliomyelitis, can be caused by other enteroviruses (34). The extremely low MS risk in EBV-negative individuals suggests that by far most MS cases are caused by EBV and could thus potentially be prevented by a suitable vaccine. The addition of MS to the list of diseases that an EBV vaccine could target strengthens the rationale to accel-

erate ongoing research with the primary goal of preventing infectious mononucleosis and posttransplantation lymphoproliferative disease (35).

One of the most effective treatments for MS is anti-CD20 monoclonal antibodies, which deplete circulating memory B cells, the primary site of persistent latent EBV infection (36). This, and preliminary results obtained with EBV-specific T cell therapy (37), suggest that EBV, in addition to causing MS, contributes to MS clinical course, which could thus be potentially modified by antivirals. Directly targeting EBV could have major advantages compared with anti-CD20-based therapies, which have to be administered by intravenous infusion and may increase the risk of infections (36).

## REFERENCES AND NOTES

- D. S. Reich, C. F. Lucchinetti, P. A. Calabresi, *N. Engl. J. Med.* **378**, 169–180 (2018).
- N. L. Fransen et al., *Brain* **143**, 1714–1730 (2020).
- D. A. Thorley-Lawson, *Nat. Rev. Immunol.* **1**, 75–82 (2001).
- E. L. Thacker, F. Mirzaei, A. Ascherio, *Ann. Neurol.* **59**, 499–503 (2006).
- L. I. Levin et al., *JAMA* **293**, 2496–2500 (2005).
- B. Serafini et al., *J. Exp. Med.* **204**, 2899–2912 (2007).
- M. A. Moreno et al., *Neurol. Neuroimmunol. Neuroinflamm.* **5**, e466 (2018).
- A. Hassani, J. R. Corbo, S. Al-Salam, G. Khan, *PLOS ONE* **13**, e0192109 (2018).
- H. Lassmann, G. Niedobitek, F. Aloisi, J. M. Middeldorp; NeuroproMiSe EBV Working Group, *Brain* **134**, 2772–2786 (2011).
- C. L. Perdue, A. A. Eick-Cost, M. V. Rubertone, *Mil. Med.* **180** (suppl.), 10–12 (2015).
- L. I. Levin, K. L. Munger, E. J. O'Reilly, K. I. Falk, A. Ascherio, *Ann. Neurol.* **67**, 824–830 (2010).
- D. H. Crawford et al., *Clin. Infect. Dis.* **43**, 276–282 (2006).
- J. B. Dowd, A. E. Aiello, D. E. Alley, *Epidemiol. Infect.* **137**, 58–65 (2009).
- H. H. Balfour Jr. et al., *J. Infect. Dis.* **208**, 1286–1293 (2013).
- M. Lipsitch, E. Tchetgen Tchetgen, T. Cohen, *Epidemiology* **21**, 383–388 (2010).
- V. Grut et al., *Eur. J. Neurol.* **28**, 3072–3079 (2021).
- H. Tremlett, R. A. Marrie, *Mult. Scler.* **27**, 6–12 (2021).
- M. Khalil et al., *Nat. Rev. Neurol.* **14**, 577–589 (2018).
- B. Delcoigne et al., *Neurology* **94**, e1201–e1212 (2020).
- K. Bjornevik et al., *JAMA Neurol.* **77**, 58–64 (2020).
- G. J. Xu et al., *Science* **348**, aaa0698 (2015).
- M. J. Mina et al., *Science* **366**, 599–606 (2019).
- T. J. VanderWeele, P. Ding, *Ann. Intern. Med.* **167**, 268–274 (2017).
- S. Sawcer et al., *Nature* **476**, 214–219 (2011).
- W. J. Jabs, M. Paulsen, H. J. Wagner, H. Kirchner, H. Klüter, *Clin. Exp. Immunol.* **116**, 468–473 (1999).
- T. Olsson, L. F. Barcellos, L. Alfredsson, *Nat. Rev. Neurol.* **13**, 25–36 (2017).
- H. Zdimierova et al., *Eur. J. Immunol.* (2020).
- A. Ascherio, K. L. Munger, *Semin. Neurol.* **36**, 103–114 (2016).
- J. M. A. Wijnands et al., *Eur. J. Neurol.* **26**, 1032–1036 (2019).
- W. A. Sibley, C. R. Bamford, K. Clark, *Lancet* **1**, 1313–1315 (1985).
- Y. Xu et al., *Brain* **144**, 2390–2400 (2021).
- T. R. Nielsen et al., *Arch. Neurol.* **64**, 72–75 (2007).
- G. N. DeLorenzo et al., *Arch. Neurol.* **63**, 839–844 (2006).
- K. Messacar et al., *Lancet* **385**, 1662–1671 (2015).
- J. I. Cohen, *Adv. Exp. Med. Biol.* **1045**, 477–493 (2018).
- S. L. Hauser et al., *N. Engl. J. Med.* **376**, 221–234 (2017).
- M. P. Pender et al., *JCI Insight* **3**, e124714 (2018).
- Data for: K. Bjornevik et al., Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis, Harvard Dataverse (2021); <https://doi.org/10.7910/DVN/KYH9SF>.

## ACKNOWLEDGMENTS

We thank M. Rubertone and A. Cost of the Armed Forces Health Surveillance Center, Silver Spring, MD, for control and sample identification and retrieval; N. Howard, Department of the Navy, Secretary of the Naval Council of Review Boards, Washington, DC, for MS case identification in the US Navy and Marines; D. Armitage, US Army Physical Disability Agency, Washington, DC, and L. Kent, SAIC, US Army Physical Disability Agency, Washington, DC, for MS case identification in the US Army; L. R. Hodge, US Air Force Personnel Center, San Antonio, TX, for MS case identification in the US Air Force; S. Bernstein, Uniformed Services University of the Health Sciences, Bethesda, MD, for project coordination and management; and D. Leppert, University of Basel, Basel, Switzerland, for serum neurofilament measurements and interpretation of results. The views expressed are those of the authors and should not be construed to represent the positions of the Department of the Army, Department of the Navy, Department of Defense, Department of the Air Force, or Uniformed Services University of the Health Sciences. **Funding:** This work was supported by grants NS046635, NS042194, and NS103891 from the National Institute of Neurological Disorders and Stroke, National Institutes of Health (NIH); National Multiple Sclerosis Society award PP-1912-35234 and German Research Foundation (DFG) research fellowship CO 2129/1-1 to M.C.; and NIH Directors Early Independence Award DP5-OD028145 to M.J.M. S.J.E. is an investigator with the Howard Hughes Medical Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. **Author contributions:** Conceptualization: K.B., M.C., K.L.M., A.A.; Data curation: D.W.N., A.I.S.; Investigation: J.K., M.J.M., Y.L., S.J.E.; Writing: K.B., M.C., K.L.M., A.A., B.C.H., J.K., M.J.M., Y.L., S.J.E., D.W.N., A.I.S. Statistical analysis: K.B., M.C., B.C.H.; Supervision: K.L.M., A.A. **Competing interests:** B.C.H. has received research support from Analysis Group, Celgene (Bristol-Myers Squibb), Verily Life Sciences, Merck-Serono, Novartis, and Genzyme. J.K. has received speaker fees, research support, travel support, and/or served on advisory boards of the Swiss MS Society, Swiss National Research Foundation (320030\_189140/1), University of Basel, Progressive MS Alliance, Bayer, Biogen, Celgene, Merck, Novartis, Octave Bioscience, Roche, and Sanofi. M.J.M. is on the advisory board of ImmunelD and an adviser to Detect and Quantum-SI. S.J.E. is a founder of TSCAN Therapeutics, MAZE Therapeutics, Mirimus, and ImmunelD. S.J.E. serves on the scientific advisory board of Homology Medicines, TSCAN Therapeutics, MAZE, and ImmunelD, and is an adviser for MPM, none of which affect this work. S.J.E. is an inventor on a patent application filed by the Brigham and Women's Hospital (US20160320406A) that covers the use of the VirScan library to identify pathogen antibodies in the blood. K.B., Y.L., A.I.S., D.W.N., A.A., K.L.M., and M.C. declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this study are present in the main paper or the supplementary materials and have been deposited at the Harvard Dataverse (38).

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abj8222](https://science.org/doi/10.1126/science.abj8222)  
Materials and Methods  
Figs. S1 to S4  
Tables S1 and S2  
References (39–52)  
Data S1  
MDAR Reproducibility Checklist

3 June 2021; accepted 2 December 2021  
10.1126/science.abj8222



## SOLAR CELLS

# Conformal quantum dot-SnO<sub>2</sub> layers as electron transporters for efficient perovskite solar cells

Minjin Kim<sup>1†</sup>, Jaeki Jeong<sup>2†</sup>, Haizhou Lu<sup>2†</sup>, Tae Kyung Lee<sup>3</sup>, Felix T. Eickemeyer<sup>2</sup>, Yuhang Liu<sup>2</sup>, In Woo Choi<sup>1</sup>, Seung Ju Choi<sup>1</sup>, Yimhyun Jo<sup>1</sup>, Hak-Beom Kim<sup>1</sup>, Sung-In Mo<sup>1</sup>, Young-Ki Kim<sup>4</sup>, Heunjeong Lee<sup>5</sup>, Na Gyeong An<sup>6</sup>, Shinuk Cho<sup>5</sup>, Wolfgang R. Tress<sup>7</sup>, Shaik M. Zakeeruddin<sup>2</sup>, Anders Hagfeldt<sup>8</sup>, Jin Young Kim<sup>6</sup>, Michael Grätzel<sup>2\*</sup>, Dong Suk Kim<sup>1\*</sup>

Improvements to perovskite solar cells (PSCs) have focused on increasing their power conversion efficiency (PCE) and operational stability and maintaining high performance upon scale-up to module sizes. We report that replacing the commonly used mesoporous-titanium dioxide electron transport layer (ETL) with a thin layer of polyacrylic acid-stabilized tin(IV) oxide quantum dots (paa-QD-SnO<sub>2</sub>) on the compact-titanium dioxide enhanced light capture and largely suppressed nonradiative recombination at the ETL-perovskite interface. The use of paa-QD-SnO<sub>2</sub> as electron-selective contact enabled PSCs (0.08 square centimeters) with a PCE of 25.7% (certified 25.4%) and high operational stability and facilitated the scale-up of the PSCs to larger areas. PCEs of 23.3, 21.7, and 20.6% were achieved for PSCs with active areas of 1, 20, and 64 square centimeters, respectively.

Efforts to realize metal halide perovskite solar cells (PSCs) with power conversion efficiencies (PCEs) of >23% have focused on formamidinium-rich lead iodide (FAPbI<sub>3</sub>) formulations (1–7) because their narrower bandgap is closer to the Shockley-Queisser optimum than for methylammonium-based or mixed-halide perovskites (8). By fully using the broad absorption spectrum of FAPbI<sub>3</sub>, a certified PCE of 25.21% with a short-circuit current density ( $J_{sc}$ ) of >26 mA/cm<sup>2</sup> was obtained for the mesoporous-structure PSCs (7). However, the mesoporous-TiO<sub>2</sub> (m-TiO<sub>2</sub>) electron transport layer (ETL) may show unwanted photocatalytic effects under ultraviolet (UV) light illumination, and the low electron mobility of m-TiO<sub>2</sub> limits the charge transport (9–11).

Among alternative metal oxide ETLs (10–17) for PSCs, SnO<sub>2</sub>-based PSCs could potentially be more efficient and stable given that SnO<sub>2</sub> is UV resistant and has a higher carrier mobility

than TiO<sub>2</sub>, which facilitates electron extraction and transport (10–12). Several techniques, such as spin coating (11, 16), atomic layer deposition (1), and chemical bath deposition (CBD) (12, 17), have been used to deposit the SnO<sub>2</sub> ETLs. Spin-coated SnO<sub>2</sub> ETL from a SnO<sub>2</sub> colloidal quantum dot (QD-SnO<sub>2</sub>) solution onto the indium-doped tin oxide (ITO) substrate enabled a certified PCE of >23% for the corresponding planar-structure PSCs (2, 16). Recently, a thin SnO<sub>2</sub> ETL on fluorine-doped tin oxide (FTO) deposited with a well-controlled CBD method enabled PSCs with a certified PCE of 25.19%, because of the improved carrier properties of SnO<sub>2</sub> ETL (17). However, compared with the m-TiO<sub>2</sub>-based PSCs, the SnO<sub>2</sub>-based PSCs still suffered from a relatively low  $J_{sc}$  of <26 mA/cm<sup>2</sup>, which is attributed to the optical losses arising from reflection and destructive interference of the incident light waves at the interfaces.

One approach to reduce these optical losses is to use the textured surface of FTO as the front contact that scatters the incoming radiation, destroying the coherence of the incoming light and affording light trapping by increasing the optical path length (18). The enhanced light absorption by the perovskite benefits the photocurrent delivered by the photovoltaic (PV) cell. Similar strategies have been used for textured crystalline silicon-based PSCs (19). However, early efforts to deposit a thin, uniform, and high-quality SnO<sub>2</sub> ETL using a solution process were incompatible with the underlying textured FTO surface (1, 12, 20–22), causing optical losses. The highest reported  $J_{sc}$  of SnO<sub>2</sub>-based PSCs of ~25.2 mA/cm<sup>2</sup> (17, 20) still limits the overall PV performance.

Here we introduce an architecture for the ETL of PSCs that consists of a compact-TiO<sub>2</sub> (c-TiO<sub>2</sub>) blocking layer covered by a thin layer of polyacrylic acid (PAA)-stabilized QD-SnO<sub>2</sub> (paa-QD-SnO<sub>2</sub>) deposited in a contiguous and

conformal manner on the textured FTO. The uniform bilayer of paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> largely improved the perovskite's absorption of sunlight and formed an outstanding electron-selective contact with the perovskite film. The quantum size effect increased the bandgap of the QD-SnO<sub>2</sub> from 3.6 eV for bulk SnO<sub>2</sub> to ~4 eV (21, 23) and produced a corresponding upward shift of its conduction band edge energy. This shift aligned it well with the conduction band edge of the perovskite so that electron capture by the SnO<sub>2</sub>-based ETL proceeded with minimal energy losses (5, 11, 16, 21).

PAA, a polymer binder, was added to the SnO<sub>2</sub> QD solution to attach the colloidal QD-SnO<sub>2</sub> firmly to the c-TiO<sub>2</sub> surface, providing a contiguous, thin, and conformal SnO<sub>2</sub> layer that fully covered the c-TiO<sub>2</sub> layer underneath. The carboxyl groups of PAA undergo strong hydrogen and coordinative bonding with the metal oxide surface, facilitating the lamination process, especially for production on a large scale (24–26). By choosing FTO substrates with suitable diffuse transmittance and reflectance, the textured paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer enabled a PCE of 25.7% (certified 25.4%) with a  $J_{sc}$  of 26.4 mA/cm<sup>2</sup> and high stability for the corresponding PSCs. We further demonstrate that the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer could be applied to realize large PSC modules with an active area up to 64 cm<sup>2</sup> while maintaining a PCE of >20%.

We investigated the microstructures of the spin-coated QD-SnO<sub>2</sub> layer on the c-TiO<sub>2</sub> using commercially available SnO<sub>2</sub> colloidal QDs with and without PAA. Unless otherwise noted, the QD-SnO<sub>2</sub> solution was diluted by deionized water (1:20) in this study. Figure S1, A to C, shows the top-view scanning electron microscope (SEM) images of the c-TiO<sub>2</sub>, QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>, and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>, respectively. Because of the textured surface, the c-TiO<sub>2</sub> layer was not fully covered by the spin-coated QD-SnO<sub>2</sub> (fig. S1B). In contrast, a uniform, conformal paa-QD-SnO<sub>2</sub> layer was formed (fig. S1C). Atomic force microscopy images (fig. S1, D to F) further confirmed the uniform morphology of the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (fig. S1F), which is different from that of the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (fig. S1E).

Figure 1, A and B, shows the cross-sectional transmission electron microscopy (TEM) images of the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayers on FTO substrates, respectively. The QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer presented an ununiform distribution over the FTO surface with a thickness that varied from ~30 (vertex region) to ~70 nm (valley region), while the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer had a uniform and conformal distribution over the FTO surface with a uniform thickness of ~30 nm. The different distribution between QD-SnO<sub>2</sub> and paa-QD-SnO<sub>2</sub> layers can also be seen clearly from the cross-sectional SEM images (fig. S2, A and B),

<sup>1</sup>Ulsan Advanced Energy Technology R&D Center, Korea Institute of Energy Research, Ulsan 44776, Republic of Korea. <sup>2</sup>Laboratory of Photonics and Interfaces, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. <sup>3</sup>Photovoltaics Research Department, Korea Institute of Energy Research (KIER), Daejeon 34129, Republic of Korea. <sup>4</sup>Central Research Facilities (UCRF), Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea. <sup>5</sup>Department of Physics and Energy Harvest Storage Research Center, University of Ulsan, Ulsan 44610, Republic of Korea. <sup>6</sup>Department of Energy Engineering, School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea. <sup>7</sup>Novel Semiconductor Devices Group, Institute of Computational Physics, Zurich University of Applied Sciences, 8401 Winterthur, Switzerland. <sup>8</sup>Department of Chemistry, Ångström Laboratory, Uppsala University, 751 20 Uppsala, Sweden.

\*Corresponding author. Email: michael.gratzel@epfl.ch (M.G.); kimds@kier.re.kr (D.S.K.)

†These authors contributed equally to this work.

suggesting that PAA played a key role in forming the uniform and contiguous layer of paa-QD-SnO<sub>2</sub> ETL, as shown in fig. S3.

Element mapping with energy-disperse x-ray spectroscopy of Ti (Fig. 1, C and D) and Sn (Fig. 1, E and F) revealed a coverage of TiO<sub>2</sub> and SnO<sub>2</sub> over the FTO surface for both QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayers. The selected-area electron diffraction generated by TEM for the paa-QD-SnO<sub>2</sub>@c-

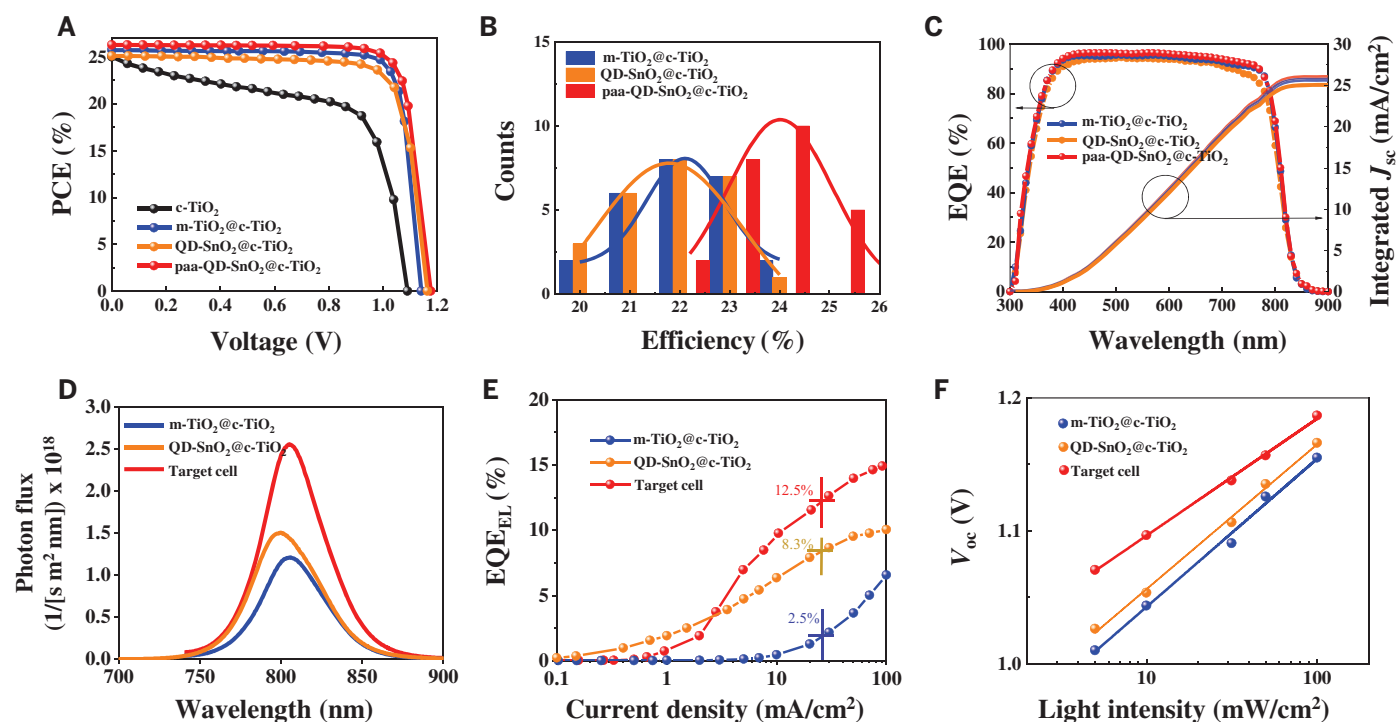
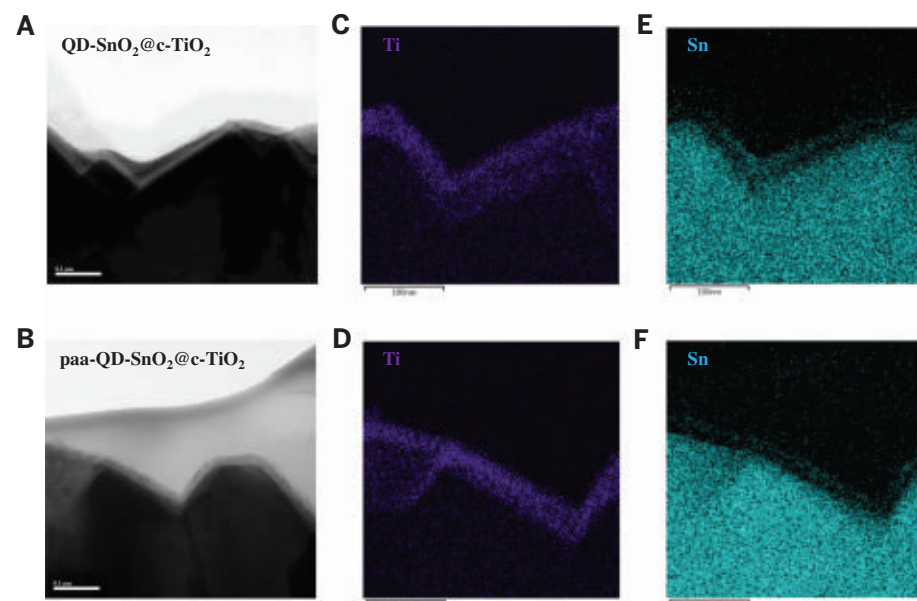
TiO<sub>2</sub> bilayer (fig. S4) showed that both QD-SnO<sub>2</sub> and c-TiO<sub>2</sub> were polycrystalline. The paa-QD-SnO<sub>2</sub> had a particle size of ~4 nm (fig. S4C), which is also confirmed by the TEM images (fig. S5, A to C) and dynamic light scattering analysis (fig. S5, D and E).

The interactions between PAA and QD-SnO<sub>2</sub> were studied by the x-ray photoelectron spectroscopy (XPS) (fig. S6) and Fourier transform infrared spectroscopy (FTIR) measurements

(fig. S7). It is clear from the XPS measurements that both QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayers showed characteristic peaks attributed to Sn (fig. S6A); however, the Sn 3d peaks of paa-QD-SnO<sub>2</sub> shifted to high binding energy by ~0.2 eV compared with that of QD-SnO<sub>2</sub>, indicating that PAA was bonded to the QD-SnO<sub>2</sub>. No obvious difference was observed for the O 1s characteristic peaks (fig. S6, B to D). FTIR measurements showed

**Fig. 1. Microstructures of the ETLs.**

(A and B) The cross-sectional TEM images of QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (A) and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (B) over the FTO substrates. Scale bars, 0.1  $\mu$ m. (C and D) EDS elemental analysis of Ti for both QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (C) and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (D) over the FTO surface. Scale bars, 100 nm. (E and F) EDS elemental analysis of Sn for both QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (E) and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (F) over the FTO surface. Scale bars, 100 nm.



peaks at  $\sim 2954$  and  $\sim 1716$   $\text{cm}^{-1}$  that arise from C–H and C=O stretching vibrations of PAA (fig. S7A) and a characteristic peak at  $\sim 574$   $\text{cm}^{-1}$  of the Sn–O vibration (fig. S7B). The Sn–O peak shifted to  $\sim 594$   $\text{cm}^{-1}$  for the paa-QD-SnO<sub>2</sub>, and the C–H and C=O stretching vibrations shifted to  $\sim 3012$  and  $\sim 1628$   $\text{cm}^{-1}$ , respectively (fig. S7C), demonstrating that PAA interacted with QD-SnO<sub>2</sub> (27).

We chose FAPbI<sub>3</sub> as the perovskite layer, and details of its fabrication can be found in the supplementary materials (SM) or a previous report (3). Top-view SEM images of the FAPbI<sub>3</sub> films spin-coated on different ETLs, including c-TiO<sub>2</sub>, m-TiO<sub>2</sub>@c-TiO<sub>2</sub>, QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>, and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>, are shown in fig. S8, A to D, respectively. Compact and dense surface morphologies were observed for all the perovskite films. Figure S9A showed identical x-ray diffraction (XRD) peak positions for all samples at  $14.1^\circ$  and  $28.2^\circ$ , which correspond to the (001) and (002) crystal planes of  $\alpha$ -FAPbI<sub>3</sub> (3–5). All the perovskite films had identical full width at half maximum for the main (001) peak (fig. S9B). We conclude that the morphology and crystallinity of the perovskite films are not affected substantially by the different ETLs.

The PSCs with an active area of 0.08  $\text{cm}^2$  were fabricated in a conventional n-i-p structure, FTO/ETL/perovskite/OAI/spiro-MeOTAD/Au, where OAI is octylammonium iodide and spiro-MeOTAD is 2,2',7,7'-tetrakis[N,N-di(4-methoxyphenyl)amino]-9,9'-spirobifluorene. We performed the quasi-steady-state current-voltage (QSS-*I*/*V*) measurement (Fig. 2A), which was used for the certification of PSCs by the National Renewable Energy Laboratory (NREL) and Newport Inc. Details of the QSS-*I*/*V* measurements are given in the SM. All the PSCs were fully aged in the ambient condition for 100 hours before the measurements.

The c-TiO<sub>2</sub>-based cell had a low PCE of 17.27% under the QSS-*I*/*V* measurement. For the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based PSC, a PCE of 23.74% with a  $J_{\text{sc}}$  of 25.74  $\text{mA}/\text{cm}^2$ , an open-circuit voltage ( $V_{\text{oc}}$ ) of 1.142 V, and a fill factor (FF) of 80.79% was obtained, which are consistent with the previous reports (3, 7). Compared with the mesoporous-structured devices, the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based cell had a higher  $V_{\text{oc}}$  of 1.164 V but a lower  $J_{\text{sc}}$  of 25.12  $\text{mA}/\text{cm}^2$ , resulting in a PCE of 23.29%. The  $J_{\text{sc}}$  of the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based PSC is similar to the values reported for the SnO<sub>2</sub>-based cells in the literature (17, 20). With the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> ETL, the PSC exhibited a high PCE of 25.18% with a considerably improved  $J_{\text{sc}}$  of 26.28  $\text{mA}/\text{cm}^2$ ,  $V_{\text{oc}}$  of 1.177 V, and FF of 81.49%, matching the 25.39% efficiency certified by Newport Inc. (figs. S10 and S11).

The conventional *J*-*V* measurements under both forward and reverse scans were also performed (fig. S12). The detailed PV parameters

are summarized in table S1. A substantially different PV value was obtained for the c-TiO<sub>2</sub>-based cell under QSS-*I*/*V* (Fig. 2A) and conventional *J*-*V* measurements (fig. S12A), whereas the other cells showed similar PV results, indicating that the c-TiO<sub>2</sub>-based cell is unstable and that the single c-TiO<sub>2</sub> layer is not a proper ETL for PSCs (28, 29). The contrasting feature of the c-TiO<sub>2</sub>-based cell compared with the other cells in Fig. 2A and fig. S12A is discussed in note 1 of the SM. Hereafter, the c-TiO<sub>2</sub>-based cells will not be discussed. A statistical distribution of the PCE of all the PSCs (Fig. 2B) shows that the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based PSC had the highest averaged values. Details of the statistical PV parameters of all different ETL-based cells are shown in fig. S13. The paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based cell is the target of the discussions that follow.

The  $J_{\text{sc}}$  of the PSCs measured under the solar simulator was verified with external quantum efficiency (EQE) measurements. The target cell had slightly higher EQE than the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based PSC over the entire absorption spectrum (Fig. 2C), resulting in a higher integrated  $J_{\text{sc}}$  of 26.01  $\text{mA}/\text{cm}^2$ ; however, a relatively lower integrated  $J_{\text{sc}}$  of 25.06  $\text{mA}/\text{cm}^2$  was obtained for the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based device than for the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based cell (25.69  $\text{mA}/\text{cm}^2$ ). The high  $J_{\text{sc}}$  of the target cell was attributed to the desired light scattering that prolongs the optical length, enhancing the light absorption by the perovskite with the conformal structured paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer over the FTO substrate. This is confirmed with the highest diffuse transmittance (haze) of paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> when compared with the other ETLs shown in fig. S14A. The higher transmittance of paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> than the QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> (fig. S14B) may be traced back to the thinner film thickness, as shown in the optical simulations (fig. S14, C and D), which could also contribute to the high  $J_{\text{sc}}$ . We further compared the  $J_{\text{sc}}$  obtained for the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based PSCs using different substrates (fig. S15). The Asahi FTO glass with high diffuse transmittance was the most suitable substrate for achieving a high  $J_{\text{sc}}$ .

The effect of the ETL composition on the photon flux emitted by the PSCs measured in steady state at an excitation photon flux equivalent to 1 sun is shown in Fig. 2D. The investigated devices were complete solar cells without the Au back contact. Compared with the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>- and QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based devices, the target cell had much higher photoluminescence (PL) intensity, reaching a PL quantum yield (PLQY) of 7.5%. This indicates a reduced nonradiative recombination at the interface between perovskite and paa-QD-SnO<sub>2</sub> ETL. Details of the measurements and calculations of PLQY are shown in SM note 2. From the PLQY measurements, we derived

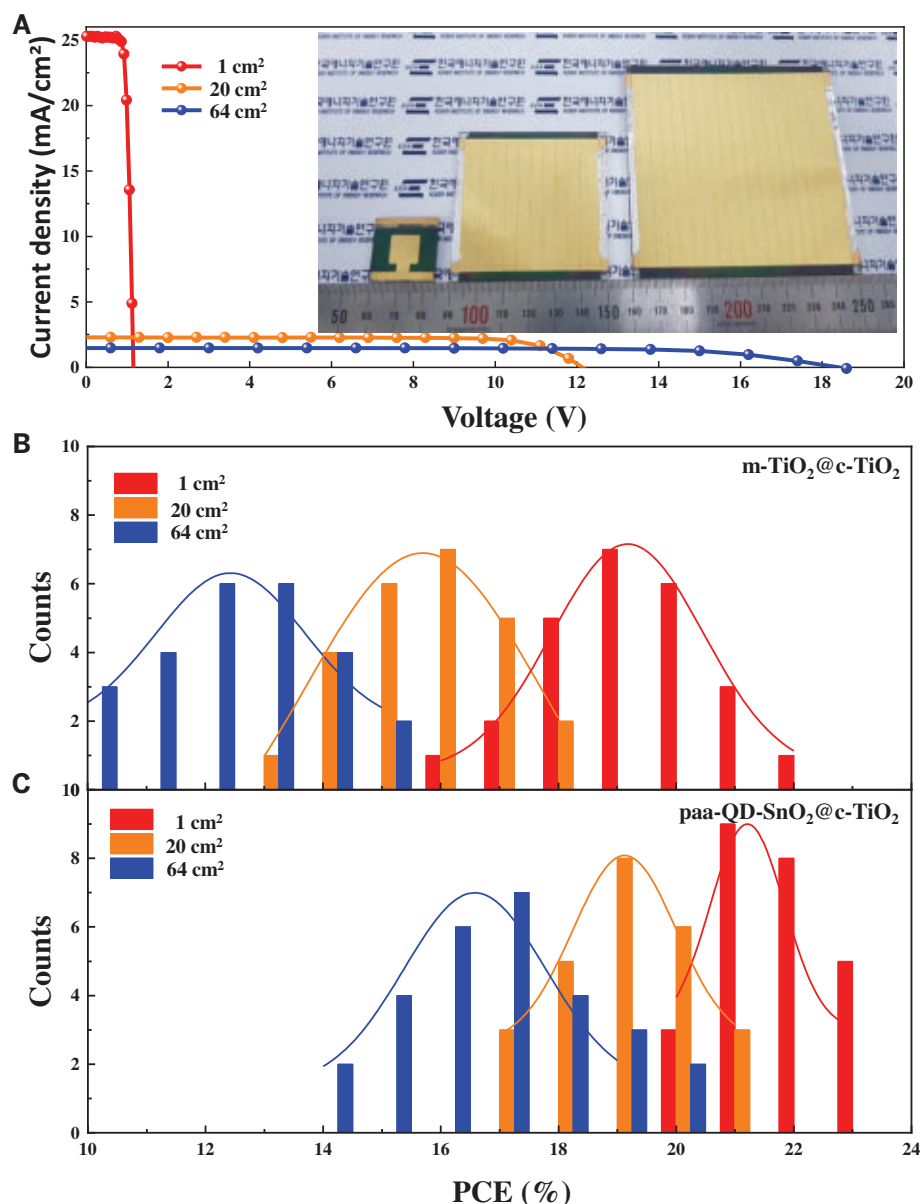
the quasi-Fermi level splitting ( $\Delta E_{\text{F}}$ ) in the perovskite under 1 sun illumination (SM note 3) and compared the ( $\Delta E_{\text{F}}$ ) with the  $V_{\text{oc}}$  measured from the same device.  $\Delta E_{\text{F}}$  and  $V_{\text{oc}}$  showed the same trend, indicating that the  $V_{\text{oc}}$  increase can be partly attributed to the reduced nonradiative recombination. The  $\Delta E_{\text{F}}/q - V_{\text{oc}}$  offset (where  $q$  is the elementary charge), however, is different. For the target cell, it is 10 mV lower than that of the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based cells, indicating a better energetic alignment at the interfaces (30).

Ultraviolet photoelectron spectroscopy (UPS) measurements on the surface of different ETLs (fig. S16) showed that the conduction band of paa-QD-SnO<sub>2</sub> matched better with perovskite than the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>, which could facilitate the charge transfer from perovskite to the ETL (fig. S17). Detailed analysis of the UPS data is shown in SM note 4. Figure S18, A and B, shows the time-resolved PL (TRPL) spectra of the perovskite films on different ETLs measured from both the perovskite and the glass sides. The perovskite film deposited on paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> had the fastest decay among all the samples. Because it has a low non-radiative recombination rate, the fast decay is dominated by the favored interfacial electron transfer. The electron mobility measurements (fig. S18C) revealed that the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> layer had the highest carrier mobility. All these results indicate a superior charge extraction from perovskite to the paa-QD-SnO<sub>2</sub> ETL.

Regarding the reduced nonradiative recombination, we determined the EQE of electroluminescence (EQE<sub>EL</sub>) on representative PSC devices. Figure 2E shows an EQE<sub>EL</sub> of 12.5% for the target cell at an injection current density of 26  $\text{mA}/\text{cm}^2$  corresponding to the  $J_{\text{sc}}$  under 1 sun illumination, whereas the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>- and QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based cells have an EQE<sub>EL</sub> of 2.5 and 8.3%, respectively, under the same conditions. Thus, the nonradiative recombination rate in the PSC was reduced by 80% simply by replacing the m-TiO<sub>2</sub> with a paa-QD-SnO<sub>2</sub> electron-selective contact layer. The highest obtained  $V_{\text{oc}}$  of 1.22 V (fig. S19), which is near the  $V_{\text{oc}}$  predicted from EQE<sub>EL</sub> according to the reciprocity theorem (31, 32), measured without the metal mask reached 98% of the radiative limit  $V_{\text{oc}}$  (1.25 V) (2, 5). We also conducted transient photovoltage measurements for the PSCs (fig. S20A). The target cell showed a slower  $V_{\text{oc}}$  decay than the reference cells, indicating a slower charge recombination rate. The dark *J*-*V* curves (fig. S20B) showed the lowest reverse saturation current for the target cell, pushing the onset of the dark current to the highest voltages, which also reflected the lowest interfacial nonradiative recombination.

Figure 2F shows the light intensity-dependent  $V_{\text{oc}}$  measurements for the PSCs. For both the reference and target cells, the





**Fig. 3. Performance of the large-size PSCs.** (A) *J-V* curves of the large-size PSCs. (Inset) Photo of the large-size PSCs. (B and C) Statistical distributions of the PCEs for the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based (B) and paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based (C) PSCs with pixel sizes of 1, 20, and 64 cm<sup>2</sup>.

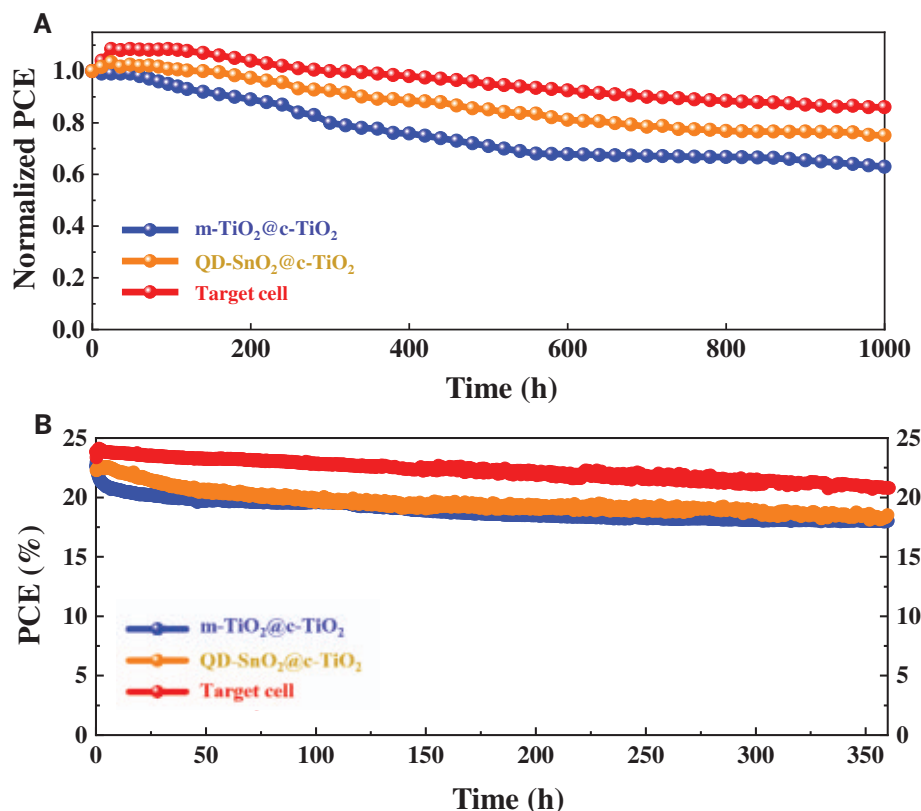
measured  $V_{oc}$  was linearly dependent on the logarithm of the light intensity. The diode ideality factor  $n_{id}$ , deduced from the slope described by  $n_{id}k_B T/q$ , where  $k_B$  is the Boltzmann constant and  $T$  is temperature, was 1.83, 1.79, and 1.46 for the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>- and QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based cells and the target cell, respectively. The reduced  $n_{id}$  contributed to the increased FF of the target cell, as the FF critically depends on the  $n_{id}$  (33). The reduced  $n_{id}$  is also consistent with the PLQY, TRPL, EQE<sub>EL</sub>, transient photovoltage decay, and dark *J-V* measurements, unambiguously supporting the conclusion of reduced nonradiative recombination of the target PSC using paa-QD-SnO<sub>2</sub>

ETL. The decreased nonradiative recombination manifests itself by much stronger photo- and electroluminescence (Fig. 2, C and D) as well as a lower ideality factor (Fig. 2E) enabling very high fill factor of 83.8% to be reached by our target device (table S2). Further strong support for our conclusion comes from the observation of a slower transient photovoltage decay (fig. S20A) and higher electron mobility (fig. S20B) as well as the dark *J-V* measurements of the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based cell compared with the control devices. The substantial reduction of the trap-assisted nonradiative recombination is the main reason for the reduced  $V_{oc}$  deficit of our target

cell of ~310 mV as compared with ~350 mV for the mesoporous-structure cell [bandgap of our perovskite film was calculated to be 1.53 eV (3)], which is one of the lowest values reported in the field of PSCs.

Scale-up of the PSCs to module size is another requirement for their commercial exploitation. We used paa-QD-SnO<sub>2</sub>@TiO<sub>2</sub> ETL to fabricate perovskite solar mini-modules with active areas up to 64 cm<sup>2</sup>. Details of the fabrication process for the solar modules can be found in the SM and movie S1. Figure 3A shows the *J-V* curves and images of the perovskite mini-modules with different active areas. The highest PCEs achieved for the PSCs with active areas of 1, 20, and 64 cm<sup>2</sup> were 23.3, 21.7, and 20.6%, respectively. Movie S2 shows a typical measurement for the 64 cm<sup>2</sup> perovskite solar modules. Figure S21A illustrates a solar module with subcells connected in series with a magnified view of the contact connections, and a geometrical FF (GFF) was calculated to be 95.6% according to the SEM images (fig. S21B) of the interconnections. Details of the PV parameters are summarized in table S2. The decrease of the PCE with increased device size is mainly caused by a decreased FF. The  $V_{oc}$  of the module (table S2) divided by the number of stripes is 18.5 V/16 = 1.156 V, which is equal to the  $V_{oc}$  of the 1 cm<sup>2</sup> cell. Hence, there is no additional  $V_{oc}$  loss in the module. Therefore, we attribute the FF decrease to the increased series resistance, including transport layer resistances, contact resistances, and interconnect resistance. We sent the PSC modules to an independent laboratory (OMA Company, Republic of Korea) for certification. PCEs of 21.66% (fig. S22) and 20.55% (fig. S23) were confirmed for the PSC mini-modules with active areas of 20 and 64 cm<sup>2</sup>, respectively, which agreed well with the measurements in our laboratory, and are compared with other reported values in fig. S24.

We further compared the statistical distribution of the PCEs for the mesoporous-based (Fig. 3B) and target (Fig. 3C) perovskite mini-modules. For the perovskite modules with the same active area (1, 20, or 64 cm<sup>2</sup>), the target modules had higher averaged PCE and narrower PCE distributions than the mesoporous-based modules. The averaged PCE of the 64 cm<sup>2</sup> PSC module increased by ~30% by simply replacing the m-TiO<sub>2</sub> with paa-QD-SnO<sub>2</sub>, indicating that the paa-QD-SnO<sub>2</sub> could be uniformly coated on the large-size substrates. The statistical distribution of the PV parameters (fig. S25) further revealed that the  $V_{oc}$  and FF of the mesoporous-based modules decreased significantly more than that of the target PSC modules. The shunt resistance of our solar modules retained a typical value of >4000 ohms-cm<sup>2</sup> (table S3), indicating that losses from leakage currents across shunts



**Fig. 4. Stability of the PSCs with different ETLs.** The stability shown here represents the best stability results of our four tested samples. **(A)** Shelf life of the unencapsulated PSCs tested in ambient air at 25% RH and 25°C with different ETLs. **(B)** The operational stability test of the unencapsulated PSCs under MPP tracking conditions in a N<sub>2</sub> environment.

produced during scale-up are negligible. We attribute the small decrease of  $V_{oc}$  and FF for the target PSC modules to the low contact resistance (fig. S26, A and B) because of the easier laser scribing as well as low series resistance and low interfacial defects of the paa-QD-SnO<sub>2</sub>@c-TiO<sub>2</sub> bilayer.

Figure 4A shows the 1000-hour shelf life tested with unsealed devices in ambient air at 25% relative humidity (RH) and 25°C. The target PSC retained 80% of the maximum PCE (25.7%) after 1000-hour storage, whereas the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>- and QD-SnO<sub>2</sub>@c-TiO<sub>2</sub>-based PSCs decreased to almost 30 and 40% of the initial efficiency, respectively, indicating that the target cell is more resistant to the moisture and oxygen than the widely used mesoporous structured PSCs. We also performed the operational stability test for the unsealed devices in a N<sub>2</sub> environment under maximum power point (MPP) tracking conditions using a light-emitting diode lamp with a calibrated light intensity of 100 mW/cm<sup>2</sup>. More details about the MPP tests can be found in the SM and in previous reports (5, 7). Figure 4B shows that the target cell had higher PCE than the reference cell during the 350-hour MPP tracking test. From the detailed PV parameters during the MPP tracking measure-

ments (fig. S27), it is clear that the higher performance of the target PSC is due to the higher and rather stable  $V_{oc}$ , which decreased less compared with the reference cells. The main degradation of the target cell is the decline of FF (fig. S27), which is attributed to the de-doping of lithium from the hole-transport layer (HTL) (5, 34).

We further tested the sealed devices under ambient conditions. Figure S28 shows that 70.5% of the initial efficiency of the target cell was retained after 700-hour light-soaking test, whereas the efficiency of the m-TiO<sub>2</sub>@c-TiO<sub>2</sub>-based cell declined 45% over the same testing time. To assess the suitability of the cells for real applications, we performed an MPP stability test for the encapsulated target cell under ambient conditions. Figure S29 shows that 95% of the initial efficiency was retained for the sealed target cell after 100 hours MPP tracking and 2 hours dark recovery under ambient conditions. The substantial decline of FF is still the main reason for the performance loss. Future studies of additive-free HTLs will be conducted to enhance the stability of PSCs.

#### REFERENCES AND NOTES

1. S. H. Turren-Cruz, A. Hagfeldt, M. Saliba, *Science* **362**, 449–453 (2018).
2. Q. Jiang et al., *Nat. Photonics* **13**, 460–466 (2019).

3. M. Kim et al., *Joule* **3**, 2179–2192 (2019).
4. H. Min et al., *Science* **366**, 749–753 (2019).
5. H. Lu et al., *Science* **370**, eabb8985 (2020).
6. H. Lu, A. Krishna, S. M. Zakeeruddin, M. Grätzel, A. Hagfeldt, *iScience* **23**, 101359 (2020).
7. J. Jeong et al., *Nature* **592**, 381–385 (2021).
8. G. E. Eperon et al., *Energy Environ. Sci.* **7**, 982–988 (2014).
9. T. Leijtens et al., *Nat. Commun.* **4**, 2885 (2013).
10. Q. Jiang, X. Zhang, J. You, *Small* **14**, e1801154 (2018).
11. Q. Jiang et al., *Nat. Energy* **2**, 16177 (2017).
12. E. H. Anaraki et al., *Energy Environ. Sci.* **9**, 3128–3134 (2016).
13. R. Chen et al., *J. Am. Chem. Soc.* **141**, 541–547 (2019).
14. S. S. Shin et al., *Nat. Commun.* **6**, 7410 (2015).
15. S. S. Shin et al., *Science* **356**, 167–171 (2017).
16. G. Yang et al., *Nat. Photonics* **15**, 681–689 (2021).
17. J. J. Yoo et al., *Nature* **590**, 587–593 (2021).
18. S. Manzoor, Z. J. Yu, A. Ali, W. Ali, Z. C. Holman, "Improved light incoupling in planar solar cells via improved texture morphology of PDMS scattering layer," *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)* (2017), pp. 1228–1232.
19. Y. Hou et al., *Science* **367**, 1135–1140 (2020).
20. Y. Zhang et al., *ACS Energy Lett.* **5**, 360–366 (2020).
21. M. M. Tavakoli, F. Giordano, S. M. Zakeeruddin, M. Grätzel, *Nano Lett.* **18**, 2428–2434 (2018).
22. T. Bu et al., *Nat. Commun.* **9**, 4609 (2018).
23. A. Das et al., *J. Mater. Chem. C* **2**, 164–171 (2014).
24. B. Hu et al., *J. Power Sources* **416**, 125–131 (2019).
25. L. Wei, Z. Hou, *J. Mater. Chem. A* **5**, 22156–22162 (2017).
26. P. Parikh et al., *Chem. Mater.* **31**, 2535–2544 (2019).
27. D. Yang et al., *Nat. Commun.* **9**, 3239 (2018).
28. J. Ji et al., *iScience* **23**, 101013 (2020).
29. M. M. Tavakoli, P. Yadav, R. Tavakoli, J. Kong, *Adv. Energy Mater.* **8**, 1800794 (2018).
30. P. Caprioglio et al., *Adv. Energy Mater.* **9**, 1901631 (2019).
31. U. Rau, *Phys. Rev. B* **76**, 085303 (2007).
32. W. Tress et al., *Adv. Energy Mater.* **5**, 1400812 (2015).
33. M. A. Green, *Solar Cells* **7**, 337–340 (1982).
34. Y. Wang et al., *Science* **365**, 687–691 (2019).

#### ACKNOWLEDGMENTS

We thank B. I. Carlsen, O. Ouellette, and M. Wei for discussions. We thank S. Song for the help of optical simulations. **Funding:** This work was supported by Development Program of the Korea Institute of Energy Research (KIER) (C1-2401 and C1-2402); Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1A6A1A03038697); and the NRF funded by the Ministry of Science, ICT, and Future Planning (2020M1A2A2080746 and 2020M1A2A208075011). M.G. acknowledges financial support from the European Union's Horizon 2020 research and innovation program under grants 881603 and 764047. **Author contributions:** M.G. and D.S.K. designed and supervised the project. A.H., J.Y.K., and S.M.Z. advised on the research. M.K., J.J., and H.L. studied and constructed the concept and analyzed the experimental data. H.L., M.K., W.R.T., and J.J. wrote the manuscript. M.G. and D.S.K. revised the manuscript. T.K.L. contributed to the optical measurements. H.L., F.T.E., and W.R.T. performed the PLQY and EQE<sub>EL</sub> measurements and analysis. Y.L. contributed to the characterization of the chemicals. I.W.C., S.J.C., N.G.A., S.C., and Y.J. characterized the perovskite film with UV-Vis absorption, XPS, and XRD. S.-I.M. analyzed the FTIR. H.-B.K. performed the UPS measurements. M.K., J.J., and H.L. performed the stability test. Y.-K.K. analyzed the high-resolution TEM. All authors contributed to the discussions about the manuscript and the reviewers' comments. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

#### SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh1885  
Materials and Methods  
Supplementary Text  
Figs. S1 to S29  
Tables S1 to S3  
References (35–50)  
Movies S1 and S2

22 February 2021; accepted 7 December 2021  
10.1126/science.abh1885

## PEROVSKITES

# Three-dimensional direct lithography of stable perovskite nanocrystals in glass

Ke Sun<sup>1†</sup>, Dezhi Tan<sup>2\*†</sup>, Xinyuan Fang<sup>3,4†</sup>, Xintao Xia<sup>1</sup>, Dajun Lin<sup>3,4</sup>, Juan Song<sup>5</sup>, Yonghong Lin<sup>6</sup>, Zhaojun Liu<sup>6</sup>, Min Gu<sup>3,4</sup>, Yuanzheng Yue<sup>7</sup>, Jianrong Qiu<sup>1,8\*</sup>

Material composition engineering and device fabrication of perovskite nanocrystals (PNCs) in solution can introduce organic contamination and entail several synthetic, processing, and stabilization steps. We report three-dimensional (3D) direct lithography of PNCs with tunable composition and bandgap in glass. The halide ion distribution was controlled at the nanoscale with ultrafast laser-induced liquid nanophase separation. The PNCs exhibit notable stability against ultraviolet irradiation, organic solution, and high temperatures (up to 250°C). Printed 3D structures in glass were used for optical storage, micro-light emitting diodes, and holographic displays. The proposed mechanisms of both PNC formation and composition tunability were verified.

Compositional tuning of the optical properties of perovskites (1, 2) is usually performed in solution to create materials for high-performance devices with long-term stability (3–5), such as mixed chloride-bromide and bromide-iodide perovskites for spectrally stable and high-efficiency blue and red light-emitting diodes (LEDs), respectively (5, 6). Despite recent advances in optoelectrical performance, low structural stability has been an obstacle for practical perovskite devices (6), and numerous strategies such as surface passivation or device encapsulation have been developed (7). In these approaches, stabilization requires additional processing steps at the thin film or device level and is not integral to tuning the nanocrystal (NC) properties.

The postsynthetic incorporation of NCs into glass has led to advanced photonic functionalities (8, 9). However, the three-dimensional (3D) tailoring of the chemical composition and the bandgap of NCs inside glass, and, in turn, the tuning of the functionalities of NC-based photonic devices, is challenging. Recently, an ultrafast laser has been used to fabricate 3D functional structures in transparent solids (10–13), but the internal composition tunability of functional structures is rather limited.

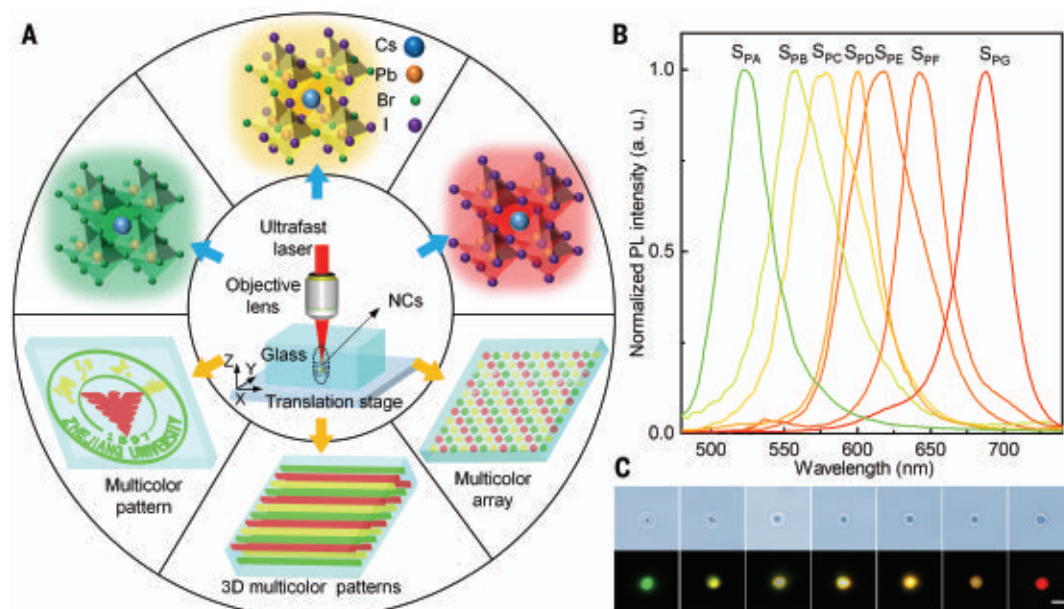
We report a different strategy for engineering the local chemistry of NCs. Specifically, ultrafast-laser pulses inject energy within an ultrashort amount of time, which leads to strong thermal accumulation and thereby increases the local pressure and temperature

above the liquidus of the studied glass system to induce localized liquid nanophase separation (14–16), so that 3D direct lithography of composition-tunable perovskite NCs (PNCs) inside glass is realized (Fig. 1A). The mechanism of the composition tuning of PNCs through liquid nanophase separation was clarified. In addition, our approach enabled the PNCs to be well protected against high-power ultraviolet (UV) light irradiation, organic solution, or temperatures up to 250°C.

We used oxide glasses containing cesium, lead, and halide elements as our medium for direct lithography of PNCs. As a typical oxide glass, borophosphate glass with the molar composition of 40B<sub>2</sub>O<sub>3</sub>–15P<sub>2</sub>O<sub>5</sub>–10Al<sub>2</sub>O<sub>3</sub>–10ZnO–5Na<sub>2</sub>O–5K<sub>2</sub>O–7Cs<sub>2</sub>O–3PbX<sub>2</sub>–5NaX (where X is Cl, Br, or I) was prepared using a melt-quenching method. The high mobility of cesium, lead, and halide ions promotes perovskite nanophase separation from the glass matrix and the subsequent formation of compositionally tunable PNCs (4, 5, 16). We achieved the

<sup>1</sup>State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou, China. <sup>2</sup>Zhejiang Lab, Hangzhou, China. <sup>3</sup>Institute of Photonic Chips, University of Shanghai for Science and Technology, Shanghai, China. <sup>4</sup>Centre for Artificial-Intelligence Nanophotonics, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, China. <sup>5</sup>College of Materials Science and Engineering, Jiangsu University, Zhenjiang, China. <sup>6</sup>Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China. <sup>7</sup>Department of Chemistry and Bioscience, Aalborg University, 9220 Aalborg, Denmark. <sup>8</sup>CAS Center for Excellence in Ultra-intense Laser Science, Chinese Academy of Sciences, Shanghai 201800, China. \*Corresponding author. Email: wtcdz@zju.edu.cn (D.T.); qjr@zju.edu.cn (J.Q.)

†These authors contributed equally to this work.



**Fig. 1. Direct lithography of composition-tunable PNCs in glass.** (A) Schematic illustration of direct lithography of colored PNCs and patterns. (B) PL spectra of CsPb(Br<sub>1-x</sub>I<sub>x</sub>)<sub>3</sub> PNCs written in one piece of glass. S<sub>PA</sub> to S<sub>PG</sub> represent the samples of CsPb(Br<sub>1-x</sub>I<sub>x</sub>)<sub>3</sub> written with different laser parameters that are shown in table S1. a.u., arbitrary units. (C) Optical images (top) and PL mappings (bottom) of CsPb(Br<sub>1-x</sub>I<sub>x</sub>)<sub>3</sub> NCs. The scale bar is 10 μm, and the excitation wavelength is 405 nm.



desired ultrafast-laser direct lithography process by optimizing the pulse duration, repetition rate, and pulse energy (table S1). The photoluminescence (PL) of the as-prepared PNCs was tuned across the wavelength range from 520 to 690 nm (Fig. 1B). The light emissions at 520 and 690 nm were attributed to the exciton recombination in  $\text{CsPbBr}_3$  and  $\text{CsPbI}_3$  NCs, respectively, and the emissions between these two wavelengths originated from the mixed halide  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs (1, 2), where  $x$  was determined using Vegard's law (fig. S1). The presence of PNCs was confirmed by both the transmission electron microscopy images (fig. S2) and Raman spectra (fig. S3), and the mean size of the PNCs was determined to be between 1 and 4 nm.

We realized control over the dynamical process of liquid nanophase separation by adjusting the ultrafast-laser irradiation time ( $t_i$ ) (Fig. 2A). Here, the halide ion migration rate depends on the complexation between  $\text{Pb}^{2+}$  and halide ions and the radius and weight of ions (17, 18). In comparison with  $\text{I}^-$ , a greater

complexation between  $\text{Pb}^{2+}$  and  $\text{Br}^-$ , lighter ionic weight, and smaller radius allowed for faster diffusion of  $\text{Br}^-$  and easier formation of Br-rich liquid perovskite through nanophase separation. Continuous irradiation allowed more  $\text{I}^-$  ions to diffuse into the liquid perovskite region from the liquid-glass domains and enabled tuning of the emission of the final PNCs (Fig. 2B and supplementary text S2) from green to red by extending  $t_i$ .

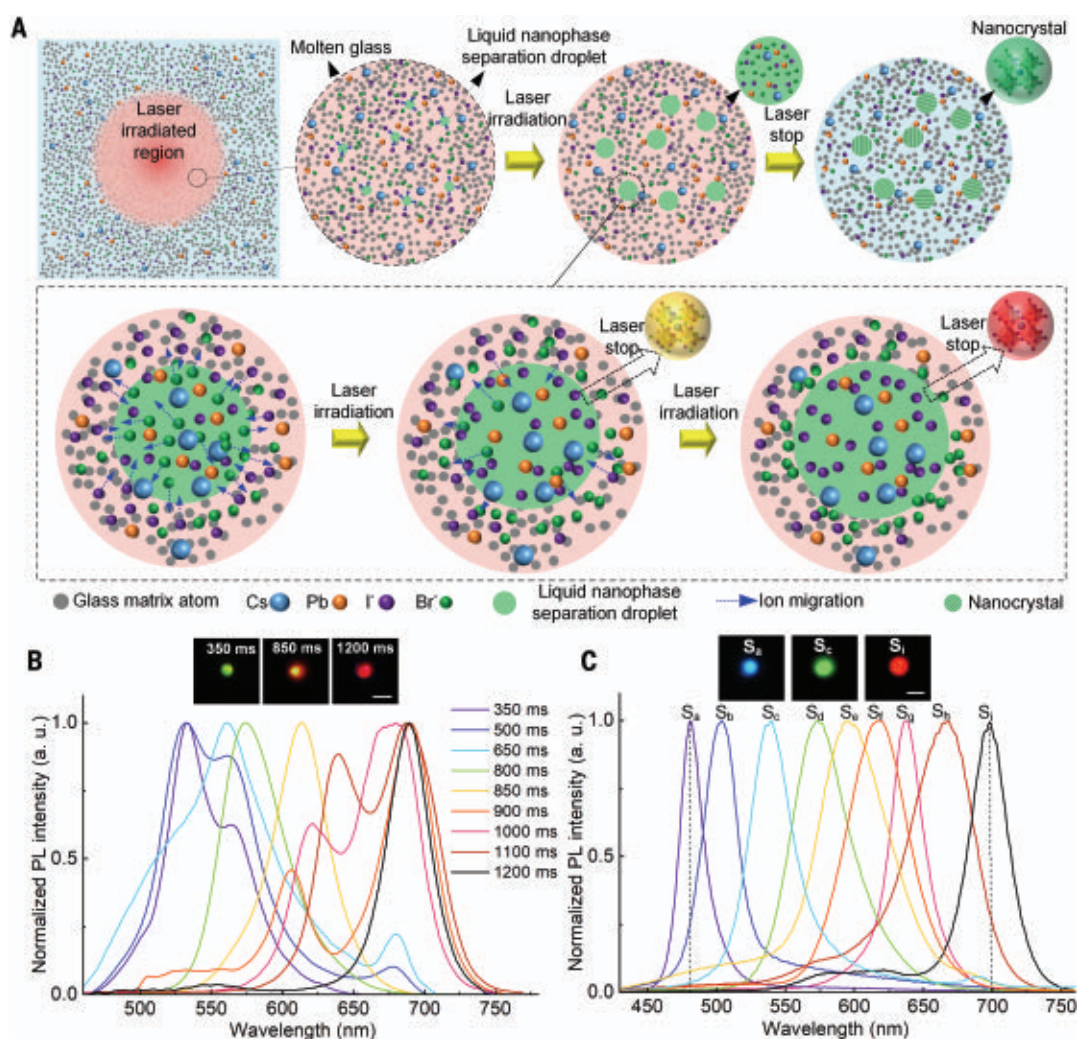
To validate our approach,  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x)_3$  NCs were generated in glass, and the emission was tuned across a wide wavelength range from 450 to 514 nm (fig. S4) by controlling the laser parameters (table S3). Furthermore, we succeeded in engineering the composition and bandgap of PNCs in the  $\text{Cl}^-$ - $\text{Br}^-$ - $\text{I}^-$  codoped glass. Thus, the full-color printing of PNCs (fig. S5) with PL tuned in a range from 480 to 700 nm was achieved (Fig. 2C) and reflected the transformation of  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x)_3$  into  $\text{CsPbI}_3$ , thus confirming the PNC composition engineering. The PL of PNCs written in both  $\text{Cl}^-$ - $\text{Br}^-$  (fig. S6) doped glasses and  $\text{Cl}^-$ - $\text{Br}^-$ - $\text{I}^-$

(fig. S7) doped glasses (fig. S7) were continuously modulated by changing  $t_i$ , and, specifically, the main PL peak shifted to longer wavelengths with an increase in  $t_i$ . Regulation of the halide ion distribution in PNCs in glass was not possible through a conventional homogeneous heat treatment (fig. S8).

Phase separation occurs in glass-forming systems if a chemical potential gradient exists. Based on our experimental findings, here we propose the mechanism of PNC formation through nanophase separation by taking  $\text{Br}^-$ - $\text{I}^-$  doped glass as an example (Fig. 2A). First, formation of immiscible phases resulted in liquid phase separation at the nanoscale level. Thus, separation of the Br-rich halide phase from the glass matrix phase occurred at temperatures above the liquidus temperature of the glass composition (19). Second, continuous ultrafast-laser irradiation not only increased the size of the liquid perovskite domains (fig. S9) but also induced the site exchange of I to Br owing to the chemical potential gradient (Fig. 2A) (20). As the laser irradiation proceeded,  $\text{I}^-$

**Fig. 2. Dynamical control over the halide ion migration and full-color printing of PNCs.**

(A) Schematic of ultrafast laser-induced liquid nanophase separation and formation of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs in the  $\text{Br}^-$ - $\text{I}^-$  doped glass. (B) PL mappings and PL spectra of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs as a function of  $t_i$ . The ultrafast-laser repetition rate is 125 kHz, the pulse duration is 885 fs, and the pulse energy is 400 nJ. (C) PL mappings and PL spectra of PNCs written in the  $\text{Cl}^-$ - $\text{Br}^-$ - $\text{I}^-$  codoped glass.  $S_0$  to  $S_9$  represent the PNC samples written in the  $\text{Cl}^-$ - $\text{Br}^-$ - $\text{I}^-$  codoped glass with different laser parameters that are shown in table S2.



ions gradually diffused from the surrounding liquid to the relatively ordered liquid perovskite domains, finally leading to the formation of  $\Gamma^-$ -containing liquid perovskite nanodomains. The nanophase separation lowered the energy barrier for formation of the domains with a preordered perovskite-like structure (21). Third, the preordered liquid perovskite domains became more ordered and created the crystallization nuclei that subsequently grew into PNCs through diffusion and reaction in a confined manner during the cooling process (22).

During ultrafast-laser direct lithography, the temperature of the laser-impacted domains increased rapidly with the number of pulses and remained stable at the maximum after several tens of pulses (typically less than 100 pulses that correspond to a  $t_i$  of 1 ms for a 100-kHz ultrafast laser). A quenching process occurred after shutting off the ultrafast-laser irradiation (23). The ultrafast laser-induced temperature ( $>1000^\circ\text{C}$ ; fig. S10) in the modified area is above the liquidus of the glass composition (fig. S11). Thus, the dependence of the emission wavelength of PNCs on  $t_i$  verified the occurrence of liquid nanophase separation. The mean size of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs increased from 1.9 to 3.6 nm with an increase in  $t_i$  from 350 to 1200 ms (fig. S9), which provided a clear signature of the continuous localized liquid nanophase separation.

The distinct evolution of PNCs (Fig. 2B and figs. S12 to S14) with adjusted ultrafast-laser parameters could be related to the differences in temperature (fig. S10), pressure, and irradiance of the ultrafast laser. For example, the phase diagram depended on the pressure, and the liquid phase separation could be facilitated by increasing the pressure up to a gigapascal level (14–16, 24, 25). These features account, in part, for why ultrafast-laser heating could drive liquid nanophase separation, whereas the normal heat treatment could not.

The instability of PNCs can occur through chemical- and thermal-induced decomposition as well as light-induced phase segregation (7, 17, 26). We investigated the stability of PNCs against UV irradiation, heat treatment, or solvent (ethanol) exposure. All of the PNCs that emitted green, yellow, orange, and red PL were stable, and no change in PL intensity was observed after UV irradiation for 12 hours (Fig. 3A). Furthermore, there was also no PL peak shift when  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs were irradiated by UV light with the power density ( $I_{\text{UV}}$ ) of  $2\text{ W/cm}^2$  (Fig. 3B) and even  $32\text{ W/cm}^2$  (figs. S15 and S16), implying an absence of phase segregation. As a reference, UV light with an  $I_{\text{UV}}$  of  $0.1\text{ W/cm}^2$  can induce substantial phase segregation in mixed-halide perovskites (4, 5, 17).

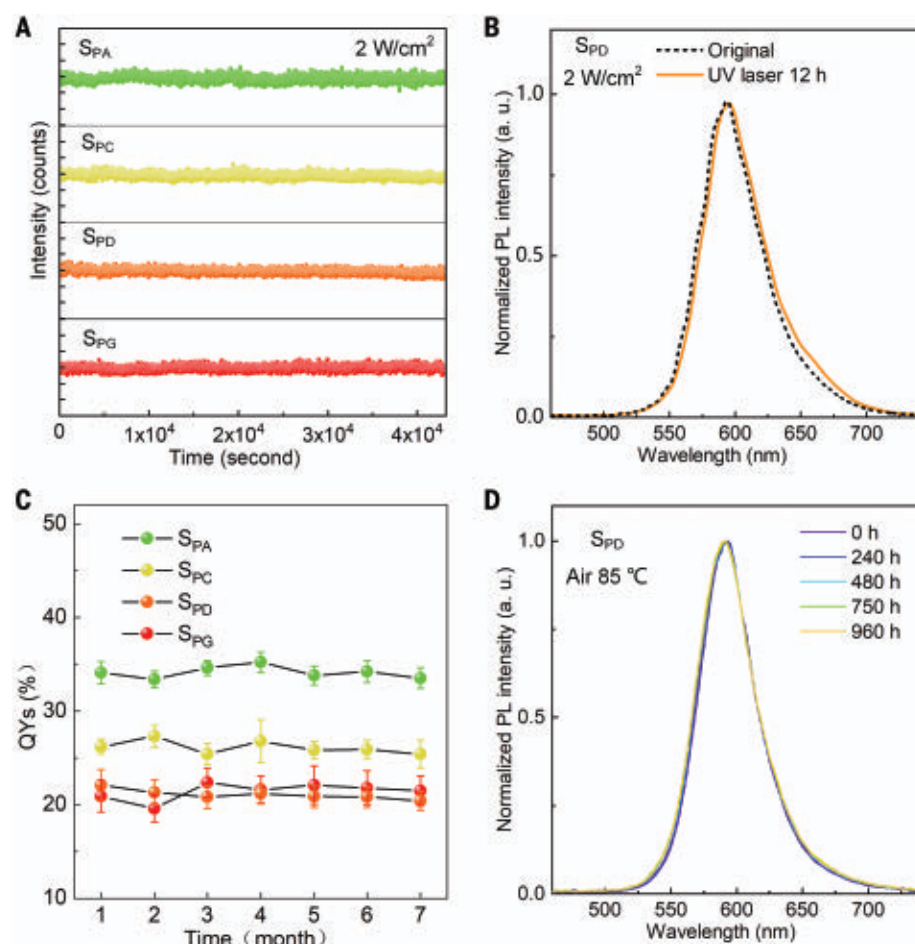
The PNCs remained stable when dispersed in ethanol without a change in the PL quan-

tum yield after 6 months (Fig. 3C). The PL intensity and position of PNCs also remained as the initial characteristics after they were heat-treated at  $85^\circ\text{C}$  for 960 hours (Fig. 3D and fig. S17) and even after heat treatment at  $250^\circ\text{C}$  for 2 hours under atmospheric conditions (fig. S18). The high stability of PNCs originates from the effective protection of the glass matrix that prevents the as-written PNCs from being attacked by molecules in the surrounding environment at various temperatures (supplementary text S5) (7). Furthermore, the factors, such as nanoconfinement, ultrafast laser-induced residual strain, high surface-to-volume ratio, high cohesive energy, and limited carrier diffusion length given the small size of PNCs could lead to strong suppression of the ion diffusion and UV light-induced phase segregation (10, 17, 27, 28).

Writing composition-tunable PNCs allows for applications in multidimensional information encoding and anticounterfeiting (Fig. 1A and fig. S19). For example, green, yellow, and

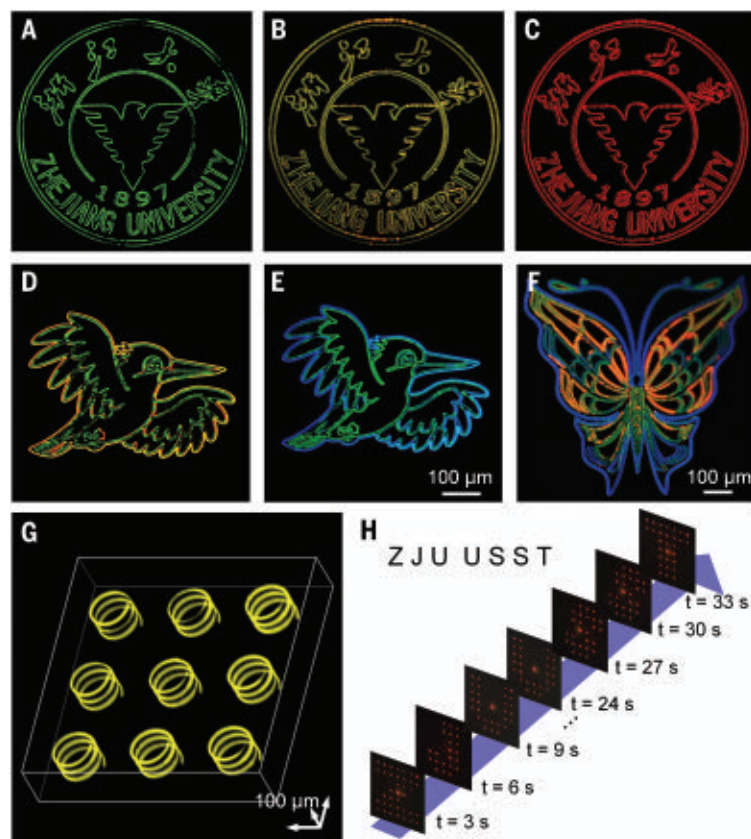
red logos of Zhejiang University were directly written in glass (Fig. 4, A to C). Figure 4, D and E, shows the colorful patterns that are produced with  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs and  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x)_3$  NCs, respectively, in the corresponding glasses. We also demonstrated full-color printing of PNCs in the  $\text{Cl}^-$ - $\text{Br}^-$ - $\text{I}^-$  codoped glass (Fig. 4F) and 3D microhelix PNC patterns (Fig. 4G).

Micrometer-scale LEDs ( $\mu$ -LEDs) for high-resolution display have been fabricated with wet chemistry-derived NCs (29, 30). For standard NC-based devices, dots with different emission wavelengths were printed or transferred on substrates, and the NC preparation and device manufacturing were complicated (29). Thus, the cost of device fabrication is high, the stability of NCs can be low, and NC patterning is difficult. These drawbacks severely limit the wide applications of NC-based devices. In addition, although considerable efforts were made to exploit glasses as light-emitting materials and devices (10, 31, 32), it has not been



**Fig. 3. Stability of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs.** (A) PL intensity of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs illuminated for 12 hours by UV laser with an  $I_{\text{UV}}$  of  $2\text{ W/cm}^2$ . (B) PL spectra of  $\text{S}_{\text{PD}}$  NCs before and after irradiation. (C) PL quantum yields (QYs) of  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs dispersed in ethanol after 6 months. Error bars represent standard deviation. (D) PL spectra of  $\text{S}_{\text{PD}}$  NCs after heat treatment at  $85^\circ\text{C}$  for 960 hours.





**Fig. 4. Direct lithography of PNC patterns and devices.** (A to C) Zhejiang University logos based on  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs with varied compositions in one piece of glass. (D to F) Typical PL images of the multicolor patterns produced with  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NCs in the  $\text{Br}^-/\text{I}^-$  doped glass (D),  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x)_3$  NCs in the  $\text{Cl}^-/\text{Br}^-$  doped glass (E), and  $\text{CsPb}(\text{Cl}_{1-x}\text{Br}_x\text{I}_y)_3$  NCs in the  $\text{Cl}^-/\text{Br}^-/\text{I}^-$  codoped glass (F). The colorful patterns were produced by adjusting the parameters during writing. (G) 3D microhelix  $\text{CsPb}(\text{Br}_{1-x}\text{I}_x)_3$  NC arrays. The patterns were excited by UV light. (H) Demonstration of a dynamic holographic display. The letters at the top left represent the images shown at different times ( $t = 3, 6, 9, 24, 27$ , and  $33$  s, respectively) in the demonstration of a dynamic holographic display. The excitation wavelength is  $532$  nm. Scale bars in (A) to (F) are  $100$   $\mu\text{m}$ .

possible to endow a single glass chip with widely and continuously tunable colors or to print NC-based devices such as  $\mu$ -LEDs and displays inside solids (5, 29).

The PNC-based glass can be used as the key component of photonic devices, not just as a transparent protection layer or substrate (33). We demonstrated one-step 3D printing of colored PNC patterns with a dot size  $<10$   $\mu\text{m}$  (fig. S20) that can be used for building micrometer-sized devices. Combining commercially available UV or blue LED arrays, this technology could apply to the fabrication of  $\mu$ -LEDs (fig. S21), high-resolution displays, and even white LEDs (figs. S22 and S23) (2, 29, 34). Moreover, the PNC arrays were used as a holographic display device. Through the switching of holograms, the dynamic holographic display based on such a PNC-based device was achieved (Fig. 4H, fig. S24, and movie S1). Three holographic images (letters “Z,” “J,” and “U”) were further reconstructed simultaneously in multiple planes along the light propagation

direction, indicating that the excitation of specific PNC patterns enabled realization of a 3D holographic display (fig. S25).

Because liquid nanophase separation only occurs at a local position inside glass, the 3D direct lithography represents a completely dry fabrication technique that advances the fabrication of structures and devices with high throughput and high scalability. It excludes contamination with organic components (reagents and solvents) during material synthesis and device processing. Furthermore, the high-temperature stability indicates that the PNC glass-based devices can be applied for long-term applications (35). The present study demonstrates that the composition and bandgap of PNCs could be tailored over the entire tunable range in one single solid chip.

#### REFERENCES AND NOTES

1. L. Protesescu et al., *Nano Lett.* **15**, 3692–3696 (2015).
2. X. Li et al., *Adv. Funct. Mater.* **26**, 2435–2445 (2016).
3. D. P. McMeekin et al., *Science* **351**, 151–155 (2016).

4. J. Xu et al., *Science* **367**, 1097–1104 (2020).
5. Y. Hassan et al., *Nature* **591**, 72–77 (2021).
6. X.-K. Liu et al., *Nat. Mater.* **20**, 10–21 (2021).
7. Y. Wei, Z. Cheng, J. Lin, *Chem. Soc. Rev.* **48**, 310–350 (2019).
8. A. Llordés, G. Garcia, J. Gazquez, D. J. Milliron, *Nature* **500**, 323–326 (2013).
9. A. L. Efros, L. E. Brus, *ACS Nano* **15**, 6192–6210 (2021).
10. X. Huang et al., *Nat. Photonics* **14**, 82–88 (2020).
11. W. Yang, P. G. Kazansky, Y. P. Svirko, *Nat. Photonics* **2**, 99–104 (2008).
12. D. Z. Tan, Z. Wang, B. Xu, J. Qiu, *Adv. Photonics* **3**, 024002 (2021).
13. D. Wei et al., *Nat. Photonics* **12**, 596–600 (2018).
14. Y. Katayama et al., *Science* **306**, 848–851 (2004).
15. S. Aasland, P. F. McMillan, *Nature* **369**, 633–636 (1994).
16. E. D. Zanotto, *Ceram. Int.* **46**, 24779–24791 (2020).
17. M. C. Brennan, A. Ruth, P. V. Kamat, M. Kuno, *Trends Chem.* **2**, 282–301 (2020).
18. S. J. Yoon, K. G. Stamplecoskie, P. V. Kamat, *J. Phys. Chem. Lett.* **7**, 1368–1373 (2016).
19. P. F. James, *J. Mater. Sci.* **10**, 1802–1825 (1975).
20. G. Nedelcu et al., *Nano Lett.* **15**, 5635–5640 (2015).
21. Y. Yue, *J. Non-Cryst. Solids* **345–346**, 523–527 (2004).
22. M. Allix, L. Cormier, in *Springer Handbook of Glass*, J. D. Musgraves, J. Hu, L. Calvez, Eds. (Springer, 2019), chap. 4.
23. D. Z. Tan, B. Zhang, J. R. Qiu, *Laser Photonics Rev.* **15**, 2000455 (2021).
24. L. Henry et al., *Nature* **584**, 382–386 (2020).
25. K. H. Kim et al., *Science* **370**, 978–982 (2020).
26. M. Liu et al., *Nat. Photonics* **15**, 379–385 (2021).
27. X. Kong et al., *Nanoscale* **10**, 8320–8328 (2018).
28. X. Wang et al., *Nat. Commun.* **10**, 695 (2019).
29. Z. Liu et al., *Light Sci. Appl.* **9**, 83 (2020).
30. Y. Wang, I. Fedin, H. Zhang, D. V. Talapin, *Science* **357**, 385–388 (2017).
31. X. Liu, J. Zhou, S. Zhou, Y. Yue, J. Qiu, *Prog. Mater. Sci.* **97**, 38–96 (2018).
32. T. Hu et al., *Light Sci. Appl.* **10**, 56 (2021).
33. J. H. Choi et al., *Nat. Photonics* **5**, 763–769 (2011).
34. M. Zhao et al., *Light Sci. Appl.* **8**, 38 (2019).
35. J. Zhang, M. Gecevicius, M. Beresna, P. G. Kazansky, *Phys. Rev. Lett.* **112**, 033901 (2014).

#### ACKNOWLEDGMENTS

We thank Y. Zhang (Qilu University of Technology) for performing differential scanning calorimetry measurements. **Funding:** This work was financially supported by the National Natural Science Foundation of China (grant nos. U20A20211, 51902286, 61775192, 61905215, 51772270, and 62005164) and the Shanghai Frontiers Science Center Program 2021-2025 (NO.20). **Author contributions:** K.S., D.T., and X.F. contributed equally to this work. D.T. conceived the idea. J.Q. organized, coordinated, and supervised the project. K.S. and X.X. performed the experiments and collected the data. X.F. and D.L. carried out the holographic display experiment. Y.L. and Z.L. carried out the LED device experiment. J.S. performed the theoretical calculations. D.T., Y.Y., and J.Q. interpreted the results and proposed the mechanism of the composition engineering of perovskite nanocrystals in glass with a laser. D.T. wrote the manuscript. M.G. supervised the holographic display experiment and reviewed the manuscript. D.T., Y.Y., and J.Q. discussed and revised the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are available in the main text or the supplementary materials.

#### SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj2691  
Materials and Methods  
Supplementary Text  
Figs. S1 to S41  
Tables S1 to S6  
References (36–58)  
Movie S1

2 May 2021; resubmitted 13 October 2021  
Accepted 2 December 2021  
10.1126/science.abj2691



## SOCIAL PSYCHOLOGY

# Early concepts of intimacy: Young humans use saliva sharing to infer close relationships

Ashley J. Thomas<sup>1,2,3\*</sup>, Brandon Woo<sup>1,3</sup>, Daniel Nettle<sup>4</sup>, Elizabeth Spelke<sup>1,3</sup>, Rebecca Saxe<sup>2,3</sup>

Across human societies, people form “thick” relationships characterized by strong attachments, obligations, and mutual responsiveness. People in thick relationships share food utensils, kiss, or engage in other distinctive interactions that involve sharing saliva. We found that children, toddlers, and infants infer that dyads who share saliva (as opposed to other positive social interactions) have a distinct relationship. Children expect saliva sharing to happen in nuclear families. Toddlers and infants expect that people who share saliva will respond to one another in distress. Parents confirm that saliva sharing is a valid cue of relationship thickness in their children’s social environments. The ability to use distinctive interactions to infer categories of relationships thus emerges early in life, without explicit teaching; this enables young humans to rapidly identify close relationships, both within and beyond families.

To become a competent member of society, humans must learn how the people around them are related to each other (1–3). Across cultures, people distinguish a special category of relationships, which we will call “thick” (4–7). Thick relationships feature strong levels of attachments, obligations, mutual responsiveness, and a feeling of oneness that is conceived in terms of shared bodily substance (5–7); they are often, but not always, between close genetic relatives (8–12). The fact that only some relationships are thick presents young humans with a problem: How do they identify which ones? For older children, distinct relationship categories can be explicitly verbally labeled (13). Anthropologists have claimed that younger children and even infants must be sensitive to how relationships are embodied in distinctive interactions (14, 15). For example, interactions that involve deliberate consensual transfer of saliva, such as kissing or sharing food utensils, distinctively occur in thick relationships (16–18). Here, we applied experimental techniques from developmental science to test whether young children, toddlers, and infants do indeed infer that two individuals who share saliva are likely to be in a thick relationship.

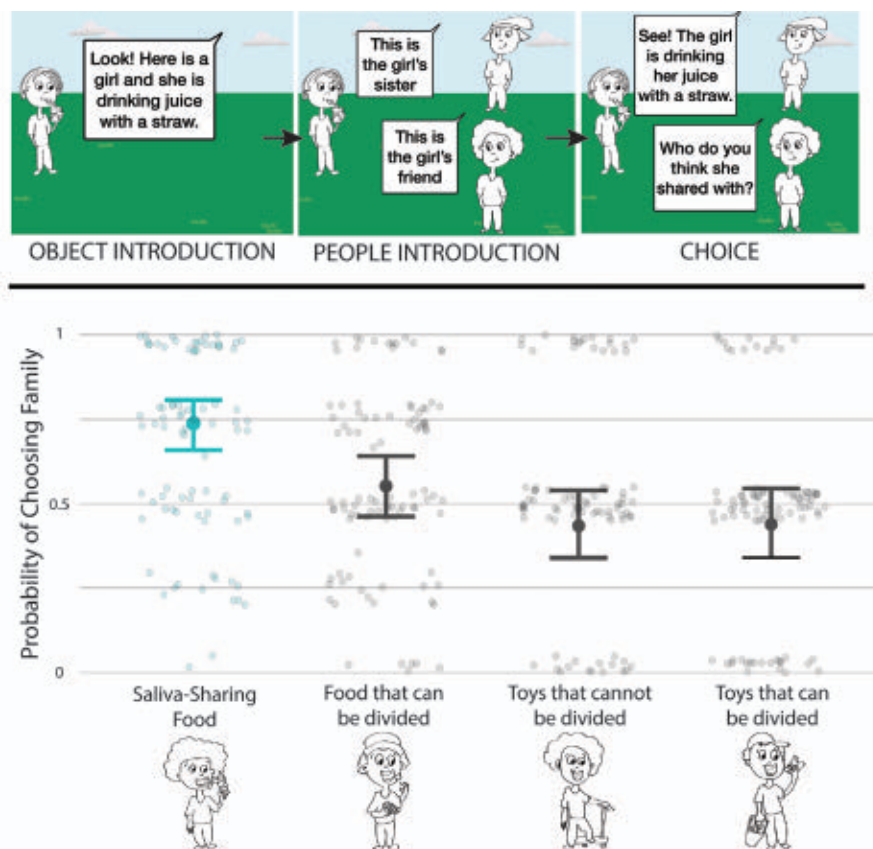
In a first experiment, when presented with interactions between cartoon people, young children (experiment 1,  $N = 113$ , 5 to 7 years old, from an American urban environment) predicted that sharing utensils, or licking the same food item, would occur within nuclear families, whereas sharing toys and partitionable food would occur equally within friend-

ships and families ( $\chi^2 = 72.74$ ,  $P < 0.001$ ; Fig. 1) (18–20). Thus, young children recognize that saliva-sharing interactions distinctively occur within nuclear families.

In the next experiments, we tested whether toddlers and infants would predict that when two individuals have shared saliva, those in-

dividuals will be more emotionally responsive in future interactions (17). This experimental design was inspired by classic studies of vervet monkeys who heard a familiar juvenile in distress and looked toward that juvenile’s mother, as if expecting her to respond (21). We used this design to test whether young humans use a brief observation of saliva sharing to infer a thick relationship between novel individuals whose genetic relatedness is unknown.

Toddlers (experiment 2A,  $N = 26$ , 16.5 to 18.5 months old) and infants (experiment 2B,  $N = 20$ , 8.5 to 10 months old) saw a central puppet alternately eat from the same orange slice with one actress (implying saliva sharing) and play ball with another actress (Fig. 2). Then they saw the puppet seated between the two actresses, expressing distress. We measured which actress participants looked toward first, and longer, as though expecting the actress to react to the puppet’s distress. Both toddlers and infants looked first, and longer, toward the actress who had shared food and saliva with the puppet [first look: 2A toddlers, 20/26,  $BF_{10}$  (Bayes factor) = 10.796; 2B infants, 16/20,  $BF_{10} = 10.306$ ; proportion look: 2A toddlers, mean = 0.774,  $BF_{10} = 149.377$ ; 2B



**Fig. 1. Materials and results for experiment 1.** Top: Sample images and text for one item of the storyboard task. Bottom: Solid dots are average probability estimates of choosing family as opposed to friend in each condition. The bars are 95% credible intervals for each condition (controlling for multiple comparisons and participant age). Open dots are response rates from each individual child. Note that there were four items in each food condition and two items in each toy condition.

<sup>1</sup>Department of Psychology, Harvard University, Cambridge, MA, USA. <sup>2</sup>Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>3</sup>Population Health Sciences Institute, Newcastle University, Newcastle, UK. <sup>4</sup>NSF Center for Brains, Minds and Machines, Cambridge, MA, USA.

\*Corresponding author. Email: ajthomas@mit.edu

infants, mean = 0.675,  $\text{BF}_{10} = 6.084$ ]. These results were replicated in an independent sample of toddlers in the test condition of experiment 2C ( $N = 23$ ) and an independent sample of infants in the test condition of experiment 2D ( $N = 24$ ) (first look: 2C<sub>test</sub> toddlers, 19/23,  $\text{BF}_{10} = 39.471$ ; 2D<sub>test</sub> infants, 22/24,  $\text{BF}_{10} = 2431$ ; proportion look: 2C<sub>test</sub> toddlers, mean = 0.788,  $\text{BF}_{10} = 3853$ ; 2D<sub>test</sub> infants, mean = 0.734,  $\text{BF}_{10} = 132$ ).

Further experiments revealed that toddlers and infants looked toward the food- and saliva-sharing actress (i) only when the central puppet expressed distress and (ii) only when the puppet in distress was the actress' own thick relation. When the central puppet was removed, leaving only the two actresses, toddlers and infants looked at both actresses equally (Fig. 2) (18). When the central puppet was replaced by a new puppet, who then expressed distress, neither infants nor toddlers looked first or longer at the food sharer (first look: 2C<sub>control</sub> toddlers, 11/26, BF<sub>01</sub> = 3.108;

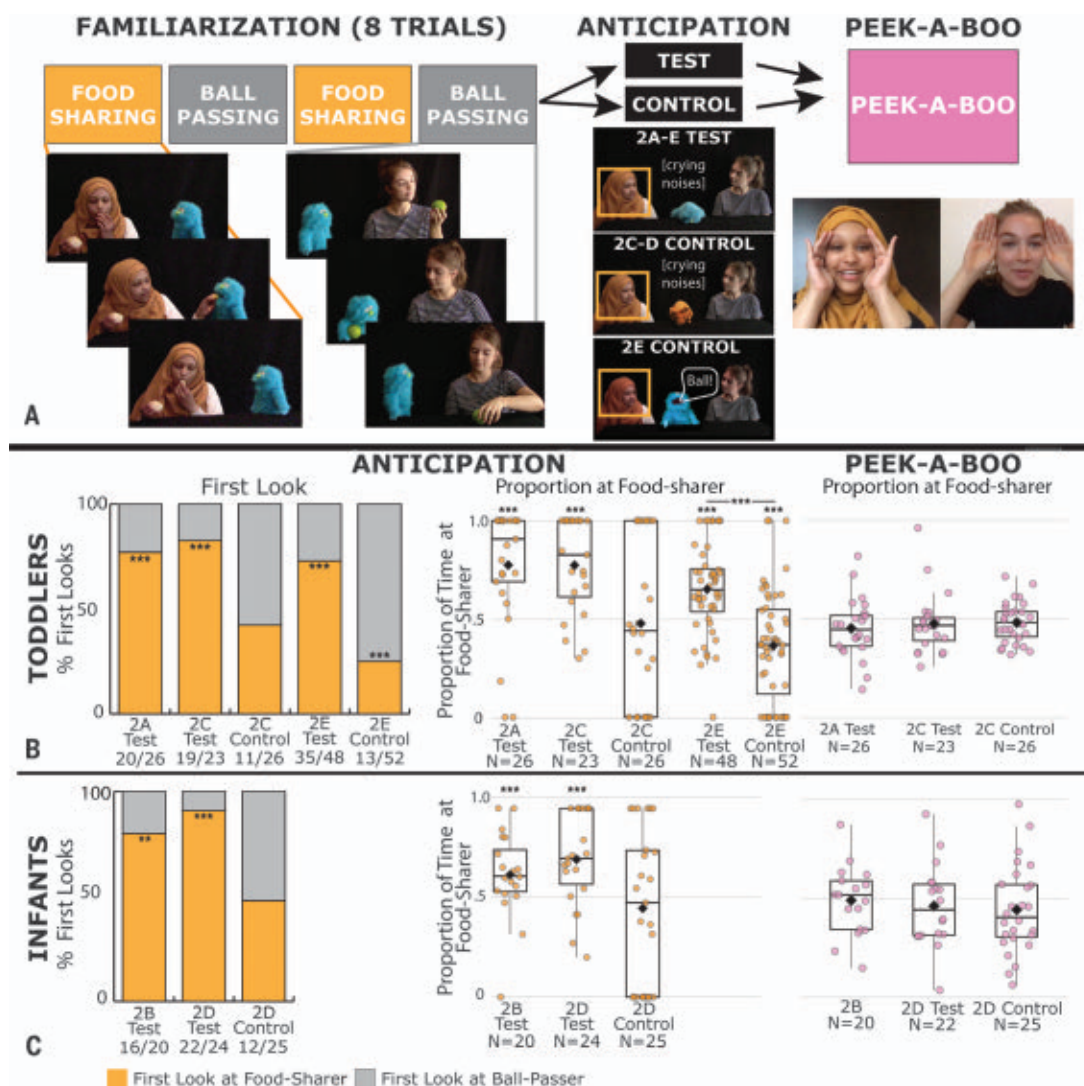
2D<sub>control</sub> infants, 12/25, BF<sub>01</sub> = 4.03; proportion look: toddlers, mean = 0.496, BF<sub>01</sub> = 4.903; infants, mean = 0.470, BF<sub>01</sub> = 4.46). These results suggest that toddlers' and infants' expectations concern the relationship, not the individuals' traits.

For experiment 2E, we recruited a larger, more economically, geographically, and racially diverse sample of toddlers ( $N = 118$ ; age 14.5 to 19 months) (18). The familiarization events were the same. Then, the central puppet either expressed distress as before ( $2E_{\text{test}}$ ) or asked for the ball ( $2E_{\text{control}}$ ). When the central puppet expressed distress, toddlers looked first, and longer, toward the food sharer rather than the ball passer ( $2E_{\text{test}}$ , first look: 35/48,  $BF_{10} = 59.5$ ; proportion look: mean = 0.65,  $BF_{10} = 1002$ ). By contrast, when the puppet requested the ball, toddlers looked first and longer at the ball passer ( $2E_{\text{control}}$ , first look: 13/52,  $BF_{10} = 267$ ; proportion look: mean = 0.37,  $BF_{10} = 54$ ). These conditions differed decisively (first look:  $BF_{10} > 1000$ ; proportion of time:  $BF_{10} >$

1000). Thus, toddlers from a diverse range of households expect that two people who share food and saliva will respond to each other's distress, but not that they will be socially more responsive to one another in general.

In experiment 3, we isolated sharing saliva, without food, as the visible evidence of a thick relationship. Interacting with one puppet, a central actress put her finger in her own mouth, rotated it, put her finger in the puppet's mouth, rotated it, and finally returned her finger to her own mouth. When interacting with a second puppet, the actress performed the same rotating finger actions touching her own and the puppet's forehead. We then measured which puppet infants and toddlers looked toward, when the central actress expressed distress.

Toddlers (age 16.5 to 18.5 months) looked first, and longer, toward the puppet from the mouth-to-mouth interaction when the actress expressed distress (experiment 3A,  $N = 26$ , first look: 20/26,  $BF_{10} = 10.796$ ; proportion look: mean = 0.746,  $BF_{10} = 477.6$ ) (Fig. 3). These



**Fig. 2. Displays and results for experiment 2.** (A) Experimental design flowchart and stills from videos used in experiment 2. The order of the familiarization trials (i.e., food sharing or ball passing first) and the identity of the food sharer were counter-balanced across participants. Participants were randomly assigned to the test or control condition. (B and C) Left: Percentage of participants who looked first toward the food sharer (orange) or ball passer (gray). Center: Proportion of time spent looking at the food sharer during the pause. Black diamonds are means; bars are medians. Right: Proportion of time that participants spent looking at the food sharer during the peek-a-boo trial. \*\*\*Bayes factor of  $>10$ . \*\*Bayes factor of  $>8$ .



results were replicated in an independent sample of toddlers (3C<sub>test</sub>,  $N = 31$ , first look: 25/31,  $BF_{10} = 91.15$ ; proportion look: mean = 0.748,  $BF_{10} = 14,856$ ). Infants' (age 8.5 to 10 months) first looks were distributed equally between both puppets (3B,  $N = 21$ , first look: 11/21,  $BF_{01} = 3.700$ ; replication 3D<sub>test</sub>,  $N = 26$ , 14/26,  $BF_{01} = 3.886$ ). Yet infants did look longer toward the puppet from the mouth-to-mouth interaction while the actress expressed distress (3B, mean = 0.631,  $BF_{10} = 1.55$ ; replication 3D<sub>test</sub>, mean = 0.716,  $BF_{10} = 26.812$ ).

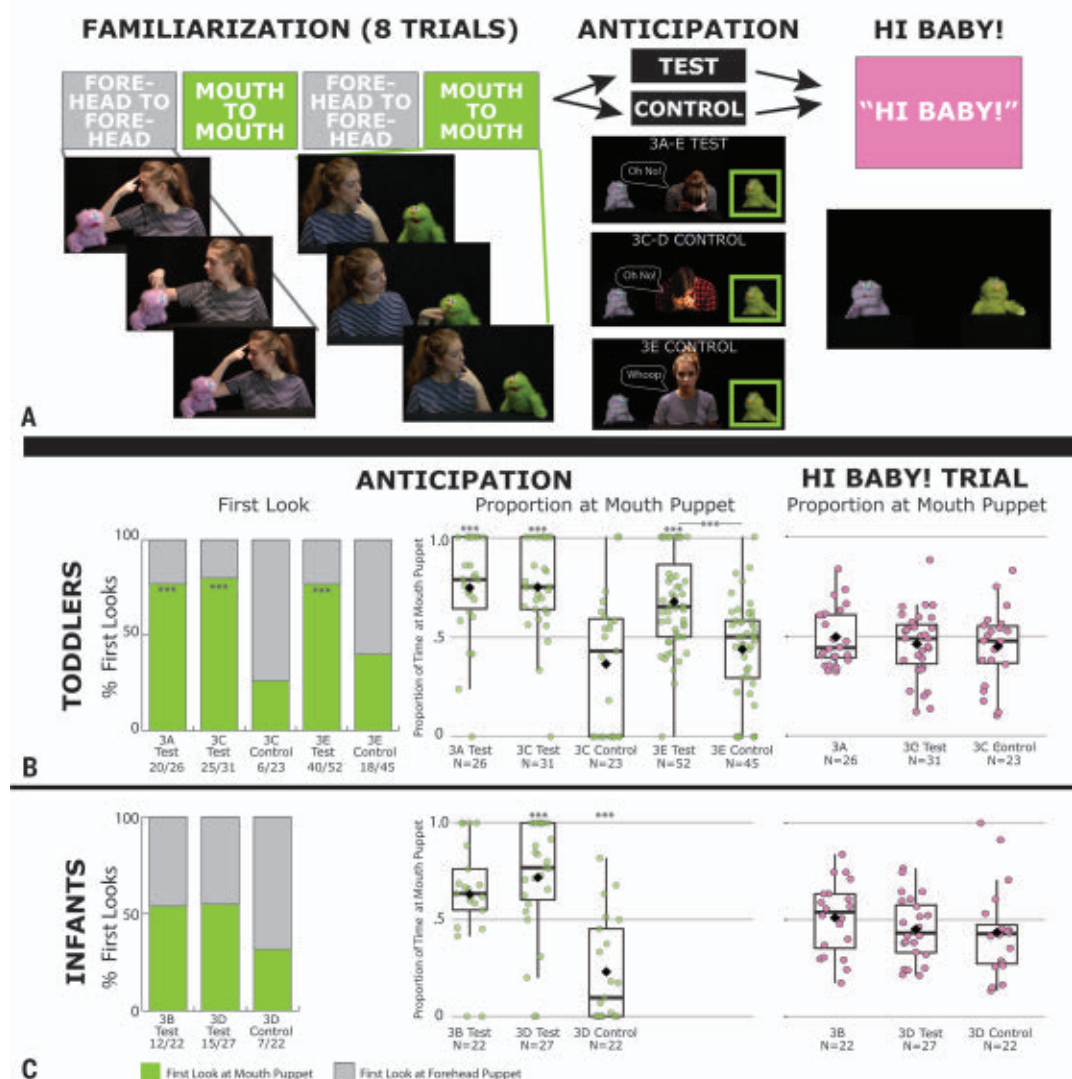
Further experiments revealed that toddlers and infants looked toward the mouth-to-mouth puppet (i) only when the actress in distress was the person in the initial interactions, and (ii) only when the central actress expressed distress. In the control conditions of experiments 3C ( $N = 23$ ) and 3D ( $N = 22$ ), the actress was replaced by a new actor who expressed distress. Here infants and toddlers seemed to expect that the other puppet, who had the forehead-to-forehead interaction with

the first actress, was more likely to react (toddlers 3C<sub>control</sub>, first look: 6/23,  $BF_{10} = 3.462$ ; proportion look: mean = 0.364,  $BF_{10} = 1.023$ ; infants 3D<sub>control</sub>, first look: 8/23,  $BF_{01} = 1.403$ ; proportion look: mean = 0.252,  $BF_{10} = 88.72$ ). An independent group of toddlers (experiment 3E) (18) looked toward the mouth-to-mouth puppet when the central actress expressed distress (3E<sub>test</sub>, first look: 40/52,  $BF_{10} = 823$ ; proportion look: mean = 0.690,  $BF_{10} = 71,669$ ), but not when the actress uttered a nonsense word (3E<sub>control</sub>, first look: 18/45,  $BF_{01} = 2.43$ ; proportion look: mean = 0.426,  $BF_{01} = 0.84$ ). These two conditions differed decisively (first look:  $BF_{10} = 418$ ; proportion look:  $BF_{10} > 1000$ ). Thus, toddlers from a wide range of households expect saliva sharing to selectively predict responses to distress.

The results of experiments 2 and 3 suggest that when toddlers and infants observe two unfamiliar individuals sharing saliva, they infer that those people are in a thick relationship. A separate survey of parents ( $N = 129$ ,

experiment 4) (18) of infants and toddlers (age 8 to 19 months) from the same population suggested that this inference would be valid. The parents expressed comfort with their child having positive social interactions (e.g., playing, reading, hugging) with people in many different relationships, but they expressed comfort with saliva-sharing interactions (i.e., sharing a utensil, drinking from the same cup, kissing on the face) only in relationships the participants assessed as thick ( $BF_{10} > 1000$ ; Fig. 4). In summary, saliva-sharing interactions provide externally observable cues of thick relationships, and young humans can use these cues to make predictions about subsequent social interactions.

Substantial prior research has shown that infants have an "intuitive psychology," supporting inferences about individuals' traits (e.g., cooperative, fair) (22–24), mental states (e.g., goals, perceptions) (25–27), and group membership (e.g., ethnicity, language) (28–32). By contrast, representations of social relationships



**Fig. 3. Displays and results for experiment 3.** (A) Experimental design flowchart and stills from videos used in experiment 3. (B and C) Left: Percentage of participants who looked first toward the puppet and who had engaged in the mouth-to-mouth interaction (green) or forehead-to-forehead interaction (gray). Center: Proportion of time spent looking at the mouth-to-mouth puppet during the pause. Black diamonds are means; bars are medians. Right: Proportion of time that participants spent looking at the mouth-to-mouth puppet during the "hi baby, hi" trial. \*\*\*Bayes factor of  $>10$ .

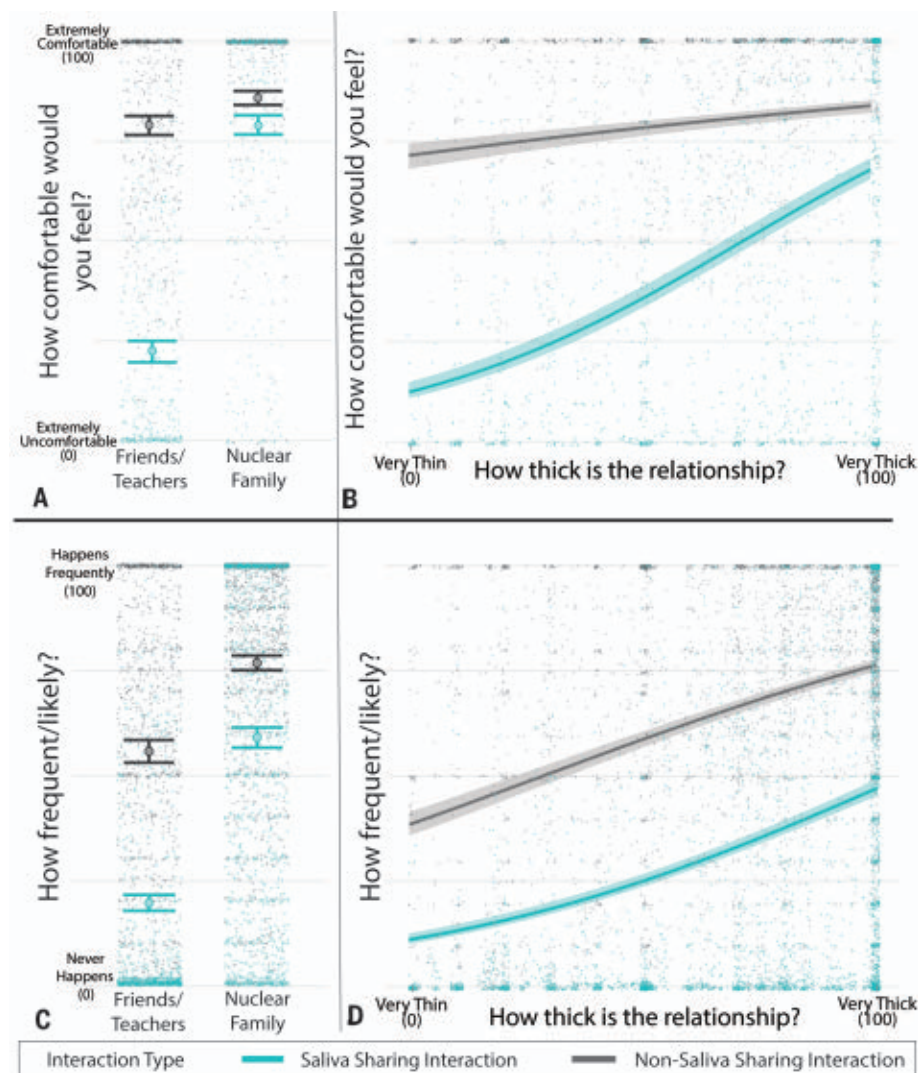


and social structures, which are a central concern of sociology and anthropology, have been relatively understudied using experimental techniques in infants (33–35). The key distinction revealed here, between thick and thin relationships, may be one foundation of infants' "intuitive sociology" (1, 2).

School-aged children's judgments about saliva sharing are likely reinforced by explicit prohibitions (particularly during a pandemic), but similar intuitions appear to originate earlier and to generalize beyond the content of verbal rules motivated by hygiene (18). We hypothesize that an early intuitive distinction between thick and thin relationships allows infants to rapidly learn the distinctive behaviors that occur in these relationships in their social environment (14, 15). These rapidly bootstrapped representations would be useful for parsing the small set of thick, intimate relationships from the larger set of thin, cooperative relationships in typical human social networks (14, 15).

Our experiments have limitations. We have not established whether thick and thin relationships are conceived as qualitatively distinct categories [e.g., (5, 7, 36)] or as ends of a continuum, with close friends or confidants having a mixture of features (37, 38). Saliva sharing is likely only one example of the set of interactions that children and infants can use to distinguish thick relationships from other cooperative relationships. Interactions distinctively occurring in thick relationships include confiding and emotional comforting, consensual exchange of blood, and touching of genitals (15). Infants may also have expectations about the social categories of the people in thick relations [for example, older women may be more likely than older men to be in thick relationships with infants (39)]. Conversely, some interactions involving transfer of saliva are aggressive and demeaning, such as spitting on a person (18). We have not established whether toddlers and infants can distinguish between saliva-sharing interactions that are consensual versus coercive.

Our evidence that young children, toddlers, and infants make distinctive inferences about thick relationships has broad theoretical implications. Anthropologists have stressed that thick relationships, characterized by strong attachment, obligation and mutual responsiveness, have highly variable mappings onto genetic relatedness (10, 40). We have identified a proximal psychological mechanism that might allow infants to parse this variability (1, 3). A consistent inferential mechanism, operating in diverse social and ecological settings, could enable young humans to rapidly acculturate to their local and culturally specific relationship and kinship structures. In sum, children, toddlers, and infants recognize saliva-sharing interactions as distinctive evi-



**Fig. 4. Results for experiment 4.** (A to D) Parent-reported comfort [(A and B)] and frequency [(C and D)] of saliva-sharing interactions, versus other caregiving and prosocial interactions, in their child's social environment as a function of relationship thickness. We operationalize relationship thickness as nuclear family [versus friends and teachers; (A) and (C)] or by directly asking parents to report the thickness of many of the child's relationships [e.g., including aunt, grandparent, etc.; (B) and (D)].

dence of a thick relationship. The pattern of who does, and who does not, share saliva may help infants to distinguish those who are kin (e.g., parents, siblings, grandparents) versus non-kin (e.g., daycare teachers, nannies) among their many caregivers (41, 42). Young humans may use observations of saliva sharing to inform their earliest understanding of the conceptual structure of family.

#### REFERENCES AND NOTES

1. L. Kaufmann, F. Clément, *Topoi* **33**, 459–475 (2014).
2. L. Thomsen, S. Carey, in *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us* (Oxford Univ. Press, 2013), p. 17.
3. L. A. Hirschfeld, *Pers. Soc. Psychol. Rev.* **5**, 107–117 (2001).
4. A. Margalit, *On Betrayal* (Harvard Univ. Press, 2017).
5. M. S. Clark, J. Mills, *J. Pers. Soc. Psychol.* **37**, 12–24 (1979).
6. T. S. Rai, A. P. Fiske, *Psychol. Rev.* **118**, 57–75 (2011).
7. M. Sahlins, *Stone Age Economics* (Taylor & Francis, 2017).

8. D. M. Schneider, *A Critique of the Study of Kinship* (Univ. of Michigan Press, 1984).
9. M. D. Sahlins, *The Use and Abuse of Biology: An Anthropological Critique of Sociobiology* (Univ. of Michigan Press, 1976).
10. D. Jones, *Evol. Hum. Behav.* **24**, 303–319 (2003).
11. L. Cronk, D. Steklis, N. Steklis, O. R. van den Akker, A. Aktipis, *Evol. Hum. Behav.* **40**, 281–291 (2019).
12. J. Carsten, *Am. Ethnol.* **22**, 223–241 (1995).
13. L. A. Hirschfeld, *Int. J. Behav. Dev.* **12**, 541–568 (1989).
14. D. Cohen, A. K.-Y. Leung, *Eur. J. Soc. Psychol.* **39**, 1278–1289 (2009).
15. A. P. Fiske, L. Thomsen, S. M. Thein, *Eur. J. Soc. Psychol.* **39**, 1294–1297 (2009).
16. A. Sorokowska et al., *Pers. Soc. Psychol. Bull.* **47**, 1705–1721 (2021).
17. L. Miller, P. Rozin, A. P. Fiske, *Eur. J. Soc. Psychol.* **28**, 423–436 (1998).
18. See supplementary materials.
19. K. R. Olson, E. S. Spelke, *Cognition* **108**, 222–231 (2008).
20. A. C. Spokes, E. S. Spelke, *Front. Psychol.* **7**, 440 (2016).
21. R. M. Seyfarth, D. L. Cheney, in *Animal Social Complexity: Intelligence, Culture, and Individualized Societies*, F. B. M. de Waal, P. L. Tyack, Eds. (Harvard Univ. Press, 2003), pp. 207–229.
22. J. K. Hamlin, *Curr. Dir. Psychol. Sci.* **22**, 186–193 (2013).

23. L. J. Powell, E. S. Spelke, *Open Mind* **2**, 61–71 (2018).
24. A. Geraci, L. Surian, *Dev. Sci.* **14**, 1012–1020 (2011).
25. A. L. Woodward, *Cognition* **69**, 1–34 (1998).
26. S. Carey, *The Origin of Concepts* (Oxford Univ. Press, ed. 1, 2009).
27. S. Liu, T. D. Ullman, J. B. Tenenbaum, E. S. Spelke, *Science* **358**, 1038–1041 (2017).
28. K. D. Kinzler, E. Dupoux, E. S. Spelke, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12577–12580 (2007).
29. Z. Liberman, A. L. Woodward, K. D. Kinzler, *Cogn. Sci.* **41** (suppl. 3), 622–634 (2017).
30. Y. Bar-Haim, T. Ziv, D. Lamy, R. M. Hodes, *Psychol. Sci.* **17**, 159–163 (2006).
31. L. J. Powell, E. S. Spelke, *Proc. Natl. Acad. Sci. U.S.A.* **110**, E3965–E3972 (2013).
32. F. Ting, Z. He, R. Baillargeon, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6025–6034 (2019).
33. L. Thomsen, W. E. Frankenhuis, M. Ingold-Smith, S. Carey, *Science* **331**, 477–480 (2011).
34. A. J. Thomas, B. W. Sarnecka, *Curr. Biol.* **29**, 2183–2189.e5 (2019).
35. Z. Liberman, A. L. Woodward, K. R. Sullivan, K. D. Kinzler, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9480–9485 (2016).
36. A. P. Fiske, *Psychol. Rev.* **99**, 689–723 (1992).
37. J. B. Silk, in *Genetic and Cultural Evolution of Cooperation*, P. Hammerstein, Ed. (MIT Press, 2003), pp. 37–54.
38. M. L. Small, *Someone to Talk To* (Oxford Univ. Press, 2017).
39. H. P. Alvarez, *Am. J. Phys. Anthropol.* **113**, 435–450 (2000).
40. G. P. Murdock, *Ethnology* **9**, 165–208 (1970).
41. S. B. Hrdy, J. M. Burkart, *Philos. Trans. R. Soc. London Ser. B* **375**, 20190499 (2020).
42. S. B. Hrdy, *Nat. Hist.* **110**, 50 (2001).

#### ACKNOWLEDGMENTS

We thank C. Lu, S. Ravikumar, M. Austin, V. Kudrnova, S. Alansaari, W. Pepe, A. Harris, and M. O. Ali for assistance with data collection and video coding; L. Mullertz, S. Dablouk, and R. Van Dine for acting in the videos; E. Chen for assistance with trimming videos, recruitment, and editing drafts of this manuscript; and L. Schulz, T. Ullman, S. Liu, H. Olson, L. Powell, and M. Hung for comments on earlier drafts of the manuscript. Statistical support was provided by data science specialist S. Worthington at the Institute for Quantitative Social Science, Harvard University. **Funding:** NIH National Research Service Award 1F32HD096829 (A.J.T.); Patrick J. McGovern Foundation (R.S.); Guggenheim Foundation (R.S.); Social Sciences and Humanities Research Council Doctoral Fellowship 752-2020-0474

(B.W.); NSF Center for Brains Minds and Machines award CCF-1231216 (E.S., A.J.T., B.W., R.S.); Siegel Foundation award S4881 (E.S.) **Author contributions:** Conceptualization: A.J.T., R.S., E.S., D.N. Methodology: A.J.T., R.S., E.S., B.W., D.N. Formal Analysis: A.J.T. Resources: E.S., R.S. Investigation: A.J.T. Visualization: A.J.T. Funding acquisition: A.J.T., R.S., E.S. Project administration: A.J.T. Supervision: A.J.T., R.S., E.S. Writing—original draft: A.J.T., R.S. Writing—review and editing: A.J.T., R.S., D.N., B.W., E.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available on OSF (<https://osf.io/a8htx>). A subset of participant videos are available on Databrary, <https://nyu.databrary.org/volume/1253>.

#### SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh1054  
Materials and Methods  
Supplementary Text  
Fig. S1  
References (43–53)

16 February 2021; accepted 7 December 2021  
10.1126/science.abh1054

## RESEARCH TECHNOLOGY

# High-speed fluorescence image-enabled cell sorting

Daniel Schraivogel<sup>1</sup>, Terra M. Kuhn<sup>2†</sup>, Benedikt Rauscher<sup>1†</sup>, Marta Rodríguez-Martínez<sup>1†</sup>, Malte Paulsen<sup>3†</sup>, Keegan Owsley<sup>4</sup>, Aaron Middlebrook<sup>4</sup>, Christian Tischer<sup>5</sup>, Beáta Ramasz<sup>3</sup>, Diana Ordoñez-Rueda<sup>3</sup>, Martina Dees<sup>2</sup>, Sara Cuylen-Haering<sup>2\*</sup>, Eric Diebold<sup>4\*</sup>, Lars M. Steinmetz<sup>1,6,7\*</sup>

Fast and selective isolation of single cells with unique spatial and morphological traits remains a technical challenge. Here, we address this by establishing high-speed image-enabled cell sorting (ICS), which records multicolor fluorescence images and sorts cells based on measurements from image data at speeds up to 15,000 events per second. We show that ICS quantifies cell morphology and localization of labeled proteins and increases the resolution of cell cycle analyses by separating mitotic stages. We combine ICS with CRISPR-pooled screens to identify regulators of the nuclear factor  $\kappa$ B (NF- $\kappa$ B) pathway, enabling the completion of genome-wide image-based screens in about 9 hours of run time. By assessing complex cellular phenotypes, ICS substantially expands the phenotypic space accessible to cell-sorting applications and pooled genetic screening.

**F**luorescence microscopy and flow cytometry are instrumental technologies used in almost all areas of biological and biomedical research. Although flow cytometric cell sorting simplifies the isolation of cells in a rapid, sensitive, and high-throughput manner, it is limited to a low-dimensional parameter space and lacks subcellular resolution (*1*). This method is therefore unable to capture phenotypes associated with processes involving varying signal localization, such as protein trafficking, cellular signaling, or

protein mislocalization during disease (*2, 3*). Fluorescence microscopy, on the other hand, enables high-resolution readouts of cellular morphology and protein localization but lacks the ability to isolate cells with specific phenotypes at high speed (*4*). Combining the spatial resolution of fluorescence microscopy with flow cytometric cell sorting has broad implications and would inspire new experimental strategies through the rapid identification and isolation of cells with specific (sub)cellular phenotypes.

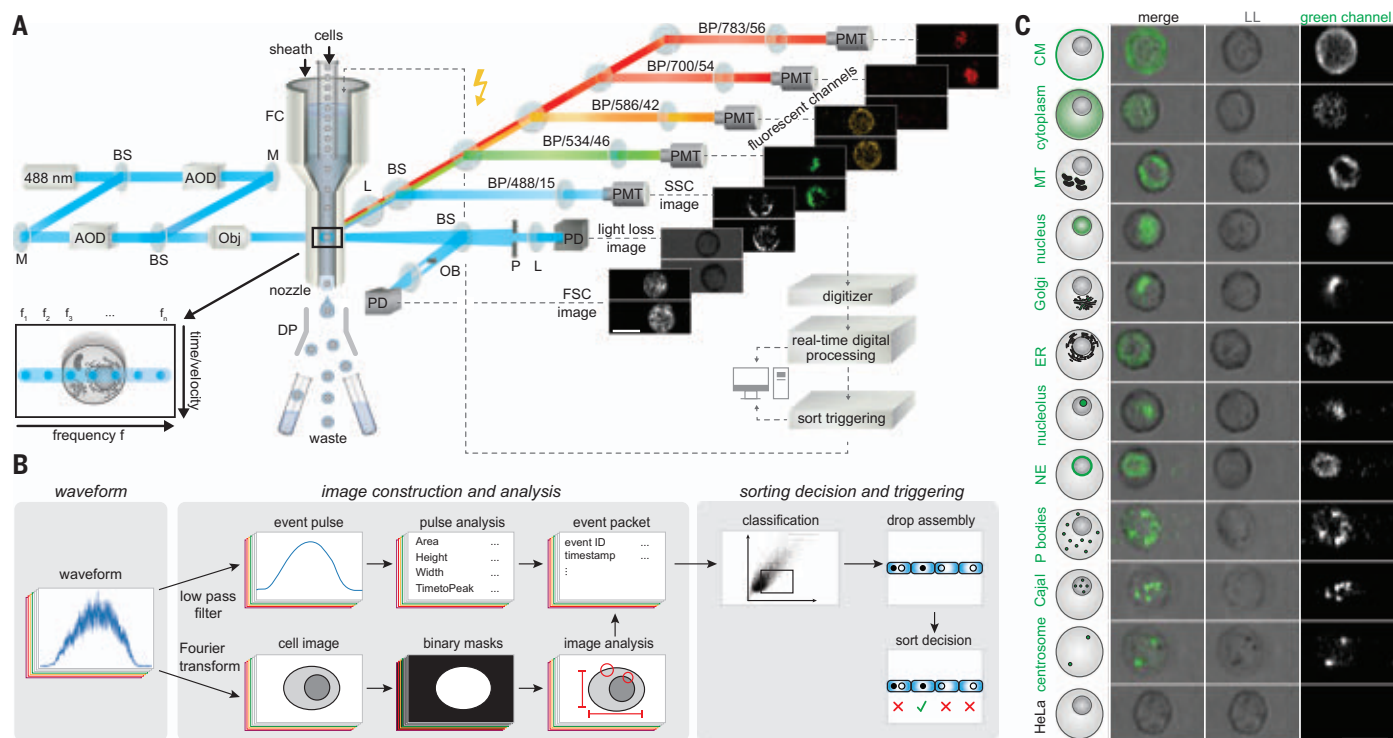
Although flow- and microfluidics-based cytometers with imaging capabilities have been developed, these approaches were unable to sort cells, came with drastically reduced throughput, or depended on nonhuman interpretable pattern recognition from raw data without image reconstruction (*5–14*). Furthermore, image-enabled cell sorting has so far relied on technically challenging and custom-built solutions. To date, no system has been developed that integrates traditional flow cytometry and microscopy, operates at speeds compatible with genetic screening approaches and short-lived

dynamic phenotypes, and can be operated in nonspecialized laboratories.

Here, we present a fully integrated image-enabled cell sorter (ICS) by combining (i) fluorescence imaging using radiofrequency-tagged emission (FIRE), a fast fluorescence imaging technique (*15*), with (ii) a traditional cuvette-based droplet sorter and (iii) new low-latency signal processing and sorting electronics (Fig. 1, A and B; for a detailed description and characterization of ICS technology, please see the materials and methods and fig. S1; for a description of the performance attributes of ICS, please see the supplementary text). To enable blur-free imaging at a high nominal flow speed of 1.1 m/s, ICS uses the FIRE approach to produce an array of 104 laser spots across 60  $\mu$ m within the core stream of the sorter cuvette, each modulated at a unique radiofrequency (Fig. 1A). The array of spots excites modulated fluorescent and scattered light from particles or cells as they flow through the optical interrogation region in the cuvette. Emitted light is collected, and the signal output is digitized and processed using low-latency, field-programmable gate arrays, allowing real-time image analysis and image-derived sort decisions. This is different from other image-enabled flow cytometers without cell-sorting capabilities (*5–8, 11–13*) (see the supplementary text for a comparison between technologies). To reconstruct a row of pixels from the FIRE signal for visualization of the event, the amplitude of the signal at a unique modulation frequency is assigned to a pixel value in a specific horizontal coordinate in the cuvette; in the direction of flow, the pixels are assigned a vertical location based on their temporal value, which forms a two-dimensional image of an event (Fig. 1A). The system collects scatter and fluorescent signals, as well as a light loss signal (analogous to bright-field images produced by traditional microscopes), which allows visualization of events in real

<sup>1</sup>Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. <sup>2</sup>Cell Biology and Biophysics Unit, EMBL, Heidelberg, Germany. <sup>3</sup>Flow Cytometry Core Facility, EMBL, Heidelberg, Germany. <sup>4</sup>BD Biosciences, San Jose, CA, USA. <sup>5</sup>Advanced Light Microscopy Core Facility, EMBL, Heidelberg, Germany. <sup>6</sup>Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. <sup>7</sup>Stanford Genome Technology Center, Palo Alto, CA, USA.  
\*Corresponding author. Email: [lars.steinmetz@embl.de](mailto:lars.steinmetz@embl.de) (L.M.S.); [eric.diebold@bd.com](mailto:eric.diebold@bd.com) (E.D.); [sara.cuylen-haering@embl.de](mailto:sara.cuylen-haering@embl.de) (S.C.-H.).  
†These authors contributed equally to this work  
‡Present address: Novo Nordisk Foundation Center for Stem Cell Medicine, reNEW, Copenhagen, Denmark.





**Fig. 1. Functionality of the ICS. (A)** Schematic representation of the ICS optical and flow hardware components. Excitation beam path: The acousto-optic deflector (AOD) splits a single laser beam ( $\lambda = 488 \text{ nm}$ ) into an array of beamlets, each having different optical frequency and angle. A second AOD tunes the optical frequency of a reference beam, which is then overlapped with the array of beamlets. The overlapping beams intersect the flow cell (FC) of a cuvette sorter. Inset left side: The array of FIRE beams (dark cyan) are shown overlapping with the reference beam (light cyan). Because of their differing optical frequencies, the overlapping beams exhibit a beating behavior, which causes each beamlet to carry a sinusoidal modulation at a distinct frequency  $f_{1-n}$ . Emission beam path: Images are generated from digitized signals on a per-event basis and include light loss, forward scatter (FSC), and side scatter (SSC) images, and four different fluorescent channels. Example images: HeLa cells expressing the Golgi marker GalNAcT2-green fluorescent protein (GFP) (green) were stained with cell surface marker CD147 PE-CF594 (orange) and DRAQ5 nuclear dye (red). FSC, SSC, and light loss images are shown in grayscale. BS, beam splitter; M, mirror; Obj, objective; DP, deflection plates; OB, obscuration bar; P, pinhole; L, lens; BP, band pass; PMT, photomultiplier tube; PD, photodiode. Scale bar,  $20 \mu\text{m}$ . **(B)** Overview of the ICS low-latency data-processing pipeline. Each

photodetector produces a pulse with high-frequency modulations encoding the image (waveform). Fourier analysis is performed to reconstruct the image from the modulated pulse. An image-processing pipeline produces a set of image features (image analysis), which are combined with features derived from a pulse-processing pipeline (event packet). Real-time sort classification electronics then classify the particle on the basis of image features, producing a sort decision that is used to selectively charge the droplets (dotted gray line in A). **(C)** ICS-based imaging of HeLa cells expressing GFP- or mNG-tagged fluorescent proteins or stained with organelle-specific green fluorescent dyes. One representative image is shown per organelle; the full datasets containing 10,000 images each are shared as described in the data and materials availability section. The following dyes or protein fusions were used: cell membrane (Cellmask dye), cytoplasm (GFP fused to HIV Rev nuclear export sequence), mitochondria (Mitotracker dye), nucleus (H2B-mNG), Golgi apparatus (GalNAc2-GFP), endoplasmic reticulum (ER, ERtracker dye), nucleolus (eGFP-Ki-67), nuclear envelope (LamB1-GFP), P bodies (eGFP-DDX6), Cajal bodies (eGFP-COIL), and centrosomes (anti-pericentrin antibody). P bodies and Cajal bodies were recorded from fixed cells, centrosomes from fixed and metaphase-stalled cells; fixation resulted in decreased contrast in the light loss (LL) image. Scale bar, 20  $\mu$ m.

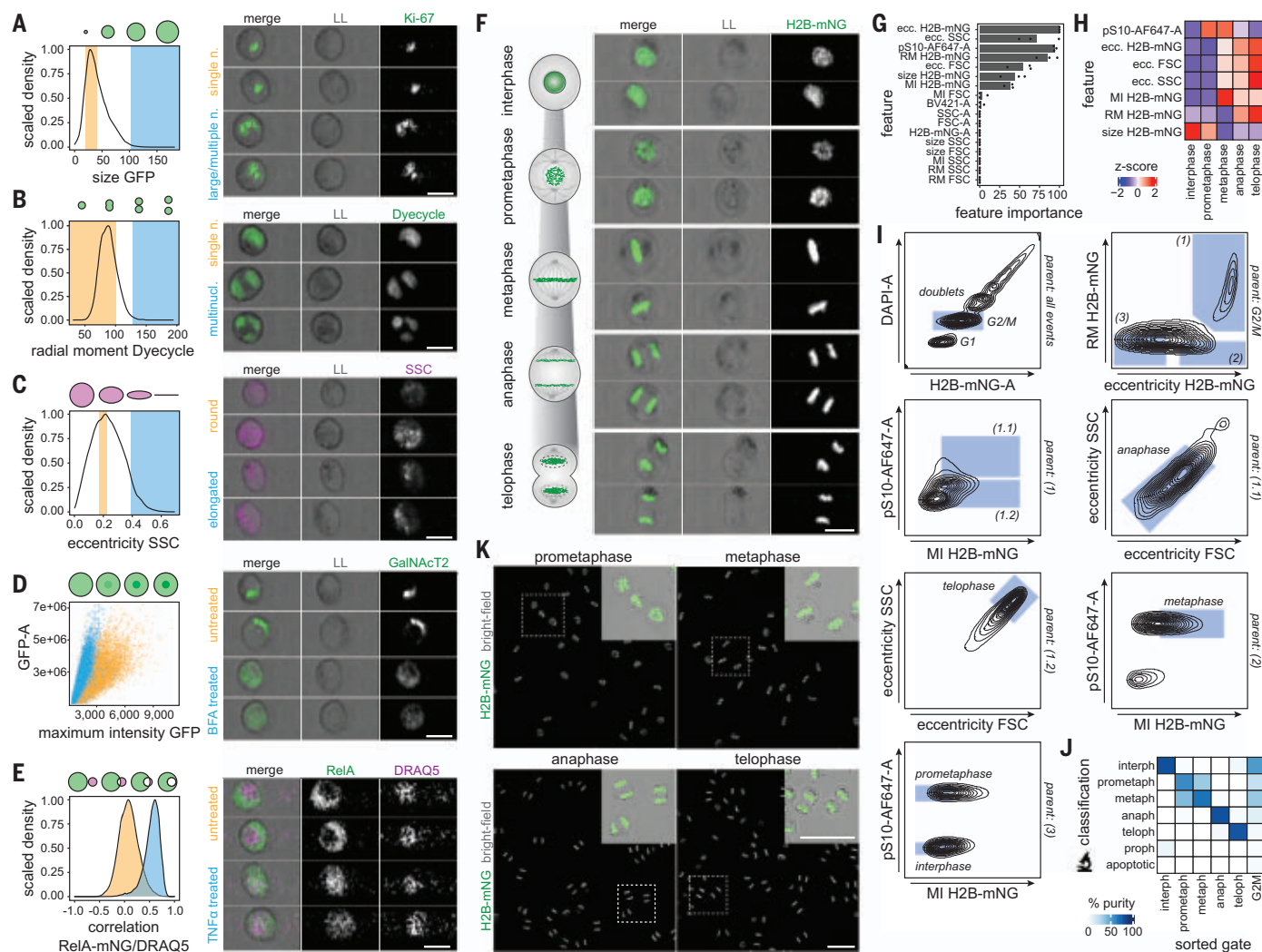
time. This contrasts with ghost cytometry, which is unable to reconstruct images from raw data in real time (10). The combination of FIRE with a cuvette-based droplet-sorter design, along with the integrated low-latency electronics, enables sorting rates at speeds of up to 15,000 events/s (fig. S1, A to C), which is comparable to traditional cell sorters and approximately one order of magnitude faster than image-activated cell sorting (9, 14). Image acquisition and high sorting rates allow immediate human interpretation of the generated data, the capture of dynamic short-lived spatial phenotypes, and the retrieval of sufficient cell numbers for downstream assays such as genome-scale screens.

To illustrate the utility of ICS for blur-free visualization of fast-flowing cells and subcellular protein distribution, we imaged a range of well-known organelles and structures of different sizes, shapes, and distributions. We were able to visualize the cell membrane, cytoplasm, membrane-enclosed organelles (nucleus, endoplasmic reticulum, Golgi apparatus, and mitochondria), and small membrane-less organelles (P bodies, Cajal bodies, and centrosomes) (Fig. 1C and fig. S2). We further demonstrated imaging of 13 cell lines of variable size and origin (fig. S3), showing the broad applicability of ICS.

For cell sorting, a set of intuitive spatial image parameters were extracted in real time

from each image channel (Fig. 1B; for details of the image parameters, please see the materials and methods and fig. S4A). Image parameters were treated identically to conventional pulse parameters (area, width, and height) by the sorting electronics, allowing the combination of spatial information and traditional flow cytometry features for analysis and sorting. We demonstrate the ability of ICS parameters to quantify spatial features and to differentiate cells in a variety of applications that previously could only be distinguished using microscopy. We were able to separate cells with single or multiple/enlarged nucleoli (Fig. 2A and fig. S4C), single or multiple nuclei (Fig. 2B and fig. S4D), and distinguish cells based on cellular

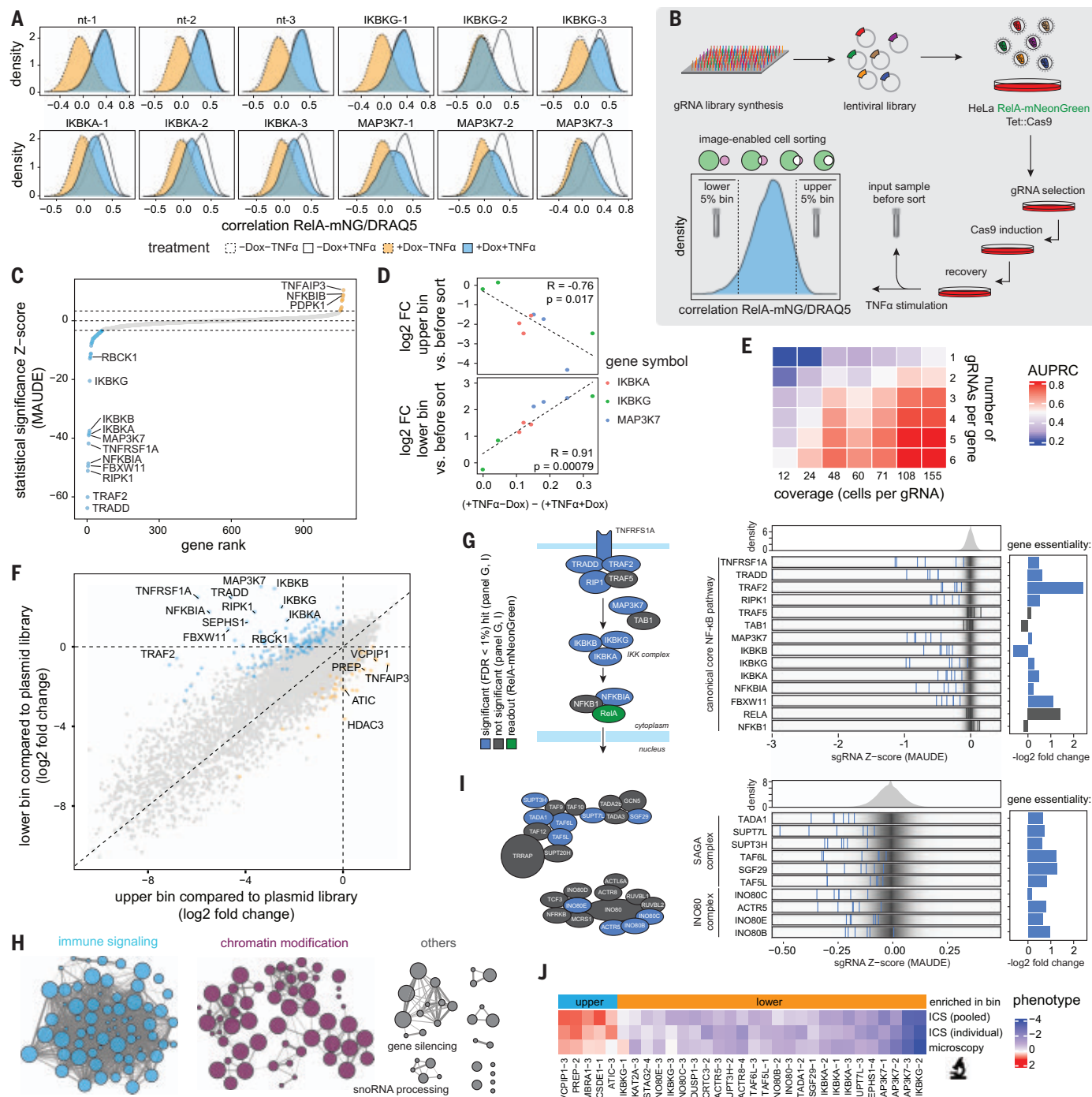




**Fig. 2. ICS measurements quantify spatial cellular processes and isolate phenotypes of interest.**

(A) HeLa cells expressing eGFP-Ki-67 were gated for singlets and live cells, and the ICS size parameter of the eGFP-Ki-67 signal was used to distinguish between cells with single small nucleoli and those with multiple or large nucleoli. Size is defined by the number of pixels above a user-defined threshold. n, nucleolus. Scale bar, 20  $\mu$ m. (B) HeLa cells stained with the nuclear dye DyeCycle Green were gated for singlets and live cells, and the radial moment of DyeCycle Green was used to differentiate cells with single or multiple nuclei. Radial moment is the mean-square distance of the signal from the centroid. n, nucleus. Scale bar, 20  $\mu$ m. (C) HeLa cells were gated for singlets and live cells, and the eccentricity calculated from the side scatter image was used to distinguish round from elongated cells. Eccentricity was computed by first finding the magnitudes of the spread along the two principal components of the image, then taking their ratio. Scale bar, 20  $\mu$ m. (D) HeLa cells expressing the Golgi marker GalNAc2-GFP were gated for singlets and live cells and either treated with brefeldin A (BFA) or left untreated. The maximum intensity of the GalNAc2-GFP channel was used to distinguish treated from untreated cells, whereas the overall GFP intensity (y axis) was largely unaffected by the treatment. Maximum intensity is the value of the brightest pixel. A, area. Scale bar, 20  $\mu$ m. (E) HeLa cells expressing RelA-mNG were treated with TNF $\alpha$  or left untreated and stained with the cell-permeable nuclear dye DRAQ5. Cells were then gated for singlets and live cells, and the correlation between RelA-mNG and DRAQ5 was used to differentiate between the treated (nuclear RelA) and untreated (cytoplasmic RelA) conditions. Correlation is the Pearson's correlation score between the intensities of the pixel values from

two imaging channels. Scale bar, 20  $\mu$ m. (F) HeLa cells expressing H2B-mNG were synchronized to increase the frequency of rare mitotic stages and released into mitosis without chemical perturbation. Then, cells were fixed for labeling with an antibody recognizing phosphorylated serine 10 on histone H3 (pS10H3) to allow microscopic validation after sorting. Samples were stained with 4',6-diamidino-2-phenylindole (DAPI) for univariate cell cycle analysis. Representative images of individual cells within the G<sub>2</sub>/M population reveal captures of major mitotic stages. LL, light loss. Scale bar, 20  $\mu$ m. (G) A decision tree model was trained to distinguish the mitotic stages of manually classified datasets ( $n = 100$  per stage, three replicate recordings and classifications). Shown are the results of a feature importance analysis of ICS measurements representing the summarized reduction in the loss function attributed to each feature at each split in the tree. RM, radial moment; ecc, eccentricity; MI, maximum intensity. (H) Feature values from (G) were standardized, and median values for cells and from three replicates of classified datasets are shown as a heatmap. Only features that vary between the mitotic stages are shown [variable importance >0 in (G)]. (I) On the basis of the identified features in (H), a hierarchical gating strategy was built that enriches for interphase, prometaphase, metaphase, anaphase, and telophase stages. (J) A total of 5000 cells were sorted for microscopic validation based on the gating strategy established in (I), and manual classification from confocal z stacks of the sorted cells was performed. Shown are mean percentages of three independent replicates. Prometaphase cells were generated by two consecutive sorts (see the materials and methods). interph., interphase. (K) Representative single-slice confocal fluorescence microscopy images from sorted cells from (J) with bright-field/H2B-mNG overlays as inlays. Scale bar, 50  $\mu$ m.



**Fig. 3. ICS detects the effects of CRISPR perturbations and enables pooled genetic screens of protein localization. (A)** Effects of individual CRISPR perturbations on RelA nuclear translocation. HeLa cells with Tet-inducible Cas9 and stably expressing RelA-mNG were transduced with gRNA-1, gRNA-2, and gRNA-3 targeting the core NF- $\kappa$ B pathway proteins IKK $\beta$ , IKK $\alpha$ , and MAP3K6, respectively, or with nontargeting (nt) control gRNAs. gRNA expression was induced with doxycycline (Dox) or left uninduced. Correlation between RelA-mNG and DRAQ5 was quantified using ICS as a measurement for RelA nuclear translocation in the presence or absence of TNF $\alpha$ . **(B)** Overview of the pooled CRISPR screening setup and readout using ICS. Positive regulators of RelA nuclear translocation are enriched in the lower bin and depleted from the upper bin. Tet::Cas9, tetracycline/doxycycline-inducible Cas9. **(C to E)** Results

of the ICS-based CRISPR screen using an NF- $\kappa$ B pathway-focused library ( $n = 1068$  genes). (C) The screen was performed at different library coverages, and reads from collected samples were combined in silico to a high-coverage (359 cells per gRNA per sorted bin) dataset. Hits were called using the software MAUDE (26). Genes are ranked by their statistical significance and selected positive/negative regulators are highlighted. The horizontal dashed lines indicate an FDR of 1%, whereas genes with FDR  $<1\%$  are marked in cyan and orange, respectively. (D) Comparison of phenotypes measured in individual perturbation experiments from (A) ( $x$  axis) or the pooled screen ( $y$  axis) using the same gRNAs. For the pooled screen, differences in gRNA abundance in the upper (top panel) and lower (bottom panel) sorted bins compared with the input sample were determined from the high-coverage dataset in (C).  $R$  values represent

Pearson correlation coefficients. FC, fold change. (E) Screen hits as determined at different library coverages (12 to 155 cells per gRNA per sorted bin) using between one and six gRNAs per gene were compared with a high-coverage reference sample (359×, six gRNAs per gene) by precision-recall analysis. Heatmap shows AUPRC values for different levels of library coverage and different numbers of gRNAs per gene. (F to J) Results of the ICS-based genome-wide screen ( $n = 18,408$  genes). (F) Scatter plot of fold changes visualizing gRNA abundance changes in upper ( $x$  axis) and lower ( $y$  axis) sorted bins compared with the plasmid library. Cyan and orange dots indicate statistically significant positive and negative regulators, respectively (FDR <1% according to MAUDE). (G) Genome-wide CRISPR screen identified core canonical NF- $\kappa$ B pathway components. Left panel: Schematic of the core canonical NF- $\kappa$ B signaling pathway. Right top panel: Distribution of the gRNA Z-score for the whole genome-wide library. Right panels: gRNA Z-score for individual gRNAs per gene overlaid with a gradient (grayscale) depicting overall Z-score distribution. Right bar chart: Gene essentiality as determined by the  $\log_2$  FC of the gRNA abundance in the unsorted cell population

compared with the plasmid library. (H) GO network of hits with FDR <1%, colored by modules identified from protein–protein interactions using STRING-db (45). Gray lines connect associated GO terms, edges represent GO terms. Names of individual edges were omitted, clusters that were not associated with immune signaling or chromatin modification were collected in a third class called “others.” (I) Screen results for SAGA and INO80 protein complex components. Left panel: Schematic illustration of the SAGA and INO80 protein complexes. Right panels: As described in (G). (J) Selected hits from the genome-wide screen (one gRNA per gene; we picked the gRNA that showed the strongest Z-score in the pooled genetic screen) were validated using two orthologous methods (individual validation using ICS, and individual validation using microscopy). The top row in the heatmap shows the phenotypes measured in the genome-wide screen (MAUDE Z-score). The phenotype in the second and third rows of the heatmap represents the standardized difference in signal medians between the knockout and control gRNA cell populations. Nuclear RelA abundance was quantified using microscopy by measuring the correlation between RelA-mNG and DRAQ5.

shape (Fig. 2C and fig. S4E). We also demonstrate the ability of ICS to reveal drug-induced organelle responses, such as the effect of brefeldin A on Golgi integrity (Fig. 2D and fig. S4F). Finally, we demonstrate the advantage of multicolor fluorescence imaging for quantification of protein localization through spatial correlation of two signals. We quantified the translocation of the nuclear factor  $\kappa$ B (NF- $\kappa$ B) pathway component RelA from the cytoplasm to the nucleus upon tumor necrosis factor  $\alpha$  (TNF $\alpha$ )-induced pathway activation, as detected by an increase in correlation between RelA and the nuclear dye DRAQ5 (Fig. 2E and fig. S4G). These experiments illustrate the utility of ICS parameters for quantification, and ultimately sorting, of a broad spectrum of phenotypes.

To demonstrate the cell-sorting functionality of the ICS, we applied it to the mitotic cell cycle, a dynamic process associated with multiple complex phenotypic changes. Traditional flow cytometry can separate three cell cycle stages, G<sub>1</sub>, G<sub>2</sub>/mitosis, and S phase, but fails to distinguish cells in different mitotic stages. Although chemicals that block mitosis can be used to enrich certain stages (notably excluding anaphase and telophase) (16–18), these approaches can alter gene expression and post-translational modifications. We demonstrate that ICS can isolate the mitotic stages of HeLa cells by using H2B-mNeonGreen (mNG) to visualize chromatin and the intensity of phosphorylated serine 10 on histone H3 (pS10H3) as a marker associated with mitotic chromatin condensation (19). We investigated cells from the G<sub>2</sub>/mitosis phases of the univariate cell cycle and created a training dataset by manually classifying 100 cells from each stage throughout mitosis (Fig. 2F; for a description of the criteria used to distinguish mitotic stages, please see the materials and methods). Classified events were organized on a trajectory in chronological order (fig. S5A). We used this training dataset to identify the most differing image-, scatter- and intensity-based param-

eters between stages by fitting a decision tree model and performing feature importance analysis (20) (Fig. 2G and fig. S5B). Image-derived parameters dominated the most differentiating parameters, such as maximum intensity, radial moment, and eccentricity of the H2B-mNG signal that differentiated among metaphase, anaphase, and telophase cells (Fig. 2H). We used these features to establish a hierarchical gating strategy for cell sorting and performed independent microscopic validation of the isolated populations (Fig. 2I). We found that ICS isolated highly pure populations, including G<sub>2</sub> interphase (96% purity), prometaphase (64%), metaphase (78%), anaphase (94%), and telophase (93%) (Fig. 2, J and K, and fig. S5, C to E). With these advances, we increased the resolution of flow cytometric cell cycle analyses to the level of distinguishing individual mitotic stages (including the thus-far inaccessible anaphase and telophase stages), yielding a method for robust enrichment of high numbers of cells in the absence of chemical blockers and from the same source sample. Isolated cells can be used in numerous downstream applications, such as the comparison of stage-specific changes in transcriptome, chromatin architecture, or protein modifications.

Pooled functional genomic screens with microscopic readouts have so far been limited in throughput and depended on technically challenging methods (21–25). ICS allows high-speed cell isolation based on fluorescence spatial information, and therefore has the potential to increase the scale and speed of microscopy-based screens and reduce technical complexity, duration, and cost. We tested the compatibility of ICS with pooled CRISPR screens by examining the nuclear translocation of RelA upon NF- $\kappa$ B pathway activation, a process that is invisible to traditional flow cytometry. To measure RelA translocation upon CRISPR-mediated perturbation, we quantified RelA-mNG/DRAQ5 spatial correlation (Fig. 2E) in HeLa cells expressing Tet-inducible Cas9 and fluorescently tagged RelA (23) (HeLa RelA) (fig. S6, A to C).

We validated the approach using individual CRISPR knockouts of three core NF- $\kappa$ B pathway components, IKBKA, IKBKG, and MAP3K7, and found consistent defects in RelA nuclear translocation upon gene knockout, demonstrating that ICS sensitively captures the effects of these perturbations (Fig. 3A and fig. S6D). Next, we proceeded with a pooled screen in which a population of Cas9-expressing cells is transduced with a mixture of guide RNAs (gRNAs). We transduced HeLa RelA cells with an NF- $\kappa$ B pathway-focused library targeting 1068 genes, including 37 NF- $\kappa$ B core canonical pathway components. Cells were then treated with TNF $\alpha$ , and the 5% lower (cytoplasmic RelA) and upper (nuclear RelA) bins of the RelA-mNG/DRAQ5 correlation parameter were isolated (Fig. 3B and fig. S7A). Sorting was conducted with an average event rate of 4000 events/s, a speed comparable to current flow-based technology for large cells such as HeLa, enabling a 100× coverage of a 1000 gRNA library in <9 min. Bulk sorts were performed at different library coverage to determine optimal library coverage and gRNA number per gene. We generated a “ground-truth” high-coverage (359-fold) dataset by pooling all reads from the differently sized samples, followed by gRNA hit calling (26). Among the most significant hits, we identified known NF- $\kappa$ B pathway components, demonstrating that ICS can identify bona fide regulators of the NF- $\kappa$ B pathway (Fig. 3C and table S1). We found strong correlation between the individual and pooled perturbations, indicating that both perturbation strategies rank genes similarly (Fig. 3D and fig. S7B). Next, we investigated how the number of gRNAs per gene and library coverage affect hit-calling performance. High performance [area under the precision recall curve (AUPRC) >0.7; 70% of hits detected at <1% false discovery rate (FDR)] was achieved with only 100 cells per gRNA and three gRNAs per gene (Fig. 3E and fig. S7, C and D). Performance increased with library coverage and number of gRNAs per gene, because sporadic false



hits caused by gRNA dropouts in the low-coverage samples decreased (fig. S7, E and F). Independent screen replicates showed high reproducibility ( $0.77 \leq R \leq 0.87$ ; fig. S7G).

To fully exploit the high-speed capabilities of ICS, we next sought to identify NF- $\kappa$ B pathway regulators globally in a genome-wide screen. We generated a new genome-wide CRISPR/Cas9 library targeting 18,408 protein-coding genes with fully adjustable numbers of gRNAs per gene (fig. S8 and supplementary text). Using six gRNAs per gene and a  $100\times$  library coverage, we identified 169 hits (FDR  $<1\%$ ), encompassing 133 positive and 36 negative regulators (Fig. 3F, fig. S9A, table S2, and supplementary text). A down-sampling-based analysis confirmed that three gRNAs per gene ranked genes similarly to the full library of six gRNAs (fig. S9, B and C). Among these hits, we identified all core canonical NF- $\kappa$ B pathway components except for three pathway genes, TRAF5, TAB1, and NFKB1, consistent with previous reports of these genes not being essential for pathway functionality (23, 27–29) (Fig. 3G). To identify potential new regulators, we performed a Gene Ontology (GO) term-based network analysis, which showed marked enrichment of a cluster of processes centered around chromatin modification (Fig. 3H). Among the underlying genes, we identified the histone deacetylase HDAC3, which induces RelA nuclear export during pathway shutdown (30). We also found previously unknown regulators including multiple components of the SAGA chromatin-acetylation complex (31) and the INO80 chromatin-remodeling complex (32), indicating a previously unknown role of these complexes in NF- $\kappa$ B pathway regulation (Fig. 3I and supplementary text). For hit validation, we assessed the top 10 previously unknown positive and negative candidates, the 10 identified members of the SAGA and INO80 complexes, and three known NF- $\kappa$ B pathway components. Individual CRISPR knockouts followed by quantification of RelA nuclear translocation using both ICS and microscopy revealed strong agreement ( $0.857 \leq R \leq 0.908$ ) between these measurements and confirmed the observations from the pooled genetic screen (Fig. 3J and fig. S9, D and E). In addition, our validation experiments indicate that ICS can reach similar accuracy and ranks genes similarly to fluorescence microscopy (Fig. 3J and fig. S9D). With the applied event rate of 4000 events/s, ICS is significantly faster compared with recently developed microscopy-based methods for pooled genetic screens (22–25) (for comparison, see the supplementary text) and enabled the completion of a genome-wide screen (three gRNAs per gene,  $100\times$  coverage) within only 9 hours of run time.

In conclusion, ICS substantially expands the phenotypic space accessible to cell-sorting ap-

plications and functional genomic screening. This method meets the requirements of high-speed cell sorting, multicolor fluorescence imaging, and full integration into a device that can be operated in nonspecialized laboratories. This will ensure broad availability and inspire new experimental strategies in diverse areas, including basic research, cell-based diagnostics, cell atlas efforts (3), and high-content image-based screening (2, 33, 34). With the potential to include downstream (multi)omics readouts (35–41), ICS provides a fundamentally new capability for probing deep into the molecular mechanisms underlying cell physiology and protein localization.

## REFERENCES AND NOTES

1. A. Cossarizza et al., *Eur. J. Immunol.* **49**, 1457–1973 (2019).
2. M. Boutros, F. Heigwer, C. Laufer, *Cell* **163**, 1314–1325 (2015).
3. P. J. Thul et al., *Science* **356**, eaal3321 (2017).
4. V. Espina et al., *Nat. Protoc.* **1**, 586–603 (2006).
5. A. S. Rane, J. Rutkauskaitė, A. DeMello, S. Stavakis, *Chem* **3**, 588–602 (2017).
6. K. Goda et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11630–11635 (2012).
7. H. E. Muñoz et al., *Anal. Chem.* **90**, 11280–11289 (2018).
8. T. Miura et al., *Biomed. Opt. Express* **9**, 3424–3433 (2018).
9. N. Nitta et al., *Cell* **175**, 266–276.e13 (2018).
10. S. Ota et al., *Science* **360**, 1246–1251 (2018).
11. G. Holznier et al., *Cell Rep.* **34**, 108824 (2021).
12. T. C. George et al., *Cytometry A* **59**, 237–245 (2004).
13. H. Mikami et al., *Nat. Commun.* **11**, 1162 (2020).
14. A. Isozaki et al., *Lab Chip* **20**, 2263–2273 (2020).
15. E. D. Diebold, B. W. Buckley, D. R. Gossett, B. Jalali, *Nat. Photonics* **7**, 806–810 (2013).
16. P. Clute, J. Pines, *Nat. Cell Biol.* **1**, 82–87 (1999).
17. G. W. Zieve, D. Turnbull, J. M. Mullins, J. R. McIntosh, *Exp. Cell Res.* **126**, 397–405 (1980).
18. J. Jackman, P. M. O'Connor, *Curr. Protoc. Cell Biol.* **Chapter 8**, Unit 8.3 (2001).
19. C. Prigent, S. Dimitrov, *J. Cell Sci.* **116**, 3677–3685 (2003).
20. P. Wei, Z. Lu, J. Song, *Reliab. Eng. Syst. Saf.* **142**, 399–432 (2015).
21. E. C. Wheeler et al., *Nat. Methods* **17**, 636–642 (2020).
22. C. Wang, T. Lu, G. Emanuel, H. P. Babcock, X. Zhuang, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10842–10851 (2019).
23. D. Feldman et al., *Cell* **179**, 787–799.e17 (2019).
24. G. Kanfer et al., *J. Cell Biol.* **220**, e202006180 (2021).
25. X. Yan et al., *J. Cell Biol.* **220**, e202008158 (2021).
26. C. G. de Boer, J. P. Ray, N. Hacohen, A. Regev, *Genome Biol.* **21**, 134 (2020).
27. K. Tada et al., *J. Biol. Chem.* **276**, 36530–36534 (2001).
28. C. Wang et al., *Nature* **412**, 346–351 (2001).
29. L.-F. Chen, W. C. Greene, *Nat. Rev. Mol. Cell Biol.* **5**, 392–401 (2004).
30. L.-F. Chen, W. Fischle, E. Verdin, W. C. Greene, *Science* **293**, 1653–1657 (2001).
31. D. Helmlinger, L. Tora, *Trends Biochem. Sci.* **42**, 850–861 (2017).
32. L. Chen et al., *J. Biol. Chem.* **286**, 11283–11289 (2011).
33. Y. Feng, T. J. Mitchison, A. Bender, D. W. Young, J. A. Tallarico, *Nat. Rev. Drug Discov.* **8**, 567–578 (2009).
34. Z. E. Perlman et al., *Science* **306**, 1194–1198 (2004).
35. D. Schraivogel et al., *Nat. Methods* **17**, 629–635 (2020).
36. D. A. Jaitin et al., *Cell* **167**, 1883–1896.e15 (2016).
37. P. Datlinger et al., *Nat. Methods* **14**, 297–301 (2017).
38. A. Dixit et al., *Cell* **167**, 1853–1866.e17 (2016).
39. G. X. Y. Zheng et al., *Nat. Commun.* **8**, 14049 (2017).
40. P. Shahi, S. C. Kim, J. R. Haliburton, Z. J. Gartner, A. R. Abate, *Sci. Rep.* **7**, 44447 (2017).
41. M. Stoeckius et al., *Nat. Methods* **14**, 865–868 (2017).
42. B. Rauscher, C. Tischer, K. Owsley, D. Schraivogel, L. Steinmetz, *Zenodo* (2021).
43. U. Sarkans et al., *Nucleic Acids Res.* **46**, D1266–D1270 (2018).
44. J. Spidlen, K. Breuer, C. Rosenberg, N. Kotecha, R. R. Brinkman, *Cytometry A* **81**, 727–731 (2012).
45. D. Szklarczyk et al., *Nucleic Acids Res.* **47**, D607–D613 (2019).

## ACKNOWLEDGMENTS

We thank L. Velten, J.-K. Heriche, R. Kumar, A. Kreshuk, and T. Alexandrov for input on computational analyses; R. Pepperkok, S. Reither, A. Hauth, F. Steudle, D. Gerlich, J. Zuber, M. Knop, Y. Hayashi, E. Schiebel, J. Ellenberg, and J. Kornienko for providing cell lines, antibodies, and constructs; M. Rogon for network analysis support; M. Krause, D. Schichler, A. Hughes, and P. Jakob for experimental support; BD Biosciences, a unit of Becton, Dickinson and Company; the BD CellView team that developed the BD CellView Imaging Technology that enabled ICS (contributing team members are listed in the supplementary text); J. Horta and D. Fantin for managerial support; J. Kim and D. Martin for instrument support and maintenance; the EMBL Advanced Light Microscopy Facility (ALMF) for support; the EMBL Genecore for next-generation sequencing services; V. Benes for advice; the EMBL Flow Cytometry Core facility for flow cytometry support, advice, and instrument maintenance; and L. Velten, K. Zeier, J. Horta, and M. Bao (Life Science Editors) for input on the manuscript. **Funding:** This work was supported by grants from the European Research Council (Advanced Investigator Grants AdG-294542 and AdG-742804 to L.M.S.), the German Research Foundation (DFG project number 402723784 to S.C.-H.), and the Human Frontier Science Program (GDA00045/2019 to S.C.-H.). D.S. was supported by a fellowship from the EMBL Interdisciplinary Postdoc (EIPDOP) program (Marie Skłodowska-Curie Actions COFUND grant agreement 664726). T.M.K. was supported by a postdoctoral fellowship from the European Molecular Biology Organization (EMBO ALTF 1154-2020). C.T. was supported by the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation (grant 2020-225265). M.P. was supported by the Novo Nordisk Foundation (grants NNF17CC0027852 and NNF21CC0073729). **Author contributions:** D.S., M.P., A.M., S.C.-H., and L.M.S. conceptualized the project. D.S., T.M.K., M.R.-M., M.P., and M.D. performed experiments. T.M.K. and S.C.-H. collected and analyzed microscopy data. B.Rau. designed CRISPR libraries. D.S. and M.R.-M. performed functional genomics screens. B.Ram. and D.O. supported flow cytometric experiments. B.Ram. performed purity sorts. D.O. performed instrument QC. B.Rau., K.O., and D.S. performed bioinformatic analysis. C.T. wrote Fiji plugins. K.O., A.M., and E.D. developed BD CellView Imaging Technology. D.S., B.Rau., M.R.-M., T.M.K., M.P., K.O., A.M., S.C.-H., E.D., and L.M.S. wrote the manuscript. All authors read and commented on the manuscript. **Competing interests:** K.O. and/or E.D. are inventors on patents 9423353, 9983132, 10078045, 10324019, 10006852, 10408758, 10823658, 10578469, 10288546, 10684211, 11002658, 10620111, 11105728, 10976236, 10935482, and 11055897 held or licensed for use by Becton, Dickinson and Co. that cover BD CellView Imaging Technology. K.O., A.M., and E.D. are employees at BD Biosciences. BD CellView, BD FACSMelody, BD FACSAria, BD FACSCorus (and any others used) are trademarks or registered trademarks of Becton, Dickinson and Company. **Data and materials availability:** NGS data from gRNA library and targeted genome sequencing were deposited at the Gene Expression Omnibus (GSE167944). Documented code to reproduce all analyses and figures was deposited at GitHub (<https://github.com/benediktrauscher/ICS>) and Zenodo (42). Fiji tools and source code thereof were deposited at GitHub (<https://github.com/embl-cba/ICS>) and Zenodo (42). ICS image data were deposited on the BiImage Archive (43) (S-BSST644, available from <https://www.ebi.ac.uk/biostudies/>). Flow cytometry data were deposited at <https://flowrepository.org> (44) (FR-FCM-Z4M5). Metadata and archiving information for ICS data are provided in table S3. Plasmid “phage UbiC tagRFP-T-DDX6” is available under a material transfer agreement from Addgene. Cell line “HeLa Tet:Cas9 RelA-mNeonGreen” is available under a material transfer agreement from Broad Institute (P. Blainey).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj3013  
Materials and Methods  
Supplementary Text  
Figs. S1 to S10  
Tables S1 to S7  
References (46–102)  
MDAR Reproducibility Checklist

27 May 2021; accepted 8 December 2021  
10.1126/science.abj3013

## 2D MATERIALS

# Visualizing broken symmetry and topological defects in a quantum Hall ferromagnet

Xiaomeng Liu<sup>1†</sup>, Gelareh Farahi<sup>1†</sup>, Cheng-Li Chiu<sup>1†</sup>, Zlatko Papic<sup>2</sup>, Kenji Watanabe<sup>3</sup>, Takashi Taniguchi<sup>4</sup>, Michael P. Zaletel<sup>5</sup>, Ali Yazdani<sup>1\*</sup>

The interaction between electrons in graphene under high magnetic fields drives the formation of a rich set of quantum Hall ferromagnetic (QHFM) phases with broken spin or valley symmetry. Visualizing atomic-scale electronic wave functions with scanning tunneling spectroscopy (STS), we resolved microscopic signatures of valley ordering in QHFM phases and spectral features of fractional quantum Hall phases of graphene. At charge neutrality, we observed a field-tuned continuous quantum phase transition from a valley-polarized state to an intervalley coherent state, with a Kekulé distortion of its electronic density. Mapping the valley texture extracted from STS measurements of the Kekulé phase, we could visualize valley skyrmion excitations localized near charged defects. Our techniques can be applied to examine valley-ordered phases and their topological excitations in a wide range of materials.

Quantum Hall ferromagnets are broken-symmetry states in which the exchange interaction between electrons in Landau levels gives rise to quantum Hall phases with polarized or coherent superposition of spin, valley, or orbital degrees of freedom (1). In the presence of a magnetic field, a variety of two-dimensional electronic systems—including those in semiconductors (1, 2), graphene (2), and an increasing number of moiré flat-band materials—host a diversity of quantum Hall ferromagnetic (QHFM) phases (3–8). Thus far, these interacting and topological phases of matter have been examined macroscopically, usually through study of their transport properties. However, the microscopic features of the electronic wave functions of these phases can directly reveal the nature of their broken symmetry (9, 10) and, more important, can determine the nature of the excitations they host. A particularly interesting aspect of broken-symmetry states is their topological excitations, such as skyrmions (11–13), which determine the stability of such phases, and whose interactions may lead to the formation of more exotic quantum phases, such as the skyrmion superconductivity recently proposed in moiré materials (14–16).

Monolayer graphene's SU(4) isospin space, consisting of spin and valley, gives rise to a rich array of QHFM phases, which have been studied using transport and thermodynamic measurements (2). Particularly intriguing is

the electrically insulating phase at the charge neutrality point at high magnetic fields (17), because with two of four isospin flavors occupied, Pauli exclusion prevents spin and valley from being simultaneously polarized. Theoretical efforts have predicted a rich phase diagram of four possible broken-symmetry QHFM states at charge neutrality (18): a charge density wave (CDW) phase, which is sublattice- and valley-polarized and spin-unpolarized; the spin ferromagnet (FM), which is a quantum spin Hall insulator; the canted antiferromagnet (CAF), in which spins on different sublattices point in near-opposite directions; and an intervalley coherent (IVC) state with a Kekulé reconstruction, which is spin-unpolarized. A recent theory also proposed the coexistence of CAF and IVC (19). Although transport studies have constrained aspects of the phase diagram (20, 21), the nature of the ground state of graphene at charge neutrality has remained unresolved in the absence of microscopic measurements that probe the order parameter. Also unexplored are the plethora of topological excitations that these phases have been predicted to host, such as a variety of skyrmions, which may have complex flavor textures and may even harbor fractional charge on the scale of the magnetic length (22–25). Here, we used spectroscopic mapping to visualize the broken-symmetry states in graphene as a function of carrier concentration, including at charge neutrality, where we find evidence for localized valley skyrmions within the Kekulé phase. Our work demonstrates the power of spectroscopic imaging to detect valley ordering and topological excitations of valley orders; the method is applicable to a wide range of two-dimensional materials and their heterostructures.

The monolayer graphene devices used for our studies are fabricated on hexagonal boron nitride (hBN) substrates, with either graphite (devices A and C) or silicon back gates (device B) (see Fig. 1 for the experimental setup and an

optical image of device A). All samples show similar spectroscopic properties, except that the fractional quantum Hall features are visible only in the graphite gate samples (devices A and C) (26). Figure 1, B and C, shows measurements of differential conductance  $dI/dV$  as function of sample bias  $V_B$  measured over a wide range of filling factors  $\nu$  ( $\nu = 2\pi n l_B^2$ , where  $l_B = \sqrt{\hbar/eB}$  is the magnetic length,  $n$  is the carrier density,  $\hbar$  is the reduced Planck's constant,  $e$  is elementary charge, and  $B$  is the magnetic field); the filling factor is controlled by the back gate voltage  $V_g$ . The Landau levels (LLs) can be identified by their peaks in  $dI/dV$ ; the energy corresponds to  $E_N = \hbar\omega_c \sqrt{N}$ , where  $N$  is the LL orbital index and  $\hbar\omega_c \sim 110$  mV is the extracted cyclotron energy from fitting Fig. 1D. This cyclotron gap corresponds to that calculated with a renormalized Fermi velocity of  $1.26 \times 10^6$  m/s, similar to the values found in previous studies (27). As the filling factor increases, the Fermi energy is pinned within a LL as it is being filled and then jumps to the next LL at  $\nu = \pm 2, \pm 6, \pm 10$ . For the incompressible states formed at these fillings, we find that energy gaps across the Fermi energy are enlarged by a factor of  $\sim 2$  relative to the expected cyclotron gap (fig. S1). This effect, which does not depend on setpoint conditions, is likely caused by the graphene's bulk insulating behavior when the chemical potential lies within these gaps [see discussion in (26)].

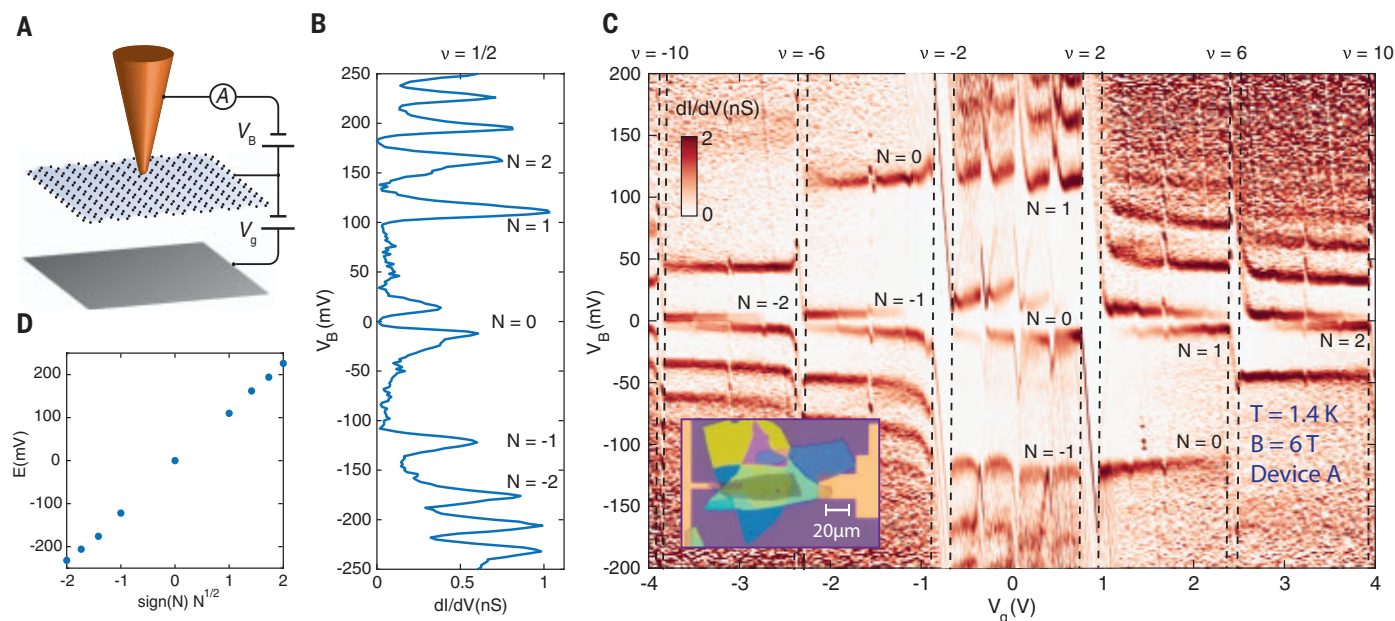
Symmetry-breaking states driven by electron-electron interaction are clearly demonstrated in our spectroscopic measurements by gaps at all the intermediate integer fillings (Fig. 1C). The sizes of the gaps in our experiment at symmetry-breaking states and single-particle quantum Hall states were larger than those observed in transport and thermodynamic studies. We find that tip-induced band bending is negligible in most of our measurements, which likely contributes to our ability to observe symmetry-breaking gaps. Although we occasionally find tips that show a signature of band bending in spectroscopic measurements (26) similar to previous studies (27–38), datasets we obtained with improved tip conditions demonstrate the following differences: (i) Our data (Fig. 1C) does not show any Coulomb diamond features associated with a tip-induced quantum dot, as seen in previous studies (34). (ii) Our sample is not doped by impurities and our measurements are not influenced by a tip-sample work function mismatch, as shown by the observation that charge neutrality occurs near zero gate voltage. (iii)  $V_B$  does not influence carrier density in the probed area; the dashed lines in Fig. 1C marking incompressible states are nearly vertical, therefore showing that tip gating is negligible. (iv) At partial fillings, the LLs are always pinned to the Fermi energy with their jumps aligned with the occurrence of the

<sup>1</sup>Joseph Henry Laboratories and Department of Physics, Princeton University, Princeton, NJ 08544, USA. <sup>2</sup>School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK. <sup>3</sup>Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan. <sup>4</sup>International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan. <sup>5</sup>Department of Physics, University of California, Berkeley, CA 94720, USA.

\*Corresponding author. Email: yazdani@princeton.edu

†These authors contributed equally to this work.





**Fig. 1. Experimental setup and large gate range spectra.** (A) Schematic of the STM measurement setup. The orange cone represents the tip, the light blue plane denotes the graphene, and the gray plane denotes the bottom gate. The bottom gate voltage  $V_g$  tunes the carrier density of graphene;  $V_B$  changes the bias voltage between the tip and graphene. (B) Spectrum of device A at  $\nu = 1/2$

showing LL peaks of different orbital numbers  $N$ . (C) Tunneling spectra of device A as a function of bias voltage and gate voltage measured at  $B = 6$  T,  $T = 1.4$  K at a fixed tip height. Inset: Optical image of device A. The left gold pad contacts the graphite gate; the right contact connects with graphene. (D) The energy of LLs extracted from the data in (B), displaying good agreement with  $E_N = \hbar\omega_c\sqrt{N}$ .

incompressible states, which suggests that there is no density mismatch between the probed area and the bulk of the sample. It is possible that our tip effective radius is small relative to the magnetic length, so that the work function mismatch between the tip and sample (which would typically lead to band bending) traps at most one electron charge below the tip, rather than producing a well-defined change in filling factor in a larger region.

Beyond resolving the presence of broken-symmetry states, our experiments also show a direct signature of fractional quantum Hall (FQH) phases in spectroscopic measurements. Focusing on the scanning tunneling spectroscopy (STS) properties between  $\nu = -2$  and  $2$ , as shown in Fig. 2A, we resolve enlarged gaps at partial filling of the zeroth LL (ZLL) corresponding to the fractional quantum Hall states at  $\nu = \pm 2/3, \pm 1/3$ . We corroborate the formation of FQH states in our devices by performing transport measurement while the tip height is reduced from the tunneling condition to directly contact the monolayer graphene (Fig. 2B). In this Corbino geometry, measurements of the conductance of our sample show dips at fractional fillings associated with the formation of FQH states. The observation of rich fractional states including  $\nu = 4/9$  in our samples, at a modest magnetic field (6 T) and at relatively elevated temperature (1.4 K), attests to their high quality, making them comparable to the fully hBN-encapsulated and dual graphite-gated devices used for the highest-

quality transport measurements. Probing FQH phases in scanning tunneling microscope (STM) measurements paves the way to explore these topological phases and their exotic excitations, including realization of methods for imaging anyons (39) or probing fractional edge states locally.

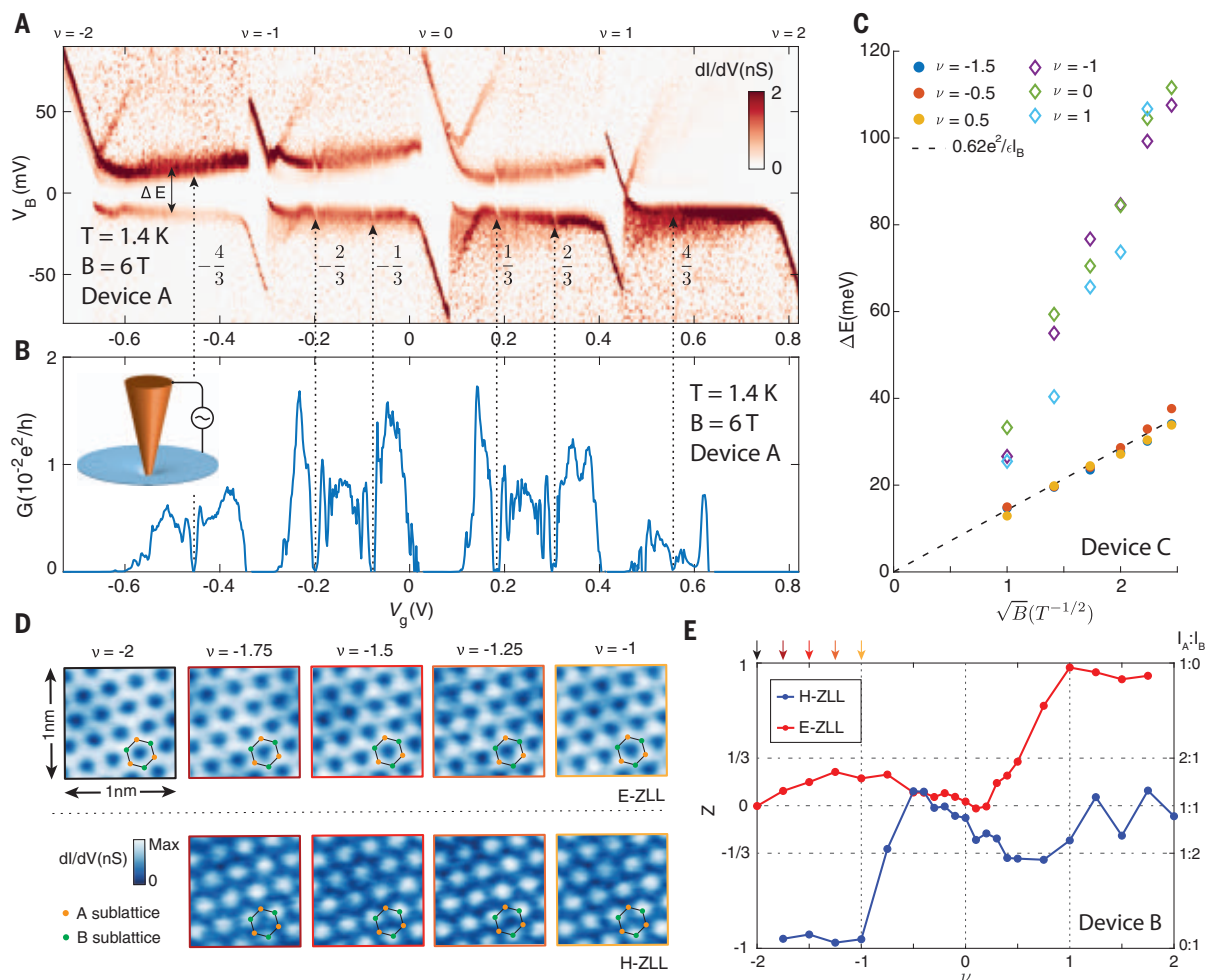
The spectroscopic measurements of the partially filled ZLL (Fig. 2A), including when the sample transitions through the FQH phases, always show splitting of the ZLL with a gap across the Fermi energy. This behavior is indicative of a Coulomb gap commonly observed when tunneling in and out of a two-dimensional electron gas at high magnetic fields (40–42). The strong correlations among electrons in the flat LLs dictate that additional energy is required for addition or removal of electrons from the system, resulting in a gap at the Fermi level that scales with the Coulomb energy  $E_c = e^2/\epsilon l_B$ , where  $\epsilon$  is the effective dielectric constant. The field dependence of this gap at partial filling follows the expected  $\sqrt{B}$  behavior (Fig. 2C), tracing Coulomb energy  $E_c$  with a 0.62 scale factor, which agrees with the value obtained from our exact diagonalization calculations (26).

To directly visualize the broken valley symmetry of graphene's ZLL, we perform spectroscopic mapping of the electron and hole excitations of the ZLL (E-ZLL and H-ZLL, respectively) with  $V_B$  at the split ZLL peaks below or above the Coulomb gap. These spectroscopic  $dI/dV$  maps are performed with the STM tip at a constant height above the graphene,

and hence they are directly proportional to the electron/hole excitation probability densities on the graphene atomic lattice. At filling  $\nu = -2$ , the  $dI/dV$  map of electron excitations shows only graphene's honeycomb lattice, whereas at partial fillings between  $\nu = -2$  and  $-1$ , the  $dI/dV$  maps of hole excitations show sublattice polarization. A key feature of graphene's ZLL is that the electron states at the  $K$  or  $K'$  valleys correspond to the A or B sublattice sites, respectively (2, 43). Therefore, the sublattice polarization observed in these maps—for example, for hole excitation at  $\nu = -1$ —is indicative of valley polarization in the ZLL, which agrees with the expectation of a spin- and valley-polarized ground state  $|K'\uparrow\rangle$  at quarter-filling (44). The electron excitation at this filling shows partial polarization of the orthogonal state comprising  $|K'\downarrow\rangle$ ,  $|K\uparrow\rangle$ , and  $|K\downarrow\rangle$ . Our measurements at fillings  $-2 < \nu < -1$  indicate that the ground state in this range also remains valley-polarized, thereby demonstrating that FQH states in this filling range are single-component and that valley symmetry breaking precedes the formation of FQH states (45).

Although valley polarization in the filling range  $-2 < \nu \leq -1$  is dictated by interactions, we demonstrate that the sublattice asymmetry energy plays an important role in choosing which valley is occupied. In Fig. 2E, we extract the sublattice polarization  $Z = (I_A - I_B)/(I_A + I_B)$ , where  $I_A$  and  $I_B$  are the intensities of  $dI/dV$  signals at the A and B sublattices (26), and plot them for the ZLL as a function of filling.





**Fig. 2. Symmetry breaking and fractional quantum Hall states of the zeroth Landau level.** (A) Tunneling spectrum of the zeroth Landau level between  $\nu = -2$  and  $2$  measured in device A. (B) Corbino transport measurement done on device A when contacting graphene with the tip by reducing tip height by 2 nm ( $B = 6$  T,  $T = 1.4$  K). Fractional states are detected from  $\nu = 1/3$  up to  $\nu = 4/9$ . The gate voltages at which fractional features appear coincide with the tunneling measurement in (A). (C) The separations of the split ZLL peak as a function of

$\sqrt{B}$ , measured on device C. The splittings at half fillings scale with Coulomb energy (black dashed line). (D)  $dI/dV$  maps taken on the electron excitation of the ZLL (E-ZLL) and the hole excitation of the ZLL (H-ZLL) peaks at quarter-fillings between  $\nu = -2$  and  $-1$  in device B. The hexagon pattern is the underlying graphene atomic lattice. The H-ZLL peak is fully sublattice-polarized in this filling range. (E) Sublattice polarization  $Z$  as a function of filling factors for H-ZLL and E-ZLL peaks extracted by Fourier transformation of  $dI/dV$  maps. Arrow colors correspond to those in (D).

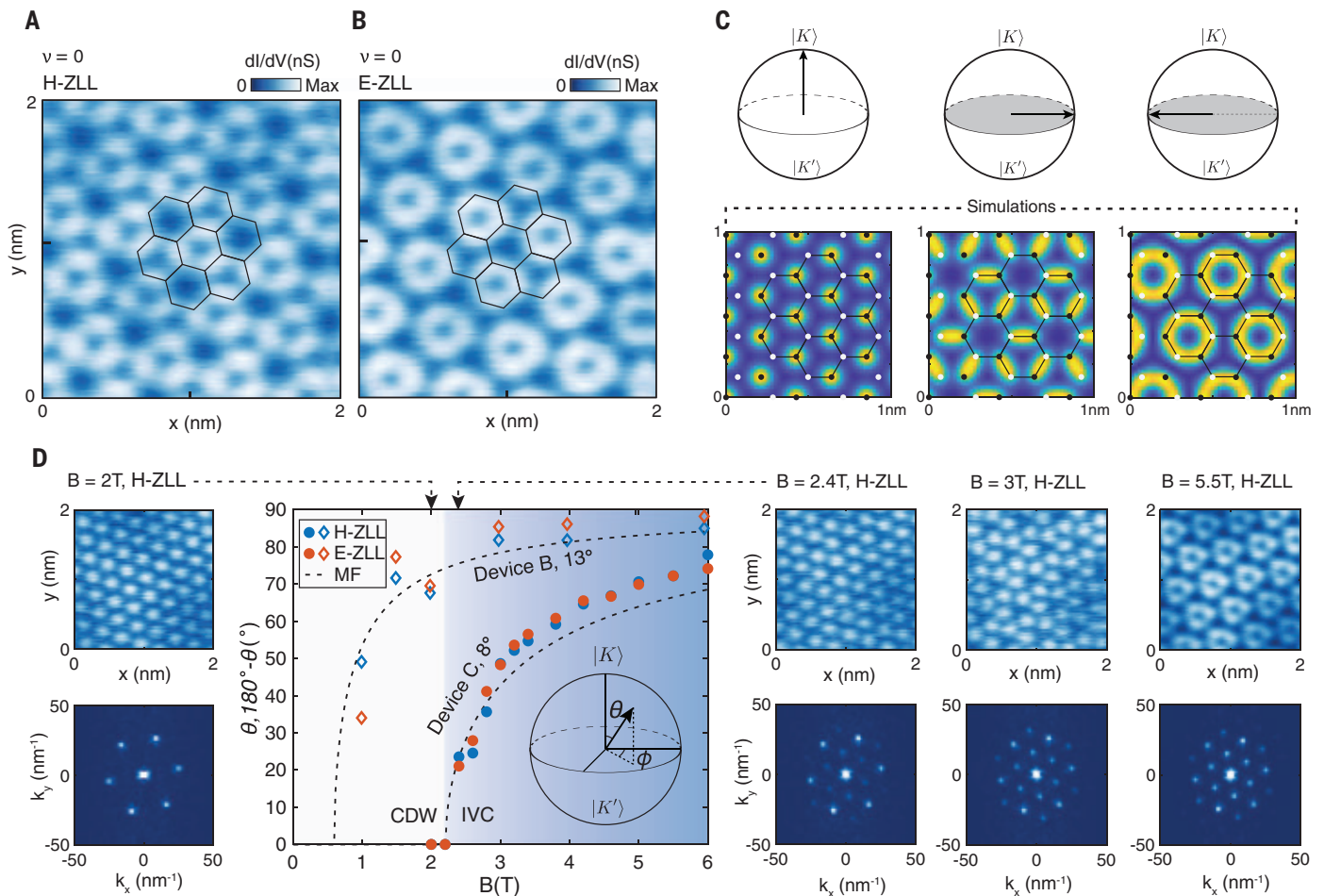
Complementary to fillings in the range  $-2 < \nu \leq -1$ , where we find full polarization of the hole excitation, we find that for the range  $1 \leq \nu < 2$ , the electron excitation maps probing the unoccupied states are fully polarized in the A sublattice. We find the occupied states, probed by the hole excitations, to be always polarized in the B sublattice regardless of the filling factor, as is evident from the blue line in Fig. 2E, which is almost entirely below zero. This behavior indicates that although interactions drive the symmetry breaking, the B sublattice is favored by an apparent AB sublattice asymmetry, likely originating from partial alignment with the hBN substrate.

We turn our attention to spectroscopic imaging at charge neutrality to show that electron interactions induce an intervalley coherent electronic state in half-filled ZLL at high fields.

Spectroscopic maps of  $\nu = 0$  at 6 T (Fig. 3, A and B, device B) show a spatially varying electronic density with a periodicity that is  $\sqrt{3}$  larger than that of the graphene lattice. This has been reported previously for graphene multilayers claimed to be decoupled, albeit without gate control (46). Such reconstruction of the unit cell, also referred to as the Kekulé distortion, is expected when an IVC phase forms. This state, which is one of the four anticipated phases at charge neutrality, has a real-space electronic wave function with probability density at both sublattices. To understand the real-space patterns for electron and hole excitations of this phase, we describe its valley order using a vector on a Bloch sphere:  $|\psi\rangle = \cos(\theta/2)|K\rangle + \sin(\theta/2)\exp(i\phi)|K'\rangle$ , with polar angle  $\theta$  and azimuthal angle  $\phi$ . For states with ordering vector pointing to the poles ( $\theta = 0, 180^\circ$ ), electron

densities correspond to full valley and sublattice polarization, forming a CDW state. In contrast, when the ordering vector lies along the equator of the Bloch sphere ( $\theta = 90^\circ$ ), we have equal weight on both sublattices, with the azimuthal angle  $\phi$  characterizing the phase coherence of the wave functions between the two sublattices. Computing the probability density  $\langle\psi|\psi\rangle$ , we find that the IVC state as described by  $\phi = 0^\circ$  and  $180^\circ$  (Fig. 3C) reproduces the Kekulé patterns seen experimentally for hole and electron excitation in Fig. 3, A and B, respectively. Naturally, the hole excitation has a real-space structure and valley polarization orthogonal to those of the electron excitation of the same state.

A more detailed analysis of the ordering vector as a function of the magnetic field reveals a continuous quantum phase transition



**Fig. 3. Intervalley coherent state at the charge neutrality point.** (A and B)  $dI/dV$  maps at the charge neutrality point, measured at  $B = 6$  T in device B. The hexagons represent the graphene lattice. The  $dI/dV$  maps show a Kekulé reconstruction that triples the area of the unit cell. (C) Bloch sphere plot and corresponding simulated probability density of valley polarization for CDW (left) and IVC with  $\phi$  of  $0^\circ$  (center) and  $180^\circ$  (right). (D) Main panel: Polar angle  $\theta$  as a function of the magnetic field in devices B and C extracted from  $dI/dV$  maps. Plots are shown for  $\theta$  (E-ZLL peaks) and  $180^\circ - \theta$  (H-ZLL peaks). The complementary behavior of H-ZLL and E-ZLL peaks confirms their orthogonal

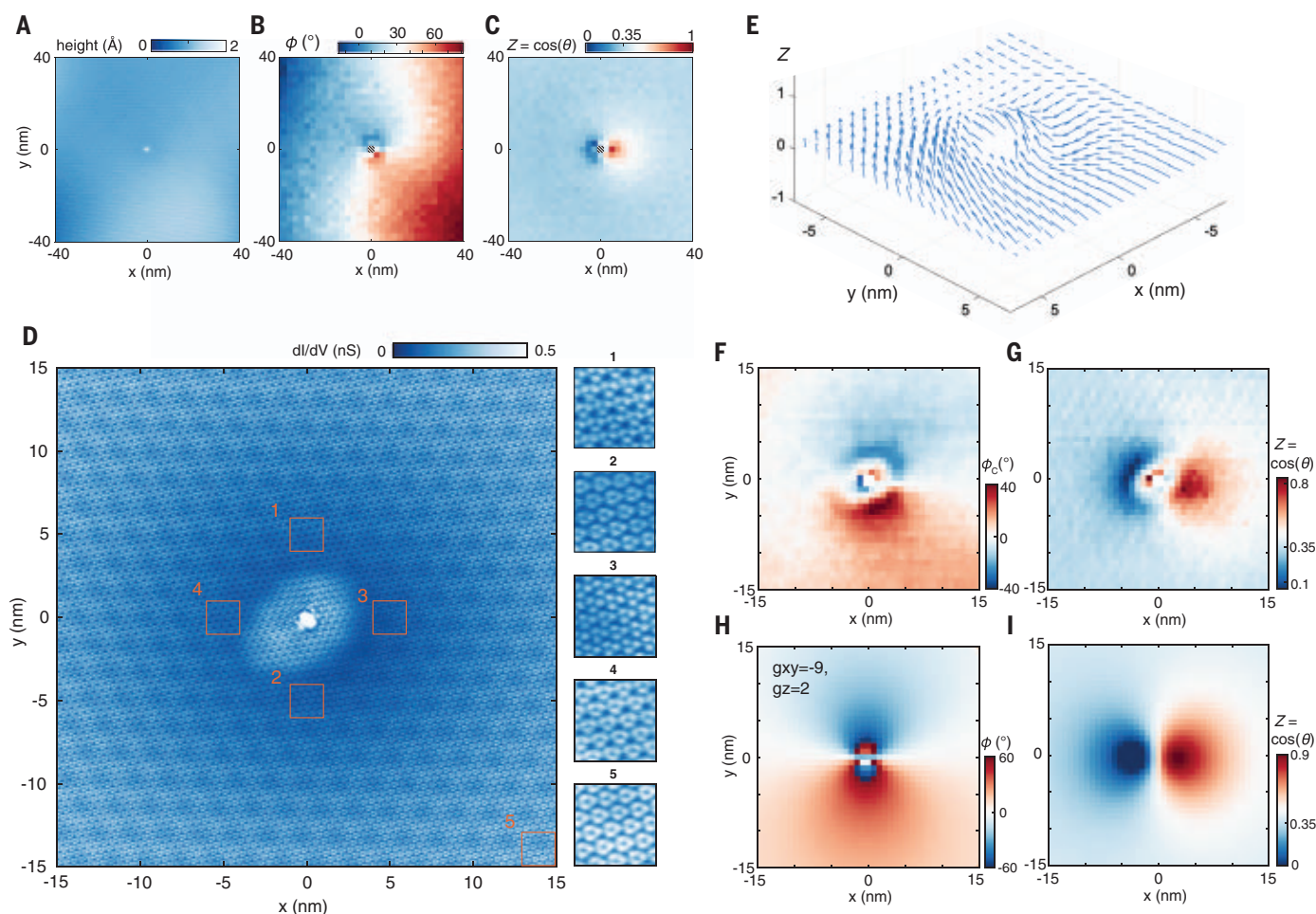
nature. Device B has a  $13^\circ$  misalignment angle between the graphene and the hBN substrate; in device C this angle is  $8^\circ$ . The color shading of the background indicates the transition from CDW to IVC in device C. The mean field (MF) behavior for  $\theta$  is shown as dashed lines, with critical fields of 2.2 T (device C) and 0.6 T (device B). Top side panels:  $dI/dV$  maps of the H-ZLL at a few representative magnetic fields in device C. Bottom side panels: Fourier transform of the  $dI/dV$  maps in the corresponding top panels. At  $B = 2$  T, only Fourier peaks of the graphene lattice are visible, whereas at  $B = 2.4$  T, Fourier peaks of the Kekulé pattern appear and increase in intensity with increasing magnetic field.

between the IVC Kekulé phase and the valley- and sublattice-polarized CDW state. We study this transition by extracting the ordering vector's polar angle  $\theta$  from the Fourier transforms of real-space  $dI/dV$  maps and examine it as a function of the magnetic field. With increasing field,  $\theta$  shows a continuous transition from the CDW phase ( $\theta = 0$ ) to an IVC state with  $\theta$  approaching  $90^\circ$  in both devices (Fig. 3D). A critical field (2.2 T for device C) can be identified where  $\theta$  becomes nonzero while intervalley coherence emerges, as detected by the appearance of Kekulé wave vectors in the FFT of  $dI/dV$  maps. We find that both the critical field and  $\theta$  at 6 T measured in the two devices correlate with the influence of sublattice asymmetry imposed by the hBN substrate. The less aligned sample (device B,  $13^\circ$  misalignment between

graphene and hBN lattice), with smaller sublattice asymmetry, shows a smaller critical field and approaches a pure IVC state with  $\theta = 90^\circ$  at a lower field than sample C. This behavior is consistent with the competition between the AB sublattice asymmetry, which favors one sublattice over the other, and valley anisotropy induced by short-range electron-electron and electron-phonon interactions (18), which favors valley polarization of  $\theta = 90^\circ$ . The magnetic field controls the strength of the interactions and in turn the valley anisotropy energy, thereby tuning  $\theta$  like the order parameter of a continuous phase transition, a behavior well captured by a mean-field description (26) (Fig. 3D, dashed lines).

Finally, we show that measurements of the spatial variation of the ordering vectors in the IVC phase can be used to directly visualize

the presence of topological excitations in this state. The spatial variations are extracted by performing local Fourier analysis on the  $dI/dV$  maps, where large areas of the sample show spatially independent  $\theta$  and a constant gradient for  $\phi$ . Uniform gradients in  $\phi$  are expected in the presence of either strain or dilute short-range disorder (25). However, near charged defects on the graphene surface, likely caused by atomic adsorbates (Fig. 4A), we see markedly different behavior. Near this defect, we find that  $\phi$  displays a swirl-like spatial variation (Fig. 4B), and the variation of  $\theta$  plotted as sublattice polarization  $Z = \cos(\theta)$  (Fig. 4C) displays a dipole-like feature. Analysis of higher-resolution electron excitation maps near this defect (Fig. 4D) shows the variations close to the defect of  $\phi$  (with the linear gradient



**Fig. 4. Valley skyrmion of the IVC state near a charged defect.** (A) Topography of the point defect found on device C. (B and C) Azimuthal angle  $\phi$  and  $Z$  polarization extracted from  $dI/dV$  maps of the E-ZLL peak (26). (D)  $dI/dV$  map of the E-ZLL zoomed in the area near the defect shown in (A). Side panels: Magnified images of a few representative areas with matching labels. (E) Valley texture extracted from (D), visualized by arrays of arrows representing valley

polarization in a Bloch sphere at each point. (F and G) Azimuthal  $\phi$  and  $Z$  polarization extracted from (D). A linear background is subtracted from  $\phi$  to produce  $\phi_c$  (26). (H and I) Azimuthal angle and  $Z$  polarization extracted from a simulated map of electron density computed for a canted antiferromagnetic (CAF) skyrmion using the same Fourier procedure [see (26) for details of the calculations].

background subtracted) and  $Z$  more clearly (Fig. 4, F and G). A visual representation of the valley ordering vector texture near this defect is shown in Fig. 4E. This valley texture is consistent with that predicted for a CAF skyrmion excitation of the Kekulé phase (23). This topological excitation forms when the valley polarization of one spin species flips by  $180^\circ$  at its center, whereas the other spin species is devoid of any valley texture. The two key signatures of this skyrmion excitation are the dipole behavior in  $Z$ , which is equivalent to a meron-antimeron pair (Fig. 4I), accompanied by a dipole in  $\phi$  oriented perpendicularly to the  $Z$  dipole (Fig. 4H). Simulating the valley texture using the nonlinear sigma model (NLSM) (26), we find excellent agreement between the results from the model calculations (Fig. 4, H and I) and our experimental results (Fig. 4, F and G). With our choice of model parameters, the calculation captures not only the qualita-

tive behavior of  $\phi$  and  $Z$  but also the size of the skyrmion, which is  $\sim 10$  nm in both theory and experiment. This CAF skyrmion carries an electric charge of  $\pm e$ , which is likely what caused their localization near a charged defect of the opposite sign. Our experiments show that besides the CAF skyrmion, other types of valley textures are also possible (fig. S5). Further work can map the zoo of predicted topological excitations in this and other QHFM phases of graphene (22, 23). From a broader perspective, the microscopic approach to studying valley ordering can be applied to other two-dimensional systems, such as twisted bilayer graphene.

After submission, we have become aware of a related STM study of the  $v = 0$  state (47).

#### REFERENCES AND NOTES

1. Z. F. Ezawa, *Quantum Hall Effects: Field Theoretical Approach and Related Topics* (World Scientific, 2008).
2. B. I. Halperin, J. K. Jain, *Fractional Quantum Hall Effects: New Developments* (World Scientific, 2020).

3. K. P. Nuckolls *et al.*, *Nature* **588**, 610–615 (2020).
4. I. Das *et al.*, *Nat. Phys.* **17**, 710–714 (2021).
5. Y. Saito *et al.*, *Nat. Phys.* **17**, 478–481 (2021).
6. J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Nature* **592**, 43–48 (2021).
7. Y. Choi *et al.*, *Nature* **589**, 536–541 (2021).
8. S. Wu, Z. Zhang, K. Watanabe, T. Taniguchi, E. Y. Andrei, *Nat. Mater.* **20**, 488–494 (2021).
9. B. E. Feldman *et al.*, *Science* **354**, 316–321 (2016).
10. M. T. Randeria *et al.*, *Nat. Phys.* **14**, 796–800 (2018).
11. S. L. Sondhi, A. Karlhede, S. A. Kivelson, E. H. Rezayi, *Phys. Rev. B* **47**, 16419–16426 (1993).
12. N. Nagaosa, Y. Tokura, *Nat. Nanotechnol.* **8**, 899–911 (2013).
13. H. Zhou, H. Polshyn, T. Taniguchi, K. Watanabe, A. F. Young, *Nat. Phys.* **16**, 154–158 (2020).
14. S. Chatterjee, N. Bultinck, M. P. Zaletel, *Phys. Rev. B* **101**, 165141 (2020).
15. S. Chatterjee, M. Ippoliti, M. P. Zaletel, arXiv 2010.01144 [cond-mat.str-el] (2020).
16. E. Khalaf, S. Chatterjee, N. Bultinck, M. P. Zaletel, A. Vishwanath, *Sci. Adv.* **7**, eabf5299 (2021).
17. J. G. Checkelsky, L. Li, N. P. Ong, *Phys. Rev. Lett.* **100**, 206801 (2008).
18. M. Kharitonov, *Phys. Rev. B* **85**, 155439 (2012).
19. A. Das, R. K. Kaul, G. Murthy, arXiv 2109.07515 [cond-mat.mes-hall] (2021).



20. A. F. Young *et al.*, *Nature* **505**, 528–532 (2014).
21. H. Zhou *et al.*, arXiv 2102.01061 [cond-mat.mes-hall] (2021).
22. Y. Lian, M. O. Goerbig, *Phys. Rev. B* **95**, 245428 (2017).
23. J. Atteia, Y. Lian, M. O. Goerbig, *Phys. Rev. B* **103**, 035403 (2021).
24. K. Nomura, S. Ryu, D. H. Lee, *Phys. Rev. Lett.* **103**, 216801 (2009).
25. C. Y. Hou, C. Chamon, C. Mudry, *Phys. Rev. B* **81**, 075427 (2010).
26. See supplementary materials.
27. J. Chae *et al.*, *Phys. Rev. Lett.* **109**, 116802 (2012).
28. D. L. Miller *et al.*, *Science* **324**, 924–927 (2009).
29. D. Walkup *et al.*, *Phys. Rev. B* **101**, 035428 (2020).
30. S. Kim *et al.*, *Nat. Commun.* **12**, 2852 (2021).
31. G. Li, A. Luican, E. Y. Andrei, *Phys. Rev. Lett.* **102**, 176804 (2009).
32. Y. J. Song *et al.*, *Nature* **467**, 185–189 (2010).
33. D. L. Miller *et al.*, *Nat. Phys.* **6**, 811–817 (2010).
34. S. Jung *et al.*, *Nat. Phys.* **7**, 245–251 (2011).
35. A. Luican, G. Li, E. Y. Andrei, *Phys. Rev. B* **83**, 041405 (2011).
36. A. Luican-Mayer *et al.*, *Phys. Rev. Lett.* **112**, 036804 (2014).
37. F. Ghahari *et al.*, *Science* **356**, 845–849 (2017).
38. C. Gutiérrez *et al.*, *Science* **361**, 789–794 (2018).
39. Z. Papić, R. S. K. Mong, A. Yazdani, M. P. Zaletel, *Phys. Rev. X* **8**, 011037 (2018).
40. J. P. Eisenstein, L. N. Pfeiffer, K. W. West, *Phys. Rev. Lett.* **69**, 3804–3807 (1992).
41. O. E. Dial, R. C. Ashoori, L. N. Pfeiffer, K. W. West, *Nature* **448**, 176–179 (2007).
42. O. E. Dial, R. C. Ashoori, L. N. Pfeiffer, K. W. West, *Nature* **464**, 566–570 (2010).
43. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162 (2009).
44. A. F. Young *et al.*, *Nat. Phys.* **8**, 550–556 (2012).
45. C. R. Dean *et al.*, *Nat. Phys.* **7**, 693–696 (2011).
46. S.-Y. Li, Y. Zhang, L.-J. Yin, L. He, *Phys. Rev. B* **100**, 085437 (2019).
47. A. Coissard *et al.*, arXiv 2110.02811 [cond-mat.mes-hall] (2021).
48. X. Liu *et al.*, Replication Data for “Visualizing Broken Symmetry and Topological Defects in a Quantum Hall Ferromagnet.” Harvard Dataverse (2021); doi:10.7910/DVN/JZWE2C.

## ACKNOWLEDGMENTS

We thank B. I. Halperin, A. H. MacDonald, N. P. Ong, and P. Kim for helpful discussions. **Funding:** Supported by NSF-DMR-1904442 and ONR-N00014-21-1-2592; Gordon and Betty Moore Foundation EPIQS initiative grant GBMF9469; NSF-MRSEC through the Princeton Center for Complex Materials NSF-DMR-2011750; DOE-BES grant DE-FG02-07ER46419; the Princeton Catalysis Initiative; the Elemental Strategy Initiative, Japan, grant JPMXP0112101001, JSPS KAKENHI grant JP20H00354, and CREST (JPMJCR15F3). JST (K.W. and T.T.); and the Army Research Office through the MURI program (grant W911NF-17-1-0323) (M.P.Z.). A.Y. acknowledges the hospitality of the Aspen Center for Physics, which is supported by NSF grant PHY-1607611, and Trinity College, where his stay was supported by a QuantEmX grant from ICAM and the Gordon and Betty Moore Foundation through grant GBMF9616. **Author contributions:** X.L., G.F., C.L.-C., and A.Y. designed the experiment. G.F., X.L., and C.L.-C. fabricated the sample. X.L., G.F., and C.L.-C. performed the measurements and analyzed the data. M.P.Z., Z.P., and X.L. conducted the theoretical analysis. K.W. and T.T. provided hBN crystals. X.L., G.F., C.L.-C., A.Y., and M.P.Z. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data from this study are available at the Harvard Dataverse (48).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm3770  
Materials and Methods  
Supplementary Text  
Figs. S1 to S7  
References (49–59)

13 September 2021; accepted 17 November 2021  
Published online 2 December 2021  
10.1126/science.abm3770

## MOLECULAR MOTORS

# Structural and functional insight into regulation of kinesin-1 by microtubule-associated protein MAP7

Luke S. Ferro<sup>1†</sup>, Qianglin Fang<sup>1\*††</sup>, Lisa Eshun-Wilson<sup>1†</sup>, Jonathan Fernandes<sup>2†</sup>, Amanda Jack<sup>3</sup>, Daniel P. Farrell<sup>4</sup>, Mert Golcuk<sup>5</sup>, Teun Huijben<sup>6</sup>, Katelyn Costa<sup>7</sup>, Mert Gur<sup>5</sup>, Frank DiMaio<sup>4</sup>, Eva Nogales<sup>1,3,8,9\*</sup>, Ahmet Yildiz<sup>1,3,9,10\*</sup>

Microtubule (MT)–associated protein 7 (MAP7) is a required cofactor for kinesin-1–driven transport of intracellular cargoes. Using cryo–electron microscopy and single–molecule imaging, we investigated how MAP7 binds MTs and facilitates kinesin-1 motility. The MT-binding domain (MTBD) of MAP7 bound MTs as an extended  $\alpha$  helix between the protofilament ridge and the site of lateral contact. Unexpectedly, the MTBD partially overlapped with the binding site of kinesin-1 and inhibited its motility. However, by tethering kinesin-1 to the MT, the projection domain of MAP7 prevented dissociation of the motor and facilitated its binding to available neighboring sites. The inhibitory effect of the MTBD dominated as MTs became saturated with MAP7. Our results reveal biphasic regulation of kinesin-1 by MAP7 in the context of their competitive binding to MTs.

Kinesin and dynein are molecular motors that deliver intracellular cargoes by walking along microtubules (MTs) (1, 2). Intracellular cargoes are differentially regulated by structural MT-associated proteins (MAPs) that decorate the MT surface (3). Distinct cellular localizations of MAPs correlate with their regulatory roles in intracellular traffic (3). Overexpression of tau disrupts kinesin-1 (hereafter kinesin)–driven transport of synaptic vesicles in axons (4, 5), whereas the knockdown of tau rescues defects in axonal transport in Alzheimer’s disease models (6). Unlike tau, MAP7 is a required cofactor for kinesin-driven transport in cells (7–9). The MAP7 projection domain binds to kinesin’s coiled-coil stalk in vitro (9, 10), recruits kinesin to MTs, and activates its motility (9, 11). Transient interactions with the MAP7 projection domain may enable kinesin to hop from one MAP to another, increasing its apparent run length by disfavoring detachment from the MT (9).

To understand how MAP7 regulates kinesin, we determined the cryo–electron microscopy (cryo-EM) structure of MTs decorated with full-length (FL) MAP7 (Fig. 1, A and B; fig. S1A; tables S1 and S2; and movie S1). The reconstruction revealed a 53-residue-long  $\alpha$  helix

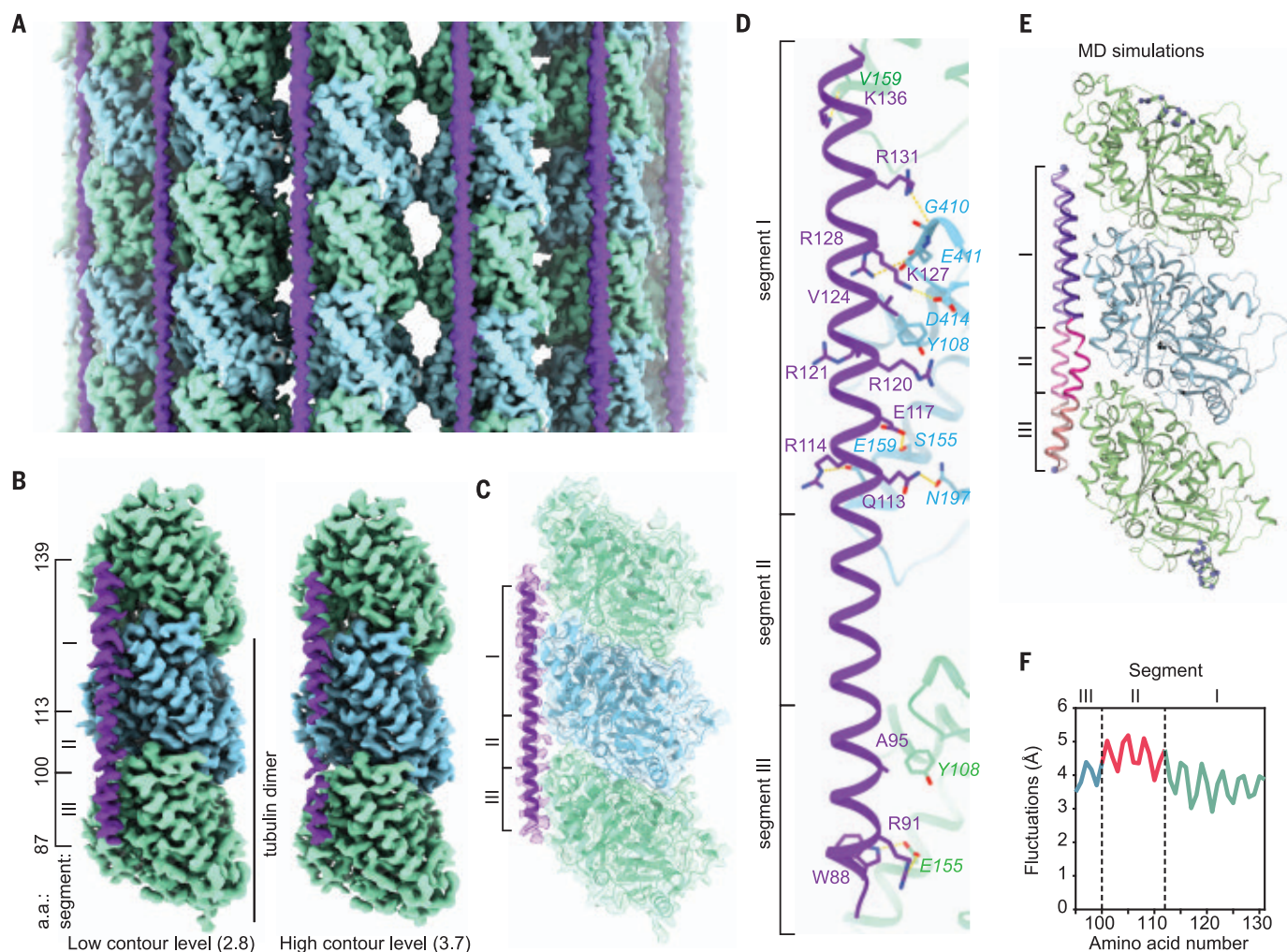
that runs parallel to the MT axis about the length of a tubulin dimer (Fig. 1, A to C, and fig. S1B). Unlike MAP2, MAP4, and tau, which bind along the outer ridges of the protofilaments (12–14), MAP7 runs halfway between the outer ridge and the site of lateral contact. The costructure of MAP7’s MT binding domain (MTBD; residues 60 to 170) and FL tau on the MT (fig. S2) illustrated their distinct MT footprints and confirmed that the helical segment corresponds to the MTBD of MAP7. We identified a single MAP7 sequence register that corresponds to a well-conserved segment of the MTBD (residues 87 to 139; fig. S3, A to C) through Rosetta modeling (15) and validated this registry by determining the structure of MTs decorated with a shorter MAP7 construct (residues 83 to 134; fig. S3D). Because MAP7 MTBD could potentially form a helix longer than the length of a tubulin heterodimer (fig. S2A), we cannot exclude the possibility of a larger footprint of MAP7 on the MT (13) (fig. S4; see materials and methods).

The  $\alpha$ -helical density for MAP7 is not uniform (Fig. 1B). Segment I (residues 113 to 139)—the best-resolved region (Fig. 1B)—interacts extensively with tubulin (Fig. 1D); Q113 and E117 of MAP7 are within hydrogen-bonding distance of N197 and S155 of  $\beta$ -tubulin, respectively; R114 and K127 of MAP7 engage in electrostatic interactions with E159 and D414 of  $\beta$ -tubulin, respectively. Y108 of  $\beta$ -tubulin inserts into a hydrophobic pocket formed by R120, R121, and V124 of MAP7. We also identified potential hydrogen bonds between R128, R131, and K136 of MAP7 and the mainchain oxygens of E411 and G410 of  $\beta$ -tubulin and V159 of  $\alpha$ -tubulin, respectively. Segment III (residues 87 to 99) interacts with  $\alpha$ -tubulin. Segment II (residues 100 to 113) faces a cavity at the intratubulin dimer and has the weakest density in our map (Fig. 1B). All-atom molecular dynamics (MD) simulations verified these pairwise

<sup>1</sup>Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. <sup>2</sup>Department of Chemistry, University of California, Berkeley, CA, USA. <sup>3</sup>Biophysics Graduate Group, University of California, Berkeley, CA, USA. <sup>4</sup>Department of Biochemistry, University of Washington, Seattle, WA, USA. <sup>5</sup>Department of Mechanical Engineering, Istanbul Technical University, Istanbul, Turkey. <sup>6</sup>Department of Imaging Physics, Delft University of Technology, Delft, Netherlands. <sup>7</sup>Press West Illustrations, Boston, MA, USA. <sup>8</sup>Howard Hughes Medical Institute, University of California, Berkeley, CA, USA. <sup>9</sup>Molecular Biophysics and Integrative Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>10</sup>Physics Department, University of California, Berkeley, CA, USA.  
**\*Corresponding authors.** yildiz@berkeley.edu (A.Y.); enogales@lbl.gov (E.N.); qfang168@berkeley.edu (Q.F.)

†These authors contributed equally to this work.

‡Present address: School of Public Health, Sun Yat-sen University, Shenzhen, China.



**Fig. 1. MAP7 binds the MT between the outer protofilament ridge and the site of lateral contact.** (A) Cryo-EM map (without symmetry expansion) of an MT decorated with MAP7;  $\alpha$ -tubulin,  $\beta$ -tubulin, and MAP7 are shown in green, blue, and purple, respectively. (B) Improved MAP7-MT cryo-EM map after symmetry expansion and protofilament-based density subtraction (see methods). MAP7 binds across both inter- and intradimer interfaces. Weaker density is seen for the region over the intradimer interface (segment II), indicative of more flexibility and weaker interaction. Only one repeat of MAP7 and its neighboring tubulins are shown for clarity. a.a., amino acids. (C) Ribbon diagram for MAP7 and tubulin with the improved

cryo-EM density map shown in gray. (D) Details of the interacting residues between MAP7 and its neighboring tubulins. Tubulin residues are shown in italics. (E) Initial and final MAP7-tubulin conformations obtained from an example all-atom MD simulation. Beads represent constrained atoms in MAP7 and tubulin. (F) Average root mean squared fluctuations of MAP7 C $\alpha$  atoms from the cryo-EM structure coordinates within 200 ns ( $n = 4$  simulations). 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.

interactions, and identified additional potential interactions between MAP7 and tubulin (fig. S5 and movie S2). Segment II fluctuated more than other segments in simulations because it only makes transient contacts with tubulin (Fig. 1, E and F, and fig. S5), explaining why this segment has weak density in our cryo-EM map.

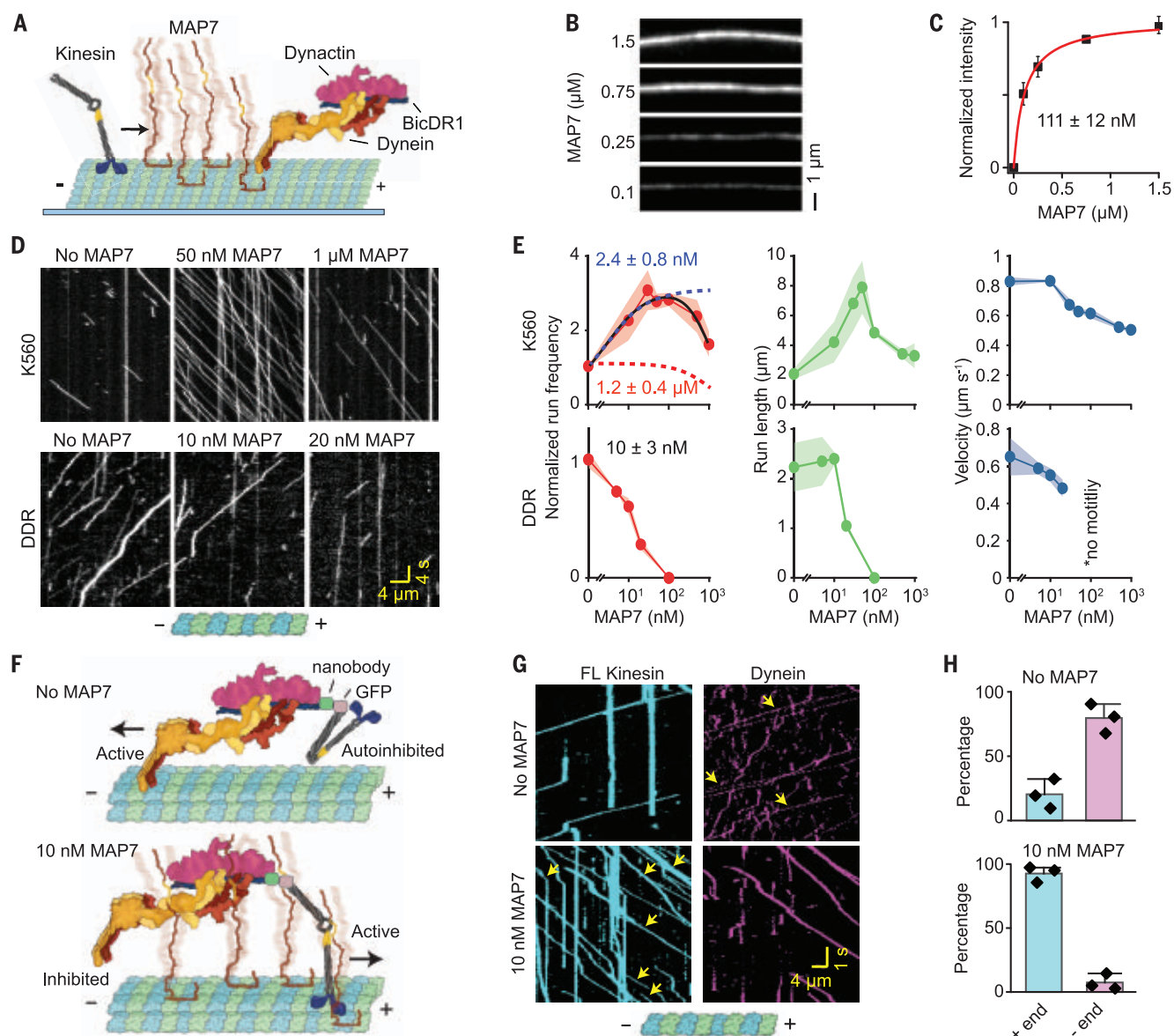
We next determined how MAP7 affects the motility of kinesin and the dynein-dynactin-BicDR1 complex [(DDR), hereafter dynein] (16) (Fig. 2A). FL MAP7 uniformly decorated MTs with a dissociation constant ( $K_D$ ) of  $111 \pm 12$  nM ( $\pm$ SE) under physiological salt (Fig. 2, B and C,

and fig. S6). As previously reported (9, 11), the addition of 50-nM MAP7 rescued FL kinesin from autoinhibition (17) and substantially increased its run frequency and length (fig. S7). MAP7 also enhanced the motility of constitutively active kinesin (K560, residues 1 to 560) (Fig. 2, D and E and movie S3) (9, 11). Unlike run length and frequency, kinesin velocity decreases even at low MAP7 concentrations (Fig. 2E), which could be a result of pausing at MTBD obstacles or binding to the projection domain. Although previously, dynein has been reported not to be inhibited by 5-nM MAP7 (10); how-

ever, we found that it was inhibited by MAP7 with a half-maximal inhibition constant ( $IC_{50}$ ) of  $10 \pm 3$  nM (Fig. 2, D and E). MAP7 decoration of MTs also switched the direction of an assembly that links FL kinesin to dynein (Fig. 2F) (18). 80% of kinesin-dynein assemblies were minus-end-directed on undecorated MTs, whereas 93% moved toward the plus end upon addition of 10 nM MAP7 (Fig. 2, G and H, fig. S6C, and movie S4).

Unexpectedly, kinesin run frequency started to decrease when MAP7 concentration was increased further (100 to 1000 nM; Fig. 2, D and





**Fig. 2. MAP7 differentially regulates kinesin and dynein motility.**

(A) Schematic of kinesin and dynein motility on MTs coated with MAP7. (B) MT decoration of fluorescently labeled MAP7 under different concentrations. (C) Fluorescence intensity (mean  $\pm$  SD) of MAP7 was fit to the Langmuir isotherm (solid curve) to calculate  $K_D$  ( $\pm$  SE). From left to right,  $n = 40, 44, 50, 68,$  and  $109$  MTs (two technical replicates). (D) Kymographs of K560 and dynein motility in the presence of MAP7. Assays were performed in 150 mM potassium acetate (KAc) and 0.1% methylcellulose. (E) Run frequency, run length, and velocity of K560 and dynein at different MAP7 concentrations (mean  $\pm$  SEM). K560 run frequency was fit to the biphasic Hill equation

(solid black curve) to reveal the half-maximal activation (blue dashed curve) and inhibition (red dashed curve) concentrations ( $\pm$  SE). Dynein run frequency was fit to the Langmuir equation (not shown) to calculate  $\text{IC}_{50}$  ( $\pm$  SE). From left to right,  $n = 281, 463, 532, 836, 381,$  and  $433,$  and  $233$  for K560, and  $386, 235, 213,$  and  $146$  for dynein; two technical replicates). (F) Schematic of kinesin and dynein assemblies assembled onto BicDR1. The addition of MAP7 switches the active motor. (G) Kymographs of kinesin-dynein assemblies with and without 10 nM MAP7. Yellow arrows show kinesin-dynein colocalizers. (H) Directionality of kinesin-dynein assemblies (mean  $\pm$  SD,  $n = 356$  and  $580$  for 0- and 10 nM MAP7, respectively, three replicates).

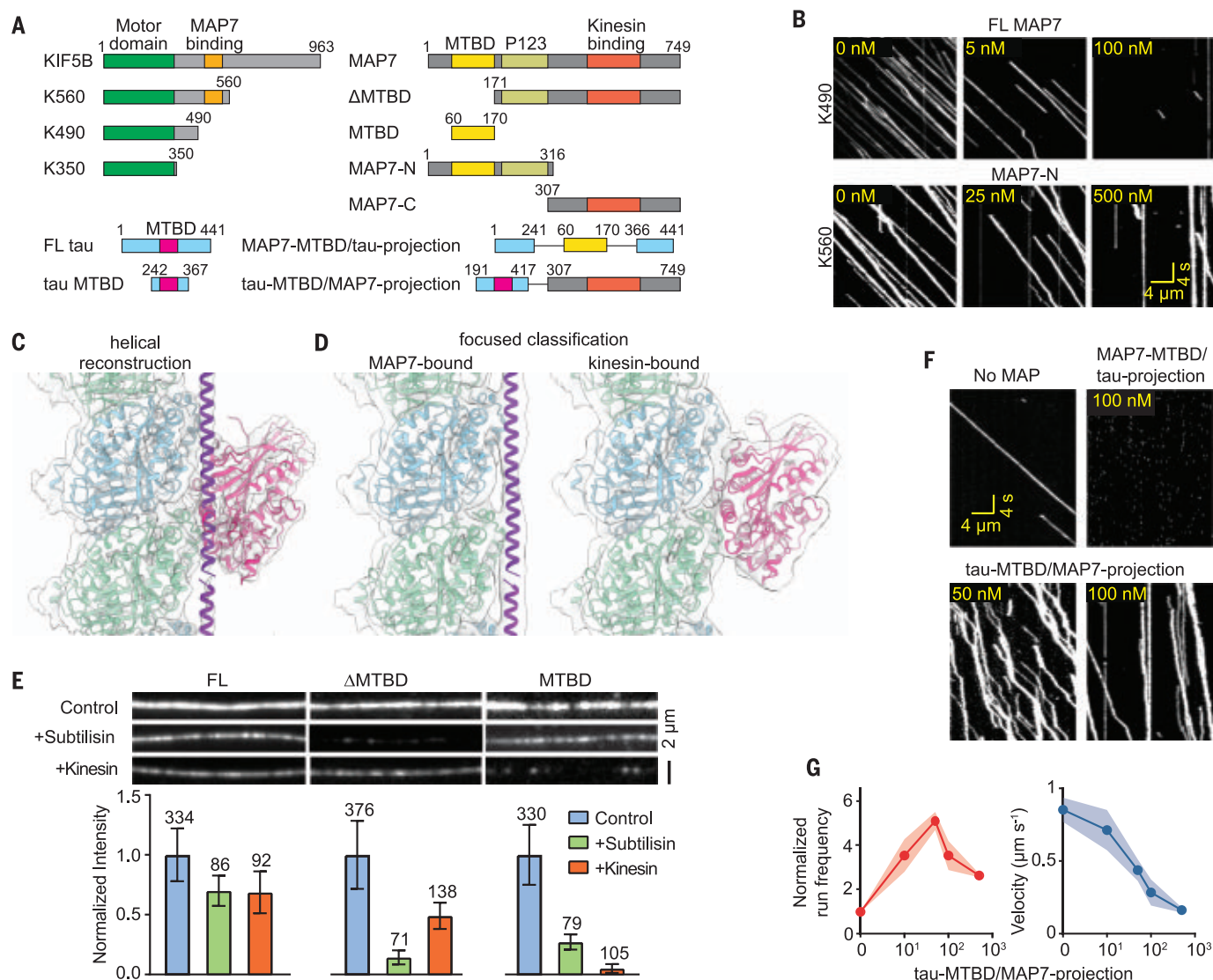
E, and movie S3). We reasoned that the non-linear relationship between MAP7 decoration and kinesin motility could arise when the motor is subjected to simultaneous activation and inhibition that dominate at different concentrations (Fig. 2E) (19). To determine the source of these opposing inputs, we disrupted the kinesin-MAP7 interaction by truncating either the MAP7-binding domain of kinesin

(K490, residues 1 to 490) or the kinesin-binding domain of MAP7 (MAP7-N and MAP7-MTBD; Fig. 3A). In all cases, MAP7 inhibited kinesin (Fig. 3B, fig. S8, and movies S5 and S6) as strongly as it inhibited dynein (Fig. 2E). Thus, MTBD inhibits whereas the projection domain activates kinesin (10).

MTBD may inhibit kinesin by competing for the same tubulin binding site, because

superimposing the MT-bound structure of kinesin (20) onto our model reveals an apparent clash between kinesin and segment II of MAP7 (fig. S9A). Alternatively, the flexible segment II may accommodate kinesin binding by shifting away from the intradimer interface, as proposed for DCX and MAP4 (14, 21). To distinguish between these possibilities, we determined the structure of MTs incubated





**Fig. 3. MT tethering enables kinesin to move on MAP7-decorated MTs despite their overlapping binding sites.** (A) Schematic of kinesin, MAP7, and tau constructs. (B) Kymographs showing kinesin motility in the presence of FL and truncated MAP7. (C) Average cryo-EM map of an MT decorated with both MAP7 and rigor kinesin. Atomic model of MAP7 and previously reported structure of kinesin on MT (PDB code: 4HNA) (20) were fitted into the cryo-EM map. (D) Focused 3D classification resulted in two distinct density maps showing either MAP7-bound

(left) or kinesin-bound (right) tubulin, indicating competitive binding. (E) Fluorescent signal (top) and normalized intensity (bottom) of 1  $\mu$ M LD655-labeled MAP7 constructs on MTs that are either untreated (control), subtilisin-treated, or predecorated with 1  $\mu$ M rigor kinesin (mean  $\pm$  SEM; *n* values are given for each bar). (F) Kinesin motility in the presence of chimeric MAPs. (G) Run frequency and velocity of kinesin in the presence of Tau-MTBD/MAP7-projection (mean  $\pm$  SEM, *n* = 27, 112, 211, 102, and 55 from left to right, two technical replicates).

with rigor kinesin [K350, residues 1 to 350, E236A mutation (K350<sup>E236A</sup>)] and MAP7 (Fig. 3C). Because distinct binding sites of MAP7 and kinesin can be artifactually averaged during reconstruction (fig. S9B), we performed focused 3D classification around the putative shared binding site of MAP7 and kinesin (fig. S10; see methods). The classification resulted in two distinct maps of the binding site, one occupied by MAP7 only, and the other occupied by kinesin only (Fig. 3D). Thus, kinesin and MAP7 cannot simultaneously bind to the same tubulin dimer.

Consistent with cryo-EM, MTBD was unable to bind MTs predecorated with K350<sup>E236A</sup>

(Fig. 3E). However, FL MAP7 or a construct lacking the MTBD ( $\Delta$ MTBD; fig. S11, A to C) still bound to K350<sup>E236A</sup>-decorated MTs, presumably through the P123 domain (residues 175 to 316) (8, 9). MT binding of  $\Delta$ MTBD was nearly abolished upon cleaving the flexible tails of tubulin (Fig. 3E and fig. S11D), explaining why P123 was invisible in our structure. To verify that kinesin is inhibited at high levels of MAP7 decoration as a result of binding site overlap, we also replaced the MT-binding regions of MAP7 with that of tau, which also overlaps with the tubulin binding site of kinesin and inhibits kinesin motility

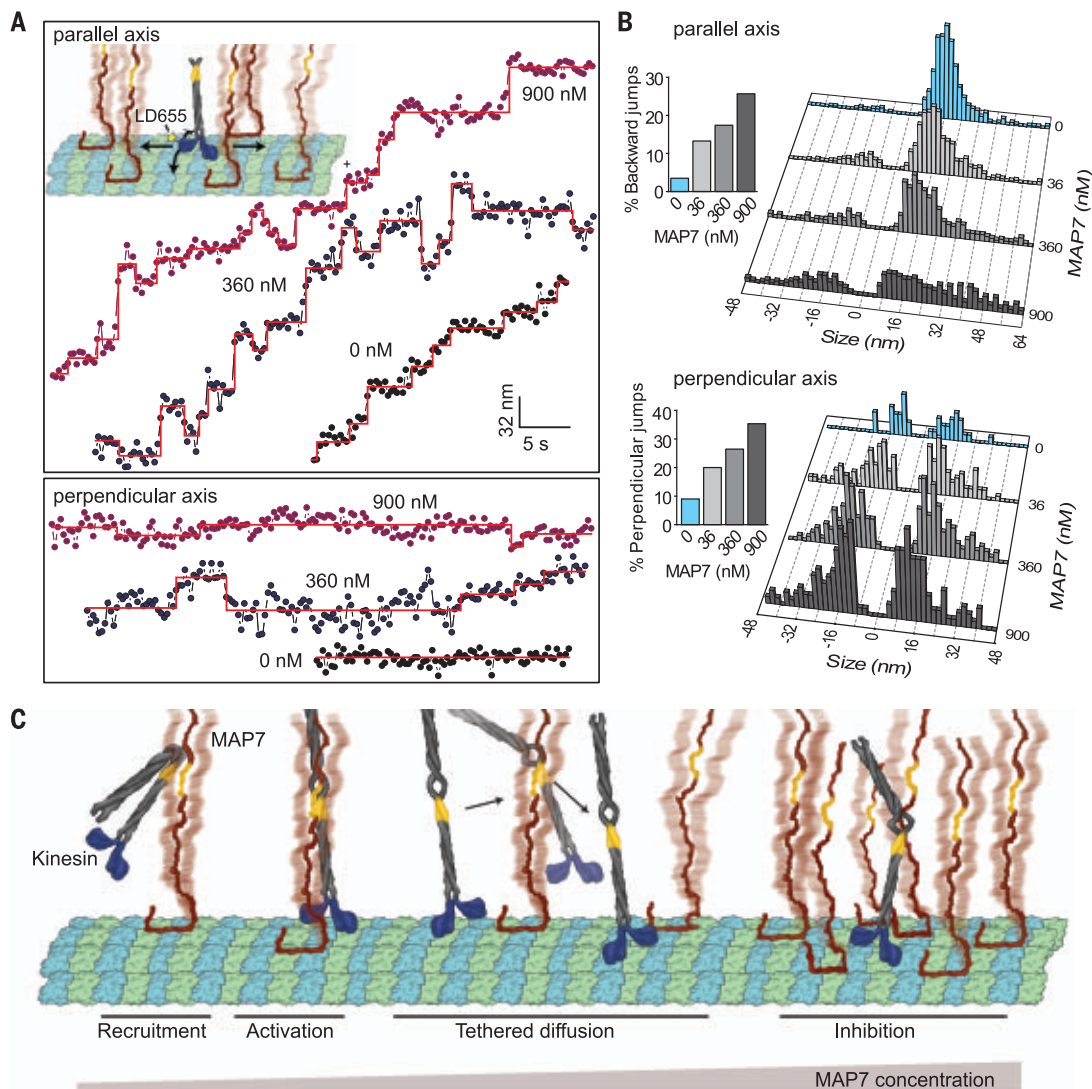
(fig. S12) (13, 22). Similar to FL MAP7, binding of this chimeric MAP to the MT resulted in biphasic regulation of kinesin motility (Fig. 3, F and G).

To reveal how kinesin “walks” along MAP7-decorated MTs despite competing for the same binding site, we fluorescently labeled the motor domain of K560 and tracked kinesin stepping with nanometer precision under limiting adenosine triphosphate (ATP) conditions (Fig. 4A). On undecorated MTs, kinesin took 16-nm steps in the forward direction (23), whereas steps in the sideways and backward directions were rare (9 and 3%, respectively; Fig. 4, A

#### Fig. 4. Kinesin bypasses MTBD obstacles through tethered diffusion.

(A) (Inset) K560 was labeled with LD655 at its N terminus, and its stepping was tracked in parallel (straight arrows) and perpendicular (curved arrows) axes of MTs under different MAP7 concentrations. Representative traces of K560 motility along parallel (top) and perpendicular (bottom) directions. Horizontal lines represent a fit to a step-finding algorithm. (B) Histograms reveal the percentage of instantaneous jumps in backward and perpendicular directions under different MAP7 concentrations (bar graphs). From top to bottom,  $n = 437$ , 580, 502, and 331 for parallel and 43, 145, 195, and 181 for perpendicular directions.

(C) Model for regulation of kinesin by MAP7. The MAP7 projection domain rescues kinesin from autoinhibition and tethers the motor to the MT. When kinesin encounters an MTBD obstacle, it dissociates from the MT, remains tethered to MAP7, and rebinds to an available tubulin site on another protofilament. Kinesin is inhibited at high MAP7 concentrations due to the scarcity of available binding sites.



and B, and fig. S13A) (24). At 36 nM MAP7, we observed a modest increase in the probability of sideways and backward stepping (Fig. 4, A and B). At nearly saturating MAP7 concentrations, we observed 16- to 64-nm displacements in our trajectories, as well as increased stepping in sideways (35%) and backward (26%) directions (Fig. 4, A and B). Because kinesin cannot take such large steps on its own (Fig. 4, A and B), these large displacements likely represent transient detachment and reattachment of the motor to the MT.

In the absence of the projection domain, kinesin was stuck on MTs decorated with 36-nM MTBD and dissociated from MTs at increased ATP concentrations (fig. S13, B and C). Furthermore, the addition of the C-terminal half of MAP7 (MAP7-C) was unable to stimulate kinesin motility on MTs decorated with the N-terminal half of MAP7 (MAP7-N, fig. S14). Thus, the projection domain is needed to be

tethered to the MTBD to enable to bypass MTBD obstacles.

On the basis of our results and previous reports (9, 11), we propose a model for kinesin stepping on MAP7-decorated MTs. MAP7 recruits kinesin-1 to the MT and activates subsequent motility (9, 11). MT binding of MAP7 also obstructs kinesin stepping along the protofilament, which results in kinesin dissociation from the MT. However, the MAP7 projection domain tethers kinesin to the MT surface, allowing it to rebound the MT at nearby sites not blocked by its MTBD. This “tethered diffusion” of kinesin appears as large forward, sideways, or backward displacements in our trajectories. When the MT surface is nearly saturated with MAP7, the frequency and length of kinesin runs are reduced because of the scarcity of empty tubulin sites to which the motor can rebound after it dissociates (Fig. 4C). Unlike kinesin-1, MAP7 inhibits dynein motility because its

MTBD overlaps with the dynein binding site (fig. S15) but its projection domain does not tether dynein to MTs.

This biphasic regulation mechanism may enable precise control of kinesin-1-driven transport by varying MAP7 density on cellular MTs. Although MAP7 is required for kinesin-1-driven processes in many cell types (10, 25, 26), dense MAP7 localization was shown to slow and pause organelle transport at branch junctions in rat neurons (8). Increased accumulation of MAP7 may reroute cargos to their destinations by facilitating the detachment of kinesin from its MT track and rebinding to neighboring MTs at these junctions (8).

#### REFERENCES AND NOTES

1. S. Bodakuntla, A. S. Jijumon, C. Villablanca, C. Gonzalez-Billault, C. Janke, *Trends Cell Biol.* **29**, 804–819 (2019).
2. A. Roll-Mecak, *Dev. Cell* **54**, 7–20 (2020).
3. L. F. Gummy et al., *Neuron* **94**, 347–362.e7 (2017).
4. A. Ebneth et al., *J. Cell Biol.* **143**, 777–794 (1998).



5. T. Ishihara *et al.*, *Neuron* **24**, 751–762 (1999).
6. K. A. Vossell *et al.*, *Science* **330**, 198 (2010).
7. K. Barlan, W. Lu, V. I. Gelfand, *Curr. Biol.* **23**, 317–322 (2013).
8. S. R. Tyrmanskyj, B. H. Yang, K. J. Verhey, L. Ma, *eLife* **7**, e36374 (2018).
9. P. J. Hooikaas *et al.*, *J. Cell Biol.* **218**, 1298–1318 (2019).
10. B. Y. Monroy *et al.*, *Nat. Commun.* **9**, 1487 (2018).
11. B. Y. Monroy *et al.*, *Dev. Cell* **53**, 60–72.e4 (2020).
12. J. Al-Bassam, R. S. Ozer, D. Safer, S. Halpain, R. A. Milligan, *J. Cell Biol.* **157**, 1187–1196 (2002).
13. E. H. Kellogg *et al.*, *Science* **360**, 1242–1246 (2018).
14. H. Shigematsu *et al.*, *J. Cell Biol.* **217**, 4155–4163 (2018).
15. Y. Song *et al.*, *Structure* **21**, 1735–1742 (2013).
16. J. T. Canty, R. Tan, E. Kusakci, J. Fernandes, A. Yildiz, *Annu. Rev. Biophys.* **50**, 549–574 (2021).
17. K. J. Verhey, N. Kaul, V. Soppina, *Annu. Rev. Biophys.* **40**, 267–288 (2011).
18. A. R. Chaudhary *et al.*, *J. Biol. Chem.* **294**, 10160–10171 (2019).
19. M. L. Valenstein, A. Roll-Mecak, *Cell* **164**, 911–921 (2016).
20. B. Gigant *et al.*, *Nat. Struct. Mol. Biol.* **20**, 1001–1007 (2013).
21. J. S. Liu *et al.*, *Mol. Cell* **47**, 707–721 (2012).
22. R. Dixit, J. L. Ross, Y. E. Goldman, E. L. Holzbaur, *Science* **319**, 1086–1089 (2008).
23. A. Yildiz, M. Tomishige, R. D. Vale, P. R. Selvin, *Science* **303**, 676–678 (2004).
24. R. Schneider, T. Korten, W. J. Walter, S. Diez, *Biophys. J.* **108**, 2249–2257 (2015).
25. H. H. Sung *et al.*, *Dev. Cell* **15**, 866–876 (2008).
26. E. Gallaud *et al.*, *J. Cell Biol.* **204**, 1111–1121 (2014).

## ACKNOWLEDGMENTS

We thank A. P. Carter and R. Tan for helpful discussions, L. Nocka for SEC-MALS experiments, B. LaFrance for MT image analysis advice, P. Grob, and J. Atherton for MT processing in RELION, D. Toso, J. Remis, and A. Chintangal for microscopy and computational support, the QB3 Macrolab for competent cell lines and TEV protease purification, the UC Berkeley Cell Culture Facility for providing the insect cells, the Cal-Cryo facility at UC Berkeley for EM imaging, and the Marconi100 for MD simulations.

**Funding:** This work was supported by grants from the National Institute of General Medical Sciences (GM094522 (A.Y.), GM123655-03 (L.F.), GM051487 (E.N.), GM127018 (E.N.), and the National Science Foundation (MCB-1617028 and MCB-1055017, A.Y.), PRACE (2019215144, MG), and Istanbul Technical University BAP (MGA-2021-42803, MG). E.N. is a Howard Hughes Medical Institute Investigator. **Author contributions:** L.F., L.E., Q.F., E.N., and A.Y. conceived the project and analyzed the data. L.F. and J.F. purified the proteins and performed single-molecule experiments. Q.F. and L.E. performed cryo-EM sample preparation, data collection, and analysis. Q.F., D.P.F., and F.D. performed Rosetta modeling. A.J. performed stepping measurements. M. Go. and M. Gu. performed MD simulations. T.H. collected and analyzed data for K490. K.C. created scientific illustrations. L.F., Q.F., L.E., M. Gur, E.N., and A.Y. wrote the manuscript, with further edits from all authors.

**Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in data S1. Materials are available from A.Y. under a material agreement with the University of California, Berkeley. The coordinates for MAP7 bound to tubulin are available at the Protein Data Bank (PDB) with accession code 7SGS. All cryo-EM maps are available at the EMDB with accession codes EMD-25120 (FL MAP7), EMD-25119 (MAP7-MTBD and FL tau), EMD-25118 (MAP7 and kinesin), EMD-25117 (MAP7<sup>83-134</sup>).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abf6154  
Materials and Methods  
Figs. S1 to S16  
Tables S1 and S2  
References (27–57)  
MDAR Reproducibility Checklist  
Movies S1 to S6  
Data S1

7 November 2020; resubmitted 28 July 2021  
Accepted 8 December 2021  
10.1126/science.abf6154

## CORONAVIRUS

# SARS-CoV-2 vaccine protection and deaths among US veterans during 2021

Barbara A. Cohn<sup>1†</sup>, Piera M. Cirillo<sup>1,2†</sup>, Caitlin C. Murphy<sup>3†</sup>, Nickilou Y. Krigbaum<sup>1,2</sup>, Arthur W. Wallace<sup>2,4\*</sup>

We report severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccine effectiveness against infection (VE-I) and death (VE-D) by vaccine type in 780,225 veterans in the Veterans Health Administration, covering 2.7% of the US population. From February to October 2021, VE-I declined for all vaccine types, and the decline was greatest for the Janssen vaccine, resulting in a VE-I of 13.1%. Although breakthrough infection increased risk of death, vaccination remained protective against death in persons who became infected during the Delta variant surge. From July to October 2021, VE-D for age <65 years was 73.0% for Janssen, 81.5% for Moderna, and 84.3% for Pfizer-BioNTech; VE-D for age ≥65 years was 52.2% for Janssen, 75.5% for Moderna, and 70.1% for Pfizer-BioNTech. Findings support continued efforts to increase vaccination, booster campaigns, and multiple additional layers of protection against infection.

The mRNA vaccines BNT162b2 (Pfizer-BioNTech) and mRNA-1273 (Moderna) and the viral vector vaccine JNJ-78436735 (Janssen) have effectively prevented clinically recognized disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) since their rollout in the United States in late 2020 (1, 2). Vaccines have also reduced the incidence of asymptomatic infection and associated infectivity (3). However, by July 2021, the United States experienced a surge in cases of COVID-19, dominated by the B.1.617.2 (Delta) variant (4, 5). Initial reports, including follow-up of the Pfizer-BioNTech and Moderna trials (6–8), suggested sustained vaccine protection (9), but three reports by the US Centers for Disease Control and Prevention (CDC) in August 2021 (10–12) demonstrated that protection against infection had declined in mid-summer as the Delta variant rose to dominance; protection against hospitalization and death remained high (13–15). Breakthrough infections, illness, hospitalizations, and deaths have since continued to emerge in vaccine recipients.

This phenomenon has been most comprehensively monitored in Israel, where high levels of transmission of the Delta variant led to a resurgent outbreak in mid-June 2021 (16) despite a successful nationwide campaign to vaccinate the population (17). Israel authorized boosters of the Pfizer-BioNTech vaccine for adults age ≥60 years in July 2021 and extended this authorization to adults age ≥50 years in August 2021 (18). Rates of infec-

tion and severe illness subsequently declined in those who received a booster (19). Largely on the basis of these data, as well as data from the UK (20, 21), the US Food and Drug Administration (FDA) authorized boosters of the Pfizer-BioNTech vaccine for older (age ≥65 years) and higher-risk adults in September 2021 (22); the FDA similarly authorized boosters of the Moderna vaccine in October 2021, as well as boosters for all recipients of the Janssen vaccine (23).

The debate over boosters in the United States (24) has laid bare the limitations of its public health infrastructure: National data on vaccine breakthrough are inadequate. The CDC transitioned in May 2021 from monitoring all breakthrough infections to focus on identifying and investigating only hospitalized or fatal cases attributable to any cause, including causes not related to COVID-19 (25). Some data on vaccinations, infections, and deaths are collected through a patchwork of local health departments (10), but these data are frequently out of date and difficult to aggregate at the national level. We addressed this gap and examined SARS-CoV-2 infection and deaths by vaccination status in 780,225 veterans during the period 1 February 2021 to 1 October 2021, encompassing the emergence and dominance of the Delta variant in the United States.

The distribution of SARS-CoV-2 infection by demographics, comorbidity, and vaccination status is shown in table S1 for 1 February 2021 to 1 October 2021 ( $n = 780,225$  subjects). The percentage of polymerase chain reaction (PCR) test positivity is higher in veterans who were unvaccinated (25.8%), female (15.8%), Hispanic (13.9%), American Indian/Alaska Native (14.7%) or Native Hawaiian/Pacific Islander (14.2%), age <50 years at time of reverse transcription PCR (RT-PCR) assay (19.1%), and had a lower comorbidity score (16.2% for Charlson Comorbidity Index = 0) (26); 33,514 positive PCR

<sup>1</sup>Public Health Institute, Oakland, CA, USA. <sup>2</sup>San Francisco Veterans Affairs Medical Center, San Francisco, CA, USA.

<sup>3</sup>School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA. <sup>4</sup>Department of Anesthesiology and Perioperative Care, University of California, San Francisco, San Francisco, CA, USA.

\*Corresponding author. Email: art.wallace@va.gov

†These authors contributed equally to this work.



tests occurred in 498,148 fully vaccinated veterans. The distribution of vaccine type by demographic is shown in table S2. Vaccine type differed by age: Younger (age <50 years) veterans were more likely to have received the Janssen vaccine than either Moderna or Pfizer-BioNTech.

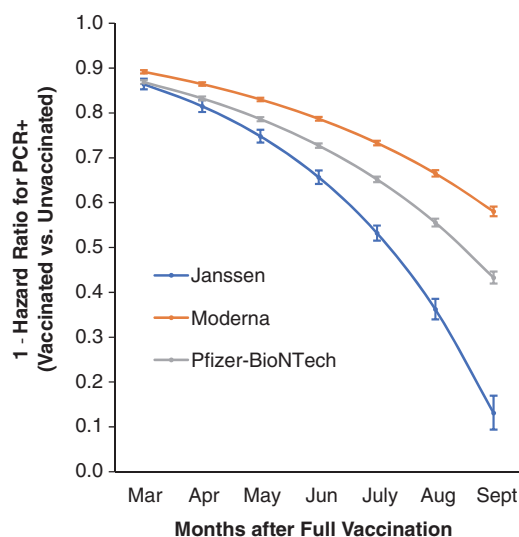
For the period 1 February 2021 to 1 October 2021, vaccine effectiveness against infection (VE-I) declined over time ( $P < 0.01$  for time dependence) (Table 1), even after adjusting for age, sex, and comorbidity. VE-I declined for all vaccine types (Fig. 1), with the largest declines for Janssen followed by Pfizer-BioNTech and Moderna. Specifically, in March, VE-I was 86.4% [95% confidence interval (CI), 85.2 to 87.6%] for Janssen, 89.2% (95% CI, 88.8 to 89.6%) for Moderna, and 86.9% (95% CI, 86.5 to 87.3%) for Pfizer-BioNTech. By September, VE-I had declined to 13.1% (95% CI, 9.2 to 16.8%) for Janssen, 58.0% (95% CI, 56.9 to 59.1%) for Moderna, and 43.3% (95% CI, 41.9 to 44.6%) for Pfizer-BioNTech.

As shown in Fig. 2, risk of infection accelerated in both unvaccinated and fully vaccinated veterans beginning in July 2021 and through September 2021, which is consistent with the time dependence observed in the Cox proportional hazards models. This pattern was similar across age groups, and risk of infection was highest for unvaccinated veterans. Veterans who were fully vaccinated with the Moderna vaccine had the lowest risk of infection, followed closely by those who received the Pfizer-BioNTech vaccine, then those who received the Janssen vaccine.

Risk of death after SARS-CoV-2 infection was highest in unvaccinated veterans regardless of age and comorbidity (Fig. 3). However, breakthrough infections were not benign, as shown by the higher risk of death in fully vaccinated veterans who became infected compared with vaccinated veterans who remained infection-free.

We observed similar results when examining the time period corresponding to the dominance of the Delta variant (fig. S1). Specifically, among those with a positive PCR test on or after 1 July 2021, vaccination was protective against death, although with some differences by age and vaccine type. For age <65 years, vaccine effectiveness against death (VE-D) was 81.7% (95% CI, 75.7 to 86.2%) for any vaccine, 73.0% (95% CI, 52.0 to 84.8%) for Janssen, 81.5% (95% CI, 70.7 to 88.4%) for Moderna, and 84.3% (95% CI, 76.3 to 89.7%) for Pfizer-BioNTech. For age ≥65 years, VE-D was 71.6% (95% CI, 68.6 to 74.2%) for any vaccine, 52.2% (95% CI, 37.2 to 63.6%) for Janssen, 75.5% (95% CI, 71.8 to 78.7%) for Moderna, and 70.1% (95% CI, 66.1 to 73.6%) for Pfizer-BioNTech.

Benefits of vaccination in reducing risk of SARS-CoV-2 infection and death are clearly



**Fig. 1. Time-dependent vaccine effectiveness against SARS-CoV-2 infection as estimated from Cox proportional hazards models, adjusted for age, race, ethnicity, sex, and comorbidity score.** Vaccine effectiveness presented as  $(1 - \text{hazard ratio} \times 100)$  and 95% CIs. Effectiveness for each month was estimated from contrasts by using product terms for vaccination status by time to most recent RT-PCR assay.

supported by this study of more than 780,225 US veterans. However, VE-I declined as risk increased in both unvaccinated and vaccinated veterans, coincident with the emergence and dominance of the Delta variant in the United States. Our analysis by vaccine type—including the Pfizer-BioNTech, Moderna, and Janssen vaccines—suggests a declining VE-I over time, particularly for the Janssen vaccine. Yet, despite increasing risk of infection because of the Delta variant, VE-D remained high, and compared with unvaccinated veterans, those fully vaccinated had a much lower risk of death after infection. These results demonstrate an urgent need to reinstate multiple layers of protection, such as masking and physical distancing—even among vaccinated persons—while also bolstering current efforts to increase vaccination.

Patterns of breakthrough SARS-CoV-2 infection among vaccinated veterans show a worrisome temporal trend, overlapping with the emergence of Delta as the dominant variant in the United States in July 2021. Although others have demonstrated high VE-I and VE-D in veterans during vaccine rollout through mid-March 2021 (27), our results suggest that vaccines are less effective in preventing infection associated with the Delta variant. The Delta variant is more infectious than other variants, likely because of increased viral load and transmission before symptoms (28). Other US studies (29–31), many conducted in large health care systems, similarly show declining VE-I as the Delta variant rose to dominance, with notable declines in older adults. For example, two studies conducted at Kaiser Permanente Southern California show that VE-I decreased from 95% at 14 to 60 days to 79% at 151 to 180 days after vaccination for ages 18 to 64 years (29), and from 80% at 1 month to 43% at 5 months after vaccination for ages ≥65 years (31). Declines

in protection against infection with Delta have been observed in Israel (16), the UK (20, 21), and Qatar (32, 33).

Endurance of VE-I in the face of the Delta variant in this large, population-based sample was dependent on vaccine type, and this was consistent across all age groups and time since vaccination. Most studies of VE-I have examined Moderna or Pfizer-BioNTech vaccines (16, 20, 21, 29–33), and our study adds to this literature by showing dramatic declines in VE-I for the Janssen vaccine. Similarly, we found that VE-D for the Janssen vaccine was much lower—about 50%—compared with that of the randomized trial. These findings are consistent with the better neutralizing antibody response observed after vaccination with Moderna or Pfizer-BioNTech compared with Janssen vaccines, and in response to the Delta variant (34). In addition, differences in immune response to mRNA vaccines by type of immunity support the more enduring protection against death (through cellular immunity) compared with protection against infection (which is more dependent on antibodies) (35).

Our findings on increased risk of death after breakthrough infection provide further support for continuing efforts to discover and implement effective interventions to prevent infection in all persons, including those who have been fully vaccinated. Fully vaccinated veterans were more likely to survive when infected with SARS-CoV-2 (breakthrough infections) compared with unvaccinated veterans who were also infected; this was true even for older age groups, those with more chronic conditions, and during and after the Delta surge in July 2021. However, breakthrough infections still carried some risk, as evidenced by the higher risk of death in vaccinated veterans who were subsequently infected compared with those who were vaccinated but remained infection-free. Breakthrough infections are also

**Table 1. Vaccine effectiveness against SARS-CoV-2 infection by month after vaccination; estimated from Cox proportional hazards models; and adjusted for age, race, ethnicity, sex, and comorbidity score.**

Adjusted hazard ratio <1.0 indicates lower risk of infection for vaccine, shown compared with unvaccinated. For vaccinated veterans, infection was assessed 15 days after the last vaccine that established full vaccination status. For unvaccinated veterans, infection was assessed beginning in 1 February 2021, coincident with broadscale vaccine eligibility in the VA. Time dependence was tested in Cox proportional hazards models by including product terms for vaccination status (Janssen, Moderna, Pfizer-BioNTech, or unvaccinated) by log(time)—Janssen\*log(time), Moderna\*log(time), Pfizer-BioNTech\*log(time)—and adjusted for age, sex, race, ethnicity, and comorbidity (Charlson Comorbidity score, overweight, type II diabetes, chronic obstructive pulmonary disease, bronchitis, acute respiratory failure, and chronic lung disease). Significance levels for all product terms were  $P < 0.0001$ . Vaccination status is modeled as time-varying, assigning follow-up time for veterans before the date of full vaccination as “unvaccinated time” and time after the date of full vaccination as “vaccinated time”; vaccination is defined as (i) a single Janssen vaccine, (ii) two Moderna vaccines, or (iii) two Pfizer-BioNTech vaccines.

	Adjusted hazard ratio	95% confidence interval		P value
Janssen versus unvaccinated*				
March	0.14	0.12	0.15	<0.01
April	0.19	0.17	0.20	<0.01
May	0.25	0.24	0.27	<0.01
June	0.34	0.33	0.36	<0.01
July	0.47	0.45	0.49	<0.01
August	0.64	0.62	0.66	<0.01
September	0.87	0.83	0.91	<0.01
Moderna versus unvaccinated*				
March	0.11	0.10	0.11	<0.01
April	0.14	0.13	0.14	<0.01
May	0.17	0.17	0.17	<0.01
June	0.21	0.21	0.22	<0.01
July	0.27	0.26	0.27	<0.01
August	0.33	0.33	0.34	<0.01
September	0.42	0.41	0.43	<0.01
Pfizer-BioNTech versus unvaccinated*				
March	0.13	0.13	0.14	<0.01
April	0.17	0.16	0.17	<0.01
May	0.21	0.21	0.22	<0.01
June	0.27	0.27	0.28	<0.01
July	0.35	0.34	0.35	<0.01
August	0.44	0.44	0.45	<0.01
September	0.57	0.55	0.58	<0.01

\*Associations at each month were estimated from contrasts by using product terms for vaccination status by time in Cox proportional hazards models, including indicator terms for vaccination status (Janssen, Moderna, or Pfizer-BioNTech) product terms and age, sex, race, ethnicity, and comorbidity (Charlson Comorbidity score, overweight, type II diabetes, chronic obstructive pulmonary disease, bronchitis, acute respiratory failure, and chronic lung disease).

a concern for transmission, and the Delta variant in particular results in high viral loads in the nose similar to that of infections in unvaccinated persons (36). Because viral load is a key determinant of transmissibility (37), the benefit of vaccination is less for the Delta variant compared with the earlier Alpha variant (38), suggesting that additional, alternative prevention practices will be essential to reduce infection. Higher risk of death after breakthrough infection implies higher rates of hospitalizations, and these prevention practices will likely also conserve medical resources.

Infection prevention in all persons will have the added, worldwide benefit of reducing the

potential for deleterious evolution of the viral genome as the infection is transmitted from person to person (37, 39). However, rates of vaccination—among other viral, social, political, and behavioral parameters—will determine the future evolution of the virus (37). Viral evolution may result in more lethal or infectious variants, or variants that escape protection the vaccine, and should be constricted by reducing infection rates.

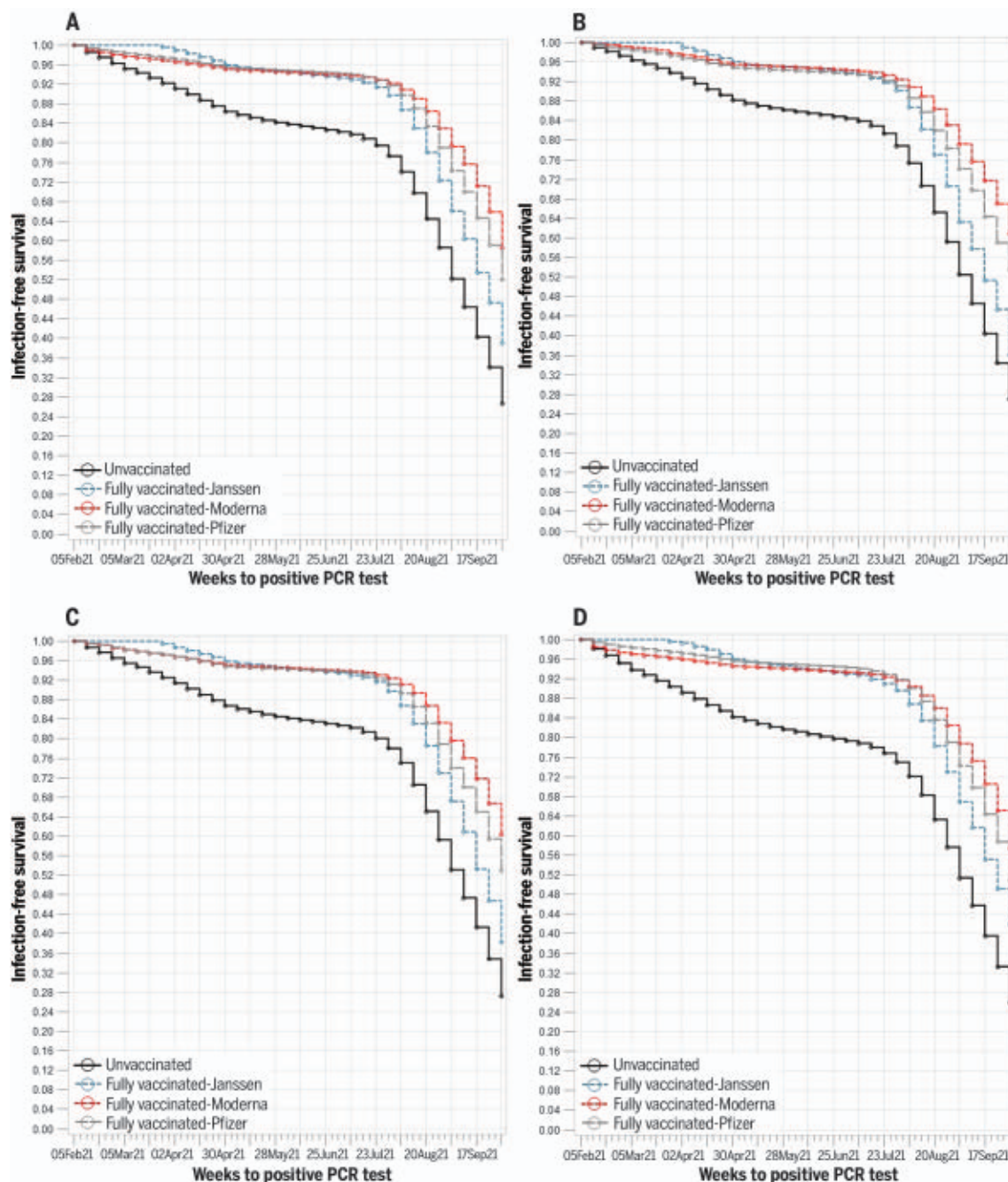
It is not yet known whether breakthrough infections increase risk of “long COVID” [otherwise known as post-acute sequelae of COVID-19 (PASC)], a constellation of debilitating and lingering symptoms after infection. These symp-

toms can lead to physiologic disruption of multiple organ systems; substantial disruption of daily life, employment, and mental health; and a higher burden on the health care system (40, 41). Long COVID has been observed as a consequence of both mild and severe infection (42), raising the possibility that survivors of breakthrough infections may also be at risk for long COVID. Therefore, prevention of breakthrough infections may avoid the overwhelming, long-term consequences of long COVID due to widespread infection.

As of this report, the scientific community continues to debate booster vaccines in the United States. The FDA authorized Pfizer-BioNTech boosters in September 2021 and Moderna and Janssen boosters in October 2021, and the CDC has made similar recommendations. Although our study does not directly address the benefits and risks of booster vaccines, findings may be interpreted in the context of this ongoing debate. First, VE-I declined most precipitously for the Janssen vaccine, and a booster with one of the mRNA vaccines may result in more durable protection for those initially vaccinated with Janssen. This is further supported by the available, albeit limited, evidence that suggests a stronger antibody response when Janssen vaccination is followed by an mRNA booster (43). Second, although their risk of death is much lower, younger veterans (age <65 years) experienced the greatest relative reduction in risk of death associated with vaccination, which suggests that this age group in addition to older adults may benefit from a booster. Early results of the first randomized trial on boosters demonstrates that a booster of Pfizer-BioNTech is 95.6% effective against infection compared with two shots and a placebo (43). Some unknowns remain—namely, how effective booster vaccines are against Delta and other emerging variants and how long immunity from a booster may last.

A strength of our study is the use of large-scale, national US Department of Veterans Affairs (VA) data, covering 2.7% of the US population and collected in real time. After transitioning to focus on breakthrough hospitalizations and deaths, the CDC now reports COVID-19 cases, associated hospitalizations, and deaths by vaccination status and age group (available at <https://covid.cdc.gov/covid-data-tracker>) as weekly rates per 100,000 persons; these data are derived from a network of acute-care hospitals in 14 states and 16 health departments that links case surveillance to immunization systems. Although informative, data lag behind by about 2 months and do not illustrate risk of hospitalization or death after a breakthrough infection. The VA Corporate Data Warehouse was essential to our timely analysis of breakthrough infections and deaths up until 1 October 2021, and moving forward,

**Fig. 2. Kaplan-Meier curves illustrating cumulative risk of SARS-CoV-2 infection, by vaccination status and age. (A) All ages. (B) Age <50 years. (C) Age 50 to 64 years. (D) Age ≥65 years. The survival function estimates time to infection detected by most recent RT-PCR assay.**



these data may be used as a tool to comprehensively monitor vaccine effectiveness because other variants are likely to emerge.

Our results should be interpreted in the context of limitations. There are many approaches to evaluating vaccine effectiveness (such as test-negative, case-control, and cohort registry). We required a recent RT-PCR assay to be included in the analysis, a feature of test-negative designs that may minimize confounding owing to health-seeking behavior. However, there may still be differences in testing intervals and frequency by vaccination status. The specific setting or reason for testing is not known, and it

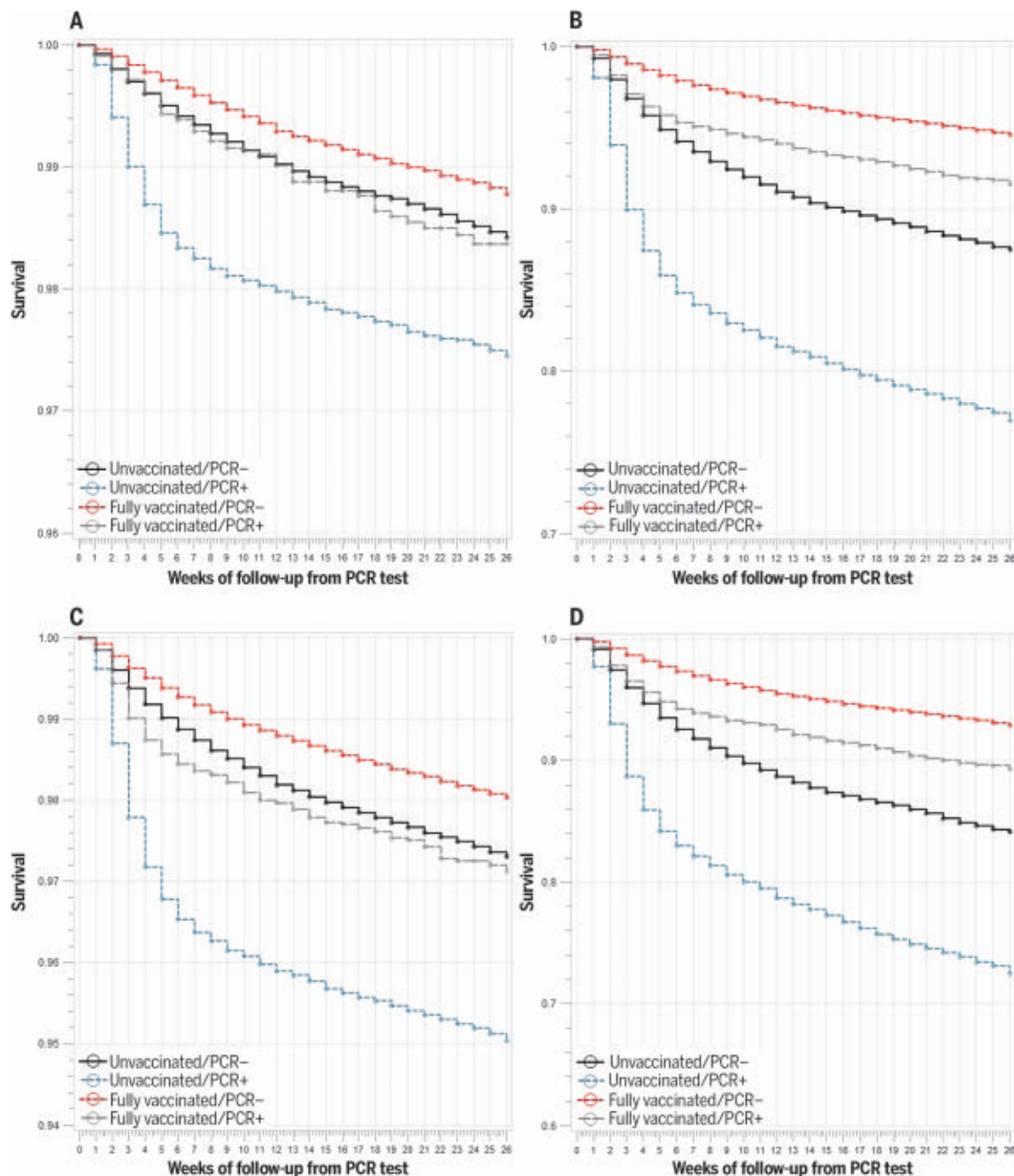
is also possible that persons with asymptomatic infections may not have been tested and therefore not included in the analysis. Our sample has proportionately fewer women, although a large number are still included. We did not have information on genotyping of infections to determine the proportion caused by the Delta variant. Patterns of survival for those with a negative PCR test by vaccination status suggests that there are underlying differences in unvaccinated compared with vaccinated persons, and that we did not measure or account for these in our analysis; these differences may contribute to the different risks of

death we observed. For example, recent polls suggest that unvaccinated Americans are less willing to adopt COVID-19 precautions, such as mask-wearing and social distancing (44). Last, we did not examine VE against hospitalization but used death as a surrogate for clinically severe infection. Our finding that VE-D remained high during the Delta surge is consistent with US studies that show sustained protection against hospitalization (15, 30, 45).

Although vaccination remains protective against SARS-CoV-2 infection, protection waned as the Delta variant emerged in the



**Fig. 3. Kaplan-Meier curves illustrating cumulative risk of death due to any cause, by vaccination status and RT-PCR assay. (A) Age <65 years. (B) Age ≥65 years. (C) Charlson Comorbidity Index score <3. (D) Charlson Comorbidity Index score ≥3.**



United States, and this decline did not differ by age. The Janssen vaccine showed the greatest decline in VE-I. Breakthrough infections were not benign; vaccinated persons who were subsequently infected had a higher risk of death compared with that of vaccinated persons who remained infection-free. Vaccination still provided protection against death in infected persons, and this benefit was observed for the Moderna, Pfizer-BioNTech, and Janssen vaccines during the Delta surge, although the benefit was greater for Moderna and Pfizer-BioNTech compared with Janssen vaccines. Our findings support the conclusion that

COVID-19 vaccines remain the most important tool to prevent infection and death. Vaccines should be accompanied by additional measures for both vaccinated and unvaccinated persons, including masking, hand washing, and physical distancing. It is imperative to implement public health interventions, such as strategic testing for control of outbreaks, vaccine passports, employment-based vaccine mandates, vaccination campaigns for eligible children as well as adults, and consistent messaging from public health leadership in the face of increased risk of infection from the Delta and other emerging variants.

#### REFERENCES AND NOTES

1. T. Pilišvili et al., *MMWR Morb. Mortal. Wkly. Rep.* **70**, 753–758 (2021).
2. A. Christie et al., *MMWR Morb. Mortal. Wkly. Rep.* **70**, 858–864 (2021).
3. M. G. Thompson et al., *MMWR Morb. Mortal. Wkly. Rep.* **70**, 495–500 (2021).
4. C. M. Brown et al., *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1059–1062 (2021).
5. R. Herlihy et al., *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1084–1087 (2021).
6. S. J. Thomas et al., *N. Engl. J. Med.* 10.1056/NEJMoa2110345 (2021).
7. N. Doria-Rose et al., *N. Engl. J. Med.* **384**, 2259–2261 (2021).
8. H. M. El Sahly et al., *N. Engl. J. Med.* 10.1056/NEJMoa2113017 (2021).
9. J. M. Polinski et al., *medRxiv* 21263385 [Preprint] (2021).

10. E. S. Rosenberg *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1306–1311 (2021).
11. S. Nanduri *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1163–1166 (2021).
12. A. Fowlkes *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1167–1169 (2021).
13. M. W. Tenforde *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1156–1162 (2021).
14. H. M. Scobie *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1284–1290 (2021).
15. W. H. Self *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1337–1343 (2021).
16. Y. Goldberg *et al.*, *medRxiv* 21262423 [Preprint] (2021).
17. N. Dagan *et al.*, *N. Engl. J. Med.* **384**, 1412–1423 (2021).
18. "Israel expands COVID vaccine booster campaign to over 50s, health workers," *Reuters*, 13 August 2021.
19. Y. M. Bar-On *et al.*, *N. Engl. J. Med.* **385**, 1393–1400 (2021).
20. J. Lopez Bernal *et al.*, *N. Engl. J. Med.* **385**, 585–594 (2021).
21. K. B. Pouwels *et al.*, *medRxiv* 21262237 [Preprint] (2021).
22. FDA, "FDA Authorizes Booster Dose of Pfizer-BioNTech COVID-19 Vaccine for Certain Populations"; [www.fda.gov/news-events/press-announcements/fda-authorizes-booster-dose-pfizer-biontech-covid-19-vaccine-certain-populations](https://www.fda.gov/news-events/press-announcements/fda-authorizes-booster-dose-pfizer-biontech-covid-19-vaccine-certain-populations).
23. FDA, "Coronavirus (COVID-19) Update: FDA Takes Additional Actions on the Use of a Booster Dose for COVID-19 Vaccines"; [www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-additional-actions-use-booster-dose-covid-19-vaccines](https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-additional-actions-use-booster-dose-covid-19-vaccines).
24. P. R. Krause *et al.*, *Lancet* **398**, 1377–1380 (2021).
25. CDC COVID-19 Vaccine Breakthrough Case Investigations Team, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 792–793 (2021).
26. N. Brusselaers, J. Lagergren, *Methods Inf. Med.* **56**, 401–406 (2017).
27. Y. Young-Xu *et al.*, *JAMA Netw. Open* **4**, e2128391 (2021).
28. Y. Wang *et al.*, *EclinicalMedicine* **40**, 101129 (2021).
29. K. Bruxvoort *et al.*, *medRxiv* 21264199 [Preprint] (2021).
30. A. Purnik *et al.*, *medRxiv* 21261707 [Preprint] (2021).
31. S. Y. Tartof *et al.*, *Lancet* **398**, 1407–1416 (2021).
32. H. Chemaitelly *et al.*, *N. Engl. J. Med.* **10.1056/NEJMoa2114114** (2021).
33. H. Chemaitelly *et al.*, *Nat. Med.* **27**, 1614–1621 (2021).
34. T. Tada *et al.*, *bioRxiv* 452771 [Preprint] (2021).
35. R. R. Goel *et al.*, *Science* **10.1126/science.abm0829** (2021).
36. K. B. Pouwels *et al.*, *Nat. Med.* **10.1038/s41591-021-01548-7** (2021).
37. R. M. Anderson *et al.*, *Interface Focus* **11**, 20210008 (2021).
38. D. W. Eyre *et al.*, *medRxiv* 21264260 [Preprint] (2021).
39. M. J. M. Niesen *et al.*, *medRxiv* 21259833 [Preprint] (2021).
40. C. H. Sudre *et al.*, *Nat. Med.* **27**, 626–631 (2021).
41. B. van den Borst, *Lancet Reg. Health West. Pac.* **12**, 100208 (2021).
42. J. L. Hirschtick *et al.*, *Clin. Infect. Dis.* **10.1093/cid/ciab408** (2021).
43. Pfizer, "Pfizer and BioNTech Announce phase 3 trial data showing high efficacy of a booster dose of their COVID-19 vaccine," press release 21 October 2021; [www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-phase-3-trial-data-showing](https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-phase-3-trial-data-showing).
44. L. Saad, "Americans Getting Out More, but Cautiously," *Gallup* 7 June 2021; <https://news.gallup.com/poll/350666/americans-getting-cautiously.aspx>.
45. K. L. Bajema *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1294–1299 (2021).
46. P. Cirillo, N. Krigbaum, Code for methods for SARS-CoV-2 vaccine protection and deaths among U.S. veterans during 2021. *Zenodo* (2021); doi:10.5281/zenodo.5609444.

## ACKNOWLEDGMENTS

We acknowledge the invaluable efforts of the Veterans Affairs data architects, managers, and clinicians who assembled the Centralized Interactive Phenomics Resource (CIPHER), rapidly compiling a library of numerous COVID-19-related phenotypes that are the basis for this research. Our work was supported by using resources and facilities of the Department of Veterans Affairs (VA) Informatics and Computing Infrastructure (VINCI). VA HSR RES 14-457. We deeply appreciate the steady service and support of the VA Informatics and Computing Infrastructure (VINCI) staff. Without the efforts of these teams, this study would not have been possible. We are grateful for the veterans

who have so selflessly served their country. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government. Institutional Review Board (IRB): This research is covered under the University of California, San Francisco IRB 10-03609 Reference 320151. **Funding:** This work was supported by the Mercatus Center at George Mason University (Fast Grants 2207) and the University of California Office of the President (Emergency COVID-19 Research Seed Funding R00RG3118). **Author contributions:** Conceptualization: B.A.C. and P.M.C. Methodology: P.M.C., C.C.M., and B.A.C. Statistical analysis: P.M.C. Funding acquisition: A.W.W., P.M.C., and N.Y.K. Data interpretation: all authors. Writing, original draft: C.C.M., P.M.C., and B.A.C. Writing, review and editing: all authors. **Competing interests:** C.C.M. reports consulting for Freenome. A.W.W. reports consulting for ECOM Medical, Obelab, Sensifree, and Shifamed. B.A.C., P.M.C., and N.Y.K. declare that they have no competing interests. **Data and materials availability:** Data and materials availability: The data that support the findings of this study are available from the Department of Veterans Affairs (VA). Code is available at (46). Data are made freely available to researchers behind the VA firewall with an approved study protocol. Summary data can be accessed from a commercial

source Data Lake Analysis for Real-World Evidence Solutions—STATinMED: <https://statinmed.com/data>. More information is available at <https://www.virec.research.va.gov> or by contacting the VA Information Resource Center at [VIREC@va.gov](mailto:VIREC@va.gov). 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.abm0620](https://science.org/doi/10.1126/science.abm0620)  
Material and Methods

Fig. S1

Tables S1 to S3

MDAR Reproducibility Checklist

24 August 2021; accepted 2 November 2021

Published online 4 November 2021

10.1126/science.abm0620

## CORAL REEFS

# Protecting connectivity promotes successful biodiversity and fisheries conservation

Luisa Fontoura<sup>1</sup>, Stephanie D'Agata<sup>1,2,3</sup>, Majambo Gamoyo<sup>4</sup>, Diego R. Barneche<sup>5,6</sup>, Osmar J. Luiz<sup>7</sup>, Elizabeth M. P. Madin<sup>8</sup>, Linda Eggertsen<sup>9</sup>, Joseph M. Maina<sup>1,10\*</sup>

The global decline of coral reefs has led to calls for strategies that reconcile biodiversity conservation and fisheries benefits. Still, considerable gaps in our understanding of the spatial ecology of ecosystem services remain. We combined spatial information on larval dispersal networks and estimates of human pressure to test the importance of connectivity for ecosystem service provision. We found that reefs receiving larvae from highly connected dispersal corridors were associated with high fish species richness. Generally, larval "sinks" contained twice as much fish biomass as "sources" and exhibited greater resilience to human pressure when protected. Despite their potential to support biodiversity persistence and sustainable fisheries, up to 70% of important dispersal corridors, sinks, and source reefs remain unprotected, emphasizing the need for increased protection of networks of well-connected reefs.

**E**cological networks of larval dispersal support the long-term resilience of marine assemblages through population replenishment and gene flow (1, 2). The spatially asymmetric nature of larval dispersal driven by species-specific life history traits and oceanographic conditions shapes coral reef connectivity patterns (3). Reefs acting as

"sources" of fish larval export can help stabilize and restore fisheries in connected "sinks" (4). Dispersal corridors connect populations between sources and sinks, thus promoting gene flow and supporting biodiversity persistence (5, 6). Discerning functionally important connectivity attributes on coral reefs is vital for maximizing biodiversity and fisheries benefits that largely contribute to the well-being of human populations (7, 8). We address three fundamental gaps concerning protection of ecological connectivity on coral reefs: (i) the relative importance of distinct larval connectivity attributes in supporting reef fish species richness (biodiversity persistence) and biomass (sustainable fisheries); (ii) fish community responses along gradients of larval connectivity, human pressure, and fisheries management; and (iii) the state of connectivity conservation for coral reefs.

We applied a Bayesian hierarchical modeling framework to test the association between fish larval connectivity and ecosystem services provision, quantified with fish species richness and fish standing biomass across a gradient

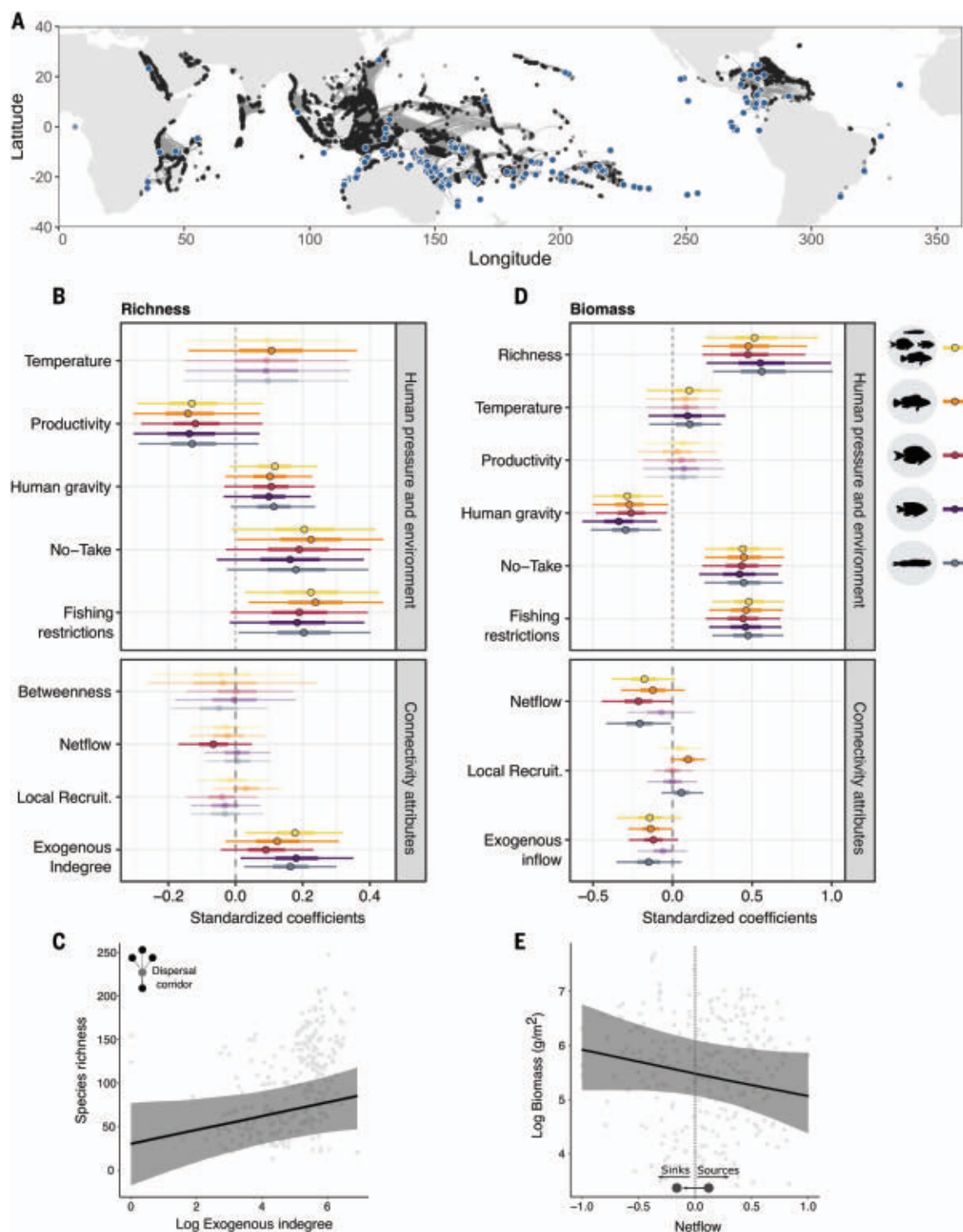
<sup>1</sup>Department of Earth and Environmental Sciences, Macquarie University, Sydney, NSW 2109, Australia. <sup>2</sup>Marine Programs, Wildlife Conservation Society, Bronx, NY, USA.

<sup>3</sup>ENTROPIE (IRD, University of La Reunion, CNRS, University of New Caledonia, Ifremer), 97400 Saint-Denis, La Reunion c/o IUEM, 29280 Plouzané, France. <sup>4</sup>Coastal and Marine Resources Development, Mombasa, Kenya. <sup>5</sup>Australian Institute of Marine Science, Crawley, WA 6009, Australia.

<sup>6</sup>Oceans Institute, The University of Western Australia, Crawley, WA 6009, Australia. <sup>7</sup>Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, NT, Australia. <sup>8</sup>Hawai'i Institute of Marine Biology, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Kane'ohe, HI 96744, USA.

<sup>9</sup>Department of Earth Sciences, Uppsala University, SE-621 67 Visby, Sweden. <sup>10</sup>Centre for Environmental Law, Macquarie University, Sydney, NSW 2019, Australia.

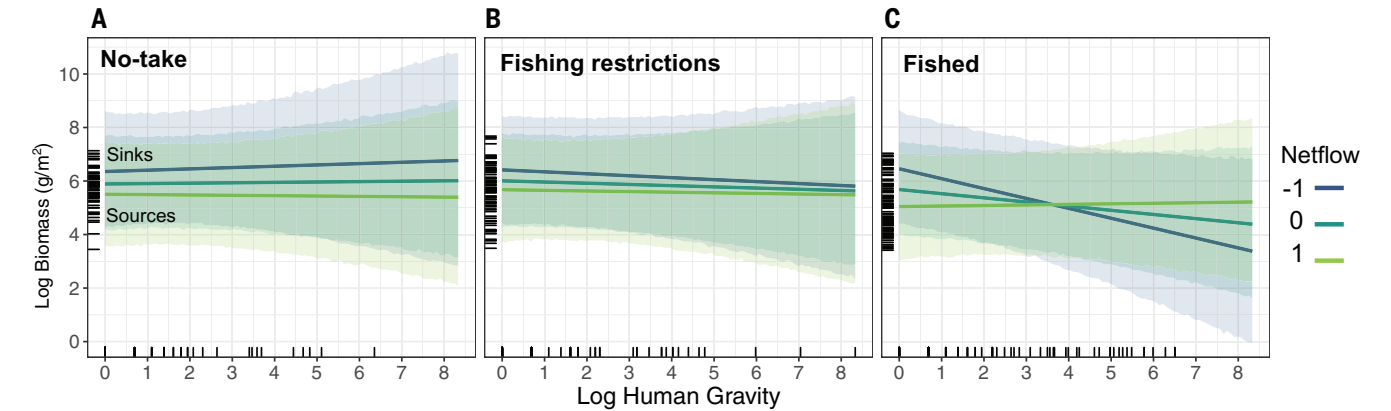
\*Corresponding author. Email: [joseph.mbuim@mq.edu.au](mailto:joseph.mbuim@mq.edu.au)



**Fig. 1. Global patterns of fish larval connectivity on coral reefs and association with total fish biomass and species richness.** (A) Fish larval dispersal simulated for coral reefs globally with ocean data from 2000 to 2005. Connectivity attributes and human pressure were estimated for 272 reef sites (blue points) to determine the role of connectivity in promoting fish biomass and species richness. (B) Bayesian models with fish species richness and (D) biomass as response variables, and the estimated standardized effect sizes for human, environmental, and connectivity

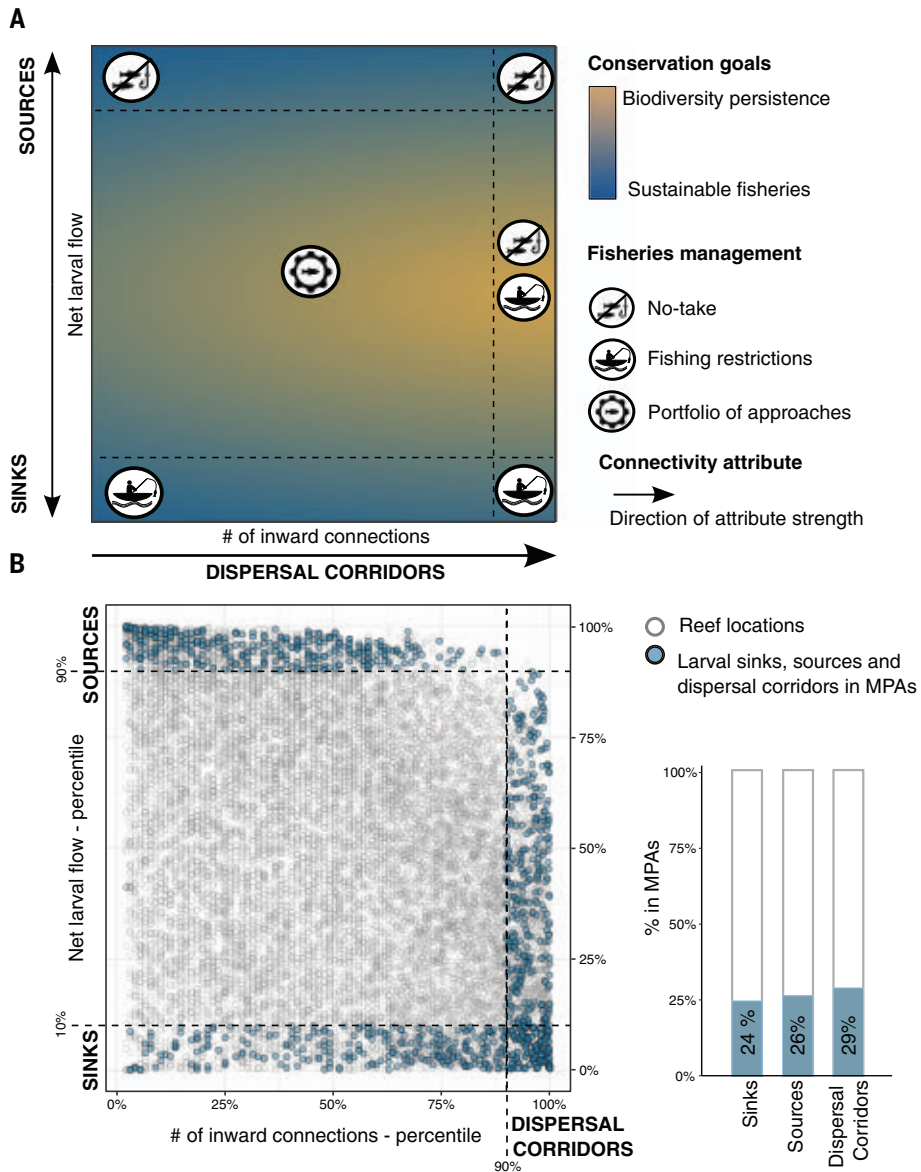
covariates. Horizontal error bars represent 95% Bayesian credible intervals, with the inner, thicker lines defining the 50% credible intervals. In (B) and (D), the y axes are the model covariates of total connectivity—average connectivity across the four fish groups—and connectivity based on individual fish groups. Model-predicted relationships between the number of inward connections of dispersal corridors (C) and species richness and net larval flow and fish biomass (E). Models were fitted with both random slopes and intercepts across provinces (fig. S4).

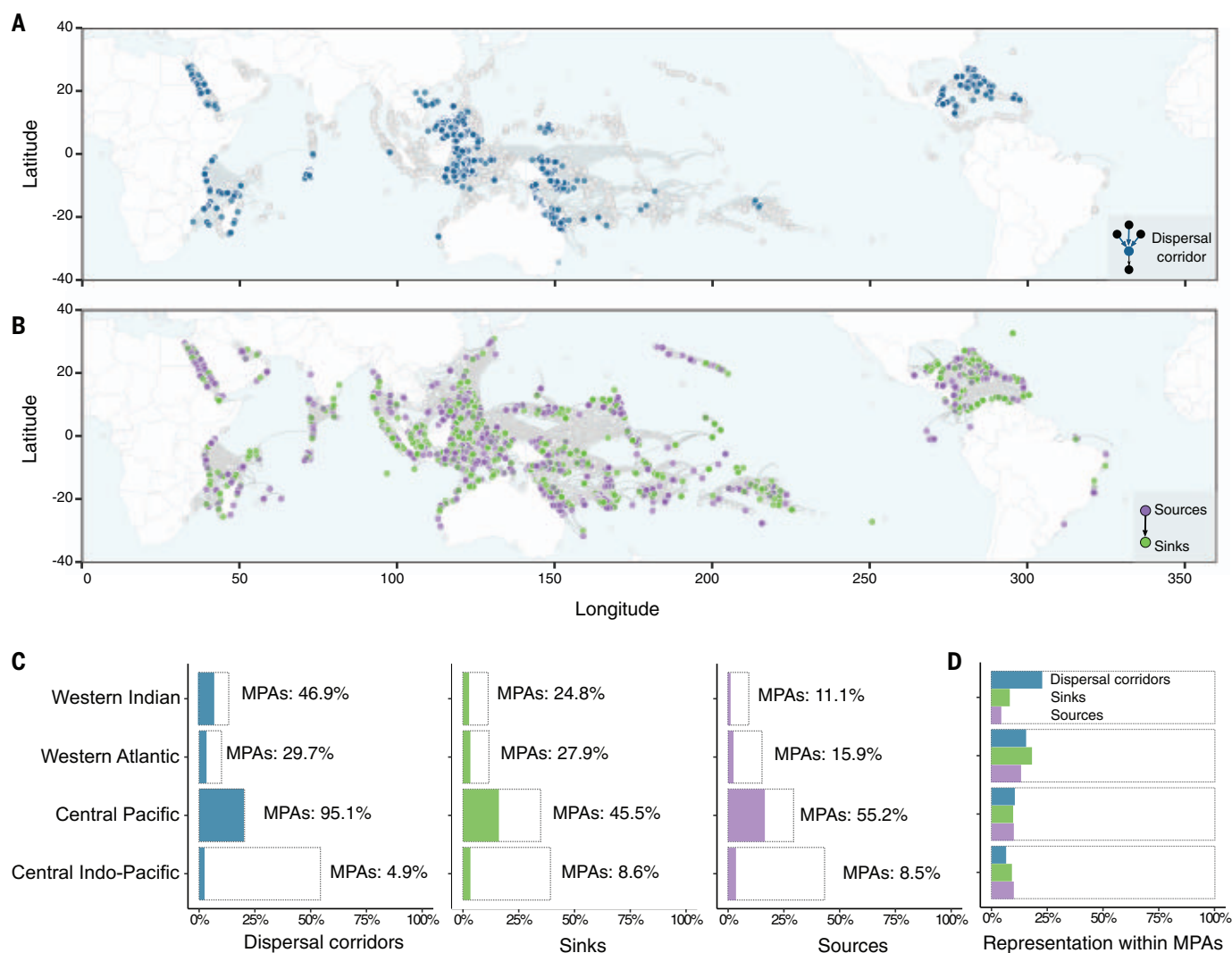




**Fig. 2. Model-predicted relationships between fish biomass and human pressure on reefs along the net larval flow gradient in three management categories.** Management levels represent (A) no-take marine reserves, (B) areas with fishing restrictions (e.g., marine parks), and (C) fished areas. Lines represent three levels of net larval flow (netflow): -1 (absolute larval sinks), 0, and 1 (absolute larval sources). Uncertainty bands represent 95% confidence intervals.

**Fig. 3. A conceptual framework for applying connectivity to optimal placement of MPAs and OECMs for biodiversity and fisheries management and the global status of connectivity conservation on coral reefs.** (A) Illustration of sweet spot locations for no-take and restricted fishing zones. The color gradient illustrates the potential of sources, sinks, and dispersal corridors to support sustainable fisheries and biodiversity persistence relative to the strength of their connectivity attributes. Dispersal corridors, when protected, can promote biodiversity persistence through gene flow and population resilience. No-take source areas support sustainable fisheries in sink areas through fish larval export. Fishing restrictions in sink areas can sustain fisheries benefits through larval subsidies from source areas. Dispersal corridors that function as strong sinks can be fished with restrictions to support local fisheries. In all other reef areas, a portfolio of fisheries management approaches can be applied. (B) Reef locations ( $n = 14,804$ ) within the 90th and 10th percentile of net larval flow are considered critical larval sources and sinks, respectively. Those within the 90th percentile along the inward connections gradient are regarded as essential dispersal corridors. The proportion of currently protected sources, sinks, and dispersal corridors are represented by colored bars.





**Fig. 4. Geographical representation of critical dispersal corridors, larval sources and sinks, and their conservation status across four biogeographical regions.** Darker points represent critical dispersal corridors (A) in addition to sources and sinks (B), as defined in Fig. 3. (C) Dotted lines indicate the

percentage of functionally important reefs across the four biogeographical regions and colored bars indicate the percentage of these reef cells within MPAs. (D) Bars indicate the representation of critical dispersal corridors, sinks, and sources within regional MPA networks.

of human pressure and fisheries restrictions (9). To account for variation in species' reproductive and larval traits (10), we calibrated larval dispersal models with biological parameters describing four fish groups with different ecological roles (fig. S1). We estimated a suite of connectivity attributes for each and across fish groups. We collated five socioenvironmental factors associated with 272 coral reefs as predictors of total fish biomass and species richness (Fig. 1A). Connectivity attributes describe the relative probability of reefs to export, receive, and retain larval subsidies. Endogenous connectivity attributes are based on reefs' direct connections, whereas exogenous connectivity attributes are based on reefs' indirect connections (table S1).

Biogeographic patterns of reef fish biodiversity are partly shaped by reef connectivity (11). Likewise, we found that higher fish species

richness was associated with highly connected dispersal corridors, particularly of small-bodied species with short pelagic larval durations (Fig. 1, B and D, and table S3). Small-bodied reef fishes contribute disproportionately to coral reef species richness relative to larger fish species (12). Thus, protecting dispersal corridors that are functionally important in maintaining larval connectivity—particularly connectivity of small-bodied fish populations—is likely to disproportionately benefit biodiversity conservation. Notably, reefs with several incoming connections are embedded in a complex network of well-connected reefs through larval dispersal (fig. S2). This emphasizes the need to identify and protect exogenous connections.

In addition to the known effects of species richness, temperature, and human pressure on fish biomass (13, 14), we found that connectivity was influential, as suggested by recent

evidence (15) (Fig. 1C). Adding connectivity attributes as covariates in the hierarchical model increased the explained variance from 33 to ~51% and the model's predictive accuracy (table S2). Overall, fish biomass was higher for reefs with a greater probability of accumulating larval subsidies from adjacent, connected reefs. Net larval flow (i.e., netflow), defined as the gradient between absolute larval sinks and absolute larval sources, was associated more strongly with fish biomass (table S4). Fish biomass in absolute sink reefs was approximately twice as high as in absolute source reefs (Fig. 1E). This finding over such a large spatial scale corroborates the long-held understanding that the accumulation of larvae subsidies favors fish population replenishment and long-term resilience of sink locations (2). By contrast, reefs with the greatest potential for exporting larvae may be more sensitive to fishing pressure and

require higher fishing restrictions to maintain high biomass and support sustainable fisheries in connected sink reefs (16).

The positive association between larval sinks and fish biomass was more evident when accounting for connectivity patterns of species that reproduce year-round, namely cryptobenthic fish and resident spawners (Fig. 1C). Resident spawners often include species targeted by fishing (e.g., surgeonfish, small snappers). By contrast, short-lifespan cryptobenthic fish are not a target but constitute up to 60% of biomass consumed by piscivorous fish (17). Therefore, our results suggest that constant larval inflow at sink locations may support fishery benefits by promoting year-round population replenishment of fisheries-targeted species and key piscivorous prey species.

The association between fish biomass and human pressure varied with the sink-source gradient and management categories (Fig. 2 and fig. S3). Fish biomass was relatively lower in larval sources than sinks in no-take and restricted reefs (Fig. 2, A to B). However, fished larval sinks demonstrated higher sensitivity to human pressure above an apparent human pressure threshold (Fig. 2C). Unsustainable harvest and higher fishing pressure on sink reefs undermines the potentially positive impacts of larvae inflow on fish biomass (18). Therefore, managing fisheries (e.g., area- or gear-based regulations) in larval sinks may facilitate the persistence of fish biomass and provide ecosystem goods and services for human coastal populations that depend on local fisheries (19). The extent to which larval sinks can contribute to local food security may also depend on the management status of connected reefs that serve as their larval sources (16). These contrasting associations between fish biomass and human pressure under different connectivity and fisheries management scenarios underscore the importance of assessing a reef's inherent connectivity attributes and the local socioecological context.

Linking distinct connectivity attributes with marine protected area (MPA) goals is critical for making informed management decisions, particularly for coral reefs where sustainability goals of biodiversity conservation and fisheries sustainability compete (7, 20). Findings on distinct yet complementary roles of sinks, sources, and dispersal corridors in predicting

species richness and fish biomass can inform placement of MPAs and other effective area-based conservation measures (OECMs) for optimizing biodiversity persistence and fisheries benefits (Fig. 3A).

Despite the expansion of MPAs over the past decade (21), we found considerable shortfalls in implementing connectivity conservation and poor placement of MPAs. Approximately 70% of the most critical dispersal corridors, larval sources, and sinks are unprotected (Fig. 3B and table S5). Furthermore, we found low representation of these functionally important reefs within the current spatial arrangement of the MPAs (~11% globally; table S5). Globally, 29% of dispersal corridors, 26% of larval sinks, and 24% of sources are currently within MPAs, but large disparities exist between biogeographical regions (Fig. 4). Conservation efforts to protect connectivity were lowest in the Indo-Pacific region, which has the largest proportion of functionally important reefs. In this global biodiversity hotspot where more than 40% of human populations depend on local fisheries (22), only 5 to 8.5% of key dispersal corridors, larval sinks, and sources are currently protected (Fig. 4C). Implementing connectivity conservation in these regions may have a disproportionately large, positive effect on the persistence of biodiversity and ecosystem services.

Despite limitations in biophysical models at large spatial scales (9), we found that connectivity attributes differed in their relative roles and importance for biodiversity maintenance and fisheries. Well-connected dispersal corridors were associated with species richness, whereas source-sink systems were more strongly associated with fish biomass. Given that ~70% of functionally important coral reefs are currently unprotected globally, these gaps highlight opportunities for implementation of connectivity conservation by strategic placement of MPAs and OECMs as part of the expansion proposed by the post-2020 biodiversity conservation policy (23).

## REFERENCES AND NOTES

1. S. Planes, G. P. Jones, S. R. Thorrold, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5693–5697 (2009).
2. H. B. Harrison, M. Bode, D. H. Williamson, M. L. Berumen, G. P. Jones, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 25595–25600 (2020).
3. G. R. Almany *et al.*, *Nat. Ecol. Evol.* **0148**, (2017).
4. H. B. Harrison *et al.*, *Curr. Biol.* **22**, 1023–1028 (2012).
5. E. Boulanger, A. Dalongeville, M. Andreello, D. Mouillot, S. Manel, *Ecography* **43**, 1167–1179 (2020).
6. R. A. Magris *et al.*, *Conserv. Lett.* **11**, e12439 (2018).
7. S. D. Gaines, C. White, M. H. Carr, S. R. Palumbi, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18286–18293 (2010).
8. A. J. Woodhead, C. C. Hicks, A. V. Norström, G. J. Williams, N. A. Graham, *Funct. Ecol.* **33**, 1023–1034 (2019).
9. Materials and methods are available in the supplementary materials.
10. E. A. Trembl *et al.*, *Integr. Comp. Biol.* **52**, 525–537 (2012).
11. V. Parravicini *et al.*, *Ecography* **36**, 1254–1262 (2013).
12. D. R. Barneche *et al.*, *Glob. Ecol. Biogeogr.* **28**, 315–327 (2019).
13. J. E. Cinner *et al.*, *Proc. Natl. Acad. Sci.* **115**, 6116–6125 (2018).
14. J. E. Duffy, J. S. Lefcheck, R. D. Stuart-Smith, S. A. Navarrete, G. J. Edgar, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6230–6235 (2016).
15. J. S. Goetze *et al.*, *Glob. Change Biol.* **27**, 3432–3447 (2021).
16. D. J. Marshall, S. Gaines, R. Warner, D. R. Barneche, M. Bode, *Front. Ecol. Environ.* **17**, 407–413 (2019).
17. S. J. Brandt *et al.*, *Science* **364**, 1189–1192 (2019).
18. C. Barceló, J. W. White, L. W. Botsford, A. Hastings, *ICES J. Mar. Sci.* **78**, 1860–1871 (2021).
19. A. Le Port *et al.*, *Proc. Biol. Sci.* **284**, 20171300 (2017).
20. J. Hilty *et al.*, “Guidelines for conserving connectivity through ecological networks and corridors,” (no. 30 in Best Practice Protected Area Guidelines Series, IUCN, 2020); <https://doi.org/10.2305/IUCN.CH.2020.PAG.30.en>.
21. S. L. Maxwell *et al.*, *Nature* **586**, 217–227 (2020).
22. FAO, Fishery and Aquaculture Country Profiles: Indonesia (2011); [www.fao.org/fishery/en/facp/idn?lang=en](http://www.fao.org/fishery/en/facp/idn?lang=en).
23. Convention of Biological Diversity, “Zero Draft of the Post-2020 Global biodiversity framework” (2021); [www.cbd.int/conferences/post2020/post2020-prep-01/documents](http://www.cbd.int/conferences/post2020/post2020-prep-01/documents).
24. L. Fontoura *et al.*, *Zenodo* (2021); doi: 10.5281/zenodo.5762799.

## ACKNOWLEDGMENTS

We thank M. Bode, E. Asamoah, and four anonymous reviewers for their valuable comments. **Funding:** L.F. was supported by an Australian Government Research Training Program (RTP) Scholarship (2017/002). Funding for S.D. was provided by the Laboratory of Excellence “Corail” (LIVELIHOOD project, grant EPHE IRD PD A02020), France. **Author contributions:** L.F. developed the concept and hypotheses of the study with J.M. and S.D.; M.G. and J.M. conducted biophysical modeling; L.F. defined the biological model parametrization and conducted network analysis; J.M., D.R.B., and S.D. implemented the statistical analyses. L.F. led the manuscript with J.M. and S.D. All the authors contributed equally to reviewing the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data and code for analysis reproducibility are available at Zenodo (24).

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abg4351](https://science.org/doi/10.1126/science.abg4351)  
Materials and Methods  
Figs. S1 to S5  
Tables S1 to S6  
References (25–66)  
MDAR Reproducibility Checklist  
Data S1

7 January 2021; resubmitted 26 July 2021  
Accepted 6 December 2021  
10.1126/science.abg4351





## EMPOWER WITH EVIDENCE

AAAS | ANNUAL MEETING

AAAS, publisher of **Science** journals, thanks sponsors  
and supporters of the 2022 AAAS Annual Meeting:



### BE A PART OF THE CONVERSATION



and more...

### Important Meeting Update!

The 2022 AAAS Annual Meeting will gather online **ONLY**, February 17-20.  
Registration and Program are now available!

For more information, visit [aaas.org/meetings](https://aaas.org/meetings) and join the conversation online  
with our official hashtag: **#AAASmtg**



## Dear Colleague,

On behalf of our colleagues on the AAAS Board of Directors, and the entire leadership and staff of AAAS, we wish you a happy, productive, and, above all, healthy new year.

As always, it is a busy time of year for AAAS, with the next six weeks focused on bringing you the thought provoking and valuable scientific sessions, workshops, networking opportunities, and more at the AAAS Annual Meeting in February. Throughout our planning, we have been closely monitoring the developments in the pandemic and, while we were once quite optimistic about the prospects of gathering safely in-person, the conditions of the past several weeks have required us to reevaluate. As such, we have made the difficult decision to forego the in-person component of the AAAS Annual Meeting in Philadelphia and convert to an entirely virtual convening.

This is certainly not the path we would have liked. For those of you who were planning to join us in Philadelphia, we were excited to once again experience the camaraderie that can only come from being together. However, the health and safety of our attendees and staff is our top priority, and, under the decision-making framework that we established at the beginning of the pandemic, we could not reasonably guarantee that health and safety in a way that did not compromise the continuity of operations for a successful in-person meeting.

We will be in touch soon with more detailed information, including refunds for in-person registrations. In the meantime, we ask that you hold onto any specific questions you might have until our teams can finalize the many details involved in this transition.

Advance (discounted) registration for the virtual meeting is available through January 24. We hope you will join us for the compelling content and powerful professional development in our new-and-improved virtual meeting space, February 17-20. The virtual platform, including Scientific Session Spotlight Videos, will be available beginning in late January.

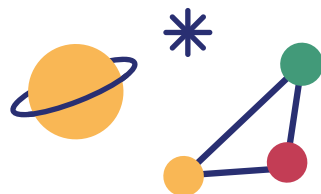
We are grateful to our partners in Philadelphia—including and especially Drexel University, our official Host University—who have been wonderful during the multi-year planning process for this meeting. We look forward to featuring many unique and exciting elements of Philadelphia in our virtual space, and thank the entire community for their support.

We hope that each of you stays safe and healthy during these times, and we look forward to seeing you online in the weeks ahead.

Sincerely,

**Susan G. Amara, Ph.D.**  
President, AAAS  
Chair, 2022 Annual Meeting

**Sudip S. Parikh, Ph.D.**  
Chief Executive Officer, AAAS  
Executive Publisher, *Science Journals*



# Join us Online for the 2022 AAAS Annual Meeting

Before the live event in February, be on the lookout for Spotlight Videos, highlighting the work of individual panelists

## Choose your own access...

### Deluxe Access registrants can:

- Join **Scientific Sessions** online and have access to the library of **Spotlight Videos**
- Take part live in lectures **from renowned speakers** in academia, industry and policy
- Enjoy STEM documentaries
- Participate virtually with features included in the Basic Access registration

Deluxe Access includes scheduled live events *and* on-demand viewing of Scientific Sessions and lectures at your convenience.

### Basic Access registrants can:

- **Attend live workshops** and learn new ways to forward your research trajectory
- Wander the virtual halls of the **E-poster Session** and talk to finalists about their work during special sessions
- Tune into **other special sessions**

Basic Access will *only* have access to scheduled live events.  
Interested in on-demand viewing? **Upgrade to Deluxe Access!**

Look on last page of this spread for registration pricing details.

## CONNECT WITH US! #AAASmtg



aaas.org/meetings



@AAASmeetings



/AAAS.Science







# Plenary Sessions



## AAAS President Address

**Susan Amara**

President, AAAS  
Chair, 2022 Annual Meeting



## Addressing Energy Access, Equity and Resiliency in the 21st Century

**Deepakraj M. Divan**

Director, Center for Distributed Energy, Georgia Tech



**Katalin Karikó**

Senior Vice President, BioNTech RNA Pharmaceuticals  
Adjunct Professor, University of Pennsylvania School of Medicine



## “Does Science Communication Still Work?”

A debate moderated by **Holden Thorp**

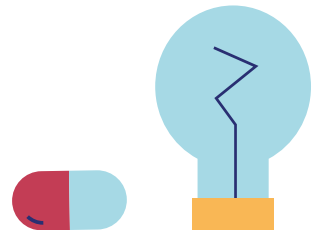
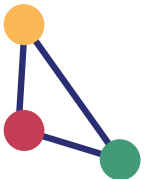
Editor-in-Chief of **Science** Journals

## Featuring panelists:

**Kathleen Hall Jamieson**  
University of Pennsylvania

**Katie Mack**  
North Carolina State University

**Joelle Simpson**  
Children's National Hospital



# Topical Lectures



## Bioinspired Ocean Exploration

**John O. Dabiri**

California Institute of Technology



## From the Possibility to the Certainty of a Supermassive Black Hole

**Andrea M. Ghez**

University of California, Los Angeles



## Games, Stories, and Science for Successful Public Engagement

**Katie Hinde**

Arizona State University



## Rising to Resilience in Water: Can We Get There from Here, and What Will It Take?

**Felicia Marcus**

Stanford University



## The Disappearing Chip: The Physics behind Nanoscale Electronics

**Nadya Mason**

University of Illinois at Urbana-Champaign



## A Care Conundrum at the Intersection of Class, Access, and Health

**Robin Nelson**

Arizona State University



## Unraveling the Big Lie: Participatory Disinformation and its Threat to Democracy

**Kate Starbird**

Human Centered Design and Engineering



## Against Crisis Science: Research Futures for Climate and Energy Justice

**Kyle Whyte**

University of Michigan



## Eyewitness Memory Is Reliable, But the Criminal Justice System Is Not

JOHN P. MCGOVERN AWARD  
LECTURE IN BEHAVIORAL  
SCIENCES

**John Wixted**

University of California  
San Diego

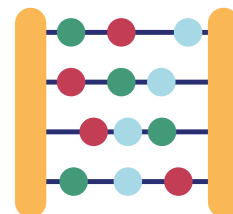


## Evidentiary Constraint: When Science Policy Meets Activism

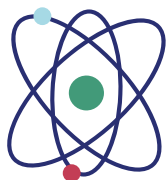
SARTON MEMORIAL LECTURE  
IN THE HISTORY AND  
PHILOSOPHY OF SCIENCE

**Audra J. Wolfe**

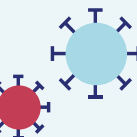
Writer and Editor



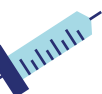
*As of January 5, 2022*



# Join us Online for the 2022 AAAS Annual Meeting



The 2022 AAAS Annual Meeting will Empower With Evidence by featuring groundbreaking multi-disciplinary research—not only to advance knowledge, but to understand, and, ultimately, resolve many of the problems we face today.



## CONNECT WITH US! #AAASmtg



[aaas.org/meetings](https://aaas.org/meetings)



@AAASmeetings



/AAAS.Science

### REGISTRATION RATES

General Attendee

\$240

Postdoc

\$155

Retired Professional

\$240

Student

\$75

### DELUXE ACCESS

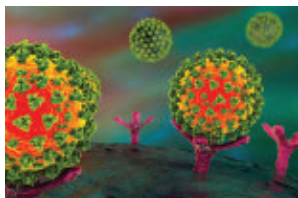
The Basic Access registration is available free of charge.

The 2022 AAAS Annual Meeting will take place online, February 17-20. The virtual meeting platform will be available in late January for streaming Spotlight Videos and more!

Stay up-to date by visiting [aaas.org/meetings](https://aaas.org/meetings) and join the conversation online with our hashtag: **#AAASmtg**







### COVID-19 Omicron Variant Research Tools

AMS Biotechnology (AMS BIO) announces new additions to its portfolio of tools for COVID-19 research. AMS BIO's Spike-neutralizing monoclonal antibodies (clone G10xA5 and G10xA1)

recognize the SARS-CoV-2 Spike receptor-binding domain (RBD) (B.1.1.529, Omicron variant) protein and neutralize its interaction with angiotensin-converting enzyme 2 (ACE2). Also new is AMS BIO's Spike S1 RBD: ACE2 Inhibitor Screening Assay Kit. This key colorimetric assay kit is designed for screening and profiling inhibitors or neutralizing antibodies of the interaction between the Omicron variant SARS-CoV-2 Spike RBD and human ACE2. AMS BIO also offers a series of recombinant antigens and pseudotyped lentivirus with key Spike mutations in order to evaluate the efficacy of antibodies and vaccines.

#### AMS Biotechnology

For info: +1-617-945-5033

[www.amsbio.com/sars-cov-2-spike-mutants](http://www.amsbio.com/sars-cov-2-spike-mutants)

### SARS-CoV-2 Variant Neutralization Antibody Assays

Bio-Rad Laboratories has released Bio-Plex Pro Human SARS-CoV-2 Variant Neutralization Antibody Assays, for research use only (RUO). The assays allow scientists to measure neutralizing antibodies quickly and efficiently against wild-type and significant variants of SARS-CoV-2 virus. Scientists developing vaccines and therapeutics can determine the efficacy of their product from development through all clinical phases. Epidemiologists can perform seroprevalence studies to better understand the longevity of the neutralizing antibody response from infection or vaccination. Researchers can increase their throughput when using the multiplex Bio-Plex Pro Assays, as compared to traditional cell-based assays and ELISA tests. The assays are offered in an 11-plex complete all-in-one kit format consisting of the wild-type RBD and S1 and 9 variant antigens, as well as customizable singleplex assays, 2-plex Delta variant coupled beads, and a custom Assay Developer Kit to create assays to any new SARS-CoV-2 variant—such as Omicron—for maximum flexibility.

#### Bio-Rad Laboratories

For info: +1-800-424-6723

[bio-rad.com/nabsarscov2](http://bio-rad.com/nabsarscov2)

### Fluorescent-Labeled Spike Proteins

Fluorokines fluorescent-labeled Spike proteins from R&D Systems feature Alexa Fluor dyes, which easily detect ACE2 expression on cells without requiring a secondary detection step. R&D Systems Fluorokines are designed to simplify detection of a specific chimeric antigen receptor (CAR) on CAR T cells or other cell types of interest. The fluorophores are conjugated to the bioactive protein via amine labeling, and the resulting Fluorokines are rigorously tested to ensure consistent labeling of each lot. The entire manufacturing process is controlled to reduce lot-to-lot variability and delivers a consistent fluorochrome-to-protein (F/P) ratio. Finally, each protein is quality-control tested by flow cytometry.

#### R&D Systems

For info: +1-800-343-7475

[www.rndsystems.com](http://www.rndsystems.com)

### T-Cell Tests

Hyris partnered with prestigious academic and research teams to develop rapid T-cell tests to track patient immunity levels to specific infections. This kit enables swift evaluation of T-cell immune responses in COVID-19-inoculated or convalescent patients, as well as measuring individual immunity levels for vaccinated people. The solution leverages the portability and reliability of the Hyris System, allowing deployment of T-cell testing anywhere, any time, with real-time results available immediately through Hyris dedicated software. The system deploys reverse transcription PCR technology, the “gold standard” according to the U.S. Centers for Disease Control and Prevention and the World Health Organization's effective testing guidelines. At the core of the system is the HYRIS bCUBE, a miniaturized, state-of-the-art device for nucleic acid analysis. With ultracompact instrument design and superior analytical performance, the Hyris bCUBE has direct connectivity to the proprietary HYRIS bAPP.

#### Hyris

For info: +44-(0)-2039361227

[www.hyris.net/index.php/t-cell](http://www.hyris.net/index.php/t-cell)

### ELISA Kits for SARS-CoV-2

Proteintech offers well-validated, highly sensitive indirect and sandwich ELISA kits for detecting immunoglobulins against SARS-CoV-2 proteins and the cytokines involved in cytokine release syndrome. These kits are developed and validated to detect immunoglobulin G and immunoglobulin M immunoglobulins against the RBD of SARS-CoV-2 Spike protein (S protein) and Nucleocapsid protein (N Protein). Proteintech's wide range of cytokine-specific ELISA kits are extensively validated using in-house developed, human cell-expressed cytokines. Emerging evidence suggests that hyperactivity of the immune system after SARS-CoV-2 infection can trigger excessive release of cytokines such as IL-6, TNF alpha, and IFN gamma. This rapid immune escalation is called the “cytokine storm syndrome” and can lead to severe lung damage and death.

#### Proteintech

For info: +1-888-478-4522

[www.ptglab.com](http://www.ptglab.com)

### SARS-CoV-2 PCR Products and Services

PCR Biosystems is committed to providing researchers, developers, and diagnostic kit manufacturers with the PCR products and services they need in the global response to COVID-19. We have massively increased manufacture of critical reagents for rapid and sensitive reverse transcription quantitative PCR (RT-qPCR). We manufacture under a rigorous ISO 9001 and ISO 13485 quality system. To support high-throughput diagnostic workflows, we've introduced 50-mL and 500-mL bulk sizes of our RT-qPCR mixes into our standard portfolio to provide greater consistency, convenience, and value in COVID-19 testing. With over 180 years of combined PCR experience on our team, we can optimize assay conditions and increase testing turnaround times specific to your setup. Please contact us to find out more.

#### PCR Biosystems

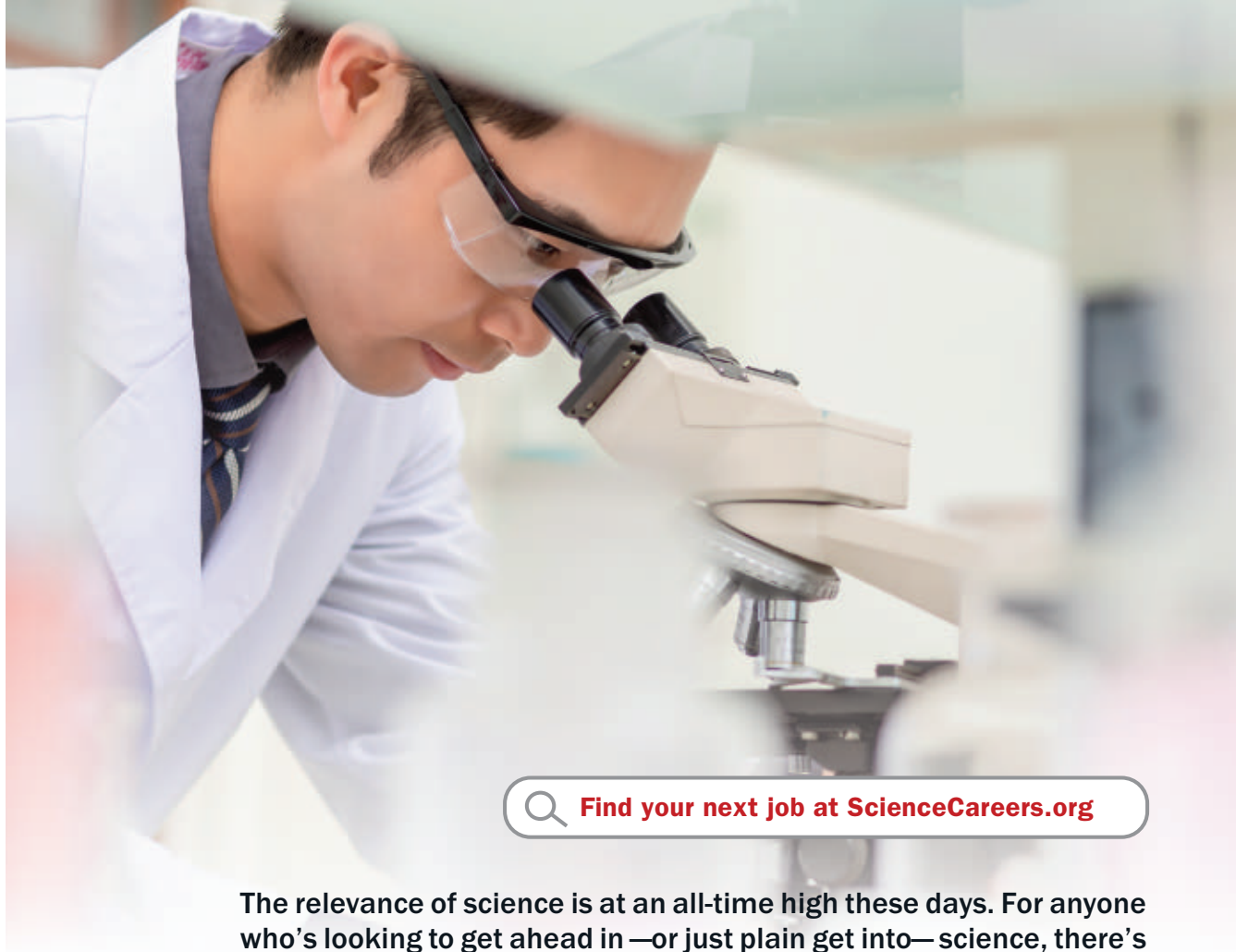
For info: +44-(0)-20-3930-8101

[pcrbio.com/usa/products/covid-19](http://pcrbio.com/usa/products/covid-19)

Electronically submit your new product description or product literature information! Go to [www.science.org/about/new-products-section](http://www.science.org/about/new-products-section) for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

# ARE YOU THE KIND WHO WORKS TO HELP ALL HUMANKIND?



Find your next job at [ScienceCareers.org](https://www.sciencecareers.org)

The relevance of science is at an all-time high these days. For anyone who's looking to get ahead in—or just plain get into—science, there's no better, more trusted resource or authority on the subject than *Science Careers*. Here you'll find opportunities and savvy advice across all disciplines and levels. There's no shortage of global problems today that science can't solve. Be part of the solution.

## ScienceCareers

FROM THE JOURNAL SCIENCE  AAAS

# ScienceCareers

FROM THE JOURNAL SCIENCE  AAAS

## Confused about your next career move?



**Download Free Career  
Advice Booklets!**

[ScienceCareers.org/booklets](http://ScienceCareers.org/booklets)





By Georgia Fousteri

## A mentor's journey

**A**t the start of our lab meeting, I did something unconventional: I read my students a poem. “As you set out for Ithaka; hope your road is a long one; full of adventure, full of discovery,” begins “Ithaka” by the Greek poet Constantine Cavafy. Then I asked what the words meant to them. We discussed how the journey in the poem is a metaphor for the human journey through life. We all may be striving to reach our own personal Ithaka, but the destination isn’t what matters most; we should take time to enjoy the journey along the way. The exercise was meant to help create a positive environment for my trainees—and it was part of my own personal journey as a mentor.

Before then, our meetings had always adhered to a more traditional format: One lab member would present a talk on their research and the rest of us would ask questions and give them feedback. The discussions were always interesting and exciting to me. But I worried the environment was intimidating to my trainees. Apart from the presenter and me, few took part in the discussion.

Years earlier, I’d been a trainee myself in a lab that held anxiety-inducing meetings. The lab head would spend the bulk of the time grilling the presenter with a series of pointed questions. Presenters would need to know minute details about articles they were citing and present their research using picture-perfect analyses and graphs.

I often learned a lot in those 2- to 3-hour meetings—how to think critically, the need to pay meticulous attention to detail. The lab head was a busy person, and that was a rare and valuable moment when we had their focused attention. But when it was my turn to present, I would sometimes suffer from sleepless nights.

I tried to create a more collegial environment when I started my own lab: I made a point of being friendly and welcoming while asking my trainees challenging questions about their science. But 6 years into my stint as lab head, I felt I hadn’t created the lab environment I wanted.

A turning point came during the pandemic. It hit home to me that my life span isn’t endless and that I need to make the most of the time I have. The isolation also gave me time for reflection. I realized that in some ways, I had reached the Ithaka I’d been striving for: a tenured position. It was time for my next big challenge. I decided that I wanted to become a better mentor—to make a mark by training the next generation of scientists.



**“My greatest wish is to see my trainees become their best selves ... and reach their career goals.”**

In my personal life, I had benefited from listening to talks by motivational speakers and self-help experts. So I decided to start there. At our next lab meeting, I showed my students a video about happiness, hoping it would help them in their personal life and give us something light to discuss. I wasn’t sure how they’d respond, but I was pleased to discover that they were interested in discussing the advice in the video afterward. At the end, I asked them whether they’d like to do similar activities in future meetings, and they said yes.

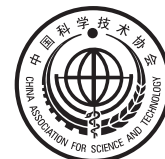
From then on, I have led a 20-minute team building session at the start of every lab meeting. My lab members and I play parts in inspirational scripts, practice meditation, or do some other activity—including discussing poetry. These sessions have eased tensions and opened minds, helping foster more participation later in the meeting.

More recently, I’ve also started to adjust how I oversee the main part of the meeting, which still revolves around a presentation by a lab member. I now ask the presenter to send me their slides a few days in advance, and I sit down with them to discuss their methods and findings. This way, I can give them constructive feedback ahead of their presentation—not in front of their peers—and leave more time in the actual meeting to discuss what their data mean.

I do not know where this journey with my students will take us. But so far, I feel hopeful about the lab environment we’re creating. My greatest wish is to see my trainees become their best selves, understand who they are and what they want, and reach their career goals. That’s a legacy I can be proud of. ■

Georgia Fousteri is a group leader at the San Raffaele Diabetes Research Institute. Send your career story to [SciCareerEditor@aaas.org](mailto:SciCareerEditor@aaas.org).

# CALL FOR PAPERS



# Research

 OPEN ACCESS

*Research* is a Science Partner Journal (SPJ) distributed by the **American Association for the Advancement of Science (AAAS)** in association with Science and Technology Review Publishing House, the publishing house under the leadership of **China Association for Science and Technology (CAST)**. *Research* provides an international platform for academic exchange, research collaboration, and technological advancements. The journal will publish fundamental research in the life and physical sciences as well as important findings or issues in engineering and applied science.

**Submit your manuscripts to *Research* today!**

Learn more: [spj.sciencemag.org/research](http://spj.sciencemag.org/research)

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](http://spj.sciencemag.org)



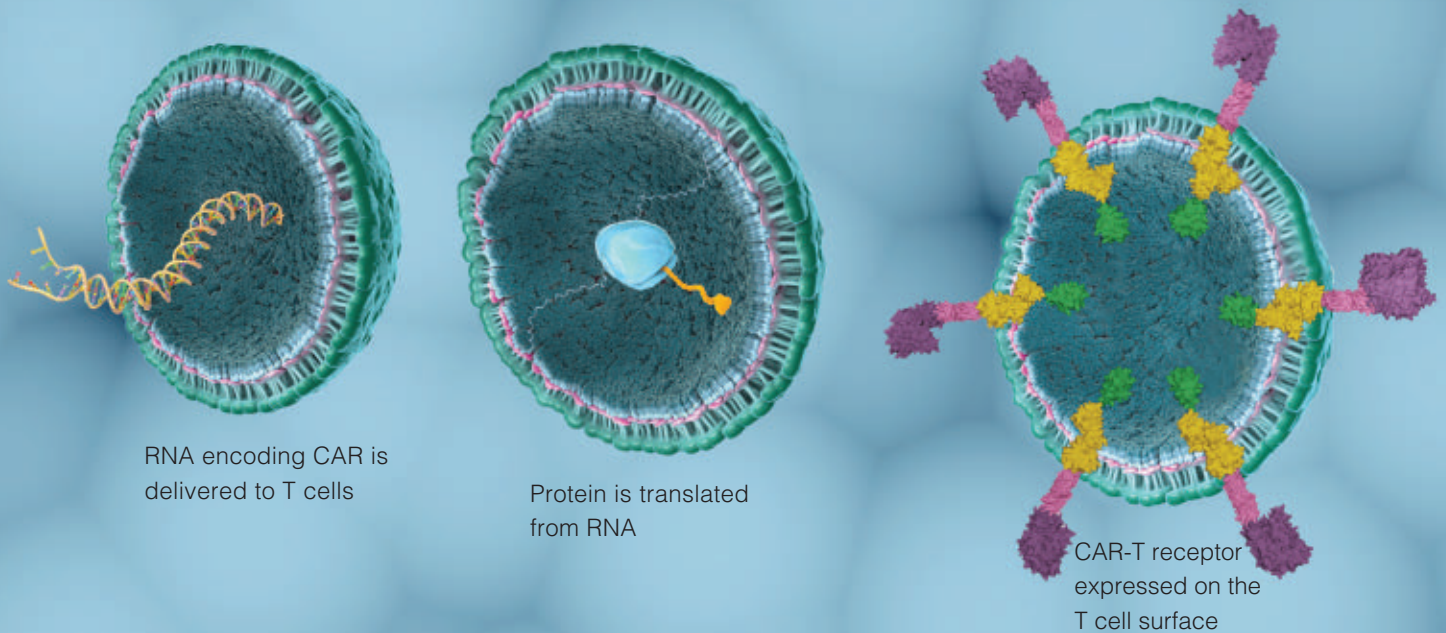
@SPJournals



@SPJournals

# CAR-T Cells Via mRNA Transfection

## Chimeric Antigen Receptors Encoded by mRNA



### CAR-T Production Services

Utilizing mRNA electroporation, rather than viral transduction, to achieve expression of the CAR has become a widely researched topic in alternative solutions to gene delivery in CAR-T cell therapy. Transient CAR-T cells offer several distinct advantages over autologous, virally transduced CAR-T cells, including decreased risk of toxicity due to a shorter half-life, and a cheaper, less labor-intensive gene transfer procedure.

At ProMab, we can synthesize any mRNA sequence and transiently express it in T cells. We provide multiple means of validation with our comprehensive cell-based assay services, such as FACS and real-time cell analysis (RTCA), with downstream options for in vivo analysis. Check out our platform to see how we can develop your CAR-T cells.

All products are for research only

Discover more | [www.promab.com](http://www.promab.com)



📍 2600 Hilltop Dr, Building B, Suite C320, Richmond, CA 94806  
☎ 1.866.339.0871 | ✉ [info@promab.com](mailto:info@promab.com)

