**The dubious ethics of clinical trials that withhold standard care** p. 729

**Weighing black holes with photometry** pp. 734 & 789

**The promise of therapeutics made in plants** p. 740

# Science

$15
13 AUGUST 2021
sciencemag.org

AAAS

## ICE AGE WANDERER

A tusk records the history of a mammoth's life p. 806

# 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**.

*"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. 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.

NEW ENGLAND BioLabs® Inc.

*be* INSPIRED
*drive* DISCOVERY
*stay* GENUINE

# CONTENTS

13 AUGUST 2021 • VOLUME 373 • ISSUE 6556

## NEWS

### IN BRIEF

720 News at a glance

### IN DEPTH

**723 Time grows short to curb warming, report warns**
IPCC science analysis concludes human role 'unequivocal' and impact 'unprecedented' *By C. O'Grady*

**724 Streams that flow only part of the year are getting even drier**
Analysis of intermittent U.S. waterways finds many are shriveling earlier and remaining dry for much longer *By E. Stokstad*

**725 Science lost and lessons learned: A lab plots its comeback**
A microbiology team regroups, with a more virtual lab and a bigger focus on mental health *By D. Grimm*

**726 'Mini–Manhattan Projects' for energy innovation wind down**
But hub model for bridging basic and applied research lives on *By A. Cho*

**727 Genetics papers from China face ethical scrutiny**
Questions about consent and potential for abuse trigger investigations *By D. Normile*

### FEATURES

**729 Failure to protect?**
A study of asthmatic children, most of them Black, shows how a common clinical trial design can expose vulnerable participants to serious risks *By C. Piller*
PODCAST

## INSIGHTS

### PERSPECTIVES

**734 How massive is that black hole?**
The flux of radiation emissions from accretion disks correlates with black hole mass *By P. Lira and P. Arevalo*
REPORT p. 789

**736 When did terrestrial plants arise?**
Microfossils suggest that co-option of algal genes may affect land plant origination time *By P. G. Gensel*
REPORT p. 792

**737 The cell of origin for Barrett's esophagus**
Undifferentiated cells that closely resemble gastric cells could be a biomarker for surveillance *By K. Geboes and A. Hoorens*
RESEARCH ARTICLE p. 760

**738 Taking the long view on metabolism**
Measured energy expenditure across the human life span reveals distinct metabolic phases *By T. W. Rhoads and R. M. Anderson*
REPORT p. 808

**740 Plant-made vaccines and therapeutics**
Advances in technology and manufacturing could boost the uptake of molecular farming *By H. Fausther-Bovendo and G. Kobinger*

**742 Searching for life on Mars and its moons**
Sample-return missions will look for extraterrestrial life and biomarkers on Mars and Phobos *By R. Hyodo and T. Usui*

**743 Making machine learning trustworthy**
Safety, transparency, and fairness are essential for high-stakes uses of machine learning *By B. Eshete*
PODCAST

**745 Richard C. Lewontin (1929–2021)**
Groundbreaking evolutionary geneticist *By A. Berry and D. A. Petrov*

### POLICY FORUM

**746 Integrate biodiversity targets from local to global levels**
A shared Earth approach links biodiversity and people *By D. O. Obura* et al.

ILLUSTRATIONS (TOP TO BOTTOM): ESA/HUBBLE, M. KORNMESSER/CC BY 4.0; STEPHAN SCHMITZ

# REVEALING THE NEXT BREAKTHROUGH IN IMMUNOLOGY

## APPLY NOW FOR THE $30,000 GRAND PRIZE

**APPLICATIONS DUE: OCTOBER 1, 2021**

The Michelson Philanthropies & ***Science*** Prize for Immunology recognizes research that will have a significant impact on vaccine and immunotherapy development. Together with AAAS/***Science***, Michelson Philanthropies awards breakthrough research from the past three years. The applications are judged on innovation and potential impact.

With these prizes, scientists (35 and younger) have the opportunity to advance their research projects, are provided with a renowned platform in ***Science*** to publish their work and connect them with other innovators through the Michelson Prizes alumni network.

Young researchers from a wide range of disciplines are encouraged to apply.

APPLY TODAY:
www.sciencemag.org/michelson

MICHELSON PHILANTHROPIES & Science
PRIZE FOR IMMUNOLOGY

## BOOKS *ET AL.*

**749 The matter of mind control**
Brainwashing case studies illuminate the history of coercive persuasion
*By S. Marks*

**750 The alternative to despair is to build an ark**
H. G. Wells's "world encyclopedia" has merit beyond its seeming similarities to Wikipedia
*By Y. Benkler*

## LETTERS

**751 China's algal bloom suffocates marine life**
*By X. Guo* et al.

**752 Eastern Europe's fraught waterway plans**
*By I. Kitowski and G. Grzywaczewski*

**752 Australia threatens to weaken forest laws**
*By D. Lindenmayer and C. Taylor*

# RESEARCH

## IN BRIEF

**754** From *Science* and other journals

## REVIEW

**757 Neuroscience**
Nicotinic acetylcholine receptor redux: Discovery of accessories opens therapeutic vistas *J. A. Matta* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABG6539

## RESEARCH ARTICLES

**758 Microbiology**
Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution
*D. Dar* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABI4882

**759 Coronavirus**
Ultrapotent antibodies against diverse and highly transmissible SARS-CoV-2 variants
*L. Wang* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABH1766
REPORT p. 818

**760 Cancer genomics**
Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition *K. Nowicki-Osuch* et al.
PERSPECTIVE p. 737

**768 CRISPR biology**
Structural basis for target site selection in RNA-guided DNA transposition systems
*J.-U. Park* et al.

**774 Plant science**
Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes *F. Sabbadin* et al.

## REPORTS

**779 Ultracold molecules**
Observation of microwave shielding of ultracold molecules *L. Anderegg* et al.

**783 Polymer chemistry**
Chemically recyclable thermoplastics from reversible-deactivation polymerization of cyclic acetals *B. A. Abel* et al.

**789 Black holes**
A characteristic optical variability time scale in astrophysical accretion disks
*C. J. Burke* et al.
PERSPECTIVE p. 734

**792 Paleobotany**
A fossil record of land plant origins from charophyte algae *P. K. Strother and C. Foster*
PERSPECTIVE p. 736

**797 Superconductivity**
Multicomponent superconducting order parameter in $UTe_2$ *I. M. Hayes* et al.

**801 2D materials**
Boridene: Two-dimensional $Mo_{4/3}B_{2-x}$ with ordered metal vacancies obtained by chemical exfoliation *J. Zhou* et al.

**806 Paleontology**
Lifetime mobility of an Arctic woolly mammoth *M. J. Wooller* et al.

**808 Metabolism**
Daily energy expenditure through the human life course *H. Pontzer* et al.
PERSPECTIVE p. 738

**813 Microbiota**
High-fat diet–induced colonocyte dysfunction escalates microbiota-derived trimethylamine *N*-oxide
*W. Yoo* et al.

**818 Coronavirus**
Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants
*M. Yuan* et al.
RESEARCH ARTICLE p. 759

## DEPARTMENTS

**719 Editorial**
Clarion call from climate panel
*By Steven Sherwood and Brian Hoskins*

**826 Working Life**
Keep quiet about homophobia or open up?
*By Brian Mustanski*


## ON THE COVER

Reproduction of a life-size oil painting of an adult male woolly mammoth navigating a mountain pass in Arctic Alaska. Little is known about the movement patterns of these extinct giants. Isotopic records from a 17,100-year-old mammoth tusk reveal that the animal covered an extensive geographic range during its lifetime. However, as the ice age ended and the Arctic environment began to change, maintaining this level of mobility would have been increasingly difficult. See page 806.

*Illustration: James Havens/ The Havens Studio, Alaska*


*Science* Staff ..... 718
New Products ..... 824
*Science* Careers ..... 825


**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 © 2021 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
**Change of address:** Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

ILLUSTRATION: ROBERT NEUBECKER

# Science AAAS

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

**Executive Editor** Monica M. Bradford

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

**DEPUTY EDITORS** Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Di Jiang, Marc S. Lavine (Canada), 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, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Katris, Nida Masiulis, 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), Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronmel Navas, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

**News Editor** Tim Appenzeller

**NEWS MANAGING EDITOR** John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Kelly Servick, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Menna Ibrahim, Sofia Moutinho, Anil Oza, Alex Viveros **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, Michael Price, Tania Rabesandratana (Barcelona), Joshua Sokol, 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** Alberto Cuadra **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 ILLUSTRATORS** Valerie Altounian, Chris Bickel **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 Pajerowski **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell **INTERN** Claire Hogan

**Chief Executive Officer and Executive Publisher** Sudip Parikh

**Publisher, Science Family of Journals** Bill Moran

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

**ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** 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 MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATE** Anne Mason **SPJ ASSOCIATE MANAGER** Samantha Bruno Fuller **SPJ ASSOCIATE** Casey Buchta

**DIRECTOR, INSTITUTIONAL LICENSING** Iquo Edim **MARKETING MANAGER** Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan, Claudia Paulsen-Young **CUSTOMER SUCCESS MANAGER** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer

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

**DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS** Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **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**
sciencemag.org/authors/science-information-authors

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

**MEDIA CONTACTS**
scipak@aaas.org

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

**INSTITUTIONAL SALES AND SITE LICENSES**
sciencemag.org/librarian

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

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

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

**MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS**
sciencemag.org/subscriptions

**MEMBER BENEFITS**
aaas.org/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

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



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

**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 Arlotta,** *Harvard U.*
**David Awschalom,** *U. of Chicago*
**Clare Baker,** *U. of Cambridge*
**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* (**S**)
**Christian Büchel,** *UKE Hamburg*
**Dennis Burton,** *Scripps Res.*
**Carter Tribley Butts,** *UC Irvine*
**György Buzsáki,** *NYU School of Med.*
**Mariana Byndloss,** *Vanderbilt U. Med. Ctr.*
**Annmarie Carlton,** *UC Irvine*
**Simon Cauchemez,** *Inst. Pasteur*
**Ling-Ling Chen,** *SIBCB, CAS*
**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. Dangl,** *UNC*
**Chiara Daraio,** *Caltech*
**Nicolas Dauphas,** *U. of Chicago*
**Christian Davenport,** *U. of Michigan*
**Frans de Waal,** *Emory U.*
**Claude Desplan,** *NYU*
**Sandra Díaz,** *U. Nacional de Córdoba*
**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.*
**Gwenn 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*
**Taekjip Ha,** *Johns Hopkins U.*
**Sharon Hammes-Schiffer,** *Yale U.*
**Wolf-Dietrich Hardt,** *ETH Zürich*
**Louise Harra,** *U. Coll. London*
**Jian He,** *Clemson U.*
**Carl-Philipp Heisenberg,** *IST Austria*
**Ykä Helariutta,** *U. of Cambridge*
**Janet G. Hering,** *Eawag*
**Heather Hickman,** *NIAID, NIH*
**Hans Hilgenkamp,** *U. of Twente*
**Kai-Uwe Hinrichs,** *U. of Bremen*
**Deirdre Hollingsworth,** *U. of Oxford*
**Randall Hulet,** *Rice U.*
**Auke Ijspeert,** *EPFL*
**Akiko Iwasaki,** *Yale U.*
**Stephen Jackson,** *USGS & U. of Arizona*
**Erich Jarvis,** *Rockefeller U.*
**Peter Jonas,** *IST Austria*
**Matt Kaeberlein,** *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 Koechlin,** *École Normale Supérieure*
**Alex L. Kolodkin,** *Johns Hopkins U.*
**Julija Krupic,** *U. of Cambridge*
**Paul Kubes,** *U. of Calgary*
**Gabriel Lander,** *Scripps Res.* (**S**)
**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*
**Anne Magurran,** *U. of St. Andrews*
**Asifa Majid,** *U. of York*
**Oscar Marín,** *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 Medellín,** *U. Nacional Autónoma de México*
**Jane Memmott,** *U. of Bristol*
**C. Jessica Metcalf,** *Princeton U.*
**Baoxia Mi,** *UC Berkeley*
**Tom Misteli,** *NCI, NIH*
**Alison Motsinger-Reif,** *NIEHS, NIH* (**S**)
**Suresh Naidu,** *Columbia U.*
**Danielle Navarro,** *U. of New South Wales*
**Daniel Nettle,** *Newcastle U.*
**Daniel Neumark,** *UC Berkeley*
**Beatriz Noheda,** *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.* (**S**)
**Daniel Pauly,** *U. of British Columbia*
**Ana Pêgo,** *U. do Porto*
**Samuel Pfaff,** *Salk Inst.*
**Julie Pfeiffer,** *UT Southwestern Med. Ctr.*
**Philip Phillips,** *UIUC*
**Matthieu Piel,** *Inst. Curie*
**Kathrin Plath,** *UCLA*
**Martin Plenio,** *Ulm U.*
**Katherine Pollard,** *UCSF*
**Elvira Poloczanska,** *Alfred-Wegener-Inst.*
**Julia Pongratz,** *Ludwig Maximilians U.*
**Philippe Poulin,** *CNRS*
**Jonathan Pritchard,** *Stanford U.*
**Lei Stanley Qi,** *Stanford U.*
**Trevor Robbins,** *U. of Cambridge*
**Joeri Rogelj,** *Imperial Coll. London*
**Amy Rosenzweig,** *Northwestern U.*
**Mike Ryan,** *UT Austin*
**Miquel Salmeron,** *Lawrence Berkeley Nat. Lab*
**Nitin Samarth,** *Penn State U.*
**Erica Ollmann Saphire,** *La Jolla Inst.*
**Joachim Saur,** *U. zu Köln*
**Alexander Schier,** *Harvard U.*
**Wolfram Schlenker,** *Columbia U.*
**Susannah Scott,** *UC Santa Barbara*
**Anuj Shah,** *U. of Chicago*
**Vladimir Shalaev,** *Purdue U.*
**Jie Shan,** *Cornell U.*
**Beth Shapiro,** *UC Santa Cruz*
**Jay Shendure,** *U. of Wash.*
**Steve Sherwood,** *U. of New South Wales*
**Brian Shoichet,** *UCSF*
**Robert Siliciano,** *JHU School of Med.*
**Lucia Sivilotti,** *U. Coll. London*
**Alison Smith,** *John Innes Ctr.*
**Richard Smith,** *UNC* (**S**)
**Mark Smyth,** *QIMR Berghofer*
**John Speakman,** *U. of Aberdeen*
**Tara Spires-Jones,** *U. of Edinburgh*
**Allan C. Spradling,** *Carnegie Institution for Sci.*
**V. S. Subrahmanian,** *Dartmouth*
**Ira Tabas,** *Columbia U.*
**Eriko Takano,** *U. of Manchester*
**Patrick Tan,** *Duke-NUS Med. School*
**Sarah Teichmann,** *Wellcome Sanger Inst.*
**Rocio Titiunik,** *Princeton U.*
**Shubha Tole,** *Tata Inst. of Fundamental Res.*
**Maria-Elena Torres Padilla,** *Helmholtz Zentrum München*
**Kimani Toussaint,** *Brown U.*
**Barbara Treutlein,** *ETH Zürich*
**Wim van der Putten,** *Netherlands Inst. of Ecology*
**Henrique Veiga-Fernandes,** *Champalimaud Fdn.*
**Reinhilde Veugelers,** *KU Leuven*
**Bert Vogelstein,** *Johns Hopkins U.*
**David Wallach,** *Weizmann Inst.*
**Jane-Ling Wang,** *UC Davis* (**S**)
**Jessica Ware,** *Amer. Mus. of Natural Hist.*
**David Waxman,** *Fudan U.*
**Chris Wikle,** *U. of Missouri* (**S**)
**Terrie Williams,** *UC Santa Cruz*
**Ian A. Wilson,** *Scripps Res.* (**S**)
**Hao Wu,** *Harvard U.*
**Wei Xie,** *Tsinghua U.*
**Yu Xie,** *Princeton U.*
**Jan Zaanen,** *Leiden U.*
**Kenneth Zaret,** *UPenn School of Med.*
**Bing Zhu,** *Inst. of Biophysics, CAS*
**Xiaowei Zhuang,** *Harvard U.*
**Maria Zuber,** *MIT*

EDITORIAL

# Clarion call from climate panel

Unprecedented flooding, searing temperatures, and raging fires across Europe, Asia, and North America this summer have created a stark backdrop for this week's release of the sixth physical science assessment report (AR6) of the Intergovernmental Panel on Climate Change (IPCC). These reports, initiated in 1990, arrive about every 7 years at the request of the countries of the United Nations Framework Convention on Climate Change. They form the basis for UN discussions and have become a crucial means to take stock of the latest scientific developments. The reports' future projections about climate change have remained fairly stable over the years and have, sadly, proven quite accurate. So, what does the new report add?

Above all, AR6 expresses greater confidence in familiar findings, owing to stronger evidence. A notable example concerns "equilibrium climate sensitivity," a measure of how much global warming ultimately occurs if the atmospheric carbon dioxide ($CO_2$) concentration doubles. Based on improved understanding of cloud processes and climate changes that have already occurred, AR6 concludes that this figure is "likely" (a two-thirds chance or greater) to lie between 2.5° and 4°C—halving the spread of 1.5° to 4.5°C in previous reports. Global temperatures had stalled in the period before the 2013 assessment (AR5) but have since surged, reaching 1.1°C above that of preindustrial times. Atmospheric $CO_2$ has reached concentrations not seen for at least 2 million years, and the new report expresses high confidence that oceans, plants, and soils will become less efficient at absorbing future carbon emissions.

As always, uncertainty remains. The latest climate models predict a wider range for climate sensitivity, with projected values implausibly weak in some cases but implausibly strong in others. This disagreement is largely a result of increased complexity in model representations of cloud feedbacks in the midlatitude storm-track regions. AR6 shrewdly deals with this inconsistency by focusing on what happens at a given level of global warming (say, 2°C), separating this from the question of when that warming level would be reached.

The report also provides new clarity on aspects like changes in extreme rainfall and drought. Almost all robustly observed regional trends in these events are upward and are projected to continue. One sobering finding is that even if global warming is limited to 2°C, heat events that once occurred twice per century will happen every 3 to 4 years—and will tend to coincide with droughts, compounding the impacts. Much better regional information is provided than in previous reports. However, the lack of adequate data in many regions, including most of Africa, is apparent and should be addressed.

The report dives into important new territory by emphasizing "low-probability, high-impact events" that are hard to quantify but unwise to ignore. For example, although the expected range of future sea level is similar to previous predictions, AR6 indicates that rises of 2 m or more by the end of the century cannot be ruled out. Nor can the possibility of abrupt responses and "tipping points" in the climate system. These are stark warnings compared with previous reports. As the authors note, the probabilities of forest dieback, ocean-circulation changes, and other disturbing scenarios increase with global temperature.

> "...this may be the last report that can... keep the climate targets of the 2015 Paris Agreement within reach."

Although the IPCC reports provide an invaluable resource and periodic wake-up call, they come at a price. This report was written by 234 authors over 3 years, with similar effort invested in two more reports on adaptation and mitigation due next year. The process is arduous: Over 75,000 review comments were individually addressed. The world's climate modeling centers invest heavily in simulations following common protocols, which is growing steadily more taxing for them.

If another assessment is commissioned on schedule, it will arrive not much before 2030. By then, if emissions persist at current rates—that is, even if emissions growth is halted—nearly all the remaining "global carbon budget," which gives a 50-50 chance of keeping global warming below 1.5°C, will have been exhausted. So, this may be the last report that can meaningfully influence policy to keep the climate targets of the 2015 Paris Agreement within reach. AR6 is intended to inform discussions at the UN Climate Change Conference of the Parties (COP26) meeting in November. Our children and grandchildren are waiting to see what comes out of it.

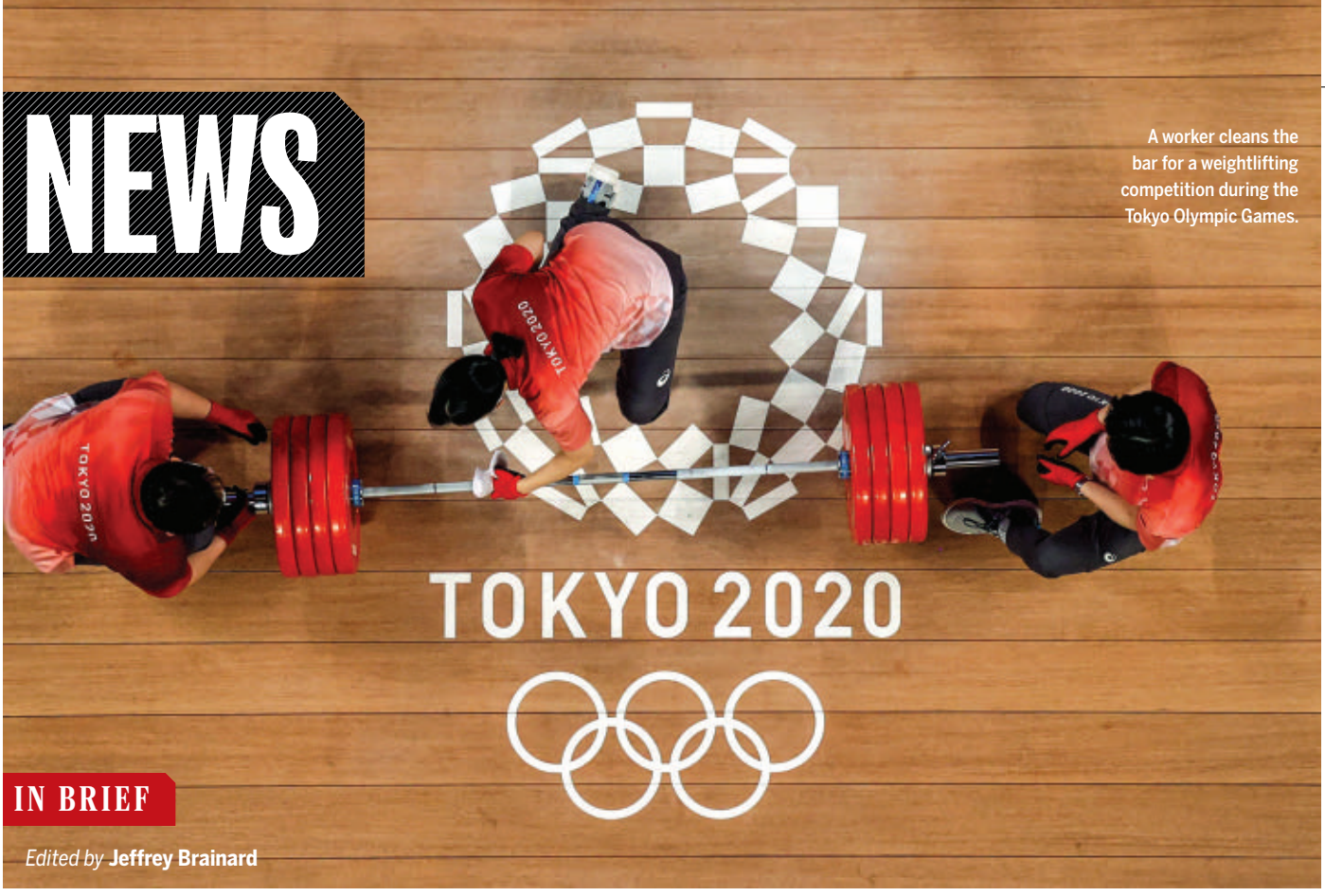
**–Steven Sherwood and Brian Hoskins**

**Steven Sherwood** is a professor at the Australia Research Council Centre of Excellence for Climate Extremes at the University of South Wales, Sydney, Australia. s.sherwood@unsw.edu.au

**Brian Hoskins** is chair of the Grantham Institute at Imperial College London, London, UK, and a professor in the Department of Meteorology at the University of Reading, Reading, UK. b.hoskins@imperial.ac.uk

Published online 10 August 2021; 10.1126/science.abl8490

# NEWS




A worker cleans the bar for a weightlifting competition during the Tokyo Olympic Games.

## IN BRIEF

*Edited by* **Jeffrey Brainard**

PUBLIC HEALTH

## COVID-19 largely spares Olympic games, but not Tokyo

Organizers of the Tokyo Olympic Games said they successfully minimized the spread of COVID-19 during the 2-week event. A total of 484 people of an estimated 50,000 involved in the games tested positive for SARS-CoV-2 between 1 July, when quarantining of foreign visitors began, and this week when *Science* went to press. Most (257) were contractors supporting the games, followed by 168 Olympics personnel and volunteers, 30 journalists, and 29 athletes, who were automatically barred from competition. Most who tested positive are Japanese residents. Japan's government banned members of the public from attending the games, and organizers' other precautions included daily testing of athletes. More than 70% of foreign athletes and staff members at the games had been vaccinated against COVID-19, *The Washington Post* reported. Only about 33% of the Japanese population has been fully vaccinated, and a majority of survey respondents opposed holding the games. Tokyo residents had a record number of new cases during the Olympics, but Japanese Prime Minister Yoshihide Suga insisted the surge was unrelated because the games' precautions were stringent.

### Scientists share Olympic glory

ATHLETICS | Among throngs of elite athletes, several researchers pedaled, jumped, or sprinted their way to medals during the Tokyo Olympic Games. Among them: Austrian mathematician Anna Kiesenhofer won a gold medal in women's road cycling in a dramatic finish. Kiesenhofer is a postdoctoral fellow at the Swiss Federal Institute of Technology Lausanne who focuses on nonlinear partial differential equations in mathematical physics. In track and field, Hugues Fabrice Zango, a Ph.D. candidate in electrical engineering at the University of Lille, became the first Olympic medalist from Burkina Faso, winning the bronze in the men's triple jump. He aspires to teach at a university in his home country. And epidemiologist Gabby Thomas of the United States earned two medals: a silver in the women's 4x100-meter relay and a bronze in the 200 meters. Thomas is interested in health disparities, particularly among Black people, and is pursuing a master's degree at the University of Texas, Austin.

Mathematician Anna Kiesenhofer won gold in road cycling.

PHOTOS: (TOP TO BOTTOM) CHRIS GRAYTHEN/POOL/AFP/GETTY IMAGES; MICHAEL STEELE/GETTY IMAGES

## Long Covid tracked in children

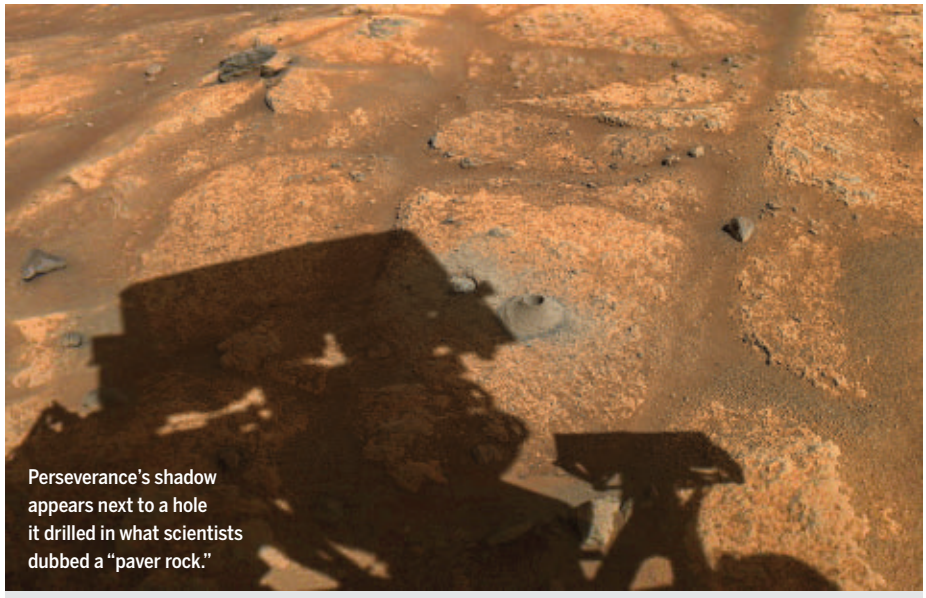
EPIDEMIOLOGY | About 4% of U.K. children sickened by SARS-CoV-2 had symptoms that lasted at least 4 weeks and were suggestive of Long Covid, a study has found. In adults, the condition has been marked by extreme fatigue, heart palpitations, neurological problems, and more. In one of the largest studies of children to date, 77 of 1734 children ages 5 to 17 with a positive test and acute symptoms still had them after 4 weeks, and 25, or 1.8%, of 1379 still had symptoms after 8 weeks. Older children were more likely than younger ones to have persistent symptoms, the research team reported on 3 August in *The Lancet Child & Adolescent Health*.

## Protein vaccine hits setback

COVID-19 | Novavax announced last week that the U.S. government has ordered it to stop making its promising protein-based coronavirus vaccine in the United States. The government also ceased funding new manufacturing there; it can resume only when the small Maryland firm meets the U.S. Food and Drug Administration's (FDA's) standards for ensuring the quality and potency of its vaccine, the company revealed in a quarterly Securities and Exchange Commission filing on 5 August. It added that it will not file for an emergency use authorization (EUA) from FDA until the fourth quarter, instead of the third quarter as planned. But Novavax, which has received $1.75 billion in U.S. funding for developing and making its vaccine, will continue to manufacture it in other countries. It announced last week that, with its partner the Serum Institute of India, it has filed for EUAs in India, Indonesia, and the Philippines—its first such filings. Novavax and the Serum Institute have agreements to provide more than 1.6 billion doses of the vaccine worldwide, 1.1 billion of them destined for low- and middle-income countries.

## Marburg strikes West Africa

INFECTIOUS DISEASES | West Africa is again facing the threat of a deadly hemorrhagic fever. A man who died in Guéckédou prefecture in Guinea on 2 August suffered from Marburg disease, laboratory tests in the Guinean capital Conakry and Dakar, Senegal, have shown. The Marburg virus is closely related to the Ebola virus, but its outbreaks are rarer and typically smaller; the biggest one, in Angola in 2004–05, ended after 252 cases. There are no approved vaccines or treatments. The new case occurred close to Guinea's borders with Liberia and Sierra Leone, in the same area where West Africa's massive Ebola epidemic of 2013–16 began. Contact tracing is ongoing, but given the country's fragile health care system and the burden of the COVID-19 pandemic, the World Health Organization says the risk for the country and the region is high.



Perseverance's shadow appears next to a hole it drilled in what scientists dubbed a "paver rock."

PHOTO: NASA/JPL-CALTECH

PLANETARY SCIENCE

## Mars rover's first drilling comes up empty

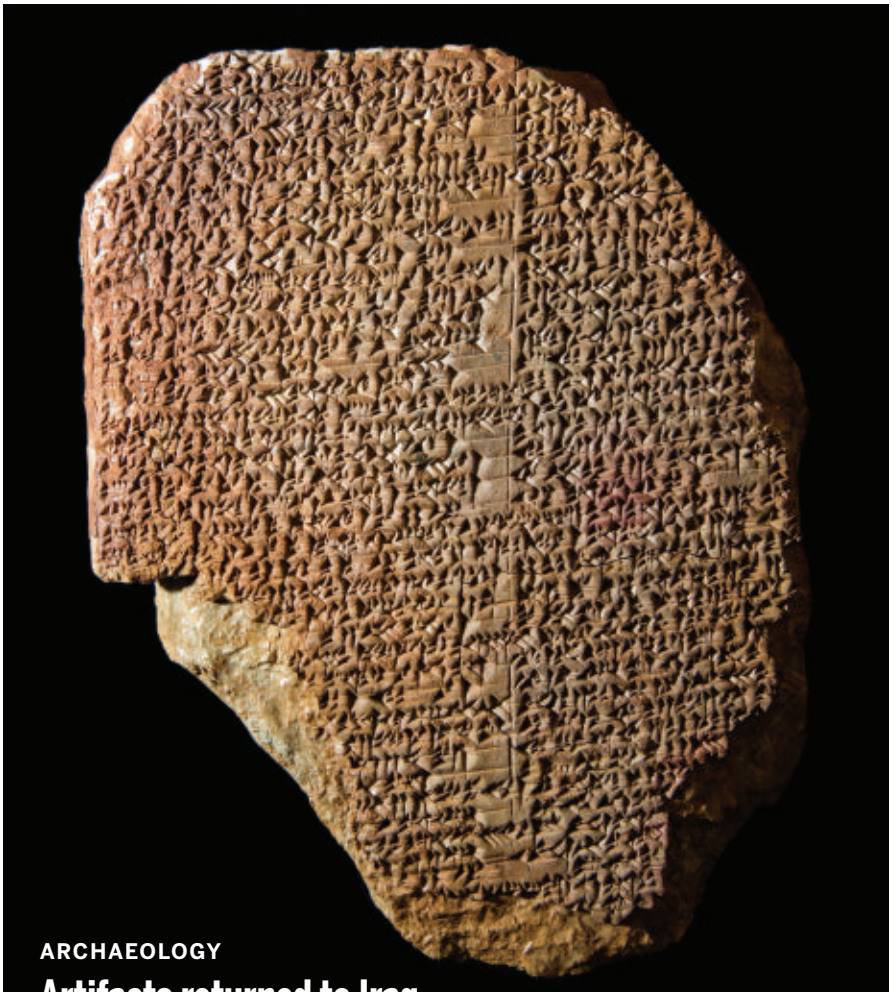
Puzzled NASA engineers are studying why the Perseverance rover last week apparently failed to collect rocks and dust after it drilled Mars's surface for the first time. Perseverance recovered the tube meant to contain samples from the drill hole in Jezero crater, but data indicated the tube was empty. NASA says it has not identified any equipment malfunctions, and engineers think unexpected characteristics of the rock might explain the failure. To diagnose what went wrong, NASA will attempt to photograph the drill hole up close. The rover carries a total of 43 titanium sample tubes, which were to be filled with material from various sites; some will be stored in a cache until another mission can return them to Earth, where they would be studied for signs of ancient life. "While this is not the 'hole-in-one' we hoped for, there is always risk with breaking new ground," said Thomas Zurbuchen, NASA's associate administrator for science. "I'm confident we have the right team working this, and we will persevere toward a solution to ensure future success." Previous drilling on Mars by other NASA probes has also encountered snags. As *Science* went to press, the next drilling attempt by Perseverance had not been scheduled.

## U.K. agency pushes open access

PUBLISHING | The United Kingdom's leading science funder announced last week a new policy that will allow authors to post articles it finances in free-to-read repositories upon publication, with no embargo period. The move by UK Research and Innovation (UKRI) is being resisted by publishers, who say it threatens to peel away subscribers and damage the industry financially; under an existing UKRI policy, authors had to wait up to 12 months. The new rules, which take effect in April 2022, also allow authors to pay journals for "gold" open access, which makes a paper free to read on the publisher's website—but only if the journal is transitioning to exclusively open access for all research papers. And the papers must bear a license allowing for text mining and free and liberal distribution of the work. The move brings UKRI's policy into alignment with Plan S, an open-access effort led by European research funders, including UKRI. The United Kingdom currently has one of the highest rates of open-access publication in the world, with about three-quarters of recently published papers free to read upon publication.

## Facebook, researchers square off

POLITICAL SCIENCE | A simmering fight between Facebook and a team of researchers monitoring its political ads boiled over last week when the social media giant suspended their accounts. The scientists running New York University's Ad Observatory project, begun in the runup to the November 2020 U.S. elections, contend that the data being collected from Facebook's website, including the names of advertisers, should be considered public information. But the company says the research violates its terms of service and the privacy of its customers. (Facebook had previously asked the researchers to stop, but said it would not take any steps until after the election.) On 4 August, Facebook disabled the accounts of graduate student Laura Edelson and engineering professor Damon McCoy, prompting an outcry from scientists. On 6 August, three Democratic senators asked Facebook CEO Mark Zuckerberg to explain the company's decision and for a tally of researchers and journalists whose accounts have been suspended this year.



ARCHAEOLOGY

## Artifacts returned to Iraq

The United States has returned to Iraq more than 17,000 looted ancient artifacts and plans to give back a rare, 3500-year-old clay tablet bearing part of the Babylonian Epic of Gilgamesh, Iraqi officials told Reuters last week. The items had been smuggled out of Iraq during the chaos of the 2003 U.S. invasion and the takeover of parts of Iraq by the Islamic State group 10 years later. The Hobby Lobby craft store chain and Cornell University had acquired most of the artifacts returned in late July. Hobby Lobby's owner had intended to give the artifacts to the Museum of the Bible in Washington, D.C., where the Gilgamesh tablet (above)–containing part of the world's first known literary work, written in the Akkadian language—has been displayed. In 2017, Hobby Lobby agreed to pay the U.S. government a $3 million fine for possessing smuggled artifacts. The repatriation is considered the largest in Iraqi history.

## Yemeni academics dismissed

INTERNATIONAL RELATIONS | Saudi Arabia is purging Yemeni scientists and other scholars from universities and hospitals in southern provinces near its war-torn neighbor. Saudi Arabia is aiding the Yemeni government's fight against Houthi rebels, but the vast majority of those dismissed are non-Houthis. At Najran University, all 106 Yemeni faculty members received written notices on 8 August that their job contracts are terminated as of 15 August. Similarly precipitous pink slips were handed out to hundreds more Yemeni academics and medical professionals at other Saudi institutions, says an expat Yemeni scientist who helped organize a change.org petition drawing attention to their plight. "It seems the Saudi government wants to clear the entire region of Yemenis," he says. Saudi officials have not commented. For a foreign worker in the kingdom, losing a job also means losing the right to reside there. The displaced scholars and their families fear political persecution if they were to return to Yemen, the petition notes, and have few other options as "only a handful of countries are granting visas to Yemeni nationals."

## Continent-size radio array planned

ASTRONOMY | The U.S. National Science Foundation (NSF) this week said it awarded the National Radio Astronomy Observatory $23 million to design a powerful radio telescope with 263 dish antennas spread across North America. The Next Generation Very Large Array would have resolving power—the ability to see fine detail—more than 10 times greater than NSF's current Very Large Array in New Mexico. It is expected to address fundamental questions in all major areas of astrophysics and to complement existing arrays and planned ones such as the Square Kilometre Array. The project faces hurdles ahead, though: It awaits a decision by the U.S. National Academies of Sciences, Engineering, and Medicine whether to include it in the forthcoming Astronomy and Astrophysics Decadal Survey, a key blueprint for U.S. spending in those fields, and the array also needs additional funding from Congress. Construction could begin by 2026 and full scientific operations by 2035.

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

PHOTO: MUSEUM OF THE BIBLE/GREEN COLLECTION/ALL RIGHTS RESERVED. © MUSEUM OF THE BIBLE, 2017

Firefighters battle a blaze in Greece on 5 August. Climate change is stoking such extreme events.

PHOTO: MICHAEL VARAKLAS/AP IMAGES

CLIMATE CHANGE

# Time grows short to curb warming, report warns

## IPCC science analysis concludes human role 'unequivocal' and impact 'unprecedented'

*By* **Cathleen O'Grady**

Every time the U.N. Intergovernmental Panel on Climate Change (IPCC) issues a new report surveying the science of global warming, the alarm sounds louder. Now, 8 years after its last report, the message of IPCC's latest assessment, released this week, is urgent and unequivocal: The window for mitigating the worst projected impacts of global climate change is closing. Average global temperatures are now 1.1°C above preindustrial records, and under every scenario for future greenhouse gas emissions that the panel examined, average warming of 1.5°C—a target set by the 2015 Paris climate accord—will very likely be reached within the next 20 years.

"This is a critical decade for keeping the 1.5°C target within reach," says Jane Lubchenco, deputy director for climate and the environment at the White House Office of Science and Technology Policy. The projections mean countries should come to the U.N. Climate Change Conference, scheduled for November, with the most "aggressive, ambitious" targets for reducing emissions possible, she says.

It is "unequivocal" and "established fact" that human activities are causing Earth to warm, says the report, which was assembled by hundreds of scientists from around the world, and climate change's impact on the planet is "unprecedented." That blunt language reflects, in part, the substantial scientific advances that have occurred since IPCC issued its last major assessment in 2013. Climate models are more detailed and powerful, and observational records cover more ground—including the rapidly warming Arctic—as well as eight additional years of warming. They have enabled researchers to better "see the climate change signal developing," says Nerilie Abram, a climate scientist at Australian National University.

Growing evidence from ancient climates has also helped researchers constrain estimates of what is called climate sensitivity—the amount of warming expected if concentrations of atmospheric carbon dioxide ($CO_2$) rise twice as high as in preindustrial times, to 560 parts per million (ppm). (Current levels are about 415 ppm and climbing fast.) The panel now estimates that a $CO_2$ doubling would boost temperatures by 2.5°C to 4°C, a narrower range than the previously estimated 1.5°C to 4.5°C.

In the best case scenario, with the world cutting net emissions to zero by 2050, $CO_2$ will fall short of doubling and warming is projected to peak midcentury at 1.6°C above preindustrial levels. Even in this scenario, Arctic sea ice is likely to vanish completely in at least one summer by 2050. In the worst case scenario, warming will very likely reach 2.4°C by midcentury and rise to 4.4°C—and potentially 5.7°C—by 2100. At higher emissions levels, "low-likelihood, high-impact" consequences—such as mass ice sheet loss in the Antarctic or the stalling of ocean currents—become more likely. The probability of these abrupt, irreversible changes is not well-understood, the report says, but they cannot be ruled out. Current emissions are on IPCC's mid- to higher trajectories, Imperial College London climate scientist Joeri Rogelj said at a briefing, and countries' pledges still fall short of achieving the lowest emissions scenario. "Let's be clear," he said, "about the work that still needs to be done."

For the first time, the report elaborates on how each increment of warming could play out in different regions, stoking extreme events such as flooding, heat waves, droughts, and fire. Past reports focused on averages, Abram notes, but "people don't live in the global average." One forecast is that climate change will give extra potency to existing natural variability in tem-

peratures and precipitation. With 1.5°C of warming, for example, high daytime temperatures that would be rare without climate change could occur four times a decade; at 4°C of warming, such heat extremes could come nearly every year. Such projections illustrate that "every little bit of warming counts," says Claudia Tebaldi, a climate scientist at Pacific Northwest National Laboratory and an author of the report. But she and others emphasize that targets like 1.5°C should not be seen as precipices beyond which there is no redemption or hope.

It is now "established fact" that warming is fueling extreme events, the report says. And since IPCC's last report, scientific advances have made it possible to link climate change to specific events, such as the recent heat wave in northwestern North America, says Francisco Doblas Reyes, a climate scientist at the Barcelona Supercomputing Center and a report author. It's clear that "climate change is here now," he says. "No region is spared," adds Sonia Seneviratne, a climate scientist at ETH Zürich and a report author. New extremes in heat, precipitation, or drought have been observed in nearly every global region. "We are starting to see events which would have had near zero probability of happening without human-induced climate change," Seneviratne says.

Some global changes are already locked in, the report notes, regardless of how much humanity reduces emissions in coming decades. Melting of glaciers and ice sheets and thawing of permafrost is now "irreversible" for decades or centuries to come, it says. The warming, acidification, and deoxygenation that are already damaging the world's oceans will continue for centuries to millennia. Continued sea level rise, now estimated at 3.7 millimeters each year between 2006 and 2018, is also inevitable: Over the next 2000 years, oceans will likely rise by 2 to 3 meters if the planet warms by 1.5°C, and up to 22 meters with 5°C of warming.

Because of the ongoing COVID-19 pandemic, the hundreds of scientists who wrote the assessment had to meet online to wrestle with how to convey the extent of the climate crisis and the urgent need to act. It was uncanny to see "the echoes of one crisis in another," Tebaldi says. And for many researchers, the work isn't done. The science assessment—the sixth produced by IPCC since 1990—is just the first of three major reports that IPCC's 195 member states will release over the coming year. The next reports will examine how climate policies can reduce emissions, and what actions will be needed to adapt to extreme events such as flooding, heat waves, and drought. ■

RIVER SCIENCE

# Streams that flow only part of the year are getting even drier

## Analysis of intermittent U.S. waterways finds many are shriveling earlier and remaining dry for much longer

*By* **Erik Stokstad**

Small streams that dry up for part of the year are easy to overlook. But these intermittent streams are everywhere, making up more than half of Earth's waterways. They help purify surface water and provide crucial habitat for creatures such as the Sonoran Desert toad, fairy shrimp, and Wilson's warbler. Now, a study has found that ephemeral streams across the continental United States have become less reliable over the past 40 years, likely as a result of climate change. Some are dry for 100 days longer per year than in the 1980s. "That's really shocking," says Sarah Null, a watershed scientist at Utah State University.

The findings, reported last month in *Environmental Research Letters*, come from a study of data collected between 1980 and 2017 by flow gauges on 540 intermittent streams around the United States. Most of the gauges were on small waterways in river headwaters, but a few tracked large rivers that are intermittent in places, such as the Rio Grande, which flows sporadically in New Mexico and Texas. The sample covered just a small fraction of intermittent streams, the authors note, and left out some states, such as Nebraska and Maine, that don't have any long-term gauges on these streams. Still, the analysis revealed some eye-opening regional shifts, says Sam Zipper, one of the authors and an ecohydrologist with the Kansas Geological Survey.

More than half of the gauges showed changes in the streams' flow patterns since 1980. Some now shrivel earlier in the year and remain dry for longer, for example, or they dwindle more quickly than before. At some 7% of gauges, dry periods expanded by 100 days or more.

That is bad news for the many plants and animals that time their reproduction to the availability of water, especially in deserts. "Just because the channel is dry does not make it biologically dead," says river scientist Ellen Wohl of Colorado State University. It also has implications for water quality, as microbes in damp sediments can remove nitrogen pollution even after the last puddles have disappeared.

The drying trend is clearest in arid regions, such as the Southwest. But even in the Southeast, which is relatively wet, streams are drying earlier and staying dry longer. In contrast, in the northern United States ephemeral rivers are now flowing longer. One possible reason: Winters are warmer and shorter, meaning frozen landscapes thaw earlier, allowing streams to flow.

Climate change is altering intermittent streams, such as this one in California's Death Valley.

PHOTO: TIM FITZHARRIS/MINDEN

In some cases, human activities such as operating dams, irrigation, and groundwater pumping could be contributing to dewatering. But a warming climate appears to be "the overarching organizer" of the shifts, Zipper says. "I definitely didn't expect the pattern to be so regionally clear."

Broader monitoring of intermittent streams would help researchers and policymakers better understand the sometimes subtle impacts that climate change is having on water quantity and quality, scientists say. "We should have many more gauges in small streams," says Albert Ruhi, a freshwater ecologist at the University of California, Berkeley.

Others say the results highlight the need for stronger legal protections for intermittent tributaries that form the headwaters of many rivers. Many such streams were excluded from federal environmental laws under former President Donald Trump's administration. (President Joe Biden's administration is now reviewing those exclusions.) Such streams can seem inconsequential, Wohl says, but, "If you start chopping off the first joint of each finger, you're going to lose functionality in your hand pretty fast." ■

ILLUSTRATION: KATTY HUERTAS

THE NEW NORMAL

# Science lost and lessons learned: A lab plots its comeback

## A microbiology team regroups, with a more virtual lab and a bigger focus on mental health

*By* **David Grimm**, *in Philadelphia*

The main door of Sunny Shin's lab is plastered with pictures of happier times: Shin photoshopped onto the cover of a Wheaties box, grad students chomping on corn cobs, a group photo on the lawn of a beach house. "We used to do a yearly retreat to the Jersey Shore," says Shin, a midcareer microbial immunologist here at the University of Pennsylvania (UPenn). "Hopefully we can go back in 2022."

When the COVID-19 pandemic hit, Shin oversaw 12 people: seven Ph.D. students, three postdocs, an undergrad, and a lab manager. She was thinking about each of them when the memo came down from an administrator in mid-March 2020: All non–COVID-19 work must stop, and most rodents must be culled because few people would be around to care for them. "It was heartbreaking," Shin said at the time, as her lab manager began the agonizing task of euthanizing 200 mice, some with unique genomes that had taken years to breed.

Shin was concerned for the future of her research on Legionnaires' disease—and, more importantly, for the future of her people. "I worry that the pandemic will affect the career trajectories of junior scientists for years to come," she says.

Some of those fears have come true, as universities across the country assess the damage of a lost year—research programs derailed, job opportunities vanished, and promising researchers lost to alternate careers. At the same time, "It's incumbent upon us to learn something valuable from this experience, and to use it to improve the lives of our students," says Daniel Kessler, chair of UPenn's cell and molecular biology graduate group. Among the legacies of the pandemic, UPenn and other schools are finding new ways to support trainees, with both their careers and their mental health.

When UPenn closed its labs and Shin's team faced 3 months of near-isolation, her immediate priority was their well-being. She dedicated the first 20 minutes of every (now virtual) lab meeting to how people were feeling. "Sunny tried to set everyone's mind at ease," says Tzvi Pollock, a fifth-year graduate student at the time. "She said everyone experiences setbacks, whether it be from a pandemic or Hurricane Sandy. That really reassured me."

Pollock says he needed the reassurance. During the pandemic, the depression and anxiety he had battled for years returned. He couldn't work or sleep. He had frequent panic attacks. "It absolutely wrecked me," he says.

He and his lab mates tried to write papers or plan experiments from home. But there was only so much they could do without access to mice and cell lines. Things didn't get much easier when UPenn reopened at half capacity last summer. Core resources such as tissue culture rooms and animal facilities filled up fast. Young graduate students like Víctor Vázquez Marrero had trouble shadowing older ones because they weren't always allowed in the same room. And Nawar Naseer, then a fourth-year grad student, chose to cram in a week of labwork between Friday and Sunday, when she had access to the facilities she needed. "Even though I was back in lab 50% of the time," Pollock says, "I was only about 30% productive."

The challenges continued at home. Shin's daughter was 6 years old when the pandemic struck. For months, Shin and her husband worked in shifts. About half of those who responded to a U.S. National Institutes of Health (NIH) survey in October 2020 said caretaking responsibilities made it "substantially more difficult to be productive," with women reporting more issues than men.

Meanwhile, nearly 70% of the postdocs surveyed said the pandemic would negatively impact their careers, with those caring for young children expressing the most anxiety. One of Shin's postdocs left academia, taking a job in industry that didn't involve bench research. "The COVID-19 pandemic is dismantling the pipeline of investigators who are essential to the future of biomedical science," fretted a recent article in *The New England Journal of Medicine*.

Worried about supporting her postdocs if she couldn't renew her grants, Shin took a 6-month sabbatical during which the university, rather than grants, covered most of her salary. But instead of taking that time off from her duties, she kept working and funneled the saved money to her postdocs. "My top priority was to keep the lab funded to buy them enough time to bounce back."

Her efforts joined a broader response. Universities like UPenn have allowed postdocs to apply for extensions, and young faculty to add a year to their tenure clocks. Meanwhile, in February NIH announced it

would extend many of the grants given to grad students, postdocs, and young faculty.

Still, Michael Lauer, NIH's deputy director for extramural research, acknowledges the agency won't be able to help everyone. "We're still in a hypercompetitive environment," for grants, a situation that predates the pandemic, he says. "If we give money to one person, that's less money for someone else."

In all, Shin estimates her lab was set back as much as 9 months. But things are slowly returning to normal. She brought in five undergrads, for a total of 15 people—more than before the pandemic. She was able to rebreed or reorder most of her mice. And her team is back to working full days in the lab.

Some pandemic retrofits will stick around. Lab meetings are in person again, but Shin still starts them by asking how everyone is doing. Lab members also now have the option of attending remotely, and about one-third do. "Some people just focus better in a virtual setting," Shin says. "Others feel more comfortable asking questions."

The lab has also embraced the instant messaging platform Slack as a way to keep in touch and check in on each other when they're not together. That situation is more typical now, as trainees have realized that they can analyze data and write papers from home. "The peer pressure to put in face time has abated," Shin says.

For its part, the university has put more emphasis on mental well-being. Pollock says that during the pandemic, "Penn had posters everywhere for all of the places you could call," but because of what he argues is still a large stigma around mental health, "people weren't availing themselves of those resources."

In response, the university has created a peer support network that allows students to reach out to friends and colleagues instead of going straight to a therapist. UPenn is also building a trainee advocacy alliance, which will partner students with peers and faculty trained in active listening and mentoring, who can help students navigate the university's mental health resources. "We want to normalize that if you need help, it's OK to ask," Kessler says. "When a student enters our school, they will now have access to multiple support systems from Day One."

Pollock is getting back to normal himself, feeling productive again and mentoring two undergrads this summer. He plans to stay in academic science, and when he gets a lab of his own he knows, now more than ever, whom he wants to emulate. "Over the past year, it's been so clear that Sunny cares more about me as a person than as a producer of data," he says. "I want to be the kind of mentor she was to me." ■

Steven Chu, who was then U.S. secretary of energy, visits the Joint Center for Artificial Photosynthesis in 2012.

PHOTO: U.S. DEPARTMENT OF ENERGY/FLICKR

ENERGY RESEARCH

# 'Mini–Manhattan Projects' for energy innovation wind down

## But hub model for bridging basic and applied research lives on

*By* **Adrian Cho**

The Department of Energy (DOE) will soon wipe away a legacy of Steven Chu, the Nobel Prize–winning physicist who served as secretary of energy from 2009 to 2013 under former President Barack Obama. According to the department's budget request for next year, DOE intends to wind down most of its Energy Innovation Hubs, multidisciplinary, multi-institutional centers that Chu devised to solve crucial energy-related problems and invigorate the sclerotic department.

Chu compared the hubs to the Manhattan Project, the World War II scramble to make an atomic bomb, and like the bomb project, they were meant to be ad hoc, temporary efforts. Some DOE bureaucrats disliked the way the hubs crossed organizational boundaries, but observers say they succeeded in making DOE's research more responsive and relevant. "The vision for the hub was, and still is, a great one," says Eric Isaacs, president of the Carnegie Institution for Science and former director of DOE's Argonne National Laboratory. In fact, DOE appears to have embraced the once-controversial model and has started several new projects that hew to it. "They look like a hub, and they walk like a hub, but they don't have this unfortunate malodorous name," says Alex King, a materials scientist retired from DOE's Ames Laboratory.

Chu, a former president of AAAS, which publishes *Science*, borrowed the basic parameters for the hubs from three bioenergy research centers started by DOE under former President George W. Bush. Each hub would receive $25 million a year for 5 years, with the possibility of a renewal. Instead of focusing on a research topic, each would strive to develop a practical solution for a single big problem, Chu said, uniting "under one roof" everybody from scientists doing basic research to engineers developing a prototype. By 2013, DOE had initiated five hubs focused on challenges ranging from converting sunlight to fuel to modeling nuclear reactors to improve their performance.

The whole point of the hubs was to breach a long-standing boundary within DOE, says Cherry Murray, a physicist at the University of Arizona and director of DOE's Office of Science from 2013 to 2015. DOE's basic and applied research are disconnected because they're funded out of different congressional budget lines. The two rub elbows at DOE's 17 national laboratories, but "the interface isn't perfect," Murray says. "So the hubs are just trying to bring that [connection] into a funding mechanism" to drive the innovation of new technologies.

Each hub was managed by either the Office of Science or an applied office, and each was encouraged to have industrial collaborators who could make sure what scientists were doing was relevant, says Bill Madia, vice president emeritus at Stanford University. "That was a little bit outrageous," he says. "You will never see the Office of Science put out a solicitation [for an ordinary grant] saying, 'You better bring in GE for a cost-sharing arrangement.'"

One hub quickly crashed and burned. The Energy Efficient Buildings Hub in Philadelphia shut down in 2013 when Congress pulled the plug, citing management issues. Another nailed its goals. The Consortium for Advanced Simulation of Light Water Reactors (CASL), based at Oak Ridge National Laboratory, produced high-resolution, physics-based software that industry has used to simulate new and existing reactors.

The Joint Center for Energy Storage Research (JCESR) at Argonne set out to increase the energy density of batteries by a factor of five, compared with the lithium-ion battery that powered the 2012 Nissan Leaf, at one-fifth the cost. "Spoiler alert, we got three times the energy density at a fifth the cost" in four novel batteries, says George Crabtree, a materials scientist at Argonne and director of JCESR. The hub also spun out multiple startups, he says, including one that is building an iron-based grid storage battery for an electric utility in Minnesota.

> **"The vision for the hub was, and still is, a great one."**
>
> **Eric Isaacs**, Carnegie Institution for Science

The most ambitious of the hubs illustrated the key challenge: to keep a hub from losing sight of its goal and morphing into just another research center. The Joint Center for Artificial Photosynthesis (JCAP), based at both the California Institute of Technology (Caltech) and Lawrence Berkeley National Laboratory, aimed to make a self-contained photocell that would harness sunlight to convert water and carbon dioxide into fuels with an efficiency of 10%, 10 times that of natural photosynthesis.

JCAP researchers surpassed that goal with a prototype that converts water to hydrogen gas, ultimately reaching an efficiency of 19.3%, says Harry Atwater, an applied physicist at Caltech and JCAP director. But the cell wasn't durable enough to be practical, Atwater says, and JCAP turned to the harder problem of transforming carbon dioxide into specific hydrocarbons such as ethylene, which it has not yet solved. "JCAP was too aspirational and too science-y," Murray says.

External pressures can also alter a hub's character. The Critical Materials Institute (CMI) at Ames Laboratory aimed to find replacements for certain materials, mostly rare earth elements, or new sources of them. "Right out of the gate, we had a really good run of successes in developing technologies and transferring them to industry," says King, who directed CMI from 2013 to 2018. For example, he says, CMI developed a red phosphor for fluorescent lighting that does not require rare europium.

In 2015, however, Republicans took control of the Senate, and discouraged DOE from doing research they thought industry should do for itself. Industrial partners also became reluctant to work openly with CMI, King says, explaining they feared a public relations problem. "If you're working with [CMI], it must mean you have a supply chain problem," he says, "and that's going to affect your stock." As a result, CMI began to lose its focus and to resemble a more typical research center, King says.

Now, DOE is winding down Chu's hubs. CASL closed down last year, and JCAP and CMI are wrapping up. JCESR with receive its last funding next year, according to the DOE budget request. (Crabtree says his understanding is that JCESR will be funded into 2023.)

Yet DOE is hardly abandoning the hub concept. Last year, the department kicked off the National Alliance for Water Innovation, which aims to develop "technology that enables 90% of [waste] waters to be reused," says Meagan Mauter, a chemical and environmental engineer at Stanford and the alliance's research director. DOE has also started the hublike Liquid Sunlight Alliance at Caltech and the Center for Hybrid Approaches in Solar Energy to Liquid Fuels at the University of North Carolina, Chapel Hill, to follow on JCAP's work.

DOE has even extended the hubs concept beyond energy research. Last year, the department initiated five quantum information science centers, each funded for $25 million a year for 5 years and aimed at developing a particular quantum technology, such as a quantum internet. "We took [the hubs concept] and turbocharged it" by encouraging even closer ties with industry, says Paul Dabbar, who helped launch the quantum centers when he was DOE's undersecretary for science from 2017 to 2021 under former President Donald Trump.

Whether the Biden administration will be as enthusiastic for hubs remains to be seen, as the Senate has yet to confirm the White House's nominees for director of the Office of Science and undersecretary of energy. But, for now, even as the hubs' name dies, the concept has found new life. ■

RESEARCH ETHICS

# Genetics papers from China face ethical scrutiny

## Questions about consent and potential for abuse trigger investigations

*By* **Dennis Normile**

When Yves Moreau, a bioinformatician at KU Leuven in Belgium, noticed a 2017 paper in *Human Genetics* that described the "male genetic landscape of China" based on a set of almost 38,000 Y-STR sequences, he saw a red flag. Y-STR stands for Y-chromosomal short tandem repeat polymorphism, bits of repetitive DNA often used in forensic investigations. Some of the samples came from Uyghurs and other minorities in China, and Moreau was skeptical that they had given informed consent for the use of their genetic data or understood that China might use it to profile their people. In June 2020, he asked the journal's editors to retract the "indefensible" paper.

Springer Nature, its publisher, launched an investigation that is still ongoing. So last month, Moreau stepped up the pressure: He wrote to the journal's entire editorial board to complain about the lack of progress. For Moreau, the paper is just one of many studies, primarily in forensic genetics, that deserve scrutiny because of consent problems in China and the potential for abuse of the data. He says he has flagged about 28 papers at six journals over the past couple of years.

And his campaign is gaining traction. Eight of 25 members of the editorial board of *Molecular Genetics & Genomic Medicine*, published by Wiley, recently resigned to protest the lack of progress in investigating a number of papers flagged by Moreau, as *The Intercept* reported last week. A former editor-in-chief of *Human Genetics*, geneticist Robert Nussbaum, has added his voice to Moreau's, complaining to the editors that the investigation of the 2017 paper "seems to have been going on a long time." Springer Nature's executive editor for medicine and life sciences, Andrea Pillmann, says it is investigating about 50 other papers, 29 of

which already have an editor's note of concern attached to them. The company has put checks in place "to help us to identify potentially concerning submissions in future," Pillmann says. Meanwhile, the Charité University Hospital in Berlin has come under fire for hosting the genetic database used in several papers under investigation.

Human rights activists welcome Moreau's efforts. "It is important for journals concerned with research ethics to take into account the state violence that is endemic throughout the Uyghur and Tibetan regions," says Darren Byler, a sociocultural anthropologist who studies Uyghur issues at Simon Fraser University, Vancouver.

A Uyghur man sits in a teahouse in China's Xinjiang province, where government surveillance is intense.

Moreau has long been concerned about threats to privacy posed by the use of genetic data. Forensic use of DNA databases has evolved from a narrowly focused law enforcement tool to a threat to personal privacy, he says. The potential for abuse is, at the moment, most clearly seen in China's Xinjiang Uyghur Autonomous Region, he adds. Since late 2017, evidence has grown that China is systematically oppressing Uyghur and other Muslim minorities in Xinjiang. Some call the tactics—including mass internment, forced labor, suppression of religion, and efforts to slash birth rates—crimes against humanity. China claims the country is simply "countering violent terrorism and separatism," as Foreign Minister Wang Yi told the United Nations Human Rights Council in February.

As part of surveillance efforts in Xinjiang, authorities have collected biometric data, including facial scans and fingerprints—used to verify identities at the region's ubiquitous checkpoints—and DNA data. Moreau says DNA profiling does not directly enable mass internments or forced labor. Rather the impact is psychological, reinforcing citizens' feelings of constant surveillance. Byler adds that DNA profiling "could be used to enforce a 'zero illegal births' policy by tracking maternity and paternity," and to find matches for organ harvesting from prisoners, "of which there is some limited evidence."

As authorities have tightened their grip on Xinjiang, Chinese researchers have stepped up research into the region's culture and genetics, says Huang Futao, a higher education scholar at Hiroshima University. Some of this research is supported by the State Administration for Science, Technology and Industry for National Defense, Huang says, and focuses on topics related to maintaining social safety and stability. The results have often been published in international journals.

Yet, "It can be very difficult to judge the validity of informed consent in China," says bioethicist Jing-Bao Nie of the University of Otago, Dunedin: "Explicit and especially implicit pressure [to give consent] often exists in various forms."

Moreau's efforts have already led to the retraction of two papers by Chinese authors in Springer Nature's *International Journal of Legal Medicine*. One, a study of STR haplotypes in Uyghur, Kazakh, and Hui men published online in March 2019, was found to have been "undertaken without the approval of [the authors'] institutional ethics committee," according to the May retraction notice. A review by the editors found that the second, an evaluation of genetic markers in four different Chinese populations published in April 2019, had ethics approval for the participation of Chinese Han individuals, but not for the Tibetan, Uyghur, and Hui participants. It was retracted in November 2020.

The *Human Genetics* paper is particularly problematic because of the sheer volume of data and the participation of coauthors from Chinese law enforcement organizations in the study, Moreau says. The paper, co-authored by Chinese and German researchers, states that the study "complies with the ethical principles of the 2000 Helsinki Declaration," which covers research in humans. But corresponding author Michael Nothnagel of the Cologne Center for Genomics now concedes that some of the data may have been collected in ways "that did not meet relevant ethical standards." Nothnagel says the authors are working with the editors and Springer Nature to resolve the issue; he did not respond to an email seeking additional details.

The data used in the paper are drawn from the Y chromosome haplotype reference database (YHRD), based at Charité, which includes Y chromosome data from more than 300,000 individuals uploaded by contributors from around the world and is used by researchers and law enforcement agencies. Moreau says there is no way to verify compliance with informed consent standards for the data, which are at least partly anonymized.

In a letter posted on its website on 6 August, Gen-ethical Network, a Berlin-based organization promoting ethical use of genetic technologies, called for an investigation into allegations of YHRD's "unethical handling of genetic data from minorities." The letter, cosigned by three other groups, cited not only the issues raised by Moreau about Uyghurs, but also similar problems with genetic research on Roma people in Europe. YHRD managers did not respond to an email from *Science* seeking comment.

Moreau's concerns go beyond informed consent. Any research that enables genetic profiling "is harmful in the hands of an authoritarian regime," he says. And he worries such studies reflect badly on the field: "Public trust in human genetics depends on our community's ability to transparently abide by its moral duties."

Nie sees little chance of change on the ground in China, where rising nationalism often eclipses ethical concerns. "I doubt that the issue of informed consent and privacy will be improved in the near future in China," Nie says. That puts a greater responsibility on international journals, Moreau says. ■

PHOTO: KEVIN FRAYER/GETTY IMAGES

FEATURES

# FAILURE TO PROTECT?

## A study of asthmatic children, most of them Black, shows how a common clinical trial design can expose vulnerable participants to serious risks

*By* **Charles Piller**

Juan Celedón, a respected pulmonary researcher at the University of Pittsburgh, wanted to address an urgent national problem. Severe asthma attacks send hundreds of thousands of U.S. children to the hospital every year. For decades, researchers have suspected extra vitamin D—essential for bone growth and healthy development, and also an immune modulator in children and adults—might help them. In 2016, Celedón and colleagues launched a major trial to test that premise.

With $4.3 million in funding from the National Heart, Lung, and Blood Institute (NHLBI) and support from a vitamin company and a drug firm, they enrolled asthmatic kids who had low or deficient levels of vitamin D—many from urban, minority communities; most were Black. Half of the 400 planned participants would receive a daily high-dose vitamin D supplement for about 1 year. The other half would serve as controls. The researchers also made a decision that cast a shadow over the trial—and inflamed a controversy confronting many other trials of similar design. Instead of treating the randomly chosen control group with a more modest dose of vitamin D—as many medical authorities advise for children with a deficiency of the vitamin—the researchers chose to give them a placebo.

When Seattle physician Bruce Davidson,

ILLUSTRATION: STEPHAN SCHMITZ

a pulmonary specialist who has studied vitamin D and asthma, heard about the "Vit-D-Kids" trial, he was stunned. The children, as asthmatics treated with corticosteroids, already faced possible bone health problems and diminished growth, and any vitamin D deficiency would place them at greater risk for fractures. Yet Davidson discovered that informed consent forms posted online failed to inform parents of enrolled children about those dangers.

Davidson, who had worked on a comparable vitamin D trial that rejected a placebo arm as unethical, repeatedly voiced his concerns about Vit-D-Kids to the scientists running it, institutional review boards (IRBs) that approved it, and NHLBI. But trial researchers called the risks minimal. The placebo was justified because vitamin D testing is not routine, they argued. If not for the trial, the kids' vitamin deficiency probably wouldn't have been detected, so they were no worse off in the study.

Davidson's objections drew some media attention during the trial and led to small changes to its enrollment criteria and consent forms, but Vit-D-Kids pressed ahead. It wasn't a success. Trial enrollment grew to nearly 200 children but was halted early for "futility" in 2019 after a data safety monitoring board (DSMB) concluded during an interim review of results that vitamin D supplementation had failed to prevent asthma attacks. Yet the researchers kept an unspecified number of kids, even if very deficient in the vitamin, on a placebo for up to six more months—to ensure "an orderly closeout," James Kiley, director of NHLBI's Division of Lung Diseases, later told a U.S. politician who asked about the decision.

"That approach was stunning and callous," Davidson says. And possibly harmful. At least nine kids, across both arms of the trial, broke bones during the trial—nearly double the number expected among asthmatic children over a comparable period. The fractures were neither disclosed as possible adverse events when the study was published in *JAMA* last year nor noted in another public summary of the trial results.

A *Science* investigation of Vit-D-Kids reviewed thousands of pages of trial protocols and consent forms; previously undisclosed DSMB deliberations; emails from the trial's principal investigator, NHLBI officials, and *JAMA* editors; and letters between National Institutes of Health (NIH) officials and a concerned member of Congress. Those documents and others reveal new aspects of the trial that concern asthma researchers like Davidson, medical ethicists, and specialists in clinical trial design who reviewed the materials at *Science*'s request.

"The advantages to society of this trial, because of the poor design, were likely none. And the risks did outweigh the benefits," says Charles Natanson, a physician and expert on trial design at the NIH Clinical Center. "This trial did not, in my opinion, meet the standards set forth in *The Belmont Report*," which in 1979 established ethical guidelines, adopted by the U.S. government, for protecting human subjects. Keeping children on a placebo after the trial was stopped for futility stood out as an "unconscionable" error, adds Ruth Macklin, a medical ethicist at Albert Einstein College of Medicine.

Davidson and others suggest the study's focus on minority children—which Kiley called "appropriate" in a statement to *Science*—only elevates their concerns about using a placebo. Jill Fisher of the University of North Carolina, Chapel Hill, who wrote a book on racial inequality in clinical trials, says Vit-D-Kids ignored the ethical imperative to treat children at known risk from vitamin D deficiency because of inadequate diet, poverty, and a lack of Sun exposure in inner cities. "We shouldn't say, 'It's unfortunate that there are these health and nutritional disparities, but it serves the interests of science to have placebo-controlled trials,'" she says. It is "structural racism" to scientifically exploit such inequalities, Fisher adds.

Kiley and Celedón declined to be interviewed but provided statements after being sent a list of questions. The trial, its protocol, and consent forms "underwent rigorous review before and after it was funded," Celedón wrote in a brief note emailed to *Science*, adding: "All regulatory bodies, including a [DSMB] and [IRBs] at seven pediatric hospitals, stated that the trial was both safe and ethical." In NHLBI's statement to *Science*, Kiley wrote, "The highest priority was to keep every child in the study safe."

Vit-D-Kids could easily be dismissed as a controversial outlier, but Natanson and others suggest it instead exemplifies the growing number of studies in humans that inappropriately reject control groups receiving "usual care"—current best practice treatments used by doctors. In a hunt for compelling results, many researchers favor using sharply divergent treatment arms in a trial. But such extreme comparisons mean doctors can't learn whether a new treatment is better or worse than usual care, says Natanson, who has analyzed the issue in trials involving critically ill patients. He has found that many such trials mislabel one or more arms as "usual care," sometimes endangering participants and misinforming physicians, which he calls "a big problem."

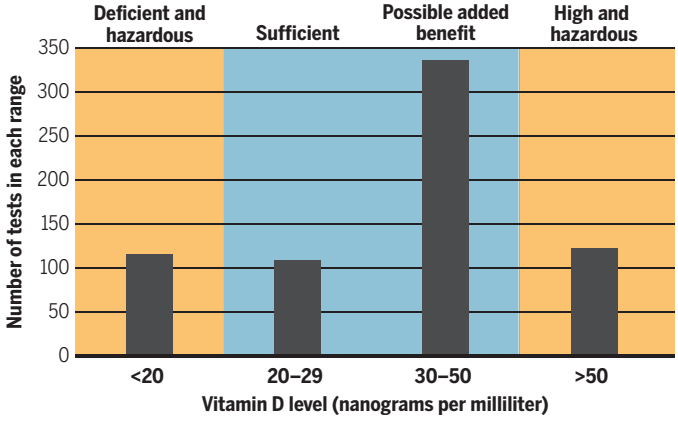
**MORE THAN A DECADE AGO**, NIH supported an ambitious trial that foreshadowed the usual-care issue in Vit-D-Kids. The study's researchers meant to solve a life-and-death medical conundrum affecting premature infants: They depend on supplemental oxygen to stay alive, but too much can cause blindness. The trial aimed to identify an oxygen level that would save lives with the fewest side effects.

It assigned more than 1300 preemies to be maintained at a relatively low blood oxygen range (85% to 89%) or a higher range (91% to 95%). Most consent forms said either range represented "usual" or "standard" care and that babies in both groups had the same likelihood of dying. Prominent supporters, including NIH Director Francis Collins, argued that both infant groups met the standard of care as practiced at trial sites.

But in a 2016 analysis in *PLOS ONE*, whose authors included Natanson, researchers examined other trials of oxygen management in preemies and concluded the bottom of the trial's low-oxygen range was not considered usual care in multiple countries. The trial departed from usual care in another respect, not noted on the consent forms: By design, the oxygen monitors displayed

## Hazard zones

The Vit-D-Kids trial tested children's vitamin D blood levels multiple times over the course of about 1 year. This chart categorizes readings from all 683 tests, estimated from text and results graphs in its published report. In more than one-third of the tests, kids had levels that were potentially hazardous according to the National Institutes of Health and other authorities.




CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) E. FORNO *ET AL.*, *JAMA*. 324(8):752 (2020;) 10.1001/JAMA.2020.12384

inaccurate readings to prevent caregivers from knowing a baby's study group.

Scores of bioethicists and clinicians—and the federal Office for Human Research Protections (OHRP)—said the informed consent forms inadequately described the risks. About 20% of babies in the low-oxygen range died, compared with 16% in the high-oxygen group.

Several other U.S. trials (see table, below) also became magnets for criticism that they violated usual-care safeguards seen as crucial, even if sometimes complex to define. Each trial had eminent backers who said accepted practices can be ambiguous, fueling divisive debates. "Usual care in Seattle may differ from usual care in San Antonio. This applies particularly to uses of technology and high-cost interventions," said Edward Campion, an editor at *The New England Journal of Medicine* when it published the infant oxygen study, at a public hearing.

Those issues came to a head again in Vit-D-Kids. Vitamin D supplementation has long been contentious. It is said to be a remedy for diabetes, cancer, heart disease, and other ailments, but clinical trials often failed to show such benefits, particularly for the high-dose pills touted by the supplement industry. How much vitamin D a growing child needs also is debated, but most authorities say kids need levels in the blood of 20 to 29 nanograms per milliliter (ng/ml) to minimize risks of bone fractures and impaired immune response, and to protect lifelong skeletal health. Guidelines from the American Academy of Pediatrics and Pediatric Endocrine Society call anything below that threshold "deficient" or "insufficient" and recommend supplements. The Vit-D-Kids protocol also cites an Institute of Medicine report that agrees with those assessments. And NIH labels 20 ng/ml "inadequate" and below 12 ng/ml "deficient."

With sites at big-city hospitals, Vit-D-Kids originally recruited asthmatic kids, ages 6 to 16, who had vitamin D levels between 10 and 29 ng/ml. Many kids were below 20 ng/ml—levels the study itself, in its protocol, deemed "deficient" and inadequate for musculoskeletal health. Yet that protocol, posted publicly online and included with the *JAMA* publication, also described the risk of leaving those children untreated as "no greater than encountered in daily life by healthy community-dwelling children." But the kids participating in the study, afflicted with asthma and receiving powerful steroid drugs to treat it, were far from healthy.

In his statement to *Science*, Kiley defended Vit-D-Kids by saying vitamin D–deficient children were excluded in favor of those with "vitamin D levels in the low to low-normal range, who normally would not be treated with supplemental vitamin D." He cited a 2016 global consensus report on rickets, a condition linked to vitamin D deficiency that causes soft bones and bow legs. Yet even the rickets report defined vitamin D levels of 12 to 20 ng/ml as "insufficient."

Vit-D-Kids compared inadequate treatment and overtreatment, according to Natanson. Kids in the treatment arm were given daily vitamin D supplements of 4000 international units, or seven times their recommended daily allowance, and some reached serum levels above 100 ng/ml. NIH guidelines say anything above 50 ng/ml is potentially hazardous; studies have suggested such levels encourage some cancers and cardiovascular problems or increase risk of death overall.

The trial protocol noted the high-dose supplement was tested against a placebo to avoid a "false negative" outcome. "They wanted to maximize their chances of finding a difference," says physician Michael Carome, a former top regulator at OHRP who directs health research for the consumer advocacy group Public Citizen.

## 'Usual-care' controversies

Trials that compare alternate treatments but lack a usual-care group reflecting standard medical practice have grown increasingly common—and drawn a backlash from ethicists and clinical research experts. They say the approach sometimes endangers trial participants and may not offer clear clinical guidance. The trials below are among those that drew scrutiny.

**Trial:** Surfactant Positive Airway Pressure and Pulse Oximetry Randomized Trial, 2005–09

**Sponsor:** National Institute of Child Health and Human Development

**Goal:** Determine the optimum oxygen level for premature infants.

**Controversy:** Critics charged that the trial paired high levels of oxygen with levels below usual care and set oxygen monitors to provide false results—misleading parents and clinicians about the risk of death. Supporters argued that evidence to define usual care was unclear and both groups fell within that standard.

---

**Trial:** Established Status Epilepticus Treatment Trial, 2016–19

**Sponsor:** National Institute of Neurological Disorders and Stroke

**Goal:** Determine which of three anticonvulsant drugs most effectively treats seizures lasting more than 5 minutes.

**Controversy:** Critics said the study endangered subjects by randomly assigning them to a drug, when under usual care each patient's history would be a key consideration. Because the researchers were initially blinded as to which drug was given, it was difficult for them to adjust treatment if patients did not respond. Researchers argued the trial was essential to determine objectively which drug worked best.

---

**Trial:** Crystalloid Liberal or Vasopressors Early Resuscitation in Sepsis Trial, ongoing

**Sponsor:** National Heart, Lung, and Blood Institute (NHLBI)

**Goal:** Determine optimal treatment—drugs that constrict blood vessels plus limited fluids, or a larger volume of fluids alone—for people with life-threatening septic shock.

**Controversy:** Critics said both treatments deviated sharply from usual care for septic patients, increasing participants' risk of death. Trial sponsors said they fell within the accepted range of care.

---

**Trial:** Myocardial Ischemia and Transfusion, ongoing

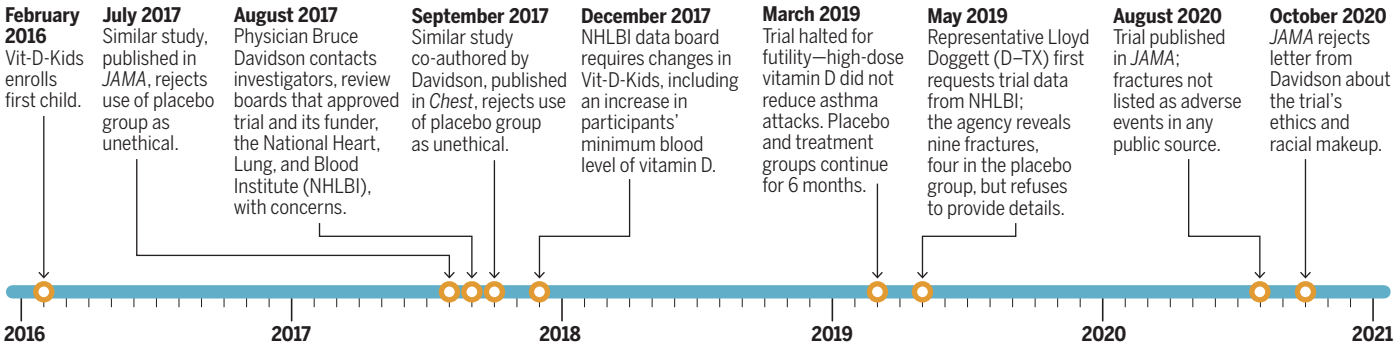
**Sponsor:** NHLBI

**Goal:** Determine the optimal blood transfusion strategy for heart attack patients with anemia, using red blood cell levels to decide when more blood is needed.

**Controversy:** Critics said informed consent forms misled participants and the trial placed those in the more restricted transfusion group at greater risk. Supporters said the trial would help resolve uncertainty about the best approach.

## Risky choices, rejected warnings

About 18 months after the Vit-D-Kids trial began, its researchers were challenged about the study's ethical approach and methodology, which critics said left many vulnerable children deficient in vitamin D, an essential nutrient.




*Science* also reviewed versions of the informed consent forms used by Vit-D-Kids, some of which Davidson acquired through a Freedom of Information Act (FOIA) request. Experts in trial design say those forms stressed potential benefits over harms and were overly complex and confusing.

For example, instead of discussing vitamin D deficiency, the forms used the more benign-sounding term "low vitamin D," says Columbia University cardiologist Raymond Givens, who studies institutional racism in medicine and medical publishing. Ethicist Harriet Washington, whose book *Medical Apartheid* discusses clinical experiments on Black Americans, also notes parents often misunderstand essential research terms.

The informed consent forms for Vit-D-Kids called rickets, not bone fractures, the primary risk for children who received placebos. But rickets affects infants and very young children—far younger than those enrolled in Vit-D-Kids. The consent form should have clearly stated that any child in the placebo group with insufficient vitamin D would face a higher risk of broken bones, says Boston University medical ethicist George Annas. But if such an explicit concern had been noted, he suspects few parents would have signed the consent form because "it almost sounds like child abuse."

In his statement, Kiley wrote to *Science* that the fracture risk from too little vitamin D didn't apply to the trial's kids because they used only inhaled steroids, not injected or oral forms, which in adults cause bone to demineralize. But up to 14 children on the placebo took systemic steroids at least twice during the trial—enough to increase the fracture risk, according to an authoritative asthma study.

**AFTER LEARNING** many details of Vit-D-Kids, Davidson in August 2017 sought advice from Frank Greer, an emeritus professor of pediatrics at the University of Wisconsin, Madison, and member of the Vit-D-Kids DSMB. In an email Davidson gave to *Science*, Greer expressed alarm about giving a placebo and concerns about possible adverse effects in the high-dose group. (NHLBI recused Greer from discussions of the matter after learning of Davidson's communication with him.)

Davidson then complained to NHLBI officials. "[Davidson] makes very valid points and … [the investigators] need to address this in a very substantive way," Kiley subsequently wrote to colleagues in an email Davidson obtained via a FOIA request. "I am inclined to put a clinical hold on this study if we cannot get [a DSMB] review done next week."

After months of deliberation, the DSMB in early 2018 approved changes to the trial, excluding any future enrollees with vitamin D levels below 14 ng/ml, compared with the prior cutoff of 10 ng/ml, and adding new wording to the consent form. The revised version said, "The risk to bone health is unclear if the vitamin D level is 14–19 ng/ml. However, many doctors would treat with vitamin D at this level." The board acted "out of an abundance of caution," Kiley wrote in his recent statement.

Carome calls the revisions a tacit acknowledgment that the consent form used to recruit many of the kids "failed to disclose important information that parents would have needed to make a fully informed decision." He says the new form continued to obfuscate risk by implying that treating vitamin D deficiency was a gray area. Moreover, parents were never told what their children's specific vitamin D levels were, either at enrollment or later in the trial. Care decisions, including whether to supplement a kid's vitamin D on their own, were effectively out of their hands.

**WHEN *JAMA* PUBLISHED** the results of Vit-D-Kids in August 2020, it looked like just another failed vitamin supplement trial. The placebo group and treatment groups experienced about the same number of adverse events—mainly asthma-triggered hospitalizations or emergency department visits, according to the paper. *JAMA* had in 2018 rejected a commentary Davidson submitted criticizing the trial's use of a placebo arm; after the study's publication, he sent a letter to the editor outlining that concern and questions about the trial's racial mix. *JAMA* rejected that as well.

*JAMA*'s editors declined interviews, but wrote in a statement they were "aware" of Davidson's concerns and that the study was designed to "ensure the safety" of participants. They noted the paper reported that trial investigators stopped giving placebos to several kids, and referred them to an endocrinologist, after their vitamin D levels dropped below the study's minimum requirement.

But Celedón and colleagues did not report in the *JAMA* paper, or in results posted to ClinicalTrials.gov, that kids in the trial broke bones. Kiley disclosed that at least nine fractures had occurred only when Representative Lloyd Doggett (D–TX), who chairs an NIH oversight panel and had been contacted by Davidson, asked about the trial.

Five fractures had occurred among kids given vitamin D and four in the placebo group, which is nearly twice the rate expected for asthma sufferers in that age group. But Kiley told Doggett that the trial's oversight board found no safety issues. (In an interview, DSMB chair Dennis Ownby of Augusta University said his panel was told that the rate of fractures was very low, but does not recall independently verifying that information.) Kiley refused to share details of those fractures with Doggett for unspecified "reasons of patient privacy and scientific integrity."

Davidson, Natanson, and other trial design experts *Science* contacted were troubled that Vit-D-Kids failed to report the fractures and their details publicly. Although the comparable number of fractures in each arm might suggest the placebo group was not at greater risk, they note the breaks are impossible to interpret without information on their severity and precipitating events, such as contact

GRAPHIC: N. DESAI/*SCIENCE*

sports versus low-stress activities.

"Letting kids spend 48 weeks as low as 10 ng/ml—and doing this primarily to minority kids while warning them about irrelevant rickets instead, then covering up bone fractures—is awful," Davidson says. "Those kids with the lower vitamin D levels need to be located, carefully assessed, and treated if necessary. They need explanations and oversight for a while to minimize their future risk."

**THE MAKEUP** of the 192 trial participants, which included 100 Black kids, intensified the debate. Kiley in his statement defends the trial's demographics, noting that Black children "are twice as likely as white children to be affected by asthma." Yet even with higher asthma rates, Black children were vastly overrepresented in the trial, in comparison with their numbers at study sites like Pittsburgh and Boston.

That concerns Givens and others. "[If] most of [a trial's] subjects will be nonwhite, and a large proportion low-income and perhaps lacking advanced education among parents, there is a need for heightened attention to the ethics and appropriateness of the trial," he says, including intensive community outreach before recruitment in poor or minority populations.

Macklin calls it "surprising—if not appalling—that IRBs in these major U.S. medical centers are willing to approve studies in which disadvantaged children are randomly assigned to a placebo group, justified by the argument that they are not being made worse off than they would be if not enrolled in the study." Annas compares Vit-D-Kids to "no-care as usual care" trials in resource-poor nations, such as NIH-funded trials in African nations in the 1990s. In pregnant women, the antiviral zidovudine, also known as AZT, was tested against a placebo to see whether it blocked mother-to-child HIV transmission. The drug was already the standard of care during pregnancy for HIV-infected women elsewhere.

Washington said she was dismayed to learn the trial protocol and consent forms prominently discussed analyzing kids' genes for clues connecting low vitamin D to asthma but glossed over obvious inner-city risk factors such as air pollution, stress, and poverty, which could also lead to vitamin deficiency and asthma. The hunt for genetic explanations above social linkages "reinforces the belief that … biological dimorphism drives a lot of illnesses," she says. "It's unethical. It's deeply illogical. And it fits a racist pattern."

Kiley and Celedón declined to comment on that issue. Ownby says he also sees asthma as a disease "primarily of lower socioeconomic status," closely linked to disadvantages experienced by Black people and other people of color and to the social issues Washington notes. But he adds that "to try to look at them all at once just isn't within the scope of most NIH grant studies. There's just not enough money." Studying genetic questions, particularly for those most affected by asthma, is also important and ethical, Ownby says.

To Carome, Vit-D-Kids offers new evidence that, overall, "our IRB system is broken." He doubts that any hospital panel that greenlit the study asked how its own physicians would normally treat asthmatic children deficient in vitamin D. None would fail to give baseline supplements, he says. It's a sign that IRBs tend to uncritically accept NIH funding as a stamp of approval, Carome adds.

Critics of Vit-D-Kids say the study, though an extreme case, also points to troubling trends amid an explosion of comparative effectiveness research (CER) trials, which examine the benefits and harms of treatments. U.S. funding for CER clinical trials and related trial support rocketed from an annual average of about $34 million between 2009 and 2011 to $284 million since then—largely because of the government-created nonprofit Patient-Centered Outcomes Research Institute (PCORI), which specializes in assessing treatments side by side.

Meanwhile, criticisms of such studies have mounted. In 2018 Public Citizen challenged what it called "reckless" flaws in an NIH-backed study of treatments for septic shock, a life-threatening effect of infection. The group said the trial randomly assigned subjects to combinations of fluids and drugs in ways that departed sharply from usual care tailored to each patient's condition. In a public statement, Carome called patients in the ongoing trial "unwitting guinea pigs in a physiology experiment that will not advance medical care for sepsis and likely will harm many."

Trial organizers challenged that verdict but last year OHRP and NHLBI forced changes in the trial's protocol to allow individualized care—improvements Public Citizen commended but called insufficient. (At the time, NIH barred Natanson and another NIH scientist from commenting on the sepsis trial.)

Natanson recently analyzed CER trials of critically ill people published over 1 year in three "high-impact journals with a reputation for a rigorous review process—the best clinical trial journals in the country." The yet-unpublished study examined roughly 50 trials, identifying those that improperly excluded a usual-care arm. He estimates that "up to half of the [CER] trials done in critically ill subjects have this problem." A huge proportion of trials of nonemergency interventions, like Vit-D-Kids, also excludes a usual-care arm, including more than 70% of PCORI's CER trials. Natanson says scientific results, patient safety, and the informed consent process all suffer when a usual-care comparator is absent. But he acknowledges that trial funders and IRBs struggle when no bright line differentiates experimental interventions from usual care. Making those distinctions can require laborious observational studies and surveys.

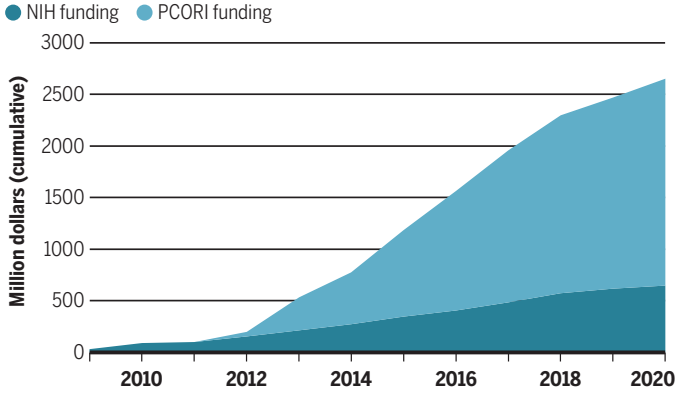
PCORI itself says usual-care comparators are often important or necessary. But given frequent challenges in defining standard practice, the group actively discourages usual-care arms—"unless they represent legitimate and coherent clinical options." That's an abdication, Natanson says. "It's much easier to say, 'I have two ideas, two strategies.' … It's much more difficult to say, 'What is usual care? How is it practiced? How can I design the trial so that … at least one arm is receiving usual care?'"

A real-world benchmark can be essential to evidence-based medicine—whether a trial examines oxygen levels for preemies or vitamin D for asthmatic kids, Natanson adds. "It's just common sense. Why study two things inside of a trial that nobody does outside of the trial?" ■

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

## Treatments on trial

National Institutes of Health (NIH) funding for clinical trials comparing different treatments rose starting in 2009. Spending on such comparative effective research increased sharply with the creation of the Patient-Centered Outcomes Research Institute (PCORI).




CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) PCORI/NIH REPORTER

ILLUSTRATION: ESA/HUBBLE, M. KORNMESSER/CC BY 4.0

ASTRONOMY

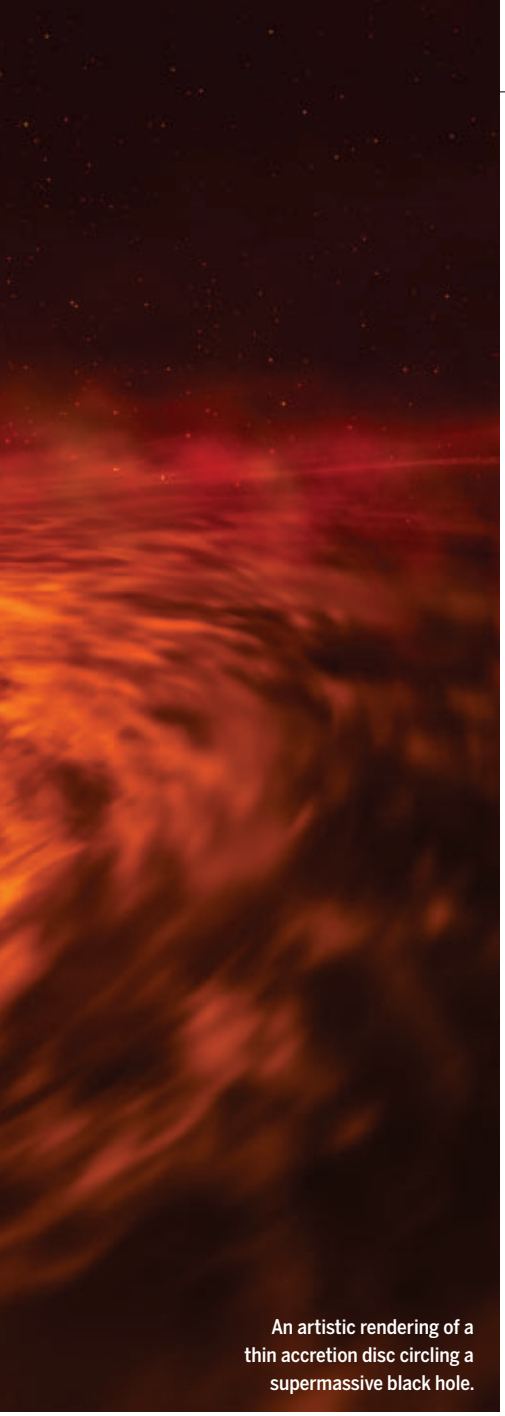
# How massive is that black hole?

## The flux of radiation emissions from accretion disks correlates with black hole mass

*By* **Paulina Lira**[1] *and* **Patricia Arevalo**[2]

A black hole is a point in space that is a cosmic sink—the gravitational attraction is so strong that not even light can escape. A supermassive black hole (SMBH) is enormous, with a mass on the order of millions to billions of times that of the Sun (*1*). It's not clear how such an entity arises. It could be the result of the merger of smaller black holes or the collapse of either a stellar cluster or large gas clouds. Every large galaxy is thought to contain a SMBH. To understand how they arise, we need to know how massive they are. On page 789 of this issue, Burke *et al.* (*2*) describe a method to make this determination on the basis of radiation emissions from the accretion disks of SMBHs. The approach also shows a connection between SMBHs and much less massive objects, such as white dwarf stars.

As matter is gravitationally attracted (accreted) toward a massive object, such as a



An artistic rendering of a thin accretion disc circling a supermassive black hole.

star or black hole, it cannot fall toward the central body in a straight line unless it was initially at rest and nothing perturbs its trajectory. Any deviation from a straight path will force the falling material to start orbiting the massive object or swing past it and leave it behind. Gas that settles into an orbit around a massive central object has notable properties: Through viscous interactions of its individual particles, it can slow down the falling speed and dissipate this kinetic energy as well as angular momentum. Viscosity will allow matter to gradually spiral down and feed the central object. This system is called an accretion disk, and it is the mechanism that feeds growing protostars (*3*), stars with close binary companions (*4*, *5*), and SMBHs in the center of galaxies (*6*, *7*). Through this mechanism, it is believed that SMBHs grew from small "seeds" in the early Universe to the colossi observed today.

Accretion disks around SMBHs can be extremely bright because of the high temperatures that are reached as the gas slows down through viscosity. These disks, not much larger than the Solar System, can outshine the whole galaxy that hosts them, in the optical, ultraviolet, and x-ray regions of the electromagnetic spectrum. It is therefore at these wavelengths that astronomers can study the accretion processes and determine the properties of the disks and of their central SMBHs.

Methods to measure the mass of a SMBH are often indirect, such as scaling relations with large-scale properties of their host galaxies (*8*, *9*). Also, methods used now often can only be applied to the nearest objects (about tens of megaparsecs from Earth) (*10*), require very special conditions (such as having very dense disks of molecular gas orbiting the black hole and positioned exactly edge-on as seen from Earth) (*11*), or require large amounts of telescope time (*12*). It also can be difficult to apply these methods to fainter accretion disks. Thus, new methods to determine SMBH masses are very much needed.

Burke *et al.* present a new method to determine the mass of SMBHs by examining the temporal variations of the emission from their accretion disks. Because accretion disks are small and extremely energetic, they are prone to instabilities that change the amount of radiation that they produce in a stochastic manner, a trademark that has been known since the discovery of the so-called quasars (very energetic and distant accreting SMBHs) (*13*). The stochasticity of the variable emission makes it difficult to isolate characteristic time scales of an accretion disk, such as a particular time associated with the orbital periods or any other "beat" signal from the structure of the disk (such as density waves or an orbiting hot spot in the disk that creates a periodic luminosity change). The exact nature of the observed variations is not well understood and is an area of very active research.

The method presented by Burke *et al.* uses the flux variability in radiation emission of 67 well-observed accreting SMBHs to determine the time scale—days, weeks, months, or years—on which the fluctuations become noticeably smaller. They found that this "damping" time scale correlates well with the masses of the SMBHs (that were determined by other means) over an impressive range from 10,000 to 10 billion solar masses. The damping time scale (of the order of several hundreds of days) can only be measured accurately if the lengths of the observations are about 10 times longer than the time scale itself. This requires dedicated monitoring campaigns that track the brightness changes in these sources stretching for many years. This apparently important restriction is at present being overcome by large monitoring surveys, which are an accumulating time series of the luminosity of millions of objects in the sky, making this method applicable on a massive scale.

As with the general SMBH flux variability, researchers have been struggling to interpret the meaning of this damping time scale. The findings of Burke *et al.* indicate that correlation of this time scale with physical properties such as the mass of the SMBH should provide further insight on the nature of both phenomena. As Burke *et al.* propose, other properties will likely play a role in determining the time scales of the fluctuations, such as the rate at which mass has been transferred to the SMBH by the accretion disk and the SMBH spin, which determines the innermost radius of the disk (*14*).

One of the most interesting aspects of the study of Burke *et al.* is that it extends its findings to much less massive objects, such as white dwarf stars, which emit radiation through a similar accretion disk mechanism and can be regarded as miniature accreting SMBHs. The authors found that stellar black holes and SMBHs follow almost the same damping time scale–mass relation. That the same relation extends through many orders of magnitude suggests that the physics of accretion disks is, at least in some aspects, scalable and proves that the variations that we see are indeed produced intrinsically by the accretion process. ■

[1]Department of Astronomy, Faculty of Physical and Mathematical Science, University of Chile, Santiago, Chile. [2]Department of Physics and Astronomy, University of Valparaiso, Valparaiso, Chile. Email: plira@das.uchile.cl; patricia.arevalo@uv.cl

## REFERENCES AND NOTES

1. Event Horizon Telescope Collaboration *et al.*, *Astrophys. J.* **875**, L1 (2019).
2. C. J. Burke *et al.*, *Science* **373**, 789 (2021).
3. R. Cesaroni, D. Galli, G. Lodato, C. M. Walmsley, Q. Zhang, in *Protostars and Planets*, V. B. Reipurth *et al.*, Eds. (Univ. Arizona Press, 2007), p. 197–212.
4. K. Mukai, *Publ. Astron. Soc. Pac.* **129**, 062001 (2017).
5. C. Done, M. Gierlinski, A. Kubota, *Astron. Astrophys. Rev.* **15**, 1 (2007).
6. A. Koratkar, O. Blaes, *Publ. Astron. Soc. Pac.* **111**, 1 (1999).
7. M. Giustini, D. Proga, *Astron. Astrophys. Rev.* **630**, A94 (2019).
8. K. Gebhardt *et al. Astrophys. J.* **539**, L13 (2000).
9. L. Ferrarese, D. Merritt, *Astrophys. J.* **539**, L9 (2000).
10. L. Ferrarese, H. Ford, *Space Sci. Rev.* **116**, 523 (2005).
11. M. Miyoshi *et al.*, *Nature* **373**, 127 (1995).
12. B. M. Peterson *et al.*, *Astrophys. J.* **613**, 682 (2004).
13. H. J. Smith, D. Hoffleit, *Astron. J.* **68**, 292 (1963).
14. C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation* (Princeton Univ. Press, 1973).

10.1126/science.abk3451

PALEONTOLOGY

# When did terrestrial plants arise?

## Microfossils suggest that co-option of algal genes may affect land plant origination time

*By* **Patricia G. Gensel**

Embryophytes—or land plants—are fundamental to life on Earth and influence systems that include carbon cycling, sedimentation rates, and rock weathering (*1*, *2*). Although much is known about their probable mid-Ordovician appearance [about 450 million years ago (Ma)] (*3*) and later diversification on the basis of fossils from late Silurian and Early Devonian periods (about 420 Ma) (*4*, *5*), the time of their origination has been unclear. This knowledge is important when considering geological processes and evolutionary interactions. There has been a discrepancy in the time of land plant origination between molecular clock estimations (based on genes and RNA) and fossil record estimates (based on morphology). On page 792 of this issue, Strother and Foster (*6*) describe fossilized spores whose characteristics raise the possibility that land plants arose by co-opting algal genes, along with acquiring de novo genes, and that the former would account for the molecular clock predating the fossil record.

Molecular clock estimates (*7*, *8*) suggest that embryophytes have existed since the Late Precambrian (about 600 to 550 Ma), whereas the earliest macrofossil of a vascular plant is from the mid-late Silurian (420 Ma). Molecular studies indicate that the closest sister lineage to embryophytes is certain streptophytic green algae (Charophytes) (see the figure). Some charophytes are aeroterrestrial—they live on land and share some ultrastructural and biochemical features with embryophytes. The Streptophyta and Chlorophyta (chlorophytic green algae) form the clade called Viridiplantae ("green plants") and are part of the Archaeplastida (Kingdom Plantae sensu lato) in the tree of life (*9*, *10*).

Assemblages of microfossils called cryptospores—from the Cambrian (581 to 485 Ma) to the Ordovician Period (485 to 443 Ma) (*3*)—bolster the macrofossil record of time of embryophyte origin. Recognizing which cryptospores represent embryophytes or aeroterrestrial streptophytic algae is often difficult, but criteria are emerging (*3*). Cryptospore data from the Early Paleozoic (about 540 Ma) suggest the presence of embryophytes in the Mid-Ordovician.

Understanding phylogenetic relationships of land plants and their close relatives is as important as the analysis of aeroterrestrial streptophytic algae characteristics. The various configurations of spores are consistent with a highly variable cytoskeletal architecture (*11*). Until the Mid-Ordovician, cryptospore tetrads and dyads or groups of them are morphologically similar to later-occurring ones but form very irregularly arranged tetrads or dyads (*3*). From the Mid-Ordovician, they were very regularly arranged, suggesting canalization of meiotic processes (*11*). The ultrastructure of the cryptospore wall resembles that of either embryophyte or charophyte algal walls, but not that of chlorophyte algal spores. The discovery of cryptophytes (*5*), which are millimeter-sized axial, sporangiate plants producing cryptospore dyads and tetrads, from late Silurian to earliest Devonian, indicates that at least some represent the spores of embryophytes.

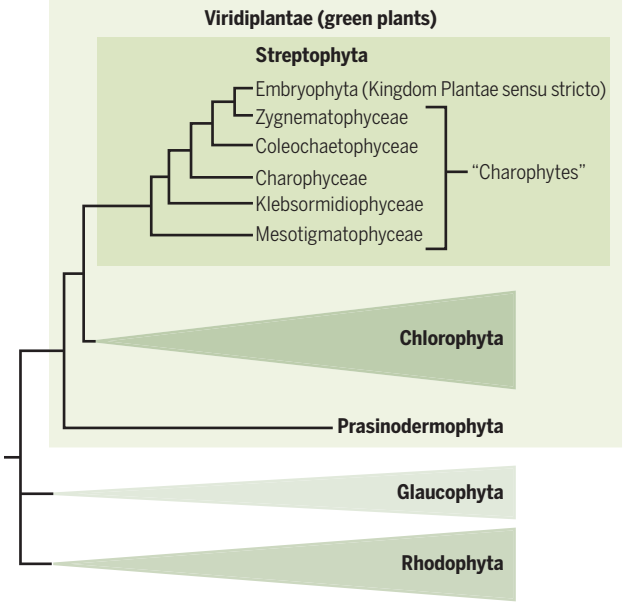
Other evidence supporting microfossil-based estimates of origination includes a presumed remnant of a sporangium (a structure containing spores) from the Late Ordovician (*12*). Comparable cryptospore morphologies and ultrastructure from some late Silurian and earliest Devonian plants and in a few extant bryophytes (nonvascular embryophytes) strengthens the argument that some dyads and tetrads represent spores of the earliest plants.

Bower (*13*) postulated a stepwise acquisition of characteristics or adaptations leading to land plants and argues for the development of the walled spore before evolving a sporophyte plant body. This involves the deposition of sporopollenin onto the meiotic products or spores of an "ancestral plant" (possibly aeroterrestrial streptophytic algae) (*3*). Sporopollenin, a polymer that coats spore walls in terrestrial plants (and possibly cryptospore walls), is deemed important for their survival in aerial or subaerial conditions. Other changes leading to the rise of land plants would include a delay in meiosis and mitotic division of nuclei or cells, resulting in either multinucleated cells or multiple cells or cells within the fertilized egg (zygote). Loss of meiosis (sterilization) of some of those nuclei or cells would ultimately lead to producing an embryo and multicellular sporophyte and sporangia, as are now found in land plants (*3*). Again, the cryptospore record of the Earlier Paleozoic is consistent with this scenario.

Whereas a clear distinction exists between the highly variable groupings of cryptospores from the Cambrian to Mid-Ordovician period (as algal) and those after this period (as nonalgal), Strother and Foster present data from the Early

### Phylogenetic relationships within the green plant clade

Recent genomic studies find certain streptophytic algae (part of the "charophytes" grade) as close sisters to land plants (embryophytes). Some streptophytic algae are land-dwelling (aeroterrestrial) or have characteristics that allowed adaptation to aerial or subaerial environments.




GRAPHIC: N. DESAI/*SCIENCE*

Biology Department, University of North Carolina, Chapel Hill, NC 27599, USA. Email: pgensel@bio.unc.edu

Ordovician showing co-occurrence of these two groups of spores. At first glance, this appears to support an earlier origination of plants (*14*). It also closes the gap between molecular and fossil estimates of the origin time. However, the authors raise another possibility on the basis of accepting the fossil record and using molecular studies to build on the idea of stepwise adaptations to land by plants from a probable aeroterrestrial streptophytic algae ancestor. Genomic studies demonstrate sequences in charophycean algae comparable to those in embryophytes (*9, 10*, *15*). Strother and Foster propose that portions of the algal genome were co-opted into the plant developmental genome, along with de novo genes. They further submit that algal-derived genes or gene sequences in plants might provide a molecular signal for phylogenetic analyses

**"There has been a discrepancy... between molecular clock estimations...and fossil record estimates..."**

that underlie molecular clock studies. This is as reasonable an explanation for the time discrepancy as other hypotheses (*14*).

Comparison of algal and plant genomes is still in early stages, but considerable information already exists (*15*). It is limited by not having fully sequenced genomes or transcriptomes from closely related streptophytes, basal plants such as liverworts or other bryophytes, and basal vascular plants. Such data, particularly those relating to meiosis and spore formation, multicellularity, and other features of embryophytes, as well as additional fossil information, are needed to test these hypotheses further. ■


**REFERENCES AND NOTES**

1. M. R. Gibling, N. S. Davies, *Nat. Geosci.* **5**, 99 (2012).
2. W. J. McMahon, N. S. Davies, *Science* **359**, 1022 (2018).
3. P. K. Strother, W. A. Taylor, in *Transformative Paleobotany*, M. Krings, C. J. Harper, N. R. Cuneo, G. W. Rothwell, Eds. (Elsevier, 2018), pp. 3–20.
4. P. G. Gensel, *Fern Gaz.* **20**, 217 (2017).
5. D. Edwards, J. L. Morris, J. B. Richardson, P. Kenrick, *New Phytol.* **202**, 50 (2014).
6. P. K. Strother, C. Foster, *Science* **373**, 792 (2021).
7. J. L. Morris *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2274 (2018).
8. Y. Nie *et al.*, *Syst. Biol.* **69**, 1 (2020).
9. L. Li *et al.*, *Nat. Ecol. Evol.* **4**, 1220 (2020).
10. S. Wang *et al.*, *Commun. Biol.* **4**, 412 (2021).
11. R. C. Brown, B. E. Lemmon, *New Phytol.* **190**, 875 (2011).
12. C. H. Wellman, P. L. Osterloff, U. Mohiuddin, *Nature* **425**, 282 (2003).
13. F. O. Bower, *The Origin of a Land Plant: A Theory Based on the Facts of Alternation* (Macmillan, 1908).
14. T. Servais *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **534**, 109280 (2019).
15. S. Wang *et al.*, *Nat. Plants* **6**, 95 (2020).


10.1126/science.abl5297

CANCER

# The cell of origin for Barrett's esophagus

## Undifferentiated cells that closely resemble gastric cells could be a biomarker for surveillance

*By* **Karen Geboes**[1] *and* **Anne Hoorens**[2]

Barrett's esophagus (BE) is a premalignant condition with increased risk for the development of esophageal adenocarcinoma (EAC). Surveillance relies on regular endoscopy with biopsies to detect dysplasia (disordered cellular growth) and diagnose cancer at an early, treatable stage. BE is a metaplastic response at the gastroesophageal boundary to chronic tissue injury by acid reflux. Metaplasia is the transformation of one differentiated cell type to another differentiated cell type: BE is characterized by replacement of normal squamous epithelium by columnar epithelium with gastric and intestinal features. The origin of metaplastic columnar cells is unclear, and different candidate precursor cells have been suggested. Better knowledge about the pathogenesis of BE may lead to better diagnosis, stratification, and treatment, and might help to develop targeted chemopreventive strategies in the future. On page 760 of this issue, Nowicki-Osuch *et al.* (*1*) find good evidence to support gastric cells as the origin of BE.

Endoscopic surveillance of people with BE to detect dysplasia or EAC is currently recommended (*2*, *3*). However, it can be difficult to confidently recognize BE by endoscopy, sampling errors may occur even with optimal adherence to surveillance protocols, there is interobserver variability in evaluating dysplasia, and definitions have been evolving over time and between regions. The squamo-columnar junction (SCJ) is the site of transition between squamous epithelium of the esophagus and columnar epithelium of the stomach and coincides with the gastroesophageal junction (GEJ) in a normal esophagus. The gastric cardia (GC) is the part of the stomach just below the GEJ. BE should only be diagnosed when there is a clear endoscopically visible change from squamous to columnar epithelium above the GEJ, characterized by a salmon color and coarse texture.

Diagnosis of BE also requires histological confirmation. Histologically, specialized intestinal metaplasia is defined as columnar epithelium with the presence of goblet cells, which are epithelial cells that secrete mucus and which are normally found in the intestinal epithelium (*4*) (see the image). However, several studies have found a similar risk of progression to dysplasia or EAC in patients with and without goblet cells in columnar-lined esophagus (*4*, *5*). Indeed, the absence of goblet cells may be a consequence of sampling error, warranting a new endoscopy. Conversely, goblet cells may develop both in the distal esophagus and proximal stomach and are histologically and histochemically identical. So, BE, especially short segments, may be difficult to diagnose and monitor. Additionally, although BE is a premalignant condition, most patients with BE will not progress to EAC, and patients with EAC do not always have evidence of BE at the time of diagnosis.

The difficulty in determining the cellular origin of BE is in part due to the inability to observe the process of metaplastic conversion in vivo and the lack of reliable physiological animal models (*6*). Epigenetic or genetic changes that alter protein expression, function, and/or activity in squamous esophageal cells or in stem or progenitor cells, such that they are reprogrammed to differentiate into columnar cells, are driven by the inflammatory environment created by chronic acid reflux. There have been several proposed mechanisms for the occurrence of metaplasia. Transdifferentiation is the conversion of one mature differentiated cell type to another, without undergoing an intermediate pluripotent state (*7*). It seems unlikely that this is the origin of BE because full phenotypic conversion of cultured mature squamous cells has not yet been demonstrated. Furthermore, the evidence of new squamous epithelium, which develops after endoscopic removal of BE, weakens the theory of transdifferentiation.

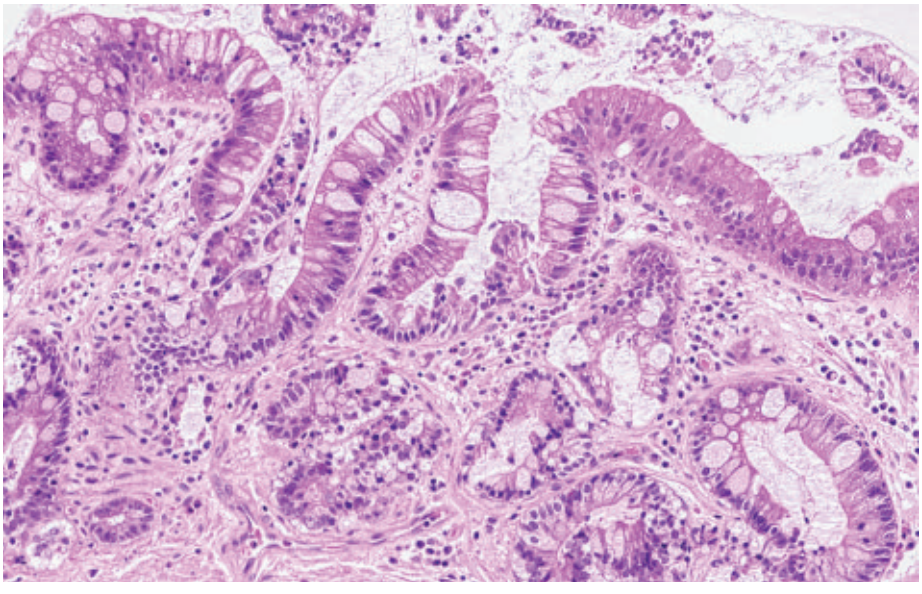
Another theory suggests that the metaplastic epithelium arises from a change

[1]Department of Gastroenterology, Universitair Ziekenhuis Gent, Gent, Belgium. [2]Department of Pathology, Universitair Ziekenhuis Gent, Gent, Belgium. Email: karen.geboes@uzgent.be

in the commitment of pluripotent stem cells that are responsible for the constant refreshing of the esophageal epithelium (*8*). Transcommitment starts with reprogramming immature progenitor cells that subsequently differentiate into another cell type. Whether the cells of origin are esophageal progenitor cells, progenitor cells in the esophageal submucosal glands, proximally shifting progenitor cells from the GC, residual embryonic cells at the SCJ, or circulating bone marrow–derived stem cells, the cells at the origin of BE would have to undergo molecular reprogramming that leads to a change in the cells' phenotypic commitment (*8*, *9*).

Nowicki-Osuch *et al*. analyzed freshly isolated human cells from superficial and submucosal compartments from the esophagus, GEJ, and GC of both healthy individuals and patients. They performed comprehensive phenotyping and multiomic profiling of different epithelial cell types in normal esophagus, gastric epithelium, BE, and EAC. An undifferentiated BE cell type was identified that showed expression of markers of intestinal and BE stem cells, and it was found that BE cells most closely resembled GC cells. The similarities between GC and BE cells do not prove causality. However, the findings of Nowicki-Osuch *et al*. add to the evidence of shared features between BE and metaplasia in the stomach. It has been proposed before that a spasmolytic polypeptide–expressing metaplasia (SPEM) precursor may account for the proposed "gastric origin" (goblet) cell that migrates up from the stomach into the esophagus (*10*). Consequently, it was suggested (*10*) that surveillance definitions may need to be altered, if SPEM is the precursor lesion and in itself possesses malignant transformation potential.

Indeed, because identification of BE and surveillance for EAC have their limitations, a better understanding of the cell of origin and the processes involved in metaplastic progression will hopefully allow the identification of biomarkers for EAC. It is therefore important that Nowicki-Osuch *et al*. found EAC to express markers of undifferentiated BE cells, even when no BE was visible. Again, this may point to a specific precursor lesion with malignant transformation potential: If it can be confirmed that undifferentiated BE cells exist in EAC, independent of the presence of metaplasia, this could lead to a drastic switch in screening programs, whereby greater attention is paid to detection of reflux and resulting molecular changes instead of detection of BE. ■



This hematoxylin and eosin stain of Barrett's esophagus shows the metaplastic change from squamous epithelium to columnar epithelium with goblet cells, which are normally found in the intestinal epithelium.


**REFERENCES AND NOTES**

1. K. Nowicki-Osuch *et al*., *Science* **373**, 760 (2021).
2. P. Sharma *et al*., *Gastroenterology* **131**, 1392 (2006).
3. J. R. Triggs, G. W. Falk, *Gastrointest. Endosc. Clin. N. Am.* **31**, 59 (2021).
4. W. M. Weinstein, A. F. Ippoliti, *Gastrointest. Endosc.* **44**, 91 (1996).
5. N. Vakil, S. V. van Zanten, P. Kahrilas, J. Dent, R. Jones, *Am. J. Gastroenterol.* **101**, 1900 (2006).
6. B. V. Naini, R. F. Souza, R. D. Odze, *Am. J. Surg. Pathol.* **40**, 45 (2016).
7. D. H. Wang, *Cell. Mol. Gastroenterol. Hepatol.* **4**, 157 (2017).
8. D. H. Wang, R. F. Souza, *Adv. Exp. Med. Biol.* **908**, 183 (2016).
9. J. Que, K. S. Garman, R. F. Souza, S. J. Spechler, *Gastroenterology* **157**, 349 (2019).
10. R. U. Jin, J. C. Mills, *Dig. Dis. Sci.* **63**, 2028 (2018).


10.1126/science.abj9797

METABOLISM

# Taking the long view on metabolism

## Measured energy expenditure across the human life span reveals distinct metabolic phases

*By* **Timothy W. Rhoads**[1] *and* **Rozalyn M. Anderson**[1,2]

Metabolism is not just about energy—how the body handles nutrient fuel and converts it to useable energetic currency. Metabolism also encompasses synthesis, modification, and exchange of the building blocks for all aspects of cellular function and acts as a sensor and regulator of cellular activities, in which individual moieties within metabolic pathways influence cellular responses. A substantial amount of the energy taken in each day is required to simply sustain life; the energetic demands of physical activity are superimposed on a vastly integrated machinery. Metabolic status has been linked to innumerable diseases and disorders, including those most prevalent with age (*1*–*3*). On page 808 of this issue, Pontzer *et al*. (*4*) analyze energy expenditure in more than 6400 males and females from 29 countries across the globe, aged between 8 days and 95 years, and show distinct metabolic phases during development and aging.

An understanding of energy expenditure across the life span must grapple with the diversity of humans, including sex, race, body composition, and their environment. Estimates of energy expenditure can be captured with indirect calorimetry that measures gas exchange and heat production of sequestered individuals, or by the doubly labeled water (DLW) method in free living individuals. The DLW technique is based on the relative bodily elimination rates of isotopes of oxygen and hydrogen (*5*). In the time since methods were developed for application in humans (*6*), the

[1]Department of Medicine, School of Medicine and Public Health, University of Wisconsin-Madison; Madison, WI, USA. [2]Geriatric Research, Education, and Clinical Center, William S. Middleton Memorial Veterans Hospital; Madison, WI, USA. Email: rozalyn.anderson@wisc.edu

PHOTO: LIZHE ZHUANG

use of DLW has been steadily growing. Associated costs of isotope dosing have limited most studies to relatively small cohorts, but there has been commitment among the research community to share data and to develop integrative methods so that large-cohort data analysis might be undertaken (*7*).

In the study of Pontzer *et al*., energy expenditure was adjusted to fat-free mass to account for differences in body size, revealing intrinsic shifts in metabolic status over the course of development, maturation, and aging. The authors identify inflection points that are the boundaries for four distinct phases. It seems clear from their data that infants and adolescents form two different metabolic categories. It has been said before, but children are not just small adults (*8*). That young people represent separate metabolic status categories has important implications for recommendations about diet and physical activity, not to mention pharmaceutical dose recommendations for younger persons. The remaining two phases cover adulthood and advanced age. Adjusted energy expenditure is notably stable from 20 years of age up to about 60 years of age, at which point a gradual decline is observed (see the figure).

The decline from age 60 is thought to reflect a change in tissue-specific metabolism, the energy expended on maintenance. It cannot be a coincidence that the increase in incidence of noncommunicable diseases and disorders begins in this same time frame (*9*). These findings indicate that life stage needs to be carefully considered when choosing disease models. This is particularly important for research on the etiology of age-associated diseases and disorders (*10*). Pathways and factors that are readily targetable in young growing animals may not be as sensitive or even responsive in older animals, and young models fail to capture the aged environment and may miss interactions that emerge as a result of intrinsic differences in metabolic status.

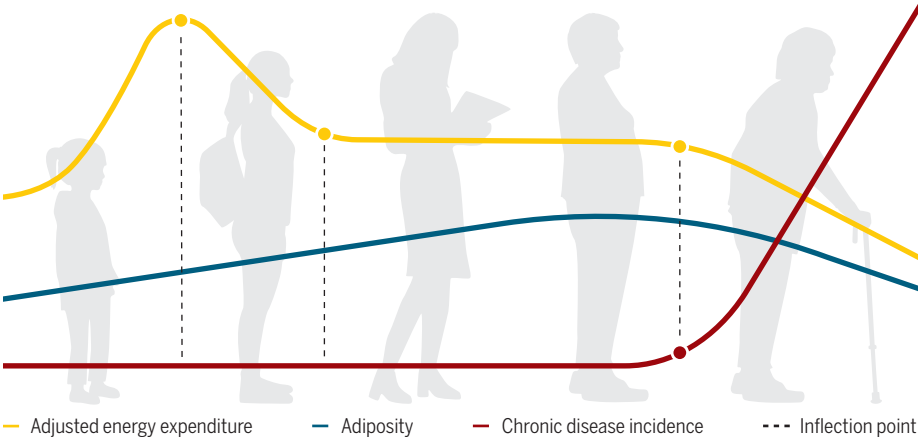
The impact of body size on metabolic rate has been discussed and explored for decades (*11*). Total energy expenditure is sex dimorphic, with lower levels in females than in males; however, accounting for fat-free mass removes this distinction. It is important that contributions from physical activity and tissue-specific metabolic rates, both of which change over the human life span, must be accounted for if computational models are to fit the observed data. Although not the focus of the work, Pontzer *et al*. identified substantial heterogeneity in body composition among individuals. Challenges arising from heterogeneity among individuals are reflected in the growing enthusiasm for precision medicine (*12*). It is abundantly clear that one size does not fit all. By adjusting for fat-free mass, this study peels some of this variance away to reveal intrinsic shifts in metabolism that are matched to life phase. There is considerable heterogeneity in how and when aging manifests in terms of disease incidence (*13*). It would be interesting to explore how mid-life disposition informs outcomes in advanced age and how well disease burden among individuals links to age-associated changes in their tissue-specific metabolism.

The causal factors in age-related vulnerability to disease no doubt reside in the documented changes in cellular biology, tissue physiology, and systemic homeostasis. Studies of laboratory animals have honed in on metabolism as a central theme in aging and in delayed aging through caloric restriction (*14*). Differences in metabolism are predicted to affect derivation of energy from nutrient sources, foundational material for synthesis of cellular machinery and communication relays, and the ability to optimize cellular activities according to prevailing conditions, whether external or internal. It will come as no surprise then that recent efforts to identify pharmacological agents that positively affect health in aging converge on metabolism (*15*). The Pontzer *et al*. study provides important new insights into human metabolism; the unprecedented scale and scope of the study is matched by the outstanding collaborative spirit that made it possible. ■

**"Pathways and factors that are readily targetable in young growing animals may not be as sensitive or even responsive in older animals, and young models fail to capture the aged environment..."**

## Life span of metabolism

Measures of energy expenditure (adjusted for fat-free mass) identify three inflection points over the human life span. Energy expenditure increases during infancy and childhood and then declines through adolescence, a plateau phase lasts throughout adulthood, and a second declining phase occurs from 60 years of age. The marked rise in incidence of chronic disease from late middle age aligns with the shift in energy expenditure and loss of adiposity, suggesting that metabolism may be a driver in aging biology.




GRAPHIC: A. MASTIN/*SCIENCE*


### REFERENCES AND NOTES

1. N. N. Pavlova, C. B. Thompson, *Cell Metab.* **23**, 27 (2016).
2. S. Costantino, F. Paneni, F. Cosentino, *J. Physiol.* **594**, 2061 (2016).
3. S. Camandola, M. P. Mattson, *EMBO J.* **36**, 1474 (2017).
4. H. Pontzer *et al*., *Science* **373**, 808 (2021).
5. J. R. Speakman, *Am. J. Clin. Nutr.* **68**, 932S (1998).
6. D. A. Schoeller, *J. Nutr.* **118**, 1278 (1988).
7. J. R. Speakman *et al*., *Cell Rep. Med.* **2**, 100203 (2021).
8. P. F. Saint-Maurice, Y. Kim, G. J. Welk, G. A. Gaesser, *Eur. J. Appl. Physiol.* **116**, 29 (2016).
9. NCD Countdown collaborators, *Lancet* **392**, 1072 (2018).
10. B. K. Kennedy *et al*., *Cell* **159**, 709 (2014).
11. M. Kleiber, *Physiol. Rev.* **27**, 511 (1947).
12. M. A. Haendel, C. G. Chute, P. N. Robinson, *N. Engl. J. Med.* **379**, 1452 (2018).
13. D. J. Lowsky, S. J. Olshansky, J. Bhattacharya, D. P. Goldman, *J. Gerontol. A Biol. Sci. Med. Sci.* **69**, 640 (2014).
14. P. Balasubramanian, P. R. Howell, R. M. Anderson, *EBioMedicine* **21**, 37 (2017).
15. L. Partridge, M. Fuentealba, B. K. Kennedy, *Nat. Rev. Drug Discov.* **19**, 513 (2020).



### ACKNOWLEDGMENTS

T.W.R. and R.M.A. are supported by NIH/NIA grants AG040178, AG057408, and AG067330; the Department for Veterans Affairs BX003846; and the Simons Foundation.

10.1126/science.abl4537

MEDICINE

# Plant-made vaccines and therapeutics

## Advances in technology and manufacturing could boost the uptake of molecular farming

By **Hugues Fausther-Bovendo**[1] and **Gary Kobinger**[1,2]

Therapeutic proteins such as vaccines, antibodies, hormones, and cytokines are generally produced in bacteria or eukaryotic systems, including chicken eggs and mammalian or insect cell cultures, with high production yield according to well-defined regulatory guidelines (*1*). The use of plants for the production of therapeutic proteins, called molecular farming, was proposed as an alternative biomanufacturing method in 1986. The first and only plant-derived therapeutic protein for human use was approved in 2012 for the treatment of Gaucher disease. In 2019, a plant-produced influenza virus vaccine completed phase 3 clinical trials, with encouraging results (*2*). More recently, phase 3 trials for an adjuvanted plant-made vaccine (CoVLP) against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (NCT04636697) began in March 2021. These successes have revived interest in plant-produced pharmaceuticals for human use, which could include edible drugs.

Molecular farming was initially proposed because growing plants only requires light, water, and soil (or artificial support). Procurement of greenhouses is cheaper than bioreactor suites, which are required for bacterial, mammalian, and insect cell culture systems, making molecular farming particularly attractive for developing countries. Furthermore, production can be scaled up or down based on the number of plants grown. Additionally, unlike in the traditional production systems, zoonotic pathogens are unable to infect plants and therefore cannot be a source of contaminant in molecular farming–derived products.

Technological progress, such as codon optimization, the inclusion of organelle-specific promoters, humanization of N-glycans, and transient transfection systems, has increased the performance of molecular farming, with yields above 1 mg/g of fresh plant weight. In comparison, yields between 5 and 20 g/liter are commonly reached in mammalian Chinese hamster ovary cell cultures that are used to manufacture monoclonal antibodies. Transient transfection systems have also increased the production speed, with harvest possible within days after transfection of adult plants as opposed to months when stable expression was sought. The production speed of molecular farming is a critical asset. Indeed, plant-produced vaccines can readily be made against new pathogens or emerging strains, with the first batch of vaccine candidates typically produced within 3 weeks (*3*). The speed of molecular farming is particularly suited for personalized medicine in which pharmaceuticals need to be tailored to individual patients, such as for cancer treatment (*4*).

Constructing facilities capable of purifying therapeutic proteins produced in plants under good manufacturing practice (GMP)–compliant conditions has also supported clinical evaluation of plant-made vaccines and therapeutics (*5*, *6*). Moreover, the simplicity and low costs associated with plant growth are an advantage, whereas expensive GMP facilities are required for the extraction, purification, and fill processes of molecular farming. By contrast, the lower yield, limited regulatory guidelines available, and limited manufacturing capacity worldwide have dampened enthusiasm toward molecular farming (*1*).

For vaccine purposes, plant-produced proteins have several advantages over their counterparts produced in bacterial, mammalian, or insect systems. Unlike bacteria, plants are capable of posttranslational modifications. Plants express different glycans, which renders plant-derived proteins more immunogenic than their mammalian counterparts (*7*). In plant-made vaccines, virus-like particles (VLPs) are produced that comprise the target pathogen's protein(s) of interest (the immunogen) and plant components within the particles. These plant components of VLPs, such as lectins, glycans, saponins, and heat shock proteins, have adjuvant properties (*7*) that can further potentiate the immune response against plant-made vaccines and may reduce the need for adjuvants in vaccine formulation.

This increased immune stimulation may lead to hypersensitivity (allergic reactions) toward plant components. However, several clinical trials, including phase 1 trials for CoVLP (NCT04450004) and phase 3 trials for plant-made vaccines against influenza virus (NCT03301051, NCT03739112), have alleviated this concern. In these large phase 3 clinical trials, involving more than 20,000 adults aged 18 to 94 years, a mixture of four separate VLPs, each containing the hemagglutinin proteins from a selected seasonal influenza virus strain, produced in tobacco plants, have shown efficacy similar to that of existing influenza vaccines (*2*). Notably, immunization with this quadrivalent VLP formulation was not associated with more adverse events or an increase in hypersensitivity reactions compared with individuals receiving a licensed influenza virus vaccine.

The plant-made vaccines against influenza virus and SARS-CoV-2 are expected to be the first therapeutic proteins produced in whole plants for human use. The glucocerebrosidase enzyme, produced as an injectable protein drug called taliglucerase alfa for the treatment of Gaucher disease, is produced in a carrot cell culture rather than in actual plants (*8*). Several plant-produced veterinary vaccines have also been developed, and one is approved by the US Department of Agriculture for immunization of chickens against Newcastle disease. However, there is competition to reduce the selling price of farm animals, so the cost of plant-made veterinary vaccines needs to be substantially reduced to ensure their widespread adoption (*5*).

Although the increased immunogenicity of plant-made protein is beneficial for vaccines, it can be detrimental for therapeutic proteins, potentially reducing their in vivo efficacy and contributing to adverse events. Despite these limitations, monoclonal antibodies against HIV (NCT01403792) and Ebola virus (NCT02363322, NCT02363322) have reached clinical stages of development. In human trials, intravenous (Ebola virus) or intravaginal (HIV) administrations of these antibodies were generally well tolerated. Mild adverse events, such as hypotension and fever, were observed following intravenous delivery (*9*, *10*). Life-threatening reactions, such as anaphylactic shock, have only been reported in a single recipient of 238 treated with antibodies against HIV or Ebola virus (*10*, *11*). The administration of antihistamine and antipyretic (fever-reducing) agents before intravenous infusion of plant-produced pharmaceuticals has mitigated the occurrence of severe adverse events (*11*).

The use of plants engineered to express human glycans also reduces the immunogenic-

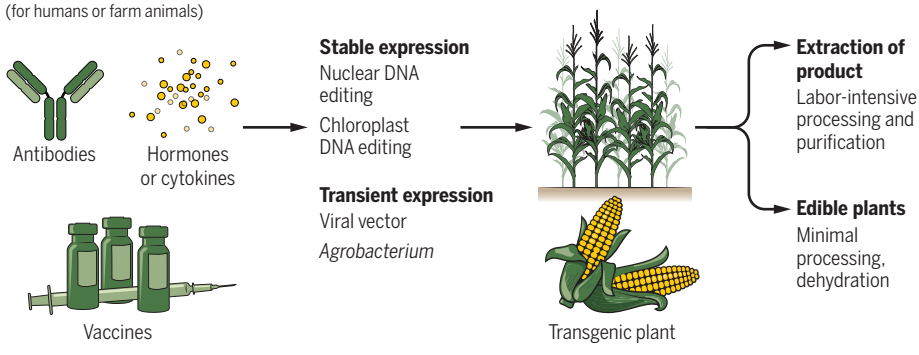
[1]Faculty of Medicine, Université Laval, Quebec, QC, Canada. [2]Galveston National Laboratory, Galveston, TX, USA. Email: gary.pignac-kobinger.1@ulaval.ca

ity of plant-made proteins. It is worth noting that independently of the production system, the intravenous injection of grams of antibodies is associated with some adverse events, which are observed even with antibodies produced in mammalian cells. Another hurdle for plant-made monoclonal antibodies is limited manufacturing capacity. Unlike vaccines, monoclonal antibody regimens usually require larger doses (50 to 150 mg/kg) and sometimes with repeated administrations. To achieve the required scale of plant-made antibodies, a substantial expansion of current manufacturing capacity, which is currently limited to less than 15 GMP facilities worldwide, is required. Funding opportunities, similar to that of the US Defense Advanced Research Projects Agency (DARPA), which financed the development of three large GMP facilities for the production of plant-made vaccines in the US, are needed (*5*).

The use of edible plants such as cereal crops, tomatoes, corn, and rice is under development for oral delivery of plant-made therapeutic proteins (*12*). Parenteral administration of tumor necrosis factor (TNF) antagonists is used for treating autoimmune diseases such as rheumatoid arthritis. In a phase 1 clinical trial, oral administration of lyophilized tobacco plant–derived cells expressing a TNF antagonist was safe and increased the amount of immunosuppressive regulatory T cells ($T_{regs}$) in human volunteers, demonstrating the feasibility of this approach (*13*). In addition to reducing adverse events, the oral route can favor the induction of tolerance to suppress autoimmune or allergic responses. To date, the induction of tolerance by using recombinant proteins within edible plants is restricted to animal models of hypersensitivity or autoimmunity. However, this could provide low-cost and simple solutions to better manage and prevent the growing incidence of allergies to common substances (*12*, *14*).

## Molecular farming

Using transient expression systems, plant-made vaccine and therapeutic proteins can be produced within weeks. For oral administration, edible plants that express therapeutic proteins require minimal processing after harvest. By combining the reduced cost and ease of administering edible vaccines or therapeutic proteins with the speed of transient expression systems, molecular farming could have a considerable impact on both human and animal health.




GRAPHIC: C. BICKEL/*SCIENCE*

Oral administration of drugs is a user-friendly alternative to the intravenous route. Furthermore, oral administration can mitigate the adverse events associated with intravenous administration of pharmaceuticals. Gut immune responses are crucial for tolerance to food and self-antigens and play an important role in ensuring a balanced immune system (*12*). Moreover, most pharmaceutical proteins currently under clinical evaluation are purified before parenteral injections. Given orally, plant-made therapeutics might only require minimal processing, thus possibly skipping expensive and time-consuming steps in the manufacturing process. Products for oral delivery can also be stored in lyophilized (dehydrated) form at room temperature for an extended period, hence sharply reducing both their cost of production and storage while simultaneously facilitating their administration.

Edible vaccines are also under development. The safety and feasibility of this approach were demonstrated in proof-of-concept phase 1 clinical trials, to monitor the safety and immunogenicity of edible vaccine candidates against *Escherichia coli*, hepatitis B virus, rabies lyssavirus, and norovirus, which occurred between 1998 and 2004. In these trials, the proportion of immunized individuals who generated an immune response against the desired target was disappointingly lower than in clinical trials involving standard vaccines administered via the parenteral route. The yield of recombinant proteins produced in plants has since increased substantially, suggesting that new edible plant-made vaccines could now generate meaningful immune responses. The use of edible plant vaccines in prime-boost immunization regimens, with an injectable vaccine as a prime and an edible vaccine as a booster, has also been investigated to improve the clinical relevance of plant-produced oral vaccines against poliovirus (*12*, *14*, *15*). However, further optimization is required before clinical acceptance of these vaccine candidates (*14*). Ensuring that edible plant-made vaccines do not lead to hypersensitivity against the plant used for production is crucial, especially plants that are widely consumed such as rice, cereals, and corn (see the figure).

The plant-made quadrivalent VLP influenza vaccine will likely be licensed for use, given the favorable outcome of the phase 3 clinical trials. Ongoing clinical development will continue to inform regulatory guidelines for plant-made therapeutic proteins. However, because doses for therapeutics are much higher than for vaccines, investment in manufacturing infrastructure must increase and production costs need to further decrease to achieve large-scale manufacturing of plant therapeutic products. Until then, the speed of molecular farming will be useful for preclinical and early clinical evaluation of therapeutic candidates. Edible, plant-made therapeutics are still predominantly in the preclinical stage of development but, if successful, could create new classes of pharmaceutical products. Manufacturing of pharmaceutical proteins may remain dominated by current production systems until economic attractiveness through easy manufacture and technological progress, such as potent edible plant-made therapeutics, shifts the balance toward molecular farming. ■

### REFERENCES AND NOTES


1. S. Schillberg, N. Raven, H. Spiegel, S. Rasche, M. Buntru, *Front. Plant Sci* **10**, 720 (2019).
2. B. J. Ward *et al.*, *Lancet* **396**, 1491 (2020).
3. F. Sainsbury, *Curr. Opin. Biotechnol.* **61**, 110 (2020).
4. U. Sahin, Ö. Türeci, *Science* **359**, 1355 (2018).
5. N. Takeyama, H. Kiyono, Y. Yuki, *Ther. Adv. Vaccines* **3**, 139 (2015).
6. B. Shanmugaraj, C. J. I. Bulaon, W. Phoolcharoen, *Plants* **9**, 1 (2020).
7. V. A. Sander, M. G. Corigliano, M. Clemente, *Front. Vet. Sci.* **6**, 20 (2019).
8. A. Zimran *et al.*, *Am. J. Hematol.* **91**, 656 (2016).
9. J. K. C. Ma *et al.*, *Plant Biotechnol. J.* **13**, 1106 (2015).
10. S. Mulangu *et al.*, *N. Engl. J. Med.* **381**, 2293 (2019).
11. PREVAIL II Writing Group, *N. Engl. J. Med.* **375**, 1448 (2016).
12. M. Merlin, M. Pezzotti, L. Avesani, *Br. J. Clin. Pharmacol.* **83**, 71 (2017).
13. E. Almon *et al.*, *J. Immunol. Methods* **446**, 21 (2017).
14. S. Rosales-Mendoza, R. Nieto-Gómez, *Trends Biotechnol.* **36**, 1054 (2018).
15. H. T. Chan *et al.*, *Plant Biotechnol. J.* **14**, 2190 (2016).


### ACKNOWLEDGMENTS


The authors are supported by the International Development Research Centre (109075-001); the Canadian Department of Foreign Affairs, Trade and Development (BIO-2019-005); and the Canadian Institutes of Health Research. Thanks to L. Zeitlin for useful discussions.

10.1126/science.abf5375

PLANETARY SCIENCE

# Searching for life on Mars and its moons

## Sample-return missions will look for extraterrestrial life and biomarkers on Mars and Phobos

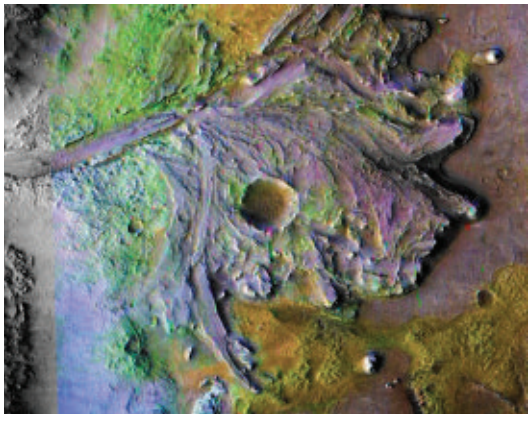
*By* **Ryuki Hyodo** *and* **Tomohiro Usui**

The scientific exploration of Mars over the past several decades has resulted in increasing evidence that the martian surface hosted habitable environments early in its history, as well as evidence of the building blocks of life in the form of organic molecules (*1*). Habitats on Mars that could harbor extant martian life have been hypothesized, such as subsurface environments, caves, and ice deposits (*2*). Mars is currently recognized as a "paleo-habitable" planet, reflecting its ancient habitability. Fully understanding the evolution of habitability and whether Mars has ever hosted life will be essential to understanding and exploring other extraterrestrial habitable environments and potential life-forms (*3*). Flagship missions of multiple space agencies in the 2020s will play essential and complementary roles and could finally provide an answer to these long-standing questions.

The planned Mars Sample-Return (MSR) mission of NASA and the European Space Agency should reveal more about the habitability of Mars by helping to determine the geologic evolution of Jezero crater and its surrounding areas, which are believed to be the site of an ancient lake (see the photo). The Mars 2020 Perseverance rover will attempt to collect samples that will allow scientists to explore the evolution of Jezero crater and its habitability over time, as well as samples that may contain evidence of biosignatures. A high-priority science objective for MSR returned-sample science is to understand the habitability of Mars and look for potential signs of both extinct and extant life (*4*).

Mars is not alone because it has two small moons, Phobos and Deimos. Throughout the history of Mars, numerous asteroidal impacts on Mars have produced martian impact ejecta, and a fraction of the ejected material has been delivered to its moons (*5*). Phobos is closer to Mars, so it has more martian ejecta than Deimos. Numerical simulations show that >$10^9$ kg of martian material could be uniformly mixed in the regolith of Phobos (the resultant martian fraction is >1000 parts per million) (*5*).

Even if martian life-forms existed and could survive the transport to Phobos without suffering from impact-shock decomposition (with a peak pressure of <5 GPa) (*5, 6*), the Phobos environment is highly inhospitable (*7*). Phobos does not have air or water, and its surface is constantly bathed in solar and galactic cosmic radiation. This indicates that martian materials on Phobos' surface almost certainly do not contain any living microorganisms.



Jezero crater on Mars is believed to be the site of an ancient lake. The Mars 2020 Perseverance rover aims to collect samples from the crater to analyze for evidence of life.

Instead, there may be dead biosignatures on Phobos, which we have called "SHIGAI" (Sterilized and Harshly Irradiated Genes, and Ancient Imprints)—the acronym in Japanese means "dead remains." SHIGAI includes any potential microorganisms that could have been alive on Mars and were recently sterilized during or after the delivery to Phobos, and the microorganisms and biomarkers that had been processed on ancient Mars before the delivery to Phobos, including potential DNA fragments. The Mars-moon system is an ideal natural laboratory for the study of interplanetary transport and sustainability of SHIGAI on airless bodies in the Solar System.

Should a martian biosphere exist, any biosignatures or biomarkers observed in the samples from Jezero crater could be widespread elsewhere on Mars and possibly occur on the surface of Phobos. Because martian ejecta has been thoroughly delivered to Phobos by impact-driven random sampling, the biosignatures and biomarkers that may be contained in the Phobos regolith could reflect the diversity and evolution of a potential martian biosphere.

Martian Moons eXploration (MMX), developed by the Japan Aerospace Exploration Agency, plans to collect a sample of >10 g from the Phobos surface and return to Earth in 2029 (*8*). Detection of a "fingerprint" of martian life and SHIGAI should be achievable through comprehensive comparative studies using martian material from the Phobos surface and samples from Jezero crater returned by MMX and MSR, respectively.

The MSR samples have the potential to contain a variety of biomarker molecules (e.g., lipids, such as hopanoids, sterols, and archaeols, and their diagenetic products) (*4*). The sample could include modern living organisms from Jezero crater, if they are present. Of course, MSR could return samples without any evidence of life because of the focus on a single location. A distinct advantage for MMX is the ability to deliver martian materials derived from several regions. The random nature of the crater-forming impacts on Mars statistically delivers all possible martian materials, from sedimentary to igneous rocks that cover all of its geological eras.

Mutual international cooperation on MSR and MMX could answer questions such as how martian life, if present, emerged and evolved in time and place. If Mars never had life at all, these missions would then be absolutely vital in unraveling why Mars is lifeless and Earth has life. Therefore, the missions may eventually provide the means to decipher the divergent evolutionary paths of life on Mars and Earth. ■

Japan Aerospace Exploration Agency, Yoshinodai, Kanagawa 252-5210, Japan. Email: hyodo.ryuki@jaxa.jp

### REFERENCES AND NOTES


1. J. L. Eigenbrode *et al.*, *Science* **360**, 1096 (2018).
2. B. L. Carrier *et al.*, *Astrobiology* **20**, 785 (2020).
3. B. L. Ehlmann *et al.*, *J. Geophys. Res. Planets* **121**, 1927 (2016).
4. D. W. Beaty *et al.*, *Meteorit. Planet. Sci.* **54**, S3 (2019).
5. R. Hyodo *et al.*, *Sci. Rep.* **9**, 19833 (2019).
6. L. E. Nyquist *et al.*, *Space Sci. Rev.* **96**, 105 (2001).
7. K. Kurosawa *et al.*, *Life Sci. Space Res.* **23**, 85 (2019).
8. M. Fujimoto, E. J. Tasker, *Nat. Astron.* **3**, 284 (2019).


### ACKNOWLEDGMENTS

We thank D. W. Beaty, M. Grady, and B. Carrier for comments and discussions on NASA-ESA MSR science. We thank H. Sugahara, M. Fujimoto, K. Kurosawa, and H. Genda for many useful discussions on Japan Aerospace Exploration Agency MMX science. This Perspective was constructed through discussions with them and submitted on their behalf.

10.1126/science.abj1512

PHOTO: NASA/JPL-CALTECH/MSSS/JHU-APL

# Making machine learning trustworthy

## Safety, transparency, and fairness are essential for high-stakes uses of machine learning

*By* **Birhanu Eshete**

Machine learning (ML) has advanced dramatically during the past decade and continues to achieve impressive human-level performance on nontrivial tasks in image, speech, and text recognition. It is increasingly powering many high-stake application domains such as autonomous vehicles, self–mission-fulfilling drones, intrusion detection, medical image classification, and financial predictions (*1*). However, ML must make several advances before it can be deployed with confidence in domains where it directly affects humans at training and operation, in which cases security, privacy, safety, and fairness are all essential considerations (*1*, *2*).

The development of a trustworthy ML model must build in protections against several types of adversarial attacks (see the figure). An ML model requires training datasets, which can be "poisoned" through the insertion, modification, or removal of training samples with the purpose of influencing the decision boundary of a model to serve the adversary's intent (*3*). Poisoning happens when models learn from crowdsourced data or from inputs they receive while in operation, both of which are susceptible to tampering. Adversarially manipulated inputs can evade ML models through purposely crafted inputs called adversarial examples (*4*). For example, in an autonomous vehicle, a control model may rely on road-sign recognition for its navigation. By placing a tiny sticker on a stop sign, an adversary can evade the model to mistakenly recognize the stop sign as a yield sign or a "speed limit 45" sign, whereas a human driver would simply ignore the visually nonconsequential sticker and apply the brakes at the stop sign (see the figure).

Department of Computer and Information Science, University of Michigan-Dearborn, Dearborn, MI, USA.
Email: birhanu@umich.edu

Attacks can also abuse the input-output interaction of a model's prediction interface to steal the ML model itself (*5*, *6*). By supplying a batch of inputs (for example, publicly available images of traffic signs) and obtaining predictions for each, a model serves as a labeling oracle that enables an adversary to train a surrogate model that is functionally equivalent to the model. Such attacks pose greater risks for ML models that learn from high-stake data such as intellectual property and military or national security intelligence.

When models are trained for predictive analytics on privacy-sensitive data, such as patient clinical data and bank customer transactions, privacy is of paramount importance. Privacy-motivated attacks can reveal sensitive information contained in training data through mere interaction with deployed models (*7*). The root cause for such attacks is that ML models tend to "memorize" ancillary parts of their training data and, at prediction time, inadvertently divulge identifying details about individuals who contributed to the training data. One common strategy, called membership inference, enables an adversary to exploit the differences in a model's response to members and nonmembers of a training dataset (*7*).

In response to these threats to ML models, the quest for countermeasures is promising. Research has made progress on detecting poisoning and adversarial inputs to limiting what an adversary may learn by just interacting with a model to limit the extent of model stealing or membership inference attacks (*1*, *8*). One promising example is the formally rigorous formulation of privacy. The notion of differential privacy promises to an individual who participates in a dataset that whether your record belongs to a training dataset of a model or not, what an adversary learns about you by interacting with the model is basically the same (*9*).

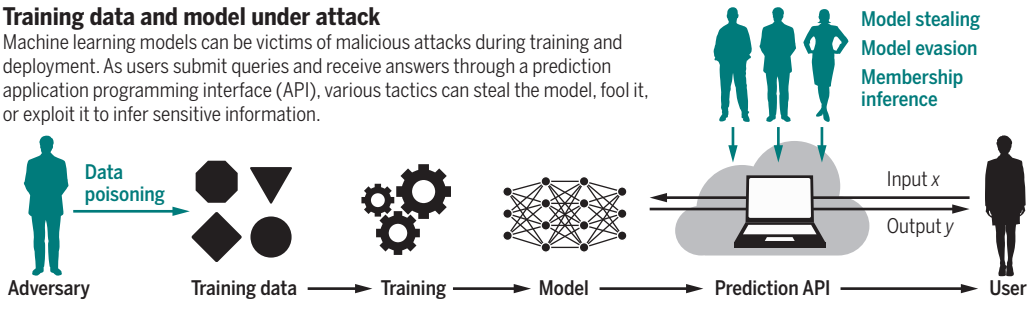
Beyond technical remedies, the lessons learned from the ML attack-defense arms race provide opportunities to motivate broader efforts to make ML truly trustworthy in terms of societal needs. Issues include

## Adversarial threats to machine learning

Machine learning models are vulnerable to attacks that degrade model confidentiality and model integrity or that reveal private information.

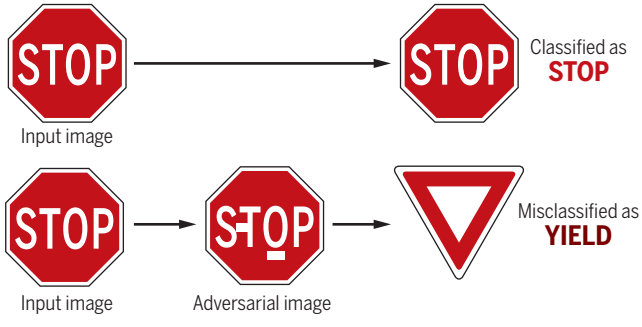
### Training data and model under attack

Machine learning models can be victims of malicious attacks during training and deployment. As users submit queries and receive answers through a prediction application programming interface (API), various tactics can steal the model, fool it, or exploit it to infer sensitive information.




### Malicious inputs

Data sets can be "poisoned" or models can be fooled. For example, an adversary can fool a model of a self-driving car by putting tiny stickers on a traffic stop sign to cause it to be interpreted as a yield sign.




GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

how a model "thinks" when it makes decisions (transparency) and fairness of an ML model when it is trained to solve high-stake inference tasks for which bias exists if those decisions were made by humans. Making meaningful progress toward trustworthy ML requires an understanding about the connections, and at times tensions, between the traditional security and privacy requirements and the broader issues of transparency, fairness, and ethics when ML is used to address human needs.

Several worrisome instances of biases in consequential ML applications have been documented (*10*, *11*), such as race and gender misidentification, wrongfully scoring darker-skin faces for higher likelihood of being a criminal, disproportionately favoring male applicants in resume screenings, and disfavoring black patients in medical trials. These harmful consequences require that the developers of ML models look beyond technical solutions to win trust among human subjects who are affected by these harmful consequences.

On the research front, especially for the security and privacy of ML, the aforementioned defensive countermeasures have solidified the understanding around blind spots of ML models in adversarial settings (*8*, *9*, *12*, *13*). On the fairness and ethics front, there is more than enough evidence to demonstrate pitfalls of ML, especially on underrepresented subjects of training datasets. Thus, there is still more to be done by way of human-centered and inclusive formulations of what it means for ML to be fair and ethical. One misconception about the root cause of bias in ML is attributing bias to data and data alone. Data collection, sampling, and annotation play a critical role in causing historical bias, but there are multiple junctures in the data processing pipeline where bias can manifest. From data sampling to feature extraction, from aggregation during training to evaluation methodologies and metrics during testing, bias issues manifest across the ML data-processing pipeline.

At present, there is a lack of broadly accepted definitions and formulations of adversarial robustness (*13*) and privacy-preserving ML (except for differential privacy, which is formally appealing yet not widely deployed). Lack of transferability of notions of attacks, defenses, and metrics from one domain to another is also a pressing issue that impedes progress toward trustworthy ML. For example, most ML evasion and membership inference attacks illustrated earlier are predominantly on applications such as image classification (road-sign detection by an autonomous vehicle), object detection (identifying a flower from a living room photo with multiple objects), speech processing (voice assistants), and natural language processing (machine translation). The threats and countermeasures proposed in the context of vision, speech, and text domain hardly translate to one another, often naturally adversarial domains, such as network intrusion detection and financial-fraud detection.

Another important consideration is the inherent tension between some trustworthiness properties. For example, transparency and privacy are often conflicting because if a model is trained on privacy-sensitive data, aiming for the highest level of transparency in production would inevitably lead to leakage of privacy-sensitive details of data subjects (*14*). Thus, choices need to be made as to the extent that transparency is penalized to gain privacy, and vice versa, and such choices need to be made clear to system purchasers and users. Generally, privacy concerns prevail because of the legal implications if they are not enforced (for example, patient privacy with respect to the Health Insurance Portability and Accountability Act in the United States). Also, privacy and fairness may not always develop synergy. For example, although privacy-preserving ML (such as differential privacy) provides a bounded guarantee on indistinguishability of individual training examples, in terms of utility, research shows that minority groups in the training data (for example, based on race, gender, or sexuality) tend to be negatively affected by the model outputs (*15*).

**"...the lessons learned from the ML attack-defense arms race provide opportunities to motivate broader efforts to make ML truly trustworthy..."**

Broadly speaking, the scientific community needs to step back and align the robustness, privacy, transparency, fairness, and ethical norms in ML with human norms. To do this, clearer norms for robustness and fairness need to be developed and accepted. In research efforts, limited formulations of adversarial robustness, fairness, and transparency must be replaced with broadly applicable formulations like what differential privacy offers. In policy formulation, there needs to be concrete steps toward regulatory frameworks that spell out actionable accountability measures on bias and ethical norms on datasets (including diversity guidelines), training methodologies (such as bias-aware training), and decisions on inputs (such as augmenting model decisions with explanations). The hope is that these regulatory frameworks will eventually evolve into ML governance modalities backed by legislation to lead to accountable ML systems in the future.

Most critically, there is a dire need for insights from diverse scientific communities to consider societal norms of what makes a user confident about using ML for high-stake decisions, such as a passenger in a self-driving car, a bank customer accepting investment recommendations by a bot, and a patient trusting an online diagnostic interface. Policies need to be developed that govern safe and fair adoption of ML in such high-stake applications. Equally important, the fundamental tensions between adversarial robustness and model accuracy, privacy and transparency, and fairness and privacy invite more rigorous and socially grounded reasonings about trustworthy ML. Fortunately, at this juncture in the adoption of ML, a consequential window of opportunity remains open to tackle its blind spots before ML is pervasively deployed and becomes unmanageable. ■

## REFERENCES AND NOTES


1. I. Goodfellow, P. McDaniel, N. Papernot, *Commun. ACM* **61**, 56 (2018).
2. S. G. Finlayson *et al.*, *Science* **363**, 1287 (2019).
3. B. Biggio, B. Nelson, P. Laskov, *Proceedings of the 29th International Conference on Machine Learning*, Edinburgh, Scotland, UK, J. Langford and J. Pineau, Eds. (Omnipress, 2012), pp. 1807–1814.
4. K. Eykholt *et al.*, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2018), pp. 1625–1634.
5. F. Tramèr, F. Zhang, A. Juels, M. K. Reiter, T. Ristenpart, *Proceedings of the 25th USENIX Security Symposium*, Austin, TX (USENIX Association, 2016), pp. 601–618.
6. A. Ali, B. Eshete, *Proceedings of the 16th EAI International Conference on Security and Privacy in Communication Networks*, Washington, DC (EAI, 2020), pp. 318–338.
7. R. Shokri, M. Stronati, C. Song, V. Shmatikov, *Proceedings of the 2017 IEEE Symposium on Security and Privacy*, San Jose, CA (IEEE, 2017), pp. 3–18.
8. N. Papernot, M. Abadi, U. Erlingsson, I. Goodfellow, K. Talwar, *arXiv*:1610.05755 [stat.ML] (2017).
9. I. Jarin, B. Eshete, *Proceedings of the 7th ACM International Workshop on Security and Privacy Analytics* (2021), pp. 25–35.
10. J. Buolamwini, T. Gebru, *Proceedings of Conference on Fairness, Accountability and Transparency*, New York, NY (ACM, 2018), pp. 77–91.
11. A. Birhane, V. U. Prabhu, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (IEEE, 2021), pp. 1537–1547.
12. N. Carlini *et al.*, *arXiv*:1902.06705 [cs.LG] (2019).
13. N. Papernot, P. McDaniel, A. Sinha, M. P. Wellman, *Proceedings of 3rd IEEE European Symposium on Security and Privacy* (London, 2018), pp. 399–414.
14. R. Shokri, M. Strobel, Y. Zick, *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, New York, NY (2021); https://www.comp.nus.edu.sg/~reza/files/Shokri-AIES2021.pdf.
15. V. M. Suriyakumar, N. Papernot, A. Goldenberg, M. Ghassemi, *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (ACM, 2021), pp. 723–734.


10.1126/science.abi5052

RETROSPECTIVE

# Richard C. Lewontin (1929–2021)

## Groundbreaking evolutionary geneticist

*By* **Andrew Berry**[1] *and* **Dmitri A. Petrov**[2]

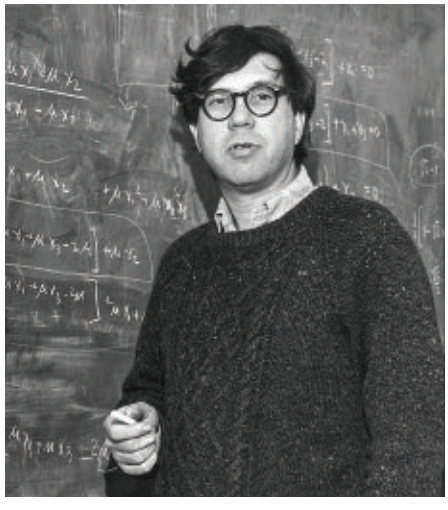
Richard Charles "Dick" Lewontin died at age 92 on 4 July, 3 days after Mary Jane, his wife of 73 years. Arguably the most influential population geneticist of the second half of the 20th century, Lewontin laid the theoretical and experimental foundations of modern evolutionary genetics. He was also a prominent social critic and philosopher of science.

Born 29 March 1929 in New York City, Lewontin earned his bachelor's degree in biology from Harvard University in 1951. He then joined Theodosius Dobzhansky's genetics and evolutionary biology lab at Columbia University, completing his MS in mathematical statistics in 1952 and his PhD in zoology in 1954. In the years that followed, he held positions at North Carolina State University in Raleigh, the University of Rochester in New York, the University of Chicago, and, from 1973 until his retirement in 1999, Harvard. Throughout his academic career, Lewontin pursued what he called the central "problematic" of population genetics: describing patterns of genetic variation in natural populations and understanding the evolutionary forces that shape them.

Because of the dearth of genetic markers, population genetics in the 1950s was strong on theory but weak on data. Then in 1966 came Lewontin and biochemist John Lee "Jack" Hubby with protein gel electrophoresis. This simple technique—detecting allelic differences in proteins that change their electrophoretic mobility—permitted the collection of protein polymorphism data in any species. On finding unexpectedly large amounts of genetic variation in populations, Lewontin and Hubby suggested that much of it might be selectively neutral. In biologist Motoo Kimura's hands, this idea evolved into the neutral theory of molecular evolution, which became the dominant paradigm in population genetics for decades to come. The unexpectedly similar levels of diversity across species revealed by these studies—the "Lewontin paradox"—remains a puzzle still.

In 1983, a graduate student in his lab, Martin Kreitman, completed Lewontin's 30-year quest to quantify genetic variation, with the first study of DNA sequence variation in natural populations. Lewontin does not appear as an author on that paper; he refused to list himself as such unless he had contributed substantially to the research.

Lewontin's theoretical work complemented his empirical work. He introduced game theory to population biology in 1961 and pioneered computer simulation in the field. With Ken-ichi Kojima, Lewontin coined the term "linkage disequilibrium" to describe the correlation among alleles at multiple loci, an idea central to both modern genetic mapping studies and population genomics. With Jesse Krakauer, Lewontin devised an early statistical test to distinguish between selective and neutral evolution, noting that demographic factors such as bottlenecks affect variation throughout the genome, whereas selection affects only targeted loci. The test had problems, but it underpins modern approaches to genome evolution. Perennially interested in the interaction between the organism and its environment, Lewontin also pioneered niche construction theory.



Lewontin worked primarily on *Drosophila* flies but also coauthored important papers on mice and crickets. In 1972, he took on humans, using relatively crude data derived from serology and allozymes to map genetic variation. His finding was important and counterintuitive: 85% of all human genetic variation can be found within a single population, with only 7% segregating among continental groupings (or "races"). In other words, genetic variation shows remarkable uniformity across the human species. Subsequent studies confirmed these results. Fifteen years later, geneticist Rebecca Cann provided context: Our species is young and only recently ventured out of Africa, meaning that all human populations are closely related.

A brilliant and unsparing critic, Lewontin found much to object to in applications of biological determinism to humans. Famously, he and Stephen Jay Gould savaged their Harvard colleague, biologist E. O. Wilson, after Wilson incorporated humans into *Sociobiology*, his 1975 survey of the adaptive basis of animal behavior. For Lewontin, the assertions of Wilson and others were naïve and dangerous oversimplifications.

With Gould, Lewontin went on to write an eloquent reminder of the pitfalls of one-dimensional adaptationist interpretation. Their wildly influential essay on the spandrels of San Marco is an appeal for pluralistic thinking. Lewontin wrote extensively (with crystal clarity) for the public both in his books and as a regular contributor to the *New York Review of Books*.

Lewontin, a Marxist who resigned from the National Academy of Sciences to protest its involvement in military research, has been accused of tainting his science with his politics. If anything, though, the "dialectical" approaches that he championed allowed him to remain open-minded yet rigorous when thinking about complex biological patterns. He was dogmatically nondogmatic.

His Marxism was arguably a key component of his scientific success. At Harvard, he designed his own research space, assigning the desirable corner office to the *Drosophila* "kitchen," where fly food was prepared. The worst job, Lewontin insisted, should be done in the nicest part of the lab. The lab's centerpiece was an enormous table for meetings, rather than the wet lab. Science, Lewontin believed, was forged through discussion. At that table, Lewontin—animated, hands waving, spectacles awry—steered conversations that rampaged across disciplines spanning science, philosophy, politics, history, and sociology.

Lewontin's legacy extends far beyond that table. He directly trained generations of scientists, including both of us, and inspired countless more. Despite his disdain for scientific celebrity, he received many awards, including the 2015 Crafoord Prize (shared with theoretical geneticist Tomoko Ohta). He also contributed to unexpected fields, collaborating in the 1950s, for example, with architect Buckminster Fuller to design geodesic domes. However, Lewontin was much more than just the sum of his intellectual contributions; he was a loyal mentor and a beloved friend. ■

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [2]Department of Biology, Stanford University, Stanford, CA, USA.
Email: berry@oeb.harvard.edu

10.1126/science.abl5430

PHOTO: THE ERNST MAYR LIBRARY AND ARCHIVES OF THE MUSEUM OF COMPARATIVE ZOOLOGY, HARVARD UNIVERSITY

POLICY FORUM

BIODIVERSITY

# Integrate biodiversity targets from local to global levels

## A shared Earth approach links biodiversity and people

By **David O. Obura[1,2,3], Yemi Katerere[4], Mariam Mayet[5], Dickson Kaelo[6], Simangele Msweli[7], Khalid Mather[8], Jean Harris[8,9], Maxi Louis[10,11], Rachel Kramer[8], Taye Teferi[12], Melita Samoilys[1,3,13], Linzi Lewis[5], Andrew Bennie[5,14], Frederick Kumah[7], Moenieba Isaacs[15], Pauline Nantongo[16]**

Decisions to be made at the 15th Conference of the Parties (COP 15) to the Convention on Biological Diversity (CBD) will shape biodiversity conservation approaches for the next 30 years, a critical time for the future of nature and people. Reflecting from our African perspective, we applaud the necessary increase in ambition to conserve nature (*1*), but we share alarm about the limited equity and justice in establishment of protected areas and impacts on people (*2*–*6*). Further, raising the burden of protection in the Global South while failing to address global economic drivers of biodiversity decline will only repeat and amplify historical cycles, and effort invested in conservation will be wasted. We see hope in new and diversified approaches to conserved areas (*7*) and the development of other, less formal conservation mechanisms. Here we offer a framework that can help to integrate these with improved conventional conservation approaches.

Historically, conservation has emphasized protected areas in the most high-biodiversity and intact regions, supported by systematic conservation planning tools set to select low-use areas for their low cost and high biodiversity benefit, and isolating those places from people in strict no-use protected areas. A disproportionate burden of such protection has fallen on the Global South, where biodiversity condition is high yet resources and public support are limited (*6*); effective protection is hampered by trade-offs with meeting peoples' needs; histories of inequitable treatment of local and Indigenous communities and small-scale food producers exist; and there is a higher dependence on nature for food, livelihoods, and culture. Only very recently has it been acknowledged that over half of the high-value land for conservation is traditionally owned, used, or occupied by Indigenous peoples and local communities (IPLCs), acting as the de facto managers (*8*). The increased conservation ambition in the past decade, and now being projected into the future, has focused on protecting very large and intact wilderness areas on land and at sea, many of which overlap with these IPLC spaces. A consequence of this overlap is increasing concern with control of territory, undermining tenure of IPLC, and weakening their links with nature.

The 2011–2020 strategy of the CBD established the 20 Aichi Targets, and although conservation actions have been essential to prevent some species extinctions and ecosystem declines, effort and resourcing have been inadequate and none of the Aichi Targets were achieved (*9*). Now the next set of 10- and 30-year goals and targets for biodiversity are under negotiation in the post-2020 global biodiversity framework (GBF). Its current draft specifies four goals and 21 targets (*10*), and its final form will be determined by countries at COP 15. One of the 21 targets is to raise the area under protection to 30% of land and sea. This builds on two of the prior Aichi subtargets, on area protection (17% of land and 10% of ocean; see the figure), that were nominally successful. However, their success was not matched by achievement of linked effective management and biodiversity representation subtargets (as many protected areas are too degraded, a large proportion of protected areas are suboptimally located, and investment in management was insufficient, to protect biodiversity), nor by equity or sustainable use targets (*9*).

Momentum supporting increased targets for area under protection to 30 and 50% of global land and sea is growing, both through scientific publications supported by new technologies and data streams (*1*) and by campaigns (e.g., Campaign for Nature, High Ambition Coalition for Nature and People) based on and supporting this science. This has amplified concern about the equity issues raised above (*5*), as well as the potential impacts of higher protection on meeting peoples' needs (*2*–*5*). An important factor in this debate is the lack of diversity in mainstream scientific and conservation voices (*5*, *7*, *11*), resulting from a combination of barriers to publishing and communicating from varied cultural and societal contexts. For example, local and Indigenous knowledge is held within communities, not in the scientific literature, and many conservation practitioners and even scientists from non-Western cultures may rarely publish in accessible and mainstream literature. This unequal representation of conservation perspectives and approaches in mainstream conservation science, and concern over perverse interpretations of area-based analyses (*2*), are particularly important where the scientific literature has a strong role in developing and supporting policy, as in multilateral processes such as the GBF. This article seeks to at least partially redress this imbalance in the literature.

Africa has a large share of remaining intact biodiversity; a small share of global funds for protection; a youthful demographic presenting both challenges and opportunities for conservation; a growing but nonetheless limited technical, scientific, and implementation capacity; and a suppressed voice in global negotiations. Africa faces particularly strong trade-offs between biodiversity protection and poverty alleviation, with many of the voices expressing concern with how protected areas have been implemented across this continent. This experience is shared across the Global South, and to succeed, the GBF must resolve these challenges in these places in socially relevant ways.

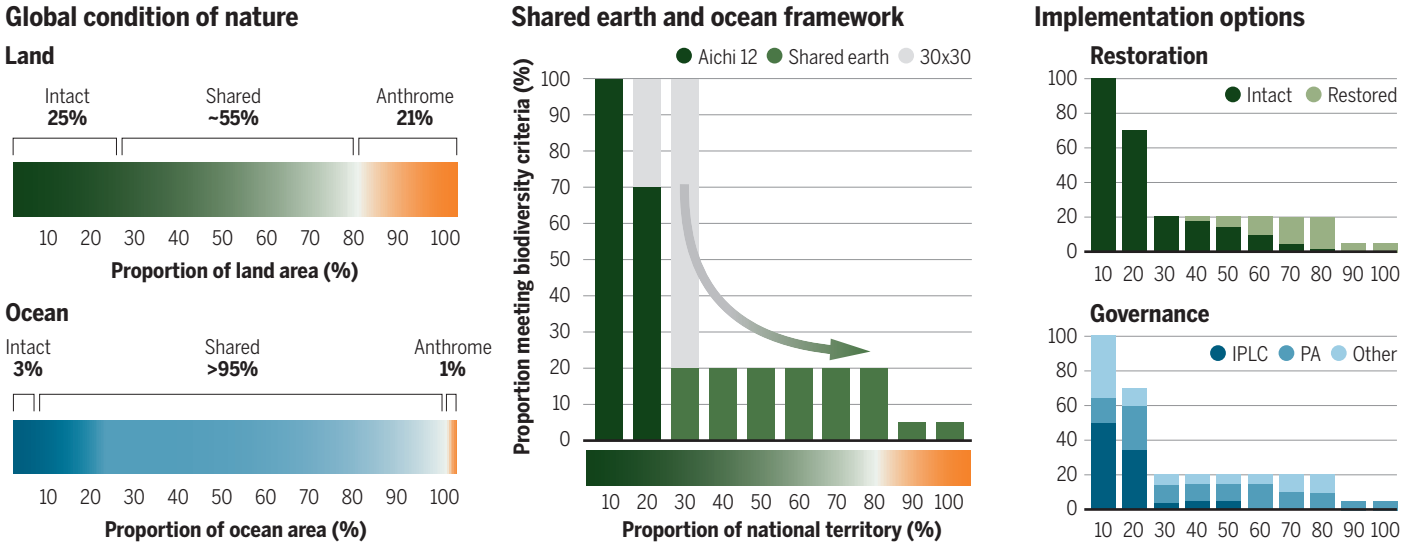
### SHARED EARTH, SHARED OCEAN

Our framework grounds the GBF through the concept of "shared" land and seascapes (*3*, *8*, *12*), connecting people and nature rather than separating them. The condition of nature varies across all land and ocean

[1]CORDIO East Africa, Mombasa, Kenya. [2]University of Queensland, Brisbane, Australia. [3]Pwani University, Kilifi, Kenya. [4]Independent: Environmental and Policy Expert, Harare, Zimbabwe. [5]Africa Centre for Biodiversity, Johannesberg, South Africa. [6]Kenya Wildlife Conservancies Association, Nairobi, Kenya. [7]African Wildlife Foundation, Nairobi, Kenya. [8]Wildlands Conservation Trust, Pietermaritzburg, South Africa. [9]Institute for Coastal and Marine Research (CMR), Nelson Mandela University, Port Elizabeth, South Africa. [10]Community Leadership Network (CLN) of Southern Africa, Windhoek, Namibia. [11]Namibian Association of Community Based Natural Resources Associations (NACSO), Windhoek, Namibia. [12]Independent: Conservation Scientist, Nairobi, Kenya. [13]University of Oxford, Oxford, UK. [14]Department of Sociology, University of the Witwatersrand, Johannesburg, South Africa. [15]Institute for Poverty, Land and Agrarian Studies (PLAAS), University of the Western Cape, Cape Town, South Africa. [16]EcoTrust, Kampala, Uganda. Email: dobura@cordioea.net

## Shared Earth, shared ocean framework

(Left) The approximate proportion of land or ocean area in 2020 in which nature is intact, variably affected in shared spaces, or fully altered in anthromes [from (*2*)]. (Middle) Schematic allocation of effective conservation actions on land across the gradient of condition (as in left panel) in a country or territory. Protection of 17% of total land under Aichi Target 12 is depicted in the 0 to 20% of land where nature is most intact. Protecting 20% of area under intact native habitat is shown in shared spaces, from 21 to 80% of territory. Protecting 5% of area under intact habitat in anthromes is shown from 81 to 100% of territory. The sum of these meets the draft global biodiversity framework target of 30% protected. Relative to the 30x30 campaign for protecting nature in areas most important for biodiversity, the shared Earth approach spreads effort and benefits of additional conservation across space. (Right) The contribution of restored versus intact habitat will increase in more altered shared spaces and in anthromes. The contribution of different governance regimes, such as by Indigenous people and local communities (IPLCs), conventional protected areas (PAs), and "other" mechanisms, may vary across spaces. Axes labels are as in the middle panel.




(see the figure), from near-intact and "natural" in remote areas with low or no human population (e.g., areas within the Amazon forest, Sahara desert, or high seas), through increasingly modified in lightly to densely populated and used shared spaces that still present functions and characteristics of a natural biome (e.g., extensive pasturelands, or sparsely to moderately populated rural areas or fished seas), to fully altered in high-intensity agricultural and urban spaces, aquaculture, and intensively modified coastlines (classed as anthropogenically altered biomes or "anthromes"). Strict boundaries between categories of such "spaces" are challenging to define, emphasizing the continuum of nature–human interactions across them (*8*, *12*). Across all of these spaces, even where nature is in a degraded state, it provides essential benefits to people (*13*), particularly to those living in poverty or with few material or financial assets. Whereas the dominant paradigm for nature conservation to date has focused on the most intact spaces for protection, we focus on the middle ground, where human interactions with nature cannot be resolved by separation.

Our approach is built on four pillars. First, it focuses at the local scale, from which decisions and measures are aggregated "from the ground up" to achieve targets. Second, it applies equity principles to ensure that needs and rights of people are met, prioritizing local institutions and rights holders, assisted by governments, organizations, scientists, and others. Third, all relevant knowledge is integrated at this level. The local focus means that local and traditional knowledge, accumulated experience, and social and cultural concerns can be in the forefront. Relevant scientific knowledge needs to be translated to this scale, using platforms that can help ensure consistency of local decisions across locations and across scales, and to address larger-scale phenomena such as connectivity. Fourth, all relevant targets of the GBF should be addressed concurrently.

Recent advances in the conservation literature connecting nature and people support this approach (*2–4*, *6–8*, *12–14*), in particular, to maintain 20% of local area under intact native habitat (*13*), notably in "shared spaces" (*12*) (see the figure). Doing this can adequately meet peoples' needs and maintain biodiversity functions, especially when applied down to a square kilometer scale so that benefits are within reach of those that need them. From this basis, multiple biodiversity targets relating to species, ecosystems, genes, and benefits to people (*15*) may be addressed together, potentially from both the 20% of intact habitat and 80% of "working" or "managed" area in the local land- or seascape (*12*) (see the figure).

In moderately and highly altered "shared spaces," such as extensive populated rural landscapes [e.g., tea-farming landscapes in central Kenya, or cocoa and mixed agroforestry landscapes in southern Cameroon (*14*)], the proportion of intact nature may be well below 20% of the landscape. But with effective restoration and appropriate time horizons to allow for rebuilding of ecological integrity and function (*15*), restored habitats could count toward conservation goals (see the figure), as well as increase their contribution of benefits to people.

In anthromes, a lower "intact habitat" threshold might be all that is possible, and decision-makers may focus on selected benefits, such as shading and temperature control, or on nonmaterial benefits provided by urban green spaces. We illustrate an arbitrary area target of 5% (see the figure), but existing guidance of 9 to 50 $m^2$ of green space per person may require variable locally set targets.

Nature is most intact and human population density lowest in intact spaces and lightly altered shared spaces, where the proportion of land or sea meeting biodiversity criteria may approach 100% (see the figure). This area can enable the summed area of high protection to approach a global target, such as 30%. Where high levels of protection are already established, our approach can help refine management to address local needs within the broader context and potentially redress equity and rights issues that may be unresolved from the past (*5*).

GRAPHIC: K. FRANKLIN/*SCIENCE*

Local government (e.g., districts, counties, states, municipalities) may provide the relevant framework for applying this approach, enabling consistent replication, locally differentiated targets, technical support, and resourcing within national systems. Within these, a wide variety of governance models could be applied to protected and conserved areas (see the figure), such as Indigenous and local community; private, local, or central government; or mixed, depending on national and local rights and tenure regimes. "Protected areas" are generally designated under national legislation and classified using International Union for Conservation of Nature categories. Less formal conservation mechanisms (e.g., Indigenous and community lands, private sanctuaries, etc.) are becoming classified under emerging standards for other effective conservation measures (OECMs) (*7*), though full inclusion and leadership by IPLCs in this process will be necessary for legitimacy. Our framework will facilitate integration of protected areas, OECMs, and other measures across intact, shared, and anthrome spaces.

This framework integrates the three pillars of the CBD and helps operationalize the goals and many of the targets of the GBF simultaneously: the equity intent of CBD objective 3 and GBF goal C (which relate to equitable access to and benefit sharing of genetic resources), extending this to all aspects of nature and its contributions to people; CBD objective 2 and GBF goal B on ensuring access to and provisioning of a range of benefits from nature; and CBD objective 1 and GBF goal A on conserving ecosystems, species, and genetic diversity, and restoring them where needed.

The framework can also help direct conservation finance to local institutions and actors, addressing goal D of the GBF. Objectives, expected results, and resourcing for conservation and protection vary along the continuum of condition of nature. Large, intact critical ecosystems—such as in wilderness areas and the high seas—play essential roles in global regulatory functions (e.g., carbon sequestration), as well as locally. Small habitat fragments in shared spaces and anthromes will be critical for daily contributions to peoples' quality of life and well-being (e.g., food security, pollination, water treatment) and in ensuring connectivity across land- or seascape mosaics, and with protected areas. The responsibilities and mechanisms for management and finance may vary across these scales. For example, small areas might be managed and supported based on functions such as maintaining integrity and connectivity, and on the spatial reach of the benefits they provide (e.g., through trade), as well as for their role in sustaining higher-level targets and global benefits (e.g., species conservation or carbon sequestration in community forests). Attention to local perceptions of justice and value, and innovative measures such as basic income grants for protected and conserved area-adjacent communities, can help balance the benefits and costs of conservation.

Looking into the future, conservation and restoration may need to focus on ecological functions and contributions to people rather than the prior intact assemblage, particularly where climatic and other changes make a historical native state impossible to reinstate. In this context, regenerative development and agroecological approaches linked to culture and Indigenous practices at land- and seascape scales may best integrate biodiversity, provisioning, and equity targets. By situating and aggregating local conservation within larger spatial frames, this approach can also facilitate biodiversity and human adaptation to climate change in landscape mosaics that facilitate climate migration.

## AN APPROACH FIT FOR THE GLOBAL BIODIVERSITY FRAMEWORK

This "shared Earth and ocean" approach meets a need identified within the GBF, for its global ambitions and elements to be translated down to the local scale, founded on inclusion and equity, where the best science and praxis can be integrated in locally relevant ways to meet multiple targets simultaneously. The interests of local communities and rights holders require locally contextualized solutions for conservation, across intact, shared, and fully altered spaces, which can be aggregated and reported to larger scales. It also offers a legitimate and people-focused way to spread effort for conservation across the whole gradient of condition of nature (*12*) rather than isolating it away from people in only the most intact regions. Like many innovative approaches, many of its elements are known and established, but with a paradigmatic shift in focus—from the ground up, and with equity and provisioning of benefits to people as priorities—the potential for success is transformed.

The ambition and complexity of efforts required to achieve the GBF will be considerably greater than for the Aichi Targets. We see merit in the full set of Targets proposed for the GBF, mirroring the indivisibility of the Sustainable Development Goals, not in any one target in isolation. For example, although protecting 30% of nature is likely necessary to adequately reduce biodiversity loss (*1*), if it is pursued without equal attention to equity, local custodianship, and provisioning of benefits to people (*3*, *4*, *15*), errors of the past will likely be repeated (*5*, *6*), as well as the level of achievement of the Aichi Targets (*9*). Critics will question if our approach will succeed better, but a renewed approach is needed, and both societal (e.g., Leaders Pledge for Nature) and scientific (*5*, *6*, *8*) voices are calling for a paradigm shift toward sustainability, meeting peoples' needs, and equity that our approach helps to implement. We call on countries, donors, and organizations to shift to this focus on shared spaces and equity from local to global scales as the best way to reconcile the challenges to delivering on the GBF and a "nature positive" outcome by 2050.

A particular challenge will be the design and establishment of the multilevel and polycentric governance systems that will be needed for success, underpinned by spatial planning frameworks and incorporating diverse conservation and sustainable production actions that address the impacts and interests of multiple actors, sectors, and jurisdictions across spatial scales. However, without addressing the drivers of biodiversity decline arising from high-consumption practices that underpin global inequalities, no conservation-focused actions will be sufficient to halt or reverse biodiversity decline or attain sustainable use at a global scale. ■

### REFERENCES AND NOTES


1. J. E. M. Watson et al., *Nature* **578**, 360 (2020).
2. M. Barnes, L. Glew, C. Wyborn, I. D. Craigie, *PeerJ Preprints* 10.7287/peerj.preprints.26486v1 (2018).
3. Z. Mehrabi, E. C. Ellis, N. Ramankutty, *Nat. Sustain.* **1**, 409 (2018).
4. J. Schleicher *et al.*, *Nat. Sustain.* **2**, 1094 (2019).
5. A. Agrawal, *et al.*, "An open letter to the lead authors of 'Protecting 30% of the Planet for Nature: Costs, Benefits and Implications'" (2021); https://openletter-towaldronetal.wordpress.com/.
6. RRI, "Rights-based conservation: The path to preserving Earth's biological and cultural diversity?" (Technical Report, Rights and Resources Initiative, 2020), p. 43.
7. H. D. Jonas *et al*, *PARKS* **27**, 71 (2021).
8. IPBES, *The Global Assessment Report on Biodiversity and Ecosystem Services - Summary for Policymakers* (2019); https://ipbes.net/sites/default/files/2020-02/ipbes_global_assessment_report_summary_for_policymakers_en.pdf.
9. GBO5, "Global Biodiversity Outlook 5" (Convention on Biological Diversity, Montreal, Quebec, Canada, 2020), p. 212.
10. CBD, "First draft of the post-2020 global biodiversity framework," CBD/WG2020/3/3 (Convention on Biological Diversity, 2021), p. 12.
11. D. Veríssimo *et al.*, *Conserv. Soc.* **18**, 220 (2020).
12. H. Locke *et al.*, *Natl. Sci. Rev.* **6**, 1080 (2019).
13. L. A. Garibaldi *et al.*, *Conserv. Lett.* **14**, e12773 (2021).
14. P. A. Minang *et al.*, Eds., *Climate-Smart Landscapes: Multifunctionality in Practice* (World Agroforestry Centre, Nairobi, 2015).
15. S. Díaz *et al.*, *Science* **370**, 411 (2020).


### ACKNOWLEDGMENTS

We acknowledge comments from E. Coppenger, S. Anderson, E. Tambara, E. Omondi, C. Gordon, S. Mangubhai, and H. Dublin.

10.1126/science.abh2234

BOOKS *et al.*

Heiress Patty Hearst is escorted to court by two federal marshals in 1976.

PHOTO: BETTMANN/GETTY IMAGES

PSYCHOLOGY

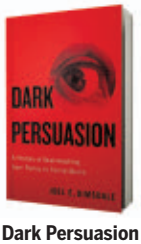
# The matter of mind control

## Brainwashing case studies illuminate the history of coercive persuasion

*By* **Sarah Marks**

In 1976, Patty Hearst, the granddaughter of American publishing magnate William Randolph Hearst, was found guilty of bank robbery, a crime she committed after enduring a sustained period of time as a captive of the domestic terrorist organization known as the Symbionese Liberation Army. The trial, with its glittering cast of expert witnesses, became a test case for psychological theories of brainwashing.

In addition to commenting on Hearst's intelligence and differentiating her behavior from those typically displayed by malingerers, the experts invoked the experiences of US prisoners of war (POWs) in Korea and the concept of "debility, dependency, and dread" to explain how Hearst, in their view, was not acting of her own free will when she committed the robbery. Instead, they argued, she was in thrall to the coercive persuasion of her captors. The jury was unconvinced and sentenced Hearst to 35 years in prison. However, President Jimmy Carter found the arguments persuasive and commuted her sentence after 22 months, and President Bill Clinton pardoned Hearst in 2001. This trial, and the debates surrounding it, is one of 10 key moments in the history of the idea of brainwashing examined by psychiatrist Joel E. Dimsdale in his new book, *Dark Persuasion*.




**Dark Persuasion**
*Joel E. Dimsdale*
Yale University Press, 2021. 304 pp.

A tragedy close to home inspired Dimsdale to dive deeply into this topic. In 1997, as the comet Hale-Bopp approached, his neighbors committed suicide at the instructions of the leaders of the Heaven's Gate cult, who were convinced that death was necessary to free members of their "bodily vehicles" and allow them to ascend to heaven.

The word "brainwashing" was coined in the early years of the Cold War by journalist Edward Hunter to describe reeducation techniques used for indoctrination in Communist China and was subsequently invoked to make sense of the defection of 21 American POWs to China after the Korean War. Elements of brainwashing can be traced back much further to techniques used in religious conversion and inquisition, but the fear unleashed by the provocative term precipitated the Central Intelligence Agency's infamous MKUltra experiments, which made use of the hallucinogen LSD, sensory deprivation, and electroconvulsive therapy during interrogations to coerce confessions and reached their bleak crescendo with a series of highly abusive experiments carried out on psychiatric patients by McGill University psychiatrist Donald Ewen Cameron between 1957 and 1964.

The book avoids making sweeping claims about the nature of brainwashing, but after a chronological assessment of each case study, Dimsdale presents a comparative table of the key features at play, including techniques such as sleep manipulation, coercion and manipulation, intentional surreptitiousness, and participation in activities not in the subject's best interests. That some combination of these are present in varying degrees in all of the cases cited suggests that Dimsdale sees them as necessary and sufficient criteria for describing brainwashing with some degree of certainty. Some of the 21st-century examples he discusses briefly in the book's last section, including controversies surrounding deep brain stimulation and the rise of social media platforms and conspiracy theories, do not quite fit these criteria, however.

While there is novelty in the synthesis of these case studies, *Dark Persuasion* does not offer much new material, and Dimsdale has not unearthed any substantial unexpected archival finds or generated new oral histories. However, his account of the Hearst trial is carefully researched and supplemented with a wealth of scientific papers from the time. And while the causal link he suggests between Pavlov's experiments with traumatized dogs and the interrogation and torture processes that led to false confessions in Stalin's show trials is based on slender evidence, Dimsdale's observations about the proximity of these events are nevertheless astute.

But historical rigor is not necessarily the point of this highly readable and compelling book. Dimsdale's goal is to prompt reflection on what he sees as the overlooked reality of coercive persuasion at a broader level and the ever-present threat that it poses to individuals and to society at large—a threat, he warns, that is becoming ever-amplified by new technologies and mass media. In this aim, he succeeds admirably. ■

The reviewer is at the Department of History, Classics and Archaeology, Birkbeck, University of London, Bloomsbury, London WC1E 7HX, UK. Email: s.marks@bbk.ac.uk

10.1126/science.abj8872

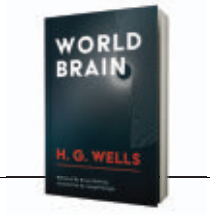
CLASSICS REVISITED

# The alternative to despair is to build an ark

## H. G. Wells's "world encyclopedia" has merit beyond its seeming similarities to Wikipedia

*By* **Yochai Benkler**

**World Brain**
*H. G. Wells*
*Foreword by Bruce Sterling and introduction by Joseph Reagle*
MIT Press, 2021. 176 pp.




Between 1936 and 1938, H. G. Wells delivered a series of lectures first published under the title *World Brain* in 1938. The standard reading of that text over the past few decades has framed Wells as the utopian futurist projecting from microfilm, radio, and telephone a world encyclopedia freely available to all. Predicting, various authors have written, the World Wide Web or Wikipedia, *World Brain* naïvely imagined that such a facility would provide the global community with the knowledge base we would need to create world peace. After 5 years of research on propaganda and misinformation, I read *World Brain* very differently.

In one of the early talks collected in the book, Wells told his audience "I do not agree with that inevitability of another great war." But by 1938, he conceded: "The Flood is coming anyhow, and the alternative to despair is to build an ark. My other name is Noah, but I am like someone who plans an ark while the rain is actually beginning."

Fascism and communism loomed large in the world of propaganda that Wells decried. But his point was both deeper and more banal than these twin threats. "Big business in America," he wrote, "appears to be completely bankrupt of political and social philosophy. Probably it never had any. It had simply a set of excuses for practices that were for a time extremely profitable and agreeable." The reader cannot but think of the "merchants of doubt," as Naomi Oreskes dubbed them, selling climate denial or the harmlessness of tobacco, sugar, or opiates (*1*).

Yet Wells spent fewer pages decrying political or business propaganda, or exalting the benefits of microfilms, than he did criticizing our stagnant education and research systems. Overworked, underpaid elementary school teachers and university professors in silly gowns are weighted down by bureaucratic burdens and follow centuries-old teaching practices, he wrote, particularly on "the collegiate side of a contemporary university, the superficial finishing-school exercises of sportive young people mostly of the wealthier classes." "The universities," he lamented, "go out to meet the tremendous challenges of our social and political life, like men who go out in armour with bows and arrows to meet a bombing aeroplane. They are pushed aside by men like Hitler, Mussolini creates academies in their despite, Stalin sends party commissars to regulate their researches."

What, then, is Wells's ark? "What I am saying," he writes, "is…that without a World Encyclopaedia to hold men's minds together in something like a common interpretation of reality, there is no hope whatever of anything but an accidental and transitory alleviation of any of our world troubles."

Authoritative reports from bodies like the IPCC align well with Wells's vision.

Wells's "world encyclopedia" is primarily organizational, not technological. He envisioned thousands of experts continuously engaged in a global series of workshops and conferences refining a body of authoritative knowledge. His world brain, read so, is less like Wikipedia and more like a general-purpose Intergovernmental Panel on Climate Change (IPCC)—a collaborative global network of expert researchers from universities and public bodies working to produce an authoritative statement on a matter of core global concern designed to guide action.

We face an epistemic crisis. The world is awash in falsehoods rooted in similar dynamics to those Wells described in his own time: religious fundamentalism, corporate opportunism, and the politics of racial and ethnonationalist hate. Hundreds of millions of people reject evolution and science; reject vaccines in the teeth of a devastating pandemic; and refuse to believe in climate science, droughts, fires, floods, and famines notwithstanding.

Wells urges us to build the ark that we can rather than despair or waste time on designing arks that we cannot build. There is no world in which the few, the expert, declare what is true and enforce it on the many. Wells himself wrote, "A professor-ridden world might prove as unsatisfactory under the stress of modern life and fluctuating conditions as a theologian-ridden world."

What universities and research institutes can do is to produce the most conscientious "world encyclopedia" possible—a statement, continuously updated, of what we know to be true, what we know to be untrue, and what we believe to be in reasonable doubt—and to translate that constantly updated consensus into teaching materials for diverse levels of education and other broadly intelligible and freely accessible formats. Such an effort would need to be publicly funded and operate internationally, so as to be resistant to corporate manipulation and the influence of any single nation's interests.

In some fields, such as the physical and biological sciences, this task will be hard. In others, such as history, sociology, or economics, it may be impossible to offer definitive answers to some questions. But clever tricks to fix social media will not address a problem with roots deeper and broader than the tweet-length blink of our technological moment. And we must do what we can. The waters are rising. ■

The reviewer is at Harvard Law School, Cambridge, MA 02138, USA. Email: ybenkler@law.harvard.edu

### REFERENCES AND NOTES

1. N. Oreskes, E. M. Conway, *Merchants of Doubt* (Bloomsbury, 2010).

10.1126/science.abk0210

PHOTO: FABRICE COFFRINI/AFP/GETTY IMAGES

LETTERS

Algae covers the beach in Qingdao, Shandong Province, China.

PHOTO: RUISHAN CHEN

*Edited by* **Jennifer Sills**

# China's algal bloom suffocates marine life

Every summer since 2007, algal blooms have grown in China's Yellow Sea (*1*). This year, covering about 1746 km$^2$, the bloom is 2.3 times larger than the country's previous record-holding bloom in 2013 (*2*). Such massive quantities of algae block sunlight from entering the ocean and deplete oxygen levels, suffocating marine life (*3*). The algae also pose challenges for tourism and marine transport. The city of Qingdao has deployed 12,686 boats to clean the water, collecting 457,700 tons of algae by 12 July (*2*). The algae are expected to persist until mid-August (*4*), at enormous economic and biological cost. Mitigating the damage will require regional collaboration.

The massive algae bloom is the result of a complex interaction between excessive coastal use for aquaculture, climate change, and coastal eutrophication. Booming seaweed aquaculture businesses in neighboring Jiangsu province are alleged to be the primary culprit for algal increases (*5*), which are then transported to coastal areas such as Qingdao by ocean currents and wind (*6*). Seaweed aquaculture in Jiangsu has increased by more than an order of magnitude since 2000, resulting in commensurate increases in algae production (*7*). Compounding the trend, warming ocean temperatures favor the growth and expansion of algae; increases in extreme weather events, such as storms, destroy the infrastructure that algal matter attaches to, facilitating its spread to the sea. Meanwhile, coastal eutrophication increases nitrate and phosphate levels, intensifying algal blooms (*8*). Other human uses of the land, including large-scale fisheries, land reclamation, and resource extraction, compete for land in this coastal area, further exacerbating algal production. As increasing eutrophication combines with climate change and human use (*9*, *10*), algae blooms will continue to threaten Yellow Sea marine life in the years to come.

Integrated actions are needed to address future algal blooms. Water quality along the coast should be monitored and pollution controlled to reduce eutrophication. An early warning system for algal blooms should be established, including public engagement throughout the algal control process. Moreover, regional coastal and ocean industrial development should be better coordinated under China's marine functional zoning and ecological red-line policies, which divide marine areas into different types of basic functional areas and legislate to protect them. This supervision helps guide the development of the marine industry and control the expansion of aquaculture, but coastal and ocean industrial development covers different marine functional areas, requiring increased coordination between provinces (*11*). Finally, an ecological compensation mechanism should be established at the regional level (*4*). Upstream provinces that are the source of the algal blooms, such as Jiangsu province, should compensate downstream provinces that bear the ecological and economic costs, such as Shandong province. Alternatively, downstream provinces should compensate upstream provinces for reducing algal flows to below a certain threshold. Controlling China's algal blooms requires regional collaborative governance to better manage the development of the seaweed industry and its environmental impacts.

**Xiaona Guo[1], Annah Zhu[2], Ruishan Chen[1]***
[1]Shanghai Jiaotong University, Shanghai 200040, China. [2]Wageningen University, Wageningen 6706KN, Netherlands.
*Corresponding author.
Email: chenrsh04@gmail.com

## REFERENCES AND NOTES

1. Y. Xiao *et al.*, *Mar. Pollut. Bull.* **140**, 330 (2019).
2. F. Yang, Y. Hu, "Qingdao algal bloom reached the highest value in history, experts say it may exist offshore for a long time," *China News* (2021); www.chinanews.com//sh/shipin/cns/2021/07-19/news895145.shtml [in Chinese].
3. S. Dineva, *Environ. Sci. Pollut. Res.* 10.1007/s11356-021-13475-8 (2021).
4. X. Chen, C. Wang, "The algal bloom in Qingdao will last until mid-August, and political consultative committee members call for an 'international chess game' to manage it," *Newspaper of the Chinese People's Political Consultative Conference* (2021); www.rmzxb.com.cn/c/2021-07-07/2898577.shtml?n2m=1 [in Chinese].
5. E. Rees, "Algal blooms fed by climate change, farm pollution and aquaculture," *China Dialogue* (2014); https://chinadialogue.net/en/climate/7271-algal-blooms-fed-by-climate-change-farm-pollution-and-aquaculture/.
6. M.-J. Zhou *et al.*, *Estuar. Coast. Shelf Sci.* **163**, 3 (2015).
7. Q. Xing *et al.*, *Remote Sens. Environ.* **231**, 111279 (2019).
8. E. Marris, *Nature* 10.1038/news.2008.998 (2008).
9. G. M. Hallegraeff *et al.*, *Commun. Earth Environ.* **2**, 117 (2021).
10. Y. Wang *et al.*, *Harmful Algae* 10.1016/j.hal.2021.102058 (2021).
11. W. Lu *et al.*, *Mar. Pol.* **62**, 94 (2015).

10.1126/science.abl5774

# Eastern Europe's fraught waterway plans

In December 2020, the Ukrainian parliament passed the Inland Water Transport Act, paving the way for the construction of the E40 international waterway, Europe's longest water route (*1*). The proposed 2000-km waterway would connect the Baltic Sea port in Gdansk, Poland, with the Black Sea port of Kherson, Ukraine, likely within a decade. The waterway would include parts of the Vistula, Wieprz, Bug, Pina, Pripyat, and Dnieper rivers (*2*, *3*), posing a threat to the wetlands of Polesia, an ecosystem that has been referred to as Europe's Amazon (*4*). The Ukrainian government supports this project as a symbol of geopolitical connection to Europe during the country's conflict with Russia (*3*) and will likely fund the first part of the construction in its 2022 or 2023 budget, but thorough ecological assessments should take place before the project moves forward (*2*, *3*).

The proposed E40 would require construction in protected areas such as Polesie State Radioecological Reserve in Belarus and Mizhrichynsky Landscape Park and Chernobyl General Zoological Reserve (part of the Chernobyl Radioecological Reserve) in Ukraine. In April, the Ukrainian government removed the E40 from the updated draft 2030 Chernobyl Exclusion Zone Development Strategy (*5*), allowing construction work on nearby river bottoms that otherwise would have been forbidden. The planned waterway would pass within 2 km of the former Chernobyl nuclear plant (*5*, *6*). Now that an exception has been made for its construction, the project will likely bring radionuclides that were emitted during the Chernobyl disaster from river bottoms to the surface (*5*, *6*).

Poland also strongly supports the planned construction (*6*, *7*), despite the risks the project poses to the country's protected areas, including 12 Natura 2000 areas, 1 national park, 4 landscape parks, and 24 nature reserves (*2*). In addition, the E40 would deprive the 772-km Bug river—the last and the longest unregulated river in Europe (*2*, *3*, *6*, *7*)—of its ability to perform ecosystem services, threatening the drinking water supply for Warsaw's 1.8 million inhabitants by compromising the river's role in treating contaminated water flowing to Poland from Ukraine (*3*, *6*, *7*).

The E40 waterway is a serious threat to vulnerable species, such as the aquatic warbler (*Acrocephalus paludicola*) (*8*), and ecosystems that survived the communist period in Eastern Europe (*2*). In Ukraine and Poland, there is a strong political will to implement the project (*1*, *6*), but areas of global and regional environmental importance should not be put under pressure for political reasons. Scientists must work to prevent the coming habitat destruction by appealing to the governments of these countries to postpone construction until the environmental consequences are better understood.

**Ignacy Kitowski[1] and Grzegorz Grzywaczewski[2]***
[1]State School of Higher Education in Chelm, PL-22-100 Chelm, Poland. [2]University of Life Sciences in Lublin, PL-20-950 Lublin, Poland.
*Corresponding author. Email: grzegorz.grzywaczewski@up.lublin.pl

### REFERENCES AND NOTES

1. O. Shevchenko, "Ukraina przyjęła ustawę istotną dla odbudowy drogi wodnej E40" (2021); www.ecpp.org.pl/ukraina-przyjela-ustawe-istotna-dla-odbudowy-drogi-wodnej-e40 [in Polish].
2. G. Grzywaczewski, I. Kitowski, *Oryx* **53**, 4 (2019).
3. I. Kitowski, M. Oskierko, *Przeglad Geopolit*. **30** (2019) [in Polish].
4. P. Weston, "The race to save Polesia, Europe's secret Amazon," *The Guardian* (2020).
5. Save Polesia, "E40 waterway removed from Ukrainian Exclusion Zone Strategy" (2021); https://savepolesia.org/e40-waterway-removed-from-ukrainian-exclusion-zone-strategy.
6. P. Nowak, "Droga wodna E40 ma połączyć Bałtyk z Morzem Czarnym. Eksperci ostrzegają: To grozi katastrofą" (2020); https://kurierlubelski.pl/droga-wodna-e40-ma-polaczyc-baltyk-z-morzem-czarnym-eksperci-ostrzegaja-to-grozi-katastrofa/ar/c8-14940338 [in Polish].
7. Polish Society for the Protection of Birds, "Kosztowna droga wodna E40—ekspertyza ekonomiczna" (2020); https://otop.org.pl/2020/11/kosztowna-droga-wodna-e40-ekspertyza-ekonomiczna [in Polish].
8. G. Grzywaczewski, I. Kitowski, *Science* **365**, 134 (2019).

10.1126/science.abk2377

# Australia threatens to weaken forest laws

Victoria is one of Australia's most forested jurisdictions. The state supports 7.2 million ha of forest, of which 1.8 million ha are broadly allocated for logging (*1*). The native forests of Victoria are critical for water production, carbon storage, and biodiversity conservation (*2*). Victoria has taken steps to protect its natural forests by committing to phasing out all native forest logging across the state by 2030 and to substantially reducing current levels of cutting starting in 2024 (*3*). However, Victoria is now updating its forest code (*4*). Instead of strengthening much-needed protections, the state is considering changes that would weaken current regulations and put its forests in renewed jeopardy.

Motivated by rapidly dwindling timber supplies (*5*), policy-makers in Victoria have planned changes that will permit environmentally harmful practices that are currently prohibited. For example, logging is illegal within water catchments where terrain exceeds 30 degrees in slope (*6*, *7*) to limit erosion, conserve aquatic ecosystems, and ensure water quality for human consumption. Under the revised laws, these slope restrictions will be relaxed.

The current policy to protect forest on steep slopes, although scientifically sound, has been poorly enforced. Recent terrain analysis indicates that 75% of 204 logged sites (called cutblocks) in one Victorian water catchment were on areas steeper than 30 degrees; logging of these areas therefore breached current laws (*8*). A ministerial review (*9*) and legal cases have found that logging frequently breached codes of practice (*10*).

Australia has been pilloried for its record on climate inaction (*11*), biodiversity loss, and under-spending on species conservation (*12*). Weakened forest laws will further accelerate biodiversity decline. Victoria's government must resist industry pressure to degrade environmental laws and weaken codes of forest practice and instead focus on strengthening enforcement of current laws.

**David Lindenmayer and Chris Taylor**
Fenner School of Environment and Society, The Australian National University, Canberra, ACT 2601, Australia.
*Corresponding author. Email: david.lindenmayer@anu.edu.au

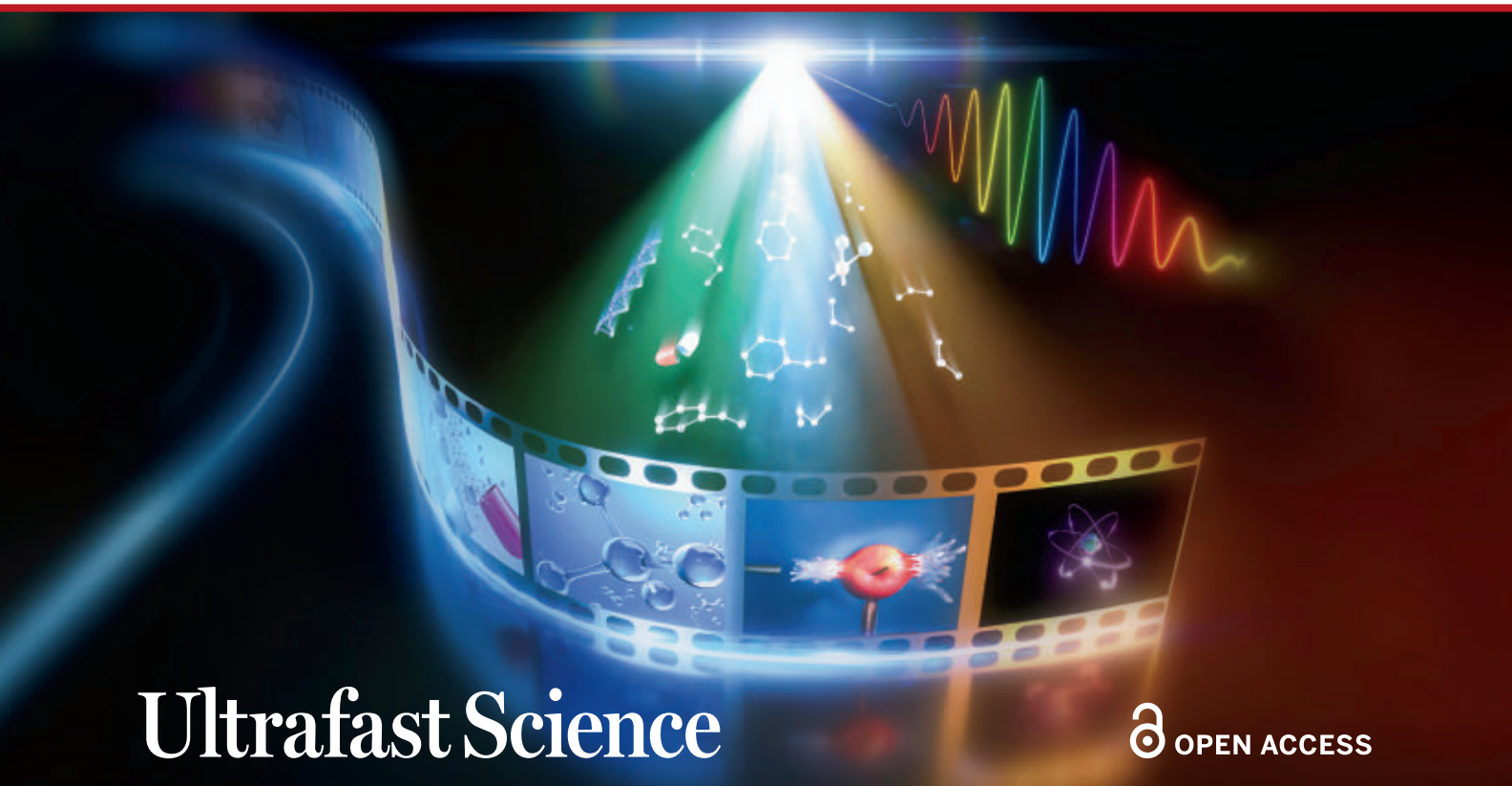
### REFERENCES AND NOTES

1. Commissioner for Environmental Sustainability Victoria, "State of the Forests 2018" (Government of Victoria, Melbourne, Australia, 2018); www.ces.vic.gov.au/reports/state-forests-2018.
2. H. Keith *et al*., *Nat. Ecol. Evol*. **1**, 1683 (2017).
3. Government of Victoria, "Timber harvesting regulation" (2021); www.vic.gov.au/timber-harvesting.
4. Victoria Department of Environment, Land, Water and Planning, "2021 proposed variation of the Code of Practice for Timber Production" (2021); https://engage.vic.gov.au/code-practice-timber-production.
5. Jaclyn Symes, "Review to protect Victoria's forests, jobs and timber industry" (2020); www.jaclynsymes.com.au/media-releases/review-to-protect-victorias-forests-jobs-and-timber-industry/.
6. Victoria Department of Environment and Primary Industries "Code of Practice for Timber Production 2014" (2014); www.forestsandreserves.vic.gov.au/__data/assets/pdf_file/0016/29311/Code-of-Practice-for-Timber-Production-2014.pdf.
7. Office of the Conservation Regulator, "Regulating timber harvesting on steep slopes" (Office of the Conservation Regulator, Melbourne, Australia, 2021).
8. C. Taylor, D. B. Lindenmayer, *Environ. Sci. Pol.* **120**, 204 (2021).
9. J. Brockington, N. Finnegen, P. Rozen, "Independent review of timber harvesting regulation: Panel report to the Secretary of the Department of Environment, Land, Water and Planning" (Victorian Government Library Service, Melbourne, Australia, 2018).
10. M. Jagot, J. E. Griffiths, R. M. Derrington, *FCAFC 92 Cost Order* (Federal Court of Australia, Melbourne, Australia, 2021).
11. M. Mazengarb, "Australia ranked dead last in world for climate action in latest UN report," Renew Economy (2021); https://reneweconomy.com.au/australia-ranked-dead-last-in-world-for-climate-action-in-latest-un-report/.
12. A. Waldron *et al*., *Nature* **551**, 364 (2017).

10.1126/science.abk3018

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS



# Ultrafast Science

OPEN ACCESS

*Ultrafast Science* is an online-only, Open Access journal published in affiliation with **Xi'an Institute of Optics and Precision Mechanics (XIOPM)** of CAS and distributed by the **American Association for the Advancement of Science (AAAS)**. The mission of the journal is to build a platform for ultrafast scientific research; present new theories, concepts, ideas, technologies and progress in this field; publish the latest research representing the forefront of the discipline and that of international concern; and promote academic exchange and development worldwide. *Ultrafast Science* will publish high-quality original research articles, comprehensive reviews, editorials, and perspectives which feature high novelty, significance and technical quality in cutting-edge and emerging topics in ultrafast science with broad interest from scientific communities.

**Submit your research to *Ultrafast Science* today!**

Learn more at **spj.sciencemag.org/ultrafastscience**

The Science Partner Journal (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit 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.

Visit us at **spj.sciencemag.org**

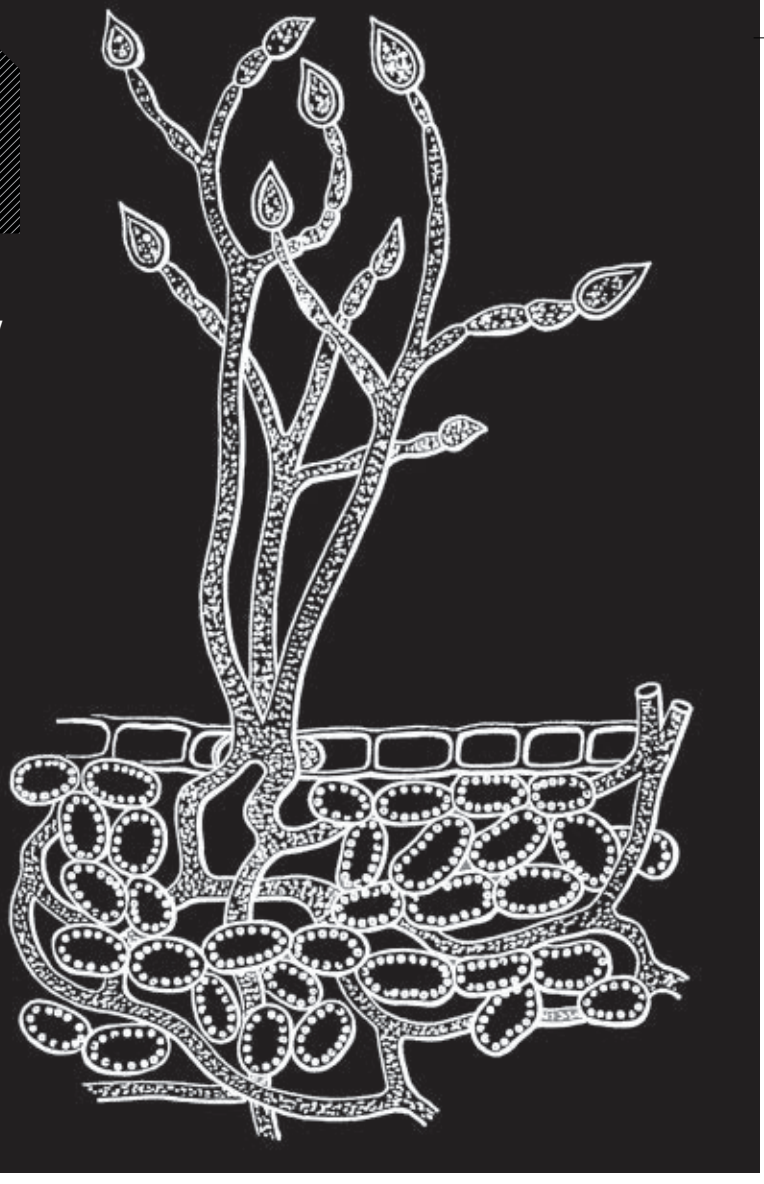
**@SPJournals** **@SPJournals**

ARTICLE PROCESSING CHARGES WAIVED UNTIL OCTOBER 2024

# RESEARCH

## IN *SCIENCE* JOURNALS

*Edited by* **Stella Hurtley**



PLANT SCIENCE

### Potato pectin falls to *Phytophthora*

*Phytophthora infestans* is a plant oomycete pathogen that drove the potato famines of the 1800s and continues to afflict potato fields today. The polysaccharide pectin makes up about a third of the cell wall in potatoes. Sabbadin *et al.* identified a family of lytic polysaccharide monooxygenases (LMPOs) that cleave pectin and are up-regulated in *P. infestans* during infection. Silencing the relevant LMPO gene successfully inhibited *P. infestans* infections. These findings open doors for disease intervention targets and for biotech applications. —PJH

*Science*, abj1342, this issue p. 774

**Antique engraved illustration of *Phytophthora infestans*, which causes potato blight**

CORONAVIRUS

### Defenses against SARS-CoV-2 variants

Our key defense against the COVID-19 pandemic is neutralizing antibodies against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus elicited by natural infection or vaccination. Recent emerging viral variants have raised concern because of their potential to escape antibody neutralization. Wang *et al.* identified four antibodies from early-outbreak convalescent donors that are potent against 23 variants, including variants of concern, and characterized their binding to the spike protein of SARS-CoV-2. Yuan *et al.* examined the impact of emerging mutations in the receptor-binding domain of the spike protein on binding to the host receptor ACE2 and to a range of antibodies. These studies may be helpful for developing more broadly effective vaccines and therapeutic antibodies. —VV

*Science*, abh1766, this issue p. 759, abh1139, this issue p. 818

SUPERCONDUCTIVITY

### Constraining symmetry

Most superconductors have only one transition point and, below a certain temperature, their electrical resistance drops to zero. In very rare cases, another superconducting transition appears at a lower temperature. By measuring its specific heat, Hayes *et al.* reveal that this two-step superconductivity occurs in the compound uranium ditelluride. Complementary optical measurements indicated the breaking of time reversal symmetry, constraining the possible symmetries of the order parameter in this material. —JS

*Science*, abb0272, this issue p. 797

PALEONTOLOGY

### A mammoth's life

Fossils have long given us glimpses of the life that came before us, but these glimpses are generally static. They tell us a bit about species that lived, but not much about how they lived. Evolving techniques are deepening our viewpoint. Wooller *et al.* examined isotopes collected from the tusk of a 17,000-year-old mammoth to elucidate its movements from birth to death. This included its time—likely with a herd—as an infant and juvenile, then as a prime-age adult, and then as a declining senior over its approximately 28-year life span. —SNV

*Science*, abg1134, this issue p. 806

METABOLISM

### A lifetime of change

Measurements of total and basal energy in a large cohort of subjects at ages spanning from before birth to old age document distinct changes that occur during a human lifetime. Pontzer *et al.* report that energy

PHOTO: MIKROMAN6/GETTY IMAGES

expenditure (adjusted for weight) in neonates was like that of adults but increased substantially in the first year of life (see the Perspective by Rhoads and Anderson). It then gradually declined until young individuals reached adult characteristics, which were maintained from age 20 to 60 years. Older individuals showed reduced energy expenditure. Tissue metabolism thus appears not to be constant but rather to undergo transitions at critical junctures. —LBR

*Science*, abe5017, this issue p. 808; see also abl4537, p. 738

BACTERIAL INFECTIONS

## Vaccines and antibiotic resistance

How can we predict or explain changes in antibiotic resistance after the introduction of new vaccines? Davies *et al.* show that the association between penicillin consumption and penicillin resistance in *Streptococcus pneumoniae* across 27 European countries can be explained by four different mathematical models of antibiotic resistance evolution. Each model encapsulates an alternative hypothesis for why antibiotic-sensitive and antibiotic-resistant bacterial strains coexist in the same population. Depending upon the model, vaccination can either inhibit or promote the spread of antibiotic resistance. —OMS

*Sci. Transl. Med.* **13**, eaaz8690 (2021).

PHYSIOLOGY

## Denser lymphatics without leakage

The lymphatic system drains excess fluid from tissues and enables immune cell migration. Exogenous administration of the lymphangiogenic growth factor VEGF-C increases the density, but also the leakiness, of lymphatics and results in off-target inflammation. Kataru *et al.* generated mice with a lymphatic endothelial cell–specific deficiency of PTEN, an enzyme that inhibits VEGF-C signaling (see the Focus by Künnapuu and Jeltsch). These mice had an expanded, non-leaky lymphatic vessel network and improved resolution of inflammation compared with VEGF-C–injected mice. —WW

*Sci. Signal.* **14**, eabc0836, eabj5058 (2021).

ULTRACOLD MOLECULES

## Shielding ultracold molecules

Ultracold molecules hold promise for a wide range of exciting applications. However, such applications are currently hampered by the limited number of ultracold molecular ensembles that can be created and by their short lifetimes. Anderegg *et al.* used a microwave dressing field to tune the collisional properties of calcium monofluoride molecules trapped in optical tweezers. This approach allowed a sixfold suppression of inelastic trap-loss collisions. This scheme paves the way for the creation of a variety of long-lived ultracold molecular ensembles. —YS

*Science*, abg9502, this issue p. 779

BLACK HOLES

## Variability time scales in active galaxies

Active galactic nuclei contain a supermassive black hole (SMBH) surrounded by an accretion disk. As disk material falls toward the SMBH, it heats up enough to emit optical light. Burke *et al.* investigated how such optical emission varies over time in a sample of 67 active galaxies (see the Perspective by Lira and Arevalo). They observed a characteristic variability in timing that scaled with the SMBH mass. The results elucidate the physical processes within accretion disks and provide a method to estimate SMBH mass from optical variability observations. —KTS

*Science*, abg9933, this issue p. 789; see also abk3451, p. 734

# IN OTHER JOURNALS

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

AGING

## Castration delays aging

As we age, our genetic material changes, not only through DNA mutation but also by epigenetic modification. Indeed, chronological age can be estimated based on analysis of DNA methylation. Male and female mammals display different average life spans, and a role for sex hormones is expected in this effect. Sugrue *et al.* established an epigenetic clock in sheep by examining methylated DNA in samples from blood and ears. They show that castration extends an animal's life span and feminizes the epigenome at specific androgen-regulated loci during aging. —BAP

*eLife* **10**, e64932 (2021).

**Castration prolongs the life of sheep by feminizing the epigenome to reduce androgen-regulated aging.**

CANCER

## Brakes off cyclin drives memory

Cyclin-dependent kinases 4 and 6 (CDK4/6) are enzymes that stimulate cell proliferation. CDK4/6 inhibitor (CDK4/6i) drugs block the signals that instruct tumor cells to divide and have been approved to treat a subset of breast cancers. Heckler *et al.* and Lelliott *et al.* found that CDK4/6i can also promote the formation of immune memory to help fight tumors. Short-term treatment of cancer cells with CDK4/6i

PHOTO: ZOONAR GMBH/ALAMY STOCK PHOTO

Growing large amounts of bioenergy crops such as shrub willows (*Salix caprea*, or pussy willow, shown) will require extensive fertilization.

BIOENERGY CROPS

## Where can we find the fertilizer?

One strategy for meeting our future energy needs without fossil carbon is the use of plantations of bioenergy crops such as fast-growing trees or grasses. However, such intensive growth operations are not inherently sustainable and will require fertilization to maintain productivity. Li *et al.* estimated that in some scenarios, the amounts of nitrogen, potassium, and phosphorous fertilizer that would be required to balance their removal by biomass is nearly half of their current use for all agricultural purposes. Fertilizer supplies are already under pressure. New strategies will be required if bioenergy crops are to be sustainable. —MAF

*Environ. Sci. Technol.* 10.1021/acs.est.1c02238 (2021).

pushed activated CD8$^+$ T lymphocytes along a memory cell trajectory, which promoted long-term tumor immunity in mice. These findings may guide the design of human clinical trials using CDK4/6i as a cancer pretreatment to kick-start the immune response before the addition of immunotherapeutic agents. —PNK

*Cancer Discov.* 10.1158/2159-8290.CD-20-1540, 10.1158/2159-8290.CD-20-1554 (2021).

PHYSICS

## Tackling the Hubbard model

The Hubbard model describes the physics of interacting particles on a lattice and is thought to contain elements essential to the superconductivity of the cuprates. Despite its apparent simplicity, solving the Hubbard model remains extremely challenging, especially at finite concentrations of hole dopants and intermediate temperatures. Wietek *et al.* used the minimally entangled typical thermal states (METTS) method to calculate the structural factors and thermodynamic properties for the two-dimensional Hubbard model with strong correlations at a finite doping and for a range of temperatures. The researchers found that a stripe order formed at low temperatures, whereas at higher temperatures, a phase resembling the experimentally observed pseudogap state took over. —JS

*Phys. Rev. X.* **11**, 0031007 (2021).

CERAMICS

## Watching grain boundaries

Grain boundaries play an important role in determining the properties of ceramics and other materials. However, tracking the migration of grain boundaries is difficult. Wei *et al.* used atomic-resolution scanning transmission electron microscopy to trigger and probe grain boundary migration in alumina. The authors found that cooperative shuffling of atoms at the grain boundary ledge is how migration proceeds, which leads to structural transitions. This new method for determining the fine details of these sorts of processes will help us better understand how microstructures develop in a range of materials. —BG

*Nat. Mater.* **20**, 951 (2021).

ECOLOGY

## Nitrogen discourages legumes

Anthropogenic nutrient enrichment of the global environment has been proceeding for many decades. Nitrogen deposition in particular leads to higher ecosystem productivity at the expense of biodiversity. In a global experiment spanning 45 sites, Tognetti *et al.* investigated how the addition of nitrogen affected the leguminous plants in grassland. Legumes under natural conditions compete successfully in low-nutrient conditions because of their ability to fix atmospheric nitrogen using their rhizobial symbionts. However, excess nitrogen deposition places them at a competitive disadvantage, leading to declines in their biomass and abundance. In turn, these declines may have adverse consequences for grassland biodiversity and food webs more generally as nitrogen deposition continues. —AMS

*Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023718118 (2021).

NITROGEN FIXATION

## Some light on diazotrophs

About half of the planet's nitrogen fixation occurs out at sea by prokaryote diazotrophs, yet we still have a poor understanding of which ones do what. Karlusich *et al.* applied a combination of image capture and nitrogenase gene (*nifH*) sequencing on globally gathered data to map diazotrophs sampled at different depths. Unexpected hotspots of diazotroph abundance were discovered in the Mozambique Channel and the south Pacific Ocean. However, in the Arctic Ocean, nitrogen fixers were found among ultrasmall bacterioplankton, whereas in the tropics, larger cyanobacterial symbionts and colony-dwelling species such as *Trichodesmium* and *Richelia* dominated. Although single-cell, free-living, noncyanobacterial diazotrophs were the most abundant overall, *Trichodesmium* dominated by sheer size; however, several species tended to co-occur in assemblies. Whether or how such assemblages interact will have to await further study. —CA

*Nat. Commun.* **12**, 4160 (2021).

PHOTO: ARTERRA PICTURE LIBRARY/ALAMY STOCK PHOTO

## ALSO IN *SCIENCE* JOURNALS

*Edited by* **Stella Hurtley**

NEUROSCIENCE

### Accessory proteins and nicotinic receptors

Acetylcholine was the first neurotransmitter identified, and nicotinic acetylcholine receptors (nAChRs) were the first neurotransmitter receptors isolated. Recent studies have identified a multitude of molecules and mechanisms that regulate nAChRs in different tissues. In a Review, Matta *et al*. discuss these discoveries and their implications for the cell biology and medicinal pharmacology of nAChRs. Many accessory factors promote the assembly and function of diverse nAChRs. Some factors are small molecules, some are proteins, some control receptor biogenesis, and some regulate channel gating. These protein chaperones and auxiliary subunits elucidate the pharmacological and physiological processes regulated by acetylcholine. —PRS

*Science*, abg6539, this issue p. 757

MICROBIOLOGY

### Spying on microbial communities, cell by cell

Within any community of organisms, gene expression is heterogeneous, which can manifest in genetically identical individuals having a different phenotype. One has to look at individuals in context and analyze patterns in both space and time to see the full picture. Aiming to fill a gap in current methods, Dar *et al*. developed a transcriptome-imaging method named parallel sequential fluorescence in situ hybridization (par-seqFISH). They applied this technique to the opportunistic pathogen *Pseudomonas aeruginosa*, focusing on biofilms where growth conditions can change at microscopic scale. Development of these communities, as revealed by mRNA composition, were followed in space and time. The results revealed a heterogeneous phenotypic landscape, with oxygen availability shaping the metabolism at a spatial scale of microns within a single contiguous biofilm segment. This tool should be applicable to complex microbial communities in the environment and the human microbiome. —MAF

*Science*, abi4882, this issue p. 758

CANCER GENOMICS

### Identifying the origin of cancer

Many cancers are classified on the basis of the organ or tissue from which they originated. However, identifying the specific cells and conditions that precede tumorigenesis can help us understand and better treat the resulting disease. Nowicki-Osuch *et al*. used a single-cell approach to investigate the cell of origin for Barrett's esophagus (BE) and the mechanisms leading to the development of esophageal adenocarcinoma (EAC) (see the Perspective by Geboes and Hoorens). Analyses of healthy human esophageal tissues, mutational lineage tracing, and organoid models revealed that BE originates from the gastric cardia and that EAC arises from undifferentiated BE cells. This analysis provides a map of the transcriptional landscape of the healthy esophagus that can be compared with mouse models of disease. —LMZ

*Science*, abd1449, this issue p. 760; see also abj9797, p. 737

CRISPR BIOLOGY

### Target site selection in CAST systems

Exciting genomic engineering possibilities exist for natural integration systems called transposons, which have co-opted CRISPR/Cas systems. An unexplained feature of these systems involves how they direct insertions in a single orientation at a precise distance from the programmed target sequence. Park *et al*. show that orientation information is communicated to the transposase TnsB using the unidirectional growth of a helical filament made up of an AAA+ protein, TnsC. ATP hydrolysis trims the filament to a minimal unit that is marked by TniQ and defined by the Cas12k protein to provide spacing information. This finding may help future engineering of these systems for therapeutic applications. —DJ

*Science*, abi8976, this issue p. 768

POLYMER CHEMISTRY

### Tough recyclable polyacetals

Cyclic acetals such as dioxolane are appealing building blocks for recyclable plastics but have proven to be difficult to polymerize controllably. Abel *et al*. show that optimal pairing of a bromomethyl ether and indium or zinc Lewis acid produces polydioxolane with high tensile strength that may be advantageous for packaging applications. Heating this plastic in strong acid easily breaks it back down to its acetal monomer, which can then be recovered by distillation from mixed plastic waste streams in high yield. —JSY

*Science*, abh0626, this issue p. 783

PALEOBOTANY

### The timing of land plant origins

Until now, the first fossil evidence of land plants was from the Devonian era 420 million years ago. However, molecular phylogenetic evidence has suggested an earlier origin in the Cambrian. Strother and Foster describe an assemblage of fossil spores from Ordivician deposits in Australia dating to approximately 480 million years ago (see the Perspective by Gensel). These spores are of intermediate morphology between confirmed land plant spores and earlier forms of uncertain relationship. This finding may help to resolve discrepancies between molecular and fossil data for the timing of land plant origins. —AMS

*Science*, abj2927, this issue p. 792; see also abl5297, p. 736

MICROBIOTA

### Gut bugs and systemic disease risk

What people eat has an immediate selective effect on the microbial populations resident in the gut. A high-fat diet is associated with the occurrence of microbes that catabolize choline and the accumulation of trimethylamine *N*-oxide (TMAO) in the bloodstream, a contributing factor for heart disease. Yoo *et al*. explored the microbial organisms and pathways that convert choline into TMAO in mice. Although gene clusters for choline metabolism are found widely among the microbiota, it is only the facultative anaerobes that become abundant in hosts on a high-fat diet. A high-fat diet impairs mitochondrial uptake of oxygen into host enterocytes and elevates nitrate in the mucus, which in turn weakens healthy anaerobic gut function. Facultative anaerobes such as the pathobiont *Escherichia coli* become dominant, which leads to an overall increase in the amount of choline catabolized into the precursor for TMAO. Whether this pathway plays a role in heart disease remains unclear. —CA

*Science*, aba3683, this issue p. 813

2D MATERIALS

### Boridene: a 2D boride

A range of two-dimensional (2D) materials, including graphene and hexagonal boron nitride, have been synthesized and studied because of the unusual properties that occur when one dimension becomes very small. MXenes are a family of materials

made of layers of inorganic transition metal carbides and nitrides that are a few atoms thick and are manufactured by selective etching. Attempts to make similar boridene materials have been challenging because of the reactive nature of boride phases and because the parent materials tend to dissolve rather than selectively etch. Zhou *et al*. synthesized boridene in the form of single-layer 2D molybdenum boride sheets by selective etching in aqueous hydrofluoric acid to produce sheets with ordered metal vacancies, opening up an additional family of materials for study. —MSL

*Science*, abf6239, this issue p. 801

T CELLS

## A virtual memory T cell spectrum

Virtual memory T (TVM) cells acquire a memory phenotype in the absence of foreign antigen and are believed to develop in response to self-antigen exposure. However, their role in protective immunity against foreign pathogens is not well understood. Using specific pathogen–free mice infected with influenza A virus, Hou *et al.* demonstrate that TVM cells rapidly infiltrate the lungs in a CXCR3-dependent manner, where they expand and promote early viral control. Compared with naïve T cells, TVM cells more efficiently gave rise to resident memory cells, with the $CCR2^{+}$ and $CCR2^{-}$ subsets poised for effector and memory responses, respectively. Thus, TVM cells undergo functional specialization, and self-reactive T cells can productively contribute to antigen-specific responses against invading pathogens. —CO

*Sci. Immunol.* **6**, eabg9433 (2021).

NANOOPTICS

## Plasmonic vortex cavities multiply OAM

Orbital angular momentum (OAM) of light can be confined to metal surfaces in the form of plasmonic vortices, where nanoscale twisted light oscillates together with the free electrons in metal. Spektor *et al*. introduced the reflection from structural boundaries as a new degree of freedom to generate and control this surface-confined angular momentum. They designed vortex cavities with chiral boundaries, in which subsequent reflections increased the vortex OAM by multiples of the chiral cavity order. They then tracked the spatiotemporal dynamics of the plasmon pulse trains within the vortex cavities using time-resolved photoemission electron microscopy with sub-femtosecond resolution. These results may have applications in quantum initialization schemes and potentially achieve vortex lattice cavities with dynamically evolving topology. —LNL

*Sci.Adv*. 10.1126/sciadv.abg5571 (2021).

REVIEW SUMMARY

NEUROSCIENCE

# Nicotinic acetylcholine receptor redux: Discovery of accessories opens therapeutic vistas

Jose A. Matta†, Shenyan Gu†, Weston B. Davini, David S. Bredt*

**BACKGROUND:** One hundred years ago, experiments on beating frog hearts identified acetylcholine (ACh) as the seminal neurotransmitter. Sixty years later, fractionation of the eel electroplax isolated nicotinic ACh receptors (nAChRs) as the first purified ion channel. We now appreciate that a family of nAChRs are differentially expressed in numerous tissues, including the brain, skeletal muscle, white blood cells, and cochlear hair cells. Paralleling this wide distribution, nAChRs mediate diverse physiological functions, including cognition, muscle contraction, immunomodulation, and sound discrimination. Neuronal nAChRs also account for the psychoactive and addictive properties of tobacco and are the primary genetic risk factors for lung cancer. Therapeutically, nAChRs provide pharmacological targets of approved medicines for cardiovascular and neurological disorders.

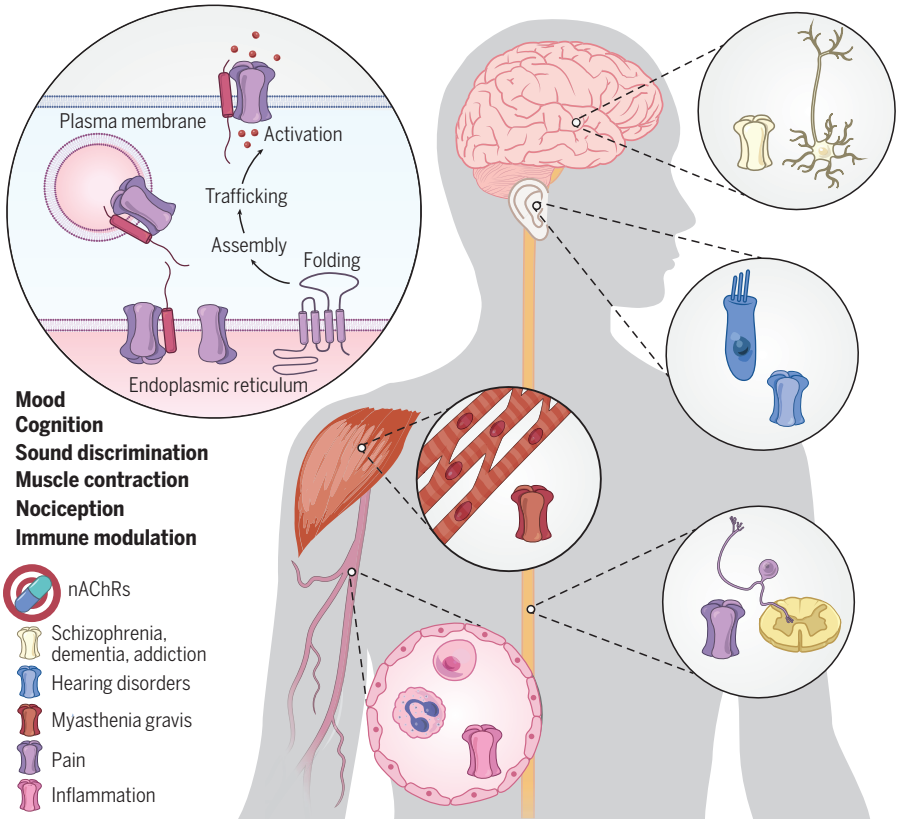
Nicotinic AChRs comprise multiple subunits whose molecular folding and surface trafficking involve complex and tightly regulated processes. As nAChRs often require tissue-specific factors for functional expression, many subtypes fail to create receptor channels in recombinant systems. Our limited understanding of nAChR assembly has impeded basic research and drug development.

**ADVANCES:** Studies in the 1970s found that smokers have increased nAChR density in the brain owing to receptor stabilization by nicotine—a process that likely contributes to tobacco addiction. Recent applications of proteomics, genetics, and expression cloning have identified a bevy of partner proteins and metabolites essential for nAChR function. These accessories act at multiple steps in nAChR biogenesis. Within the endoplasmic reticulum, chaperones mediate nAChR subunit folding and assembly. Other factors then promote nAChR trafficking to the plasma membrane. Finally, auxiliary subunits associated with surface nAChRs modulate channel activation. These chaperones and auxiliary subunits include both nAChR-specific regulators and more pleiotropic factors. On the one hand, NACHO (a neuronal endoplasmic reticulum–resident protein) serves as a client-specific chaperone for neuronal nAChRs. By contrast, transmembrane inner ear protein contributes to both hair cell nAChRs and mechanosensitive channels, which modulate cochlear amplification and transduce sound waves, respectively. Interplay between nAChR accessory components can further regulate receptor distribution and function.

**OUTLOOK:** Discovery of these molecules and mechanisms is transforming basic and translational science concerning nAChRs. Inclusion of appropriate chaperones during protein production is enabling structural studies of nAChR subtypes. Accessory components are also permitting biophysical studies of nAChR channel properties. Furthermore, understanding mechanisms that control trafficking and subunit composition is defining roles for nAChRs in biological processes and disease.

This research also provides therapeutic opportunities. The dearth of pharmacological agents for certain nAChRs results from challenges in recombinant expression of many receptor types. The ability to express complex nAChR subunit combinations in cell lines "unlocks" them for the chemical screening that initiates drug discovery. Auxiliary subunits can themselves provide pharmacological targets. Furthermore, drugging chaperone pathways may benefit myasthenia gravis and other diseases associated with aberrant nAChR levels.

Despite being the archetypal neurotransmitter receptor, much remains unknown about nAChRs. The identification of molecular partners and elucidation of regulatory mechanisms provide a cell biological renaissance and can suggest avenues for treating diseases associated with nAChR dysfunction. ■




**Accessory proteins assist nAChRs and drug discovery.** Throughout the body, nAChRs are differentially expressed in neurons, myocytes, leukocytes, and cochlear and vestibular hair cells. An array of nAChR chaperones and auxiliary subunits (inset) mediate endoplasmic reticulum folding and assembly, intracellular trafficking, and plasma membrane activation. The recent identification of receptor accessories enables drug discovery for these nAChRs, which provide compelling targets for neurological, psychiatric, immunological, and auditory disorders.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: dbredt@gmail.com
†These authors contributed equally to this work.
Cite this article as J. A. Matta *et al.*, *Science* 373, eabg6539 (2021). DOI: 10.1126/science.abg6539

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abg6539

REVIEW

NEUROSCIENCE

# Nicotinic acetylcholine receptor redux: Discovery of accessories opens therapeutic vistas

Jose A. Matta†, Shenyan Gu†, Weston B. Davini, David S. Bredt‡*

The neurotransmitter acetylcholine (ACh) acts in part through a family of nicotinic ACh receptors (nAChRs), which mediate diverse physiological processes including muscle contraction, neurotransmission, and sensory transduction. Pharmacologically, nAChRs are responsible for tobacco addiction and are targeted by medicines for hypertension and dementia. Nicotinic AChRs were the first ion channels to be isolated. Recent studies have identified molecules that control nAChR biogenesis, trafficking, and function. These nAChR accessories include protein and chemical chaperones as well as auxiliary subunits. Whereas some factors act on many nAChRs, others are receptor specific. Discovery of these regulatory mechanisms is transforming nAChR research in cells and tissues ranging from central neurons to spinal ganglia to cochlear hair cells. Nicotinic AChR–specific accessories also enable drug discovery on high-confidence targets for psychiatric, neurological, and auditory disorders.

More than 100 years ago, Loewi and Dale identified acetylcholine (ACh) as the first known neurotransmitter (*1*). We now appreciate that ACh mediates diverse physiological functions by acting on a large family of receptors. Whereas Loewi's experiments on frog hearts involved a heterotrimeric GTP-binding protein (G protein)–coupled muscarinic ACh receptor (*2*), nicotinic ACh receptors (nAChRs) are ligand-gated ion channels (*3*, *4*). These nAChRs are enriched in skeletal muscle, where they transduce nerve-muscle communication, and in the brain, where they mediate synaptic transmission (*5*). Nicotinic AChRs also occur in more-specialized locations, such as beige adipocytes (*6*), lymphocytes (*7*), and cochlear hair cells (*8*).

This widespread distribution of nAChRs creates physiological roles for these receptors in neuronal, sensory, metabolic, and immune tissues. Additionally, nAChRs can mediate pathophysiological processes. The addictive properties of nicotine involve nAChRs in brain circuits that mediate craving and reward (*9*). Mutations in nAChR subunits cause several genetic diseases, including epileptic (*10*) and neuromuscular (*11*) disorders. Polymorphisms in nAChRs are linked to lung cancer owing to increased preference for tobacco (*12*). Furthermore, numerous approved and experimental medicines for cardiovascular, psychiatric, and cognitive disorders target nAChRs (*13*, *14*).

Just as ACh was the seminal neuronal messenger, the nAChR was the first biochemically isolated neurotransmitter receptor. The receptor was purified from fish electric organs by affinity chromatography using the potent antagonist α-bungarotoxin (*5*). In common with all nAChRs, this receptor comprises a pentamer of subunits that each contain an N-terminal extracellular domain for ligand binding followed by four transmembrane (TM) regions. A large cytosolic region between TM3 and TM4 includes two structured helices MX and MA (*15*) and mediates interactions with cytoskeletal anchoring proteins. Neuronal nAChR derives from nine α (α2 to α10) and three β (β2 to β4) subunits. The most abundant nAChRs in the brain are α7 homopentamers and α4β2 heteropentamers, although other subunits can be included in these pentamers. Dozens of receptor combinations have been pharmacologically characterized.

Cellular assembly of nAChRs from their subunits involves complex, highly regulated processes (*16*). Essential factors for nAChR assembly were initially found by forward genetic screens of invertebrates. This is exemplified by *ric-3*, which is required for nAChR function in *Caenorhabditis elegans* (*17*). Accessories for mammalian nAChRs were found through analyses of developmentally regulated genes that encode α-bungarotoxin–like neuropeptides, such as lynx1, that modulate certain neuronal nAChRs (*18*). More recently, genome-wide cDNA screening (Fig. 1) has identified a trove of molecules and mechanisms that enable the assembly and function of diverse nAChRs (*19*). Here, we review the remarkable diversity of processes and pathways that regulate the assembly of nAChRs and focus on their relevance in the development of therapies for neurological, psychiatric, and sensory disorders.

## Skeletal muscle nAChR

Excitation and contraction of skeletal muscle are initiated by the motor neuron release of ACh. The embryonic muscle AChR contains $\alpha_2\beta\gamma\delta$ subunits, and this composition switches by birth to the adult form containing $\alpha_2\beta\varepsilon\delta$ (*3*, *4*). The assembled receptor concentrates at crests of junctional folds in association with a cytoskeletal anchoring protein rapsyn. Motor neuron–derived agrin induces nAChR clustering during synaptogenesis by binding LRP4 and signaling via MuSK (*20*). All nine of these genes have human mutations that cause inherited myasthenic syndromes, which are characterized by fatigable muscle weakness owing to aberrant neuromuscular nAChR transmission (*11*). Furthermore, in myasthenia gravis, autoantibodies that react with proteins at the neuromuscular junction mediate inflammation and depletion of nAChRs (*21*).

Symptomatic treatment for myasthenias associated with reduced nAChR levels involves blocking acetylcholinesterase to increase synaptic levels of ACh (*21*). Cholinesterase inhibitors provide modest benefit, but better therapies are needed. Replacing the mutated gene in congenital myasthenias may be possible but requires complex gene therapies tailored for specific patients (*22*). Interrupting the autoimmune process in myasthenia gravis can also provide a cure, but this requires lifelong immunosuppression (*21*).

Enhancing nAChR assembly and surface expression provides another conceptual strategy for treating certain myasthenias. The sequential steps in subunit oligomerization and conformational maturation that form muscle nAChR pentamers have been characterized (*16*). Forward genetic screening identified CRELD1, a protein disulfide isomerase, as an evolutionarily conserved maturational enhancer of muscle nAChRs (*23*). Whereas CRELD1 is widely expressed, muscle-specific targets would be more attractive. Drugs that enhance MuSK signaling to augment synaptic nAChR clusters are being explored for the treatment of neuromuscular disorders (*24*).

## NACHO as a client-specific chaperone for neuronal nAChRs

Whereas muscle nAChRs form active channels when transfected into heterologous cell lines, many neuronal nAChRs do not (*25*). The discovery of NACHO by genome-wide expression cloning (Fig. 1A) resolved why the α7 nAChR fails to form functional receptors in non-neuronal cells (*26*). NACHO is a neuronal endoplasmic reticulum (ER)–resident protein that mediates α7 assembly (*26*). NACHO coexpression during protein production recently facilitated the determination of the elusive α7 nAChR structure (*27*). NACHO also promotes the biogenesis and function of the broadly expressed α4β2 receptors and the more localized α6β2β3 receptors (*28*). NACHO knockout (KO) mice show hyperlocomotive activity and impaired cognitive function (*28*), consistent with altered

Neuroscience Discovery, Janssen Pharmaceutical Companies of Johnson & Johnson, San Diego, CA 92121, USA.
*Corresponding author. Email: dbredt@gmail.com
†These authors contributed equally to this work.
‡Present address: MPM Capital, Brisbane, CA 94005, USA.

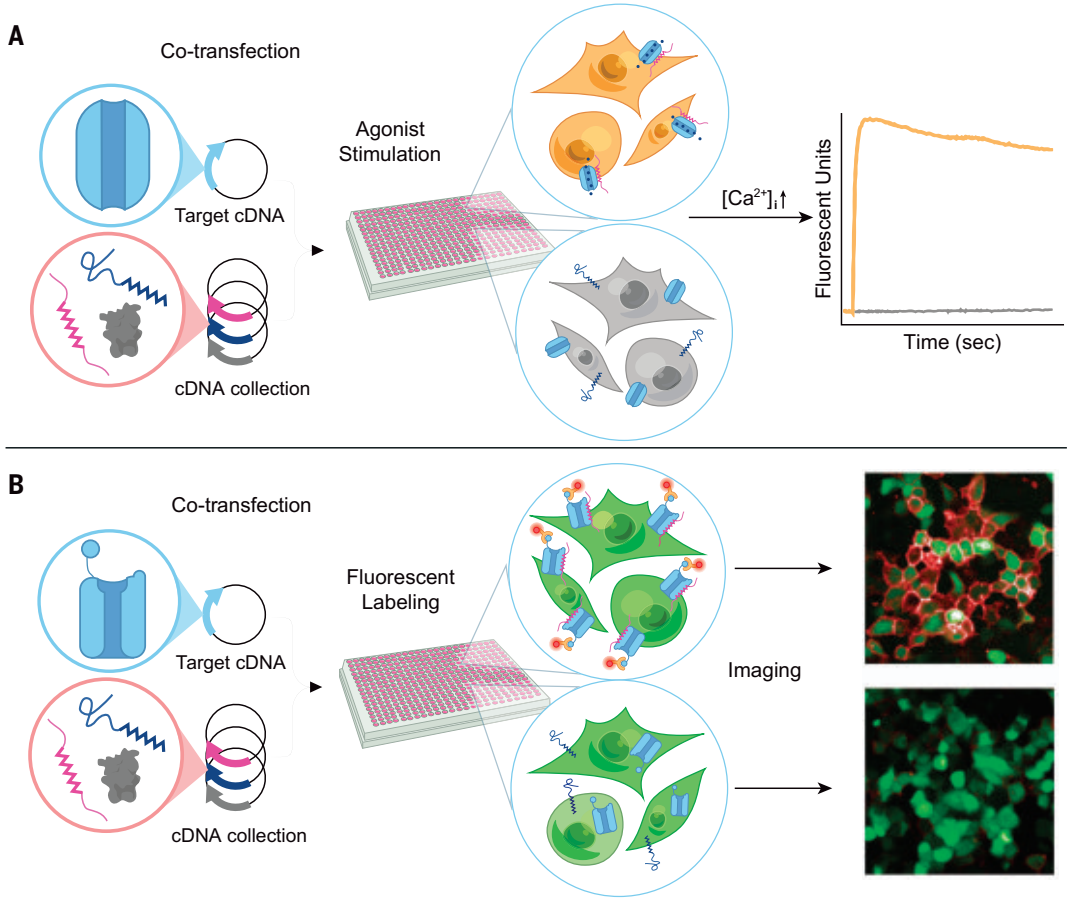



**Fig. 1. cDNA screening platform for discovery of receptor accessories.** Plasmid-encoding receptor of interest (blue) is cotransfected with individual clones from a genome-wide expression library. (**A**) Functional assay can identify proteins that enable receptor activity. (**B**) High-content imaging can identify proteins that promote receptor trafficking.

nAChR expression. By contrast, NACHO KO mice show normal activity for all other receptors analyzed, including γ-aminobutyric acid type A ($GABA_A$), 5-HT3A, AMPA, and TRPV1 (*28*). Therefore, NACHO appears to act as a "client-specific" chaperone for nAChRs.

NACHO enables receptor assembly through actions on the second TM domain of α7 (*29*). Unexpectedly, NACHO does not directly interact with α7 (*26*, *27*). Instead, NACHO binds to components of the N-glycan oligosaccharyltransferase complex and to calnexin (*29*). This fits with studies showing that α7 receptor function requires glycosylation at three asparagine residues in its N-terminal ectodomain (*30*) and that NACHO effects on α7 also require these asparagines (*29*). Furthermore, the ER chaperone calnexin is required for assembly of numerous nAChR subtypes (*31*). These results yield a model whereby NACHO links the oligosaccharyltransferase and calnexin (Fig. 2). In turn, calnexin mediates assembly by interacting reversibly with N-linked glycans on α7 and by recruiting downstream chaperones (*29*).

After NACHO-mediated assembly of α7, the mammalian homolog of RIC-3 (*32*) promotes receptor surface trafficking (*26*). Certain antiapoptotic Bcl-2 proteins can also act after NACHO to enhance α7 surface expression and channel function (*33*). These actions of RIC-3 and Bcl-2 proteins on α7 synergize with NACHO and involve distinct mechanisms (Fig. 2). Effects of NACHO require the extracellular and first two TM domains of α7 (*29*), RIC-3 engages the amphipathic MA helix in the α7 cytosolic loop (*34*), and Bcl-2 necessitates a BH3-like motif in this same MA helix (*33*).

### Regulation of nAChR assembly by drugs and metabolites

Small molecules also regulate nAChR assembly. These compounds can be separated into three groups on the basis of their mechanisms of action. The first group comprises nAChR orthosteric ligands that engage the ACh binding pocket at subunit interfaces and stabilize mature nAChRs (*35*). The second group includes menthol (*36*) and polyamines (*37*)—compounds that directly bind to nAChR at nonorthosteric sites and thereby modulate receptor assembly. The third group contains molecules that work indirectly on AChRs as more-general mediators of protein folding, including 4-phenylbutyrate, valproate, and butyrate, which are also histone deacetylase inhibitors (*38*).

Human studies in the 1980s first noted that nAChR levels in the brain are elevated in cigarette smokers (*39*). Follow-up experiments found that nicotine enhances nAChR function in animals and that nicotine posttranslationally promotes nAChR levels (*40*). This is the opposite of what is typical for G protein–coupled receptors, such as opiate receptors, which are down-regulated by chronic agonist exposure. Mechanistically, nicotine interaction with the ligand-binding site promotes nAChR pentamer formation and receptor surface expression. These effects are independent of receptor activation and are also observed with orthosteric antagonists (*38*).

In the brain, nicotine primarily enhances β2-containing receptors (*41*). As the reinforcing properties of nicotine are mediated by mesolimbic β2-containing receptors, nAChR up-regulation likely participates in nicotine addiction (*9*). Menthol also up-regulates brain nAChRs (*42*). Menthol allosterically inhibits α4β2 nAChRs (*43*) by binding both to an extracellular channel pore site and to a membrane-embedded site in the desensitized receptor (*44*). In transgenic mice, menthol increases assembly of α4- and α6-containing nAChR in dopaminergic neurons (*36*). This up-regulation of nAChRs in reward circuits may explain why smokers of menthol cigarettes have more difficulty quitting smoking than nonmenthol smokers (*45*).

Beyond these pharmacological effects, ligand-mediated assembly may also play an endogenous

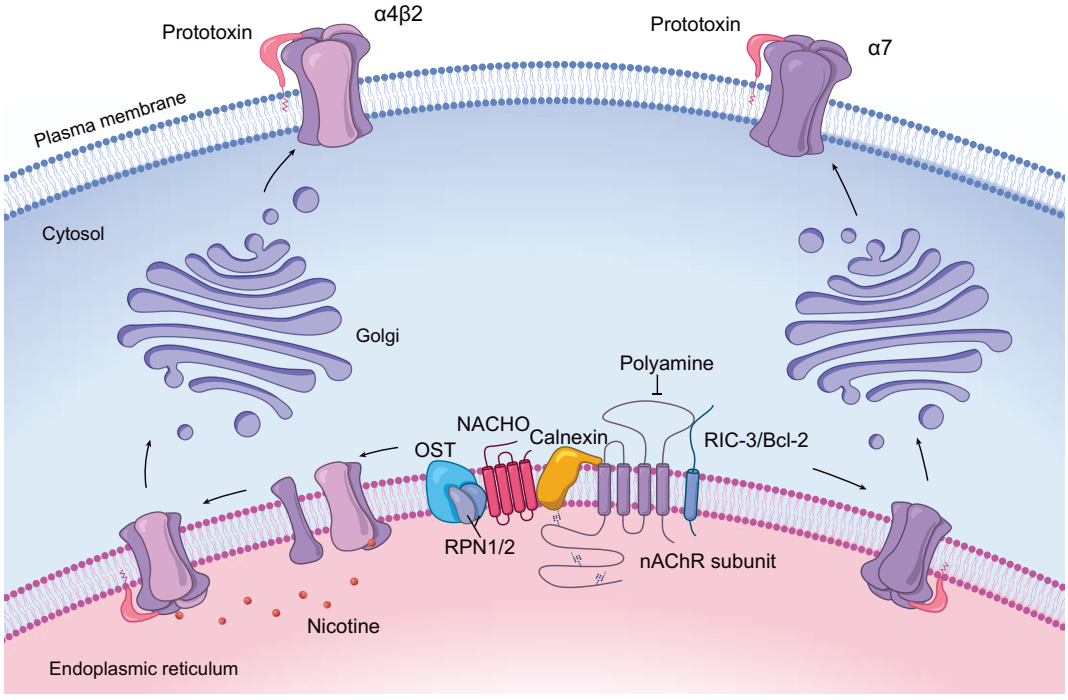



**Fig. 2. Regulation of α7 and α4β2 assembly, trafficking, and activation by protein chaperones, small molecules, and accessory components.** Within the ER, NACHO conspires with calnexin and the oligosaccharyltransferase (OST) proteins RPN1/2 to promote the folding of α7 and α4β2. In synergy with NACHO, RIC-3 and Bcl-2 proteins enhance α7 assembly and trafficking to the plasma membrane. By engaging the α4β2 ligand-binding site, nicotine enhances assembly of α4β2 receptors. Prototoxins associate with α7 and α4β2 pentamers and modulate channel trafficking and activation.

role in the maturation of certain nAChRs. Genome-wide screening for factors that enable the function of α9α10 nAChRs identified choline acetyltransferase (ChAT), the biosynthetic enzyme for ACh (*46*). ChAT-mediated assembly of α9α10 reflects chemical chaperoning by ACh. Mutating the ligand-binding site on α9 abolishes receptor maturation and surface trafficking (*46*). The fact that ACh does not efficiently penetrate cell membranes suggests that its actions to enhance α9α10 receptor levels can occur through extracellular interactions that promote the assembly of incomplete oligomers or stabilize transient surface pentamers. Indeed, even extracellular application of α-bungarotoxin, which is also not cell permeant, augments α9α10 levels (*46*). This extracellular ACh-mediated assembly mechanism may help position functional α9α10 receptors at sites of ACh release (*46*).

Expression cloning (Fig. 1B) for cDNAs that enhance α4β2 surface trafficking (*37*) identified spermidine/spermine acyltransferase (SAT1). SAT1 is the rate-limiting enzyme for polyamine catabolism, which suggests that these ubiquitous linear polycations inhibit α4β2 expression. Indeed, blocking polyamine synthesis with difluoromethylornithine (DFMO) increases levels of functional α4β2 and α7 receptors (*37*). Polyamines typically act by occluding the ion channel pore (*47*). By contrast, polyamine regulation of nAChR assembly relies on interaction with the large cytosolic region (*37*). Polymorphisms in SAT1 have consistently been associated with suicidal behavior (*48*). These findings provide additional validation to assess nicotinic agents as a treatment for suicidal ideation (*49*).

**nAChR regulation by accessory proteins**

Protein chaperones assist receptor assembly and trafficking; by contrast, auxiliary subunits associate stably with receptors and control channel function. An auxiliary subunit for a nAChR was reported in *C. elegans* (*50*). Forward genetic screening identified MOLO-1, a single-pass TM protein that positively modulates worm nAChRs. Notably, MOLO-1 physically interacts with levamisole-sensitive nAChRs and enhances channel activation (*50*). Functional activity of the EAT-2 nAChR in *C. elegans* requires the single-pass TM protein EAT-1, which serves as an essential auxiliary subunit (*51*).

Several mammalian nAChRs also have auxiliary subunits. Nicotinic AChRs containing an α6 subunit are abundant in the striatum and mediate acetylcholine-induced dopamine release, but they do not express functionally in any recombinant system (*52*). α6 was originally dubbed an orphan subunit until the coexpression of chicken α6 and human β4 subunits produced ACh-gated currents in *Xenopus* oocytes (*53*). In the striatum, α6 assembles with β2 and β3 subunits to form a conotoxin MII–sensitive presynaptic receptor (*52*, *54*). Insights regarding α6β2β3-containing nAChR assembly came from studies of NACHO KO mice, which show reduced striatal binding sites for [$^{125}$I]conotoxin MII (*28*). Furthermore, in human embryonic kidney (HEK) cells, NACHO cotransfection promotes oligomeric assembly but not functional expression of α6β2β3 receptors (*28*).

Genome-wide screening to search for proteins that conspire with NACHO to reconstitute α6β2β3 channel function identified β-anchoring and -regulatory protein (BARP), lysosomal associated membrane protein 5 (LAMP5), and sulfotransferase SULT2B1 (*55*). These proteins exert differential effects on α6β2β3 (*55*). Whereas NACHO promotes α6β2β3 assembly, LAMP5 and SULT2B1 increase cell surface receptor trafficking, and BARP enhances channel activation (Fig. 3A).

BARP binds to the β subunit of voltage-gated calcium channels and reduces channel function (*56*). By contrast, BARP enhances α6β2β3 activity by slowing channel desensitization and deactivation (*55*). BARP actions on nAChR show marked selectivity. It enhances the activation of α3β2 receptors but does not affect α4β2 or α3β4 receptors (*55*). Consistent with BARP's regulation of α6β2β3-containing nicotinic receptors, striatal synaptosomes from BARP KO mice show reduced nicotine-evoked dopamine release and reduced [$^{125}$I]conotoxin MII–binding sites (*55*). The nAChR- and calcium channel–binding domains in BARP do not overlap, which suggests that BARP may physiologically link these ion channel complexes. Indeed, nicotine-evoked dopamine release involves voltage-gated calcium channel activity downstream of nAChR activation (*57*).

Another pharmacologically important α6-containing receptor is a diheteromer containing the β4 subunit that occurs in the sensory neurons of dorsal root ganglia (*58*). BARP and IRE1α are needed to reconstitute α6β4 receptors, which are curiously not affected by NACHO (*59*). IRE1α is a sensor of the unfolded protein response, which promotes protein folding in the ER during cellular stress (*60*). Effects on α6β4 involve the canonical IRE1α

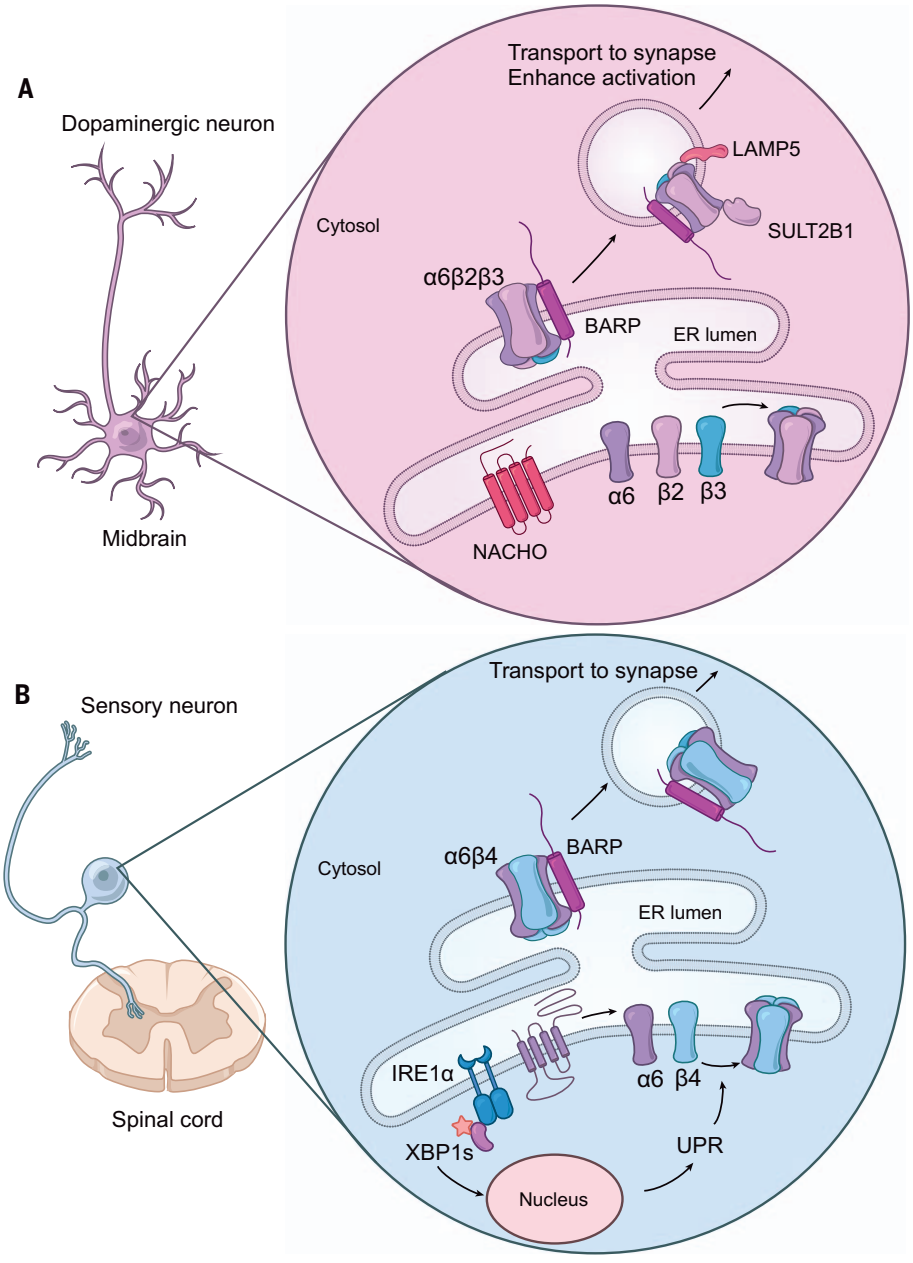



**Fig. 3. Differential regulation of α6-containing nAChRs in central dopaminergic neurons and dorsal root ganglion sensory neurons.** (**A**) For α6β2β3 receptors in dopaminergic neurons, NACHO promotes assembly, LAMP5 and SULT2B1 enhance surface trafficking, and BARP controls channel activation. (**B**) By contrast, assembly of α6β4 receptors in sensory neurons involves IRE1α, XBP1 splicing, and the unfolded protein response. BARP enhances α6β4 surface expression but not α6β4 channel activation. UPR, unfolded protein response.

pathway of autophosphorylation and splicing to form XBP1s (Fig. 3B). Through these actions, IRE1α promotes the assembly of α6β4 receptors, which fits with previous studies linking the unfolded protein response to nAChR upregulation (*61*). Once assembled, α6β4 receptors bind to BARP, which promotes their surface expression (Fig. 3B) but has no effect on receptor desensitization or deactivation properties. As discussed below, the discovery that BARP and IRE1α can reconstitute human α6β4 receptor activity has enabled pharmacological reassessment of experimental analgesics and nominates α6β4 as an attractive target for treating neuropathic pain.

The most discretely expressed nAChR comprises α9 and α10 subunits. Among our senses, only the auditory system receives central nervous system (CNS) efferent innervation, which terminates on the hair cell α9α10 nAChR (*62*). This pathway provides inhibitory feedback to enhance sound discrimination and to protect cochlear hair cells from acoustic trauma (*8*, *63*). The α9α10 nAChR is among the most–calcium-selective ion channels known, and its activation links tightly to SK2, a calcium-sensitive potassium channel that hyperpolarizes the hair cell to suppress electromotility (*64*). Endogenous α9α10 ionotropic activity has only been recorded in cochlear and vestibular hair cells, and the receptor does not function in transfected cell lines.

Genome-wide expression cloning identified factors that reconstitute functional α9α10 receptors (*46*). These experiments found that the ACh biosynthetic enzyme ChAT promotes α9α10 assembly and pointed to an essential role for ligand binding in α9α10 biogenesis. Additionally, α9α10 channel function requires an auxiliary subunit (Fig. 4). Notably, the auxiliary subunits that best enable α9α10 function are TM inner ear (TMIE) and TMEM132e, which are both recessively mutated in nonsyndromic deafnesses (*65*, *66*). Whereas TMIE and TMEM132e lack sequence homology, both are single-pass TM proteins that localize in part to the basal membrane of cochlear hair cells, where α9α10 receptors reside (*66*, *67*). In addition to regulating α9α10 receptors, both TMIE and TMEM132e are required for activity of the sound-sensing mechanotransduction (MET) channel, which suggests that TMIE and TMEM132e play multiple roles in hair cells. *Spinner* mice lacking functional TMIE are deaf (*68*) and evince abnormal inner hair cell α9α10 responses (*46*).

Another class of proteins that associate stably with nAChRs are "prototoxins"—members of the ly6/uPAR superfamily—which contain receptor-binding motifs homologous to α-neurotoxins from snake venom (*18*). A glycosylphosphatidylinositol (GPI) anchor tethers many of these three-fingered α-bungarotoxin–like proteins to the plasma membrane, allowing them to associate with nAChRs (Fig. 2). The lynx1 prototoxin inhibits α4β2 and α7 nAChRs. These effects on receptor function are multifaceted and involve both reduction of ACh affinity and acceleration of channel desensitization (*69*). Furthermore, lynx1 influences α4β2 receptor assembly, favoring the low-sensitivity (α4)3(β2)2 versus the high-sensitivity (α4)2(β2)3 stoichiometry (*70*). The related protein lynx2 also inhibits function of α4β2 and α7 receptors (*71*). Prototoxins' effects may occur intracellularly, where they influence receptor assembly (*70*) and trafficking (*72*), or at the plasma membrane, where they influence receptor function (*69*, *73*). Finally, SLURPs (secreted Ly-6/uPAR–related proteins) are nonanchored prototoxins such as SLURP1, which is secreted by keratinocytes to modify cholinergic tone and influence epidermal healing (*74*).

The large number of prototoxins, their differential association with nAChR subtypes, and their diverse functional effects provide numerous opportunities for therapeutic intervention. Lynx1 KO mice show improved associative learning and memory, consistent with

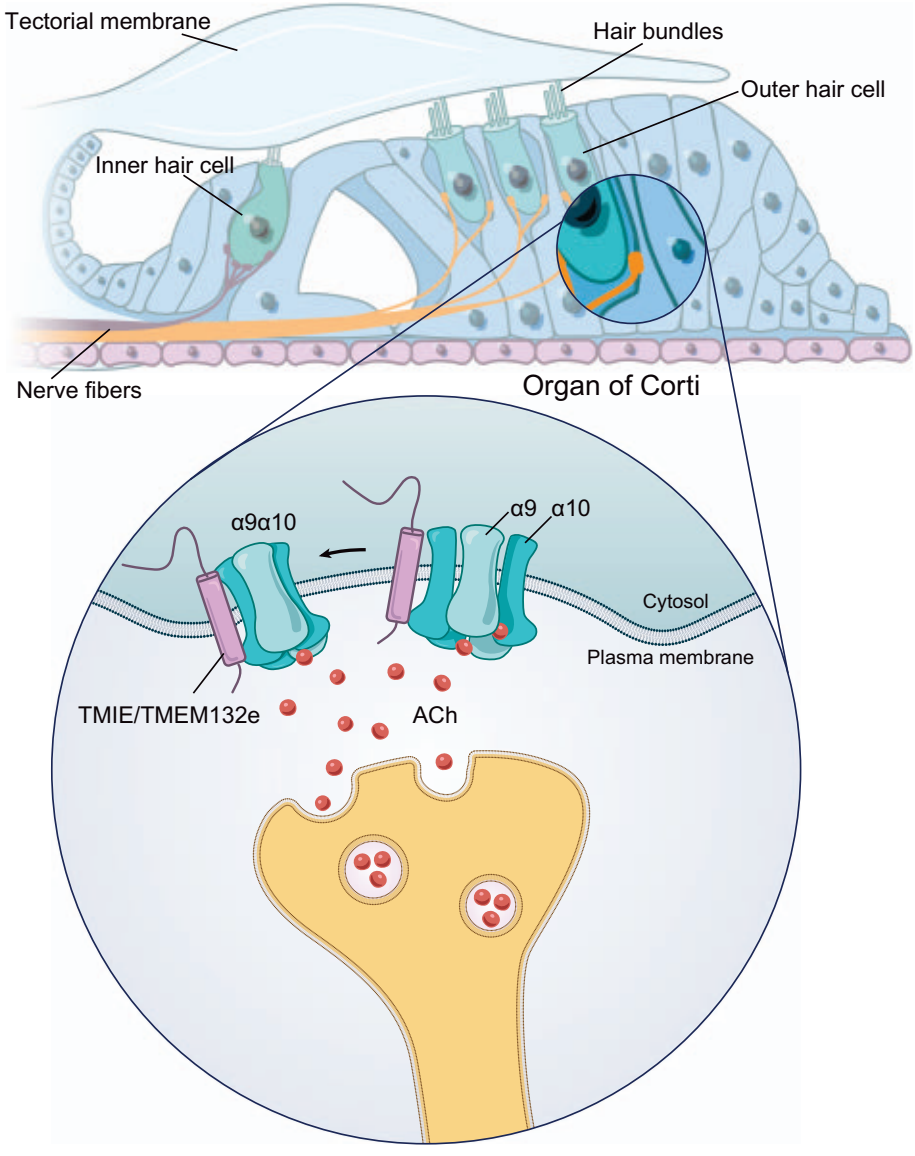



**Fig. 4. α9α10 receptor assembly in cochlear hair cells.** Stable expression of surface α9α10 pentamers requires ligand binding, which can be provided by ACh released from presynaptic nerve terminals. Channel function of α9α10 receptors additionally requires an auxiliary subunit, which can be TMIE or TMEM132e.

enhanced nicotinic cholinergic tone (*75*). These mice also display an enhanced antinociceptive response to nicotine (*76*). By contrast, lynx2 KO mice show elevated anxiety-like behavior (*71*). These data suggest that modulating prototoxins' actions may benefit diverse neuropsychiatric disorders. How prototoxins functionally interact with the recently discovered nAChR receptor accessories, such as NACHO and 14-3-3, which can also influence subunit stoichiometry (*77*), represents another future direction.

### Implications of accessory components on nAChR neurobiology

Clearly a diverse collection of molecular partners and pathways regulate nAChRs. Some factors are small molecules, some are proteins, some control receptor biogenesis, and some regulate channel activation (Table 1). Perhaps most notable is the nAChR selectivity of these accessory components. NACHO and BARP regulate numerous nAChRs. By contrast, TMIE, TMEM132e, and IRE1α affect only single nAChR types.

Why might nAChRs be regulated by such a wide array of selective modulatory factors? This complexity likely reflects the need to control nAChRs in specific tissue types, at specific cellular locations, and during specific physiological conditions. Whereas neuronal α7 nAChR assembly requires NACHO, other proteins may chaperone α7 biogenesis in non-neuronal cells. Functional α7 occurs in immune cells (*78*), which lack NACHO (*79*) but express Bcl-2 proteins (*80*). RIC-3 occurs in several non-neuronal tissues, including the skin and endothelial cells that express α7 but not NACHO (*81*). Bcl-2 proteins (*82*) and α7 (*83*) exert neuroprotective properties and may mitigate neuronal injury associated with brain trauma or neurodegenerative disorders. Accordingly, adaptive mechanisms may enhance cholinergic signaling through α7 downstream of Bcl-2.

Up-regulation of nAChRs by orthosteric ligands has important pharmacological and physiological implications. On the one hand, nicotine-mediated up-regulation of β2-containing nAChRs in the brain likely contributes to tobacco addiction. On the other hand, ACh-mediated receptor up-regulation can concentrate nAChRs at sites of ACh release. Postsynaptic α9α10 function requires hair cell cholinergic innervation. Developmental studies show that functional α9α10 receptors occur transiently in inner hair cells before the onset of hearing (*84*, *85*); by contrast, outer hair cell α9α10 receptor function appears only after hearing onset (*63*). These dynamics of postsynaptic α9α10 function closely match the timing of hair cell cholinergic innervation (*85*). Furthermore, in age-related hearing loss, mature inner hair cells regain both cholinergic input and functional α9α10 receptors (*86*). The discovery that α9α10 receptor assembly is enhanced by extracellular ACh provides a mechanism to link postsynaptic receptor assembly to sites of presynaptic cholinergic innervation (*46*). As ligand binding enhances levels of α4β2 and other nAChR types, this mechanism may play a more general role in coordinating development of cholinergic synapses.

### Implications of accessory components for nAChR neuropharmacology

Functional expression of previously elusive nAChRs now enables drug discovery for unmet medical needs ranging from chronic pain to Parkinson's disease to hearing disorders. Epibatidine, an alkaloid from frog skin, is a general nAChR agonist and a powerful analgesic (*87*). Furthermore, a pan-nAChR agonist ABT-594 showed robust efficacy for diabetic neuropathy in a phase 2 clinical trial. Unfortunately, ABT-594 had unacceptable adverse effects (*88*), and its therapeutic nAChR target was unknown. Subsequently, a genomics screen of dorsal root ganglion tissue from outbred mouse strains identified that α6 nAChR subunit mRNA levels inversely correlate with pain responses in a spared nerve injury model (*58*). Accordingly, analgesic effects of nicotine after both inflammatory and neuropathic injuries are absent in α6 KO mice. Furthermore, human postoperative pain and temporomandibular disorder are affected by a polymorphism in the *CHRNA6* (α6) promoter (*58*).

This research pointed to α6β4 as an enticing therapeutic target for treating chronic pain (*58*).

**Table 1. Molecular partners regulating nACh receptor biogenesis and function.**

| nAChR subtype | α7 | α4β2 | α6β2β3 | α3β2 | α6β4 | α9α10 |
|---|---|---|---|---|---|---|
| Primary tissue | CNS | CNS | CNS | Autonomic ganglion | Sensory neuron | Cochlear hair cell |
| Physiological role | Excitatory transmission | Excitatory transmission | Excitatory transmission | Postganglionic transmission | Nociception | Cochlear transmission |
| Protein chaperones | NACHO, Bcl-2, RIC-3 | NACHO | NACHO, SULT2B1 | NACHO | IRE1α, SULT2B1 | None known |
| Chemical chaperones | Nicotine, phenylbutyrate, valproate, polyamines | Nicotine, menthol, butyrate, polyamines | Nicotine, menthol | Nicotine | Nicotine | ACh, polyamines |
| Auxiliary subunits | Prototoxin | Prototoxin | BARP | BARP | BARP | TMIE, TMEM132e |

However, drug discovery on human α6β4 receptors was precluded because these receptors were refractory to expression in recombinant cell lines. Functional reconstitution of human α6β4 by BARP and IRE1α now unlocks this target for drug discovery (*59*). Armed with this methodology, pharmacological reassessment of nicotinic agents used in neuropathic pain trials (*88*, *89*) showed that clinical benefit best correlates with drug efficacy on the α6β4 receptor (*59*). Furthermore, nicotine- and ABT-594–mediated anti-allodynia are blunted in BARP KO mice (*59*). These results provide evidence that agonism on α6β4 can relieve neuropathic pain and is an appealing non-opioid target.

Two other therapeutically attractive nACh receptors are the α6β2β3-containing and α9α10 receptors. The ability to express these channels in mammalian heterologous cells allows for biochemical and biophysical analyses not previously possible. These receptors are also now amenable to the high-throughput screening of chemical libraries for drug discovery. In the brain, α6β2β3-containing receptors are enriched on presynaptic terminals of mesolimbic neurons and mediate acetylcholine- and nicotine-induced dopamine release in the striatum and nucleus accumbens, which makes them attractive targets for treating Parkinson's disease, schizophrenia, and drug addiction.

The α9α10 nAChR is enriched on the outer hair cells of adult cochlea and mediates efferent input from the superior olivary complex. These olivocochlear efferents provide negative feedback to decrease the amplification afforded by the outer hair cells (*8*). This process improves signal discrimination amidst background noise and protects against sound-induced hearing loss. Genetic enhancement of α9α10 receptor function protects the inner ear sensory system from acoustic trauma (*90*). Furthermore, abnormalities in this feedback pathway likely participate in the pathophysiology of tinnitus (*86*). Functional expression of α9α10 recombinant receptors will enable drug discovery that may provide solutions for these common and untreatable auditory disorders.

A few biologically important nAChR combinations, including those that contain the α2 or α5 subunits, remain refractory to expression in recombinant cell lines and lack selective pharmacological agents. Nicotinic AChRs containing α2β2 subunits occur in hippocampal oriens lacunosum-moleculare (OLM) neurons. These α2β2-containing receptors account for nicotine effects on synaptic plasticity and provide an appealing target for cognitive disorders (*91*). Especially interesting is the α5 nAChR subunit, whose variants alter tobacco preference and confer the greatest overall genetic risk for lung cancer (*92*). Notably, of all human single-nucleotide polymorphisms, those within the α5 nAChR subunit (*CHRNA5*) are second only to ApoE4 in determining human life span (*93*). Functionally expressing and pharmacologically targeting nAChRs containing the α5 subunit could provide a precision medicine to treat nicotine addiction. Discovery of accessory components represents an inflection in biological understanding of nAChRs. Achieving the plateau provides a panorama of next-generation opportunities for this prototypical receptor family.

## REFERENCES AND NOTES


1. H. Dale, Pharmacology and Nerve-endings. *Proc. R. Soc. Med.* **28**, 319–332 (1935). doi: 10.1177/003591573502800330; pmid: 19990108
2. E. C. Hulme, N. J. Birdsall, N. J. Buckley, Muscarinic receptor subtypes. *Annu. Rev. Pharmacol. Toxicol.* **30**, 633–673 (1990). doi: 10.1146/annurev.pa.30.040190.003221; pmid: 2188581
3. J. Lindstrom, Nicotinic acetylcholine receptors in health and disease. *Mol. Neurobiol.* **15**, 193–222 (1997). doi: 10.1007/BF02740634; pmid: 9396010
4. N. Le Novère, P. J. Corringer, J. P. Changeux, The diversity of subunit composition in nAChRs: Evolutionary origins, physiologic and pharmacologic consequences. *J. Neurobiol.* **53**, 447–456 (2002). doi: 10.1002/neu.10153; pmid: 12436412
5. J. P. Changeux, Discovery of the First Neurotransmitter Receptor: The Acetylcholine Nicotinic Receptor. *Biomolecules* **10**, 547 (2020). doi: 10.3390/biom10040547; pmid: 32260196
6. H. Jun *et al*., An immune-beige adipocyte communication via nicotinic acetylcholine receptor signaling. *Nat. Med.* **24**, 814–822 (2018). doi: 10.1038/s41591-018-0032-8; pmid: 29785025
7. X. Zhang *et al*., Brain control of humoral immune responses amenable to behavioural modulation. *Nature* **581**, 204–208 (2020). doi: 10.1038/s41586-020-2235-7; pmid: 32405000
8. A. B. Elgoyhen, E. Katz, The efferent medial olivocochlear-hair cell synapse. *J. Physiol. Paris* **106**, 47–56 (2012). doi: 10.1016/j.jphysparis.2011.06.001; pmid: 21762779
9. J. P. Changeux, Nicotine addiction and nicotinic receptors: Lessons from genetically modified mice. *Nat. Rev. Neurosci.* **11**, 389–401 (2010). doi: 10.1038/nrn2849; pmid: 20485364
10. M. De Fusco *et al*., The nicotinic receptor β2 subunit is mutant in nocturnal frontal lobe epilepsy. *Nat. Genet.* **26**, 275–276 (2000). doi: 10.1038/81566; pmid: 11062464
11. A. G. Engel, X. M. Shen, D. Selcen, S. M. Sine, Congenital myasthenic syndromes: Pathogenesis, diagnosis, and treatment. *Lancet Neurol.* **14**, 420–434 (2015). doi: 10.1016/S1474-4422(14)70201-7; pmid: 25792100
12. H. M. Schuller, Is cancer triggered by altered signalling of nicotinic acetylcholine receptors? *Nat. Rev. Cancer* **9**, 195–205 (2009). doi: 10.1038/nrc2590; pmid: 19194381
13. K. T. Dineley, A. A. Pandya, J. L. Yakel, Nicotinic ACh receptors as therapeutic targets in CNS disorders. *Trends Pharmacol. Sci.* **36**, 96–108 (2015). doi: 10.1016/j.tips.2014.12.002; pmid: 25639674
14. R. Hurst, H. Rollema, D. Bertrand, Nicotinic acetylcholine receptors: From basic science to therapeutics. *Pharmacol. Ther.* **137**, 22–54 (2013). doi: 10.1016/j.pharmthera.2012.08.012; pmid: 22925690
15. A. Gharpure, C. M. Noviello, R. E. Hibbs, Progress in nicotinic receptor structural biology. *Neuropharmacology* **171**, 108086 (2020). doi: 10.1016/j.neuropharm.2020.108086; pmid: 32272141
16. W. N. Green, T. Claudio, Acetylcholine receptor assembly: Subunit folding and oligomerization occur sequentially. *Cell* **74**, 57–69 (1993). doi: 10.1016/0092-8674(93)90294-Z; pmid: 8334706
17. S. Halevi *et al*., The *C. elegans ric-3* gene is required for maturation of nicotinic acetylcholine receptors. *EMBO J.* **21**, 1012–1020 (2002). doi: 10.1093/emboj/21.5.1012; pmid: 11867529
18. J. M. Miwa *et al*., lynx1, an endogenous toxin-like modulator of nicotinic acetylcholine receptors in the mammalian CNS. *Neuron* **23**, 105–114 (1999). doi: 10.1016/S0896-6273(00)80757-6; pmid: 10402197
19. M. P. Maher, J. A. Matta, S. Gu, M. Seierstad, D. S. Bredt, Getting a Handle on Neuropharmacology by Targeting Receptor-Associated Proteins. *Neuron* **96**, 989–1001 (2017). doi: 10.1016/j.neuron.2017.10.001; pmid: 29216460
20. S. J. Burden, M. G. Huijbers, L. Remedio, Fundamental Molecules and Mechanisms for Forming and Maintaining Neuromuscular Synapses. *Int. J. Mol. Sci.* **19**, 490 (2018). doi: 10.3390/ijms19020490; pmid: 29415504
21. N. E. Gilhus, Myasthenia Gravis. *N. Engl. J. Med.* **375**, 2570–2581 (2016). doi: 10.1056/NEJMra1602678; pmid: 28029925
22. R. Thompson, G. Bonne, P. Missier, H. Lochmüller, Targeted therapies for congenital myasthenic syndromes: Systematic review and steps towards a treatabolome. *Emerg Top Life Sci* **3**, 19–37 (2019). doi: 10.1042/ETLS20180100; pmid: 30931400
23. M. D'Alessandro *et al*., CRELD1 is an evolutionarily-conserved maturational enhancer of ionotropic acetylcholine receptors. *eLife* **7**, e39649 (2018). doi: 10.7554/eLife.39649; pmid: 30407909
24. M. G. Huijbers, J. J. G. M. Verschuuren, Treating muscle-specific kinase myasthenia gravis from the inside out. *Neurol. Neuroimmunol. Neuroinflamm.* **7**, e646 (2019). doi: 10.1212/NXI.0000000000000646; pmid: 31831572
25. N. S. Millar, P. C. Harkness, Assembly and trafficking of nicotinic acetylcholine receptors (Review). *Mol. Membr. Biol.* **25**, 279–292 (2008). doi: 10.1080/09687680802035675; pmid: 18446614
26. S. Gu *et al*., Brain α7 Nicotinic Acetylcholine Receptor Assembly Requires NACHO. *Neuron* **89**, 948–955 (2016). doi: 10.1016/j.neuron.2016.01.018; pmid: 26875622
27. C. M. Noviello *et al*., Structure and gating mechanism of the α7 nicotinic acetylcholine receptor. *Cell* **184**, 2121–2134.e13 (2021). doi: 10.1016/j.cell.2021.02.049; pmid: 33735609

28. J. A. Matta *et al*., NACHO Mediates Nicotinic Acetylcholine Receptor Function throughout the Brain. *Cell Rep.* **19**, 688–696 (2017). doi: 10.1016/j.celrep.2017.04.008; pmid: 28445721
29. H. J. Kweon *et al*., NACHO Engages N-Glycosylation ER Chaperone Pathways for α7 Nicotinic Receptor Assembly. *Cell Rep.* **32**, 108025 (2020). doi: 10.1016/j.celrep.2020.108025; pmid: 32783947
30. D. Chen, H. Dang, J. W. Patrick, Contributions of N-linked glycosylation to the expression of a functional α7-nicotinic receptor in *Xenopus* oocytes. *J. Neurochem.* **70**, 349–357 (1998). doi: 10.1046/j.1471-4159.1998.70010349.x; pmid: 9422381
31. S. H. Keller, J. Lindstrom, P. Taylor, Involvement of the chaperone protein calnexin and the acetylcholine receptor β-subunit in the assembly and cell surface expression of the receptor. *J. Biol. Chem.* **271**, 22871–22877 (1996). doi: 10.1074/jbc.271.37.22871; pmid: 8798466
32. S. Halevi *et al*., Conservation within the RIC-3 gene family. Effectors of mammalian nicotinic acetylcholine receptor expression. *J. Biol. Chem.* **278**, 34411–34417 (2003). doi: 10.1074/jbc.M300170200; pmid: 12821669
33. G. B. Dawe *et al*., α7 nicotinic acetylcholine receptor upregulation by anti-apoptotic Bcl-2 proteins. *Nat. Commun.* **10**, 2746 (2019). doi: 10.1038/s41467-019-10723-x; pmid: 31227712
34. M. Castillo *et al*., Dual role of the RIC-3 protein in trafficking of serotonin and nicotinic acetylcholine receptors. *J. Biol. Chem.* **280**, 27062–27068 (2005). doi: 10.1074/jbc.M503746200; pmid: 15927954
35. R. Srinivasan, B. J. Henderson, H. A. Lester, C. I. Richards, Pharmacological chaperoning of nAChRs: A therapeutic target for Parkinson's disease. *Pharmacol. Res.* **83**, 20–29 (2014). doi: 10.1016/j.phrs.2014.02.005; pmid: 24593907
36. B. J. Henderson *et al*., Menthol alone upregulates midbrain nAChRs, alters nAChR subtype stoichiometry, alters dopamine neuron firing frequency, and prevents nicotine reward. *J. Neurosci.* **36**, 2957–2974 (2016). doi: 10.1523/JNEUROSCI.4194-15.2016; pmid: 26961950
37. M. Dhara *et al*., Polyamine regulation of ion channel assembly and implications for nicotinic acetylcholine receptor pharmacology. *Nat. Commun.* **11**, 2799 (2020). doi: 10.1038/s41467-020-16629-3; pmid: 32493979
38. A. Kuryatov, J. Mukherjee, J. Lindstrom, Chemical chaperones exceed the chaperone effects of RIC-3 in promoting assembly of functional α7 AChRs. *PLOS ONE* **8**, e62246 (2013). doi: 10.1371/journal.pone.0062246; pmid: 23638015
39. R. D. Schwartz, K. J. Kellar, Nicotinic cholinergic receptor binding sites in the brain: Regulation in vivo. *Science* **220**, 214–216 (1983). doi: 10.1126/science.6828889; pmid: 6828889
40. M. J. Marks *et al*., Nicotine binding and nicotinic receptor subunit RNA after chronic nicotine treatment. *J. Neurosci.* **12**, 2765–2784 (1992). doi: 10.1523/JNEUROSCI.12-07-02765.1992; pmid: 1613557
41. C. M. Flores, S. W. Rogers, L. A. Pabreza, B. B. Wolfe, K. J. Kellar, A subtype of nicotinic cholinergic receptor in rat brain is composed of alpha 4 and beta 2 subunits and is up-regulated by chronic nicotine treatment. *Mol. Pharmacol.* **41**, 31–37 (1992). pmid: 1732720
42. A. L. Brody *et al*., Up-regulation of nicotinic acetylcholine receptors in menthol cigarette smokers. *Int. J. Neuropsychopharmacol.* **16**, 957–966 (2013). doi: 10.1017/S1461145712001022; pmid: 23171716
43. M. Hans, M. Wilhelm, D. Swandulla, Menthol suppresses nicotinic acetylcholine receptor functioning in sensory neurons via allosteric modulation. *Chem. Senses* **37**, 463–469 (2012). doi: 10.1093/chemse/bjr128; pmid: 22281529
44. R. Shahoei, E. Tajkhorshid, Menthol Binding to the Human α4β2 Nicotinic Acetylcholine Receptor Facilitated by Its Strong Partitioning in the Membrane. *J. Phys. Chem. B* **124**, 1866–1880 (2020). doi: 10.1021/acs.jpcb.9b10092; pmid: 32048843
45. K. Ahijevych, B. E. Garrett, The role of menthol in cigarettes as a reinforcer of smoking behavior. *Nicotine Tob. Res.* **12**, S110–S116 (2010). doi: 10.1093/ntr/ntq203; pmid: 21177367
46. S. Gu *et al*., Hair cell α9α10 nicotinic acetylcholine receptor functional expression regulated by ligand binding and deafness gene products. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24534–24544 (2020). doi: 10.1073/pnas.2013762117; pmid: 32929005
47. K. Williams, Interactions of polyamines with ion channels. *Biochem. J.* **325**, 289–297 (1997). doi: 10.1042/bj3250289; pmid: 9230104
48. G. Turecki, The molecular bases of the suicidal brain. *Nat. Rev. Neurosci.* **15**, 802–816 (2014). doi: 10.1038/nrn3839; pmid: 25354482
49. J. A. Gandelman, P. Newhouse, W. D. Taylor, Nicotine and networks: Potential for enhancement of mood and cognition in late-life depression. *Neurosci. Biobehav. Rev.* **84**, 289–298 (2018). doi: 10.1016/j.neubiorev.2017.08.018; pmid: 28859996
50. T. Boulin *et al*., Positive modulation of a Cys-loop acetylcholine receptor by an auxiliary transmembrane subunit. *Nat. Neurosci.* **15**, 1374–1381 (2012). doi: 10.1038/nn.3197; pmid: 22922783
51. S. Choudhary *et al*., EAT-18 is an essential auxiliary protein interacting with the non-alpha nAChR subunit EAT-2 to form a functional receptor. *PLOS Pathog.* **16**, e1008396 (2020). doi: 10.1371/journal.ppat.1008396; pmid: 32243475
52. S. R. Letchworth, P. Whiteaker, Progress and challenges in the study of α6-containing nicotinic acetylcholine receptors. *Biochem. Pharmacol.* **82**, 862–872 (2011). doi: 10.1016/j.bcp.2011.06.022; pmid: 21736871
53. V. Gerzanich, A. Kuryatov, R. Anand, J. Lindstrom, "Orphan" α6 nicotinic AChR subunit can form a functional heteromeric acetylcholine receptor. *Mol. Pharmacol.* **51**, 320–327 (1997). doi: 10.1124/mol.51.2.320; pmid: 9203638
54. M. Zoli *et al*., Identification of the nicotinic receptor subtypes expressed on dopaminergic terminals in the rat striatum. *J. Neurosci.* **22**, 8785–8789 (2002). doi: 10.1523/JNEUROSCI.22-20-08785.2002; pmid: 12388584
55. S. Gu *et al*., α6-Containing Nicotinic Acetylcholine Receptor Reconstitution Involves Mechanistically Distinct Accessory Components. *Cell Rep.* **26**, 866–874.e3 (2019). doi: 10.1016/j.celrep.2018.12.103; pmid: 30673609
56. P. Béguin *et al*., BARP suppresses voltage-gated calcium channel activity and $Ca^{2+}$-evoked exocytosis. *J. Cell Biol.* **205**, 233–249 (2014). doi: 10.1083/jcb.201304101; pmid: 24751537
57. J. M. Kulak, J. M. McIntosh, D. Yoshikami, B. M. Olivera, Nicotine-evoked transmitter release from synaptosomes: Functional association of specific presynaptic acetylcholine receptors and voltage-gated calcium channels. *J. Neurochem.* **77**, 1581–1589 (2001). doi: 10.1046/j.1471-4159.2001.00357.x; pmid: 11413241
58. J. S. Wieskopf *et al*., The nicotinic α6 subunit gene determines variability in chronic pain sensitivity via cross-inhibition of P2X2/3 receptors. *Sci. Transl. Med.* **7**, 287ra72 (2015). doi: 10.1126/scitranslmed.3009986; pmid: 25972004
59. D. Knowland *et al*., Functional α6β4 acetylcholine receptor expression enables pharmacological testing of nicotinic agonists with analgesic properties. *J. Clin. Invest.* **130**, 6158–6170 (2020). doi: 10.1172/JCI140311; pmid: 33074244
60. P. Walter, D. Ron, The unfolded protein response: From stress pathway to homeostatic regulation. *Science* **334**, 1081–1086 (2011). doi: 10.1126/science.1209038; pmid: 22116877
61. B. J. Henderson, H. A. Lester, Inside-out neuropharmacology of nicotinic drugs. *Neuropharmacology* **96**, 178–193 (2015). doi: 10.1016/j.neuropharm.2015.01.022; pmid: 25660637
62. A. B. Elgoyhen, E. Katz, P. A. Fuchs, The nicotinic receptor of cochlear hair cells: A possible pharmacotherapeutic target? *Biochem. Pharmacol.* **78**, 712–719 (2009). doi: 10.1016/j.bcp.2009.05.023; pmid: 19481062
63. P. A. Fuchs, A. M. Lauer, Efferent Inhibition of the Cochlea. *Cold Spring Harb. Perspect. Med.* **9**, a033530 (2019). doi: 10.1101/cshperspect.a033530; pmid: 30082454
64. P. Dallos *et al*., Acetylcholine, outer hair cell electromotility, and the cochlear amplifier. *J. Neurosci.* **17**, 2212–2226 (1997). doi: 10.1523/JNEUROSCI.17-06-02212.1997; pmid: 9045745
65. S. Naz *et al*., Mutations in a novel gene, *TMIE*, are associated with hearing loss linked to the *DFNB6* locus. *Am. J. Hum. Genet.* **71**, 632–636 (2002). doi: 10.1086/342193; pmid: 12145746
66. J. Li *et al*., Whole-exome sequencing identifies a variant in *TMEM132E* causing autosomal-recessive nonsyndromic hearing loss DFNB99. *Hum. Mutat.* **36**, 98–105 (2015). doi: 10.1002/humu.22712; pmid: 25331638
67. C. L. Cunningham *et al*., TMIE Defines Pore and Gating Properties of the Mechanotransduction Channel of Mammalian Cochlear Hair Cells. *Neuron* **107**, 126–143.e8 (2020). doi: 10.1016/j.neuron.2020.03.033; pmid: 32343945
68. K. L. Mitchem *et al*., Mutation of the novel gene *Tmie* results in sensory cell defects in the inner ear of spinner, a mouse model of human hearing loss DFNB6. *Hum. Mol. Genet.* **11**, 1887–1898 (2002). doi: 10.1093/hmg/11.16.1887; pmid: 12140191
69. I. Ibañez-Tallon *et al*., Novel modulation of neuronal nicotinic acetylcholine receptors by association with the endogenous prototoxin lynx1. *Neuron* **33**, 893–903 (2002). doi: 10.1016/S0896-6273(02)00632-3; pmid: 11906696
70. W. A. Nichols *et al*., Lynx1 shifts α4β2 nicotinic receptor subunit stoichiometry by affecting assembly in the endoplasmic reticulum. *J. Biol. Chem.* **289**, 31423–31432 (2014). doi: 10.1074/jbc.M114.573667; pmid: 25193667
71. A. B. Tekinay *et al*., A role for LYNX2 in anxiety-related behavior. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4477–4482 (2009). doi: 10.1073/pnas.0813109106; pmid: 19246390
72. C. A. Puddifoot, M. Wu, R. J. Sung, W. J. Joiner, Ly6h regulates trafficking of alpha7 nicotinic acetylcholine receptors and nicotine-induced potentiation of glutamatergic signaling. *J. Neurosci.* **35**, 3420–3430 (2015). doi: 10.1523/JNEUROSCI.3630-14.2015; pmid: 25716842
73. Y. Moriwaki *et al*., Endogenous neurotoxin-like protein Ly6H inhibits alpha7 nicotinic acetylcholine receptor currents at the plasma membrane. *Sci. Rep.* **10**, 11996 (2020). doi: 10.1038/s41598-020-68947-7; pmid: 32686737
74. F. Chimienti *et al*., Identification of SLURP-1 as an epidermal neuromodulator explains the clinical phenotype of Mal de Meleda. *Hum. Mol. Genet.* **12**, 3017–3024 (2003). doi: 10.1093/hmg/ddg320; pmid: 14506129
75. J. M. Miwa, K. R. Anderson, K. M. Hoffman, Lynx Prototoxins: Roles of Endogenous Mammalian Neurotoxin-Like Proteins in Modulating Nicotinic Acetylcholine Receptor Function to Influence Complex Biological Processes. *Front. Pharmacol.* **10**, 343 (2019). doi: 10.3389/fphar.2019.00343; pmid: 31114495
76. N. I. Nissen *et al*., Augmenting the antinociceptive effects of nicotinic acetylcholine receptor activity through lynx1 modulation. *PLOS ONE* **13**, e0199643 (2018). doi: 10.1371/journal.pone.0199643; pmid: 29969495
77. S. Mazzaferro, S. T. Whiteman, C. Alcaino, A. Beyder, S. M. Sine, NACHO and 14-3-3 promote expression of distinct subunit stoichiometries of the α4β2 acetylcholine receptor. *Cell. Mol. Life Sci.* **78**, 1565–1575 (2021). doi: 10.1007/s00018-020-03592-x; pmid: 32676916
78. R. L. Papke, J. M. Lindstrom, Nicotinic acetylcholine receptors: Conventional and unconventional ligands and signaling. *Neuropharmacology* **168**, 108021 (2020). doi: 10.1016/j.neuropharm.2020.108021; pmid: 32146229
79. A. Deshpande *et al*., Why Does Knocking Out NACHO, But Not RIC3, Completely Block Expression of α7 Nicotinic Receptors in Mouse Brain? *Biomolecules* **10**, 470 (2020). doi: 10.3390/biom10030470; pmid: 32204458
80. S. Cory, Regulation of lymphocyte survival by the bcl-2 gene family. *Annu. Rev. Immunol.* **13**, 513–543 (1995). doi: 10.1146/annurev.iy.13.040195.002501; pmid: 7612233
81. A. S. Vallés, F. J. Barrantes, Chaperoning α7 neuronal nicotinic acetylcholine receptors. *Biochim. Biophys. Acta* **1818**, 718–729 (2012). doi: 10.1016/j.bbamem.2011.10.012; pmid: 22040696
82. D. E. Merry, S. J. Korsmeyer, Bcl-2 gene family in the nervous system. *Annu. Rev. Neurosci.* **20**, 245–267 (1997). doi: 10.1146/annurev.neuro.20.1.245; pmid: 9056714
83. S. Kaneko *et al*., Nicotine protects cultured cortical neurons against glutamate-induced cytotoxicity via α7-neuronal receptors and neuronal CNS receptors. *Brain Res.* **765**, 135–140 (1997). doi: 10.1016/S0006-8993(97)00556-8; pmid: 9310404
84. B. J. Morley, D. D. Simmons, Developmental mRNA expression of the α10 nicotinic acetylcholine receptor subunit in the rat cochlea. *Brain Res. Dev. Brain Res.* **139**, 87–96 (2002). doi: 10.1016/S0165-3806(02)00514-X; pmid: 12414097
85. E. Katz *et al*., Developmental regulation of nicotinic synapses on cochlear inner hair cells. *J. Neurosci.* **24**, 7814–7820 (2004). doi: 10.1523/JNEUROSCI.2102-04.2004; pmid: 15356192
86. S. P. Zachary, P. A. Fuchs, Re-Emergent Inhibition of Cochlear Inner Hair Cells in a Mouse Model of Hearing Loss. *J. Neurosci.* **35**, 9701–9706 (2015). doi: 10.1523/JNEUROSCI.0879-15.2015; pmid: 26134652
87. J. W. Daly *et al*., Alkaloids from frog skin: The discovery of epibatidine and the potential for developing novel non-opioid analgesics. *Nat. Prod. Rep.* **17**, 131–135 (2000). doi: 10.1039/a900728h; pmid: 10821107
88. M. C. Rowbotham, R. W. Duan, J. Thomas, W. Nothaft, M. M. Backonja, A randomized, double-blind, placebo-controlled trial evaluating the efficacy and safety of ABT-594 in patients with diabetic peripheral neuropathic pain. *Pain* **146**, 245–252 (2009). doi: 10.1016/j.pain.2009.06.013; pmid: 19632048
89. M. C. Rowbotham *et al*., Efficacy and safety of the α4β2 neuronal nicotinic receptor agonist ABT-894 in patients

with diabetic peripheral neuropathic pain. *Pain* **153**, 862–868 (2012). doi: 10.1016/j.pain.2012.01.009; pmid: 22386472

90. L. E. Boero *et al*., Preventing presbycusis in mice with enhanced medial olivocochlear feedback. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11811–11819 (2020). doi: 10.1073/pnas.2000760117; pmid: 32393641
91. S. Siwani *et al*., OLMα2 Cells Bidirectionally Modulate Learning. *Neuron* **99**, 404–412.e3 (2018). doi: 10.1016/j.neuron.2018.06.022; pmid: 29983324
92. C. D. Fowler, Q. Lu, P. M. Johnson, M. J. Marks, P. J. Kenny, Habenular α5 nicotinic receptor subunit signalling controls nicotine intake. *Nature* **471**, 597–601 (2011). doi: 10.1038/nature09797; pmid: 21278726
93. P. K. Joshi *et al*., Variants near CHRNA3/5 and APOE have age- and sex-related effects on human lifespan. *Nat. Commun.* **7**, 11174 (2016). doi: 10.1038/ncomms11174; pmid: 27029810

## ACKNOWLEDGMENTS

The authors thank scientists in the neuroscience department at Janssen who have contributed to research on nAChRs and inspired aspects of this Review. We thank M. Miyamoto, who assisted in creating the figures for this Review. **Funding:** All authors were full-time employees of Janssen while this Review was written. **Author contributions:** All authors (J.A.M., S.G., W.B.D., and D.S.B.) contributed to writing, reviewing, and editing the manuscript; S.G. originally conceptualized the figures. **Competing interests:** All authors are coinventors on patents [Expression systems for the α9α10 nAChR and methods of use thereof (WO2020234179A1) and Expression systems for the α6β4 nAChR and methods of use thereof (US 63/178835)] held by Janssen that describe methods for enabling functional expression of nAChRs to aid in drug discovery.

10.1126/science.abg6539

RESEARCH ARTICLE SUMMARY

MICROBIOLOGY

# Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution

**Daniel Dar, Nina Dar, Long Cai*, Dianne K. Newman***

**INTRODUCTION:** Microbial populations display heterogeneous gene expression profiles that result in phenotypic differences between individual bacteria. This diversity can allow populations to survive under uncertain and fluctuating conditions such as sudden antibiotic exposure, divide costly functions across different subpopulations, and enable interactions between different phenotypes. In addition to the temporal phenotypic heterogeneity seen in planktonic cultures, microbial populations and communities often exist in multicellular biofilms that exhibit considerable heterogeneity at the microscale, both in the local physicochemistry that individuals experience and in the species composition in their neighborhoods. Phenotypic and microscale variation represent central features of microbial populations, but the landscape of possible cellular states, their spatiotemporal regulation, and their roles in many biological phenomena are still largely unknown.

**RATIONALE:** The microscale heterogeneity that defines microbial life can play important roles in community organization and function, including in antibiotic resistance and virulence. However, our understanding of these basic features has been limited by our ability to capture this heterogeneity at the relevant spatiotemporal scales. Overcoming these limitations could lead to new insights into the inner workings of microbial assemblages.

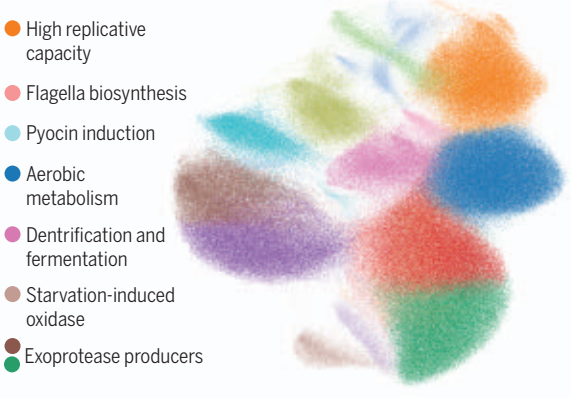
**RESULTS:** We developed par-seqFISH (parallel sequential fluorescence in situ hybridization), a high-throughput method that captures gene expression profiles of individual bacteria while also preserving their physical context within spatially structured environments. We applied this approach to the study of *Pseudomonas aeruginosa*, a model biofilm-forming bacterium and an opportunistic human pathogen. Focusing on a set of 105 marker genes representing key aspects of *P. aeruginosa* physiology and virulence, we explored the transcriptional profiles of >600,000 bacteria across dozens of growth conditions. We uncovered a diverse set of metabolic- and virulence-related cellular states and quantified their temporal dynamics during population growth. In addition to recording gene expression, we also demonstrated that par-seqFISH captures cell biological parameters such as cell size and can be further integrated with specific dyes to measure features such as chromosome copy in the same cells. Applying par-seqFISH to developing *P. aeruginosa* biofilms, we exposed the biogeographic context of cellular states, providing new insights into the spatial expression of various genes. These included, among other things, mutually exclusive expression patterns of flagella and type IV pili genes and a localized induction of pyocins, which are involved in kin selection and extracellular DNA release. Looking more closely, we found that pyocin-encoding transcripts strongly localized to the bacterial cell poles. In early biofilms, we identified extensive microscale phenotypic structuring in which anaerobic metabolic processes such as denitrification appeared to locally influence the microenvironment through byproduct production. In more mature biofilms, we detected larger-scale partitions into zones of differential metabolic activities and virulence factor biosynthesis.

**CONCLUSION:** Transcriptome imaging using par-seqFISH captures the microscale phenotypic variation of free-living and sessile bacterial populations. We report extensive heterogeneity in growing *P. aeruginosa* populations and demonstrate that individual multicellular biofilms can contain coexisting but separated subpopulations with distinct physiological activities. This multiplexed and spatially resolved method offers a generalizable tool for studying bacterial populations in space and time directly in their native contexts. Future applications in natural and clinical samples will provide insights into the conditions experienced by microbes in complex environments and the coordinated physiological responses that emerge in turn and reshape them. ■

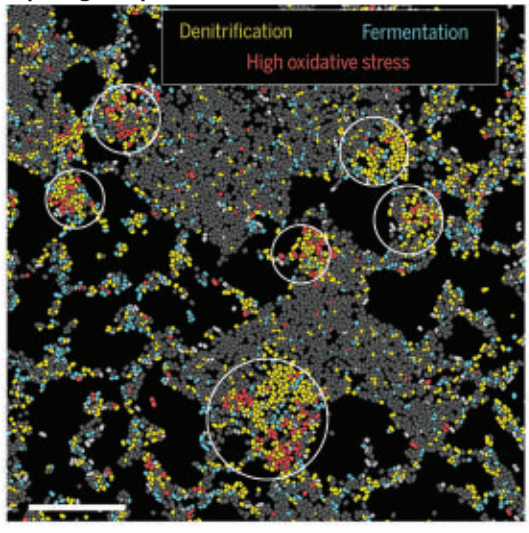
The list of author affiliations is available in the full article online.
*Corresponding author. Email: dkn@caltech.edu (D.K.N.); lcai@caltech.edu (L.C.)
Cite this article as D. Dar *et al.*, *Science* 373, eabi4882 (2021). DOI: 10.1126/science.abi4882

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abi4882

**Single-cell transcriptional profiling of planktonic and biofilm populations with par-seqFISH**




**Exploring the spatial context of cellular states**




**Transcriptome imaging using par-seqFISH reveals the dynamics and spatial organization of transcriptional programs in *P. aeruginosa* populations at single-cell resolution.** Transcriptional states of individual bacterial cells were identified using clustering analysis (left). The detected cellular states are depicted in different colors. Cell metabolic states can be directly mapped to their native biofilm context to identify emerging microenvironment dynamics (right).

RESEARCH ARTICLE

MICROBIOLOGY

# Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution

Daniel Dar[1,2]†, Nina Dar[1], Long Cai[1]*, Dianne K. Newman[1,2]*

**Capturing the heterogeneous phenotypes of microbial populations at relevant spatiotemporal scales is highly challenging. Here, we present par-seqFISH (parallel sequential fluorescence in situ hybridization), a transcriptome-imaging approach that records gene expression and spatial context within microscale assemblies at a single-cell and molecule resolution. We applied this approach to the opportunistic pathogen *Pseudomonas aeruginosa*, analyzing about 600,000 individuals across dozens of conditions in planktonic and biofilm cultures. We identified numerous metabolic- and virulence-related transcriptional states that emerged dynamically during planktonic growth, as well as highly spatially resolved metabolic heterogeneity in sessile populations. Our data reveal that distinct physiological states can coexist within the same biofilm just several micrometers away, underscoring the importance of the microenvironment. Our results illustrate the complex dynamics of microbial populations and present a new way of studying them at high resolution.**

Life exists in context. Cells within microbial populations and communities are typically closely associated with one another in multicellular biofilms, whether found within infected tissues, attached to surfaces, or forming assemblages in the deep sea (*1*, *2*). Natural microbiota and infectious bacteria generally exist in biofilm aggregates that are several dozen micrometers across and can contain many interacting species (*3*–*5*). Despite the ubiquity of the biofilm lifestyle in both natural and manmade habitats, understanding what life is like within a biofilm for individual microbes has proven challenging. Whereas single-cell–level activities have been tracked at high spatial resolution using a variety of approaches in diverse contexts (*6*–*8*), we have been unable to resolve the hundreds, if not thousands, of concurrent activities that characterize microbial life at relevant spatiotemporal scales. What we understand about microbial life literally has been limited by our ability to see.

Despite this limitation, it has become clear in recent years that extreme phenotypic heterogeneity defines the microbial experience (*9*, *10*). This is as true for isogenic populations as it is for complex biofilm communities. Clonemates sampled from the same environment often display substantial differences that are thought to result from stochastic gene expression and variable environmental factors (*9*, *11*, *12*). The detection of phenotypic diversity even in seemingly well-mixed environments such as chemostats (*11*, *13*) also serves as a powerful reminder that life at the microscale may inhabit far more diverse niches than are readily apparent. Phenotypic diversity has been rationalized as providing microbes with a fitness advantage in an unpredictable world (*9*, *14*). In addition, specialized functions have been proposed to underpin collective interactions such as division of labor (*9*, *15*–*17*). However, little is still known about the range of possible cellular phenotypic states and their roles in most biological processes.

What triggers such phenotypic plasticity? And are there underlying "rules" that govern any patterns that may exist at the microscale? In sessile communities, clonal or multispecies, biological activities give rise to changing chemical gradients that create a range of local microenvironments (*18*, *19*). Furthermore, spatial organization enables different conflicting metabolic states or species to coexist through physical separation, increasing the potential for diversity and allowing for new interactions to emerge (*10*, *20*–*23*). Indeed, natural communities often contain many interacting species that assemble into intricate spatial structures. These microscale assemblies can promote interactions between species and represent a key ecosystem feature (*23*, *24*). However, a wide gulf, limited by technology, still separates such observations from a coherent conceptual framework to explain the rules governing microbial ecology.

Recent advances in imaging methods provide a means to chart the physical associations between different species in natural environments (*4*, *25*–*27*), but interpreting these maps remains challenging without additional functional information on the physiological states and activities of relevant community members. By contrast, recent adaptations of single-cell RNA sequencing (RNA-seq) approaches to free-living bacteria provide a powerful means of exploring their phenotypic landscape (*28*–*30*). However, these approaches do not preserve the spatial context of analyzed cells and are therefore limited in their capacity to address single and multispecies biofilms. Thus, a major gap exists in our ability to account for both spatial and functional complexity, limiting progression toward a high-resolution understanding of microbial life.

Single-molecule fluorescence in situ hybridization (FISH)–based technologies have been used to measure gene expression directly within native tissues, recording both spatial and functional information. These methods can shed important light on single-cell heterogeneity but are traditionally limited to measuring the expression of only a few genes at a time (*31*–*34*). In addition to this limited throughput, single-gene measurements do not provide a means to capture coordinated cellular responses, the molecular "fingerprint" of multiple biological activities that underpin distinct physiological states. Recent advances in combinatorial mRNA labeling and sequential FISH (seqFISH) allow for hundreds or even thousands of genes to be analyzed within the same sample at a submicrometer resolution (*35*–*37*). Until now, seqFISH has been used in mammalian systems to expose the physical organization of cell states within tissues (*35*–*39*). We reasoned that the high spatial resolution of these modern transcriptome-imaging techniques also had the potential to illuminate the microscale organization of microbial populations and communities.

In this study, we adapted and further developed seqFISH for studying bacteria, measuring the expression of hundreds of genes within individual cells while also capturing their spatial context. We used *Pseudomonas aeruginosa* planktonic and biofilm populations to demonstrate how different cellular functions are coordinated in time and space. Our proof-of-concept work illustrates how the ability to observe transcriptional activities at the microscale permits insights into the spatiotemporal regulation and coordination of critical life processes, enabling hitherto unrecognized, transient physiological states to be identified and new hypotheses to be generated. These findings represent the tip of the iceberg, and the opportunities for discovery enabled by this approach promise to reveal new insights about the rules governing microbial ecology.

## A sequential mRNA-FISH framework for studying bacterial gene expression

Combinatorial mRNA labeling requires that each measured mRNA molecule be individually

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [2]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.
*Corresponding author. Email: dkn@caltech.edu (D.K.N.); lcai@caltech.edu (L.C.)
†Present address: Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 76100, Israel.

resolved. This is much more challenging in bacteria because of the small size of their cells, as many different mRNA molecules occur in close proximity and cannot be resolved using standard fluorescence microscopy. We therefore used a nonbarcoded seqFISH approach (*40*).

In seqFISH, target mRNAs are first hybridized with a set of primary, nonfluorescent probes, which are flanked by short sequences uniquely assigned per gene (Fig. 1A). Specific genes can be turned ON through a secondary hybridization with short, fluorescently labeled "readout" probes complementary to the gene-specific flanking sequences (Fig. 1A). Several genes can be measured at once using a set of readout probes labeled with different fluorophores. These short, fluorescent readout probes can be efficiently stripped and washed away from the sample without affecting the primary probes (*41*) (Fig. 1A). Thus, once expression is measured, fluorescence can be turned OFF and a new set of genes can be measured by introducing a new set of readout probes (Fig. 1B). This two-step design allows for potentially hundreds of genes to be measured sequentially, one after the other in the same sample, using automated microscopy (Fig. 1B). The individual gene mRNA-FISH data can be combined into spatially resolved multigene profiles at the single-bacterium level (Fig. 1B).

Because of the diffraction limit and the small size of bacteria, mRNA-FISH fluorescent signals (appearing as spots within cells) can contain overlapping mRNA molecules that cannot be spatially resolved using standard microscopes. Therefore, counting the number of spots within a bacterial cell severely underestimates expression levels. This problem can be overcome by integrating the fluorescence intensity per spot, which scales linearly with the number of mRNAs. Fluorescence intensity can be converted to discrete mRNA counts by measuring the characteristic intensity of a single transcript. This analog-to-digital conversion approach has been shown to provide a wide dynamic range in bacteria (*33*, *42*).

We developed seqFISH for the study of *P. aeruginosa*, an opportunistic human pathogen and a severe cause of morbidity and mortality in cystic fibrosis patients (*43*, *44*).

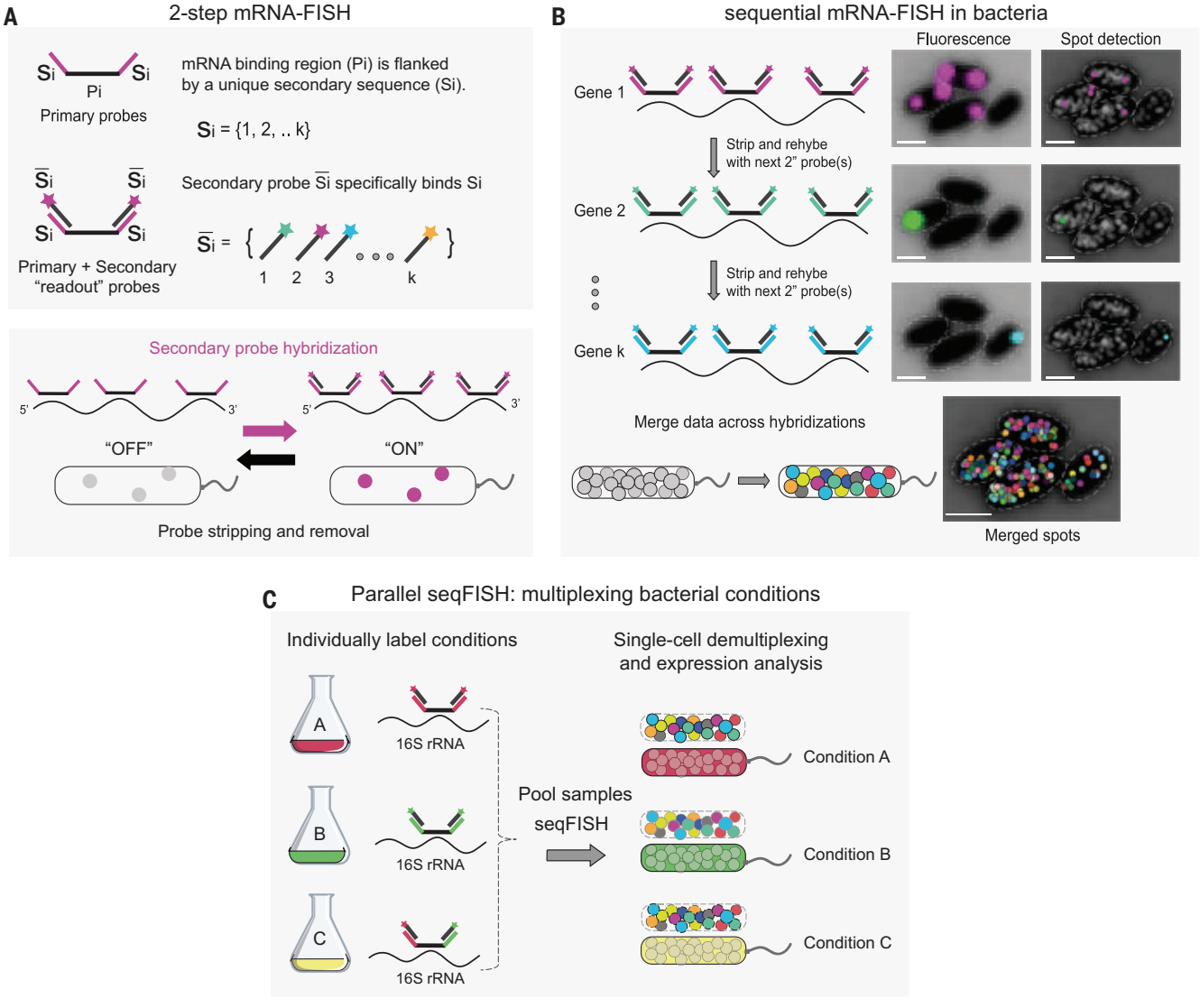



**Fig. 1. Parallel and sequential mRNA-FISH in bacteria.** (**A**) seqFISH probe design scheme. Primary probes contain unique sequences (Si) that are read by secondary probes (colored wands). Each gene is read by a unique probe and its fluorescence can be turned ON or OFF. (**B**) mRNA-FISH applied sequentially to the same sample. In each cycle, a new set of secondary readout probes are introduced. Raw fluorescence data are shown on the right, and the detected local spot maxima are shown in the spot detection image. Merged spots for many genes are shown in shuffled colors. (**C**) Combinatorial labeling can be used to encode species taxonomy using 16S rRNA or to enable the parallel study of bacteria grown in different conditions.

We generated a probe library targeting a set of 105 marker genes that capture many core physiological aspects of this pathogen (tables S1 and S2). These included genes involved in biosynthetic capacity (ribosome and RNA-polymerase subunits), anerobic physiology (fermentation and denitrification pathways), stress responses (oxidative and nutrient limitation), cellular signaling [cyclic diguanylate monophosphate (c-di-GMP)], biofilm matrix components, motility (flagella and T4P), all major quorum-sensing (QS) systems, as well as multiple antibiotic resistance and core virulence factors. In addition, to control for false positives, we designed probes targeting three different negative control genes that do not exist in *P. aeruginosa* (fig. S1).

### Parallel and sequential mRNA-FISH in single bacterial cells

To test our bacterial seqFISH approach, we first studied *P. aeruginosa* grown in well-understood batch culture conditions. We performed a growth curve experiment in lysogeny broth (LB) medium, in which key parameters such as cell density, growth rate, and oxygen levels change in a predictable manner. We collected 11 time points representing the lag phase, exponential growth phase, and stationary phase, and imaged the expression of 105 genes within them simultaneously for 2 days (Fig. 2A). Independent imaging of these 11 samples in a serial manner would have taken ~3 weeks of automated microscopy time.

To perform simultaneous imaging, we developed a multiplexing method that enables parallel seqFISH (par-seqFISH) experiments. We designed a set of primary probes targeting the 16*S* ribosomal RNA (rRNA) (Ribo-Tags), which contain unique combinations of flanking sequences (barcodes) that serve as the "readout" in a seqFISH run (Fig. 1, C and D, and table S3). In principle, this multiplexing approach can be applied to studying combinations of different species or for pooling bacteria from different growth conditions (Fig. 1C). We validated the latter application by individually labeling the 16*S* rRNAs of each of the 11 growth curve samples with unique Ribo-Tags. The samples were pooled, collectively hybridized with the 105-gene-probe library, and subjected to sequential hybridizations to measure gene expression and to decode cell identity (Fig. 2B). We acquired expression profiles for >50,000 individual *P. aeruginosa* cells, >91.8% of which were unambiguously decoded and assigned to the condition from which they originated (Fig. 2B). We estimated the false-positive decoding rate to be 0.04% (one in 2500 cells) by counting the number of hits for barcodes left out of the experiment, demonstrating both high efficiency and accuracy for par-seqFISH.

In addition to acquiring mRNA expression profiles, our imaging-based platform permits concurrent tracking of key information such as cell size and shape and can be combined with functional stains, markers, and/or immunofluorescence measurements (*45*). This opens up the possibility of correlating particular expression profiles at the single-cell level with integrative physiological or cell biological parameters. We applied a 4′,6-diamidino-2-phenylindole (DAPI) stain as a part of the par-seqFISH experiment and used DAPI fluorescence to estimate the nucleoid size and chromosome copy per cell. Comparing cells at different stages of growth showed that both nucleoid size (estimating cell size) and chromosome number distributions followed identical trends, in agreement with the *P. aeruginosa* literature (*46*) (Fig. 2, C and D). We also estimated ribosome abundance using 16*S* rRNA fluorescence. The distribution of this metric differed significantly from that of the chromosome parameters, displaying contrasting intensities at different stages of the lag phase, increased variability at the deep stationary phase, and a delay in signal decline during the shift from exponential growth to the stationary phase (Fig. 2E). Conversely, the total number of mRNAs per cell (estimated by our 105 genes) differentiated each time point along the growth curve, reaching a maxima and minima at the fastest and slowest growth rates, respectively (Fig. 2F). These data further support the accuracy of our par-seqFISH multiplexing approach and demonstrate the unique ability of this method to integrate single-cell gene expression with global parameters.

To determine whether our expression profiles faithfully captured known physiological processes that occur during culture development, we grouped the cells according to their decoded conditions and calculated their average gene expression profiles. We found a temporally resolved expression pattern associated with different stages of growth (Fig. 2G). For example, genes representing high replicative and/or biosynthetic capacity, such as those involved in RNA and protein biosynthesis, reached their peak expression during the maximal division rate but decreased between 90- and 250-fold during the stationary phase (Fig. 2G). By contrast, stress factors involved in stationary phase adaptation and nutrient limitation peaked at low division rates and higher cell densities (Fig. 2G). QS signal production, receptor expression, and target activation reflect the known hierarchical QS-regulatory network (*47*). The expression of anaerobic metabolism genes occurred in two stages: (i) early induction of the fermentation and nitrate-nitrite reduction genes in the entry to stationary phase, in which hypoxic conditions emerge, followed by (ii) expression of the remaining denitrification pathway at lower predicted oxygen levels (*48*) (Fig. 2G). Furthermore, the shift from aerobic to anaerobic metabolism was accompanied by sequential exchanges in terminal oxidase identities, from *ccoN1* to *ccoN2* and finally *ccoN4*, concomitantly with the induction of phenazine biosynthesis (*49*, *50*) (Fig. 2G).

Repeated mRNA measurements of the same genes in independent and spaced hybridization rounds were well correlated, both in average expression and at the single-bacterium level (Pearson's $R$ = 0.86, 0.89 and 0.9, for *sigX*, *rpsC*, and *rpoS*, respectively). In addition, the three negative control genes had an average false-positive rate of 0.002 transcripts per cell (fig. S1). We also found a good correlation between par-seqFISH and previous RNA-seq experiments conducted under similar conditions (*51*) (Pearson's $R$ = 0.79 to 0.84), as well as a strong correlation between close time points along the growth curve (fig. S2). Together, these results further validate the accuracy of our multiplexing method and demonstrate that our marker genes capture diverse transcriptional states across a wide range of physiological conditions.

### Transient emergence of physiologically distinct subpopulations during LB growth

Phenotypic diversity in clonal populations can generate distinct subpopulations that adjust to dynamic environmental changes and specialize in different tasks at different times, setting a fertile ground for bet-hedging behaviors and complex interactions (*9*, *15*, *52*). The single-cell resolution and high sensitivity of seqFISH has the potential to shed light on this important yet largely unexplored aspect of microbial life.

We applied uniform manifold approximation and projection (UMAP) dimensionality reduction and unsupervised clustering to identify distinct transcriptional cell states in our single-cell expression data (*29*, *53*). The cell clusters detected by this analysis charted the phenotypic landscape in LB growth from the perspective of our chosen marker genes. Analyzing the 11 time points together, we detected 20 clusters (representing different subpopulations) with diverse predicted functional capabilities. These included, among others, differential replicative capacity, exoproduct biosynthesis, and virulence factor production (Fig. 3, A and B). We found that the sampled populations of most of the growth conditions were partitioned into multiple coexisting subgroups with distinct expression profiles (Fig. 3A, fig. S2, and table S4). Our data suggest that the degree of dispersion within this expression space (estimating phenotypic diversity) varies significantly between conditions and is elevated during the stationary phase (fig. S3 and table S4).

Our growth condition–specific analysis revealed intriguing dynamics during lag phase progression. It could be expected that lag phase cultures will follow a steady ribosome accumulation as the cells progress toward exponential

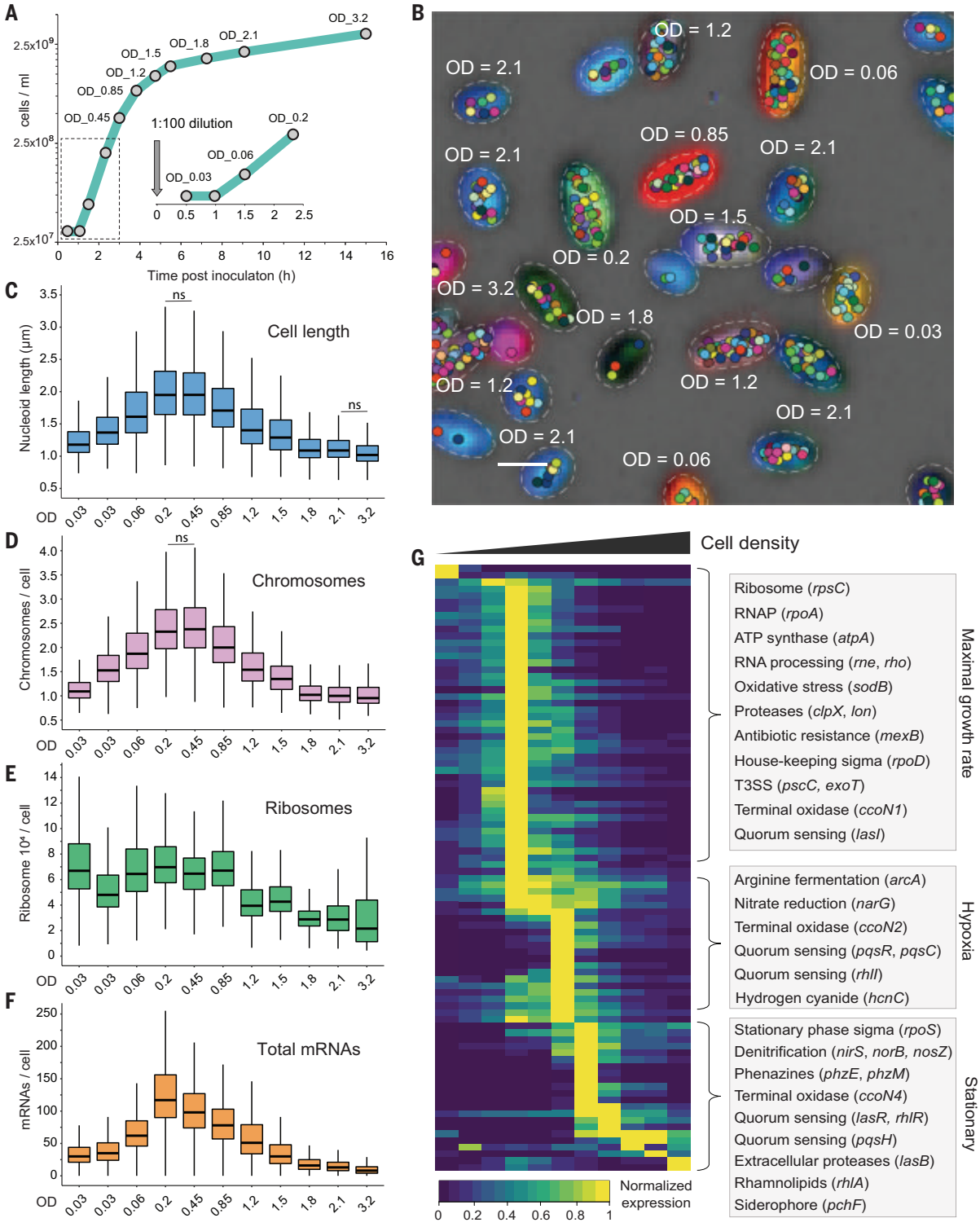



**Fig. 2. par-seqFISH of an LB growth curve experiment.** (**A**) Sampled LB growth curve. Collected time points are indicated with gray circles. Magnification shows the sampled lag phase. The presented colony-forming units per milliliter were estimated using $OD_{600}$ values [$OD_{600}$ = 1.0 reporting on ~$10^9$ cfu/ml (*104*)]. The $OD_{600}$ values are indicated over each time point. (**B**) Demultiplexed bacteria and their mRNAs. The merged, raw Ribo-Tag 16S rRNA fluorescence is shown for a representative region. Different barcodes (16S combinations) appear as different color combinations that identify the condition the cells were in when they were originally collected before sample pooling (indicated with the corresponding $OD_{600}$ value). Ellipses fitted to the segmented cell boundaries are shown. The mRNA spots (fitted position of maximal intensity) for all genes per cell are shown in unique colors per gene. Each spot may represent more than one mRNA copy. (**C** to **F**) Condition-specific distributions of nucleoid length, chromosome copy, ribosome levels, and total mRNAs detected across our gene set. Distributions contain all demultiplexed cells per condition and are significantly different from their previous time point unless otherwise noted (Wilcoxon, $P < 0.001$). (**G**) Heatmap showing average gene expression normalized to the maximal value for each gene across all conditions. Highlighted gene groups and their functions are indicated on the right.

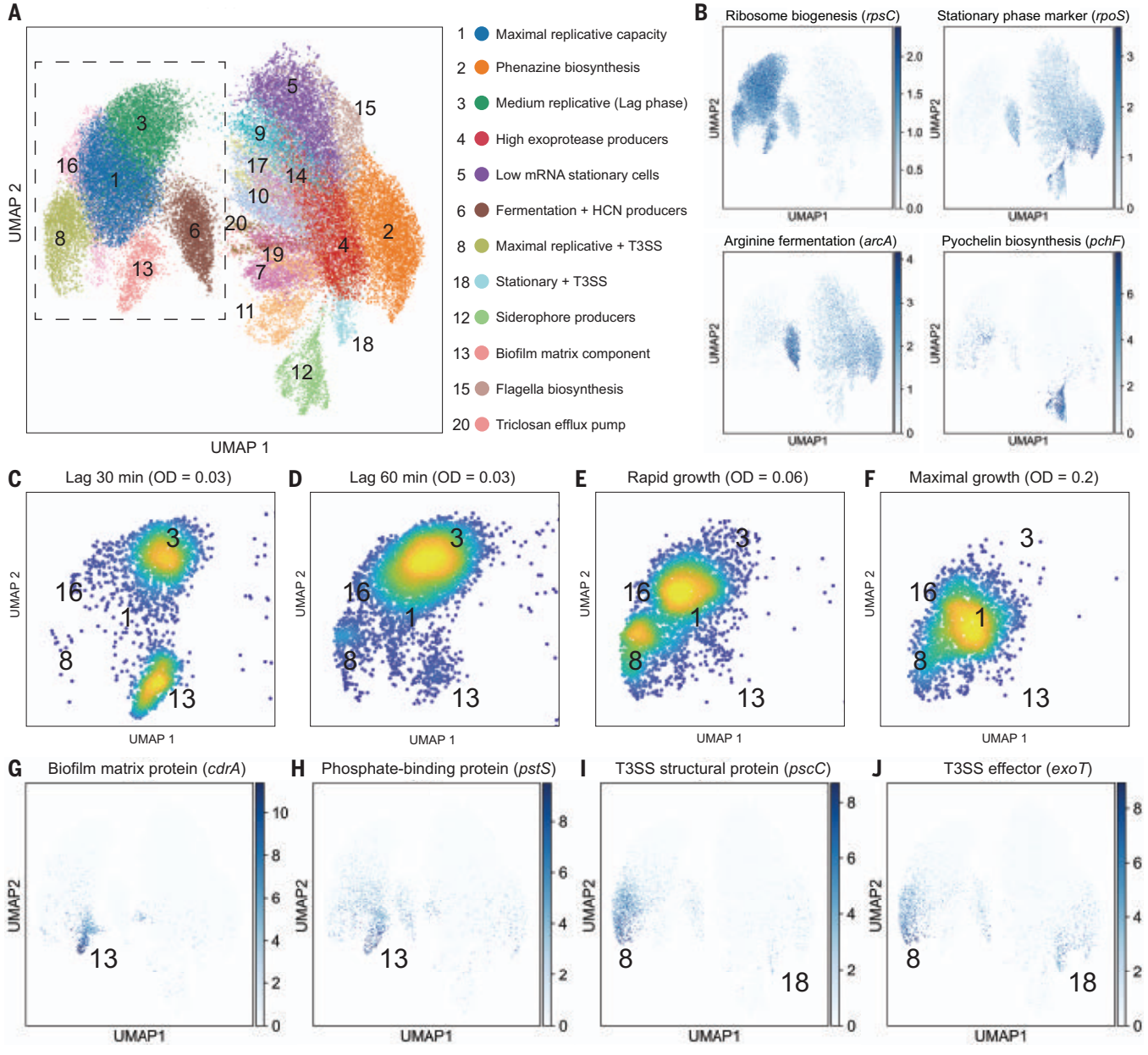



**Fig. 3. Single-bacterium analysis revealing physiologically distinct dynamic subpopulations.** (**A**) UMAP analysis using cells from all 11 time points. Identified clusters are shown in different colors and are indexed by group size. Specific groups and their enriched functions are shown on the right. (**B**) Gene expression overlays for four genes that report on metabolic state, stationary phase progression, and exoproduct biosynthesis. The color map shows the normalized expression scaled to unit variance. Cells from all 11 time points are displayed in the plot. (**C** to **F**) Density scatter plots of cells from individual conditions in a magnification of the UMAP [dashed box in (A)]. The clusters are indicated by their index. (**G** to **J**) Gene expression overlays shown as in (B).

growth and maximal ribosome content (*54*). However, we found a relative decline in the average rRNA levels: Early lag phase populations (30 min after dilution) had a higher signal than the more advanced lag culture (60 min after dilution; Fig. 2E). These differences appeared to be rooted in the transient emergence and disappearance of an early lag subpopulation with exceptionally high levels of 16*S* rRNA (cluster 13, comprising 34.6% of the population in early lag; Fig. 3, C to F; fig. S3; and table S4). In agreement with the deviation in the rRNA signal, this subpopulation also showed a proportional increase in total mRNA counts. However, its size and chromosome copy distributions were not elevated (fig. S4, cluster 13 versus cluster 3).

Beyond illuminating the extent of heterogeneity in seemingly well-mixed cultures and classifying subpopulations into particular types, seqFISH can also directly connect global cell-specific parameters such as ribosome levels or cell shape to particular gene expression signatures. For example, a closer examination of the metabolically hyperactive subpopulation revealed a 186-fold enrichment in *cdrA* expression relative to the rest of the population (Fig. 3G). The *cdrA* gene encodes a major adhesive protein component of the *P. aeruginosa* biofilm matrix (*55*, *56*). Expression of *cdrA* is commonly used as a reporter for c-di-GMP levels, a key signaling molecule involved in surface attachment (*16*). This subpopulation

also displays a 30-fold enrichment in *pstS* expression, which encodes for the phosphate-binding component of the *pstSCAB* phosphate uptake system (Fig. 3H). PstS has been previously detected in extracellular appendages of *P. aeruginosa* and has been suggested to provide an adhesion phenotype to intestinal epithelial cells (*57*). In support of this non-canonical role, *pstS* was recently suggested to confer a similar adherence phenotype in *Acinetobacter baumannii*, another human pathogenic bacterium (*58*).

A second example from our dataset of the type of fine-grained information that seqFISH can provide comes from the temporal expression of genes involved in virulence factor production. Single-cell variation in virulence factor production has been suggested as a mechanism for division of labor during infection (*17*). *P. aeruginosa* uses a variety of virulence factors to overcome the host immune response (*44*), including the type 3 secretion system (T3SS) that translocates toxins (effectors) directly into host cells (*59*). Our gene set monitors two T3SS structural genes, *pscC* and *pcrD*, and two main effector genes, *exoT* and *exoY*, all of which are encoded in different operons (*60*). We detected two different types of subpopulations with enriched T3SS-related genes, suggesting a unique division of cells into virulent and avirulent states (Fig. 3, I and J). The first group transiently appears during exponential growth and constitutes 8 to 30% of the population (Fig. 3, C to F, I, and J, and table S4). This group expresses both the secretion system genes (86-fold enrichment) and the effector genes (28-fold). By contrast, the second group appears three or four divisions later, close to the replicative minima at stationary phase, and occupies only ~2.7% of cells (table S4). This subpopulation is strongly enriched for the two effector genes (average 26-fold; Fig. 3, I and J) but only mildly so for the secretion system genes (sixfold) compared with the earlier group.

We can potentially reconcile these observations as follows: *P. aeruginosa* has been shown to contain one to three T3SS units per cell under inducing conditions (*61*). Thus, successive divisions after T3SS expression will result in rapid dilution of the T3SS$^+$ group. Assuming that the inheritance of the T3SS and effectors is uncoupled, then T3SS$^+$ stationary phase cells are likely to lose their effectors during division and thus are predicted to be "inactive." An intriguing hypothesis is that *P. aeruginosa* invests in the costly T3SS$^+$ subpopulation during "times of plenty" (rapid growth) and specifically expresses the effectors at stationary to "reload" and maintain this subpopulation after division-based dilution, just before growth arrest. Together, these examples underscore the power of seqFISH to suggest hypotheses that can be tested going forward.

### Spatial transcriptomics at a single-cell resolution in *P. aeruginosa* biofilms

Although much can be learned by applying seqFISH to planktonic cultures, in many contexts, bacteria exist in biofilms (*1*, *2*). Variation in local environmental conditions and the effect of spatially confined metabolic activities in biofilm populations can promote the emergence of chemically distinct microenvironments and phenotypes (*10*, *19*). We reasoned that seqFISH's capacity to record transcriptional activities with micrometer resolution would be particularly useful in shedding light on these processes.

The *P. aeruginosa* biofilm mode of life is particularly important in chronic infections such as those residing in the airways of individuals with cystic fibrosis (*62*, *63*). Accordingly, having used LB medium to validate bacterial seqFISH, we switched to synthetic cystic fibrosis sputum medium (SCFM) for our biofilm studies (*64*). Briefly, bacteria were incubated in coverslip-attached microwells and the medium was replaced every several hours. Using biofilms that were allowed to develop for 10 or 35 hours, we imaged hundreds of aggregates ranging in size from several bacteria to tens of thousands of tightly bound members (Fig. 4, A and B). As a reference for cellular physiological states, we also performed a planktonic growth curve experiment in SCFM. We applied par-seqFISH multiplexing to image 10 time points matching those sampled in the planktonic LB experiment. We found a similar degree of heterogeneity in SCFM- and LB medium–grown populations (Fig. 4D). We extracted the physical coordinates of individual bacterial cells within microaggregates, acquiring a microscale spatial expression profile for ~365,000 surface-attached bacteria (Fig. 4, A and B). In addition, we collected single-cell expression data for ~218,000 planktonic cells.

A basic question that we sought to answer is what is the extent to which transcriptional responses are specific to the biofilm lifestyle? We performed a joint UMAP analysis using both biofilm and planktonic samples (Fig. 4C). These different modes of growth cluster into independent groups in expression space, reflecting their significant physiological differences (Fig. 4D). Ribosome and RNA polymerase subunit expression in the planktonic experiment correlated strongly with growth rate, as observed in LB medium (Fig. 4D). Examining these marker genes in the biofilm-derived cells placed the average replicative capacity of the 10-hour (10h) and 35h biofilm populations approximately equal to those of the early-middle and late-stationary planktonic populations, respectively (Fig. 4F). Expression of the stationary phase master regulator *rpoS* further supported this classification (Fig. 4G). However, biofilm cells also have distinctive expression profiles that distinguish them from liquid cultures. For example, the matrix component gene *cdrA* was uniformly expressed in both the 10h and 35h biofilms but repressed in most planktonic cells (Fig. 4E). In addition, compared with stationary liquid cells, our data indicate that early biofilms (10h) have higher expression of *sigX* (5.1-fold), a transcription factor recently implicated in biofilm formation (*65*); *mexB* (>4.5-fold), of the *mexA-mexB-oprM* antibiotic efflux system; and an increase in the 3′-5′ exonuclease polynucleotide phosphorylase (*pnp*) (7.5-fold). Comparing the 35h biofilm with stationary cells, we found a 3.3-fold increase in the extracellular protease *lasB* but reduced expression of other proteases such *lasA* (3-fold lower), as well as *aprA* and the rhamnolipid biosynthesis gene *rhlA* (~10-fold lower). These genes are QS regulated, and our liquid cultures expressed both *lasA* and *rhlA* at later time points than *lasB*, suggesting that these differences may reflect the age of the biofilm rather than features that define the biofilm state per se.

### In situ analysis of biofilm-specific functions

The above data demonstrate that seqFISH can capture both cell states and their physical position directly within intact biofilms, providing an opportunity to examine known and new processes that contribute to biofilm development from a quantitative and highly spatially resolved perspective. To illustrate this, we focused on the expression patterns of representative genes known to define critical stages in biofilm development, such as attachment, maturation, and exclusion of competitors.

Motility systems such as the flagella and the type 4 pilus (T4P) are a major determinant of surface colonization and subsequent biofilm formation (*66*–*68*). Recent work identified an asymmetric division process coined "touch-seed-and-go," in which flagellated mother cells first attach to a surface and then produce unflagellated daughter cells that contain the T4P. This c-di-GMP–dependent phenotypic diversification enables the mother "spreader" cell to spawn multiple adherent "seed" populations (*69*). This is thought to be mainly regulated by surface sensing (*69*). However, how such motility-based division of labor affects the organization of biofilms at stages beyond surface attachment remains unknown.

We examined the spatial expression patterns of the major flagellum and T4P components, *fliC* and *pilA*, respectively, in the early surface colonization experiment (10h biofilm). An abundant "checkerboard"–like pattern was evident, in which cells expressed high levels of either *fliC* or *pilA* but generally not both (Fig. 5A). We found that the highly expressing *fliC*$^+$ and *pilA*$^+$ subpopulations (>3 SDs above the population mean) represented a total of ~4% of all cells in our experiment (2% for each subgroup), yet the double-positive

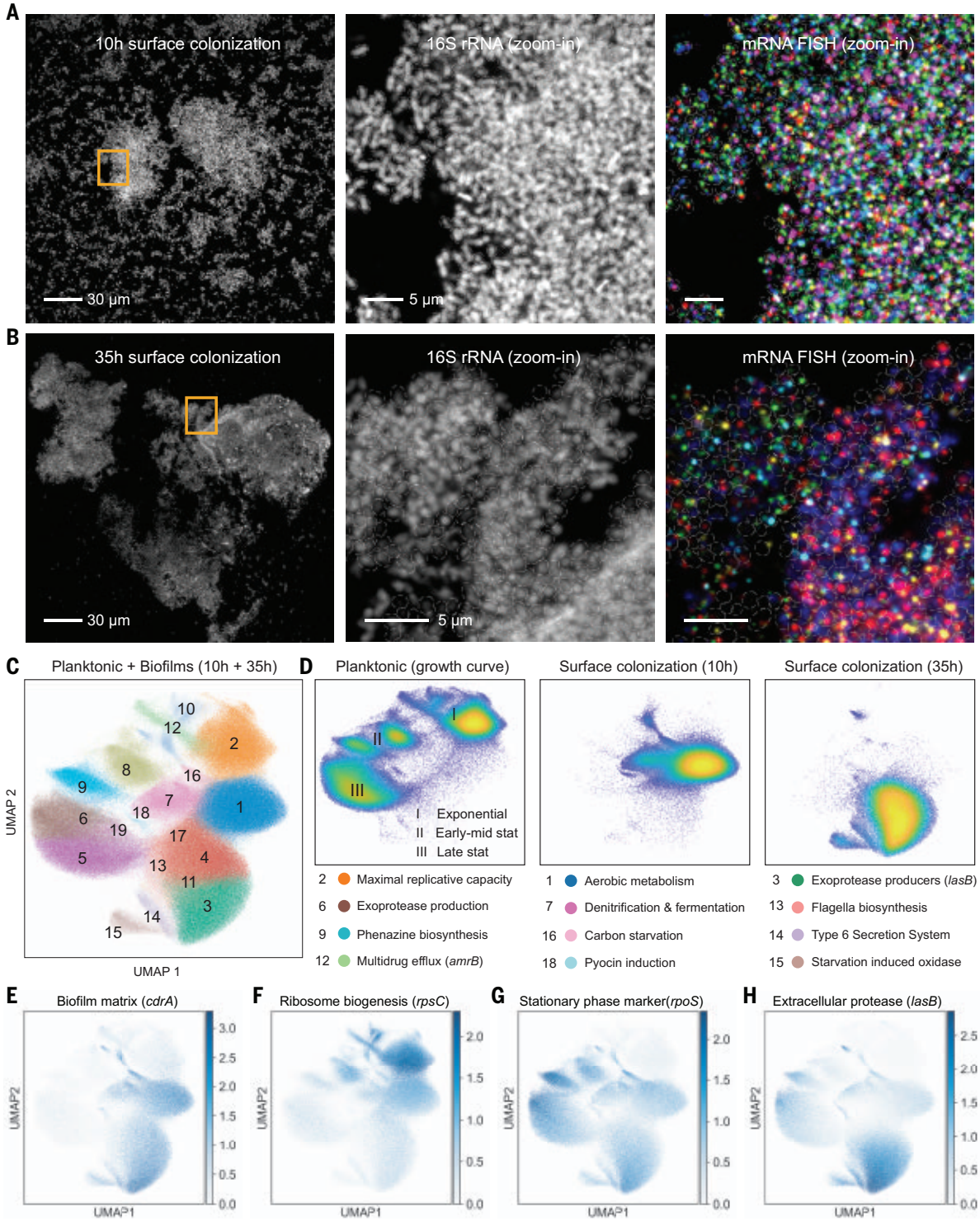



**Fig. 4. Spatial transcriptomics in *P. aeruginosa* biofilms at a single-cell resolution.** (**A**) Representative field of view collected during a 10h surface colonization experiment showing cells using 16*S* rRNA fluorescence (gray). Magnification (orange box) shows the cell segmentation masks depicted as white ellipses. The 16*S* rRNA signal and mRNA-FISH data for several genes are shown in different colors. (**B**) A 35h experiment field is shown in an identical manner to (A). Scale bar length is annotated within the figure. (**C**) Joint UMAP cluster analysis of biofilm and planktonic experiments. Planktonic cells are shown for all time points collected. (**D**) UMAP scatter plots showing cells from either planktonic or biofilm experiments as indicated. At the bottom, a highlighted set of UMAP clusters associated with each experiment is annotated with enriched functions. (**E** to **H**) UMAP overlay with specific gene data. The color map shows the normalized expression scaled to unit variance. Cells from the liquid experiment and both the 10h and 35h biofilms are displayed together.

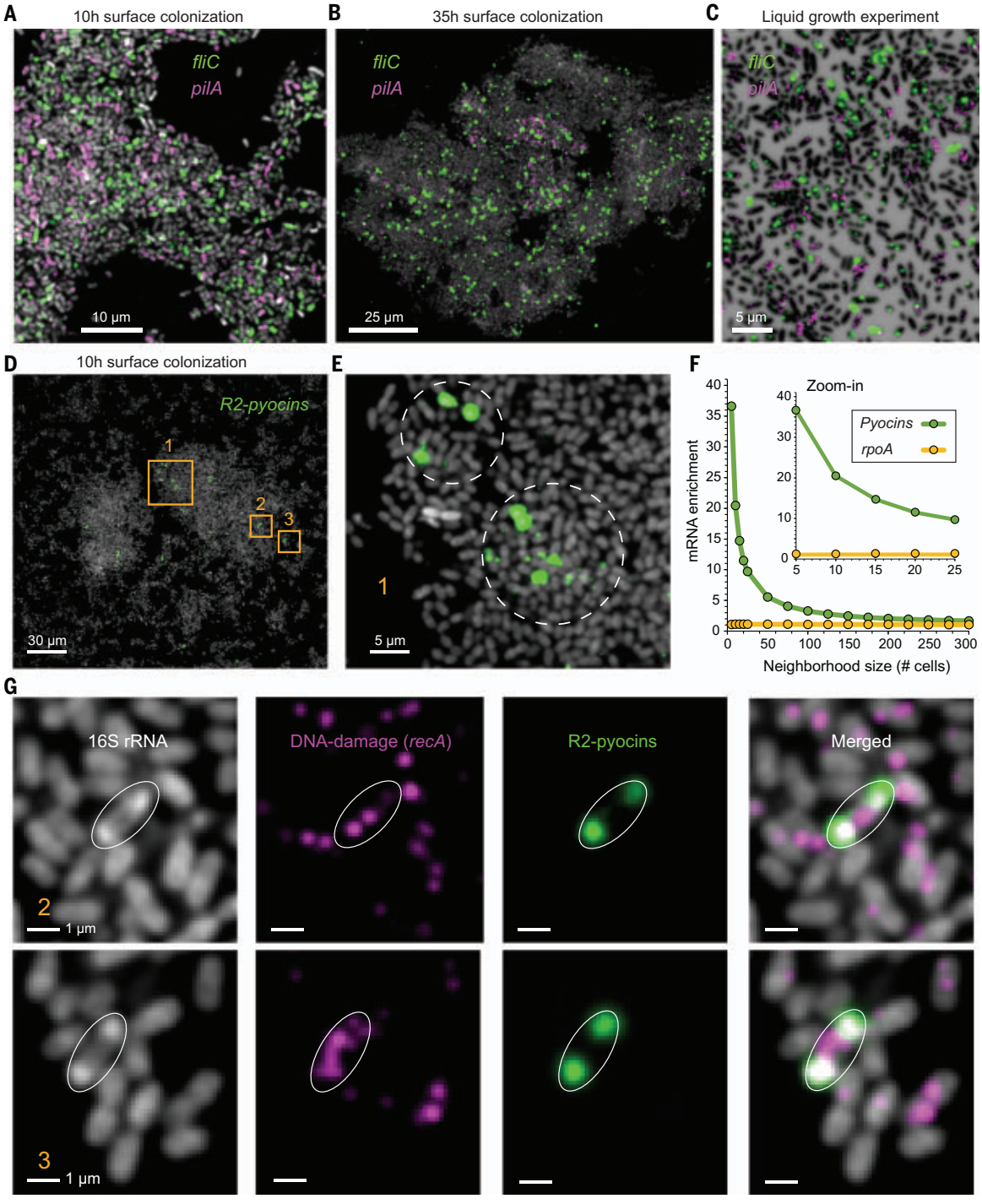



**Fig. 5. Spatial expression patterns for motility- and pyocin-related genes.** (**A** and **B**) Representative regions from the 10h and 35h biofilm experiments. Cells are shown using 16S rRNA fluorescence (gray) and overlaid with raw mRNA-FISH fluorescence for different genes as indicated. (**C**) Planktonic cells from the paired liquid experiments. Cells are shown using DAPI and gene expression as indicated. (**D** and **E**) 10h aggregate showing R2-pyocin expression. (**F**) Enrichment of R2-pyocin mRNA near strong induction sites (cell with 99.5th percentile pyocin expression). The *x*-axis shows the number of cells closest to an induction site that were analyzed (neighborhood size; center cell was excluded); the *y*-axis shows the enrichment in each neighborhood relative to the total population. A non-pyocin control gene, *rpoA*, is shown. (**G**) Examples of mRNA R-pyocin transcript and ribosome polar localization as indicated in the panel legends.

subpopulation (*fliC*$^+$/*pilA*$^+$) only constituted 0.07%. This pattern occurred uniformly across most aggregates, both in small groups (tens of cells) and in large sets containing thousands of cells. Conversely, the older 35h biofilms showed lower expression of *pilA* but contained a sparse but uniform distribution of *fliC*$^+$ cells, suggesting that biofilm-associated bacteria invest in a costly motility apparatus despite being spatially confined (Fig. 5B). Examining the expression of *fliC* and *pilA* in our paired planktonic experiment, we found a similar mutually exclusive pattern (~2% of both single-positive groups and ~0.15% of the double-positive cells) (Fig. 5C). Thus, in contrast to

the current model, our planktonic control experiment suggests that the asymmetric distribution of motility systems is unlikely to be directly regulated by surface sensing (Fig. 5C); such a conclusion would not be possible without the means to compare transcriptional activities at the single-cell level.

Beyond initial surface attachment, bacteria must establish a strong foothold for colony development and also outcompete resident microbes. One strategy that potentially addresses both needs is the use of phage tail–like bacteriocins, which are broadly called tailocins (*70*). These elements are thought to be adapted from prophages and are applied as narrow-spectrum toxins for kin exclusion (*70*, *71*). However, in contrast to antibiotics, these phage tail–like structures are released into the environment through explosive lysis events that kill the producer and spray the toxin locally to inhibit nearby competitors (*72*, *73*). This event also releases extracellular DNA that integrates into the biofilm matrix, structurally supporting biofilm maturation (*72*, *74*). How this "sacrificial" process is regulated within developing biofilms is not well understood.

Our UMAP analysis identified a subpopulation (cluster 18; Fig. 4C) exhibiting >1000-fold enrichment in expression of the R2-pyocin operon (*P. aeruginosa* tailocin), represented by the *PA14_08150* gene. This UMAP cluster was enriched by about fourfold in 10h biofilm–derived cells (0.45% of the entire population), suggesting that pyocin induction is up-regulated during surface attachment. Furthermore, we found an 11-fold higher expression of the DNA-repair gene *recA*, in agreement with its role in inducing pyocin expression (*75*). Visualizing the expression of the pyocin producers, we found that induction events were spread across various microaggregate regions but often appeared in local clusters (Fig. 5, D and E). Indeed, we found a ~37-fold average spatial enrichment in pyocin expression in the immediate vicinity of strong induction sites compared with the general population (Fig. 5F). This enrichment decayed rapidly as a function of neighborhood size, suggesting a highly localized effect (Fig. 5F).

In addition to reporting multigene expression profiles, seqFISH also reports the physical position of measured mRNA molecules at a submicrometer resolution. During this analysis, we observed that R2-pyocin transcript fluorescence generally appeared as two spots. Upon closer examination, we discovered that this mRNA was strongly localized to the two cell poles (Fig. 5G). The 16*S* rRNA fluorescence signal in these pyocin producers showed identical polarization, a rare pattern not observed in neighboring noninducing cells (Fig. 5G). These data suggest that ribosomes and the R2-pyocin transcript are mobilized after induction and spatially colocalize. By contrast, the expression of *recA* did not follow this pattern, suggesting a pyocin-specific effect (Fig. 5G). A recent study discovered an identical polar localization for two different *Pseudomonas protegens* R-tailocins at the protein level (*73*). Together, these data hint at a potentially evolutionary conserved, RNA-dependent mechanism for R-tailocin protein polar localization. We hypothesize that the spatially correlated ribosomal enrichment may provide efficient local translation and particle accumulation before cell lysis.

### Temporal evolution of metabolic heterogeneity during biofilm development.

Beyond resolving transcriptional activities that contribute to biofilm developmental processes, seqFISH can also reveal how biofilm cells metabolically respond to subtle changes in their local microenvironment. Chemical heterogeneity is a key feature of spatially structured environments, and metabolic heterogeneity characterizes mature biofilms (*10*, *18*, *19*, *76*). However, until now, it has been impossible to capture the development of fine-grained metabolic structure across multiple suites of genes at different times.

To map biofilm metabolic development, we focused on genes for which regulation and functions are well understood. In particular, we focused on catabolic genes with products that enable energy conservation under different oxygen concentrations. Oxygen is a central and dynamic factor that influences metabolic activity in bacterial biofilms (*10*, *19*, *77*, *78*). Local oxygen availability can vary significantly within structured environments and is biotically shaped within biofilms (*18*, *77*, *79*). *P. aeruginosa* can survive under anaerobic conditions by fermenting different substrates and/or denitrifying (*50*, *80*, *81*). Accordingly, monitoring the expression of these catabolic genes and others that are co-regulated with them provides a means of tracking local oxygen availability and its dynamic effects on biofilm metabolic coordination.

How quickly and over what spatial scales do biofilm cells metabolically differentiate? Following the *uspL* gene, which was strongly induced during hypoxic conditions and correlated with anaerobic fermentation and denitrification genes in our planktonic growth experiments, we observed unexpectedly heterogeneous responses to oxygen depletion over just a few micrometers in young (10h) biofilms (Fig. 6A). *uspL* expression was strongly spatially correlated with multiple anaerobic markers (fig. S5), indicating that this gene reports on local anaerobic activities. A closer examination of these putative hypoxic sites showed a frequent anticorrelation of *uspL* with multiple genes that were otherwise uniformly expressed in 10h biofilms, appearing as co-localized but reversed expression patches (Fig. 6B). Among the anticorrelated functions were the tricarboxylic acid (TCA) cycle gene *sucC* and replicative capacity genes such as those encoding RNA polymerase and ribosome subunits (Fig. 6B and figs. S4 and S5). However, exceptions to this anticorrelation were also observed (fig. S6).

Can the metabolic heterogeneity revealed by oxygen-responsive marker genes provide an entry point for the discovery of more nuanced cellular responses at the microscale? Our spatial correlation analysis revealed an intriguing association between anaerobic metabolism genes, such as those in the denitrification pathway (*narG-nirS-norB-nosZ*), and the oxidative stress response genes *katA*, *katB*, and *sodM*, which encode for the inducible catalases and an Mn-dependent superoxide dismutase, respectively (*82*–*84*) (Fig. 6, C and D, and figs. S4 and S5). Nitrite-respiring *P. aeruginosa* produce the highly toxic intermediate nitric oxide (NO) (*85*). Indeed, KatA was recently demonstrated to play a role in protection from NO-associated stress (*84*), suggesting that these subaggregate regions correspond to microenvironments with high NO levels. In agreement with this hypothesis, we found that the stress response pattern was also spatially correlated with heat-shock protease expression, including the membrane protease *ftsH*, which was found to play an important role in survival under anoxic conditions (*86*) (Fig. 6E and fig. S4). These data highlight how contrasting physiological states can be established just a few micrometers away early in biofilm development.

We hypothesize that these coordinated expression patterns for particular genes reflected the spatiometabolic distribution of distinct physiological "states" across the biofilm. To test this hypothesis, we conducted a targeted UMAP analysis using only the 10h biofilm cells (fig. S7). We identified two main anaerobic subpopulations corresponding to denitrification- and fermentation-dominated metabolic states and representing 11.8 and 7.2% of all cells in the experiment, respectively (fig. S7). In addition, we detected a smaller subpopulation of denitrifying cells (2.4% of cells) with a 5.3-fold average increase in the oxidative stress factors *katB*, *sodM*, and *ahpF*, the latter of which encodes for an alkyl hydroperoxide reductase (*87*). Relative to the main denitrifying subgroup, stressed cells had lower expression of the denitrification pathway (about fourfold) and a more than twofold reduction in replicative capacity marker levels (*rpoA*, *rpsC*, and *atpA*), in support of a potentially damaged state. Projecting these single-cell metabolic states over their respective biofilm positions showed a strong overlap with the above predicted hypoxic pockets, supporting our hypothesis and revealing that multiple metabolic states can coexist in the same patch (Fig. 6F and fig. S5).

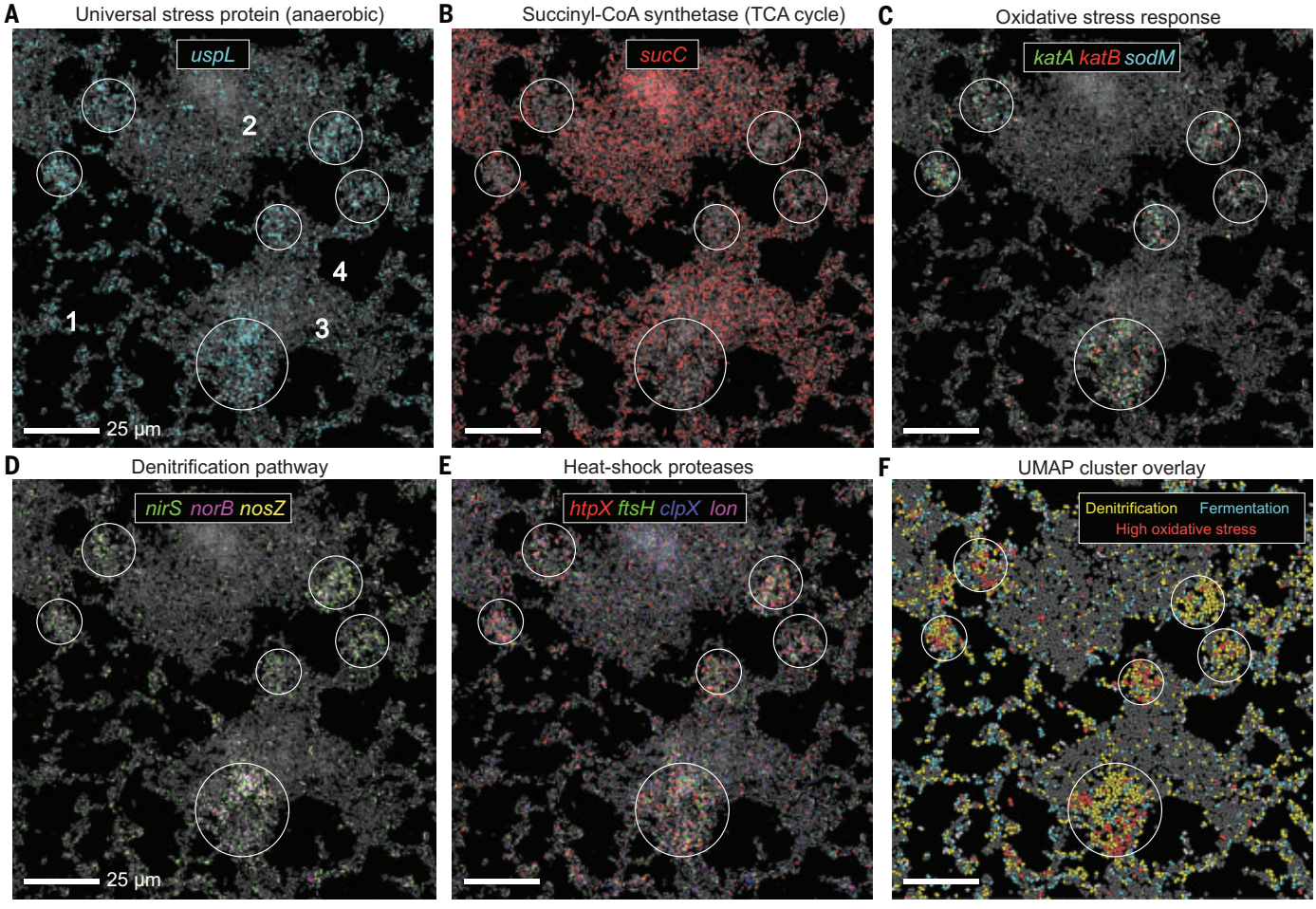



**Fig. 6. Oxygen availability shapes microscale metabolic heterogeneity in biofilms.** (**A** to **E**) Representative 10h biofilms. Cells are shown using 16*S* rRNA FISH fluorescence (gray) and overlaid with raw mRNA-FISH fluorescence for different genes as indicated in each panel. White circles highlight regions of interest. (**F**) Cells painted according to their UMAP-derived metabolic state as indicated in the panel legends (also see fig. S7, clusters 0, 8, 12, and 15), showing colocalization of multiple metabolic states within a given region.

Given the extent of transcriptional heterogeneity manifest in young biofilms, we wondered whether such heterogeneity would persist as the biofilms aged. We speculated that the higher cell densities and more committed spatial structuring of mature biofilms might favor larger-scale metabolic zonation. We therefore examined the spatial expression patterns in a 35h biofilm experiment.

In contrast to the spatial variation in aerobic and anaerobic metabolic processes seen in 10h biofilms, 35h biofilms had an ~50-fold lower average expression of the denitrification pathway genes *nar-nirs-norB-nosZ*. Indeed, these genes are known to be repressed by the *las* and *rhl* QS systems, indicating *P. aeruginosa* is programmed to shut down denitrification at high cell densities (*80*, *88*). However, in addition to this complete and co-regulated pathway, *P. aeruginosa* also encodes an independent periplasmic nitrate reductase (*nap*) (*89*). Unexpectedly, the *napA* gene was expressed in a spatially uniform manner but at a low level in the 35h aggregates, a pattern that was closely shared with the *uspL* gene, and these two genes together were expressed in 20.3% (±5.5%) of the measured cells within each individual aggregate (Fig. 7A and fig. S7). NapA has been implicated in maintaining redox homeostasis under oxygen limitation (*78*), and the *uspL* paralog *uspK* was shown to play a role in survival under such conditions (*86*, *90*). At first, these results seemed to suggest that as an aggregate cell mass grows, survival physiology on average dominates over growth-promoting processes. However, we also found substantial and large-scale spatial heterogeneity in certain genes, such as those encoding the replicative capacity markers, which were highly expressed in 17.7% (±10.9%) of aggregate cells (Fig. 7B and fig. S7), and *lasB*, which encodes a QS-regulated extracellular protease and is expressed at similarly high levels in 43.5% (±6.1%) of the cells (Fig. 7C and fig. S7). These data suggest that a single 35h microaggregate can contain regions with distinct physiological states and virulence-related activities. Finally, the fact that metabolism dynamically shapes the microenvironment leads to the prediction that differences in local nutrient availability will be reflected in heterogeneous transcriptional activities over small spatial scales (*10*). We saw evidence of this phenomenon in our data when focusing, for example, on carbon metabolism. Where replicative capacity appeared to be high and carbon was presumably replete, we saw coexpression of the TCA cycle gene *sucC* (Fig. 7, B to D). However, when carbon is limiting, bacteria can use the glyoxylate shunt (GS), which bypasses the oxidative decarboxylation steps of the TCA. The GS provides an alternative metabolic pathway for using acetate and fatty acids as carbon sources (*91*, *92*). In the GS, carbon flux is redirected by isocitrate lyase, which competes with the TCA enzyme isocitrate dehydrogenase for isocitrate. Because isocitrate dehydrogenase has a much lower Michaelis constant ($K_m$), it must be enzymatically inactivated

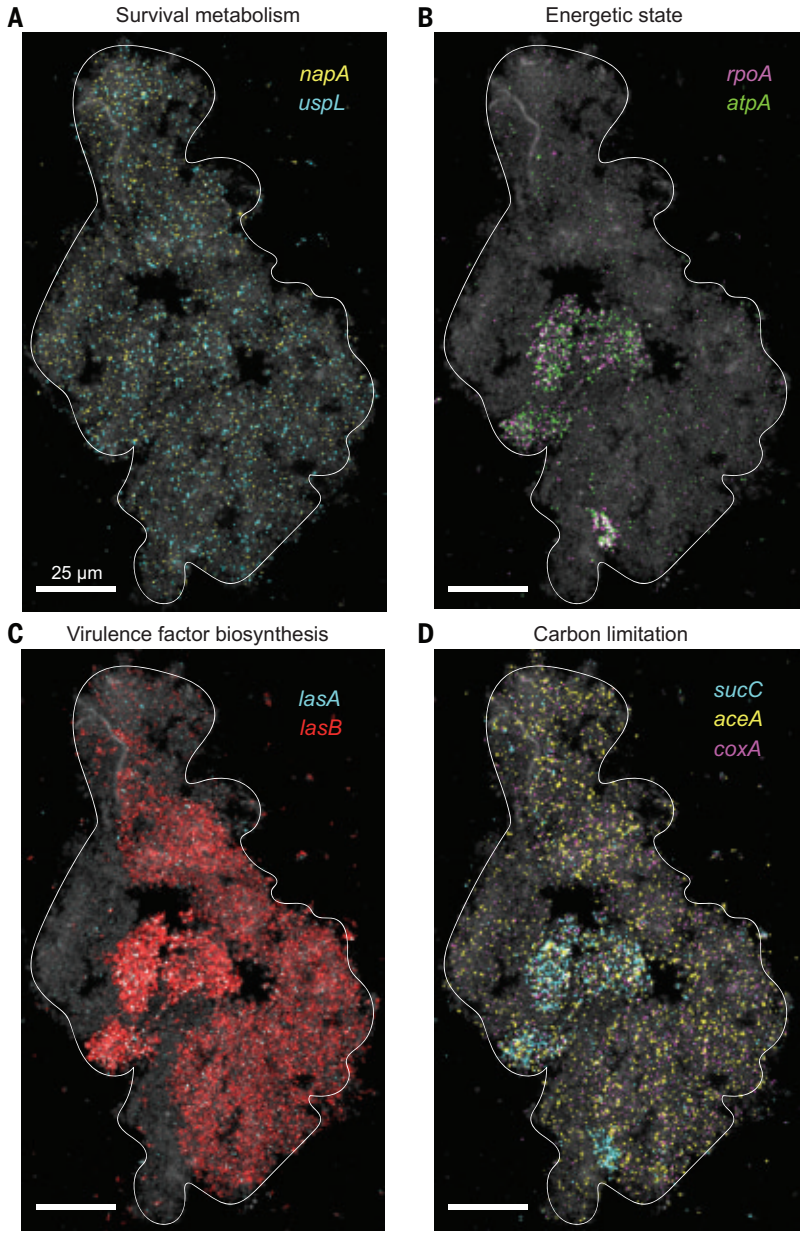



**Fig. 7. Functional zonation in a single microaggregate.** A *P. aeruginosa* 35h aggregate. Bacteria are shown using 16S rRNA FISH fluorescence (gray) and are overlaid with raw mRNA-FISH fluorescence for different genes as described in the panel legends.

by phosphorylation for the carbon flux to be redirected to the GS (*92*). However, little is still known about the transcriptional regulation of these pathways (*93*). Our gene set contains both the GS gene *aceA* and the downstream TCA cycle gene *sucC*. Although these genes are often coexpressed, we found that only the GS marker *aceA* was expressed in the predicted lower-energetic-capacity biofilm zones (Fig. 7D and fig. S7), suggesting that these subregions experience carbon limitation. In support of this hypothesis, these regions also expressed the tightly regulated terminal oxidase gene *coxA*, which is transcriptionally induced by carbon starvation, a condition in which it promotes survival (*86*, *94*) (Fig. 7D and fig. S7). Together, *aceA*- and *coxA*-expressing cells covered up to 43% of an aggregate cell mass in our experiment. This is just one example of the type of coherent spatiometabolic stratification pattern that seqFISH can reveal at any given moment in time.

## Discussion

Until now, our ability to capture the dynamic metabolic activities of microbial populations and communities at small spatial scales has been limited to tracking just a few parameters. This technical limitation has restricted our ability to observe and understand the features that define these ubiquitous associations. Our analysis of *P. aeruginosa* populations has shown that par-seqFISH can reveal a high degree of transcriptional heterogeneity spanning multiple dimensions, from the subcellular to the microscale. Moreover, by tracking the temporal and spatial dynamics of cellular states in subpopulations, our results demonstrate that spatial transcriptomics can provide new insights into how bacteria sustain functional diversity.

The high temporal and spatial resolution enabled by par-seqFISH permitted us to make unexpected discoveries. For example, in planktonic cultures, we observed the short-lived temporal emergence of two T3SS$^+$ populations: a large group, which appeared in the exponential phase and expressed all of the needed T3SS components, and a second, ~10-fold smaller group, which emerged during the midstationary phase and expressed the effectors but not the secretion system. The estimated three or four divisions that separated these subpopulation correlated with their size differences, suggesting that these two subpopulations could represent the same T3SS$^+$ population, just at different stages of growth. We hypothesize that the specific expression of effector genes in the stationary T3SS$^+$ subpopulation serves to replenish the effectors lost by the diluting effect of cell divisions. If true, then this would mean that *P. aeruginosa* not only generates heterogeneous subpopulations but can also actively maintain their functional capabilities. Such an observation would not have been possible without the ability to measure the expression of many genes within the same cell.

In *P. aeruginosa* biofilms, despite marked levels of metabolic heterogeneity, coherent coexpression patterns also emerged. We found a strong spatial correlation between denitrification genes and oxidative stress factors, suggesting that local denitrification results in NO toxicity. This hypothesis is based on the expression of the inducible peroxidase *katA*, which is known to be up-regulated by NO under anaerobic conditions and to alleviate NO toxicity (*84*). We also observed overlapping induction of other factors such as *katB* (*83*) and the superoxide dismutase *sodM* (*82*), suggesting that they may also play protective roles. These patterns were highly spatially confined, suggesting that NO toxicity did not propagate to neighboring cells, even those just a few micrometers beyond. However, it remains unclear how such hydrogen peroxide– and superoxide-detoxifying enzymes protect cells from NO. It is known that NO interacts with relevant oxidants to produce reactive nitrogen species such as peroxynitrite (*95*). Therefore, perhaps these oxidative stress–response factors act by limiting the pool of oxidants available for reactive

nitrogen species production. Various reactive nitrogen species cause diverse types of cellular damage, including the chemical modification of proteins, specifically cysteine and tyrosine residues (*95*). Our data point toward elevated expression of cellular proteases in NO-stressed regions. We therefore suggest that these proteases act to detoxify cells by eliminating damaged proteins, a hypothesis that remains to be tested in future studies.

The par-seqFISH multiplexing approach, which we developed to increase the throughput of seqFISH for single-cell analysis, could be applied in other ways and in both synthetic and natural communities. For example, because par-seqFISH is based on 16*S* rRNA labels (Ribo-Tags), it could in principle be used to encode bacterial taxonomy. Recently, a conceptually similar and exciting method for combinatorial labeling of taxonomy was introduced in a biogeographical study of the human microbiome (*25*). In principle, the par-seqFISH strategy could be readily extended to capture a similar or higher level of taxonomic complexity and add the currently missing feature of mRNA expression. A critical next step will be to develop methods to chart the environmental conditions that contextualize expression patterns observed in any given case. Extension of this approach to natural and clinical samples could provide important insights into the conditions experienced by microbes in more complex environments and the coordinated physiological responses that emerge in turn.

## Materials and methods

### *Bacterial strains and growth conditions*

*P. aeruginosa* strain UCBPP-PA14 was grown aerobically with shaking at 250 rpm in LB medium (Difco) or on LB agar plates at 37°C. SCFM was made as previously described (*64*). For the growth curve experiments, an overnight LB culture was washed twice using fresh growth medium (either LB or SCFM) and then diluted 1:100 into 100 ml of prewarmed fresh medium. The cultures were grown at 37°C with shaking at 250 rpm and collected at various time points, as indicated in Fig. 2A. The SCFM samples were collected at cell densities identical to those in the LB experiment except that the optical density at 600 nm ($OD_{600}$) = 3.2 sam4ple was omitted. Collected samples were immediately fixed in ice-cold 2% paraformaldehyde (PFA), incubated on ice for 1.5 hours in the dark, and then washed twice with 1× phosphate-buffered saline (PBS). Samples were resuspended in 70% EtOH and incubated at −20°C for 24 hours to permeabilize the cells. Surface colonization was performed by washing and diluting an LB overnight culture 1:100 into fresh SCFM and dispensing 100 μl into coverslip-attached open incubation chambers (Electron Microscopy Sciences, #70333-42). The coverslips were incubated in Parafilm-sealed sterile petri dishes at 37°C, and the medium was gently exchanged every 4 hours. A damp Kimwipe was placed in the petri dish to control medium evaporation. During the overnight stage of the 35h experiment, the medium was exchanged only once after 8 hours. Biofilm experiments were collected by gently exchanging the SCFM with 100 μl of ice-cold 2% PFA solution and incubating the sample at 4°C for 1.5 hours. The samples were washed twice with 1× PBS, resuspended in 70% EtOH, incubated overnight at 4°C, and prepared for seqFISH the following day as described below.

### *seqFISH probe design and library generation*

Primary probes were designed as 30-nucleotide (nt) stretches in a GC range of 45 to 65%. Probe sequences containing more than four consecutive base repeats were removed. The remaining probes were compared with the reference genome using BLAST, and any probe with nonspecific binding of at least 18 nucleotides was discarded. Negative control genes were selected from the P1 phage genome (NC_005856.1) using the same criteria. Each selected gene was covered by 12 to 20 nonoverlapping probes randomly selected from the gene probe set. The probes were designed as a 30-nt mRNA-binding region flanked by overhangs composed of four repeats of the secondary hybridization sequence (complementary to a designated fluorescent readout probe; table S2). Thus, it is estimated that during secondary hybridization, each mRNA was covered by 48 to 80 fluorescent readout probes (i.e., 12 to 20 × 4), consistent with previous mRNA-FISH experiments in bacteria (*33*, *42*).

A library of 1763 probes targeting 105 *P. aeruginosa* genes and three negative controls was designed (tables S1 and S2). Additional flanking sequences were added to the primary probe sequences to enable library amplification by polymerase chain reaction (PCR) (forward 5′- TTTCGTCCGCGAGTGACCAG-3′ and reverse 5′-CAACGTCCATGTCGGGATGC-3′). The primary probe set was purchased as oligoarray complex pool from Twist Bioscience and constructed as previously described (*36*) (table S2). Briefly, a set of nine PCR cycles was used to amplify the designated probe sequences from the oligo pool. The amplified PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, #28104) according to the manufacturer's instructions. The PCR products were used as the template for in vitro transcription (New England Biolabs, #E2040S), followed by reverse transcription (Thermo Fisher Scientific, #EP7051). Then, the single-stranded DNA probes were alkaline hydrolyzed with 1 M NaOH at 65°C for 15 min to degrade the RNA templates, followed by 1 M acetic acid neutralization. Next, to clean up the probes, ethanol precipitation was performed to remove stray nucleotides, phenol–chloroform extraction was performed to remove protein, and Zeba Spin Desalting Columns (7 K molecular weight cutoff) (Thermo Fisher Scientific, #89882) were used to remove residual nucleotides and phenol contaminants. Readout probes were designed as previously described and ordered from Integrated DNA Technologies (*36*).

Ribo-Tag probes were designed to target the same region in the 16*S* rRNA gene according to the criteria described above, but with 28-nt binding regions. Each probe sequence was flanked with two secondary sequences selected out a set of six that were dedicated to multiplexing (table S3). An additional 16*S* rRNA probe was generated as a standard between all multiplexed samples and was hybridized to an independent region of the 16*S* rRNA (table S3). This probe provided an additional reference and was used to register images from different channels (see below).

### *Coverslip functionalization*

Coverslips were cleaned with a plasma cleaner on a high setting (Harrick Plasma, #PDC-001) for 5 min, followed by immersion in 1% bind-silane solution (GE, #17-1330-01) made in pH 3.5 10% (v/v) acidic ethanol solution for 30 min at room temperature. The coverslips were washed with 100% ethanol three times and dried in an oven at >90 °C for 30 min. The coverslips were then treated with 100 μg $\mu l^{-1}$ poly-D-lysine (Sigma-Aldrich, #P6407) in water for at least 1 hour at room temperature, followed by three rinses with water. Coverslips were air-dried and kept at −20°C for no longer than 2 weeks before use.

### *par-seqFISH*

Independent fixed samples were individually hybridized with 16*S* rRNA labels, washed, and then pooled into a single mixture that was hybridized with the gene probe library and prepared for imaging. Approximately $10^8$ cells were collected from each sample into a microcentrifuge, pelleted by centrifugation (6000 rpm), and then resuspended in 20 μl of water with 6 nM of the designated 16*S* rRNA label (sample specific) and another 6 nM of a shared reference 16*S* rRNA probe (table S3). Each sample was then mixed with 30 μl of prewarmed primary hybridization buffer [50% formamide, 10% dextran sulfate, and 2× saline-sodium citrate (SSC)] by gentle pipetting, incubated at 37°C for >16 hours, washed twice with 100 μl of wash buffer (55% formamide and 0.1% Triton X-100 in 2× SSC; 5 min at 8000 rpm for the viscous hybridization buffer), and then incubated at 37°C in 100 μl of wash buffer for 30 min to remove nonspecific probe binding. Samples were then washed twice with 100 μl of 2× SSC and pooled together into a new microcentrifuge

in equal volumes. The mixture was pelleted and resuspended in 40 µl of water, and 10 µl of the mixture was added to 10 µl of the gene probe library mixture and mixed well with 30 µl of prewarmed primary hybridization buffer. The hybridizations were incubated for >16 hours at 37°C and then washed and prepared as described above. The final mixture was resuspended in 20 to 25 µl of 1× PBS, and 5 to 10 µl was gently spotted at the center of the coverslip and incubated at room temperature for 10 min to allow the cells to sediment and bind the surface. The coverslips were centrifuged for 5 min at 1000 rpm to create a smooth, dense cell monolayer. The cells were immobilized using a hydrogel as previously described (*36*) and stained with 10 µl $ml^{-1}$ DAPI (Sigma-Aldrich, #D8417) for 5 min before imaging so that cells could be visualized.

In biofilm experiments, the fixed and permeabilized surface-attached microaggregates were air dried, covered with a hydrogel, and hybridized with the gene library and rRNA probes in one single reaction, as described above.

### seqFISH imaging

All seqFISH experiments were performed using a combined imaging and automated fluidics delivery system as previously described (*36*). DAPI-stained samples mounted on coverslips were connected to the fluidic system. The regions of interest were registered using the DAPI fluorescence, and a set of sequential secondary hybridizations, washes, and imaging was performed.

Each hybridization round contained three unique 15-nt readouts probes, each conjugated to Alexa Fluor 647 (A647), Cy3B, or Alexa Fluor 488 (A488). All readout probes were ordered from Integrated DNA Technologies and prepared as 500 nM stock solutions. Each serial probe mixture was prepared in EC buffer [10% ethylene carbonate (Sigma-Aldrich, #E26258), 10% dextran sulfate (Sigma-Aldrich, #D4911), and 4× SSC]. Hybridizations were incubated with the sample for 20 min to allow for secondary probe binding. The samples were then washed to remove excess readout probes and to limit nonspecific binding using ~300 µl of 10% formamide wash buffer (10% formamide and 0.1% Triton X-100 in 2× SSC). Samples were then rinsed with ~200 µl of 4× SSC and stained with DAPI solution (10 µg $ml^{-1}$ DAPI and 4× SSC). Last, an antibleaching buffer solution [10% (w/v) glucose, 1:100 diluted catalase (Sigma-Aldrich, #C3155), 0.5 mg $ml^{-1}$ glucose oxidase (Sigma-Aldrich, #G2133), and 50 mM, pH 8 Tris-HCl in 4× SSC] was flowed through the samples. Imaging was performed with a Leica DMi8 microscope equipped with a confocal scanner unit (Yokogawa, #CSU-W1), a sCMOS camera (Andor Zyla 4.2 Plus), a 63× oil objective lens (Leica, 1.40 numerical aperture), and a motorized stage (ASI, #MS2000). Lasers from CNI and filter sets from Semrock were used. Snapshots were acquired using 647-, 561-, 488-, and 405-nm fluorescent channels with 0.5-µm z-steps for all experiments, with the exception of the 35h biofilm experiment, in which 1.0-µm z-steps were collected. After imaging, readout probes were stripped using 55% wash buffer (55% formamide, 0.1% Triton-X 100, and 2× SSC) that was flowed through for 1 min, followed by incubation for 15 min before rinsing with 4× SSC solution. For this protocol, serial hybridizations, imaging, and signal quenching steps were repeated for ~40 rounds to capture 16*S* rRNA for multiplexing, mRNA expression, and background signal. The integration of the automated fluidics delivery system and imaging was controlled using µManager (*96*).

### Image analysis demultiplexing and gene expression measurement

Maximal projection images were generated using ImageJ (*97*) for DAPI and 16*S* rRNA, and hybridization rounds were registered using DAPI fluorescence. Aberrations between fluorophores were corrected by alignment of 16*S* rRNA signals across all channels. Cells were segmented using the DAPI signal with SuperSegger using the 60XPa configuration (*98*) and filtered using custom scripts to eliminate odd shapes or autofluorescent or low-signal components.

For par-seqFISH demultiplexing, the background (no readouts) and 16*S* rRNA fluorescence intensity for each relevant secondary readout probe was measured within segmented cell boundaries to provide a signal-to-background score for each readout. The cells were classified according to the positive readout combinations (table S3). The number of false-positives was estimated by counting the number of cells classified into combinations left out of the experiment.

The mRNA-FISH data were analyzed using Spätzcells (*42*). Briefly, spots were detected as regional maxima with intensity greater than a threshold value that was set using the negative control genes and fit with a two-dimensional (2D) Gaussian model. The integrated intensity of the spot and the position of its estimated maxima were determined (*42*). Spots were assigned to cells using cell segmentation masks (*42*). In biofilm experiments, spots were assigned to cells in a z-section–sensitive manner. Deviating spot maxima positions that did not overlap a cell boundary were tested against the flanking z-sections to identify their cell of origin. If no cell was detected, then the spots were discarded. All predicted low-expression genes (defined as genes with spots in <30% of all cells) were identified, and the distribution of their spot intensities was fit with a Gaussian mixture model to identify the characteristic intensity of a single mRNA normalized to the number of probes used for the specific gene. The median characteristic single-mRNA signal was then calculated using all low-expression genes for each fluorophore (A647, A488, and cy3B). The variation between different genes labeled with the same fluorophore was low, with a coefficient of variation of 18 to 21%. This median characteristic value was used to transform fluorescence intensity into discrete mRNA counts per gene within each cell. The A488 characteristic signal was corrected by a factor of 1.5 to account for its lower intensity in our system. In each cell, the total intensity of each gene was calculated by summing the intensities of all spots. The total value was normalized by the characteristic value for a single mRNA in the corresponding fluorophore.

### Single-cell expression analysis and cell biological parameter calculations

Single-cell UMAP analysis was performed using Scanpy v1.7.0 (*99*). Genes detected at consistently low levels were excluded from the analysis. These included *pilY1*, *flgK*, *nasA*, *algU*, *purF*, *phzH*, *phzS*, and *pslG* (table S1). The standard Scanpy normalization and scaling, dimensionality reduction, and clustering as described in the Scanpy tutorial were followed, minus the high-variance gene selection and without a library size normalization. Fifteen neighbors and 15 and 17 PCA components were used for the LB and merged SCFM analyses, respectively. Clustering was performed using the Leiden method. Jupyter notebooks with the chosen parameters, run lines, output files, and source data are available at Zenodo (see the Acknowledgments).

Cell nucleoid size was calculated using the segmentation mask. A chromosome score was calculated as the median DAPI intensity multiplied by the nucleoid size. The median chromosome score was calculated for the last time point in our LB experiment (deep stationary; $OD_{600}$ = 3.2). Because most cells in this stage are in a nondividing state, we set this value as a reference for a single chromosome copy. We then normalized the scores of all cells in the experiment using this value, as seen in Fig. 2. In addition to using Ribo-Tags to label cells from different conditions, we also hybridized another region in the 16*S* rRNA with a probe that was shared across all samples (table S3; described above). We used this reference signal to compare the 16*S* rRNA intensity between cells from different conditions. We measured the median 16*S* rRNA signal per cells and multiplied it by the nucleoid size (which completely overlaps the 16*S* signal and estimates cell size). In *E. coli*, maximal ribosome numbers appear at the maximal growth rate and have been estimated at 72,000 (*100*). The median rRNA score was calculated for the maximal growth ($OD_{600}$ = 0.2) and normalized to 72,000 as in *E. coli* for a rough estimate (Fig. 2).

### *Image analysis in surface colonization experiments*

Images were registered as described above, and segmentation was performed using the pixel classification workflow in Ilastik (*101*). The Ilastik classification model was trained with background, cell boundaries, and cell bodies using the 16*S* rRNA signal and it performed extremely well. However, in high-density regions, segmentation often resulted in overconnectivity because of incorrect 3D overlaps and such cell clusters were disconnected. The binary masks (segmentation output) were thinned, and all 3D connected components (CCs) were recalculated, which reduced spurious connections. Then, all CCs traversing >2.5 μm were set aside for reevaluation for potential overconnections. For each such 3D component, each *z*-slice was examined individually and all 2D CCs were identified. Overly large or curved blobs, which represent segmentation artifacts that often incorrectly connect distinct cells across z-sections, were removed. In addition, orientation and overlap with components in the previous flanking z-section were calculated for each 2D detected component. If this component exhibited a significant change in its orientation (i.e., the direction it was pointing), it was disconnected from the component below. The analysis was then continued using the newly oriented component as a seed. Cell clusters that could not be properly disentangled were removed from the analysis. At the end of the analysis, the cell 3D masks were re-thickened.

Bulk neighborhood analysis was used to determine the immediate neighborhoods associated with high expression of a specific gene. For the gene of interest, the top 99th percentile of cells (99.5th for the pyocin-specific analysis), denoted as "center cells," were identified. Using the 3D centroid coordinates of center cells, their closest neighbors were identified within a specified distance (up to 10 μm for pyocins and 3 μm for the rest). Then, up to $k$ closest cells (five to 300 neighbors in the pyocin analysis to view the enrichment decay and up to five for the rest of the genes) were collected. All of the neighborhood cells selected (not including the center cells) were then analyzed in bulk together and their mean gene expression was calculated and compared with the population (minus all center cells not used). This analysis was conducted across all genes and a Pearson correlation analysis was performed to identify spatially correlating genes (fig. S5).

### *RNA-seq analysis*

Previously published RNA-seq datasets (*51*) in which wild-type *P. aeruginosa* PA14 was grown in LB medium to $OD_{600}$ values similar to those collected in our experiment ($OD_{600}$ = 1.1 and 2.0 in the RNA-seq experiment; Fig. 2A) were analyzed. The raw sequencing data were aligned to the *P. aeruginosa* reference genome using bowtie2 (*102*), and the reads were assigned to genes using featureCounts (*103*). Reads per kilobase per million values were calculated for all genes and then compared with the average seqFISH expression profile of each LB-grown sample (fig. S2).

## REFERENCES AND NOTES


1. H.-C. Flemming, S. Wuertz, Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247–260 (2019). doi: 10.1038/s41579-019-0158-9; pmid: 30760902
2. J. W. Costerton *et al*., Bacterial biofilms in nature and disease. *Annu. Rev. Microbiol.* **41**, 435–464 (1987). doi: 10.1146/annurev.mi.41.100187.002251; pmid: 3318676
3. W. H. DePas *et al*., Exposing the three-dimensional biogeography and metabolic states of pathogens in cystic fibrosis sputum via hydrogel embedding, clearing, and rRNA labeling. *mBio* **7**, e00796 (2016). doi: 10.1128/mBio.00796-16; pmid: 27677788
4. J. L. Mark Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst, G. G. Borisy, Biogeography of a human oral microbiome at the micron scale. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E791–E800 (2016). doi: 10.1073/pnas.1522149113; pmid: 26811460
5. J. A. Schaber *et al*., *Pseudomonas aeruginosa* forms biofilms in acute infection independent of cell-to-cell signaling. *Infect. Immun.* **75**, 3715–3721 (2007). doi: 10.1128/IAI.00586-07; pmid: 17562773
6. P. Jorth, M. A. Spero, J. Livingston, D. K. Newman, Quantitative visualization of gene expression in mucoid and nonmucoid *Pseudomonas aeruginosa* aggregates reveals localized peak expression of alginate in the hypoxic zone. *mBio* **10**, e02622 (2019). doi: 10.1128/mBio.02622-19; pmid: 31848278
7. R. Hatzenpichler, V. Krukenberg, R. L. Spietz, Z. J. Jay, Next-generation physiology approaches to study microbiome function at single cell level. *Nat. Rev. Microbiol.* **18**, 241–256 (2020). doi: 10.1038/s41579-020-0323-1; pmid: 32055027
8. G. L. Chadwick, F. Jiménez Otero, J. A. Gralnick, D. R. Bond, V. J. Orphan, NanoSIMS imaging reveals metabolic stratification within current-producing biofilms. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20716–20724 (2019). doi: 10.1073/pnas.1912498116; pmid: 31548422
9. M. Ackermann, A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13**, 497–508 (2015). doi: 10.1038/nrmicro3491; pmid: 26145732
10. C. R. Evans, C. P. Kempes, A. Price-Whelan, L. E. P. Dietrich, Metabolic heterogeneity and cross-feeding in bacterial multicellular systems. *Trends Microbiol.* **28**, 732–743 (2020). doi: 10.1016/j.tim.2020.03.008; pmid: 32781027
11. F. Schreiber *et al*., Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. *Nat. Microbiol.* **1**, 16055 (2016). doi: 10.1038/nmicrobiol.2016.55; pmid: 27572840
12. F. Schreiber, M. Ackermann, Environmental drivers of metabolic heterogeneity in clonal microbial populations. *Curr. Opin. Biotechnol.* **62**, 202–211 (2020). doi: 10.1016/j.copbio.2019.11.018; pmid: 31874388
13. S. H. Kopf *et al*., Heavy water and (15) N labelling with NanoSIMS analysis reveals growth rate-dependent metabolic heterogeneity in chemostats. *Environ. Microbiol.* **17**, 2542–2556 (2015). doi: 10.1111/1462-2920.12752; pmid: 25655651
14. J.-W. Veening, W. K. Smits, O. P. Kuipers, Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* **62**, 193–210 (2008). doi: 10.1146/annurev.micro.62.081307.163002; pmid: 18537474
15. A. Z. Rosenthal *et al*., Metabolic interactions between dynamic bacterial subpopulations. *eLife* **7**, e33099 (2018). doi: 10.7554/eLife.33099; pmid: 29809139
16. C. R. Armbruster *et al*., Heterogeneity in surface sensing suggests a division of labor in *Pseudomonas aeruginosa* populations. *eLife* **8**, e45084 (2019). doi: 10.7554/eLife.45084; pmid: 31180327
17. M. Diard *et al*., Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature* **494**, 353–356 (2013). doi: 10.1038/nature11913; pmid: 23426324
18. P. S. Stewart, M. J. Franklin, Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* **6**, 199–210 (2008). doi: 10.1038/nrmicro1838; pmid: 18264116
19. P. S. Stewart, Diffusion in biofilms. *J. Bacteriol.* **185**, 1485–1491 (2003). doi: 10.1128/JB.185.5.1485-1491.2003; pmid: 12591863
20. E. Wolfsberg, C. P. Long, M. R. Antoniewicz, Metabolism in dense microbial colonies: $^{13}C$ metabolic flux analysis of E. coli grown on agar identifies two distinct cell populations with acetate cross-feeding. *Metab. Eng.* **49**, 242–247 (2018). doi: 10.1016/j.ymben.2018.08.013; pmid: 30179665
21. O. Kotte, B. Volkmer, J. L. Radzikowski, M. Heinemann, Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. *Mol. Syst. Biol.* **10**, 736 (2014). doi: 10.15252/msb.20135022; pmid: 24987115
22. F. Bocci, Y. Suzuki, M. Lu, J. N. Onuchic, Role of metabolic spatiotemporal dynamics in regulating biofilm colony expansion. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4288–4293 (2018). doi: 10.1073/pnas.1706920115; pmid: 29610314
23. C. D. Nadell, K. Drescher, K. R. Foster, Spatial structure, cooperation and competition in biofilms. *Nat. Rev. Microbiol.* **14**, 589–600 (2016). doi: 10.1038/nrmicro.2016.84; pmid: 27452230
24. O. X. Cordero, M. S. Datta, Microbial interactions and community assembly at microscales. *Curr. Opin. Microbiol.* **31**, 227–234 (2016). doi: 10.1016/j.mib.2016.03.015; pmid: 27232202
25. H. Shi *et al*., Highly multiplexed spatial mapping of microbial communities. *Nature* **588**, 676–681 (2020). doi: 10.1038/s41586-020-2983-4; pmid: 33268897
26. C. Tropini, K. A. Earle, K. C. Huang, J. L. Sonnenburg, The gut microbiome: Connecting spatial organization to function. *Cell Host Microbe* **21**, 433–442 (2017). doi: 10.1016/j.chom.2017.03.010; pmid: 28407481
27. S. A. Wilbert, J. L. Mark Welch, G. G. Borisy, Spatial ecology of the human tongue dorsum microbiome. *Cell Rep.* **30**, 4003–4015.e3 (2020). doi: 10.1016/j.celrep.2020.02.097; pmid: 32209464
28. F. Imdahl, E. Vafadarnejad, C. Homberger, A.-E. Saliba, J. Vogel, Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol.* **5**, 1202–1206 (2020). doi: 10.1038/s41564-020-0774-1; pmid: 32807892
29. A. Kuchina *et al*., Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, eaba5257 (2021). doi: 10.1126/science.aba5257; pmid: 33335020
30. S. B. Blattman, W. Jiang, P. Oikonomou, S. Tavazoie, Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat. Microbiol.* **5**, 1192–1201 (2020). doi: 10.1038/s41564-020-0729-6; pmid: 32451472
31. A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008). doi: 10.1038/nmeth.1253; pmid: 18806792
32. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998). doi: 10.1126/science.280.5363.585; pmid: 9554849
33. L.-H. So *et al*., General properties of transcriptional time series in Escherichia coli. *Nat. Genet.* **43**, 554–560 (2011). doi: 10.1038/ng.821; pmid: 21532574
34. H. M. T. Choi, V. A. Beck, N. A. Pierce, Next-generation in situ hybridization chain reaction: Higher gain, lower cost, greater durability. *ACS Nano* **8**, 4284–4294 (2014). doi: 10.1021/nn405717p; pmid: 24712299
35. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015). doi: 10.1126/science.aaa6090; pmid: 25858977
36. C. L. Eng *et al*., Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019). doi: 10.1038/s41586-019-1049-y; pmid: 30911168
37. E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, L. Cai, Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014). doi: 10.1038/nmeth.2892; pmid: 24681720
38. J. R. Moffitt *et al*., Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018). doi: 10.1126/science.aau5324; pmid: 30385464
39. S. Shah, E. Lubeck, W. Zhou, L. Cai, In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* **92**, 342–357 (2016). doi: 10.1016/j.neuron.2016.10.001; pmid: 27764670

40. A. Lignell, L. Kerosuo, S. J. Streichan, L. Cai, M. E. Bronner, Identification of a neural crest stem cell niche by Spatial Genomic Analysis. *Nat. Commun.* **8**, 1830 (2017). doi: 10.1038/s41467-017-01561-w; pmid: 29184067
41. S. Shah *et al*., Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell* **174**, 363–376.e16 (2018). doi: 10.1016/j.cell.2018.05.035; pmid: 29887381
42. S. O. Skinner, L. A. Sepúlveda, H. Xu, I. Golding, Measuring mRNA copy number in individual Escherichia coli cells using single-molecule fluorescent in situ hybridization. *Nat. Protoc.* **8**, 1100–1113 (2013). doi: 10.1038/nprot.2013.066; pmid: 23680982
43. A. Y. Bhagirath *et al*., Cystic fibrosis lung environment and *Pseudomonas aeruginosa* infection. *BMC Pulm. Med.* **16**, 174 (2016). doi: 10.1186/s12890-016-0339-5; pmid: 27919253
44. S. Malhotra, D. Hayes Jr.., D. J. Wozniak, Cystic Fibrosis and *Pseudomonas aeruginosa*: The Host-Microbe Interface. *Clin. Microbiol. Rev.* **32**, e00138-18 (2019). doi: 10.1128/CMR.00138-18; pmid: 31142499
45. Y. Takei *et al*., Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590**, 344–350 (2021). doi: 10.1038/s41586-020-03126-2; pmid: 33505024
46. I. Vallet-Gely, F. Boccard, Chromosomal organization and segregation in *Pseudomonas aeruginosa*. *PLOS Genet.* **9**, e1003492 (2013). doi: 10.1371/journal.pgen.1003492; pmid: 23658532
47. J. Lee, L. Zhang, The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. *Protein Cell* **6**, 26–41 (2015). doi: 10.1007/s13238-014-0100-x; pmid: 25249263
48. A. Price-Whelan, L. E. P. Dietrich, D. K. Newman, Pyocyanin alters redox homeostasis and carbon flux through central metabolic pathways in *Pseudomonas aeruginosa* PA14. *J. Bacteriol.* **189**, 6372–6381 (2007). doi: 10.1128/JB.00505-07; pmid: 17526704
49. J. Jo, K. L. Cortez, W. C. Cornell, A. Price-Whelan, L. E. Dietrich, An orphan $cbb_3$-type cytochrome oxidase subunit supports *Pseudomonas aeruginosa* biofilm growth and virulence. *eLife* **6**, e30205 (2017). doi: 10.7554/eLife.30205; pmid: 29160206
50. H. Arai, Regulation and function of versatile aerobic and anaerobic respiratory metabolism in *Pseudomonas aeruginosa*. *Front. Microbiol.* **2**, 103 (2011). doi: 10.3389/fmicb.2011.00103; pmid: 21833336
51. S. Schulz *et al*., Elucidation of sigma factor-associated networks in *Pseudomonas aeruginosa* reveals a modular architecture with limited and function-specific crosstalk. *PLOS Pathog.* **11**, e1004744 (2015). doi: 10.1371/journal.ppat.1004744; pmid: 25780925
52. S. E. Finkel, Long-term survival during stationary phase: Evolution and the GASP phenotype. *Nat. Rev. Microbiol.* **4**, 113–120 (2006). doi: 10.1038/nrmicro1340; pmid: 16415927
53. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [stat.ML] (9 February 2018).
54. E. Bosdriesz, D. Molenaar, B. Teusink, F. J. Bruggeman, How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *FEBS J.* **282**, 2029–2044 (2015). doi: 10.1111/febs.13258; pmid: 25754869
55. B. R. Borlee *et al*., *Pseudomonas aeruginosa* uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol. Microbiol.* **75**, 827–842 (2010). doi: 10.1111/j.1365-2958.2009.06991.x; pmid: 20088866
56. C. Reichhardt, C. Wong, D. Passos da Silva, D. J. Wozniak, M. R. Parsek, CdrA: interactions within the *Pseudomonas aeruginosa* biofilm matrix safeguard it from proteolysis and promote cellular packing. *mBio* **9**, e01376 (2018). doi: 10.1128/mBio.01376-18; pmid: 30254118
57. O. Zaborina *et al*., Structure-function aspects of PstS in multi-drug-resistant *Pseudomonas aeruginosa*. *PLOS Pathog.* **4**, e43 (2008). doi: 10.1371/journal.ppat.0040043; pmid: 18282104
58. M. L. Gil-Marqués *et al*., Role of PstS in the pathogenesis of *Acinetobacter baumannii* under microaerobiosis and normoxia. *J. Infect. Dis.* **222**, 1204–1212 (2020). doi: 10.1093/infdis/jiaa201; pmid: 32324853
59. A. R. Hauser, The type III secretion system of *Pseudomonas aeruginosa*: Infection by injection. *Nat. Rev. Microbiol.* **7**, 654–665 (2009). doi: 10.1038/nrmicro2199; pmid: 19680249
60. O. Wurtzel *et al*., The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLOS Pathog.* **8**, e1002945 (2012). doi: 10.1371/journal.ppat.1002945; pmid: 23028334
61. C. Lombardi *et al*., Structural and functional characterization of the type three secretion system (T3SS) needle of *Pseudomonas aeruginosa*. *Front. Microbiol.* **10**, 573 (2019). doi: 10.3389/fmicb.2019.00573; pmid: 31001211
62. T. Bjarnsholt *et al*., *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatr. Pulmonol.* **44**, 547–558 (2009). doi: 10.1002/ppul.21011; pmid: 19418571
63. N. Høiby *et al*., The clinical impact of bacterial biofilms. *Int. J. Oral Sci.* **3**, 55–65 (2011). doi: 10.4248/IJOS11026; pmid: 21485309
64. K. L. Palmer, L. M. Aye, M. Whiteley, Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J. Bacteriol.* **189**, 8079–8087 (2007). doi: 10.1128/JB.01138-07; pmid: 17873029
65. G. Gicquel *et al*., The extra-cytoplasmic function sigma factor sigX modulates biofilm and virulence-related properties in *Pseudomonas aeruginosa*. *PLOS ONE* **8**, e80407 (2013). doi: 10.1371/journal.pone.0080407; pmid: 24260387
66. L. L. Burrows, *Pseudomonas aeruginosa* twitching motility: Type IV pili in action. *Annu. Rev. Microbiol.* **66**, 493–520 (2012). doi: 10.1146/annurev-micro-092611-150055; pmid: 22746331
67. G. A. O'Toole, R. Kolter, Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* **30**, 295–304 (1998). doi: 10.1046/j.1365-2958.1998.01062.x; pmid: 9791175
68. R. Belas, Biofilms, flagella, and mechanosensing of surfaces by bacteria. *Trends Microbiol.* **22**, 517–527 (2014). doi: 10.1016/j.tim.2014.05.002; pmid: 24894628
69. B.-J. Laventie *et al*., A surface-induced asymmetric program promotes tissue colonization by *Pseudomonas aeruginosa*. *Cell Host Microbe* **25**, 140–152.e6 (2019). doi: 10.1016/j.chom.2018.11.008; pmid: 30581112
70. M. G. K. Ghequire, R. De Mot, The tailocin tale: Peeling off phage tails. *Trends Microbiol.* **23**, 587–590 (2015). doi: 10.1016/j.tim.2015.07.011; pmid: 26433692
71. L.-M. Bobay, M. Touchon, E. P. C. Rocha, Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12127–12132 (2014). doi: 10.1073/pnas.1405336111; pmid: 25092302
72. L. Turnbull *et al*., Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* **7**, 11220 (2016). doi: 10.1038/ncomms11220; pmid: 27075392
73. J. Vacheron, C. M. Heiman, C. Keel, Live cell dynamics of production, explosive release and killing activity of phage tail-like weapons for *Pseudomonas* kin exclusion. *Commun. Biol.* **4**, 87 (2021). doi: 10.1038/s42003-020-01581-1; pmid: 33469108
74. C. B. Whitchurch, T. Tolker-Nielsen, P. C. Ragas, J. S. Mattick, Extracellular DNA required for bacterial biofilm formation. *Science* **295**, 1487 (2002). doi: 10.1126/science.295.5559.1487; pmid: 11859186
75. M. D. Brazas, R. E. W. Hancock, Ciprofloxacin induction of a susceptibility determinant in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **49**, 3222–3227 (2005). doi: 10.1128/AAC.49.8.3222-3227.2005; pmid: 16048929
76. T. L. Povolotsky, A. Keren-Paz, I. Kolodkin-Gal, Metabolic microenvironments drive microbial differentiation and antibiotic resistance. *Trends Genet.* **37**, 4–8 (2021). doi: 10.1016/j.tig.2020.10.007; pmid: 33203570
77. A. K. Wessel *et al*., Oxygen limitation within a bacterial aggregate. *mBio* **5**, e00992 (2014). doi: 10.1128/mBio.00992-14; pmid: 24736225
78. L. E. P. Dietrich *et al*., Bacterial community morphogenesis is intimately linked to the intracellular redox state. *J. Bacteriol.* **195**, 1371–1380 (2013). doi: 10.1128/JB.02273-12; pmid: 23292774
79. E. S. Cowley, S. H. Kopf, A. LaRiviere, W. Ziebis, D. K. Newman, Pediatric cystic fibrosis sputum can be chemically dynamic, anoxic, and extremely reduced due to hydrogen sulfide formation. *mBio* **6**, e00767 (2015). doi: 10.1128/mBio.00767-15; pmid: 26220964
80. S. S. Yoon *et al*., *Pseudomonas aeruginosa* anaerobic respiration in biofilms: Relationships to cystic fibrosis pathogenesis. *Dev. Cell* **3**, 593–603 (2002). doi: 10.1016/S1534-5807(02)00295-2; pmid: 12408810
81. M. Eschbach *et al*., Long-term anaerobic survival of the opportunistic pathogen *Pseudomonas aeruginosa* via pyruvate fermentation. *J. Bacteriol.* **186**, 4596–4604 (2004). doi: 10.1128/JB.186.14.4596-4604.2004; pmid: 15231792
82. D. J. Hassett, L. Charniga, K. Bean, D. E. Ohman, M. S. Cohen, Response of *Pseudomonas aeruginosa* to pyocyanin: Mechanisms of resistance, antioxidant defenses, and demonstration of a manganese-cofactored superoxide dismutase. *Infect. Immun.* **60**, 328–336 (1992). doi: 10.1128/iai.60.2.328-336.1992; pmid: 1730464
83. S. M. Brown, M. L. Howell, M. L. Vasil, A. J. Anderson, D. J. Hassett, Cloning and characterization of the katB gene of *Pseudomonas aeruginosa* encoding a hydrogen peroxide-inducible catalase: Purification of KatB, cellular localization, and demonstration that it is essential for optimal resistance to hydrogen peroxide. *J. Bacteriol.* **177**, 6536–6544 (1995). doi: 10.1128/jb.177.22.6536-6544.1995; pmid: 7592431
84. S. Su *et al*., Catalase (KatA) plays a role in protection against anaerobic nitric oxide in *Pseudomonas aeruginosa*. *PLOS ONE* **9**, e91813 (2014). doi: 10.1371/journal.pone.0091813; pmid: 24663218
85. F. Cutruzzolà, N. Frankenberg-Dinkel, Origin and impact of nitric oxide in *Pseudomonas aeruginosa* biofilms. *J. Bacteriol.* **198**, 55–65 (2016). doi: 10.1128/JB.00371-15; pmid: 26260455
86. D. W. Basta, M. Bergkessel, D. K. Newman, Identification of fitness determinants during energy-limited growth arrest in *Pseudomonas aeruginosa*. *mBio* **8**, e01170 (2017). doi: 10.1128/mBio.01170-17; pmid: 29184024
87. U. A. Ochsner, M. L. Vasil, E. Alsabbagh, K. Parvatiyar, D. J. Hassett, Role of the *Pseudomonas aeruginosa* oxyR-recG operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of katB-ankB, ahpB, and ahpC-ahpF. *J. Bacteriol.* **182**, 4533–4544 (2000). doi: 10.1128/JB.182.16.4533-4544.2000; pmid: 10913087
88. M. Toyofuku *et al*., Quorum sensing regulates denitrification in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **189**, 4969–4972 (2007). doi: 10.1128/JB.00289-07; pmid: 17449629
89. N. E. Van Alst, L. A. Sherrill, B. H. Iglewski, C. G. Haidaris, Compensatory periplasmic nitrate reductase activity supports anaerobic growth of *Pseudomonas aeruginosa* PAO1 in the absence of membrane nitrate reductase. *Can. J. Microbiol.* **55**, 1133–1144 (2009). doi: 10.1139/W09-065; pmid: 19935885
90. K. Schreiber *et al*., Anaerobic survival of *Pseudomonas aeruginosa* by pyruvate fermentation requires an Usp-type stress protein. *J. Bacteriol.* **188**, 659–668 (2006). doi: 10.1128/JB.188.2.659-668.2006; pmid: 16385055
91. S. K. Dolan, M. Welch, The glyoxylate shunt, 60 years on. *Annu. Rev. Microbiol.* **72**, 309–330 (2018). doi: 10.1146/annurev-micro-090817-062257; pmid: 30200852
92. A. Crousilles, S. K. Dolan, P. Brear, D. Y. Chirgadze, M. Welch, Gluconeogenic precursor availability regulates flux through the glyoxylate shunt in *Pseudomonas aeruginosa*. *J. Biol. Chem.* **293**, 14260–14269 (2018). doi: 10.1074/jbc.RA118.004514; pmid: 30030382
93. S. K. Dolan, G. Pereira, R. Silva-Rocha, M. Welch, Transcriptional regulation of central carbon metabolism in *Pseudomonas aeruginosa*. *Microb. Biotechnol.* **13**, 285–289 (2020). doi: 10.1111/1751-7915.13423; pmid: 31187593
94. T. Kawakami, M. Kuroki, M. Ishii, Y. Igarashi, H. Arai, Differential expression of multiple terminal oxidases for aerobic respiration in *Pseudomonas aeruginosa*. *Environ. Microbiol.* **12**, 1399–1412 (2010). pmid: 19930444
95. A. M. Stern, J. Zhu, An introduction to nitric oxide sensing and response in bacteria. *Adv. Appl. Microbiol.* **87**, 187–220 (2014). doi: 10.1016/B978-0-12-800261-2.00005-0; pmid: 24581392
96. A. Edelstein, N. Amodaj, K. Hoover, R. Vale, N. Stuurman, Computer control of microscopes using μManager. *Curr. Protoc. Mol. Biol.* **Chapter 14**, 20 (2010). pmid: 20890901
97. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012). doi: 10.1038/nmeth.2089; pmid: 22930834
98. S. Stylianidou, C. Brennan, S. B. Nissen, N. J. Kuwada, P. A. Wiggins, SuperSegger: Robust image segmentation, analysis and lineage tracking of bacterial cells. *Mol. Microbiol.* **102**, 690–700 (2016). doi: 10.1111/mmi.13486; pmid: 27569113
99. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). doi: 10.1186/s13059-017-1382-0; pmid: 29409532
100. R. Milo, P. Jorgensen, U. Moran, G. Weber, M. Springer, BioNumbers—The database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **38** (suppl_1), D750–D753 (2010). doi: 10.1093/nar/gkp889; pmid: 19854939

101. S. Berg *et al*., ilastik: Interactive machine learning for (bio) image analysis. *Nat. Methods* **16**, 1226–1232 (2019). doi: 10.1038/s41592-019-0582-9; pmid: 31570887
102. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi: 10.1038/nmeth.1923; pmid: 22388286
103. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014). doi: 10.1093/bioinformatics/btt656; pmid: 24227677
104. Y. Zhang, Z. Hu, Combined treatment of *Pseudomonas aeruginosa* biofilms with bacteriophages and chlorine. *Biotechnol. Bioeng.* **110**, 286–295 (2013). doi: 10.1002/bit.24630; pmid: 22886888
105. D. Dar, N. Dar, L. Cai, D. K. Newman, Data for: Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution, Zenodo (2021); doi: 10.5281/zenodo.4767568
106. D. Dar, N. Dar, L. Cai, D. K. Newman, Imaging data for: Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution, Zenodo (2021); doi: 10.5281/zenodo.4771778



## ACKNOWLEDGMENTS

We thank G. A. O'Toole and M. Whiteley for help with designing the gene set, M. Bergkessel and R. Sorek for critically reading the manuscript, and members of the Newman laboratory for critically reading the manuscript and for fruitful discussions and comments, particularly M. Bergkessel for assistance with RNA-Seq analysis. **Funding:** This work was supported by the National Institutes of Health (grants 1R01AI127850-01A1 and 1R01HL152190-01 to D.K.N.) and the Army Research Office (grant W911NF-17-1-0024 to D.K.N.). L.C. was supported by the Allen Frontier group. D.D. was supported by the Rothschild foundation, EMBO Long-Term, and Helen Hay Whitney postdoctoral fellowships, as well as a Geobiology Postdoctoral Fellowship from the Division of Geological and Planetary Sciences, Caltech. **Author contributions:** D.D., N.D., L.C., and D.K.N. designed the study. D.D. led the study, designed the experiments, and performed the experiments with N.D. D.D. analyzed the data. D.K.N. and L.C. supervised the study. All authors contributed to writing the manuscript. **Competing interests:** L.C. is a cofounder of Spatial Genomics, Inc. A provisional patent (No. 63/153,234) has been filed by California Institute of Technology with inventors Daniel Dar, Dianne K. Newman, Kirsten Frieda, and Long Cai entitled "Multiplexing of experimental conditions and samples in spatial genomics." **Data and materials availability:** Custom MATLAB scripts and single-cell source data from this study are available at Zenodo (*105*). Imaging data obtained during this study have also been deposited at Zenodo (*106*). All other data are presented in the main text or the supplementary materials.


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/eabi4882/suppl/DC1
Figs. S1 to S8
Tables S1 to S4
MDAR Reproducibility Checklist

12 March 2021; accepted 25 June 2021
10.1126/science.abi4882

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

# Ultrapotent antibodies against diverse and highly transmissible SARS-CoV-2 variants

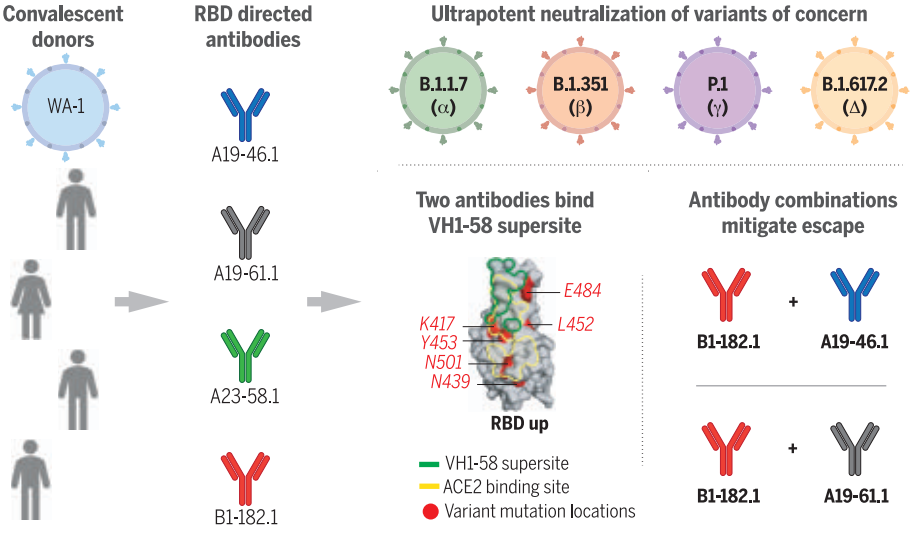
**Lingshu Wang†, Tongqing Zhou†, Yi Zhang, Eun Sung Yang, Chaim A. Schramm, Wei Shi, Amarendra Pegu, Olamide K. Oloniniyi, Amy R. Henry, Samuel Darko, Sandeep R. Narpala, Christian Hatcher, David R. Martinez, Yaroslav Tsybovsky, Emily Phung, Olubukola M. Abiona, Avan Antia, Evan M. Cale, Lauren A. Chang, Misook Choe, Kizzmekia S. Corbett, Rachel L. Davis, Anthony T. DiPiazza, Ingelise J. Gordon, Sabrina Helmold-Hait, Tandile Hermanus, Prudence Kgagudi, Farida Laboune, Kwanyee Leung, Tracy Liu, Rosemarie D. Mason, Alexandra F. Nazzari, Laura Novik, Sarah O'Connell, Sijy O'Dell, Adam S. Olia, Stephen D. Schmidt, Tyler Stephens, Christopher D. Stringham, Chloe Adrienna Talana, I-Ting Teng, Danielle A. Wagner, Alicia T. Widge, Baoshan Zhang, Mario Roederer, Julie E. Ledgerwood, Tracy J. Ruckwardt, Martin R. Gaudinski, Penny L. Moore, Nicole A. Doria-Rose, Ralph S. Baric, Barney S. Graham, Adrian B. McDermott, Daniel C. Douek, Peter D. Kwong, John R. Mascola, Nancy J. Sullivan*, John Misasi†**

**INTRODUCTION:** Worldwide appearance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern (VOCs) with increased transmissibility and resistance to therapeutic antibodies necessitates the discovery of broadly reactive antibodies. We isolated receptor binding domain (RBD) targeting antibodies that potently neutralize 23 variants, including the B.1.1.7, B.1.351, P.1, B.1.429, B.1.526, and B.1.617 VOCs. Structural and functional studies revealed the molecular basis for antibody binding and showed that antibody combinations reduce the generation of escape mutants, suggesting a potential means to mitigate development of therapeutic resistance.

**RATIONALE:** Investigation of antibody responses from convalescent subjects infected with the Washington-1 (WA-1) strain for reactivity against WA-1 and VOCs can inform improvements to vaccine design and therapeutics.

**RESULTS:** Blood from 22 convalescent subjects who recovered from SARS-CoV-2 WA-1 infection was screened for neutralizing and binding activity, and four subjects with high reactivity against the WA-1 variant were selected for antibody isolation. SARS-CoV-2 spike (S)–reactive antibodies were identified through B cell sorting with S protein–based probes. WA-1 live-virus neutralization assays identified four RBD-targeting antibodies with high potency [half-maximal inhibitory concentration ($IC_{50}$) 2.1 to 4.8 ng/ml], two of which were derived from the same IGHV1-58 germline but from different donors. Antigen-binding fragments (Fabs) of these antibodies exhibited nanomolar affinity to S (2.3 to 7.3 nM). Competition assays and electron microscopy indicated that two of the most potent antibodies blocked angiotensin-converting enzyme 2 (ACE2) and bound open conformation RBD, whereas the other two bound both up and down conformations of RBD and blocked ACE2 binding. Binding and lentivirus neutralization assays against 13 circulating VOCs or variants of interest—including B.1.1.7, B.1.351, B.1.427, B.1.429, B.1.526, P.1, P.2, B.1.617.1, and B.1.617.2—indicated that these antibodies were highly potent against VOCs despite being isolated from subjects infected with early ancestral SARS-CoV-2 viruses. Cryo-EM studies of the two most potent antibodies in complex with S revealed that these antibodies target a site of vulnerability on RBD but have minimal contacts with mutational hotspots, defining the structural basis for their high effectiveness against the emerging VOCs and further delineating an IGHV1-58 antibody supersite. To investigate potential mechanisms of escape, we applied antibody selection pressure to replication-competent vesicular stomatitis virus (rcVSV) expressing the WA-1 SARS-CoV-2 S (rcVSV-SARS2) and identified S mutations that conferred in vitro resistance. We evaluated these antibodies individually or in combinations for their capacity to prevent rcVSV-SARS2 escape and discovered that antibody combinations with complementary modes of recognition to the RBD lowered the risk of resistance.

**CONCLUSION:** Our study demonstrates that convalescent subjects previously infected with ancestral variant SARS-CoV-2 produce antibodies that cross-neutralize emerging VOCs with high potency. Structural and functional analyses reveal that antibody breadth is mediated by targeting a site of vulnerability at the RBD tip offset from major mutational hotspots in VOCs. Selective boosting of immune responses targeting specific RBD epitopes, such as the sites defined by these antibodies, may induce breadth against current and future VOCs. ■




**Isolation and characterization of convalescent donor antibodies that effectively neutralize emerging SARS-CoV-2 VOCs.** Antibodies isolated from donors infected with ancestral SARS-CoV-2 viruses showed ultrapotent neutralization of emerging VOCs. The two most potent antibodies shared usage of the IGHV1-58 gene and targeted the RBD with minimal contact to VOC mutational hotspots. Cocktails of antibodies with complementary binding modes suppressed antibody escape.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: njsull@mail.nih.gov
†These authors contributed equally to this work.
Cite this article as L. Wang *et al.*, *Science* **373**, eabh1766 (2021). DOI: 10.1126/science.abh1766
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.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abh1766

RESEARCH ARTICLE

CORONAVIRUS

# Ultrapotent antibodies against diverse and highly transmissible SARS-CoV-2 variants

Lingshu Wang[1]†, Tongqing Zhou[1]†, Yi Zhang[1], Eun Sung Yang[1], Chaim A. Schramm[1], Wei Shi[1], Amarendra Pegu[1], Olamide K. Oloniniyi[1], Amy R. Henry[1], Samuel Darko[1], Sandeep R. Narpala[1], Christian Hatcher[1], David R. Martinez[2,3], Yaroslav Tsybovsky[4], Emily Phung[1], Olubukola M. Abiona[1], Avan Antia[1], Evan M. Cale[1], Lauren A. Chang[1], Misook Choe[1], Kizzmekia S. Corbett[1], Rachel L. Davis[1], Anthony T. DiPiazza[1], Ingelise J. Gordon[1], Sabrina Helmold Hait[1], Tandile Hermanus[5,6], Prudence Kgagudi[5,6], Farida Laboune[1], Kwanyee Leung[1], Tracy Liu[1], Rosemarie D. Mason[1], Alexandra F. Nazzari[1], Laura Novik[1], Sarah O'Connell[1], Sijy O'Dell[1], Adam S. Olia[1], Stephen D. Schmidt[1], Tyler Stephens[4], Christopher D. Stringham[1], Chloe Adrienna Talana[1], I-Ting Teng[1], Danielle A. Wagner[1], Alicia T. Widge[1], Baoshan Zhang[1], Mario Roederer[1], Julie E. Ledgerwood[1], Tracy J. Ruckwardt[1], Martin R. Gaudinski[1], Penny L. Moore[5,6], Nicole A. Doria-Rose[1], Ralph S. Baric[2,3], Barney S. Graham[1], Adrian B. McDermott[1], Daniel C. Douek[1], Peter D. Kwong[1], John R. Mascola[1], Nancy J. Sullivan[1]*, John Misasi[1]†

The emergence of highly transmissible SARS-CoV-2 variants of concern (VOCs) that are resistant to therapeutic antibodies highlights the need for continuing discovery of broadly reactive antibodies. We identified four receptor binding domain–targeting antibodies from three early-outbreak convalescent donors with potent neutralizing activity against 23 variants, including the B.1.1.7, B.1.351, P.1, B.1.429, B.1.526, and B.1.617 VOCs. Two antibodies are ultrapotent, with subnanomolar neutralization titers [half-maximal inhibitory concentration ($IC_{50}$) 0.3 to 11.1 nanograms per milliliter; $IC_{80}$ 1.5 to 34.5 nanograms per milliliter). We define the structural and functional determinants of binding for all four VOC-targeting antibodies and show that combinations of two antibodies decrease the in vitro generation of escape mutants, suggesting their potential in mitigating resistance development.

Since the start of the severe acute respiratory syndrone coronavirus 2 (SARS-CoV-2) outbreak, >170 million people have been infected, and >3.7 million have died from COVID-19 (*1*). The virus is decorated with a trimeric spike protein (S), which comprises an S1 subunit that binds host cells and an S2 subunit that is responsible for membrane fusion. The S1 subunit comprises an N-terminal domain (NTD); the receptor binding domain (RBD) that binds the host angiotensin-converting enzyme 2 (ACE2) receptor; and two additional subdomains, SD1 and SD2. Shortly after the first Wuhan Hu-1 (Hu-1) genome sequence was published (*2*), S proteins based on this sequence were generated for use in antibody discovery (*3*–*5*). SARS-CoV-2 variants such as B.1.1.7 (for example, Alpha, 501Y.V1) (*6*), B.1.351 (for example, Beta, 501Y.V2) (*7*), P.1 (for example, Gamma, 501Y.V3), and B.1.617.2 (for example, Delta, 452R.V3) (*8*, *9*) contain mutations, many in S, that mediate resistance to therapeutic monoclonal antibodies, have increased transmissibility, and potentially increase pathogenicity (*10*–*14*). Vaccines designs based on the original Hu-1 outbreak strain sequence elicit antibody responses that show decreased in vitro neutralizing activity against variants (*14*–*16*). In this study, antibodies isolated from convalescent subjects who were infected by the Washington-1 (WA-1) strain, which has an identical S sequence to Hu-1, were investigated for reactivity against WA-1 and variants of concern (VOCs), and we defined the structural features of their binding to S.

[1]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [2]Department of Epidemiology, UNC Chapel Hill School of Public Health, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. [3]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. [4]Electron Microscopy Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA. [5]National Institute for Communicable Diseases (NICD) of the National Health Laboratory Service (NHLS), Johannesburg, South Africa. [6]SAMRC Antibody Immunity Research Unit, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
*Corresponding author. Email: njsull@mail.nih.gov
†These authors contributed equally to this work.

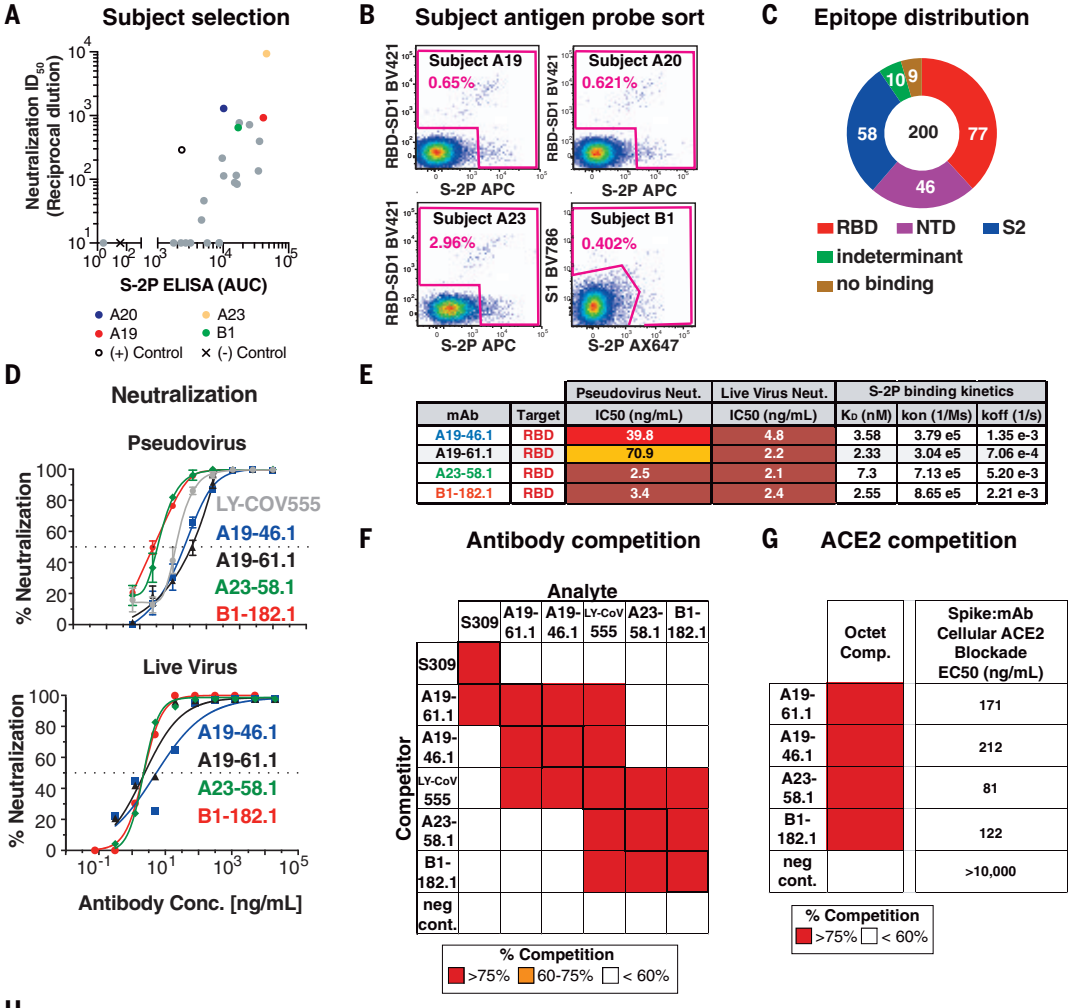
## Identification and characterization of antibodies against WA-1

We obtained blood from 22 convalescent subjects, who had experienced mild to moderate symptoms after WA-1 infection, between 25 and 55 days after symptom onset. Four subjects—A19, A20, A23, and B1—had both high neutralizing and binding activity against the WA-1 variant (Fig. 1A) and were selected for antibody isolation efforts. $CD19^+/CD20^+$/immunoglobulin $M^-$ ($IgM^-$)/$IgA^+$ or $IgG^+$ B cells were sorted for binding to a stabilized version of S (S-2P), the full S1 subunit, or the RBD plus the subdomain-1 region of S1 (RBD-SD1) (Fig. 1B and fig. S1). In total, we sorted 889 B cells, recovered 709 (80%) paired heavy- and light-chain antibody sequences, and selected 200 antibodies for expression. A meso scale discovery (MSD) binding assay was used to measure binding of these 200 antibodies to stabilized spike, the full S1 subunit, RBD, or NTD. There was a broad response across all spike domains with 77 binding RBD, 46 binding NTD, 58 inferred to bind the S2 subunit based on binding to S but not to S1, and 19 binding an indeterminant epitope or failing to recognize spike in an MSD binding assay (Fig. 1C).

Pseudovirus neutralization assays by using the WA-1 spike showed that four RBD targeting antibodies—A19-46.1, A19-61.1, A23-58.1, and B1-182.1 (table S1)—are especially potent [half-maximal inhibitory concentration (the concentration of an antibody required to inhibit virus entry by 50%) ($IC_{50}$) 2.5 to 70.9 ng/ml] (Fig. 1, D and E). WA-1 live virus neutralization (*17*) revealed similar high potent neutralization by all four antibodies ($IC_{50}$ 2.1 to 4.8 ng/ml) (Fig. 1, D and E). All four antibody Fabs exhibited nanomolar affinity for SARS-CoV-2 S-2P (2.3 to 7.3 nM), which is consistent with their potent neutralization (Fig. 1E).

Antibodies targeting the RBD can be categorized into four general classes (classes I to IV) on the basis of competition with the ACE2 target cell receptor protein for binding to S and recognition of the up or down state of the three RBDs in S (*18*). LY-CoV555 is a therapeutic antibody that binds RBD in both the up and down states, blocks ACE2 binding, and is categorized as class II. However, despite potent activity against WA-1, VOCs have been reported to contain mutations that confer resistance to LY-CoV555 (*14*, *19*, *20*) and similarly binding antibodies. We therefore examined whether the epitopes targeted by the four high-potency antibodies were distinct from LY-CoV555. We used a surface plasmon resonance (SPR)–based competition binding assay to compare the binding profile of these antibodies to LY-CoV555. Although LY-CoV555 competed with A19-46.1, A19-61.1, A23-58.1, and B1-182.1 (and vice versa), their overall competition profiles were not the same. A23-58.1 and B1-182.1 exhibit similar binding profiles, and A19-61.1 and A19-46.1 likewise display a shared competition binding profile in our SPR assay. However, the latter two antibodies can be distinguished from each other owing to A19-61.1 competition with the class III antibody S309 (Fig. 1F) (*21*), which binds an epitope in RBD that is accessible in the up or down position but does not compete with ACE2 binding (*18*).

To determine whether the antibodies block ACE2 binding, we used biolayer interferometry




| mAb | Target | Pseudovirus Neut. IC50 (ng/mL) | Live Virus Neut. IC50 (ng/mL) | S-2P binding kinetics KD (nM) | kon (1/Ms) | koff (1/s) |
|---|---|---|---|---|---|---|
| A19-46.1 | RBD | 39.8 | 4.8 | 3.58 | 3.79 e5 | 1.35 e-3 |
| A19-61.1 | RBD | 70.9 | 2.2 | 2.33 | 3.04 e5 | 7.06 e-4 |
| A23-58.1 | RBD | 2.5 | 2.1 | 7.3 | 7.13 e5 | 5.20 e-3 |
| B1-182.1 | RBD | 3.4 | 2.4 | 2.55 | 8.65 e5 | 2.21 e-3 |

**Fig. 1. Identification and classification of highly potent antibodies from convalescent SARS-CoV-2 subjects.** (**A**) Sera from 22 convalescent subjects were tested for neutralizing (*y* axis, $ID_{50}$) and binding antibodies (*x* axis, S-2P ELISA AUC), and four subjects—A19, A20, A23, and B1 (colored) with both high neutralizing and binding activity against the WA-1—were selected for antibody isolation. (**B**) Final flow cytometry sorting gate of $CD19^+/CD20^+/IgG^+$ or $IgA^+$ PBMCs for four convalescent subjects (A19, A20, A23, and B1). Shown is the staining for RBD-SD1 BV421, S1 BV786, and S-2P APC or Ax647. Cells were sorted by using indicated sorting gate (pink), and percent of positive cells that were either RBD-SD1-, S1-, or S-2P-– positive is shown for each subject. (**C**) Gross binding epitope distribution was determined by using an MSD-based ELISA testing against RBD, NTD, S1, S-2P, or HexaPro. S2 binding was inferred from S-2P or HexaPro binding without binding to other antigens. Indeterminant epitopes showed a mixed binding profile. Total number of antibodies (200) and absolute number of antibodies within each group is shown. (**D**) Neutralization curves by using WA-1 spike pseudotyped lentivirus and live virus neutralization assays to test the neutralization capacity of the indicated antibodies (*n* = 2 to 3 replicates). (**E**) Table showing antibody binding target, $IC_{50}$ for pseudovirus and live virus neutralization, and Fab:S-2P binding kinetics (*n* = 2 replicates) for the indicated antibodies. (**F**) SPR-based epitope binning experiment. Competitor antibody (*y* axis) is bound to S-2P before incubation with the analyte antibody (*x* axis) as indicated, and percent competition range bins are shown as red (>75%), orange (60 to 75%), or white (<60%) (*n* = 2 replicates). Negative control antibody is anti-Ebola glycoprotein antibody mAb114 (*37*). (**G**) Competition of ACE2 binding. The indicated antibodies (*y* axis) complete binding of S-2P to soluble ACE2 protein by using biolayer interferometry [left column, percent competition (>75% shown as red, <60% as white)] or to cell surface–expressed ACE2 by using cell-surface staining (right column, $EC_{50}$ at ng/ml shown). (**H**) Negative-stain 3D reconstructions of SARS-CoV-2 spike and Fab complexes. A19-46.1 and A19-61.1 bind to RBD in the down position, whereas A23-58.1 and B1-182.1 bind to RBD in the up position. Representative classes were shown with two Fabs bound, although stoichiometry at one to three Fabs was observed.

ACE2-competition and cell-surface binding assays to show that all four antibodies prevent the binding of ACE2 to spike (Fig. 1G and fig. S2). This suggests that A19-46.1, A23-58.1, and B1-182.1 neutralize infection by directly blocking the interaction of RBD with ACE2 and would be classified as either class I (ACE2 blocking, binding RBD up only) or II (ACE2 blocking, binding RBD up or down) RBD antibodies (*18*). A19-61.1 competition with S309 and ACE2 binding suggests that it binds at least partly outside of the ACE2 binding motif but may sterically block ACE2 binding similar to the class III antibody REGN10987. To refine the classification of these antibodies, we performed negative-stain three-dimensional (3D) reconstruction and found that A19-46.1 and A19-61.1 bound near one another with all RBDs in the down position (Fig. 1H), which is consistent with them being class II and class III antibodies, respectively. Similarly, A23-58.1 and B1-182.1 bound to overlapping regions when RBDs are in the up position, suggesting that they are class I antibodies.

### Antibody binding and neutralization against circulating variants

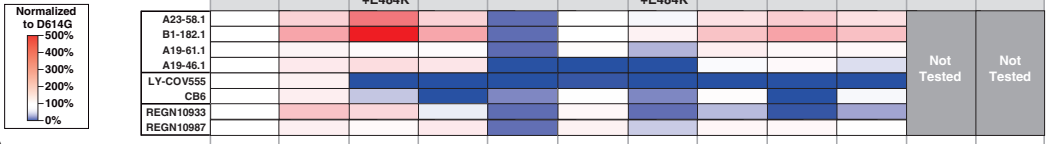
Because each donor subject was infected with a variant close to the ancestral WA-1, we evaluated antibody activity against recently emerged variants such as D614G, which has become the dominant variant across the world (*22*). Similar to LY-CoV555, neutralization potency was increased against D614G compared with WA-1, with the $IC_{50}$ and $IC_{80}$ of each experimental antibody 1.4- to 6.3-fold lower than that seen for the WA-1 ($IC_{50}$ of 0.8 to 20.3 ng/ml and $IC_{80}$ of 2.6 to 43.5 ng/ml) (Fig. 2, A and C, and fig. S3). [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. In the mutants, other amino acids were substituted at certain locations; for example, D614G indicates that

**Fig. 2. Antibody binding and neutralization of VOCs or VOIs.** (**A**) Table showing domain and mutations relative to WA-1 for each of the 10 variants tested in (B) and (C). (**B**) Spike protein variants were expressed on the surface of HEK293 T cells, and binding to the indicated antibody was measured with flow cytometry. Data are shown as MFI normalized to the MFI for the same antibody against the D614G parental variant. Percent change is indicated by a color gradient from red (increased binding, Max 500%) to white (no change, 100%) to blue (no binding, 0%). (**C**) $IC_{50}$ and $IC_{80}$ values for the indicated antibodies against 10 variants shown in (A). Ranges are indicated with white (>10,000 ng/ml), light blue (>1000 to ≤10,000 ng/ml), yellow (>100 to ≤1000 ng/ml), orange (>50 to ≤100 ng/ml), red (>10 to ≤50 ng/ml), maroon (>1 to ≤10 ng/ml), and purple (≤1 ng/ml). (**D**) Location of spike protein variant mutations on the spike glycoprotein for B.1.1.7, B.1.351, B.1.429, P.1 v2, B.1.617.1, and B.1.617.2. P681 and V1176 are not resolved in the structure, and therefore their locations are not noted in B.1.1.7 and P.1 v2.

**A**

| Domain | Mutations | D614G | B.1.1.7 | B.1.1.7 +E484K | B.1.351 v2 | B.1.427 | B.1.429 | B.1.429 +E484K | B.1.526 v2 | P.1 v2 | P.2 | B.1.617.1 | B.1.617.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NTD | T95I/G142D/E154K | | | | | | | | | | | NTD x 3 mut | |
| | T19R/G142D/Δ156-157/R158G | | | | | | | | | | | | NTD x 5 mut |
| | L5F/T95I/D253G | | | | | | | | NTD x 3 mut | | | | |
| | L18F/D80A/D215G/ΔL242-244 | | | | NTD x 6 mut | | | | | | | | |
| | L18F/T20N/P26S/D138Y/R190S | | | | | | | | | NTD x 5 mut | | | |
| | ΔH69 /V70 | | ΔH69-70 | ΔH69-70 | | | | | | | | | |
| | S13I/W152C | | | | | | S13I/W152C | S13I/W152C | | | | | |
| | ΔY144 | | ΔY144 | ΔY144 | | | | | | | | | |
| RBD | K417N or K417T | | | | K417N | | | | | K417T | | | |
| | L452R | | | | | L452R | L452R | L452R | | | | L452R | L452R |
| | T478K | | | | | | | | | | | | T478K |
| | E484K or E484Q | | | E484K | E484K | | | E484K | E484K | E484K | E484K | E484Q | |
| | N501Y | | N501Y | N501Y | N501Y | | | | | N501Y | | | |
| S1-C-term | A570D | | A570D | A570D | | | | | | | | | |
| | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G | D614G |
| | H655Y | | | | | | | | | H655Y | | | |
| | P681H or P681R | | P681H | P681H | | | | | | | | P681R | P681R |
| S2 | T716I/S982A/D1118H | | S2 x 3 | S2 x 3 | | | | | | | | | |
| | A701V | | | | A701V | | | | A701V | | | | |
| | D950N | | | | | | | | | | | | D950N |
| | T1027I | | | | | | | | | T1027I | | | |
| | Q1071H | | | | | | | | | | | Q1071H | |
| | V1176F | | | | | | | | | V1176F | V1176F | | |

**B** Cell surface binding

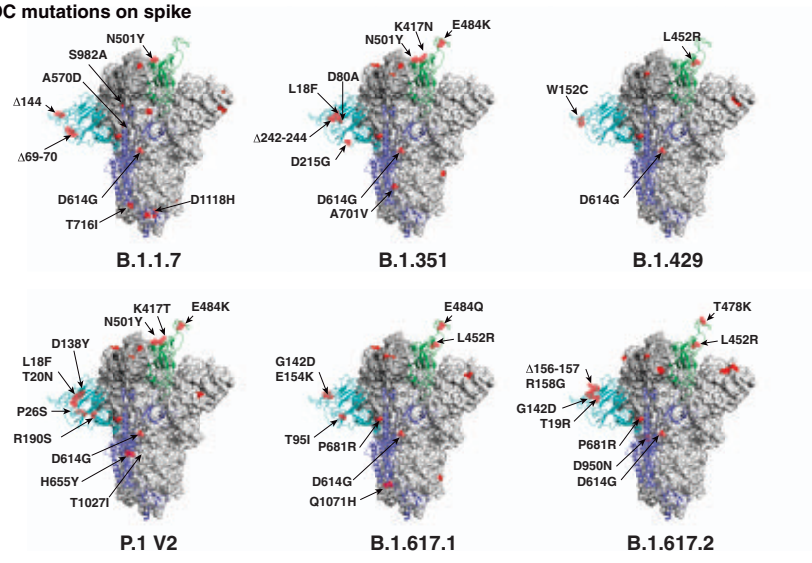



**C** Neutralization

| IC50 (ng/mL) | WA-1 | D614G | B.1.1.7 | B.1.1.7 +E484K | B.1.351 v2 | B.1.427 | B.1.429 | B.1.429 +E484K | B.1.526 v2 | P.1 v2 | P.2 | B.1.617.1 | B.1.617.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A23-58.1 | 2.5 | 1.8 | < 0.6 | 11.1 | 1.6 | 3.8 | 1.9 | 4.7 | 4.5 | 0.7 | 10.1 | 3.9 | 1.6 |
| B1-182.1 | 3.4 | 0.8 | < 0.6 | 5.0 | 0.7 | 1.6 | 1.5 | 3.4 | 2.1 | 0.3 | 4.8 | 2.9 | 1.0 |
| A19-61.1 | 70.9 | 12.8 | 11.1 | 22.0 | 10.8 | 23.4 | 15.5 | 7.2 | 7.1 | 18.7 | 17.1 | 13.1 | 28.3 |
| A19-46.1 | 39.8 | 20.3 | 11.5 | 82.0 | 57.1 | > 10,000 | > 10,000 | > 10,000 | 72.5 | 23.2 | 34.8 | > 10,000 | > 10,000 |
| LY-COV555 | 7.1 | 3.4 | 4.1 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 |
| CB6 | 26.2 | 31.0 | 40.1 | 575.8 | > 10,000 | 54.2 | 22.9 | 118.1 | 35.3 | > 10,000 | 49.4 | 36.5 | 13.9 |
| REGN10933 | 7.7 | 5.2 | 6.9 | 87.5 | > 10,000 | 7.5 | 9.0 | 62.5 | 71.3 | 1046.3 | 56.2 | 97.1 | 3.2 |
| REGN10987 | 71.0 | 20.0 | 16.8 | 10.9 | 24.4 | 47.1 | 74.9 | 113.6 | 27.5 | 4.0 | 33.9 | 65.7 | 312.8 |

| IC80 (ng/mL) | WA-1 | D614G | B.1.1.7 | B.1.1.7 +E484K | B.1.351 v2 | B.1.427 | B.1.429 | B.1.429 +E484K | B.1.526 v2 | P.1 v2 | P.2 | B.1.617.1 | B.1.617.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A23-58.1 | 10.7 | 4.9 | 3.9 | 34.5 | 9.1 | 14.7 | 6.4 | 27.0 | 12.8 | 3.2 | 29.8 | 13.9 | 3.5 |
| B1-182.1 | 8.8 | 2.6 | 2.4 | 14.8 | 2.6 | 9.2 | 3.4 | 10.6 | 9.2 | 1.5 | 17.3 | 7.9 | 3.5 |
| A19-61.1 | 163.3 | 26.1 | 18.6 | 37.6 | 19.6 | 64.9 | 31.1 | 16.7 | 14.3 | 30.3 | 28.4 | 23.8 | 41.0 |
| A19-46.1 | 128.7 | 43.5 | 25.0 | 206.3 | 157.2 | > 10,000 | > 10,000 | > 10,000 | 180.9 | 74.9 | 230.8 | > 10,000 | > 10,000 |
| LY-COV555 | 35.7 | 10.5 | 15.9 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 | > 10,000 |
| CB6 | 113.9 | 83.1 | 1512.3 | 3855.6 | > 10,000 | 380.3 | 125.0 | 241.1 | 147.2 | > 10,000 | 294.2 | 178.3 | 37.0 |
| REGN10933 | 21.1 | 15.7 | 23.3 | 1333.5 | > 10,000 | 40.7 | 21.0 | 186.6 | 341.9 | > 10,000 | 305.2 | 283.1 | 10.1 |
| REGN10987 | 855.5 | 412.2 | 232.1 | 122.5 | 101.5 | 403.3 | 464.3 | 761.2 | 179.0 | 21.3 | 176.0 | 809.5 | 1265.0 |

**D** Mapping of VOC mutations on spike



aspartic acid at position 614 was replaced by glycine.]

Next, we assessed antibody binding to D614G and nine additional cell surface–expressed spike variants that have appeared subsequent to WA-1 and that are not considered VOCs or variants of interest (VOIs) (B.1.1.7.14, B.1.258.24, Y453F/D614G, Ap.1, B.1.388, ΔH69-70/N501Y/D614G, K417N/D614G, B.1.1.345, and B.1.177.31) (*6*–*9*, *22*). Experimental antibodies were compared with four antibodies that are in clinical use [LY-CoV555, REGN10933, REGN10987, and CB6 (LY-CoV016)]. All control and experimental antibodies showed a minor reduction in binding (less than twofold) to B.1.258.24 (N439K/D614G) (figs. S3 and S4). Despite this, their neutralization capacities were minimally affected, with the exception of REGN10987 (2005 ng/ml) as reported previously (figs. S3 and S4) (*23*). Whereas none of the experimental antibodies showed large reductions in binding, LY-CoV555, CB6 (*24*), and REGN10933 (*25*) each showed >10-fold binding deficits to one or more variants (Y453F/D614G, K417N/D614G, B.1.1.345, or B.1.177.31) in these cell-based binding assays (figs. S3 and S4).

We next evaluated the capacity of each antibody to neutralize lentiviral particles pseudotyped with the same 10 variant spike proteins. Consistent with published data, REGN10933 did not neutralize Y453F/D614G or B.1.177.31 (K417N/E484K/N501Y/D614G) (*12*, *14*, *26*); CB6 did not neutralize B.1.177.31; and LY-CoV555 and REGN109333 showed potency reductions of 28- to >1400-fold for neutralization of viruses containing E484K (fig. S3) (*12*, *14*). Relative to WA-1, the A23-58.1 $IC_{50}$ neutralization was threefold lower for ΔH69-70/N501Y/D614G (0.9 ng/ml) and fivefold lower for Ap.1 (<0.6 ng/ml), and although A23-58.1 maintained high potency, neutralization against B.1.1.345 was increased fourfold (10.2 ng/ml). Neutralization by B1-182.1 maintained high potency ($IC_{50}$ < 3.2 ng/ml) for all variants and showed more than fourfold improved potency for 6 of the 10 variants tested ($IC_{50}$ < 0.8 ng/ml) (fig. S3). For A19-61.1, variant neutralization was three- to sixfold more potent than that of WA-1 (WA-1 $IC_{50}$ 70.9 ng/ml; variants $IC_{50}$ 11.1 to 23.7 ng/ml) (fig. S3). Last, neutralization by A19-46.1 was similar to that of WA-1 for all variants except B.1.1.345 and B.1.177.31, which were still highly potent despite having $IC_{50}$ values that were two to threefold less active (B.1.1.345, 95.0 ng/ml; B.1.177.311, 61.8 ng/ml; and WA-1, 39.8 ng/ml) (fig. S3). Together, these data show the capacity of these newly identified antibodies to maintain high neutralization potency against a diverse panel of 10 variant spike proteins.

### Antibody binding and neutralization of VOIs and VOCs

We analyzed neutralization of 13 circulating VOIs and VOCs, some of which have high transmissibility, including B.1.1.7, B.1.351, B.1.427, B.1.429, B.1.526, P.1, P.2, B.1.617.1, and B.1.617.2 (Fig. 2 and fig. S3) (*6*, *7*, *11*). Consistent with published data, we found that LY-CoV555, CB6, REGN10933, and REGN10987 maintained high potency against B.1.1.7 ($IC_{50}$ 0.1 to 40.1 ng/ml), and LY-CoV555 and CB6 were unable to neutralize B.1.351 v1, B.1.351 v2, P.1 v1, or P.1 v2 variants ($IC_{50}$ > 10,000 ng/ml) (Fig. 2 and fig. S3) (*12*, *14*, *26*); LY-CoV555 was unable to neutralize B.1.526 v2, B.1.617.1, and B.1.617.2; CB6 showed 5- to 27-fold worse activity against B.1.1.7+E484K and B.1.429+E484K but remained active against B.1.617.1 and B.1.617.2; REGN10933 showed 9- to 200-fold reduction in neutralization against variants with mutations at E484 (B.1.1.7+E484K, B.1.429+E484K, B.1.526 v2, P.1 v1/v2, and B.1.617.1) and maintained activity against B.1.617.2, which does not contain a mutation at E484 (Fig. 2 and fig. S3); and REGN10987 maintained or had slightly increased potency against each of the VOCs and VOIs except B.1.617.2, which showed a fourfold reduction in activity (Fig. 2 and fig. S3). In comparison, A23-58.1, B1-182.1, A19-46.1, and A19-61.1 maintained similar or improved potency ($IC_{50}$ < 0.6 to 11.5 ng/ml) against B.1.1.7 and B.1.1.7+E484K relative to WA-1 (Fig. 2 and fig. S3). The potency of A19-46.1 was within 2.5-fold or lower relative to WA-1 for all variants ($IC_{50}$ 11.5 to 101.4 ng/ml versus WA-1 39.8 ng/ml), except those containing L452R ($IC_{50}$ >10,000 ng/ml) (B.1.427, B.1.429, B.1.429+E484K, B.1.617.1, and B.1.617.2) (Fig. 2 and fig. S3). Further analyses showed that A23-58.1, B1-182.1, and A19-61.1 maintained high potency against all VOCs and VOIs ($IC_{50}$ < 0.6 to 28.3 ng/ml), including the recently identified B.1.617.1 and B.1.617.2 (Fig. 2 and fig. S3). These results indicate that despite being isolated from subjects infected with early ancestral SARS-CoV-2 viruses, each of these antibodies have highly potent reactivity against VOCs.

### Structural and functional analysis of VH1-58 antibodies

The two most potent antibodies, A23-58.1 and B1-182.1, shared highly similar gene family usage in their heavy and light chains, despite being from different donors (table S1). Both use IGHV1-58 heavy chains and IGKV3-20/IGKJ1 light chains and similarly low levels of somatic hypermutation (SHM) (<0.7%) (table S1). This antibody gene family combination has been identified in other COVID-19 convalescent subjects and has been proposed as a public clonotype (*27*–*30*). To gain structural insights on the interaction between this class of antibodies and the SARS-CoV-2 spike, we obtained cryo–electron microscopy (cryo-EM) reconstructions for structures of the Fab A23-58.1 bound to a stabilized WA-1 S at 3.39 Å resolution and of the Fab B1-182.1 bound to a stabilized WA-1 S at 3.15 Å resolution (Fig. 3, A and B; figs. S5 and S6; and table S2). This revealed that the antibody bound to spike with all RBDs in the up position, confirming the negative stain results (Fig. 1H). However, the cryo-EM reconstruction densities of the interface between RBD and Fab were poor owing to conformational variation.

To resolve the antibody-antigen interface, we performed local refinement and improved the local resolution to 3.89 Å for A23-58.1 and to 3.71 Å for B1-182.1 (figs. S5 and S6). Because both A23-58.1 and B1-182.1 recognized the RBD in a very similar way, we used the RBD-A23-58.1 structure for detailed analysis. Antibody A23-58.1 binds to an epitope on the RBD that faces the threefold axis of the spike and is accessible only in the RBD-up conformation (Fig. 3A). The interaction buried a total of 619 $Å^2$ surface area from the antibody and 624 $Å^2$ from the spike (table S3). The A23-58.1 paratope constituted all six complementarity-determining regions (CDRs) with heavy chain and light chain contributing 74 and 26% of the binding surface area, respectively (Fig. 3, C and E, and table S3). The 14-residue-long CDR H3, which is 48% of the heavy-chain paratope, kinks at $Pro^{95}$ and $Phe^{100F}$ (Kabat numbering scheme for antibody residues) to form a foot-like loop that is stabilized by an intraloop disulfide bond between $Cys^{97}$ and $Cys^{100B}$ at the arch. A glycan was observed at the CDR H3 $Asn^{96}$ (fig. S5F). The CDRs formed an interfacial crater with a depth of ~10 Å and a diameter of ~20 Å at the opening. Paratope residues inside the crater were primarily aromatic or hydrophobic. CDR H3 $Pro^{95}$ and $Phe^{100F}$ lined the bottom, and CDR H1 $Ala^{33}$, CDR H2 $Trp^{50}$ and $Val^{52}$, and CDR H3 $Val^{100A}$ lined the heavy-chain side of the crater (Fig. 3, D and E). On the light-chain side, CDR L1 $Tyr^{32}$ and CDR L3 residues $Tyr^{91}$ and $Trp^{96}$ provided 80% of the light chain–binding surface (Fig. 3, D and E). By contrast, paratope residues at the rim of the crater are mainly hydrophilic; for example, $Asp^{100D}$ formed hydrogen bonds with $Ser^{477}$ and $Asn^{487}$ of the RBD (Fig. 3D and table S3).

The A23-58.1 epitope comprised residues between β5 and β6 at the tip of RBD (Figs. 3D and 4A). With the protruding $Phe^{486}$ dipping into the crater formed by the CDRs, these residues formed a hook-like motif that is stabilized by an intraloop disulfide bond between $Cys^{480}$ and $Cys^{488}$. Aromatic residues—including $Phe^{456}$, $Tyr^{473}$, $Phe^{486}$, and $Tyr^{489}$—provided 48% (299 $Å^2$) of the epitope (Fig. 3D and table S3). $Lys^{417}$ and $Glu^{484}$, which are located at the outer edge of the epitope, contributed only 3.7% of the binding surface (Fig. 3C and table S3). Overall, the cryo-EM analysis provides a structural basis for the potent neutralization of the E484K/Q mutant by A23-58.1.

The binding modes and sequences of A23-58.1 and B1-182.1 are very similar to those of

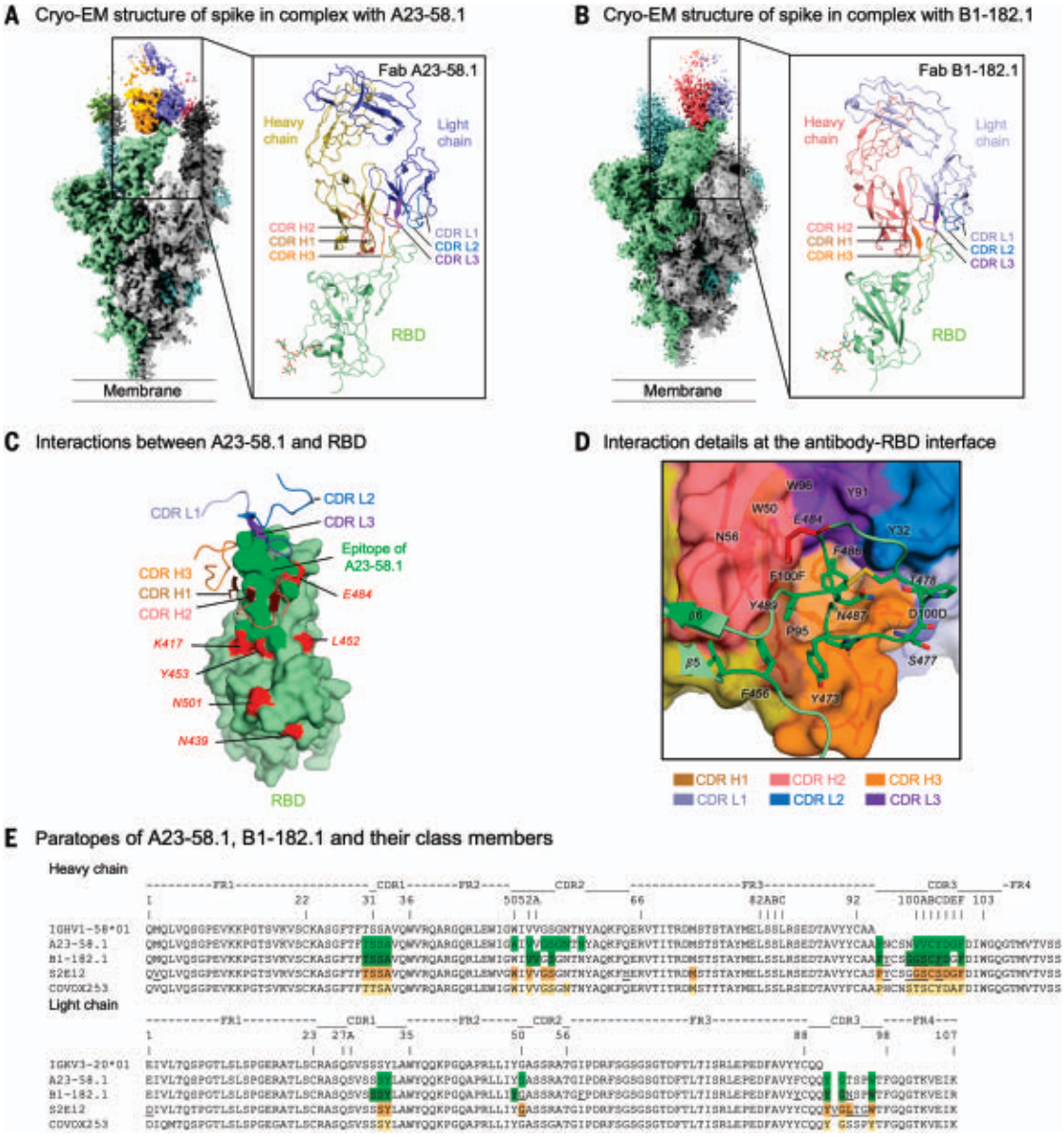



**Fig. 3. Structural basis of binding and neutralization for antibodies A23-58.1 and B1-182.1.** (**A**) Cryo-EM structure of A23-58.1 Fab in complex with SARS-CoV-2 HexaPro spike. (Left) Overall density map. Protomers are light green, gray, and cyan. One of the A23-58.1 Fab bound to the RBD is shown in orange and blue. (Right) Structure of the RBD and A23-58.1 after local focused refinement. The heavy-chain CDRs are brown, salmon, and orange for CDR H1, CDR H2, and CDR H3, respectively. The light-chain CDRs are marine blue, light blue, and purple blue for CDR L1, CDR L2, and CDR L3, respectively. The contour level of the cryo-EM map is 5.7σ. (**B**) Cryo-EM structure of B1-182.1 Fab in complex with SARS-CoV-2 HexaPro spike. (Left) Overall density map. Protomers are light green, gray, and cyan. One of the B1-182.1 Fab bound to the RBD is shown in salmon and light blue. (Right) Structure of the RBD and B1-182.1 after local focused refinement. The heavy-chain CDRs are brown, deep salmon, and orange for CDR H1, CDR H2, and CDR H3, respectively. The light-chain CDRs are marine blue, slate, and purple blue for CDR L1, CDR L2, and CDR L3, respectively. The contour level of the cryo-EM map is 4.0σ. (**C**) Interaction between A23-58.1 and RBD. All CDRs were involved in binding of RBD. Epitope of A23-58.1 is shown in bright green surface. RBD mutations in current circulating SARS-CoV-2 variants are red. K417 and E484 are located at the edge of the epitope. (**D**) Interaction details at the antibody-RBD interface. The tip of the RBD binds to a crater formed by the CDRs (shown viewing down to the crater). Interactions between aromatic and hydrophobic residues are prominent at the lower part of the crater. Hydrogen bonds at the rim of the crater are indicated with dashed lines. RBD residues are indicated with italicized font. (**E**) Paratopes of A23-58.1, B1-182.1, S2E12 (PDB ID: 7K45), and COVOX253 (PDB ID: 7BEN) from the same germline. Sequences of B1-182.1, S2E12, and COVOX253 were aligned with variant residues underlined. Paratope residues for A23-58.1, B1-182.1, S2E12, and COVOX253 were highlighted in green, dark green, light brown, and light orange, respectively.

previously reported IGHV1-58/IGKV3-20–derived antibodies, such as S2E12 (*27*), COVOX 253 (*30*), and CoV2-2196 (*31*), confirming that they are members of the same structural class (Fig. 3E). To understand why B1-182.1 is highly effective at neutralizing the emerging VOCs, we compared its binding mode with that of A23-58.1. Analysis indicated that B1-182.1 rotated about 6° along the long axis of Fab from that of A23-58.1 (Fig. 4B). This rotation on one hand increased B1-182.1 CDR L1 contacts on invariant regions of RBD to strengthen binding (Fig. 4B) and on the other hand critically reduced contact on $Glu^{484}$ to 6 $Å^2$ and main-chain only compared with ~40 $Å^2$ main- and side-chain contacts for A58.1 and S2E12 (Fig. 4B and table S3). Overall, the subtle changes in antibody mode of recognition to regions on RBD harboring variant mutations provided structural basis on the effectiveness of B1-182.1 and A23-58.1 on neutralization of VOCs.

To understand how A23-58.1 and B1-182.1 overcome mutations that cause reduced antibody potency against virus variants, we superposed the antibody-RBD complex structures of CB6 [Protein Data Bank (PDB) ID 7C01] (*24*), REGN10933 (PDB ID 6XDG) (*25*, *26*), and LY-CoV555 (PDB ID 7KMG) (*19*) with the A23-58.1 structure over the RBD region. Both REGN10933 and CB6 bind to the same side of the RBD as does A23-58.1 (Fig. 4C). However, the binding surfaces of REGN10933 and CB6

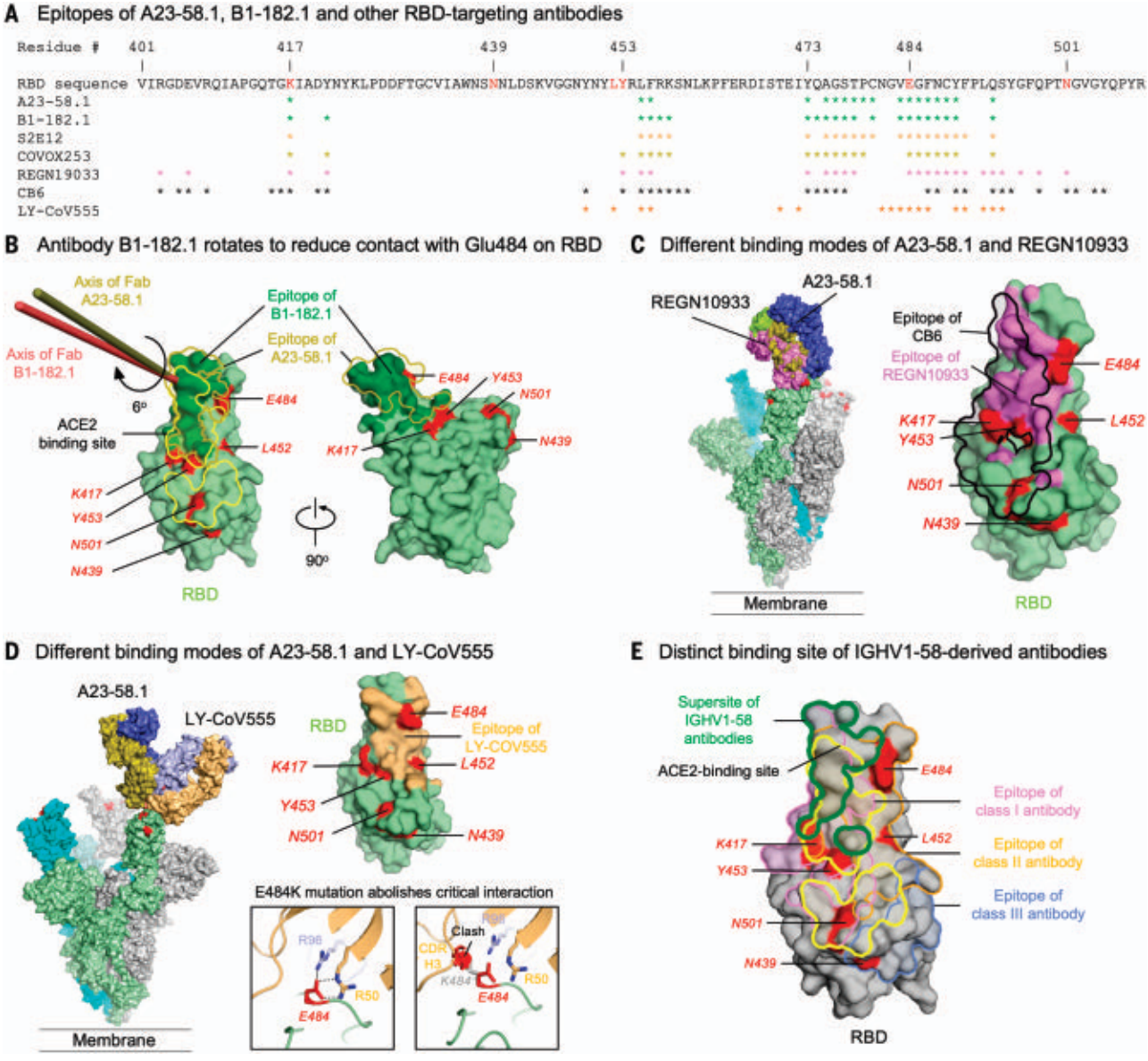



**Fig. 4. Binding modes of A23-58.1 and B1-182.1 enable neutralization to VOCs.** (**A**) Mapping of epitopes of A23-58.1, B1-182.1, and other antibodies on RBD. Epitope residues for different RBD-targeting antibodies are marked with an asterisk under the RBD sequence. (**B**) Comparison of binding modes of A23-58.1 and B1-182.1. (Left) Analysis indicated that axis of Fab B1-182.1 is rotated 6° from that of A23-58.1. (Right) This rotation resulted in a slight shift of the epitope of B1-182.1 on RBD, which reduced its contact to E484. RBD mutations of concern are red, the epitope surface of B1-182.1 is dark green, and the borders of ACE2-binding site and A23-58.1 epitope are yellow and olive, respectively. (**C**) Comparison of binding modes of A23-58.1, CB6, and REGN10933. For clarity, one Fab is shown to bind to the RBD on the spike. The shift of the binding site to the saddle of RBD encircled K417, E484, and Y453 inside the CB6 (black line) and REGN10933 epitopes (violet surface), explaining their sensitivity to the K417N, Y453F, and E484K mutations. (**D**) Comparison of binding modes of A23-58.1 and LY-CoV555. (Left) One Fab is shown to bind to the RBD on the spike. (Top right) E484 is located inside the LY-CoV555 epitope. (Bottom right) E484K/Q mutation abolishes critical contacts between RBD and CDR H2 and CDR L3; moreover, E484K/Q and L452R cause potential clashes with heavy chain of LY-CoV555, explaining its sensitivity to the E484K/Q and L452R mutations. (**E**) IGHV1-58–derived antibodies target a supersite with minimal contacts to mutational hotspots. Supersite defined by common atoms contacted by the IGHV1-58–derived antibodies (A23-58.1, B1-182.1, S2E12, and COVOX253) on RBD is indicated with the green line. Boundaries of the ACE2-binding site and epitopes of class I, II, and III antibodies represented by C102 (PDB ID 7K8M), C144 (PDB ID 7K90), and C135 (PDB ID 7K8Z) are indicated with yellow, pink, light orange, and blue boundary lines, respectively.

were shifted toward the saddle of the open RBD and encompassed residues Lys$^{417}$, Tyr$^{453}$, Glu$^{484}$, and Asn$^{501}$ (Fig. 4C); mutations K417N and Y453F thus would abolish key interactions and lead to the loss of neutralization for both REGN10933 and CB6 (Fig. 2). By contrast, LY-CoV555 approached the RBD from a different angle, with its epitope encompassing Glu$^{484}$ and Lys$^{452}$ (Fig. 4D). Structural examination indicates that E484K/Q abolishes key interactions with CDR H2 Arg$^{50}$ and CDR L3 Arg$^{96}$ of LY-CoV555. In addition, both E484K/Q (Fig. 4D) and L452R mutations cause clashes with heavy chain of LY-CoV555. When compared with epitopes of class I, II, and III antibodies (*30*), the supersite defined by common contacts of the IGHV1-58–derived antibodies (A23-58.1, B1-182.1, S2E12, and COVOX253)

had minimal interactions with residues at the mutational hotspots (Fig. 4E). These structural data suggest that the binding modes of A23-58.1 and B1-182.1 enabled their high effectiveness against the new SARS-CoV-2 VOCs.

On the basis of the structural analysis, we investigated the relative contribution of predicted contact residues on binding and neutralization (Fig. 4A). Cell surface–expressed spike binding to A23-58.1 and B1-182.1 was knocked out by F486R, N487R, and Y489R (Fig. 5A and fig. S7), resulting in a lack of neutralization for viruses pseudotyped with spikes containing these mutations (Fig. 5B). By contrast, binding and neutralization of A19-46.1 and A19-61.1 were minimally affected by these changes (Fig. 6, B and C, and fig. S7). CB6, LY-CoV555, and REGN10933 binding and neutralization were also affected by the three mutations, which is consistent with the structural analysis that these residues are shared contact(s) with A23-58.1 and B1-182.1. Taken together, the shared binding and neutralization defects suggest that the hook-like motif and CDR crater are critical for the binding of antibodies within the VH1-58 public class.

### Generation and testing of escape mutations

To explore critical contact residues and mechanisms of escape that might be generated during the course of infection, we applied antibody selection pressure to replication-competent vesicular stomatitis virus (rcVSV) expressing the WA-1 SARS-CoV-2 spike (rcVSV-SARS2) (*32*) to identify spike mutations that confer in vitro resistance against A23-58.1, B1-182.1, A19-46.1, or A19-61.1 (fig. S8). rcVSV-SARS2 was incubated with increasing concentrations of antibody, and cultures from the highest concentration of antibody with >20% cytopathic effect (CPE), relative to no infection control, were carried forward into a second round of selection to drive resistance (fig. S8) (*26*). A shift to higher antibody concentrations required for neutralization indicates the presence of resistant viruses. To gain insight into spike mutations driving resistance, we performed Illumina-based shotgun sequencing (fig. S8). Variants present at a frequency of >5% and increasing from round 1 to round 2 were considered to be positively selected resistant viruses. For A19-46.1, escape mutations were generated at four sites: Y449S (frequency 15%), N450S (frequency 16%), N450Y (frequency 14%), L452R (frequency 83%), and F490V (frequency 58%) (Fig. 6A and fig. S8). The most dominant, L452R, is consistent with the previous finding that B.1.427, B.1.429, B.1.617.1, and B.1.617.2 were resistant to A19-46.1 (Fig. 2 and fig. S3). Although F490L severely reduced neutralization by A19-46.1 ($IC_{50}$> 10,000 ng/ml), the effect of F490V was minimal, suggesting that F490V may require additional mutations for escape to occur (Fig. 6, A to C). Because Y449, N450, and L452 are immediately adjacent to S494, we tested whether S494R would also disrupt binding and neutralization (Fig. 6, A to C, and fig. S9) and found that this mutation mediates neutralization escape. Each of the identified residue locations was confirmed through binding and/or neutralization and would be expected to be accessible when RBD is in the up or down position (fig. S9), and several are shared by class II RBD antibodies (*18*, *33*) and REGN10933 (*25*, *34*).

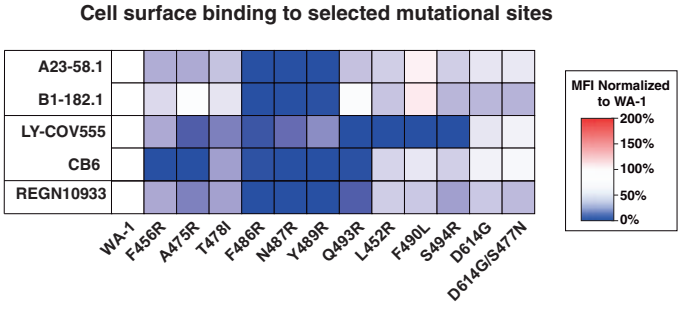
Three residues were positively selected in the presence of A19-61.1: K444E/T (frequency 7-93%), G446V (frequency 24%), and G593R (frequency 19%) (Fig. 6A). There was no overlap with those selected by A19-46.1. G593R is located outside the RBD domain (fig. S9), did not affect neutralization, and may therefore represent a false positive. The highest frequency change was K444E, which represented 57 to 93% of the sequences in replicate experiments (Fig. 6A). This residue is critical for the binding of class III RBD antibodies such as REGN10987 (*18*, *25*, *26*, *34*). Because of the proximity of S494 to K444 and G446, S494R was tested for escape potential and shown to mediate escape from A19-61.1 neutralization. These results are consistent with A19-61.1 targeting a distinct epitope from REGN10987 and other class III RBD antibodies.

For A23-58.1, a single F486S mutation (frequency 91 to 98%) was positively selected. Similarly, B1-182.1 escape was mediated by F486L (frequency 21%), N487D (frequency 100%), and Q493R (frequency 45%). Q493R had minimal impact on binding and was not found to affect neutralization (Fig. 6, B and C). However, F486, N487, and Y489 were all in agreement with previous structural analysis (Figs. 3D, 5, and 6 and fig. S9). F486 is located at the tip of the RBD hook and contacts the binding interface in the antibody crater where aromatic side chains dominantly form the hook and crater interface (Fig. 3D). Therefore, the loss in activity may occur through replacement of a hydrophobic aromatic residue (phenylalanine) with a small polar side chain (serine) (Fig. 3D).

### Potential escape risk and mitigation

To probe the relevance of in vitro–derived resistance variants to potential clinical resistance, we investigated the relative frequency

**A**




**B**

Impact of mutations on antibody neutralization

IC 50 (ng/mL)

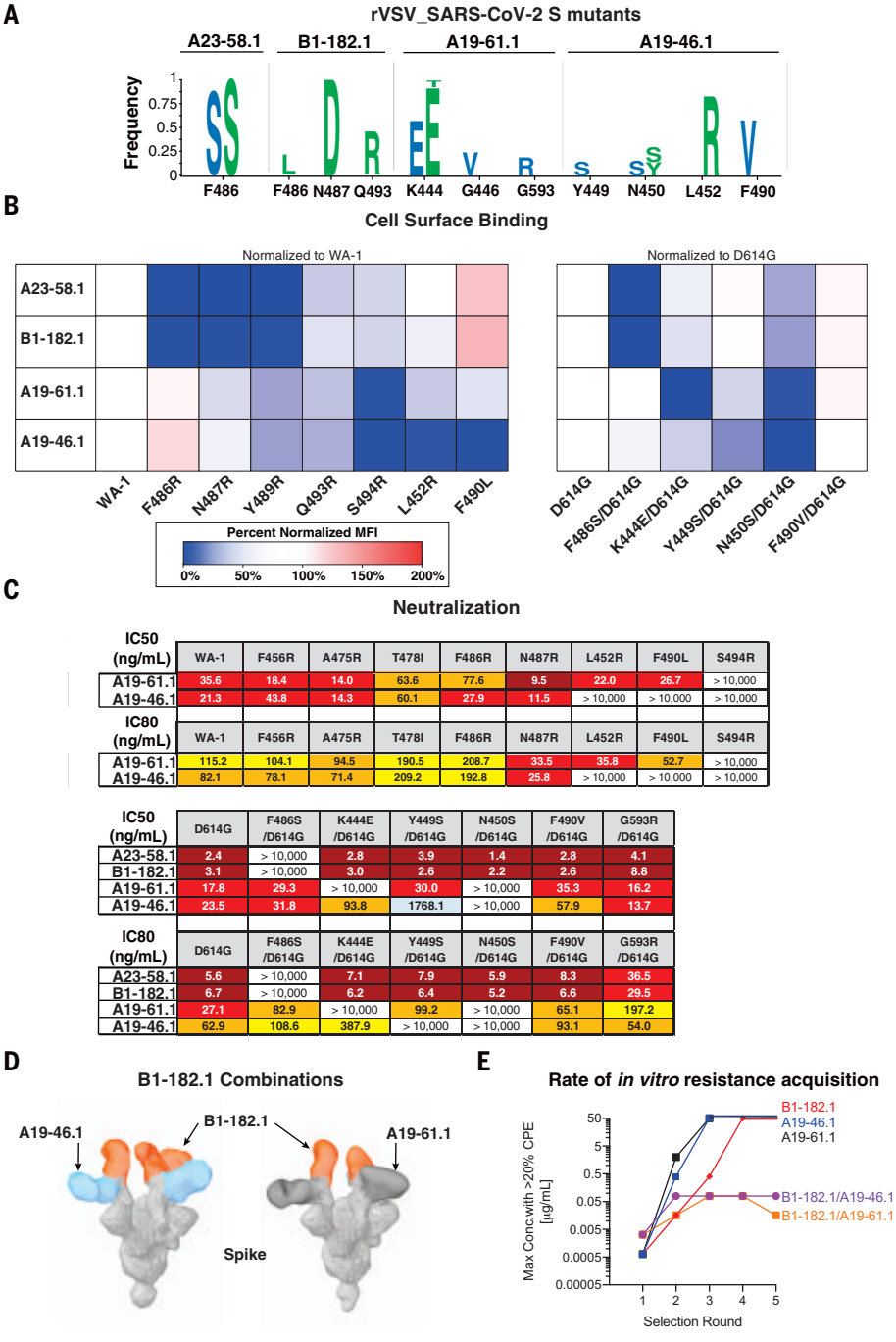
| | WA-1 | F456R | A475R | T478I | F486R | N487R | L452R | F490L | S494R | D614/S477N |
|---|---|---|---|---|---|---|---|---|---|---|
| A23-58.1 | 3.5 | 22.0 | 8.6 | 13.9 | > 10,000 | > 10,000 | 3.2 | 8.1 | 4.2 | 2.8 |
| B1-182.1 | 1.5 | 7.2 | 8.0 | 8.1 | > 10,000 | > 10,000 | 1.6 | 2.7 | 2.0 | 3.3 |
| LY-COV555 | 12.1 | 29.8 | 17.6 | 18.8 | 1181.0 | 6.3 | > 10,000 | > 10,000 | > 10,000 | 7.9 |
| CB6 | 28.0 | > 10,000 | > 10,000 | 21.3 | 468.2 | > 10,000 | 17.5 | 117.4 | 86.4 | 18.0 |
| REGN10933 | 6.1 | 23.1 | 341.5 | 7.8 | > 10,000 | > 10,000 | 5.8 | 26.9 | 6.0 | 16.8 |
| REGN10987 | 42.7 | 23.9 | 23.1 | 60.9 | 18.8 | 3.5 | 215.4 | 57.7 | 227.0 | 29.4 |

IC 80 (ng/mL)

| | WA-1 | F456R | A475R | T478I | F486R | N487R | L452R | F490L | S494R | D614/S477N |
|---|---|---|---|---|---|---|---|---|---|---|
| A23-58.1 | 8.9 | 58.3 | 29.6 | 50.9 | > 10,000 | > 10,000 | 7.0 | 15.3 | 17.1 | 7.2 |
| B1-182.1 | 7.8 | 25.6 | 26.1 | 42.0 | > 10,000 | > 10,000 | 6.1 | 4.9 | 6.9 | 7.0 |
| LY-COV555 | 29.6 | 153.0 | 38.4 | 49.7 | > 10,000 | 15.4 | > 10,000 | > 10,000 | > 10,000 | 19.2 |
| CB6 | 118.7 | > 10,000 | > 10,000 | 97.4 | 2258.5 | > 10,000 | 107.4 | 233.7 | 335.8 | 73.3 |
| REGN10933 | 39.4 | 55.4 | 1307.4 | 26.6 | > 10,000 | > 10,000 | 22.3 | 60.9 | 29.5 | 45.4 |
| REGN10987 | 160.9 | 43.8 | 129.6 | 580.0 | 160.2 | 24.9 | 1171.9 | 693.9 | 2308.5 | 192.9 |

**Fig. 5. Critical binding residues for antibodies A23-58.1 and B1-182.1.** (**A**) The indicated spike protein mutations predicted with structural analysis were expressed on the surface of HEK293 T cells, and binding to the indicated antibody was measured with flow cytometry. Data are shown as MFI normalized to the MFI for the same antibody against the WA-1 parental binding. Percent change is indicated by a color gradient from red (increased binding, max 200%) to white (no change, 100%) to blue (no binding, 0%). (**B**) $IC_{50}$ and $IC_{80}$ values for the indicated antibodies against WA-1 and the nine spike mutations. Ranges are indicated with white (>10,000 ng/ml), light blue (>1000 to ≤10,000 ng/ml), yellow (>100 to ≤1000 ng/ml), orange (>50 to ≤100 ng/ml), red (>10 to ≤50 ng/ml), and maroon (>1 to ≤10 ng/ml).

**Fig. 6. Mitigation of escape risk by using dual antibody combinations.** (**A**) Replication competent vesicular stomatitis virus (rcVSV) whose genome-expressed SARS-CoV-2 WA-1 was incubated with serial dilutions of the indicated antibodies and wells with cytopathic effect (CPE) were passaged forward into subsequent rounds (fig. S8) after 48 to 72 hours. Total supernatant RNA was harvested, and viral genomes were shotgun sequenced to determine the frequency of amino acid changes. Shown are the spike protein amino acid and position change and frequency as a logo plot. Amino acid changes observed in two independent experiments are indicated in blue and green letters. (**B**) The indicated spike protein mutations predicted with structural analysis (Fig. 3) or observed with escape analysis (Fig. 6A) were expressed on the surface of HEK293 T cells, and binding to the indicated antibody was measured with flow cytometry. Data are shown as MFI normalized to the MFI for the same antibody against the (left) WA-1 or (right) D614G parental binding. Percent change is indicated with a color gradient from red (increased binding, max 200%) to white (no change, 100%) to blue (no binding, 0%). (**C**) $IC_{50}$ and $IC_{80}$ values for the indicated antibodies against WA-1 and the mutations predicted with structural analysis (Fig. 3) or observed with escape analysis (Fig. 6A). Ranges are indicated with white (>10,000 ng/ml), light blue (>1000 to ≤10,000 ng/ml), yellow (>100 to ≤1000 ng/ml), orange (>50 to ≤100 ng/ml), red (>10 to ≤50 ng/ml), and maroon (>1 to ≤10 ng/ml). (**D**) Negative-stain 3D reconstruction of the ternary complex of spike with Fab B1-182.1 and (left) A19-46.1 or (right) A19-61.1. (**E**) rcVSV SARS-CoV-2 was incubated with increasing concentrations ($1.3 \times 10^{-4}$ to 50 μg/ml) of either single antibodies (A19-46.1, A19-61.1, and B1-182.1) and combinations of antibodies (B1-182.1/A19-46.1 and B1-182.1/A19-61.1). Every 3 days, wells were assessed for CPE, and the highest concentration well with the >20% CPE was passaged forward onto fresh cells and antibody-containing media. Shown is the maximum concentration with >20% CPE for each of the test conditions in each round of selection. Once 50 μg/ml has been reached, virus was no longer passaged forward.




of variants containing escape mutations present in the GISAID sequence database using the COVID-19 Viral Genome Analysis Pipeline (https://cov.lanl.gov) (*22*) in which, as of 7 May 2021, there were 1,062,910 entries. Of the residues noted to mediate escape or resistance to A19-46.1 (Y449S, N450S/Y, L452R, F490L/V, and S494R), only F490L (0.02%) and L452R (2.27%) were present at greater than 0.01%. For the A19-61.1 escape mutations (K444E, G446V, and S494R), only G446V has been noted in the database >0.01% (0.03%). Last, for A23-58.1 and B1-182.1, ancestral WA-1 residues F486, N487, and Y489 were present in >99.96% of sequences, and only F486L was noted in the database at >0.01% (0.03%). Although the relative lack of A19-61.1, A23-58.1, and B1-182.1 escape mutations in circulating viruses could reflect either under-sampling or the absence of selection pressure, it may also suggest that the in vitro–derived mutations may exact a fitness cost on the virus.

Viral genome sequencing has suggested that in addition to spread through transmission, convergent selection of de novo mutations may be occurring (*6*–*9*, *13*, *22*, *35*). Therefore, effective therapeutic antibody approaches might require new antibodies or combinations of antibodies to mitigate the impact of mutations. On the basis of their complementary modes of spike recognition and breadth of neutralizing activity, combination of B1-182.1 with either A19-46.1 or A19-61.1 may decrease the rate of in vitro resistance acquisition compared with each antibody alone. Consistent with the competition data (Fig. 1F), negative-stain EM 3D reconstructions show that the Fabs in both combinations were able to simultaneously engage spike with the RBDs in the

up position (Fig. 6D). Binding was observed for up to three Fabs of B1-182.1 and three Fabs of A19-46.1 or A19-61.1 per spike in the observed particles (Fig. 6D), indicating that the epitopes of A19-46.1 and A19-61.1 on the spike are accessible in both RBD up and down positions (Figs. 1H and 6D). The absence of observed RBD-down classes suggests the possibility that the combination induces a preferred mode of RBD-up engagement (RBD up versus RBD down) because of the requirement of B1-182.1 or A23-58.1 for RBD-up binding.

Next, we evaluated the capacity of individual antibodies or combinations to prevent the appearance of rcVSV SARS-CoV-2–induced cytopathic effect (CPE) through multiple rounds of passaging in the presence of increasing concentrations of antibodies. In each round, the well with the highest concentration of antibody with at least 20% CPE was carried forward into the next round. We found that wells with A19-61.1 or A785.46.1 single-antibody treatment reached the 20% CPE threshold in their 50 μg/ml well after three rounds of selection (Fig. 6E). Similarly, B1-182.1 single-antibody treatment reached >20% CPE in the 50 μg/ml wells after four rounds (Fig. 6E). Conversely, for both dual treatments (B1-182.1/A19-46.1 or B1-182.1/A19-61.1), the 20% CPE threshold was reached at a concentration of only 0.08 μg/ml and did not progress to higher concentrations, despite five rounds of passaging (Fig. 6E). Thus, combinations may lower the risk that a natural variant will lead to the complete loss of neutralizing activity and suggests a path forward for these antibodies as combination therapies.

## Discussion

Worldwide genomic sequencing has revealed the occurrence of SARS-CoV-2 variants that increase transmissibility and reduce potency of vaccine-induced and therapeutic antibodies (*10*–*16*). Recently, there has been substantial concern that antibody responses to natural infection and vaccination by using ancestral spike sequences may result in focused responses that lack potency against mutations present in more recent variants (such as K417N, L452R, T478K, E484K/Q, N501Y in B.1.351, B.1.617.1, and B.1.617.2) (*12*–*16*). Additionally, neutralization of P.1 viruses can be achieved by using sera obtained from subjects infected by B.1.351 (*36*), suggesting that shared epitopes in RBD (K417N, E484K, and N501Y) are mediating the cross-reactivity. Although the mechanism of B.1.351 and P.1 cross-reactivity is likely focused on the three RBD mutations, the mechanism of broadly neutralizing antibody responses between WA-1 and later variants is not as well established. As a first step to address the risk of reduced antibody potency against new variants, we isolated and defined new antibodies with neutralization breadth covering newly emerging SARS-CoV-2 variants, including the highly transmissible variants B.1.1.7, B.1.351, and B.1.617.2. Increased potency and breadth were mediated by binding to regions of the RBD tip that are offset from E484K/Q, L452R and other mutational hot spots that are major determinants of resistance in VOCs (*10*–*16*).

Our results show that highly potent neutralizing antibodies with activity against VOCs was present in at least three of four convalescent subjects who had been infected with ancestral variants of SARS-CoV-2 (Figs. 1 and 2 and figs. S3 and S10). Furthermore, our structural analyses, the relative paucity of potential escape variants in the GSAID genome database, the identification of public clonotypes (*27*, *28*), and each subject having mild to moderate illness all suggest that these antibodies were generated in subjects who rapidly controlled their infection and were not likely to have been generated because of the generation of a E484 escape mutation during the course of illness. Taken together, these data establish the rationale for a vaccine-boosting regimen that may be used to selectively induce immune responses that increase the breadth and potency of antibodies that target specific RBD regions of the spike glycoprotein (such as VH1-58 supersite). Because both variant sequence analysis and in vitro time-to-escape experiments suggest that combinations of these antibodies may have a lower risk for loss of neutralizing activity, these antibodies represent a potential means to achieve both breadth against current VOCs and to mitigate risk against those that may develop in the future.

## Materials and methods

### *Isolation of PBMCs from SARS CoV-2 subjects*

Human convalescent sera samples were obtained 25 to 55 days following symptom onset from adults with previous mild to moderate SARS-CoV-2 infection. Specimens were collected after subjects provided written informed consent under institutional review board approved protocols at the National Institutes of Health Clinical Center (NCT00067054) and University of Washington (Seattle) (Hospitalized or Ambulatory Adults with Respiratory Viral Infection [HAARVI] study). Whole blood was collected in vacutainer tubes, which were inverted gently to remix cells prior to standard Ficoll-Hypaque density gradient centrifugation (Pharmacia; Uppsala, Sweden) to isolate peripheral blood mononuclear cells (PBMCs). PBMCs were frozen in heat-inactivated fetal calf serum containing 10% dimethylsulfoxide in a Forma CryoMed cell freezer (Marietta, OH). Cells were stored at ≤−140°C

### *Expression and purification of protein*

For expression of soluble SARS CoV-2 S-2P protein, manufacturer's instructions were followed. Briefly, plasmid was transfected using Expifectamine into Expi293 cells (Life Technology, #A14635, A14527) and the cultures enhanced 16-24 hours post-transfection. Following 4-5 days incubations at 120 rpm, 37°C, 9% CO2, supernatant was harvested, clarified via centrifugation, and buffer exchanged into 1X PBS. Protein of interests were then isolated by affinity chromatography using Streptactin resin (Life science) followed by size exclusion chromatography on a Superose 6 increase 10/300 column (GE healthcare).

Expression and purification of biotinylated S-2P, NTD, RBD-SD1 and Hexapro used in binding assays were produced by an in-column biotinylation method as previously described (*5*). Using full-length SARS-Cov2 S and human ACE2 cDNA ORF clone vector (Sino Biological, Inc) as the template to generate S1 or ACE2 dimer proteins. The S1 PCR fragment (1~681aa) was digested with Xbal and BamHI and cloned into the VRC8400 with HRV3C-his (6X) or Avi-HRV3C-his(6X) tag on the C-terminal. The ACE2 PCR fragment (1~740aa) was digested with Xbal and BamHI and cloned into the VRC8400 with Avi-HRV3C-single chain-human Fc-his (6x) tag on the C-terminal. All constructs were confirmed by sequencing. Proteins were expressed in Expi293 cells by transfection with expression vectors encoding corresponding genes. The transfected cells were cultured in shaker incubator at 120 rpm, 37°C, 9% CO2 for 4~5 days. Culture supernatants were harvested and filtered, and proteins were purified through a Hispur Ni-NTA resin (Thermo Scientific, #88221) and following a Hiload 16/600 Superdex 200 column (GE healthcare, Piscataway NJ) according to manufacturer's instructions. The protein purity was confirmed with SDS–polyacrylamide gel electrophoresis (SDS-PAGE).

### *Probe conjugation*

SARS CoV-2 Spike trimer (S-2P) and subdomains (NTD, RBD-SD1, S1) were produced by transient transfection of 293 Freestyle cells as previously described (*4*). Avi-tagged S1 was biotinylated using the BirA biotin-protein ligase reaction kit (Avidity, #BirA500) according to the manufacturer's instructions. The S-2P, RBD-SD1, and NTD proteins were produced by an in-column biotinylation method as previously described (*5*). Successful biotinylation was confirmed using Bio-Layer Interferometry, by testing the ability of biotinylated protein to bind to streptavidin sensors. Retention of antigenicity was confirmed by testing biotinylated proteins against a panel of cross-reactive SARS-CoV and SARS CoV-2 human monoclonal antibodies. Biotinylated probes were conjugated using either allophycocyanin (APC)-, Ax647-, BV421-, BV786-, BV711-, or BV570-labeled streptavidin. Reactions were prepared at a 4:1 molecular ratio of biotinylated protein

to streptavidin, with every monomer labeled. Labeled streptavidin was added in ⅕ increments and in the dark at 4°C (rotating) for 20 min in between each addition. Optimal titers were determined using splenocytes from immunized mice and validated with SARS CoV-2 convalescent human PBMC.

### *Isolation of and sequencing of antibodies by single B cell sorting*

Cryopreserved human PBMCs from four COVID-19 convalescent donors were thawed and stained with Live/DEAD Fixable Aqua Dead Cell Stain kit (cat# L34957, ThermoFisher). After washing, cells were stained with a cocktail of anti-human antibodies, including CD3 (cat # 317332, Biolegend), CD8 (cat # 301048, Biolegend), CD56 (cat # 318340, Biolegend), CD14 (cat # 301842, Biolegend), CD19 (cat # IM2708U, Beckman Coulter), CD20 (cat # 302314, Biolegend), IgG (cat # 555786, BD Biosciences), IgA (cat # 130-114-001, Miltenyi), IgM (cat # 561285, BD Biosciences) and subsequently stained with fluorescently labeled SARS-CoV-2 S-2P (APC or Ax647), S1 (BV786 or BV570), RBD-SD1 (BV421) and NTD (BV711 or BV421) probes. Antigen-specific memory B cells ($CD3^-CD19^+CD20^+IgG^+$ or $IgA^+$ and $S\text{-}2P^+$ and/or RBD+ for the donors Subjects A19, A20 and A23, $S\text{-}2P^+$ and/or $NTD^+$ for the donor Subject B1) were sorted using a FACSymphony S6 (BD Sciences) into Buffer TCL (Qiagen) with 1% 2-mercaptoethanol (ThermoFisher Scientific). Nucleic acids were purified using RNAClean magnetic beads (Beckman Coulter) followed by reverse transcription using oligo-dT linked to a custom adapter sequence and template switching using SMARTScribe RT (Takara). PCR amplification was carried out using SeqAmp DNA Polymerase (Takara). A portion of the amplified cDNA was enriched for B cell receptor sequences using forward primers complementary to the template switch oligo and reverse primers against the IgA (GAGGCTCAGCGGGAAGACCTTGGGGCTGGTCGG) IgG, Igκ, and Igλ (*37*) constant regions. Enriched products were made into Illumina-ready sequencing libraries using the Nextera XT DNA Library Kit with Unique Dual Indexes (Illumina). The Illumina-ready libraries were sequenced by paired end 150 cycle MiSeq reads. The resulting reads were demultiplexed using an in-house script and V(D)J sequences were assembled using BALDR in unfiltered mode (*38*). Poor or incomplete assemblies or those with low read support were removed, and the filtered contigs were re-annotated with SONAR v4.2 in single cell mode (*39*). A subset of the final antibodies was manually selected for synthesis based on multiple considerations, including gene usage, SHM levels, CDRH3 length, convergent rearrangements, and specificity implied by flow cytometry.

### *Synthesis, cloning, and expression of monoclonal antibodies*

Sequences were selected for synthesis to sample expanded clonal lineages within our dataset and convergent rearrangements both among donors in our cohort and compared to the public literature. In addition, we synthesized a variety of sequences designed to be representative of the whole dataset along several dimensions, including apparent epitope based on flow data; V gene usage; SHM levels; CDRH3 length; and isotype. Variable heavy chain sequences were human codon optimized, synthesized and cloned into a VRC8400 (CMV/R expression vector)-based IgG1 vector containing an HRV3C protease site (*40*) as previously described (*36*). Similarly, variable lambda and kappa light chain sequences were human codon optimized, synthesized and cloned into CMV/R-based lambda or kappa chain expression vectors, as appropriate (Genscript). Previously published antibody vectors for LY-COV555(18) and mAb114 (*41*) were used. The antibodies: REGN10933 was produced from published sequences (*25*) and kindly provided by Devin Sok from Scripps. For antibodies where vectors were unavailable (e.g., S309, CB6), published amino acids sequences were used for synthesis and cloning into corresponding pVRC8400 vectors (*42*, *43*). For antibody expression, equal amounts of heavy and light chain plasmid DNA were transfected into Expi293 cells (Life Technology) by using Expi293 transfection reagent (Life Technology). The transfected cells were cultured in shaker incubator at 120 rpm, 37°C, 9% CO2 for 4~5 days. Culture supernatants were harvested and filtered, mAbs were purified over Protein A (GE Health Science) columns. Each antibody was eluted with IgG elution buffer (Pierce) and immediately neutralized with one tenth volume of 1M Tris-HCL pH 8.0. The antibodies were then buffer exchanged at least twice in PBS by dialysis.

### *ELISA method description*

Testing is performed using the automated enzyme-linked immunosorbent assay (ELISA) method as detailed in VRC-VIP SOP 5500 Automated ELISA on Integrated Automation System. Quantification of IgG concentrations in serum/plasma are performed with a Beckman Biomek based automation platform. The SARS-CoV-2 S-2P (VRC-SARS-CoV-2 S-2P (15-1208)-3C-His8-Strep2x2) and RBD (Ragon-SARS-CoV-2 S-RBD (319-529)-His8-SBP) Antigen are coated onto Immulon 4HBX flat bottom plates overnight for 16 hours at 4°C at a concentration of 2 μg/ml and 4μg/ml, respectively. Proteins were produced and generously provided by Dr. Dominic Esposito (Frederick National Laboratory for Cancer Research, NCI). Antigen concentrations were defined during assay development and antigen lot titration. Plates are washed and blocked (3% milk TPBS) for 1 hour at room temperature. Duplicate serial 4-fold dilutions covering the range of 1:100 to 1:1638400 (8-dilution series) of the test sample (diluted in 1%milk in TPBS) are incubated at room temperature for 2 hours followed by Horseradish Peroxidase - labeled goat anti-human antibody detection (1 hour at room temperature) (Thermo Fisher catalogue # A1881), and TMB substrate (15 min at room temperature; DAKO catalogue # S1599) addition. Color development is stopped by addition of sulfuric acid and plates are read within 30 min at 450 nm and 650 nm via the Molecular Devices Paradigm plate reader. Each plate harbors a negative control (assay diluent), positive control (SARS-CoV-2 S2-specific monoclonal antibody S-652-112 spiked in NHS and/or pool of COVID-19 convalescent sera) and batches of 5 specimen run in duplicates. All controls are trended over time.

Endpoint Titer dilution from raw OD data are interpolated using the plate background OD + 10 STDEV by asymmetric sigmoidal 5-pl curve fit of the test sample. In the rare event, the asymmetric sigmoidal 5-pl curve failed to interpolate the endpoint titer, a sigmoidal 4-pl curve is used for the analysis. Area under the curve (AUC) is calculated with baseline anchored by the plate background OD + 10 STDEV. Data analysis is performed using Microsoft Excel and GraphPad Prism Version 8.0.

### *Assignment of major binding determinant using MSD binding assay*

MSD 384-well streptavidin-coated plates (MSD, cat# L21SA) were blocked with MSD 5% Blocker A solution (MSD, cat# R93AA), using 35 ul per well. These plates were then incubated for 30 to 60 min at room temperature. Plates were washed with 1x Phosphate Buffered Saline + 0.05% Tween 20 (PBST) on a Biotek 405TS automated microplate washer. Five SARS CoV-2 capture antigens were used. Capture antigens consisted of VRC-produced S1, S-2P, S6P (Hexapro), RBD, and NTD. All antigens were AVI-tag biotinylated using BirA (Avidity, cat # BirA500) AVI-tag specific biotinylation following manufacturer's instructions except S1. For S1, an Invitrogen FluoReporter Mini-Biotin-XX Protein Labeling Kit (Thermo Fisher, cat # F6347) was utilized to achieve random biotinylation. Antigen coating solutions were prepared for S1, S-2P, S6P, RBD, and NTD at optimized concentrations of 0.5, 0.25, 1, 0.5, and 0.25 ug/ml, respectively. These solutions were then added to MSD 384-well plates, using 10 μl per well. Each full antigen set is intended to test one plate of experimental SARS CoV-2 monoclonal antibodies (mAbs) at one dilution. Once capture antigen solutions were added, plates were incubated for 1 hour at room temperature on a Heidolph Titramax 1000 (Heidolph, part # 544-12200-00) vibrational

plate shaker at 1000 rpm. During this time, experimental SARS CoV-2 mAb dilution plates were prepared. Using this initial plate, 3 dilution plates were created at dilution factors of 1:100, 1:1000, and 1:10000. Dilutions were performed in 1% assay diluent (MSD 5% Blocker A solution diluted 1:5 in PBST). Positive control mAbs S652-109 (SARS Cov-2 RBD specific) and S652-112 (SARS CoV-2 S1, S-2P, S6P, and NTD specific) and negative control mAb VRC01 (anti-HIV) were added to all dilution plates at a uniform concentration of 0.05 μg/ml. Once mAb dilution plates were prepared, MSD 384-well plates were washed as above. The content of each 96-well dilution plate was added to the MSD 384-well plates, using 10 μl per well. MSD 384-well plates were then incubated for 1 hour at room temperature on vibrational plate shaker at 1000 rpm. MSD 384-well plates were washed as above, and MSD Sulfo-Tag labeled goat anti-human secondary detection antibody (MSD, cat# R32AJ) solution was added to plates at a concentration of 0.5 ug/ml, using 10 μl per well. Plates were again incubated for 1 hour at room temperature on vibrational plate shaker at 1000 rpm. MSD 1x Read Buffer T (MSD, cat# R92TC) was added to MSD 384-well plates, using 35 μl per well. MSD 384-well plates were then read using MSD Sector S 600 imager. Gross binding epitope of S-2P or Hexapro positive antibodies was assigned into the following groups: RBD (i.e., $RBD^+$ or $RBD^+/S1^+$ AND $NTD^-$), NTD (i.e., $NTD^+$ or $NTD^+/S1^+$ and $RBD^-$), S2 (i.e., $S1^-$, $RBD^-$ and $NTD^-$) or indeterminant (i.e., mixed positive). Antibodies lacking binding to any of the antigens were assigned to the "no binding" group.

### Full-length S constructs

cDNAs encoding full-length S from SARS CoV-2 (GenBank ID: QHD43416.1) were synthesized, cloned into the mammalian expression vector VRC8400 (*42*, *43*) and confirmed by sequencing. S containing D614G amino acid change was generated using the wt S sequence. Other variants containing single or multiple aa changes in the S gene from the S wt or D614G were made by mutagenesis using QuickChange lightning Multi Site-Directed Mutagenesis Kit (cat # 210515, Agilent). The S variants, N439K, Y453F, A222V, E484K, K417N, S477N, N501Y, delH69/V70, N501Y-delH69/V70, N501Y-E484K-K417N, B.1.1.7 (H69del-V70del-Y144del-N501Y-A570D-P681H-T716I-S982A-D1118H), B.1.351.v1 (L18F-D80A-D215G-(L242-244)del-R246I-K417N-E484K-N501Y-A701V), B.1.351.v2 (L18F-D80A-D215G-(L242-244)del-K417N-E484K-N501Y-A701V), B.1.427 (L452R-D614G), B.1.429 (S13I-W152C-L452R-D614G), B.1.526.v2 (L5F-T95I-D253G-E484K-D614G-A701V), P.1.v1 (L18F-T20N-P26S-D138Y-R190S-K417T-E484K-N501Y-D614G-H655Y-T1027I), P.1.v2 (L18F-T20N-P26S-D138Y-R190S-K417T-E484K-N501Y-D614G-H655Y-T1027I-V7116F), P.2 (E484K-D614G-V7116F), B.1.617.1 (T95I-G412D-E154K-L452R-E484Q-D614G-P681R-Q1071H), B.1.617.2 (T19R-G142D-del156-157-R158G-L452R-T478K-D614G-P681R-D950N) and antibody escape mutations, F486S, K444E, Y449S, N450S and F490V were generated based on S D614G while the antibody contact residue mutations, F456R, A475R, T478I, F486R, Y489R, N487R, L452R, F490L, Q493R, S494R on S wt. These full-length S plasmids were used for pseudovirus production and for cell surface binding assays.

### Pseudovirus neutralization assay

S-containing lentiviral pseudovirions were produced by co-transfection of packaging plasmid pCMVdR8.2, transducing plasmid pHR' CMV-Luc, a TMPRSS2 plasmid and S plasmids from SARS CoV-2 variants into 293T (ATCC) cells using Fugene 6 transfection reagent (Promega, Madison, WI) (*44*–*46*). 293T-ACE2 cells, provided by Dr. Michael Farzan, were plated into 96-well white/black Isoplates (PerkinElmer, Waltham, MA) at 5000 cells per well the day before infection of SARS CoV-2 pseudovirus. Serial dilutions of mAbs were mixed with titrated pseudovirus, incubated for 45 min at 37°C and added to 293T-ACE2 cells in triplicate. Following 2 hours of incubation, wells were replenished with 150 ml of fresh media. Cells were lysed 72 hours later, and luciferase activity was measured with Microbeta (Perking Elmer). Percent neutralization and neutralization IC50s, IC80s were calculated using GraphPad Prism 8.0.2. Serum neutralization assays were performed as above excepting all human sera had an input starting serial dilution of 1:20 and neutralization was quantified as the inhibition dilution 50% ($ID_{50}$) of virus entry. Alternative method pseudovirus neutralization assay in fig. S3 utilized a first-generation lentivirus system and was performed as in Wibmer *et al.* (*12*).

### Cell surface binding

Human embryonic kidney (HEK) 293 T cells were transiently transfected with plasmids encoding full length SARS CoV-2 spike variants using lipofectamine 3000 (L3000-001, ThermoFisher) following manufacturer's protocol. After 40 hours, the cells were harvested and incubated with monoclonal antibodies (1 μg/ml) for 30 min. After incubation with the antibodies, the cells were washed and incubated with an allophycocyanin conjugated anti-human IgG (709-136-149, Jackson Immunoresearch Laboratories) for another 30 min. The cells were then washed and fixed with 1% paraformaldehyde (15712-S, Electron Microscopy Sciences). The samples were then acquired in a BD LSRFortessa X-50 flow cytometer (BD biosciences) and analyzed using Flowjo (BD biosciences). Mean fluorescent intensity (MFI) for antibody binding to S wt or D614G was set up as 100%. The MFI of the antibody binding to each variant was normalized to S wt or D614G.

### Competitive mAb binding assay using surface plasmon resonance

Monoclonal antibody (mAb) competition assays were performed on a Biacore 8K+ (Cytiva) surface plasmon resonance spectrometer. Anti-histidine $IgG_1$ antibody was immobilized on Series S Sensor Chip CM5 (Cytiva) using a His capture kit (Cytiva), per manufacturer's instructions. 1X PBS-P+ (Cytiva) was used for running buffer and diluent, unless noted. 8X His-tagged SARS-CoV-2 Spike protein containing 2 proline stabilization mutations, K986P and V987P, (S-2P) (*4*) was captured on the active sensor surface. "Competitor" mAb or a negative control mAb114 (*37*) were first injected over both active and reference surfaces, followed by "analyte" mAb. Between cycles, sensor surfaces were regenerated with 10 mM glycine, pH 1.5 (Cytiva).

For data analysis, sensorgrams were aligned to Y (Response Units, RUs) = 0, beginning at the beginning of each mAb binding phase in Biacore 8K Insights Evaluation Software (Cytiva). Reference-subtracted, relative "analyte binding late" report points (in RUs) were used to determine percent competition for each mAb. Maximum analyte binding for each mAb was first defined by change in RUs during analyte binding phase when negative control mAb was used as competitor mAb. Percent competition (%C) was calculated using the following formula: %C = 100 * {1 –[((analyte mAb binding RUs when S-2P-specific mAb is used as competitor) / (maximum analyte binding RUs when negative control mAb is used as competitor)]}.

### Competitive ACE2 binding assay using biolayer interferometry

Antibody cross-competition was determined based on biolayer interferometry using a fortéBio Octet HTX instrument. His1K biosensors (fortéBio) were equilibrated for >600 s in Blocking Buffer [1% BSA (Sigma) + 0.01% Tween-20 (Sigma) + 0.01% Sodium Azide (Sigma) + PBS (Gibco), pH7.4] prior to loading with his tagged S-2P protein (10 μg/ml in Blocking Buffer) for 1200s. Following loading, sensors were incubated for 420s in Blocking Buffer prior to incubation with competitor mAbs (30 mg/ml in Blocking Buffer) or ACE2 (266 nM in Blocking Buffer) for 1200s. Sensors were then incubated in Blocking buffer for 30s prior to incubation with ACE2 (266 nM in Blocking Buffer) for 1200s. Percent competition (PC) of ACE2 mAbs binding to competitor-bound S-2P was determined using the equation: PC = 100 – [(ACE2 binding in the presence competitor mAb)/(ACE2 binding in the absence of competitor mAb)] × 100. All the assays were performed in duplicate and with agitation set to 1000 rpm at 30°C.

### Inhibition of S protein binding to cell surface ACE2

Serial dilutions of mAb were mixed with pretitrated biotinylated S trimer (S-2P), incubated for 30 min at RT and added to BHK21 cells stably expressing hACE2 on cell surface. Following 30 min of incubation on ice, the cells were washed and incubated with an BV421 conjugated Streptavidin (cat # 563259, BD Biosciences) for another 30 min. The cells were then washed and fixed with 1% paraformaldehyde (15712-S, Electron Microscopy Sciences). The samples were then acquired in a BD LSRFortessa X-50 flow cytometer (BD biosciences) and analyzed using Flowjo (BD biosciences). MFI for S protein binding to cell surface was set up as 100%. Percent inhibition of S protein binding to cell surface ACE2 by mAb IgG and half-maximal effective concentration ($EC_{50}$) were calculated using GraphPad Prism 8.0.2.

### Live virus neutralization assay

Full-length SARS CoV-2 virus based on the Seattle Washington strain was designed to express nanoluciferase (nLuc) and was recovered via reverse genetics and described previously (17). Virus titers were measured in Vero E6 USAMRIID cells, as defined by plaque forming units (PFU) per ml, in a 6-well plate format in quadruplicate biological replicates for accuracy. For the 96-well neutralization assay, Vero E6 USAMRID cells were plated at 20,000 cells per well the day prior in clear bottom black walled plates. Cells were inspected to ensure confluency on the day of assay. Serially diluted mAbs were mixed in equal volume with diluted virus. Antibody-virus and virus only mixtures were then incubated at 37°C with 5% $CO_2$ for one hour. Following incubation, serially diluted mAbs and virus only controls were added in duplicate to the cells at 75 PFU at 37°C with 5% $CO_2$. After 24 hours, cells were lysed, and luciferase activity was measured via Nano-Glo Luciferase Assay System (Promega) according to the manufacturer specifications. Luminescence was measured by a Spectramax M3 plate reader (Molecular Devices, San Jose, CA). Virus neutralization titers were defined as the sample dilution at which a 50% reduction in RLU was observed relative to the average of the virus control wells.

Live virus neutralization assays described above were performed with approved standard operating procedures for SARS CoV-2 in a biosafety level 3 (BSL-3) facility conforming to requirements recommended in the Microbiological and Biomedical Laboratories, by the US Department of Health and Human Service, the US Public Health Service, and the US Center for Disease Control and Prevention (CDC), and the National Institutes of Health (NIH).

### Production of Fab fragments from monoclonal antibodies

To generate mAb-Fab, IgG was incubated with HRV3C protease (EMD Millipore) at a ratio of 100 units per 10 mg IgG with HRV 3C Protease Cleavage Buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5) at 4°C overnight. Fab was purified by collecting flowthrough from Protein A column (GE Health Science), and Fab purity was confirmed by SDS-PAGE.

### Determination of binding kinetics of Fab

A fortéBio Octet HTX instrument was used to measure binding kinetics of the Fab of A23-58.1, B1-182.1, A19-46.1 and A19-61.1 to SARS CoV-2 S-2P protein. SA biosensors (fortéBio) were equilibrated for >600 s in Blocking Buffer [1% BSA (Sigma) + 0.01% Tween-20 (Sigma) + 0.01% Sodium Azide (Sigma) + PBS (Gibco), pH7.4] prior to loading with biotinylated S-2P protein (1.5 mg/ml in Blocking Buffer) for 600s. Following loading, sensors were incubated for 420s in Blocking Buffer prior to binding assessment of the Fabs. Association of Fabs was measured for 300 s and dissociation was measured for up to 3600 s in Blocking Buffer. All the assays were performed with agitation set to 1000 rpm at 30°C. Data analysis and curve fitting were carried out using Octet analysis software, version 11-12. Experimental data were fitted using a 1:1 binding model. Global analyses of the complete data sets assuming binding was reversible (full dissociation) were carried out using nonlinear least-squares fitting allowing a single set of binding parameters to be obtained simultaneously for all concentrations used in each experiment.

### Negative-stain electron microscopy.

Protein samples were diluted to a concentration of approximately 0.02 mg/ml with 10 mM HEPES, pH 7.4, supplemented with 150 mM NaCl. A 4.8-µl drop of the diluted sample was placed on a freshly glow-discharged carbon-coated copper grid for 15 s. The drop was then removed with filter paper, and the grid was washed with three drops of the same buffer. Protein molecules adsorbed to the carbon were negatively stained by applying consecutively three drops of 0.75% uranyl formate, and the grid was allowed to air-dry. Datasets were collected using a Thermo Scientific Talos F200C transmission electron microscope operated at 200 kV and equipped with a Ceta camera. The nominal magnification was 57,000x, corresponding to a pixel size of 2.53 Å, and the defocus was set at -1.2 µm. Data was collected automatically using EPU. Single particle analysis was performed using CryoSPARC (47).

### Cryo-EM specimen preparation and data collection.

The stabilized SARS CoV-2 spike HexaPro (3) was mixed with Fab A23-58.1 or B1-182.1 at a molar ratio of 1.2 Fab per protomer in PBS. The final spike protein concentration was 0.5 mg/ml. n-Dodecyl β-D-maltoside (DDM) detergent was added shortly before vitrification to a concentration of 0.005%. Quantifoil R 2/2 gold grids were subjected to glow discharging in a PELCO easiGlow device (air pressure, 0.39 mBar; current, 20 mA; duration, 30 s) immediately before specimen preparation. Cryo-EM grids were prepared using an FEI Vitrobot Mark IV plunger with the following settings: chamber temperature of 4°C, chamber humidity of 95%, blotting force of –5, blotting time of 3 s, and drop volume of 2.7 µl. Datasets were collected at the National CryoEM Facility (NCEF), National Cancer Institute, on a Thermo Scientific Titan Krios G3 electron microscope equipped with a Gatan Quantum GIF energy filter (slit width: 20 eV) and a Gatan K3 direct electron detector (table S2). Four movies per hole were recorded in the counting mode using Latitude software. The dose rate was 14.65 $e^-$/s/pixel.

### Cryo-EM data processing and model fitting

Data process workflow, including Motion correction, CTF estimation, particle picking and extraction, 2D classification, ab initio reconstruction, homogeneous refinement, heterogeneous refinement, non-uniform refinement, local refinement and local resolution estimation, were carried out with C1 symmetry in cryoSPARC 2.15 (*47*) For local refinement to resolve the RBD-antibody interface, a mask for the entire spike-antibody complex without the RBD-antibody region was used to extract the particles and a mask encompassing the RBD-antibody region was used for refinement. The overall resolution was 3.39 Å and 3.15 Å for the map of A23-58.1- and B1-182.1-bound spike, 3.89 Å and 3.71 Å for the map of RBD: antibody interface after local refinement, respectively. The coordinates for the SARS-CoV-2 spike with three ACE2 molecules bound at pH 7.4 (PDB ID: 7KMS) were used as initial models for fitting the cryo-EM map. Iterative manual model building and real space refinement were carried out in Coot (*48*) and in Phenix (*49*), respectively. Molprobity (*50*) was used to validate geometry and check structure quality at each iteration step. UCSF Chimera and ChimeraX were used for map fitting and manipulation (*51*).

### Selection of rcVSV SARS CoV-2 virus escape variants using monoclonal antibodies

A replication competent vesicular stomatitis virus (rcVSV) with its native glycoprotein replaced by the Wuhan-1 spike protein (rcVSV SARS CoV-2) that contains a 21 amino acid deletion at the C-terminal region (*32*) (generous gift of Kartik Chandran and Rohit Jangra). Passage 7 virus was passaged twice on Vero cells to obtain a polyclonal stock. A single

plaque from this 9th passage was double plaque purified and expanded on Vero cells to create monoclonal virus population. The reference genome for this stock was sequence using Illumina-based sequencing as described below.

To select for virus escape variants, an equal volume of clonal population of rcVSV SARS CoV-2 was mixed with serial dilutions of antibodies (5-fold) in DMEM supplemented with 10% FCS and Glutamine to give an MOI of 0.1 to 0.001 at the desired final antibody concentration (range $5.1e^{-6}$ to 50 μg/ml and 0 μg/ml). Virus:antibody mixtures were incubated at room temperature for 1 hour. After incubation, 300 μl of virus:antibody mixtures were added to $1 \times 10^5$ Vero E6 cells in 12 well plates for 1 hour at 37°C, 5% $CO_2$. The plates were rotated every 15 min to prevent drying. After absorption, 700 μl of additional antibodies mixture was added to each well at their respective concentration. Cells were incubated for 72hrs at 37°C, 5% $CO_2$. Virus replication was monitored using cytopathic effect and supernatant was collected from the wells with cytopathic effect. Harvested supernatant was clarified by centrifugation at 3750rpm for 10 min. For the subsequent rounds of selection, clarified supernatant from the well with the highest concentration of antibody that has CPE >20% supernatant was diluted prior to being mixed with equal volume of antibodies as in the initial round of selection. Infection, monitoring and collection of supernatants was performed as in the initial round.

### *Shotgun sequencing of rcVSV SARS CoV2 supernatants*

Total RNA was extracted from clarified supernatants using QIAmp viral RNA mini extraction kit (Qiagen) following the manufacturer's recommended protocol. Purified RNA was fragmented using NEBNext Ultra II RNA Library Prep reagents, then reverse transcribed using random hexamers, and double-stranded cDNA was synthesized (New England BioLabs) as previously described (*52*). Double-stranded cDNA was purified using magnetic beads (MagBio Genomics) and barcoded Illumina-ready libraries were subsequently prepared (New England BioLabs). The libraries were sequenced as paired-end 2x150 base pair NextSeq 2000 reads.

### *Spike SNP variant calls of rcVSV antibody induced revertants*

Raw sequencing reads were demultiplexed and trimmed to remove adaptor sequences and low quality bases. They were then aligned against the reference viral genome with Bowtie (v2.4.2). Single nucleotide polymorphisms (SNPs) were called using HaplotypeCaller from the Genome Analysis Tool Kit (GATK, v4.1.9.0). The HaplotypeCaller parameter, "–sample-ploidy", was set to 100 in order to identify SNPs with a prevalence of at least 1%. SNPs for all samples were then aggregated, interrogated and translated using custom scripts. A SNP and correlated amino acid translation for the spike protein was considered positive if it was present at a frequency of greater than 0.1 (10%) and showed an increasing frequency from round 1 to round 2 of the antibody selections.

### *Multiplex SARS-CoV-2 variant binding assay*

Multiplexed Plates (96 well) precoated with SARS Cov2 spike (WA-1), SARS Cov2 RBD (WA-1), SARS Cov2 spike (B.1.351), SARS Cov2 spike (B.1.1.7), SARS Cov2 spike (P.1), SARS Cov2 RBD (B.1.351), SARS Cov2 RBD (B.1.1.7), SARS Cov2 RBD (P.1) and BSA are supplied by the manufacturer. On the day of the assay, the plate is blocked for 60 min with MSD Blocker A (5% BSA). The blocking solution is washed off and test samples are applied to the wells at 4 dilution (1:100, 1:500, 1:2500, and 1:10,000) unless otherwise specified and allowed to incubate with shaking for two hours. Plates are washed and Sulfo-tag labeled anti IgG antibody is applied to the wells and allowed to associate with complexed coated antigen – sample antibody within the assay wells. Plates are washed to remove unbound detection antibody. A read solution containing ECL substrate is applied to the wells, and the plate is entered into the MSD Sector instrument. A current is applied to the plate and areas of well surface where sample antibody has complexed with coated antigen and labeled reporter will emit light in the presence of the ECL substrate. The MSD Sector instrument quantitates the amount of light emitted and reports this ECL unit response as a result for each sample and standard of the plate. Magnitude of ECL response is directly proportional to the extent of binding antibody in the test article. All calculations are performed within Excel and the GraphPad Prism software, version 7.0. Readouts are provided as AUC.

## REFERENCES AND NOTES


1. Johns Hopkins Coronavirus Resource Center, COVID-19 Map; https://coronavirus.jhu.edu/map.html.
2. F. Wu *et al*., A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020). doi: 10.1038/s41586-020-2008-3; pmid: 32015508
3. C. L. Hsieh *et al*., Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020). doi: 10.1126/science.abd0826; pmid: 32703906
4. D. Wrapp *et al*., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020). doi: 10.1126/science.abb2507; pmid: 32075877
5. T. Zhou *et al*., Structure-based design with Tag-based purification and in-process biotinylation enable streamlined development of SARS-CoV-2 spike molecular probes. *Cell Rep.* **33**, 108322 (2020). doi: 10.1016/j.celrep.2020.108322; pmid: 33091382
6. A. Rambaut, N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D. L. Robertson, E. Volz, COVID-19 Genomics Consortium UK, Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations (2020); https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563.
7. H. Tegally *et al*., Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021). doi: 10.1038/s41586-021-03402-9; pmid: 33690265
8. N. R. Faria *et al*., Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings (2021); https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586.
9. F. Naveca *et al*., Phylogenetic relationship of SARS-CoV-2 sequences from Amazonas with emerging Brazilian variants harboring mutations E484K and N501Y in the Spike protein (2021); https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585.
10. A. P. Horby *et al*., Paper from the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) on new coronavirus (COVID-19) variant B.1.1.7. (2021); www.gov.uk/government/publications/nervtag-update-note-on-b117-severity-11-february-2021.
11. Y. J. Hou *et al*., SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **370**, 1464–1468 (2020). doi: 10.1126/science.abe8499; pmid: 33184236
12. C. K. Wibmer *et al*., SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* **27**, 622–625 (2021). doi: 10.1038/s41591-021-01285-x; pmid: 33654292
13. R. F. Garry, Mutations arising in SARS-CoV-2 spike on sustained human-to-human transmission and human-to-animal passage (2021); https://virological.org/t/mutations-arising-in-sars-cov-2-spike-on-sustained-human-to-human-transmission-and-human-to-animal-passage/578.
14. 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
15. A. Muik *et al*., Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine-elicited human sera. *Science* **371**, 1152–1153 (2021). doi: 10.1126/science.abg6105; pmid: 33514629
16. Z. Wang *et al*., mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616–622 (2021). doi: 10.1038/s41586-021-03324-6; pmid: 33567448
17. Y. J. Hou *et al*., SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, 429–446.e14 (2020). doi: 10.1016/j.cell.2020.05.042; pmid: 32526206
18. C. O. Barnes *et al*., SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020). doi: 10.1038/s41586-020-2852-1; pmid: 33045718
19. B. E. Jones *et al*., The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in nonhuman primates. *Sci. Transl. Med.* **13**, eabf1906 (2021). doi: 10.1126/scitranslmed.abf1906; pmid: 33820835
20. P. Chen *et al*., SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with COVID-19. *N. Engl. J. Med.* **384**, 229–237 (2021). doi: 10.1056/NEJMoa2029849; pmid: 33113295
21. D. Pinto *et al*., Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020). doi: 10.1038/s41586-020-2349-y; pmid: 32422645
22. B. Korber *et al*., Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827.e19 (2020). doi: 10.1016/j.cell.2020.06.043; pmid: 32697968
23. E. C. Thomson *et al*., Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171–1187.e20 (2021). doi: 10.1016/j.cell.2021.01.037; pmid: 33621484
24. R. Shi *et al*., A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020). doi: 10.1038/s41586-020-2381-y; pmid: 32454512
25. J. Hansen *et al*., Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010–1014 (2020). doi: 10.1126/science.abd0827; pmid: 32540901
26. A. Baum *et al*., Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014–1018 (2020). doi: 10.1126/science.abd0831; pmid: 32540904
27. M. A. Tortorici *et al*., Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **370**, 950–957 (2020). doi: 10.1126/science.abe3354; pmid: 32972994

28. D. F. Robbiani *et al*., Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020). doi: 10.1038/s41586-020-2456-9; pmid: 32555388
29. S. J. Zost *et al*., Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443–449 (2020). doi: 10.1038/s41586-020-2548-6; pmid: 32668443
30. W. Dejnirattisai *et al*., The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell* **184**, 2183–2200.e22 (2021). doi: 10.1016/j.cell.2021.02.032; pmid: 33756110
31. J. Dong *et al*., Genetic and structural basis for recognition of SARS-CoV-2 spike protein by a two-antibody cocktail. bioRxiv [Preprint] 1 March 2021). .doi: 10.1101/2021.01.27.428529
32. M. E. Dieterle *et al*., A replication-competent vesicular stomatitis virus for studies of SARS-CoV-2 spike-mediated cell entry and its inhibition. *Cell Host Microbe* **28**, 486–496.e6 (2020). doi: 10.1016/j.chom.2020.06.020; pmid: 32738193
33. C. G. Rappazzo *et al*., Broad and potent activity against SARS-like viruses by an engineered human monoclonal antibody. *Science* **371**, 823–829 (2021). doi: 10.1126/science.abf4830; pmid: 33495307
34. C. O. Barnes *et al*., Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* **182**, 828–842.e16 (2020). doi: 10.1016/j.cell.2020.06.025; pmid: 32645326
35. X. Shen *et al*., SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral spike vaccines. *Cell Host Microbe* **29**, 529–539.e3 (2021). doi: 10.1016/j.chom.2021.03.002; pmid: 33705729
36. T. Moyo-Gwete *et al*., Cross-reactive neutralizing antibody responses elicited by SARS-CoV-2 501Y.V2 (B.1.351). *N. Engl. J. Med.* **384**, 2161–2163 (2021). pmid: 33826816
37. S. J. Krebs *et al*., Longitudinal analysis reveals early development of three MPER-directed neutralizing antibody lineages from an HIV-1-infected individual. *Immunity* **50**, 677–691.e13 (2019). doi: 10.1016/j.immuni.2019.02.008; pmid: 30876875
38. A. A. Upadhyay *et al*., BALDR: A computational pipeline for paired heavy and light chain immunoglobulin reconstruction in single-cell RNA-seq data. *Genome Med.* **10**, 20 (2018). doi: 10.1186/s13073-018-0528-3; pmid: 29558968
39. C. A. Schramm *et al*., SONAR: A high-throughput pipeline for inferring antibody ontogenies from longitudinal sequencing of B cell transcripts. *Front. Immunol.* **7**, 372 (2016). doi: 10.3389/fimmu.2016.00372; pmid: 27708645
40. J. S. McLellan *et al*., Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343 (2011). doi: 10.1038/nature10696; pmid: 22113616
41. J. Misasi *et al*., Structural and molecular basis for Ebola virus neutralization by protective human antibodies. *Science* **351**, 1343–1346 (2016). doi: 10.1126/science.aad6117; pmid: 26917592
42. D. H. Barouch *et al*., A human T-cell leukemia virus type 1 regulatory element enhances the immunogenicity of human immunodeficiency virus type 1 DNA vaccines in mice and nonhuman primates. *J. Virol.* **79**, 8828–8834 (2005). doi: 10.1128/JVI.79.14.8828-8834.2005; pmid: 15994776
43. A. T. Catanzaro *et al*., Phase I clinical evaluation of a six-plasmid multiclade HIV-1 DNA candidate vaccine. *Vaccine* **25**, 4085–4092 (2007). doi: 10.1016/j.vaccine.2007.02.050; pmid: 17391815
44. L. Naldini, U. Blömer, F. H. Gage, D. Trono, I. M. Verma, Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11382–11388 (1996). doi: 10.1073/pnas.93.21.11382; pmid: 8876144
45. Z. Y. Yang *et al*., Evasion of antibody neutralization in emerging severe acute respiratory syndrome coronaviruses. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 797–801 (2005). doi: 10.1073/pnas.0409065102; pmid: 15642942
46. L. Wang *et al*., Evaluation of candidate vaccine approaches for MERS-CoV. *Nat. Commun.* **6**, 7712 (2015). doi: 10.1038/ncomms8712; pmid: 26218507
47. 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; pmid: 28165473
48. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi: 10.1107/S0907444904019158; pmid: 15572765
49. P. V. Afonine *et al*., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012). doi: 10.1107/S0907444912001308; pmid: 22505256
50. I. W. Davis, L. W. Murray, J. S. Richardson, D. C. Richardson, MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615-9 (2004). doi: 10.1093/nar/gkh398; pmid: 15215462
51. E. F. Pettersen *et al*., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: 10.1002/jcc.20084; pmid: 15264254
52. D. Ssemwanga *et al*., Characterization of the neutralizing antibody response in a case of genetically linked HIV superinfection. *J. Infect. Dis.* **217**, 1530–1534 (2018). doi: 10.1093/infdis/jiy071; pmid: 29579256


## ACKNOWLEDGMENTS

We thank the staff of the Clinical Trials Program of the Vaccine Research Center and the volunteers that made this research possible. We also appreciate the assistance of R. Hunegnaw for assistance with figure preparation. We are grateful to T. L. Fox and T. J. Edwards of NCEF for collecting cryo-EM data and for technical assistance with cryo-EM data processing. **Funding:** This work was funded by the intramural research program of the Vaccine Research Center, NIAID, NIH. D.R.M. is funded by a Burroughs Wellcome Fund Postdoctoral Enrichment Program Award and a Hanna H. Gray Fellowship from the Howard Hughes Medical Institute and was supported by an NIH NIAID T32 AI007151 and an NIH F32 AI152296. The manuscript was also supported in part by an NIH R01 AI157155 to RSB. This research was, in part, supported by the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E. **Author contributions:** T.J.R., E.P., A.T.D., N.A.D.-R., R.D.M., J.M., C.A.S., L.W., K.S.C., and E.M.C. designed and performed cell sorting experiments. A.R.H., A.A., R.L.D., F.L., S.D., and C.A.S. performed and analyzed sequencing data. Proteins, antibodies, and other reagents were produced by W.S., I-T.T., L.W., T.Z., A.S.O., E.P., T.J.R., J.M., O.M.A., L.A.C., A.T.D., E.S.Y., Y.Z., B.Z., A.F.N., and T.L., and J.M., L.W., T.Z., Y.T., T.S., Y.Z., W.S., E.S.Y., A.P., C.A.T., O.K.O., C.A.S., S.D., S.R.N., C.H., D.R.M., M.C., S.H.H., T.H., P.K., K.L., T.L., S.O'C., S.O'D., S.D.S., C.D.S., and D.A.W. conceived of, designed, and performed experiments, data analysis and reporting. M.R.G., A.T.W., L.N., and I.J.G. performed research subject recruitment, collection of samples, and maintenance of the sample repository. T.Z. and Y.T. led electron microscopy studies. J.M., N.J.S., J.R.M., D.C.D., B.S.G., A.B.M., P.D.K., J.E.L., M.R., N.A.D.-R., P.L.M., and R.S.B. supervised experiments. J.M., N.J.S., T.Z., L.W., and C.A.S. wrote the manuscript, with help from all authors. **Competing interests:** J.R.M., L.W., C.A.S., J.R.M, D.C.D., N.J.S., A.R.H., T.Z., P.D.K., W.S., Y.Z., E.S.Y., M.R., R.D.M., and A.P. are inventors on US patent application no. 63/147,419. **Data and materials availability:** All data are available in the main text or the supplementary materials. Atomic coordinates and cryo-EM maps of the reported structure have been deposited into the Protein Data Bank (PDB) and Electron Microscopy Data Bank (EMD) under the session codes PDB 7LRT and EMD-23499 for SARS-CoV-2 spike in complex with antibody A23-58.1, PDB 7LRS and EMD-23498 for local refinement of the RBD-antibody A23-58.1 region, PDB 7MM0 and EMD-23915 for SARS-CoV-2 spike in complex with antibody B1-182.1, and PDB 7MLZ and EMD-23914 for local refinement of the RBD-antibody B1-182.1 region. Antibody DNA sequences have been deposited in GenBank with the following accession numbers: MZ458523 for A1946.1_HC, MZ458524 for A19-46.1_lc, MZ458525 for A19-61.1_HC, MZ458526 for A19-61.1_kC, MZ458527 for A23-58.1_HC, MZ458528 for A23-58.1_kC, MZ458529 for B1-182.1_HC, and MZ458530 for B1-182.1_kC. Original materials in this manuscript are available from N.J.S. under a materials transfer agreement with the National Institutes of Health. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/eabh1766/suppl/DC1
Figs. S1 to S10
Tables S1 to S3
MDAR Reproducibility Checklist

20 February 2021; accepted 28 June 2021
10.1126/science.abh1766

RESEARCH ARTICLES

CANCER GENOMICS

# Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition

Karol Nowicki-Osuch[1]†‡, Lizhe Zhuang[1]†, Sriganesh Jammula[2], Christopher W. Bleaney[3], Krishnaa T. Mahbubani[4,5], Ginny Devonshire[2], Annalise Katz-Summercorn[1], Nils Eling[2,6], Anna Wilbrey-Clark[7], Elo Madissoon[7], John Gamble[4,8], Massimiliano Di Pietro[1], Maria O'Donovan[1], Kerstin B. Meyer[7], Kourosh Saeb-Parsy[4,8], Andrew D. Sharrocks[3], Sarah A. Teichmann[7,9], John C. Marioni[2,6,7], Rebecca C. Fitzgerald[1]*

The origin of human metaplastic states and their propensity for cancer is poorly understood. Barrett's esophagus is a common metaplastic condition that increases the risk for esophageal adenocarcinoma, and its cellular origin is enigmatic. To address this, we harvested tissues spanning the gastroesophageal junction from healthy and diseased donors, including isolation of esophageal submucosal glands. A combination of single-cell transcriptomic profiling, in silico lineage tracing from methylation, open chromatin and somatic mutation analyses, and functional studies in organoid models showed that Barrett's esophagus originates from gastric cardia through c-MYC and HNF4A-driven transcriptional programs. Furthermore, our data indicate that esophageal adenocarcinoma likely arises from undifferentiated Barrett's esophagus cell types even in the absence of a pathologically identifiable metaplastic precursor, illuminating early detection strategies.

Metaplasia is usually associated with an increased risk of malignancy and is thought to result from a transcriptional switch within existing cells or an outgrowth of minor cell types, yet its specific origin is often unclear (*1*). It commonly occurs at sites where different epithelial types meet, such as squamo-columnar junctions (SCJ) (*2*). Barrett's esophagus (BE) is an archetypal metaplastic condition comprising a mosaic of gastric and intestinal cell types (*3*, *4*). BE occurs in up to 10% of individuals with gastric reflux and starts at the gastroesophageal junction (GEJ), causing the SCJ to be displaced proximally (*5*). It increases the propensity for esophageal adenocarcinoma (EAC), which has an overall 5-year survival rate of 15% (*6*, *7*). Given that the native esophagus is squamous, the glandular phenotype of EAC is thought to be inextricably linked to BE (*6*). However, about 50% of EAC patients do not have evidence of BE at the time of diagnosis (*8*), calling this current dogma into question.

The controversy could be resolved by determining the origin of BE, which has been hypothesized to originate from many sources, including esophageal submucosal glands (SMG) or various specific cell populations at the GEJ (*9*–*16*) (fig. S1). One major impediment to research is that mouse models used for lineage tracing do not fully resemble human gastroesophageal physiology owing to a keratinized squamous forestomach and a lack of SMG. Furthermore, isolation of SMG is particularly challenging in fresh human tissue. The overall aims of this study were therefore to molecularly characterize all putative cell origins for BE and determine whether all EAC subtypes are derived from BE.

## Results

### *Determining the cellular components of human SMG*

SMG have been studied as the origin of BE, yet never specifically isolated. Using stereomicroscopy (fig. S2 and movie S1), we isolated fresh SMGs from human esophagus. We performed immunostaining with cell markers of CDH1 (pan-epithelial), KRT5 (squamous), KRT8 (columnar), and KRT7 (SMG) (*17*) and used three-dimensional (3D) confocal microscopy to identify all the cellular components: duct cells, oncocytes, mucous cells, and myoepithelial cells (Fig. 1A, fig. S3, and movies S2 and S3) (*18*).

In both the intercalated and main duct of SMG, we observed a P63$^+$KRT5$^+$KRT7$^+$ cell population (Fig. 1B) that resembles the transitional basal progenitors, previously reported to generate BE and reside on the surface of SCJ (*15*, *18*). We also found that oncocytes, characterized by eosinophilic cytoplasm and centrally located nuclei, are common in disease-free donors and sometimes form their own acini (fig. S4, A and B). This is contrary to the literature, which suggests that they are associated with gastroesophageal reflux disease and BE (*19*).

Next, to characterize all cell types within the SMG, we dissociated fresh SMG into single cells and performed single-cell RNA sequencing (scRNA-seq) (*20*). We identified four major epithelial cell types (Fig. 1, C and D, and tables S1 and S2) expressing transcripts previously not associated with SMG, including *MUC5B*, *KRT23*, and *AGR2*, which were confirmed by immunostaining (fig. S4C). Few (<0.1%) SMG cells express the cell division marker *MKI67*, indicating their quiescent status, and contrary to a previous scRNA-seq study (*11*), no significant *OLFM4* was detected (Fig. 1D and fig. S5A). A comparison of our SMG profile with esophageal cells from the Human Cell Atlas (HCA) project (*21*) and a recent SMG study (*11*) demonstrated the sparse nature of SMG in esophageal biopsies. Only a very small fraction of cells in the HCA project, and none of the cells from the previous SMG study (*11*), showed SMG phenotypes (fig. S5, B to E) (*20*).

### *A KRT7$^{high}$* population of SCJ and its lineage relationship with SMG

The transitional basal cells and residual embryonic cells of the GEJ have also been proposed as the origin for BE (*14*, *15*). We collected normal squamo-columnar junction (N-SCJ) tissue from eight disease-free donors (fig. S6) and performed scRNA-seq to study its cellular composition. In addition to the expected esophageal squamous and gastric columnar cells, we discovered a small distinct *KRT7$^{high}$* population (Fig. 2, A and B), which consisted of P63$^+$KRT5$^+$KRT7$^+$ transitional basal cells (C1 and C2), KRT7$^+$MUC4$^+$ residual embryonic cells (C3), and a *MUC5B$^{high}$* cell type (C4) (Fig. 2, C and D, and table S3). Immunostaining showed that MUC5B is expressed in the suprabasal layer and is in some cases independent from the transitional basal P63$^+$KRT5$^+$KRT7$^+$ cells (Fig. 2, E and F).

When the single-cell transcriptome profiles of N-SCJ cells were compared to those of cells from normal esophagus (NE), normal gastric cardia (NG) (*22*), and SMG from disease-free donors, the *KRT7$^{high}$* population from N-SCJ exhibited the highest similarity to SMG cells with respect to cell types present in our reference set (Figs. 2, G and H, and fig. S7A). Overall, this suggests that the transitional basal progenitors, residual embryonic cells, and the *MUC5B$^{high}$* cells might belong to a single

[1]Medical Research Council Cancer Unit, Hutchison/Medical Research Council Research Centre, University of Cambridge, Cambridge CB2 0X2, UK. [2]Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge CB2 0RE, UK. [3]Faculty of Biology, Medicine and Health, Michael Smith Building, Oxford Road, University of Manchester, Manchester, UK. [4]Cambridge Biorepository for Translational Medicine (CBTM), NIHR Cambridge Biomedical Research Centre, Cambridge, UK. [5]Department of Haematology, University of Cambridge, Cambridge, UK. [6]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [7]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [8]Department of Surgery, University of Cambridge, Cambridge, UK. [9]Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge CB3 0HE, UK.
*Corresponding author. Email: rcf29@cam.ac.uk
†These authors contributed equally to this work.
‡Present address: Irving Institute for Cancer Dynamics, Columbia University, New York, NY, USA.

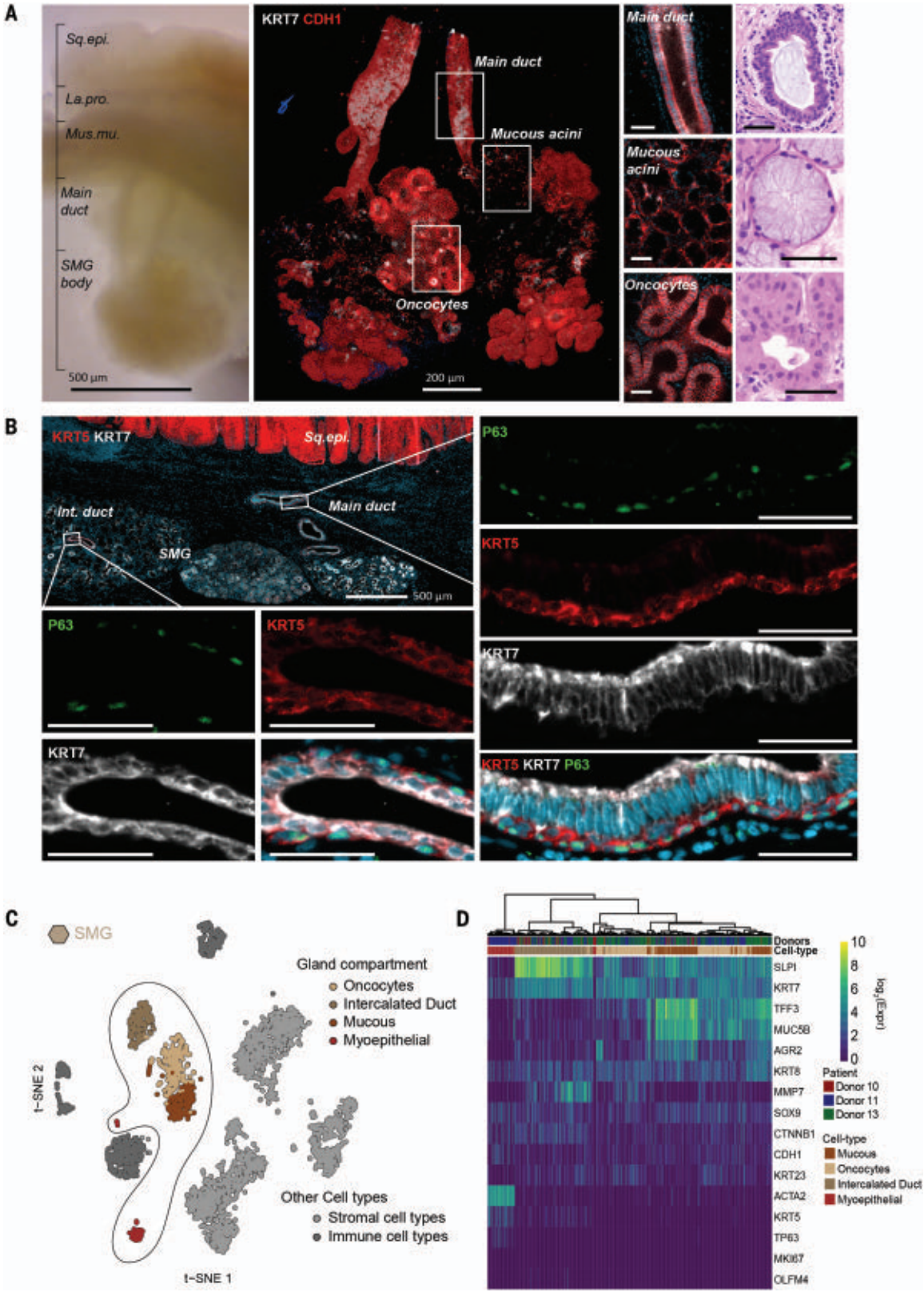



**Fig. 1. Characterization and single-cell profile of human SMG.** (**A**) Whole-mount SMG from disease-free donors, with esophagus squamous epithelium (Sq. epi.), lamina propria (La. pro.), muscularis mucosa (Mus. mu.), duct, and the SMG body. SMG were imaged by confocal microscopy; duct, mucous cells, and oncocytes were confirmed by hematoxylin and eosin (H&E) staining. (**B**) Immunofluorescence of SMG sections. Intercalated (Int.) and main duct regions of SMG were magnified to show the $P63^{+}KRT5^{+}KRT7^{+}$ cells. (**C**) t-SNE (t-distributed stochastic neighbor embedding) projection of scRNA-seq of SMG cells from three disease-free donors. Epithelial cells are circled. (**D**) Heatmap of expression of known SMG marker genes. Only epithelial cells of SMG are shown. For (A) and (B), scale bars are 50 µm unless otherwise indicated.

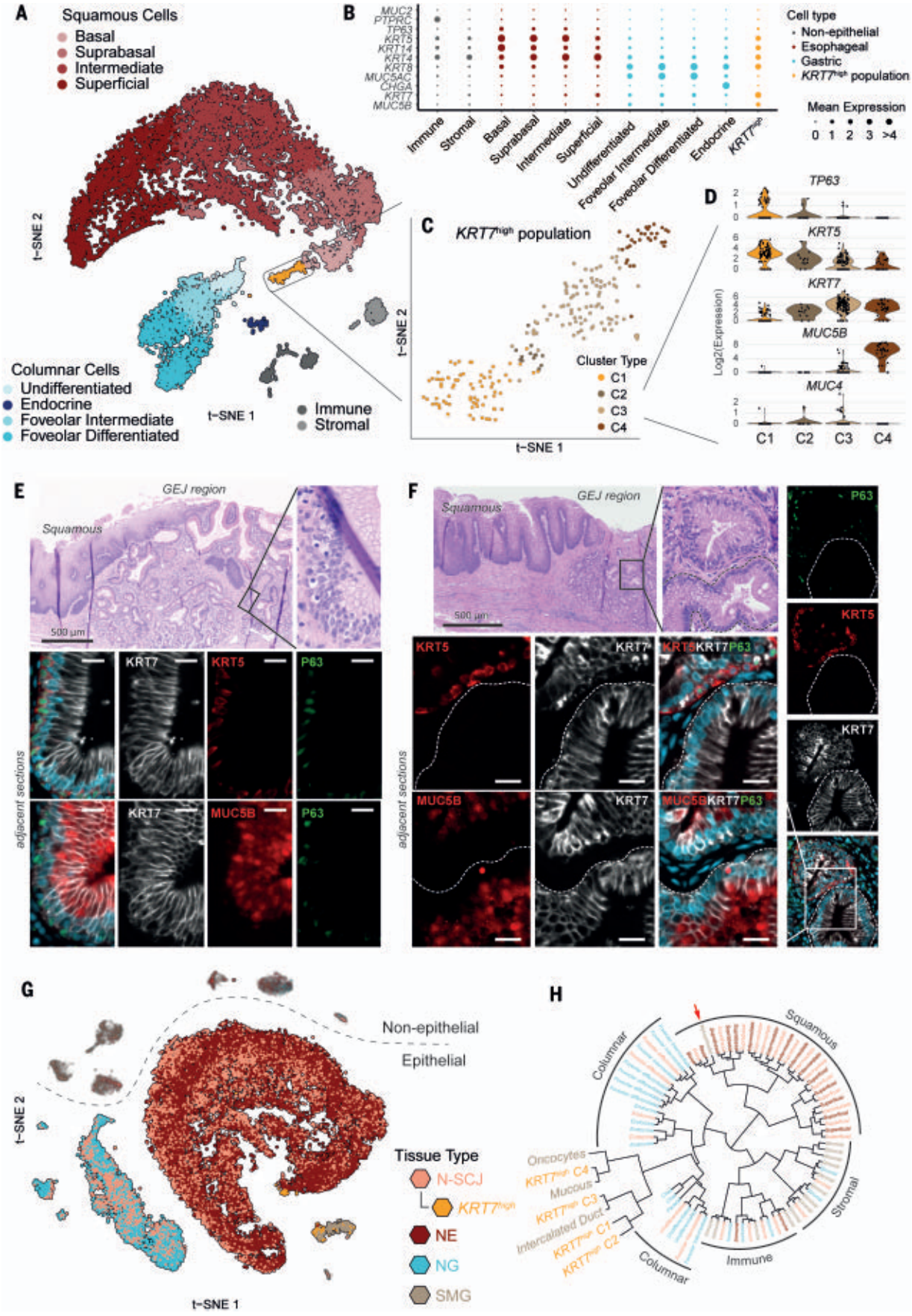

A
Squamous Cells
Basal
Suprabasal
Intermediate
Superficial
Columnar Cells
Undifferentiated
Endocrine
Foveolar Intermediate
Foveolar Differentiated
Immune
Stromal
t-SNE 1
t-SNE 2
B
MUC2
PTPRC
TP63
KRT5
KRT14
KRT4
KRT8
MUC5AC
CHGA
KRT7
MUC5B
Immune
Stromal
Basal
Suprabasal
Intermediate
Superficial
Undifferentiated
Foveolar Intermediate
Foveolar Differentiated
Endocrine
KRT7high
Cell type
Non-epithelial
Esophageal
Gastric
KRT7high population
Mean Expression
0 1 2 3 >4
C
KRT7high population
Cluster Type
C1
C2
C3
C4
t-SNE 1
t-SNE 2
D
TP63
KRT5
KRT7
MUC5B
MUC4
Log2(Expression)
C1
C2
C3
C4
E
GEJ region
Squamous
500 µm
KRT7
KRT5
P63
MUC5B
adjacent sections
F
GEJ region
Squamous
500 µm
KRT5
KRT7
KRT5KRT7P63
MUC5B
KRT7
MUC5BKRT7P63
P63
KRT5
KRT7
adjacent sections
G
Non-epithelial
Epithelial
t-SNE 1
t-SNE 2
Tissue Type
N-SCJ
KRT7high
NE
NG
SMG
H
Squamous
Columnar
Stromal
Immune
Columnar
Oncocytes
KRT7high C4
Mucous
KRT7high C3
Intercalated Duct
KRT7high C1
KRT7high C2

**Fig. 2. A *KRT7*$^{high}$ epithelial population at the GEJ region.** (**A**) t-SNE projection of scRNA-seq of cells from N-SCJ. The *KRT7*$^{high}$ population is highlighted in yellow. (**B**) Bubble plot of selected marker genes relevant to squamous, gastric, and SMG phenotypes. (**C**) t-SNE projection and clustering results of scRNA-seq of the *KRT7*$^{high}$ population. (**D**) Violin plots of gene expression in C1 to C4 cell populations. Marker genes for the transitional basal progenitor cells (P63$^+$KRT5$^+$KRT7$^+$) and residual embryonic cells (KRT7$^+$MUC4$^+$) are shown. (**E** and **F**) H&E staining and immunofluorescence of disease-free donors. H&E confirmed that these cells are not located at SMG but at the GEJ. Adjacent sections were co-stained for KRT5/KRT7/P63 and MUC5B/KRT7/P63, respectively. Note the single-layered MUC5B$^+$ epithelium (marked by a dashed line and corresponding to C4 of the *KRT7*$^{high}$ population). Scale bars are 20 μm unless otherwise indicated. (**G**) t-SNE projections of scRNA-seq of cells from N-SCJ, NE, NG, and SMG. Color denotes tissue types; epithelial cell types are shown in full color, and nonepithelial cell types are shaded out. (**H**) Similarity (Euclidean distance) between cell clusters from all of the tissue types. Clusters C1 to C4 and SMG cell types are highlighted with bigger font. The red arrow indicates myoepithelial cells. Font color denotes tissue types, same as in (G).

*KRT7*$^{high}$ lineage that are at different stages of differentiation, and it is possible that they share the same origin with SMG (fig. S7B).

### *scRNA-seq profiling of BE and normal gastroesophageal tissues*

Having established the phenotypes of SMG and N-SCJ regions, we expanded scRNA-seq analysis to 43,000 cells from 43 samples collected across seven different tissue sites from 14 donors and BE patients (Fig. 3A, fig. S8, and table S1). All major cell types in NE, NG, and ND (normal duodenum, which served as an intestinal cell reference for BE)—including squamous basal and superficial cells, gastric foveolar, endocrine, parietal, and chief cells—were successfully identified (fig. S9, A to F, and tables S4 to S6). Next, we analyzed BE samples, including the BE squamo-columnar junction (B-SCJ; fig. S6B) (*23*). Within BE, we identified columnar cells resembling gastric foveolar cells (*MUC5AC*, *KRT20*), goblet cells (*MUC2*, *TFF3*), and previously uncharacterized poorly differentiated enteroendocrine-like cells (*NEUROG3*, *CHGA*) (fig. S10 and table S7). Furthermore, we identified an undifferentiated BE cell type characterized by lack of differentiation features but showing *OLFM4* expression, a marker of intestinal and BE stem cells (*11*, *24*). This undifferentiated cell type was also confirmed by intercellular transcriptional variability (also referred to as entropy) (*25*) (fig. S11).

To reconstruct the lineage relations of all cell types, we used hierarchical clustering of cell type–specific consensus transcriptomes. Except for goblet cells, BE cells (from both B-SCJ and BE) showed the strongest similarities to gastric cardia cells (from both NG and N-SCJ). This was consistent through all stages of differentiation for gastric and BE cells (Fig. 3B and fig. S12). In addition, BE cells are distinct from any of the SMG cell types and the N-SCJ *KRT7*$^{high}$ clusters. We did not observe *KRT7*$^{high}$ or any intermediate cell populations in the B-SCJ samples, suggesting that cell transdifferentiation from NE to BE is unlikely (fig. S13) (*23*).

### *Genetic and epigenetic similarities between BE and gastric cardia*

In view of the observation that BE cell types resemble the transcriptional profiles of their NG counterparts, we used three lineage-tracing methods (*26*–*28*) to further investigate this. First, methylation profiles were generated from fresh samples of NE, NG, SMG, and BE. Principal components analysis from the top 10,000 most variable probes confirmed that NE and SMG share similar methylation profiles, suggesting a close embryological origin, and have distinctive features not detected in BE samples. BE methylation profiles resembled NG and did not overlap with NE or SMG (Fig. 3C).

Second, we studied lineage relatedness using ATAC-seq (assay for transposase-accessible chromatin with sequencing) that enabled profiling of the open chromatin landscape of tissues from two independent cohorts. We identified the top 10,000 most variable peaks of open chromatin across all the tissue types and performed hierarchical clustering of normalized reads in peaks. The results were concordant with the methylation profiling—NG and BE share the most similar features of chromatin accessibility, whereas SMG clustered with NE (Fig. 3D).

Third, we interrogated the 60× whole-genome sequencing (WGS) data of matched endoscopic biopsies from five patients (SMG not included) (*18*) to identify any spontaneously accumulated somatic mutations that were shared between BE and its putative tissue of origin (*29*). To do this, we identified clonal mutations in BE and then determined their presence in NE or NG (Fig. 3E). We found such mutations in four out of five patients, which were all shared between BE and NG (Fig. 3F and fig. S14). The results were further confirmed by targeted Sanger sequencing in a separate set of tissues from the same patients (fig. S14).

### *c-MYC and HNF4A drive the transcriptional programs of BE cells*

To investigate the transcriptional mechanisms underlying the development of BE, we performed gene set enrichment analysis (GSEA) (*30*) and causal analysis (*31*) of genes differentially expressed between different stages of BE and NG differentiation (Fig. 4, A and B). We reasoned that the differences between the undifferentiated cells of each tissue would be key for BE development. Differential analysis of NG and BE undifferentiated phenotypes identified activation of *CDX1*, *CDX2*, and *MYC* modules in BE (Fig. 4C). Further comparison of BE columnar differentiated cells with undifferentiated cells showed activation of the transcription factor *HNF4A* in the former and *MYC* in the latter (Fig. 4D and fig. S15). These results were corroborated by bulk ATAC-seq analysis. A comparison of regions that are more accessible in BE versus NG tissues identified binding motifs for a variety of transcription factors associated with gastrointestinal development enriched in BE. This included *CDX2* and *HNF4A*, further confirming the scRNA-seq analysis (fig. S16, A and B) (*32*).

At the protein level, c-MYC was detected in 61 out of 66 BE biopsies (25 patients) containing intestinal metaplasia, with expression confined to the bottom of the crypts (Fig. 4E). There was no c-MYC expression in 64 normal gastric biopsies (28 patients). Similarly, all 66 BE biopsies and none of the 64 gastric biopsies were positive for HNF4A. In contrast to c-MYC expression, the expression of HNF4A was present in the top two-thirds of the BE crypts (Fig. 4E).

To elucidate the functional relevance of these findings, we established organoid cultures from disease-free NG tissues and transduced them with exogenous c-MYC and HNF4A marked by green fluorescent protein (GFP) labeling (Fig. 4, F and G). Because NG and BE organoids are morphologically indistinguishable with no identifiable goblet cells, a gene panel was used (Fig. 4H) to evaluate the phenotype of transduced NG organoids. Ectopic expression of c-MYC led to an up-regulation of genes characteristic of undifferentiated BE cells, including *NCL* and *LEFTY1*, whereas HNF4A led to up-regulation of a wider BE gene set comprising *FBP1*, *CLDN3*, *LEFTY1*, *TFF3*, and particularly *CDX2* (Fig. 4I and fig. S17A). ATAC-seq also showed increased accessibility at the promotor regions of *CLDN3* and *CDX2* in BE (fig. S16C). The regulatory roles of c-MYC and HNF4A were also confirmed in the normal gastric cell line HFE145 (fig. S17, B and C). These results demonstrate that ectopic expression of c-MYC and HNF4A in gastric cells drives expression of genes relevant for the BE phenotype.

### *EAC likely originates from undifferentiated BE cells regardless of its clinical features*

Subtypes of EAC based on the presence or absence of BE have different prognoses, raising

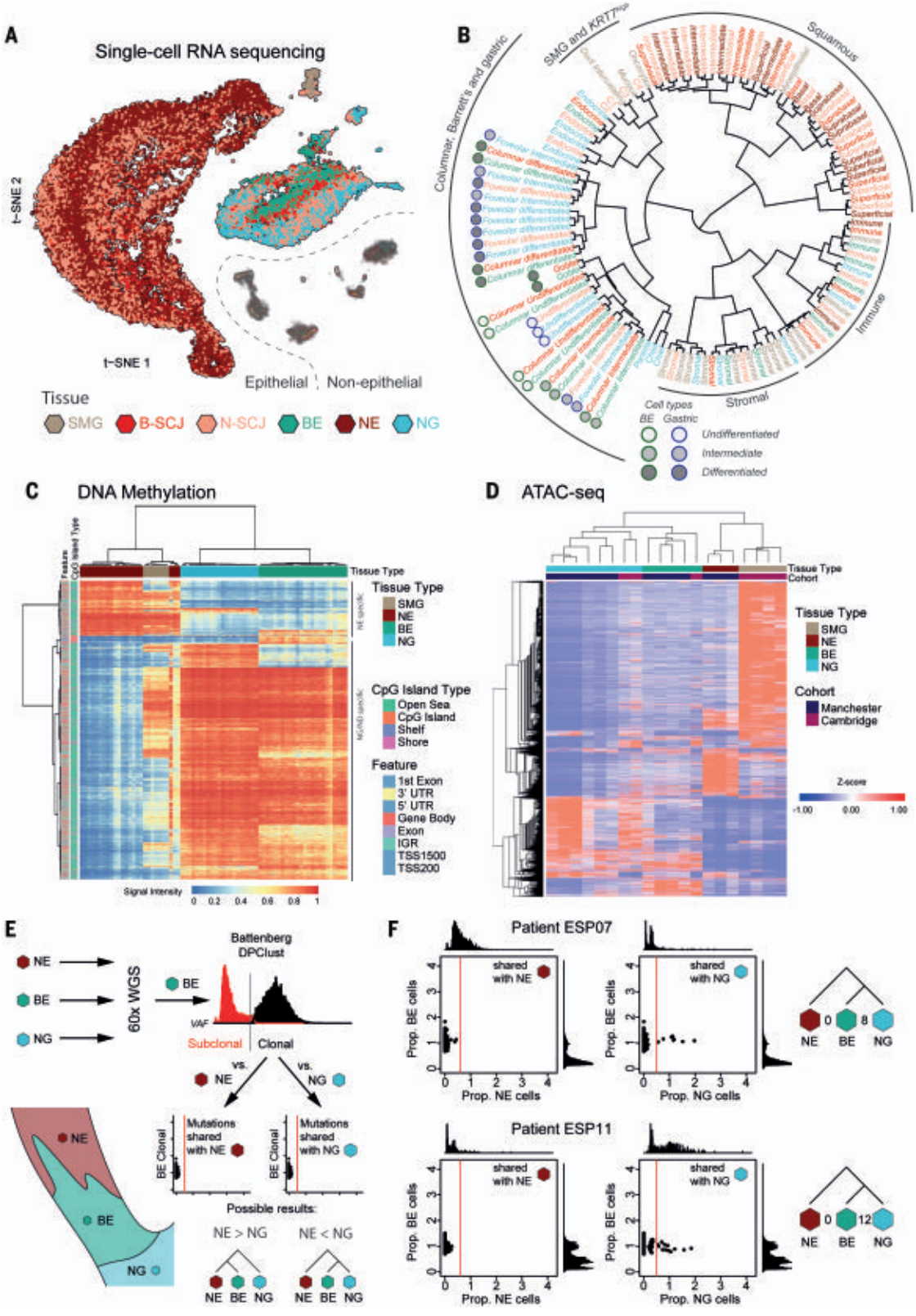



**Fig. 3. BE shows epigenetic, genetic, and transcriptional similarity with gastric cardia.** (**A**) t-SNE projection of scRNA-seq of cells from all tissue sites. (**B**) Similarity (Euclidean distance) between all the cell clusters. Font color denotes tissue sites, same as in (A); undifferentiated, intermediate, and differentiated cells of BE and NG are marked by green and blue circles, respectively. (**C**) Clustering of 10,000 highly variable probes of bulk methylation data from BE, NG, NE, and SMG. IGR, intergenic region; TSS, transcription start site; UTR, untranslated region. (**D**) Hierarchical clustering of the 10,000 most variable peaks of open chromatin regions in bulk ATAC-seq data. (**E**) Schematics of somatic mutation analysis and pairwise comparison from 60× WGS data. (**F**) Pairwise comparison of BE versus NE and BE versus NG in two BE patients; their relationship is indicated by the similarity plot based on shared somatic mutations.

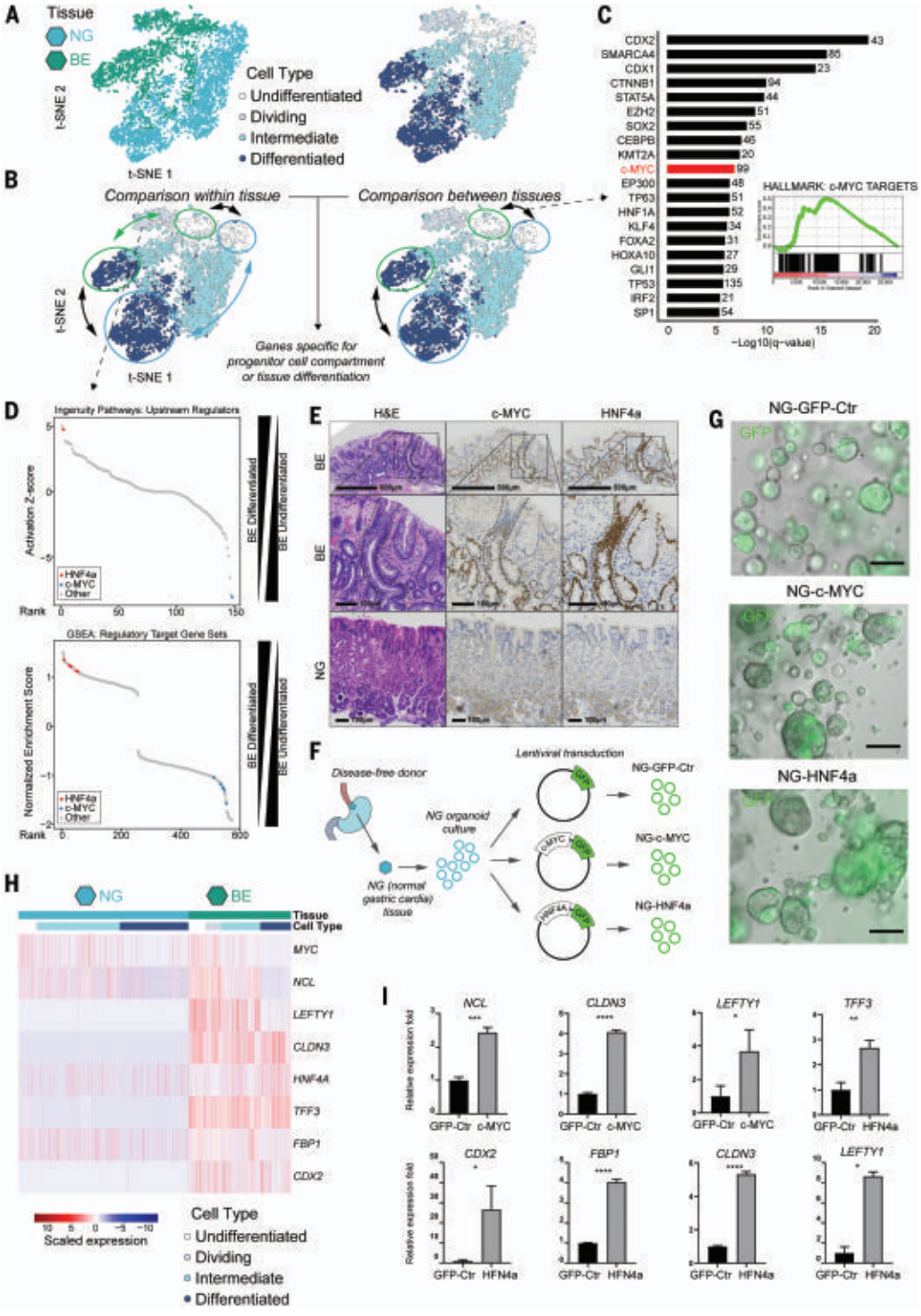



**Fig. 4. c-MYC and HNF4A drive the transcriptional programs of BE cells.** (**A**) t-SNE projection of scRNA-seq of selected epithelial cells from NG and BE with tissue sites (left) and cell subtype (right) overlaid. (**B**) Schematics of differential analysis between cell types used for identification of marker genes. (**C**) Causal analysis of pathways that are up-regulated when comparing BE undifferentiated and NG undifferentiated cells. The top 20 pathways with a positive *z*-score are shown; the number indicates the number of genes common between the dataset and the annotated pathway. The inset shows GSEA using Hallmark gene sets of BE undifferentiated and NG undifferentiated cells. (**D**) (Top) causal analysis of pathways up- and down-regulated when comparing BE undifferentiated and foveolar-like cells. Only pathways with a *q* value < 0.05 and a *z*-score of activation above 1 or below −1 are shown. (Bottom) GSEA using C3 gene sets database of BE undifferentiated and foveolar-like cells. (**E**) Representative immunohistochemical staining of MYC and HNF4A in BE and NG biopsies with an individual crypt highlighted. (**F**) Schematics of generation and transduction of NG organoids. (**G**) Fluorescence imaging of GFP indicating successful transduction of organoids. Ctr, control. (**H**) Marker genes highly expressed in BE cells versus NG. $Log_2$-transformed, normalized counts were *z*-score scaled per gene for visualization purposes. (**I**) Relative expression fold change of BE marker genes in transduced NG organoids upon overexpression of *c-MYC* or *HNF4A*. Error bars represent standard deviation. Statistical analysis: Paired *t* test; **p* < 0.05; ***p* < 0.01; ****p* < 0.001; *****p* < 0.0001.

the question as to whether all EAC arises from BE (*8*). Here, we took advantage of our scRNA-seq profiles to interrogate bulk RNA-seq data from esophageal cancer (*33*–*35*). MuSiC (multi-subject single cell deconvolution) analysis of an independent dataset of bulk RNA-seq correctly assigned scRNA-seq profiles to individual tissue types and was able to distinguish BE columnar cells and NG foveolar cells (figs. S18 and S19). Next, we compared the two phenotypically distinct types of esophageal cancer as confirmation of the method. As expected, bulk RNA-seq data of 81 ESCC (esophageal squamous cell carcinoma) and 321 EAC samples (*36*, *37*) assigned phenotypes correctly (Fig. 5, A and B). The EAC similarity to BE cells was independent of clinical features including TNM (tumor/node/metastasis) stage, the Siewert classification, and the presence of adjacent BE. The expression of differentiated BE cells (BE columnar and goblet

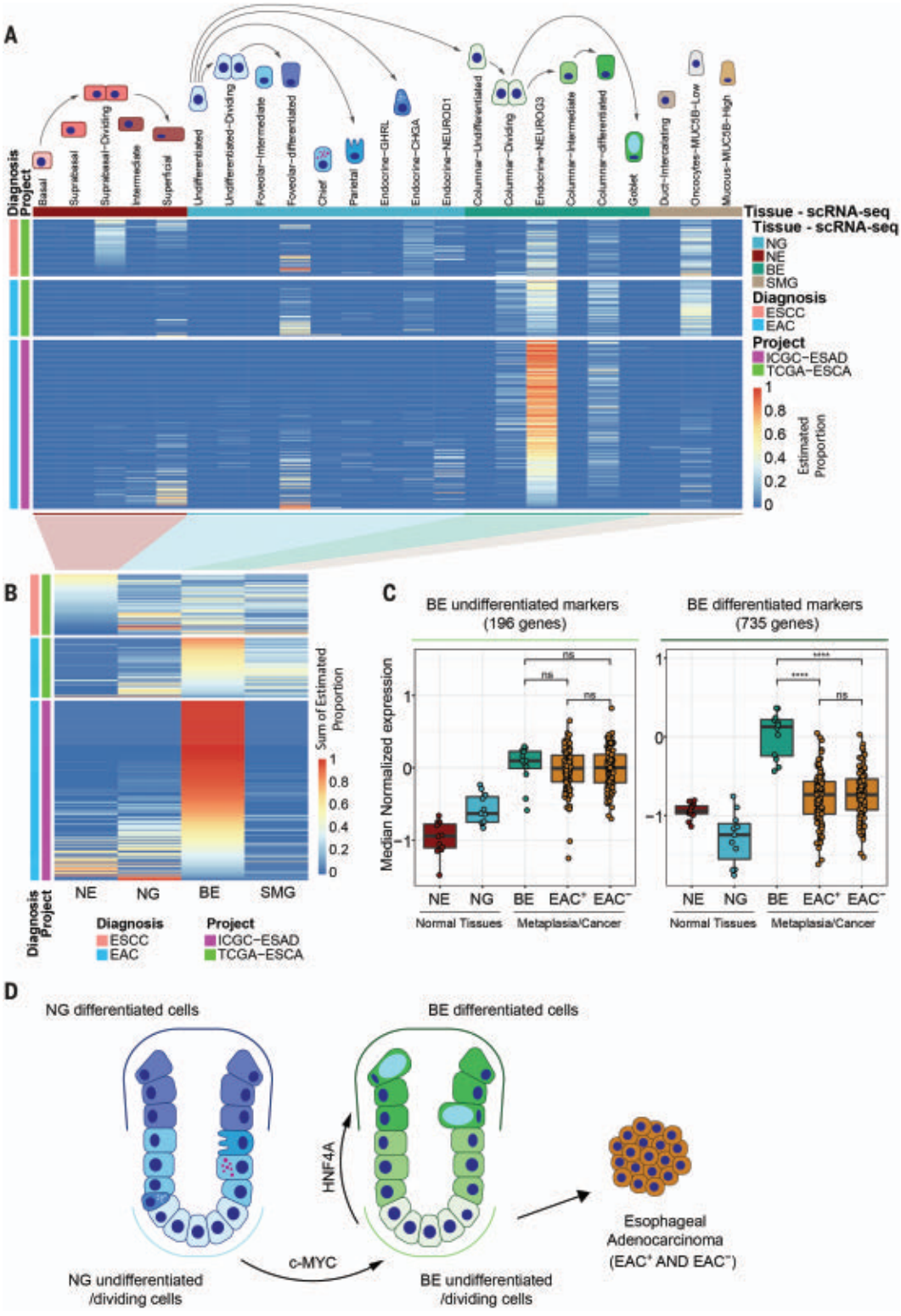



**Fig. 5. Deconvolution analyses of bulk human EAC and ESCC transcriptomes.** (**A**) Contribution of individual cell phenotypes originating from NE, BE, NG, and SMG to the phenotypes of esophageal cancers. The estimation was performed using MuSiC after removal of nonepithelial cell types. Cancer samples (rows) are sorted by the contribution of NE phenotype or BE phenotype for ESCC and EAC, respectively. ICGC-EAD, International Cancer Genome Consortium–Esophageal Adenocarcinoma; TCGA-ESCA, The Cancer Genome Atlas–Esophageal Carcinoma. (**B**) Contribution of combined scRNA-seq–based tissue phenotypes to esophageal cancer phenotypes. Individual tissue contributions are sums of estimated proportions of individual cells from that tissue type. (**C**) Contribution of undifferentiated and endocrine-like (left) or differentiated (foveolar-like and goblet; right) phenotypes of BE to transcriptomes of NE, NG, BE, and EAC samples with or without adjacent BE ($EAC^+$ or $EAC^-$). Each data point represents a median normalized expression of BE undifferentiated or differentiated marker genes for individual samples. Individual $\log_2$ expression values were normalized to the mean $\log_2$ expression of the given gene in BE samples. Boxes indicate interquartile range (IQR), and whiskers extend to the value no further than 1.5 fold of IQR from the boxes. Statistical analysis: Unpaired Wilcoxon test; ****$p < 0.0001$; ns, not significant. (**D**) Schematic of EAC development. c-MYC and HNF4A drive transcriptional programs of BE from NG, and undifferentiated cells of BE serve as the precursor to EAC.

cell markers; tables S7 to S9) decreased during the transition from BE to EAC (Fig. 5C). Conversely, the undifferentiated BE phenotype (defined as a combination of genes specific to BE undifferentiated and endocrine-like cells; tables S9 and S10) was maintained during the BE-to-EAC transition (Fig. 5C). A corresponding analysis with differentiated and undifferentiated NG cell markers did not show any similarity with EAC (fig. S20). This suggests a unified origin of EAC from a metaplastic precursor, even in cases where BE metaplasia is not apparent at diagnosis or in the pathological specimen.

## Discussion

Our results suggest that undifferentiated gastric cells from the cardia give rise to BE via transcriptional programs driven by c-MYC and HNF4A (Fig. 5D). Furthermore, transcriptional profiling of EAC showed expression of markers found in undifferentiated BE cells, regardless of whether metaplasia precursors were identifiable histologically.

The fierce debate over the origin of BE can be attributed to reliance on mouse models and limited human samples. For example, both the residual embryonic cell (*14*) and transitional basal cell (*15*) hypotheses did not consider human SMG and interpreted expansion of $KRT7^+$ cells at the mouse gastroesophageal region as indicative of BE development, yet our data demonstrate that $KRT7^+$ cells are common in the BE-free SCJ and SMG. Although caution is required before treating the mouse metaplasia-like phenotype as equivalent to clinical BE, mouse models remain a useful tool to investigate environmental cues and EAC pathogenesis (*13*, *38*, *39*). Studies of SMG have been limited by difficulty in its isolation. In formalin-fixed tissue sections, Leedham and colleagues found a single BE gland that shared the same mutational profile with the adjacent SMG duct (*9*). Recently, Owen and colleagues used scRNA-seq to profile cells from human endoscopic pinch biopsies of NE, NG, and BE (*11*) and deduced that BE arose from SMG, but without isolating the SMG specifically.

c-MYC and HNF4A have been studied in the BE context (*40*, *41*), but a NE origin was always assumed, and we demonstrated that NG should be the control tissue. Although BE and EAC were previously studied with gene expression microarrays in bulk tissues (*41*, *42*), the genesis of EAC had been difficult to address before single-cell profiling techniques. Our results suggest that EAC may derive from gastric cells via a BE-like metaplasia, even though the evolutionary trajectory and prognosis can vary between patients (*8*). Moreover, it is possible that intestinal metaplasia phenotypes, such as BE or gastric intestinal metaplasia, are an indispensable step toward neoplasia. This is in keeping with a TCGA (The Cancer Genome Atlas) study that concluded that EAC falls within a spectrum of gastroesophageal adenocarcinomas (*37*).

The strengths of this study include comprehensive multi-omic profiling of freshly isolated human cells from superficial to submucosal compartments across the GEJ in healthy and diseased individuals, which has been challenging to achieve in human. However, there are limitations to this study. Transcriptional similarity does not prove causality, and it is possible that minuscule cell populations were undetected or lost during tissue preparation. Single-cell–based, deep somatic lineage tracing with paired scRNA-seq of human samples will be informative to address these limitations when the technologies are mature.

This study provides several orthogonal lines of evidence for a gastric origin for BE, which is likely to be a requisite step for neoplastic progression at this site. We hope that these data will pave the way for further research and clinical early detection and cancer prevention strategies.

## REFERENCES AND NOTES


1. J. M. Slack, *Lancet* **328**, 268–271 (1986).
2. M. Herfs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10516–10521 (2012).
3. S. A. C. McDonald, T. A. Graham, D. L. Lavery, N. A. Wright, M. Jansen, *Cell. Mol. Gastroenterol. Hepatol.* **1**, 41–54 (2014).
4. Y. Peters *et al.*, *Nat. Rev. Dis. Primers* **5**, 35 (2019).
5. E. C. Lin, J. Holub, D. Lieberman, C. Hur, *Clin. Gastroenterol. Hepatol.* **17**, 857–863 (2019).
6. E. C. Smyth *et al.*, *Nat. Rev. Dis. Primers* **3**, 17048 (2017).
7. M. Arnold *et al.*, *Lancet Oncol.* **20**, 1493–1505 (2019).
8. T. Sawas *et al.*, *Gastroenterology* **155**, 1720–1728.e4 (2018).
9. S. J. Leedham *et al.*, *Gut* **57**, 1041–1048 (2008).
10. R. J. von Furstenberg *et al.*, *Cell. Mol. Gastroenterol. Hepatol.* **4**, 385–404 (2017).
11. R. P. Owen *et al.*, *Nat. Commun.* **9**, 4261 (2018).
12. Y. Hu *et al.*, *J. Gastrointest. Surg.* **11**, 827–834 (2007).
13. M. Quante *et al.*, *Cancer Cell* **21**, 36–51 (2012).
14. X. Wang *et al.*, *Cell* **145**, 1023–1035 (2011).
15. M. Jiang *et al.*, *Nature* **550**, 529–533 (2017).
16. L. Hutchinson *et al.*, *Stem Cells Dev.* **20**, 11–17 (2011).
17. G. Gonzalez, Q. Huang, H. Mashimo, *Dis. Esophagus* **29**, 670–680 (2016).
18. See supplementary text 1.
19. D. R. Braxton, D. C. Nickleach, Y. Liu, A. B. Farris III, *Virchows Arch.* **465**, 135–143 (2014).
20. See supplementary text 2.
21. E. Madissoon *et al.*, *Genome Biol.* **21**, 1 (2019).
22. See supplementary text 3.
23. See supplementary text 4.
24. L. G. van der Flier, A. Haegebarth, D. E. Stange, M. van de Wetering, H. Clevers, *Gastroenterology* **137**, 15–17 (2009).
25. D. Grün *et al.*, *Cell Stem Cell* **19**, 266–277 (2016).
26. K. A. Hoadley *et al.*, *Cell* **173**, 291–304.e6 (2018).
27. A. B. Stergachis *et al.*, *Cell* **154**, 888–903 (2013).
28. S. Nik-Zainal *et al.*, *Cell* **149**, 979–993 (2012).
29. See supplementary text 5.
30. A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
31. A. Krämer, J. Green, J. Pollard Jr., S. Tugendreich, *Bioinformatics* **30**, 523–530 (2014).
32. See supplementary text 6.
33. M. C. Vladoiu *et al.*, *Nature* **572**, 67–73 (2019).
34. M. D. Young *et al.*, *Science* **361**, 594–599 (2018).
35. See supplementary text 7.
36. A. M. Frankell *et al.*, *Nat. Genet.* **51**, 506–516 (2019).
37. The Cancer Genome Atlas Research Network, *Nature* **541**, 169–175 (2017).
38. N. S. Münch *et al.*, *Gastroenterology* **157**, 492–506.e2 (2019).
39. D. J. Fu *et al.*, *Nat. Commun.* **11**, 84 (2020).
40. S. Wang *et al.*, *Oncogene* **25**, 3346–3356 (2006).
41. C. Rogerson *et al.*, *Genome Res.* **29**, 723–736 (2019).
42. D. B. Stairs *et al.*, *PLOS ONE* **3**, e3534 (2008).
43. S. Jammula *et al.*, *Gastroenterology* **158**, 1682–1697.e1 (2020).
44. E. Britton *et al.*, *PLOS Genet.* **13**, e1006879 (2017).
45. K. Nowicki-Osuch, Supporting code for "Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition." Zenodo (2021); https://doi.org/10.5281/zenodo.4740422.


## ACKNOWLEDGMENTS


We thank all patients, organ donors, and their families who contributed to this study. We thank Cambridge Biorepository for Translational Medicine for collecting deceased organ donor samples. We thank the members of the OCCAMS consortium for the recruitment of patients and provision of samples for data generation. We thank P. Coupland, K. Kania, and CRUK Cambridge Institute Core Genomics facility for scRNA-seq and the Genomic Technologies Core Facility, University of Manchester, for ATAC-sequencing. We thank J. Rogan and E. Cheadle for their assistance in accessing the MCRC Biobank tissues. We acknowledge support from the Human Research Tissue Bank, Cambridge University Hospitals NHS Foundation Trust. We thank D. Shorthouse for critical assessment of the manuscript. **Funding:** This work was supported by grants from the Medical Research Council (RG84369) and CRUK (RG66287 and RG81771/84119) to R.C.F., grants from CRUK (C9545/A29580) to J.C.M., grants from Wellcome Trust (206194) to S.A.T., grants from Chan Zuckerberg Initiative (174169) to K.B.M., and grants from NIHR Cambridge BRC (RG92051) to K.S.-P. C.W.B. was supported by a CRUK-funded clinical training PhD studentship. K.T.M. was supported by a Chan Zuckerberg Initiative Award. A.K.-S. was supported by a Cambridge Cancer Centre Clinical Research Fellowship. **Author contributions:** K.N.-O. and L.Z. conceived and performed the experiments, performed the analysis, and prepared the manuscript and figures. S.J. analyzed the methylation data. C.W.B. performed ATAC-seq and analyzed the data. K.T.M., J.G., and K.S.-P. collected and managed the deceased organ donor samples. G.D. performed alignment and preliminary analysis of WGS data. A.K.-S. collected bulk RNA-seq data from BE and normal samples. N.E. designed and performed the scRNA-seq analysis and participated in preparing the earlier versions of the manuscript. A.W.-C. and E.M. collected and provided the HCA scRNA-seq data. M.D.P. collected endoscopic samples and provided guidance on other sample collections. M.O. supervised surgical samples collection and provided guidance on other sample collections. K.B.M. conceived and supervised the scRNA-seq analysis. A.D.S. conceived and supervised the analysis of ATAC-seq data. S.A.T. conceived the analysis and provided the HCA scRNA-seq data. J.C.M. conceived and supervised the analysis. R.C.F. conceived the experiments and analysis, supervised all research, and wrote the manuscript. **Competing interests:** R.C.F. holds patents related to Cytosponge-TFF3 and related assays that have been licensed by the Medical Research Council to Covidien (now Medtronic). R.C.F. is a co-founder and shareholder in an early detection and digital pathology company Cyted Ltd. The authors declare no other conflicts of interest. **Data and materials availability:** scRNA-seq data are available from the European Genome-phenome Archive (EGA) (EGAD00001005438), www.esophaguscancercellatlas.org (username: cellatlas; password: Barrett2021), and the HCA project (PRJEB31843). Bulk RNA-seq data are available from the EGA (EGAD00001006882 for BE, NE, and NG; and EGAD00001005388 for EAC). WGS data for BE, NE, NG, and ND samples are available from the EGA (EGAD00001006874). Methylation data for SMG samples are available from the EGA (EGAD00010001795), and other tissues were previously published (*43*). ATAC-seq data are available at ArrayExpress (E-MTAB-6751 and E-MTAB-5169) and were previously published (*44*). Cancer RNA-seq data were previously published (*36*, *37*). All code is deposited in Zenodo (*45*).


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/760/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S24
Tables S1 to S12
References (*46*–*91*)
MDAR Reproducibility Checklist
Movies S1 to S3

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


11 June 2020; resubmitted 26 January 2021
Accepted 7 June 2021
10.1126/science.abd1449

CRISPR BIOLOGY

# Structural basis for target site selection in RNA-guided DNA transposition systems

Jung-Un Park[1], Amy Wei-Lun Tsai[1]†, Eshan Mehrotra[1]†, Michael T. Petassi[2]‡, Shan-Chi Hsieh[2]‡, Ailong Ke[1], Joseph E. Peters[2]*, Elizabeth H. Kellogg[1]*

CRISPR-associated transposition systems allow guide RNA–directed integration of a single DNA cargo in one orientation at a fixed distance from a programmable target sequence. We used cryo–electron microscopy (cryo-EM) to define the mechanism that underlies this process by characterizing the transposition regulator, TnsC, from a type V-K CRISPR-transposase system. In this scenario, polymerization of adenosine triphosphate–bound TnsC helical filaments could explain how polarity information is passed to the transposase. TniQ caps the TnsC filament, representing a universal mechanism for target information transfer in Tn7/Tn7-like elements. Transposase-driven disassembly establishes delivery of the element only to unused protospacers. Finally, TnsC transitions to define the fixed point of insertion, as revealed by structures with the transition state mimic ADP•$AlF_3$. These mechanistic findings provide the underpinnings for engineering CRISPR-associated transposition systems for research and therapeutic applications.

CRISPR-associated proteins (Cas) are macromolecular machines that provide bacteria and archaea with adaptive immunity against bacteriophages and other invasive genetic elements. The RNA-guided DNA nuclease activity of CRISPR-Cas systems has been repurposed (most notably in the case of CRISPR-Cas9) for programmable genomic editing by making precise double-strand breaks (DSBs) in DNA complementary to the RNA guide (*1*). Although conventional CRISPR-Cas systems can generate DSBs with high fidelity at chosen DNA sites, the actual insertion of new DNA is dependent on inefficient processes such as homology-directed repair or nonhomologous end-joining. Moreover, introduction of a DSB into the host genome is dangerous, as it can lead to genome instability. Excitingly, there are examples of transposons that are naturally programmable for targeting: Tn7-like transposons that have co-opted type I (Cascade) (*2*) and type V (Cas12) (*3*) CRISPR-Cas systems on multiple independent occasions for guide RNA–directed transposition. These CRISPR-associated Tn7-like transposons have been shown to exhibit a single programmable DNA integration event at a precise distance and in a specific orientation with respect to the protospacer adjacent motif (PAM) site (*4*–*6*).

In both prototypic Tn7 and the Tn7-like transposon relatives that encode CRISPR-Cas systems, the overall mechanism of transposase recruitment and insertion into a specific target site remains elusive, with little structural information to guide mechanistic studies. Despite remarkable diversity (*7*), every RNA-directed transposition system characterized to date contains a CRISPR effector protein (Cas12k in this study), proteins dedicated to target capture (TniQ+TnsC), and a transposase (TnsB) (Fig. 1A). Previous structural studies have focused on the I-F3 Cascade-TniQ target-DNA binding complex, expressed from the element found in *Vibrio cholerae* (*8*). Although the Cascade-TniQ structure reveals the physical association between the CRISPR effector domain and TniQ, how target-DNA binding ultimately results in transposition remains a mystery.

CRISPR-associated transposons share crucial features with the prototypic Tn7 element. However, instead of a guide RNA complex, prototypic Tn7 uses TnsD (consisting of a TniQ domain integrated with a DNA binding domain) to recognize a specific attachment site (*attTn7*) in the bacterial genome for integration. An incompletely characterized interaction between TnsD and the regulator protein TnsC (a homolog of TnsC from RNA-directed transposition systems) will recruit the core TnsA+TnsB transposase bound to the ends of the element to integrate into the target DNA (*9*). TnsC is an AAA+ protein that has functional parallels with MuB from bacteriophage Mu (*10*). More broadly, in both Tn7/Tn7-like systems and Mu, these AAA+ proteins are important regulators of transposition, both mediating transposase recruitment to the target site and preventing multiple insertions from occurring [also referred to as target site immunity (*11*)]; the latter phenomenon is also reported in CRISPR-associated transposons (*4*, *5*).

Here, we used cryo–electron microscopy (cryo-EM) to characterize TnsC of the type V-K CRISPR-associated transposase system from *S. hofmanni* (referred to throughout as ShCAST). We were particularly drawn to the ShCAST system because of its relative simplicity (four protein components) and established in vitro activity (*5*). We discovered that in the presence of adenosine triphosphate (ATP), TnsC forms a continuous spiraling helical filament interacting with one DNA strand within the duplex, thereby providing a target site search mechanism that is also capable of conducting polarity information to the transposase. We demonstrate that propagation of the TnsC filament terminates when it forms a complex with TniQ, its binding partner; in so doing, it establishes a target site capture mechanism for Tn7 and the extended family of Tn7-like elements. We identify a TnsB transposase-directed process that disassembles the TnsC filament, driven by ATP hydrolysis, suggesting how a targeted protospacer is only used once, whereas future insertions are diverted to new protospacers. This same TnsB interaction with TnsC would provide a mechanism to direct the TnsB-bound transposon DNA to a TnsC-TniQ complex for transposition. Notably, we find that ADP•$AlF_3$-bound TnsC collapses to a single hexamer that would be capable of conveying precise distance information from the protospacer to the point of insertion, and points to a nucleotide-based feature for stabilizing the TnsC-TniQ complex.

## ATP hydrolysis drives ShCAST target site selection

The TnsC regulator protein conveys essential information from the guide RNA complex to the transposase. Previous work with prototypic Tn7 and the Mu transposition system indicates that the nucleotide-bound states of the TnsC and MuB adenosine triphosphatases (ATPases) are important for target site selection. Mu is a well-studied model system for transposition: MuB ATPase forms helical filaments in the presence of ATP, and MuB disassembly is stimulated by MuA transposase (*12*, *13*). ATP hydrolysis is required for proper target selection in both Mu (*12*) and prototypic Tn7 (*14*). To test nucleotide cofactor requirements for ShCAST targeting, we performed an in vitro transposition assay (see supplementary materials) (*5*). Clearly targeted in vitro transposition was detected only in the presence of ATP (Fig. 1B). Sequencing of seven independent events indicated that transposition occurred at the expected distance from the PAM; two events were simple inserts and five co-integrates were consistent with previous observations (*5*, *15*, *16*) (fig. S1A). The ShCAST proteins are incapable of simple inserts on their own, so these products must form by an alternative mechanism in the plasmid targets (e.g., a RecA-independent recombination mechanism such as template switching). Bulk analysis of the reaction products found with

[1]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA. [2]Department of Microbiology, Cornell University, Ithaca, NY 14853, USA.
*Corresponding author. Email: joe.peters@cornell.edu (J.E.P.); ehk68@cornell.edu (E.H.K.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.

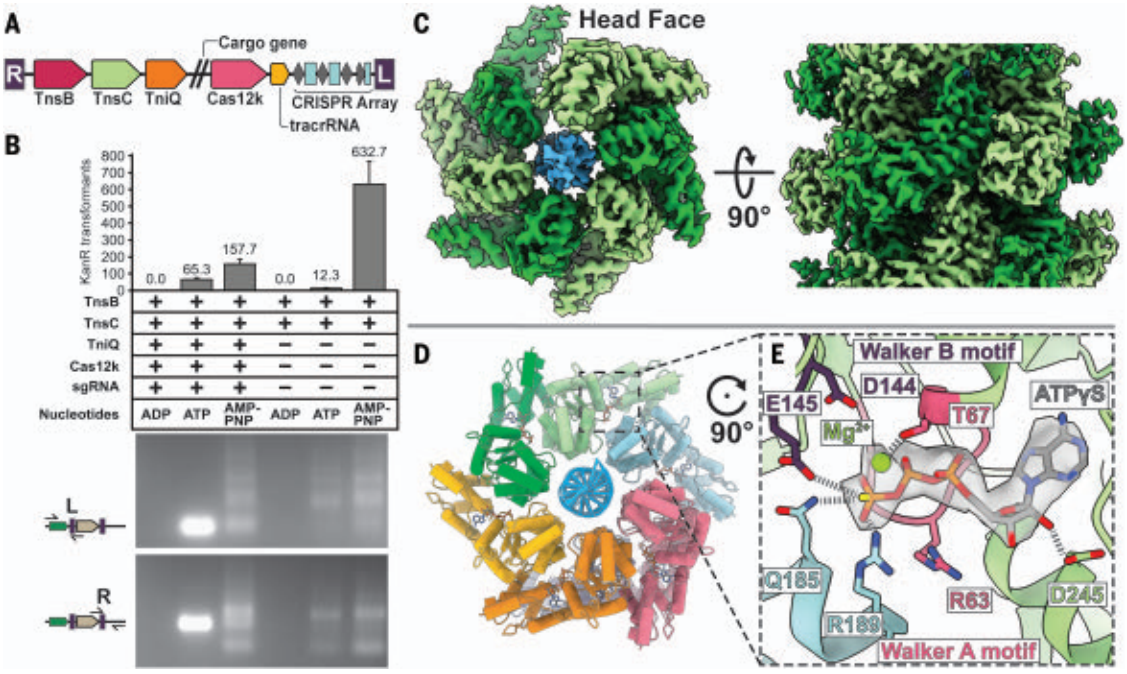



**Fig. 1. ATP is crucial for both RNA-guided transposition of ShCAST, and filament assembly of AAA+ regulator TnsC.** (**A**) The ShCAST transposon is defined by the right (R) and left (L) ends of the element, encoding Cas12k RNA-binding effector, TniQ, TnsC, and TnsB, trans-activating crRNA (tracrRNA), and CRISPR array. Double slash represents the region where cargo genes are found in the transposon. (**B**) ATP hydrolysis is required for targeted RNA-guided transposition. Colony counts from transformation of each deproteinated transposition reaction as a proxy for overall transposition activity are shown (one third of total reaction volume was used in transformation). Data are means ± SD ($n$ = 3). Reaction mixes were additionally analyzed by PCR using a transposon end primer (L or R) along with primers flanking the target site as indicated in the schematic on the left. A high rate of on-target insertion results in a single intense band; insertions distributed around the target plasmid result in a number of PCR products. A single representative image of $n$ = 3 replicates is shown. (**C**) High-resolution (3.2 Å) cryo-EM reconstruction of ATPγS-TnsC (light and dark green) forms a continuous helical filament encircling DNA (blue). The arrangement of layers (six TnsC subunits) within the filament is referred to here as a "head-to-tail" configuration. (**D**) Atomic model of TnsC helical filament consists of six subunits of TnsC arranged in a helical spiral. (**E**) The ATP-binding pocket follows canonical features of AAA+ proteins, with conserved Walker A (pink) and Walker B motif (purple) coordinating ATPγS binding. The adjacent subunit (light blue) forms intersubunit contacts that partially contribute to filament formation by forming interactions with the terminal phosphate.

adenylyl-imidodiphosphate (AMPPNP) by polymerase chain reaction (PCR) revealed a collection of products, which we confirmed to have resulted entirely from untargeted transposition by direct DNA sequencing (Fig. 1B and fig. S1B). Robust targeted transposition required the presence of all reaction components; however, we found considerable random transposition with TnsB, TnsC, and ATP only. Transposition was not found with adenosine diphosphate (ADP) and random transposition was stimulated with AMPPNP. This indicates that, similar to Mu and prototypic Tn7, ATP hydrolysis (via TnsC) is required for ShCAST to select the correct target (*14*, *17*). We also established a genetic assay that monitors full transposition events for the ShCAST system and found a combination of RNA-targeted and off-target events (25% versus 75%, respectively, in this assay) with the ShCAST system (see below) (*5*, *18*, *19*). We also discovered that, similar to MuB ATPase (*20*), ShCAST TnsC forms helical filaments in the presence of ATP, AMPPNP, and adenosine 5′-*O*-(3-thiotriphosphate) (ATPγS). To investigate the structural basis of these observations, we pursued the cryo-EM structure of TnsC.

### The atomic structure of TnsC possesses a canonical AAA+ fold and forms helical filaments

We discovered that TnsC can adopt a helical filament (fig. S2) with a $6_1$ screw axis encircling DNA in the ATP-bound (or ATPγS-bound) states (Fig. 1C and fig. S3), on average ~220 Å in length or about five full turns (fig. S2). Therefore, each TnsC layer has two potential polymerization interfaces, which we refer to as the "head" and the "tail" face (Fig. 1C). Helical search of the cryo-EM images [using iterative helical real space reconstruction (IHRSR)] defines a rise of 6.82 Å and a twist of 60° (see supplementary materials), consistent with helical layer-line analysis (fig. S2). The ATPγS-bound state is of higher resolution overall (3.2 Å for ATPγS versus 3.6 Å for ATP; fig. S4) and has a more uniform distribution as assessed by local resolution estimates (3.0 to 4.0 Å for ATPγS versus 3.5 to 6 Å for ATP; fig. S5). A full-length model (257 of 276 residues) was built into the ATPγS cryo-EM map, starting from a homology model based on distantly related AAA+ YME1 (15% sequence identity; see supplementary materials) (*21*) (Fig. 1D).

ShCAST TnsC and prototypic Tn7 TnsC each have a conserved AAA+ domain, but ShCAST TnsC lacks the N- and C-terminal extensions of prototypic Tn7 that mediate its interactions with other Tns proteins (*22*) (fig. S6). As sequence analysis suggests, the structure follows most of the features of the initiator clade of AAA+ proteins. A DALI (*23*) search reveals strong structural resemblance to the N-terminal portion of Cdc6 [global root mean square deviation (RMSD), 2.7 Å] (*24*) (fig. S7), highlighting the high degree of conservation within the AAA+ domain, even among ATPases of highly divergent function. Correspondingly, the highly conserved Walker A motif ($Gly^{60}$-Glu-Ser-Arg-Thr-Gly-Lys-$Thr^{67}$) and Walker B motif ($Met^{140}$-Leu-Ile-Ile-Asp-$Glu^{145}$) (Fig. 1E and figs. S6 and S8) form a pocket for ATP binding, and mutation of the catalytic glutamate ($Glu^{145}$) almost completely abrogates in vitro and in vivo transposition activity (fig. S9). In addition to these intrasubunit contacts, the ATP-binding pocket is completed by highly conserved $Arg^{189}$ (the arginine finger) and $Gln^{185}$ (fig. S8) from an adjacent subunit. These residues form hydrogen-bonding interactions

with the terminal phosphate of ATP, both recognizing ATP and stabilizing intersubunit contacts (Fig. 1E). Mutation of these residues to alanine causes reduced transposition activity, which may be related to an impaired ability to form helical filaments (as visualized by EM) (fig. S9, A and C). Interestingly, the $Arg^{189}$ → Ala (R189A) mutant allowed transposition at wild-type levels in vitro, but targeting to the protospacer was lost (fig. S9, A and B). Notably, $Asp^{144}$ appears too distant (4.6 Å) from the magnesium ion to facilitate a nucleophilic attack on the γ-phosphate of ATP (Fig. 1E); this suggests that a conformational change (possibly brought about by the transposase TnsB) is required to carry out ATP hydrolysis (see below). Taken together, these observations explain why TnsC, which is predominantly a monomer in solution, requires ATP to oligomerize into the observed helical filament but can be readily disassembled upon ATP hydrolysis.

### Unidirectional TnsC filamentation provides a mechanism for establishing insertion polarity

ATP-bound TnsC forms a right-handed helical filament wrapping around the DNA duplex, forming a spiral "ladder" of interactions with the sugar-phosphate backbone (Fig. 2A). Each TnsC subunit contributes two amino acid contacts, $Lys^{103}$ and $Thr^{121}$, to interact with two adjacent backbone phosphates (Fig. 2A), which suggests that ShCAST TnsC most likely exhibits little to no DNA sequence specificity, similar to MuB (*20*) and TnsC from prototypic Tn7 (*25*). In the ShCAST system, these protein-DNA contacts distort duplex DNA, similar to the AAA+ transposition protein IstB (*26*). In this case, ShCAST TnsC distorts DNA to match the helical symmetry of the filament (Fig. 2B). Strikingly, these interactions are formed preferentially with one strand of the DNA duplex (Fig. 2A), and the local resolution of the complementary strand of DNA is substantially worse than the bound strand (fig. S5, B, D, and F). This suggests that TnsC can form filaments on single-stranded DNA that resemble filaments assembled on double-stranded DNA, a hypothesis we confirmed using EM (fig. S10). We found that in vivo transposition was nearly abolished with the single $Lys^{103}$ → Ala (K103A) and $Thr^{121}$ → Ala (T121A) mutants, but unexpectedly the double mutant was reduced only 65% (fig. S9, D and E). In vitro, K103A + T121A mutant transposition activity was ~10 times wild-type levels, but only 20% of these were on-target insertions (fig. S9, A and B). The lack of specific DNA interactions may in part be compensated by the highly basic surface formed within the pore of the TnsC filament (fig. S11). The findings suggest that protein-DNA interactions at $Lys^{103}$ and $Thr^{121}$ are important for restraining transposition and directing targeting information from the effector complex to the transposase.

Protein filament polymerization is generally a unidirectional process. Consistent with this, TnsC filaments reconstituted with nonhydrolyzable analog ATPγS or frozen immediately upon reconstitution (with ATP) exhibited uniform polarity to cover the entire DNA substrate used in our analysis (i.e., each "head" of TnsC interacts with the "tail" of the adjacent TnsC layer, termed "head-to-tail") (Fig. 2C). Although ATP-dependent filaments were stable over short time frames, they appeared to be more dynamic with prolonged incubation. For example, when samples were incubated overnight, we observed a substantial number of two converging filaments, forming head-to-head filament structures where they met (20% after 12 hours versus none after 10 min; Fig. 2C and fig. S12). This is consistent with a dynamic process where the single filament we initially observed coating the entire DNA substrate could presumably partially dissociate, allowing new converging filaments to form (see supplementary materials) (fig. S12). Therefore, we hypothesize that polar growth of TnsC filaments in the 5′-to-3′ direction of the bound DNA strand is the searching mechanism that enables TnsC to search for its target site, which is defined by TniQ and Cas12k.

### TniQ interacts with TnsC to define the target site

In prototypic Tn7 and Tn7-like systems, target information is conveyed to TnsC from a TniQ domain family protein called TnsD (*27*). By contrast, in the RNA-guided transposition systems, the target site is chosen by a TniQ-associated guide RNA complex. In the case of I-F3 systems, TniQ is positioned at the programmed insertion site via its association with Cascade (*8*). We propose that TnsC filaments are perpetually searching for a target site via directional growth. This directional searching of TnsC filaments, until collision with an appropriately positioned TniQ, could explain how insertions occur only on one side of an effector complex. Therefore, diverse targeting mechanisms may have evolved by fusing TniQ to different DNA binding domains, and, in the case of guide RNA–directed systems, by associating with CRISPR-effector proteins. We believe this would serve as a unifying model accounting for diverse targeting mechanisms spanning both prototypic Tn7 and Tn7-like elements with and without CRISPR-Cas systems. To explore this further, we reconstituted a simplified, minimal system to probe the possible role of TniQ as a target site selection factor.

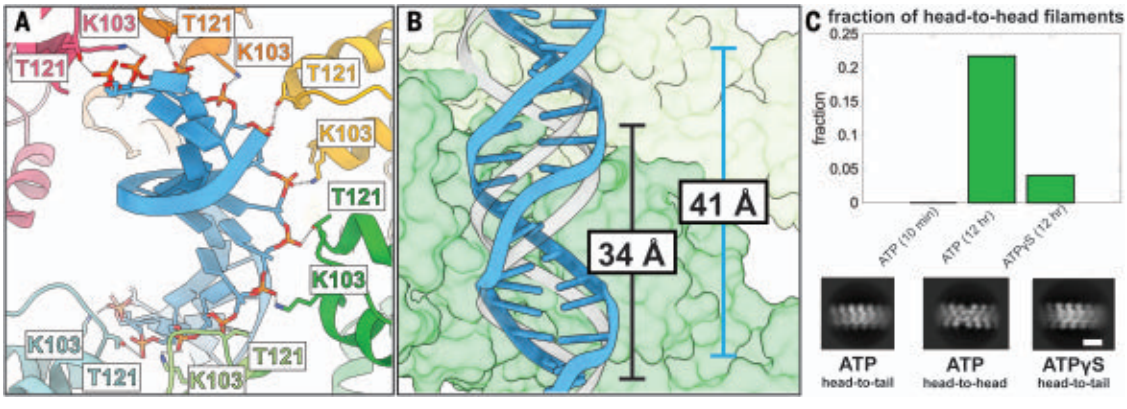



**Fig. 2. Structural analysis of TnsC-DNA interactions reveals how TnsC distorts DNA; coupled with cryo-EM–based time-course experiments, these data suggest that TnsC polymerizes in the 5′-to-3′ direction.** (**A**) Each TnsC subunit makes two major contacts with the DNA sugar-phosphate backbone (blue, sticks), forming a ladder of interactions selectively with one strand of DNA. (**B**) TnsC-DNA interactions result in DNA distortions. One full turn of B-form DNA spans 34 Å (gray); however, the spacing between layers (41 Å) stretches the DNA, distorting one full turn of the duplex DNA to match TnsC's helical spacing. (**C**) The fraction of filament collisions, otherwise referred to as head-to-head filaments, observed is influenced by both nucleotide (ATP versus ATPγS) and increase over time (more are present after 12 hours versus 10 min). Relevant 2D class averages for each sample are shown. Scale bar, 100 Å.

To directly visualize the interaction between TnsC and TniQ and their possible roles in target selection, we incubated TnsC and TniQ together in the presence of ATP and DNA, and then examined the resulting complexes by means of high-resolution cryo-EM reconstructions (3.9 Å resolution; fig. S4). We found that TniQ selectively engages with the polymerizing face, capping the TnsC filaments (Fig. 3B), consistent with the idea that the 5′-to-3′ directional propagation of the filament leads to productive interactions with the Cas12k-TniQ complex. We observe a total of two TniQ monomers; each copy interacts with two TnsCs (the TniQ-TnsC interface buried surface area is 1368 Å$^2$ versus 1052 Å$^2$ for adjacent TnsC subunits along the body of the TnsC filament) (Fig. 3B), even though sterically three TniQs can be bound to the advancing TnsC filament. Each TniQ monomer also appears to be interacting with DNA, contacting the DNA strand that is not bound by TnsC (fig. S13A). Despite the high overall quality of the cryo-EM reconstruction, the local resolution of TniQ is too low for de novo model building (6 to 8 Å; see supplementary materials) (fig. S13B). Nonetheless, homology models of TniQ's functional domains [helix-turn-helix (HTH) and zinc-finger (ZnF), respectively; fig. S13C] built from the I-F3 TniQ crystal structure (PDB 6V9P) (*28*) explain the cryo-EM density well (see supplementary materials) (Fig. 3C). Both the HTH and ZnF motifs appear to interact with the same region of TnsC. In the type I-F3 system, TniQ associates as a homodimer; however, in the ShCAST system, TniQ is naturally found as essentially a "minimal" TniQ domain, lacking a dimerization interface (Fig. 3A). Correspondingly, we do not see substantial protein-protein interactions between the two copies of TniQ. However, reminiscent of the I-F3 system, the two copies of TniQ are oriented such that the N terminus of one TniQ monomer is close to the C terminus of the other (Fig. 3B). The ATPγS TnsC atomic model explains the remaining cryo-EM density well, indicating that TniQ binding itself does not change TnsC helical parameters.

The selective interaction of TniQ with only the advancing end of TnsC filaments explains how TniQ, likely also associated with Cas12k during the guide RNA–directed process, selects target site insertion polarity. With this model of ShCAST TnsC-TniQ interaction in hand, we speculate on the possible higher-order assembly of a guide RNA–directed target site selection complex. Superimposing our docked ShCAST TniQ model onto the type I-F3 Cascade-TniQ structure (PDB 6PIJ) reveals that the spatial organization of TniQ's functional domains is conserved (global RMSD, 2.5 Å; fig. S13D). Our model additionally reveals a possible path for the double-stranded DNA downstream of the R-loop (fig. S13E), which was not visualized in previous structures (*8*).

### The TnsC ADP•AlF$_3$ structure represents a target-capture state and contains spacing information

A notable feature of Tn7 and Tn7-like elements, including guide RNA–directed systems, is that the point of insertion is displaced a fixed distance from the actual machinery of target recognition, and no particular sequence is required for end joining on the target DNA. The TnsC filaments we identify here would provide a mechanism to offset the point of transposase association and point of insertion from the recognized target sequence. But how can the precise spacing that is a hallmark of Tn7 and Tn7-like elements be dictated by an extended, continuous filament?

In prototypic Tn7 and Mu, TnsC (MuB) oligomers are disassembled by ATP hydrolysis, stimulated by the transposase TnsB (MuA) (*13*, *29*). We discovered that this feature is conserved in the ShCAST system: ATP-bound TnsC filaments are disassembled upon addition of TnsB, whereas AMPPNP-bound filaments are not (fig. S14). We predict that the interactions between TnsC and TnsB could be reminiscent of MuB-MuA interactions (*12*, *20*, *30*) and would therefore be located near the tail face of TnsC. This immediately suggested to us a link between ATP hydrolysis and the precise insertional spacing from the PAM site observed for all guide RNA–directed transposition systems to date (*4*, *5*). Although a continuous TnsC filament would be incompatible with the insertional preferences observed (i.e., fixed spacing from the PAM site), TnsC filament "trimming" by TnsB may result in a specific oligomeric configuration that neatly encodes spacing information (see below). The TniQ association across protomers of TnsC and with DNA could additionally physically resist dissociation or act in an allosteric fashion to allow TnsC to resist hydrolysis.

To investigate the hydrolytic state, we determined the cryo-EM structure of TnsC using a nucleotide analog that represents a hydrolysis transition-state mimic, ADP•AlF$_3$. Our 3.9 Å cryo-EM reconstruction (Fig. 4A and fig. S4) revealed that ADP•AlF$_3$-bound TnsC assembles only in an asymmetric structure that can be described as two hexamers oriented in a head-to-head configuration, similar to the configuration found when converging filaments meet (Fig. 2C and fig. S15A). Although the same length of DNA substrate was used for reconstitution of ADP•AlF$_3$ and ATP-bound TnsC (60 base pairs in both cases), the ADP•AlF$_3$

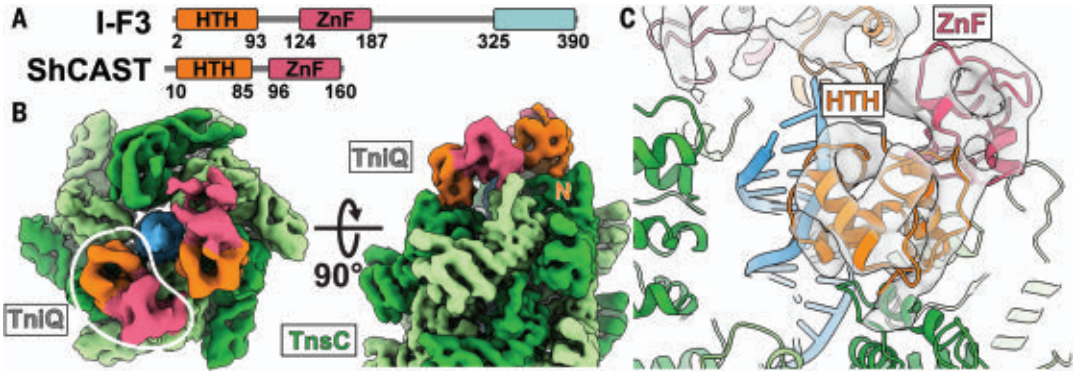



**Fig. 3. Cryo-EM structure of TniQ-TnsC reveals how target site selector protein TniQ binding at the target site can interact with polymerizing TnsC.** (**A**) TniQ from ShCAST is truncated with respect to I-F3 TniQ. Numbers indicate residue positions. Functional domains corresponding to the helix-turn-helix (HTH, orange) motif and zinc-finger ribbon (ZnF, pink) motif are indicated. The light blue domain (only in I-F3) corresponds to the C-terminal winged helix-turn-helix motif and is missing in ShCAST TniQ. (**B**) Two copies of TniQ (orange/pink) interact with the head interface of the ATP-bound TnsC filament. Each monomer of TniQ interacts with two subunits (light/dark green) of TnsC. The cryo-EM map shown is filtered according to local resolution estimates (Bsoft). "N" denotes the N terminus of TniQ. (**C**) Homology models of the HTH and ZnF motifs fit well with the observed cryo-EM density map. Cryo-EM density for ShCAST TniQ is shown; TnsC (green) and DNA (blue) are displayed in ribbon.

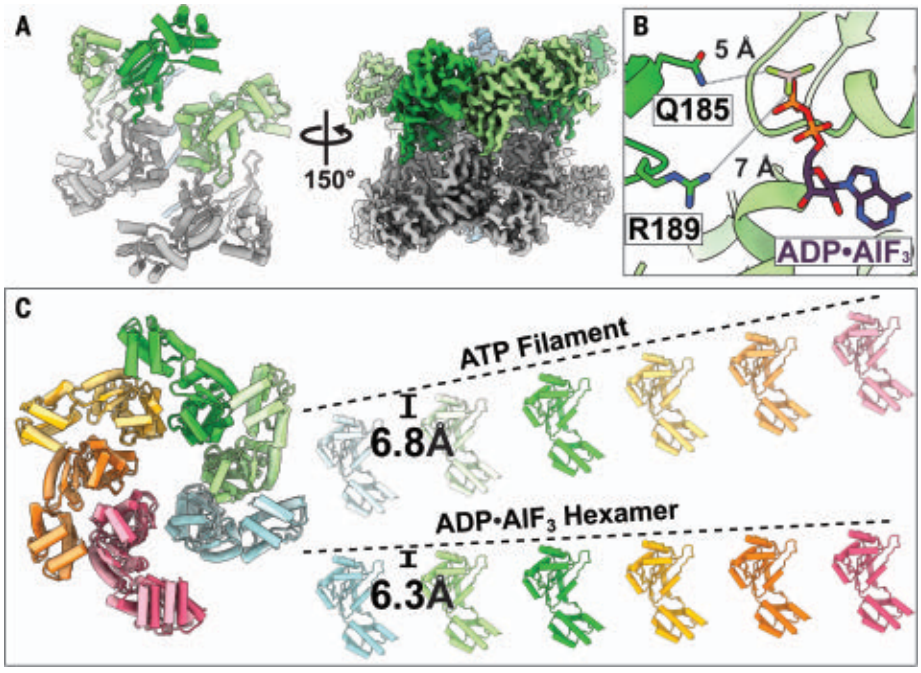



**Fig. 4. Cryo-EM structure of ADP•$AlF_3$-bound TnsC adopts a closed-off hexameric structure that is unable to support propagation of the filament.** (**A**) The ADP•$AlF_3$ 3.9 Å cryo-EM consensus density map reveals a head-to-head configuration of hexamers bound to duplex DNA. This structure cannot support the formation of more than one helical turn. (**B**) TnsC subunits are repositioned such that the terminal phosphate is too far (5 to 7 Å) to be coordinated by intersubunit contacts $Gln^{185}$ and $Arg^{189}$. (**C**) The difference in subunit position results in a smaller rise (indicated by dashed lines) in the ADP•$AlF_3$ hexamer relative to the ATP-bound TnsC filament (6.3 versus 6.8 Å per subunit), resulting in a "closed" configuration that cannot accommodate another subunit to propagate the helical filament.

particles were notably shorter (the DNA-binding footprint is 22 nucleotides total; fig. S15, B and C). This indicates that the ADP•$AlF_3$ complex represents a conformational state of TnsC that is different from the continuous helical filament.

The structural configuration had obvious implications for relating the distance from the protospacer to the point of integration. In the ATP-binding pocket of ADP•$AlF_3$ structure, the lack of intersubunit contacts ($Gln^{185}$ and $Arg^{189}$ are 5 to 7 Å from ADP•$AlF_3$) results in an altered TnsC subunit organization (Fig. 4B) and higher conformational flexibility (movie S1). This altered binding site configuration propagates to result in an overall smaller helical rise in the ADP•$AlF_3$ state (6.3 Å for ADP•$AlF_3$ versus 6.8 Å for ATP; Fig. 4C). We believe this represents the conformational changes that occur upon filament disassembly. The lack of TnsC filaments in the ADP•$AlF_3$ sample also suggests that upon ATP hydrolysis, the head-to-head configuration we observe is more stable against disassembly compared to the filament. This is intriguing because the interface between the two TnsC subunits (representing a total surface area of 1333 $Å^2$) corresponds to the previously identified TniQ binding site (Fig. 3B). We speculate that the observed head-to-head interface is substituting for the interface between TnsC and TniQ (described above) and the Cas12k-TniQ complex found during bona fide transposition. Thus, it is possible that one TnsC hexamer may remain stably bound to the Cas12k-TniQ complex after TnsB-stimulated ATP hydrolysis.

## Discussion

Previous biochemical characterization of the ShCAST system (*5*) and work presented here indicates that programmed insertion of the DNA element occurs at a fixed distance from the protospacer in a single orientation. From these structural studies, we form a comprehensive picture that reconciles ShCAST TnsC's seemingly disparate proposed roles in target site selection (Fig. 5 and movie S2) and draws strong mechanistic parallels with MuB (*12*, *31*). Our cryo-EM structures of ATP-bound TnsC reveal filaments that polymerize unidirectionally in the 5′-to-3′ direction. We hypothesize that such filaments represent a "searching" state that would encounter the Cas12k-TniQ–defined target site, with a specific polarity, on the PAM distal side of the effector. Our cryo-EM structure of TniQ-TnsC reveals the potential nature of this association at the target site: Only one face of TnsC forms productive interactions with TniQ. An ability of TnsB (possibly bound to the transposon ends needed for integration) to "follow" TnsC to the chosen target site could draw the element to the integration site marked by TniQ, as a similar process drives plasmid partitioning systems using ATPases (*32*). Interestingly, our model also accounts for the "immunity" process previously reported for the ShCAST system that prevents multiple insertions from occurring at the same protospacer (*5*). In the post-hydrolysis state, we observe that TnsC is incapable of forming a filament. Although the exact form of TnsC in the active integration complex remains to be resolved, our results suggest how TnsC filaments interact with a Cas12k-TniQ complex with the right polarity and how TnsB-mediated ATP hydrolysis defines a shortened, integration-competent state.

We have also shown that TnsC induces a distortion in its DNA substrate by enforcing the helical parameters of the TnsC filament onto the DNA. Although the implications of this slight unwinding of the DNA require further exploration, it is tempting to speculate that this distortion could be crucial for its function. DNA distortions are a generally used driving force for integrases (*33*). As such, the high potential energy stored in the distorted DNA can be harnessed by the TnsB transposase machinery to facilitate forward transposition, or alternatively, may play a role in ensuring

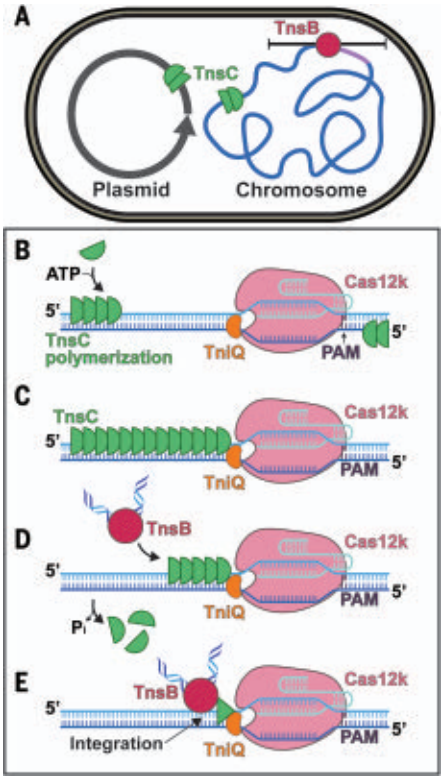



**Fig. 5. Mechanistic model describing TnsC's role in target site selection.** TnsC promotes exploration of alternative target sites by polymerizing along DNA. (**A**) ShCAST elements may exist in mobile plasmids or in attachment sites (indicated by purple segment) within bacterial chromosomes. TnsB around the previous insertion sites (red circle) triggers TnsC depolymerization (indicated with black bars), thereby rendering sites "immune" to insertion. (**B** to **E**) Results are summarized using a conceptual cartoon describing TnsC function. Movie S2 summarizes the same information using the reported structures. (B) TnsC polymerizes unidirectionally along DNA (green semicircles) on either strand in the presence of ATP in the 5′-to-3′ direction. (C) Once TnsC encounters Cas12k-TniQ, it is prevented from polymerizing further and forms a complex with TniQ. (D) TnsB (bound to the terminal ends of the transposon) is able to stimulate TnsC depolymerization and simultaneously can be recruited to the target site. (E) TnsC is disassembled to a finite oligomeric assembly (indicated by green triangle that represents a conformational change that is unable to support a continuous helical filament), which allows integration a fixed distance from the protospacer.

stable binding of TnsC at the target site. In prototypic Tn7, it has been established that TnsABC transposes at a specific position in a single orientation using a DNA distortion induced by TnsD (TniQ) (*34*). A similar spacing paradigm has been proposed with the prototypic Tn7 system, where a different mechanism is used to accomplish the same spacing feature and distribution of binding interfaces on opposite sides of TnsC (*35*). In the absence of TnsD, a DNA distortion associated with triplex DNA is sufficient to initiate TnsC recruitment (*36*), which suggests that TnsC presents a predistorted DNA to the TnsB transposase to accommodate target DNA. Our results indicate that in the Cas12k system, TnsC introduces distortions in DNA upon binding, which suggests that the roadblock of the TniQ-Cas12k complex may constrain this distortion information and pass it to the TnsB transposase to allow access to target DNA to license transposition.

The target immunity process found with some transposons protects both the element and the surrounding region from subsequent insertion events attributable to a local high concentration of the transposase. In the case of CRISPR-associated transposons, the process would also serve to divert insertions to unused protospacers. The results presented here suggest that the structural basis of target immunity is filamentation and disassembly of the AAA+ regulators in these systems, as described in other transposition systems with both Mu and prototypic Tn7. The physical association between TnsC and TniQ revealed by our cryo-EM reconstruction explains how target site selection information can be conveyed between the TnsC regulator and the RNA-binding CRISPR effector domain to result in programmable insertion. It is likely that analogous physical interactions occur in prototypic Tn7 and type I-F3 guide RNA–directed systems.

Tn7-like transposons, including the prototypic Tn7 and guide RNA–directed transposition systems, display a remarkable diversity of targeting modalities that predominantly rely on proteins with TniQ domains. This diversity in targeting pathways suggests a remarkable adaptability for TniQ, which defines the target site for the core transposition system (*2*, *3*, *37*, *38*). Notably, both TnsC and TniQ in ShCAST sequences are considerably shorter than their equivalents from other systems (in prototypic Tn7 and other Tn7-like systems), containing only the highly conserved AAA+ core of TnsC and the HTH and ZnF motifs of the TniQ domain. Thus, we suspect that the structure visualized here likely represents the conserved, minimal functional interactions that are required between TniQ and TnsC. The interactions between ShCAST TnsB transposase and its regulator TnsC, however, remain to be determined. Therefore, the TnsC-TniQ structure provides an excellent starting point for engineering new links and interactions between new target DNA recognition modules and the core transposase for more sophisticated genome-editing applications.

## REFERENCES AND NOTES


1. G. J. Knott, J. A. Doudna, *Science* **361**, 866–869 (2018).
2. J. E. Peters, K. S. Makarova, S. Shmakov, E. V. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7358–E7366 (2017).
3. G. Faure *et al.*, *Nat. Rev. Microbiol.* **17**, 513–525 (2019).
4. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, *Nature* **571**, 219–225 (2019).
5. J. Strecker *et al.*, *Science* **365**, 48–53 (2019).
6. M. Saito *et al.*, *Cell* **184**, 2441–2453.e18 (2021).
7. J. E. Peters, *Mol. Microbiol.* **112**, 1635–1644 (2019).
8. T. S. Halpin-Healy, S. E. Klompe, S. H. Sternberg, I. S. Fernández, *Nature* **577**, 271–274 (2020).
9. J. E. Peters, *Microbiol. Spectr.* **2**, 10.1128/microbiolspec.MDNA3-0010-2014 (2014).
10. K. Mizuuchi, *Annu. Rev. Biochem.* **61**, 1011–1051 (1992).
11. Z. Skelding, J. Queen-Baker, N. L. Craig, *EMBO J.* **22**, 5904–5917 (2003).
12. T. A. Baker, M. Mizuuchi, K. Mizuuchi, *Cell* **65**, 1003–1013 (1991).
13. E. C. Greene, K. Mizuuchi, *Mol. Cell* **9**, 1079–1089 (2002).
14. R. J. Bainton, K. M. Kubo, J. N. Feng, N. L. Craig, *Cell* **72**, 931–943 (1993).
15. P. A. Rice, N. L. Craig, F. Dyda, *Science* **368**, eabb2022 (2020).
16. P. L. H. Vo, C. Acree, M. L. Smith, S. H. Sternberg, *Mobile DNA* **12**, 13 (2021).
17. M. Yamauchi, T. A. Baker, *EMBO J.* **17**, 5509–5518 (1998).
18. P. L. H. Vo *et al.*, *Nat. Biotechnol.* **39**, 480–489 (2021).
19. B. E. Rubin *et al.*, bioRxiv 209189 [preprint]. 17 July 2020.
20. N. Mizuno *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, E2441–E2450 (2013).
21. C. Puchades *et al.*, *Science* **358**, eaao0464 (2017).
22. K. Y. Choi, J. M. Spencer, N. L. Craig, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2858–E2865 (2014).
23. L. Holm, *Protein Sci.* **29**, 128–140 (2020).
24. J. Liu *et al.*, *Mol. Cell* **6**, 637–648 (2000).
25. P. Gamas, N. L. Craig, *Nucleic Acids Res.* **20**, 2525–2532 (1992).
26. E. Arias-Palomo, J. M. Berger, *Cell* **162**, 860–871 (2015).
27. R. Mitra, G. J. McKenzie, L. Yi, C. A. Lee, N. L. Craig, *Mob. DNA* **1**, 18 (2010).
28. N. Jia, W. Xie, M. J. de la Cruz, E. T. Eng, D. J. Patel, *Cell Res.* **30**, 182–184 (2020).
29. Z. Skelding, R. Sarnovsky, N. L. Craig, *EMBO J.* **21**, 3494–3504 (2002).
30. Z. Wu, G. Chaconas, *J. Biol. Chem.* **269**, 28829–28833 (1994).
31. E. C. Greene, K. Mizuuchi, *Mol. Cell* **10**, 1367–1378 (2002).
32. L. C. Hwang *et al.*, *EMBO J.* **32**, 1238–1249 (2013).
33. V. Arinkin, G. Smyshlyaev, O. Barabas, *Curr. Opin. Struct. Biol.* **59**, 168–177 (2019).
34. P. N. Kuduvalli, J. E. Rao, N. L. Craig, *EMBO J.* **20**, 924–932 (2001).
35. Y. Shen *et al.*, bioRxiv 445525 [preprint]. 25 May 2021.
36. J. E. Rao, P. S. Miller, N. L. Craig, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3936–3941 (2000).
37. S.-C. Hsieh, J. E. Peters, bioRxiv 429022 [preprint]. 6 February 2021.
38. M. T. Petassi, S. C. Hsieh, J. E. Peters, *Cell* **183**, 1757–1771.e18 (2020).


## ACKNOWLEDGMENTS

We thank the Cornell Center for Materials Research facility, as well as K. Spoth and M. Silvestry-Ramos, for maintenance of electron microscopes used for this research (NSF MRSEC program, DMR-1719875); XSEDE for computational resources used for image processing (MCB200090 to E.H.K.); A. Byer, N. Ando, G. Bhabha, and S. Vos for advice on the use of nucleotide analogs; members of the Ke and Peters groups for helpful and stimulating discussions; L. Eshun-Wilson, E. Alani, and B. Crickard for valuable advice throughout this project; and N. Craig and A. Guarné for valuable feedback on the manuscript. **Funding:** Supported by NIH grants R00-GM124463 (E.H.K.), R01GM129118 and R21AI148941 (J.E.P.), and GM118174 (A.K.).

**Author contributions:** E.H.K. conceived of the project and designed cryo-EM imaging experiments; J.-U.P. and A.W.T. prepared samples for cryo-EM imaging; J.-U.P. and E.M. analyzed, processed, and refined cryo-EM images to obtain 3D reconstructions; J.-U.P., E.H.K., and E.M. built and refined atomic models; M.T.P. and S.-C.H. designed and carried out the in vivo and in vitro experiments under the guidance of J.E.P.; and all authors synthesized the ideas and contributed to figures and manuscript writing. **Competing interests:** The Peters lab has corporate funding for research that is not directly related to the work in this publication. Cornell University has filed patent applications with J.E.P. as inventor involving CRISPR-Cas systems associated with transposons that are not directly related to this work. **Data and materials availability:** Atomic models are available through the Protein Data Bank (PDB) with accession codes 7M99 (ATPγS TnsC), 7M9A (ADP•$AlF_3$ TnsC consensus), 7M9C (ADP•$AlF_3$ TnsC open), 7M9B (ADP•$AlF_3$ TnsC closed), and 7N6I (TniQ-bound TnsC); all cryo-EM reconstructions are available through the EMDB with accession codes EMD-23724 (ATP TnsC head-to-tail), EMD-23725 (ATP TnsC head-to-head), EMD-23720 (ATPγS TnsC), EMD-23721 (ADP•$AlF_3$ TnsC consensus), EMD-23722 (ADP•$AlF_3$ TnsC open), EMD-23723 (ADP•$AlF_3$ TnsC closed), and EMD-23726 (TniQ-TnsC). NGS data are available from the NCBI Sequence Read Archive (Bioproject: PRJNA737449).

**SUPPLEMENTARY MATERIALS**

science.sciencemag.org/content/373/6556/768/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S3
Movies S1 and S2
References (*39*–*67*)
MDAR Reproducibility Checklist

8 April 2021; accepted 7 July 2021
Published online 15 July 2021
10.1126/science.abi8976

PLANT SCIENCE

# Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes

**Federico Sabbadin[1]*, Saioa Urresti[2], Bernard Henrissat[3,4,5], Anna O. Avrova[6], Lydia R. J. Welsh[6], Peter J. Lindley[2], Michael Csukai[7], Julie N. Squires[6], Paul H. Walton[2], Gideon J. Davies[2], Neil C. Bruce[1], Stephen C. Whisson[6], Simon J. McQueen-Mason[1]***

**The oomycete *Phytophthora infestans* is a damaging crop pathogen and a model organism to study plant-pathogen interactions. We report the discovery of a family of copper-dependent lytic polysaccharide monooxygenases (LPMOs) in plant pathogenic oomycetes and its role in plant infection by *P. infestans*. We show that LPMO-encoding genes are up-regulated early during infection and that the secreted enzymes oxidatively cleave the backbone of pectin, a charged polysaccharide in the plant cell wall. The crystal structure of the most abundant of these LPMOs sheds light on its ability to recognize and degrade pectin, and silencing the encoding gene in *P. infestans* inhibits infection of potato, indicating a role in host penetration. The identification of LPMOs as virulence factors in pathogenic oomycetes opens up opportunities in crop protection and food security.**

Oomycetes are fungal-like microorganisms that cause severe diseases in agriculture and aquaculture and pose a recurrent threat to global food security (*1*). The *Phytophthora* genus comprises more than 140 species (*2*), including the late blight pathogen *Phytophthora infestans*, which triggered the Irish potato famine in the mid-19th century. *P. infestans* infects both potato and tomato crops, causing economic losses in excess of $6 billion annually (*1*).

The plant cell wall, which represents the first protective barrier against pathogens, is a composite material made of cellulose microfibrils embedded in a matrix of hemicellulose and pectin (*3*). Pectin forms an acidic polysaccharide network and makes up the scaffold of the middle lamella that binds plant cells to one another. Plant pathogens deploy cell wall–degrading enzymes to penetrate this protective barrier (*4*). Once the plant cell wall is breached, *P. infestans* secretes intracellular effector proteins (*5*) that manipulate the host and its defense mechanisms. Despite their importance, cell wall–degrading enzymes have received less attention than intracellular effectors, leaving a gap in our understanding of the mechanisms underlying infection. Here, we show that phytopathogenic oomycetes harbor a family of secreted copper-bound lytic polysaccharide monooxygenases (LPMOs) that cleave pectin, enabling host penetration and infection by *P. infestans*.

LPMOs are copper-containing enzymes known for their ability to degrade recalcitrant and crystalline polysaccharides such as cellulose and chitin through oxidative attack (*6*–*9*). LPMOs require external electron donors in the form of small compounds (*9*) or redox protein partners (*10*) and have been characterized from fungi, bacteria, viruses, and invertebrates (*9*, *11*), mostly in the context of plant biomass degradation.

## Identification of an expanded LPMO family in phytopathogenic oomycetes

We identified a group of proteins that have no recognized functional domains but share substantial homology and carry predicted N-terminal signal peptides for secretion. The mature proteins feature an N-terminal histidine and a further conserved histidine, potentially forming a "histidine brace" (*7*), the hallmark of LPMOs. Homology searches against the National Center for Biotechnology Information (NCBI) nr databases (for nonredundant protein sequences) revealed that this gene family is only found in oomycetes and is expanded in hemi-biotrophic and necrotrophic plant pathogenic species, averaging 46 family members in *Phytophthora* spp. (Fig. 1A). This LPMO family now appears as auxiliary activity 17 (AA17) in the online Carbohydrate-Active enZYmes (CAZy) database (*12*).

Whereas most AA17 genes code for the catalytic domain alone, sequences from *Saprolegniales* (fish pathogens) typically have putative C-terminal cellulose-binding domains (*13*) and putative protein-protein or protein-carbohydrate interaction domains (*14*), and some from *Phytophthora parasitica* feature putative chitin-binding domains (*15*) (fig. S1A). With few exceptions, AA17 genes in plant pathogenic oomycetes form genomic clusters surrounded by transposable elements, similar to bacterial high-density pathogenicity islands involved in infection (*16*) (fig. S1B).

## AA17 genes are up-regulated during plant infection

Analysis of the published transcriptome of *Phytophthora palmivora*, a relative of *P. infestans* infecting *Nicotiana benthamiana* (*17*), revealed 42 AA17-coding genes, and more than half of them are induced >10-fold during infection (fig. S2 and data S1). We performed differential gene expression analysis of the transcriptome of *P. infestans* as it infected tomato plant leaves and found that AA17s were the most-induced CAZy family apparent during early infection (Fig. 1B and data S2), with PITG_04949 being the most up-regulated (Fig. 1C).

We used reverse transcription quantitative polymerase chain reaction (RT-qPCR) to investigate the induction of LPMO-coding genes (both AA16 and AA17) during infection of potato leaves (data S3). Known transcriptional markers for developing infection (*P. infestans* putative haustorium-specific membrane protein *Pihmp1*, *P. infestans* avirulence protein 3a

[1]Centre for Novel Agricultural Products, Department of Biology, University of York, York YO10 5DD, UK. [2]Department of Chemistry, University of York, York YO10 5DD, UK. [3]Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 7257 CNRS, Université Aix-Marseille, 163 Avenue de Luminy, 13288 Marseille, France. [4]INRA, USC 1408 AFMB, 13288 Marseille, France. [5]Department of Biological Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia. [6]Cell and Molecular Sciences, James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK. [7]Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire RG42 6EY, UK.
***Corresponding author. Email: federico.sabbadin@york.ac.uk (F.S.); simon.mcqueenmason@york.ac.uk (S.J.M.-M.)**

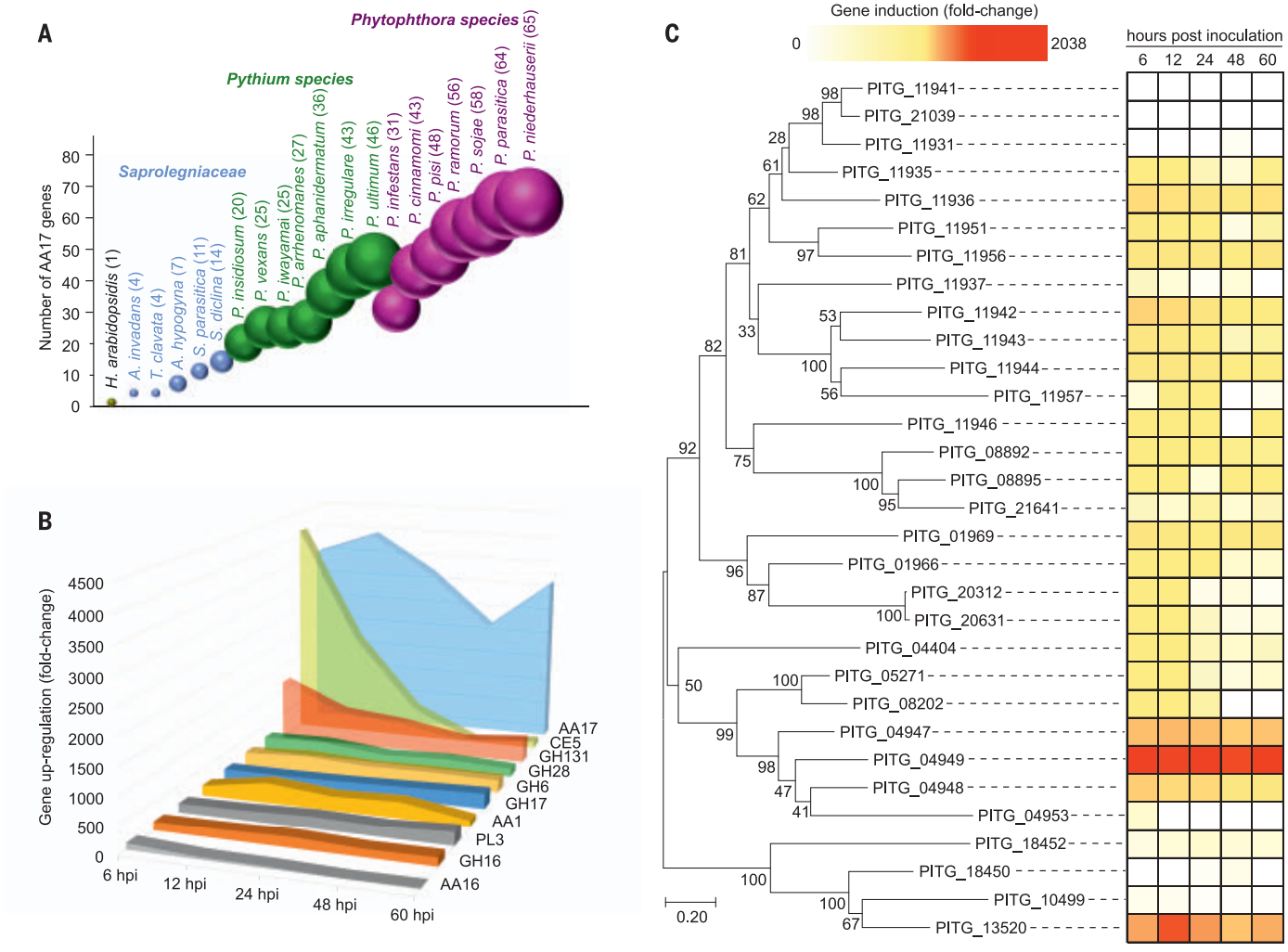



**Fig. 1. Taxonomic distribution of AA17 LPMO genes and induction during infection of tomato leaves.** (**A**) Number of AA17 genes across representative oomycete species. (**B**) Cumulative up-regulation of CAZy families, obtained from transcriptome sequencing of *P. infestans* infecting tomato leaves at 6, 12, 24, 48, and 60 hours postinoculation (hpi), relative to sporangia. (**C**) Neighbor joining phylogenetic tree of *P. infestans* AA17 LPMOs (predicted catalytic domains only) and gene induction levels on the basis of RNA sequencing (RNA-seq) data of *P. infestans* infecting tomato leaves during the time course experiment (6 to 60 hpi). The maximum fold change (2038) corresponds to gene PITG_04949 at 60 hpi. Bootstrap values (calculated using Mega7 with 1000 cycles) are shown at branching points.

*PiAvr3a*, and PITG_12808) showed up-regulation, as expected (*18*–*20*). Transcripts for six AA16 LPMOs were expressed at low levels or were undetectable. Of the 31 AA17 LPMO genes tested, the expression of 15 was detected in all replicates, and 11 exhibited greater than twofold up-regulation during infection, compared with expression in sporangia. The AA17 LPMO genes exhibiting greatest transcript abundance were PITG_04949, 04947, 20631/20312, 01966, and 13520.

Our data show that PITG_04949 (hereafter called *PiAA17C*) dominates in terms of gene induction in *P. infestans* infecting both tomato and potato leaves. *PiAA17C* is found in a gene cluster with three additional AA17 genes, PITG_04947 and PITG_04948 (henceforth called *PiAA17A* and *PiAA17B*, also induced during infection) and PITG_04953 (not substantially expressed during infection).

### *Pi*AA17C has a canonical type-2 copper center

The predicted mature proteins that are coded by AA17 genes in *P. infestans* typically feature an N-terminal LPMO domain followed by a variable polypeptide rich in serine, proline, and threonine residues. Analysis of these C-terminal extensions using IUPred2A (intrinsically unstructured/disordered proteins and domains prediction tool) (*21*) predicts them to be intrinsically disordered. We cloned the LPMO domains of *Pi*AA17A, -B, and -C; expressed them heterologously in *Escherichia coli*; and purified them using established methods (fig. S3A) (*9*). Correct protein folding and copper binding was confirmed by using thermal shift assays (fig. S3, B to D).

In vitro activity assays that used copper-loaded *Pi*AA17A, -B, and -C with a range of polysaccharides and reducing agents (ascorbic acid, gallic acid, and pyrogallol), followed by analysis by matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry (MALDI-TOF MS) and electrospray ionization mass spectrometry (ESI MS), failed to reveal any products released from the substrates tested, including those oxidized by characterized LPMOs (cellulose, cello-oligosaccharides, chitin, xylan, xyloglucan, and glucomannan) (*6*–*9*, *11*, *22*).

Despite the lack of activity by *Pi*AA17C on these substrates, electron paramagnetic resonance spectroscopy (EPR) revealed a spectral envelope consistent with a type-2 copper center with near axial symmetry ($g_z > g_y \approx g_x > g_e$, where the $g$ value is the proportionality factor)

(Fig. 2A and table S1) (*23*), confirming that *Pi*AA17C is an LPMO with a canonical histidine brace active site (*7*). The spin-Hamiltonian values are similar to those obtained for fungal AA9 LPMOs (*22*), indicating active high-valent copper-oxygen intermediates (e.g., copper-oxyl) as part of their catalytic cycle (*23*). Visible superhyperfine coupling is resolved in the perpendicular region of the spectrum and is attributed to coupling to the nitrogen nuclei in the *xy* plane, which ranges from 34 to 38 MHz, typical for an LPMO (*24*).

### The structure of *Pi*AA17C suggests an interaction with charged polysaccharides

To shed light on the potential substrate specificity and activity of *Pi*AA17C, we solved its x-ray crystal structure to 1.01-Å resolution. Structural similarity searches using Dali (*25*), which compares the C-alpha of the query structure with others available in the Protein Data Bank (PDB), gave the best match with chitin-active *Cj*LPMO10A from *Cellvibrio japonicus* [PDB ID: 5FJQ(A)], although the overall similarity is low (Dali *z*-score 9.5, 11% identity, 2.7-Å root mean square deviation over 118 aligned C-alphas). Akin to other LPMO families, the *Pi*AA17C structure reveals a central β-sandwich fold decorated with several loops and stabilized by three disulfide bonds (Fig. 2B and table S2). The active site features a histidine brace (*7*) formed by $His^{1}$ and $His^{96}$, which is required for copper coordination. The electrostatic surface potential and residue charge distribution of *Pi*AA17C are notably distinct from those of typical LPMOs active on crystalline cellulose or chitin, which display a flat surface surrounding the active site. By contrast, the His-brace of *Pi*AA17C lies within a cleft (Fig. 2C). Sequence conservation analysis (Fig. 2D) (*26*) highlighted conservation at the active site as well as several polar and negatively charged residues across the AA17 family, which suggests the ability to bind charged polysaccharides (Fig. 2E). The surface surrounding the active site also lacks the aromatic residues required for the recognition of uncharged polysaccharides in canonical LPMOs (*24*). Another unusual feature of *Pi*AA17C's structure is a large α helix (residues 36 to 51) (fig. S4) with a negatively charged groove engraved beneath it, leading to the active site. These negatively charged residues are reminiscent of the aspartate side chains involved in calcium coordination in homogalacturonate-active enzymes (*27*–*29*), and this prompted us to test our purified enzymes against negatively charged substrates.

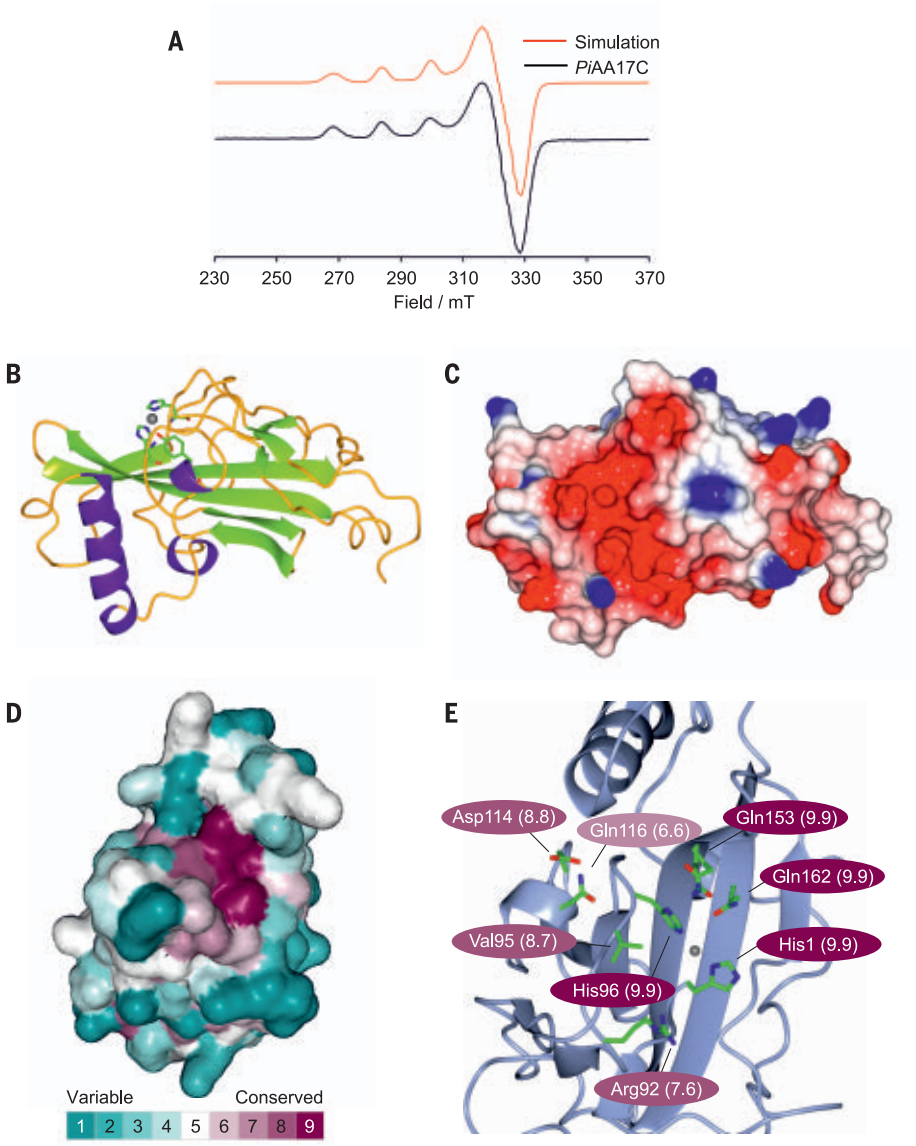



**Fig. 2. Structural and spectroscopic characterization of *Pi*AA17C.** (**A**) Continuous-wave X-band EPR spectrum of *Pi*AA17C (~0.5 mM) in sodium phosphate buffer (pH 7.0, 20 mM) collected at 150 K (black) and spectral simulation (red). (**B**) Overall structure of *Pi*AA17C, showing the antiparallel β sheet (in green) and histidine brace (sticks), featured in all LPMOs, and the long α helix (in purple), which is not present in other LPMO families (see fig. S4). (**C**) Electrostatic surface potential of *Pi*AA17C, showing the cleft surrounding the active site and the negatively charged groove (red) leading toward it. (**D**) Sequence conservation analysis (ConSurf, based on 401 AA17 sequences) of *Pi*AA17C looking down on the active site. The surface is colored by ConSurf score according to the indicated scoring scheme. (**E**) Highly conserved residues around the active site of *Pi*AA17C, based on ConSurf analysis of 401 AA17 sequences (LPMO domain only). Color shades indicate the level of conservation, calculated using ConSurf.

### AA17 LPMOs oxidatively cleave homogalacturonan

*Pi*AA17A, -B, and -C were incubated with a panel of charged polysaccharides in the presence of a range of reducing agents (ascorbic acid, gallic acid, and pyrogallol). MALDI-TOF MS and ESI MS analysis of the released products revealed distinct oxidized and native product peaks when using homogalacturonan (Fig. 3, A and B, and fig. S5, D, F, and H) and oligogalacturonides [degree of polymerization (DP) 10 to 15] (fig. S6, D, F, and H) as substrates. Homogalacturonan forms the backbone of pectin and consists of a linear chain of (1,4)-linked α-D-galacturonic acid units with variable degrees of methyl esterification (*3*). Activity on charged polysaccharides has not previously been reported for LPMOs. We observed that AA17s can accept electrons from ascorbic acid, but not from small phenolic compounds, whereas both types of molecules were used successfully with published LPMOs (*9*). Treatment of the released products with exo-polygalacturonase (GH28) acting on the nonreducing end (C4) degraded native oligogalacturonides, whereas peaks corresponding

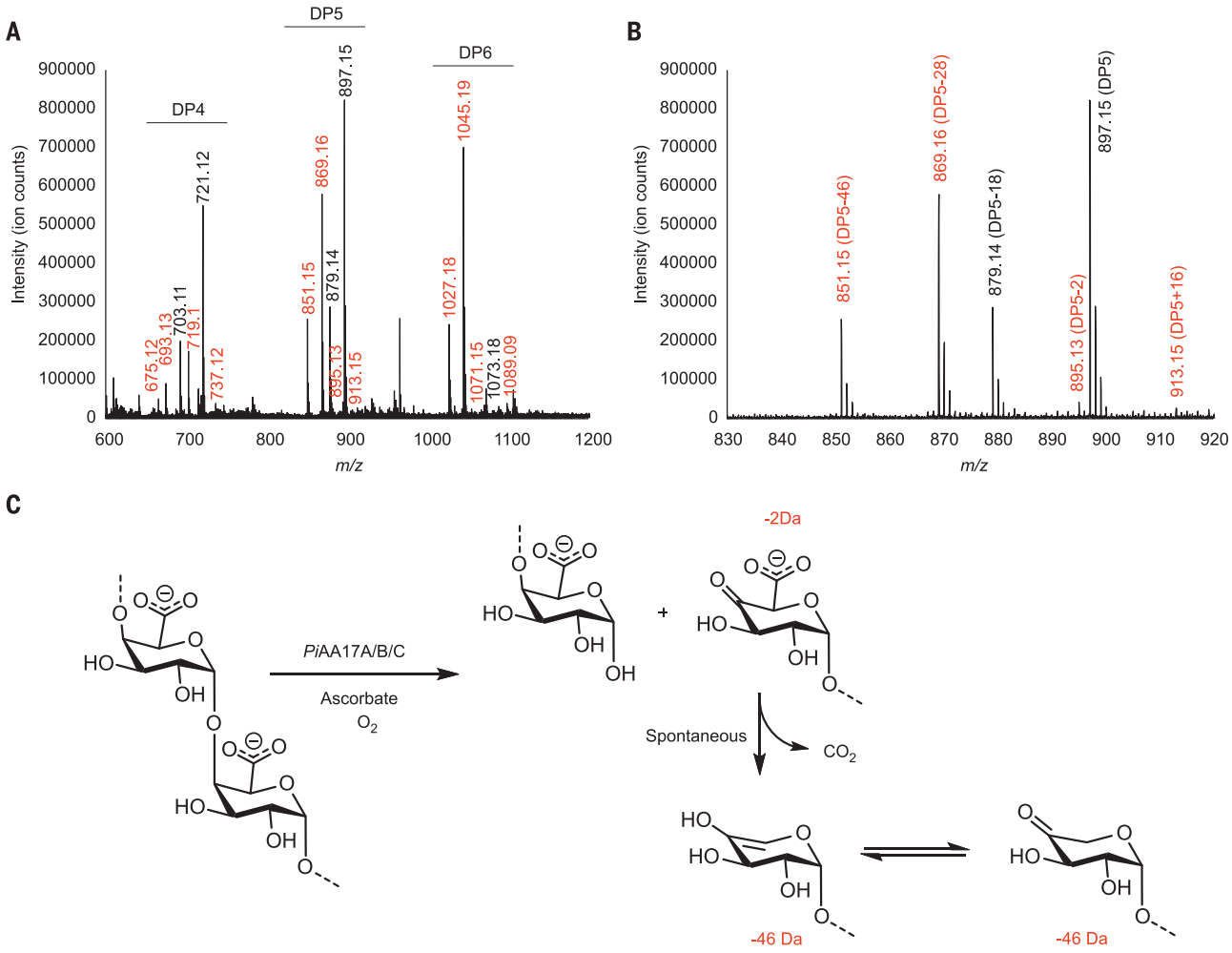



**Fig. 3. Biochemical characterization of *Pi*AA17C and proposed mechanism of action.** (**A**) Negative-mode ESI MS spectrum of products released by 2 μM *Pi*AA17C from 4 mg $ml^{-1}$ homogalacturonan in the presence of 4 mM ascorbic acid after 24-hour incubation (see materials and methods). Masses are indicated by numbers for the main peaks corresponding to native (black) and oxidized (red) oligogalacturonides. Control reactions with substrate only, substrate plus ascorbic acid, and substrate plus *Pi*AA17C did not generate detectable amounts of oligogalacturonides (see fig. S5). *m/z*, mass/charge ratio. (**B**) Expanded ESI MS spectra for DP5, showing the main peaks and their identity. *m/z* 897.15: native oligogalacturonide. *m/z* 895.13: −2 species, oxidized (ketone). *m/z* 913.15: +16 species, oxidized (gemdiol). *m/z* 879.14: −18 species, dehydrated native oligogalacturonide; *m/z* 851.15: −46 species, oxidized + decarboxylated oligogalacturonide (C4 ketone); *m/z* 869.16: −28 species, oxidized + decarboxylated oligogalacturonide (hydrated version of the C4 ketone, corresponding to 851.16 + 18 mass units, a gemdiol). The keto sugar is in equilibrium with the C4 gemdiol in aqueous solution, a feature common to keto saccharides, including C4 ketones generate by characterized LPMOs (*40*). (**C**) Proposed mechanism of action for *Pi*AA17A, -B, and -C acting on polygalacturonic acid. In the presence of ascorbic acid, the LPMO carries out a C4 oxidation, leading to the formation of a ketone in C4, followed by spontaneous decarboxylation of the resulting β-keto acid. The C4-enol undergoes tautomerization and formation of the more stable ketone form.

to the putative oxidized species were not altered (fig. S7), suggesting that *Pi*AA17C predominantly carries out C4-specific oxidation of homogalacturonan and that the resulting C4-oxidized oligogalacturonides are not accessible to the C4-acting exo-polygalacturonase. C4 oxidation also best explains the identity of two species (-46 and -28) not previously observed with characterized LPMOs. On the basis of peak masses from both MALDI-TOF MS and ESI MS analyses, we propose that *Pi*AA17A, -B, and -C carry out a C4-oxidative cleavage of polygalacturonic acid, generating a C4-ketone in a β position relative to one carboxylic group and resulting in an unstable β-keto acid that undergoes spontaneous decarboxylation and tautomerization (Fig. 3C). This mechanism is analogous to that proposed for uridine 5′-diphosphate (UDP)–glucuronic acid decarboxylase, for which the oxidation of UDP glucuronic acid in the C4 position generates a ketone that undergoes decarboxylation with formation of a UDP-4-keto pentose (*30*).

### AA17 LPMOs recognize the carboxylic groups of de-esterified pectin

We found that *Pi*AA17C was not active on esterified citrus pectin (fig. S8, E and F). However, preincubation of esterified pectin with pectin methylesterase, followed by addition of *Pi*AA17C, resulted in substrate degradation (fig. S8H), suggesting that carboxylic groups exposed through de-esterification are important for substrate recognition and cleavage by the LPMO. In addition, thermal shift analysis of *Pi*AA17A, -B, and -C revealed a specific interaction with homogalacturonan and oligogalacturonides, whereas interaction with highly esterified pectin was only detected after demethylation of the substrate (fig. S9 and table S3). The role of AA17 LPMOs in pectin degradation is supported by their co-induction with several putative pectin methylesterases (from families CE8 and CE13), as well as other pectin-active enzymes (families GH28, PL3, and PL4)

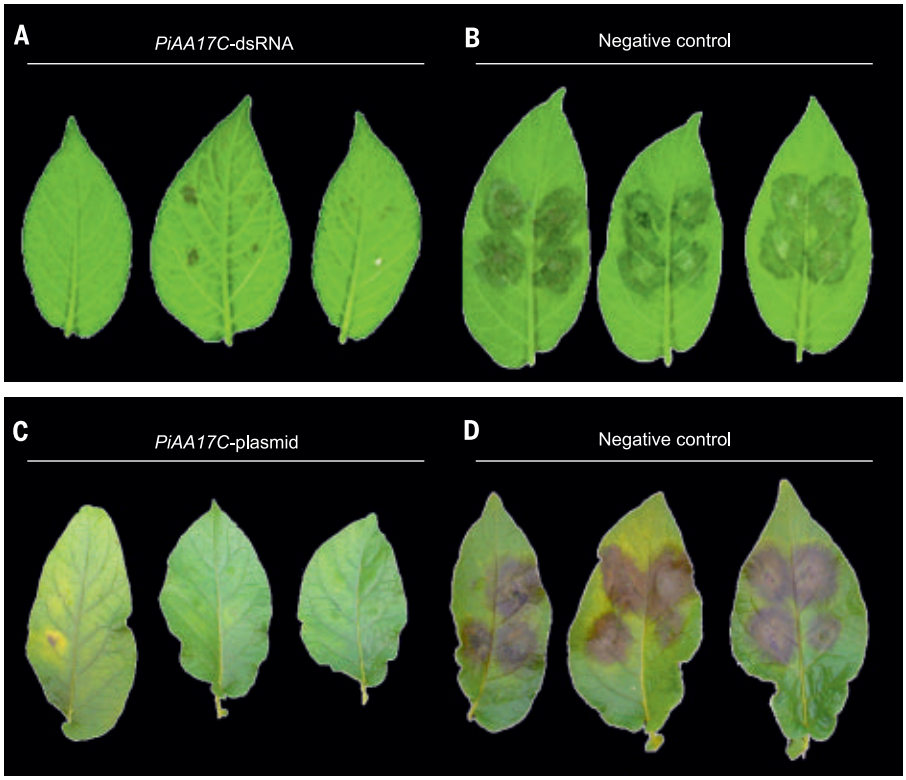



**Fig. 4. In vivo *PiAA17C* gene silencing during infection of *P. infestans* on plant leaves.** (**A** to **D**) Transient (A and B) and stable (C and D) silencing of *PiAA17C*. (A) Potato leaves infected with *P. infestans* isolate 88069 treated with dsRNA targeting *PiAA17C*. (B) Leaves infected with *P. infestans* isolate 88069 treated with nonhomologous dsRNA (negative control). (C) Leaves treated with *P. infestans* isolate 3928A stably transformed with silencing plasmid targeting *PiAA17C* (line IR6). (D) Leaves treated with wild-type *P. infestans* (negative control, isolate 3928A).

in our transcriptomic studies (data S2) and published data from other *Phytophthora* species (data S1) (*17*, *31*).

### *PiAA17C* is necessary for successful plant infection by *P. infestans*

We assessed the role of *Pi*AA17C in plant infection through transient gene silencing by delivering in vitro synthesized double-stranded RNA (dsRNA) into protoplasts of *P. infestans* (*32*), followed by colony regeneration and infection of potato leaves. Infection phenotypes recorded 72 to 96 hours postinoculation showed that the introduction of dsRNA targeting *PiAA17C* reduced virulence compared with control lines (Fig. 4, A and B). We verified this result using stable gene silencing in which *PiAA17C* was transformed into *P. infestans* as an inverted repeat (*33*), resulting in the removal of expression for the gene through the formation of heterochromatin. Six geneticin-resistant *P. infestans* lines that exhibited silencing of *PiAA17C* were selected (data S4) and showed a marked reduction in lesion size upon infection of potato leaves (Fig. 4, C and D, and fig. S10).

As seen previously for other genes silenced in *P. infestans* (*34*), the expression of nearby genes was also affected by the silencing of *PiAA17C*, with the greatest effect on the closest gene *PiAA17B* and a lesser impact on the more distant *PiAA17A* (data S4). Transcripts of *PiAA17C* are the most abundant in *P. infestans* during infection, and lesion sizes caused by the transgenic lines closely mirror the level of silencing of *PiAA17C* (fig. S10), whereas the differing levels of expression of *PiAA17A* and *PiAA17B* between replicates effectively rule them out as major contributors to the observed effect, validating the results of the transient gene silencing experiment.

## Discussion

Our results suggest that the evolutionary arms race with plants has spurred oomycetes to evolve LPMOs as virulence factors to overcome host defenses. Although AA17s likely have a role in facilitating host penetration by cleaving pectin and disrupting the plant cell wall network, we speculate that they further enhance *P. infestans* virulence by interfering with host immunity. Oligogalacturonides released by pathogen polygalacturonase enzymes are well-characterized inducers of plant responses to pathogens, and the oxidation of oligogalacturonides by berberine bridge enzyme–like proteins has been shown to hinder their recognition by the plant, which prevents the activation of immune responses (*35*). AA17s generate oxidized and decarboxylated oligogalacturonides while degrading the pectin backbone and might play a role in avoiding the plant immune response while simultaneously breaching the host cell wall.

Transcriptomic analyses have shown upregulation of some AA9-coding genes during fungal invasion of plant tissue (*36*), which suggests a role in pathogenesis. Recently, a marked induction of immunity-related genes was observed in *Arabidopsis thaliana* treated with native and oxidized cellulose oligosaccharides produced by a fungal AA9 LPMO (*37*). The different physicochemical properties of charged oligogalacturonides released by berberine bridge enzyme–like proteins and AA17s, compared with uncharged cello-oligosaccharides released by AA9s, may hold the key to the two seemingly contrasting effects on the host immune system.

The amount of ascorbic acid that we found to be effective in our assays (1 to 4 mM) is comparable to physiological levels measured in potato and tomato plants (*38*, *39*). The observation that *Pi*AA17A, -B, and -C activity is driven by ascorbic acid but not phenolic compounds may reflect an adaptation of the *P. infestans* life cycle, in which it invades fresh host tissue (rich in cellular reductants, including ascorbic acid) rather than lignified tissue (a plentiful source of phenolics, compatible with characterized LPMO families from wood-decay fungi).

Our results shed light on the complex interactions between hosts and pathogens in the plant cell wall and illustrate how targeting LPMO genes through RNA-based approaches could provide a strategy to fight crop diseases and increase agricultural productivity.

## REFERENCES AND NOTES

1. L. Derevnina *et al.*, *Philos. Trans. R. Soc. B* **371**, 20150459 (2016).
2. X. Yang, B. M. Tyler, C. Hong, *IMA Fungus* **8**, 355–384 (2017).
3. R. A. Burton, M. J. Gidley, G. B. Fincher, *Nat. Chem. Biol.* **6**, 724–732 (2010).
4. C. P. Kubicek, T. L. Starr, N. L. Glass, *Annu. Rev. Phytopathol.* **52**, 427–451 (2014).
5. S. C. Whisson *et al.*, *Nature* **450**, 115–118 (2007).
6. G. Vaaje-Kolstad *et al.*, *Science* **330**, 219–222 (2010).
7. R. J. Quinlan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15079–15084 (2011).
8. G. R. Hemsworth *et al.*, *J. Am. Chem. Soc.* **135**, 6069–6077 (2013).
9. F. Sabbadin *et al.*, *Nat. Commun.* **9**, 756 (2018).
10. T. C. Tan *et al.*, *Nat. Commun.* **6**, 7542 (2015).
11. Z. Forsberg *et al.*, *Curr. Opin. Struct. Biol.* **59**, 54–64 (2019).
12. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **42**, D490–D495 (2014).
13. A. Várnai *et al.*, *Adv. Appl. Microbiol.* **88**, 103–165 (2014).
14. H. Tordai, L. Bányai, L. Patthy, *FEBS Lett.* **461**, 63–67 (1999).
15. T. Suetake *et al.*, *J. Biol. Chem.* **275**, 17929–17932 (2000).
16. O. Gal-Mor, B. B. Finlay, *Cell. Microbiol.* **8**, 1707–1719 (2006).
17. E. Evangelisti *et al.*, *BMC Biol.* **15**, 39 (2017).
18. A. O. Avrova *et al.*, *Cell. Microbiol.* **10**, 2271–2284 (2008).
19. M. R. Armstrong *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7766–7771 (2005).
20. M. Abrahamian, A. M. V. Ah-Fong, C. Davis, K. Andreeva, H. S. Judelson, *PLOS Pathog.* **12**, e1006097 (2016).
21. B. Mészáros, G. Erdos, Z. Dosztányi, *Nucleic Acids Res.* **46**, W329–W337 (2018).
22. T. J. Simmons *et al.*, *Nat. Commun.* **8**, 1064 (2017).
23. G. R. Hemsworth, L. Ciano, G. J. Davies, P. H. Walton, *Methods Enzymol.* **613**, 63–90 (2018).

24. K. E. H. Frandsen *et al*., *Nat. Chem. Biol.* **12**, 298–303 (2016).
25. L. Holm, P. Rosenström, *Nucleic Acids Res.* **38**, W545–W549 (2010).
26. H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, *Nucleic Acids Res.* **38**, W529–W533 (2010).
27. S. J. Charnock, I. E. Brown, J. P. Turkenburg, G. W. Black, G. J. Davies, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12067–12072 (2002).
28. M. D. Yoder, N. T. Keen, F. Jurnak, *Science* **260**, 1503–1507 (1993).
29. R. Pickersgill, J. Jenkins, G. Harris, W. Nasser, J. Robert-Baudouy, *Nat. Struct. Biol.* **1**, 717–723 (1994).
30. M. Bar-Peled, C. L. Griffith, T. L. Doering, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12003–12008 (2001).
31. L. M. Blackman, D. P. Cullerne, P. Torreña, J. Taylor, A. R. Hardham, *PLOS ONE* **10**, e0136899 (2015).
32. S. C. Whisson, A. O. Avrova, P. van West, J. T. Jones, *Mol. Plant Pathol.* **6**, 153–163 (2005).
33. H. S. Judelson, S. Tani, *Eukaryot. Cell* **6**, 1200–1209 (2007).
34. A. L. Vu, W. Leesutthiphonchai, A. M. V. Ah-Fong, H. S. Judelson, *Mol. Plant Microbe Interact.* **32**, 915–927 (2019).
35. M. Benedetti *et al*., *Plant J.* **94**, 260–273 (2018).
36. G. Jagadeeswaran, L. Veale, A. J. Mort, *Trends Plant Sci.* **26**, 142–155 (2021).
37. M. Zarattini *et al*., *Commun. Biol.* **4**, 727 (2021).
38. J. S. Han, N. Kozukue, K. S. Young, K. R. Lee, M. Friedman, *J. Agric. Food Chem.* **52**, 6516–6521 (2004).
39. H. Gautier, C. Massot, R. Stevens, S. Sérino, M. Génard, *Ann. Bot.* **103**, 495–504 (2009).
40. T. Isaksen *et al*., *J. Biol. Chem.* **289**, 2632–2642 (2014).
41. P. H. Walton, P. Lindley, F. Sabbadin, G. J. Davies, EPR data for "Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes," York Research Database (2021); doi: 10.15124/0ee2a9c1-6d3b-4ee0-8ac2-c19061a40d94.



### ACKNOWLEDGMENTS

We thank J. Agirre for his help with comparisons of *Pi*AA17C with all available LPMO structures. We thank Diamond Light Source for access to beamline IO4 (proposal number mx-9948), which contributed to the results presented here. **Funding:** This work was funded by the UK Biotechnology and Biological Sciences Research Council (grants BB/L001926/1, BB/J016500/1, BB/L021633/1, BB/V000365/1, and BB/V000675/1). The York Centre of Excellence in Mass Spectrometry was created with funds from Science City York, Yorkshire Forward, and the Northern Way Initiative and by EPSRC (EP/K039660/1; EP/M028127/1). A.O.A., S.C.W., L.R.J.W., and J.N.S. acknowledge Syngenta and the Scottish Government Rural and Environment Science and Analytical Services Division. **Author contributions:** F.S. carried out analysis of RNA-seq data, gene cloning, heterologous protein expression and purification, enzyme activity assays, and mass spectrometry analysis of reaction products and prepared figures and tables. S.U. and G.J.D. conceived the x-ray crystallography studies. S.U. crystallized the proteins, collected and analyzed crystallographic data, solved the crystal structures, and made structural figures and tables. P.H.W. and P.J.L. conceived the EPR study. P.J.L. carried out EPR experiments and simulations. B.H. performed bioinformatics analyses and alignments. M.C., S.C.W., and J.N.S. conceived and performed the RNA-seq experiments and analyzed the data. S.C.W. and L.R.J.W. performed the stable gene silencing experiments. L.R.J.W. performed RT-qPCR. A.O.A. conceived and performed the transient gene silencing experiments. F.S., S.C.W., N.C.B., and S.J.M.-M. organized the data and wrote the manuscript. All authors contributed to production of the manuscript. **Competing interests:** The authors declare no competing interests. B.H. is now affiliated with the Technical University of Denmark. **Data and materials availability:** Coordinates and structure factors for the x-ray structure of *Pi*AA17C were deposited in the PDB under accession code 6Z5Y. Raw EPR data are available through the York Research Database (*41*). *P. infestans* raw RNA-seq data are available in the NCBI Sequence Read Archive under BioProject PRJNA739688, accession numbers SRR14871460 to SRR14871482. All other data are in the main paper or supplementary materials.


### SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/774/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S3
References (*42*–*58*)
MDAR Reproducibility Checklist
Data S1 to S4

22 April 2021; accepted 6 July 2021
10.1126/science.abj1342

REPORTS

ULTRACOLD MOLECULES

# Observation of microwave shielding of ultracold molecules

**Loïc Anderegg[1,2]*, Sean Burchesky[1,2], Yicheng Bao[1,2], Scarlett S. Yu[1,2], Tijs Karman[3,4], Eunmi Chae[5], Kang-Kuen Ni[1,2,6], Wolfgang Ketterle[2,7], John M. Doyle[1,2]**

**Harnessing the potential wide-ranging quantum science applications of molecules will require control of their interactions. Here, we used microwave radiation to directly engineer and tune the interaction potentials between ultracold calcium monofluoride (CaF) molecules. By merging two optical tweezers, each containing a single molecule, we probed collisions in three dimensions. The correct combination of microwave frequency and power created an effective repulsive shield, which suppressed the inelastic loss rate by a factor of six, in agreement with theoretical calculations. The demonstrated microwave shielding shows a general route to the creation of long-lived, dense samples of ultracold polar molecules and evaporative cooling.**

Applications of ultracold molecules in quantum simulation, precision measurement, ultracold chemistry, and quantum computation (*1*, *2*) have led to rapid progress in direct cooling (*3*–*7*), assembly (*8*–*16*), trapping (*17*–*19*), and control of molecules (*20*–*23*). Engineering and control of molecular interactions will enable or enhance many of these applications. In particular, collisional interactions play a critical role in the ability to cool and, therefore, control molecules. Although there has been some success with sympathetic and evaporative cooling (*24*, *25*), these efforts have been hindered by large inelastic loss rates for both reactive and nonreactive molecular species in optical traps (*26*–*29*). Suppressing these inelastic losses, and, more generally, tuning interactions, is key to effective evaporative cooling of ultracold molecules and quantum applications.

A path to achieving this suppression is shielding, whereby molecules can be repelled from short-range distances where inelastic processes occur. Various shielding schemes for atoms (*30*–*33*) and molecules have been proposed (*34*–*39*). Recently, a scheme using dc electric fields to generate repulsive interactions was demonstrated in a two-dimensional geometry for KRb molecules (*40*). Here, we report microwave shielding of $^{40}Ca^{19}F$ molecules in three dimensions using optical tweezer traps. By tuning the microwave frequency from blue to red detuned, the system switches from shielded to "antishielded," changing the inelastic collision rate by a factor of 24, in agreement with theoretical calculations.

The microwave shielding mechanism studied here works as follows: Continuous, near-resonant microwave fields dress the molecular states, generating an oscillating dipole moment in the CaF molecule that gives rise to strong, long-range dipolar interactions (*41*). With the correct microwave dressing, this interaction is repulsive. Additionally, the dipolar interaction substantially enhances the elastic collision rate, resulting in a high elastic-to-inelastic collisional ratio, which is a key feature for evaporative cooling. In the microwave shielding scheme we used, the uppermost dressed state adiabatically converts to the repulsive branch of the dipole-dipole interaction. This repulsion leads to a classical turning point at long range (*37*, *38*) (Fig. 1A), preventing molecules from reaching short range where they would be lost with high probability (*29*). There will be residual inelastic loss at long range, predicted to be a result of nonadiabatic transitions (so-called "microwave-induced loss") (*37*, *38*). Coupled-channel calculations have shown that effective shielding requires circular polarization and high Rabi frequencies of the microwave field (*37*, *38*). Circular polarization provides coupling to the repulsive branch of the resonant dipole-dipole interaction regardless of the orientation of the collision axis relative to the molecule orientation, resulting in shielding in the bulk, that is, three dimensions. Rabi frequencies could be made high enough to create a large gap between field-dressed levels, ensuring adiabaticity during the collision.

Our experiment started from a magneto-optical trap of $^{40}Ca^{19}F$ molecules (*6*). CaF may

[1]Department of Physics, Harvard University, Cambridge, MA, USA. [2]Harvard-MIT Center for Ultracold Atoms, Cambridge, MA, USA. [3]ITAMP, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA. [4]Radboud University, Institute for Molecules and Materials, Heijendaalseweg 135, 6525 AJ Nijmegen, Netherlands. [5]Department of Physics, Korea University, Seongbuk-gu, Seoul, Republic of Korea. [6]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [7]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA.
*Corresponding author. Email: anderegg@g.harvard.edu

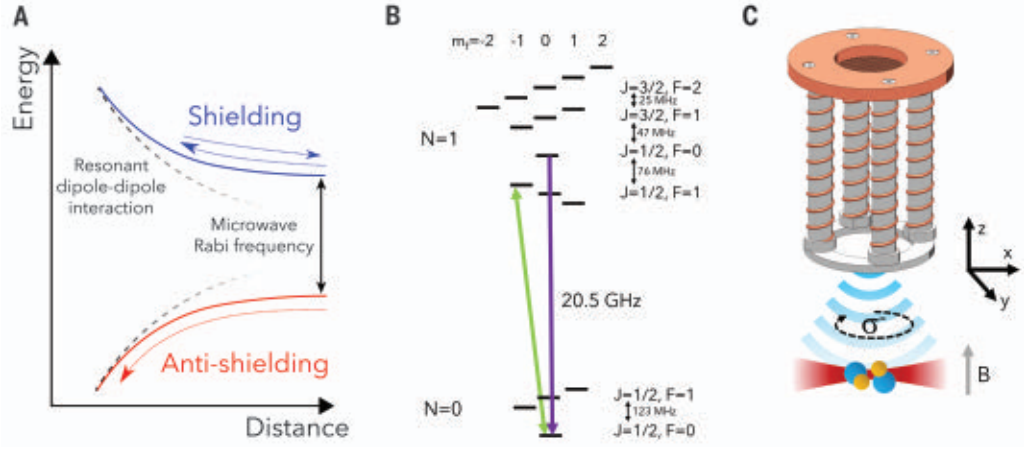



**Fig. 1. Microwave shielding overview.** (**A**) Diagram of the shielding process. The upper dressed state leads to a repulsive potential, preventing the molecules from reaching short range and undergoing loss. (**B**) CaF energy levels of the $X^2\Sigma^+$ electronic ground state showing the N = 1 and N = 0 rotational states separated by 20.5 GHz. The shielding transition is shown in green. The purple arrow shows the Landau-Zener sweep to the absolute ground state. (**C**) Experimental schematic showing the relative orientations of the helical antenna with respect to the tweezers. The tweezer light is linearly polarized along the *z* axis, parallel to the magnetic field *B*.

be laser cooled, optically manipulated, and imaged owing to its closed optical cycling transitions. We used Λ-enhanced gray molasses cooling to load molecules into a conservative 1064-nm optical dipole trap (*17*, *42*). Single molecules were then transferred into two 780-nm optical tweezer traps (*43*). The tweezers had a beam waist of about 1.6 μm and a depth of 1.8 mK. Light-assisted collisions caused by the Λ-cooling light during tweezer loading ensured no more than a single molecule in each tweezer. As detailed previously (*29*), the two tweezer traps can be merged to create a colliding pair of molecules in a single trap. This merge was accomplished by using a single 780-nm laser source to create one stationary trap and by using an acousto-optical deflector to create one steerable trap. The light for these two traps was combined and then focused down, forming two tightly focused tweezer traps [radial frequency ($\omega_r$) = $2\pi \times 91.5$ kHz] that could be merged. We measured a typical molecule temperature of 96 μK (*44*), both before and after merging. The molecules occupied many spatial modes, and therefore the collisions were three-dimensional in nature.

Once the tweezers were loaded, we applied an optical pumping pulse to populate the $|N = 1, J = 1/2, F = 0, m_f = 0\rangle$ state (Fig. 1B), where $N$ is the rotational angular momentum, $J$ is the total angular momentum, $F$ is the total angular momentum including nuclear spin, and $m_f$ is its projection onto the magnetic field. Next, we used a Landau-Zener microwave sweep to move the population to the absolute ground state $|N = 0, J = 1/2, F = 0, m_f = 0\rangle$. This transition is nominally dipole forbidden, but an applied 4-G magnetic field was mixed in the $|N = 0, F = 1, m_f = 0\rangle$ state, providing a substantial transition dipole moment. To remove any remaining population in the N = 1 rotational level, we applied a 5-ms pulse of resonant light, recoil heating the N = 1 molecules out of the tweezer trap. The two molecules, both in the ground internal state, were then merged together into a single tweezer for collisions to take place. At this point, the shielding was turned on for a variable amount of time before the tweezers were separated, and the molecules were transferred back to the N = 1 manifold for imaging. The molecules are imaged using Λ-imaging (*42*, *43*) to verify their survival.

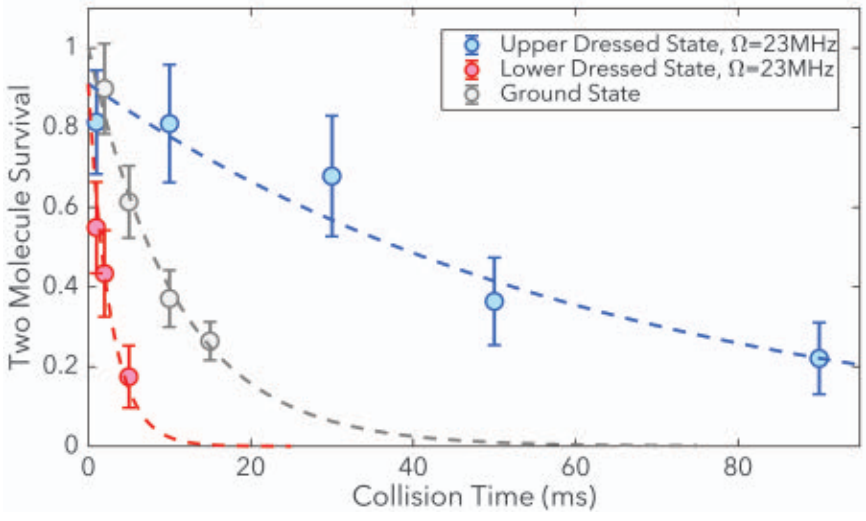



**Fig. 2. Microwave shielding of CaF collisions.** The gray trace (10.8 ms) shows the bare two-body loss of unshielded ground-state collisions. The blue trace (64 ms) shows the shielded loss rate at a Rabi frequency of 23 MHz and magnetic field of 27 G in the upper dressed state. The red trace (2.7 ms) shows the loss rate in the lower dressed state with a Rabi frequency of 23 MHz and magnetic field of 27 G. The error bars represent the standard error.

Circularly polarized microwaves were generated by a two-by-two helical antenna array (*45*) (Fig. 1C). The helical antennas were designed for axial mode operation, creating circular polarization with a helicity set by the winding of the antenna. An array was used to increase the cleanliness of the circular polarization and the overall output power. The 20.5-GHz microwaves were generated from mixing a low–phase noise 18.5-GHz source with a 2-GHz source locked to a low-noise oscillator (*44*). The 20.5-GHz signal was then amplified and split into four paths of equal length through phase-stable cables. Each antenna had a separate 5-W microwave amplifier and a mechanical phase shifter. The microwaves propagated into a stainless-steel vacuum chamber through a glass window along the *z* axis, defined as the direction of the magnetic field (Fig. 1C). We determined the polarization of the microwave field by measuring the Rabi

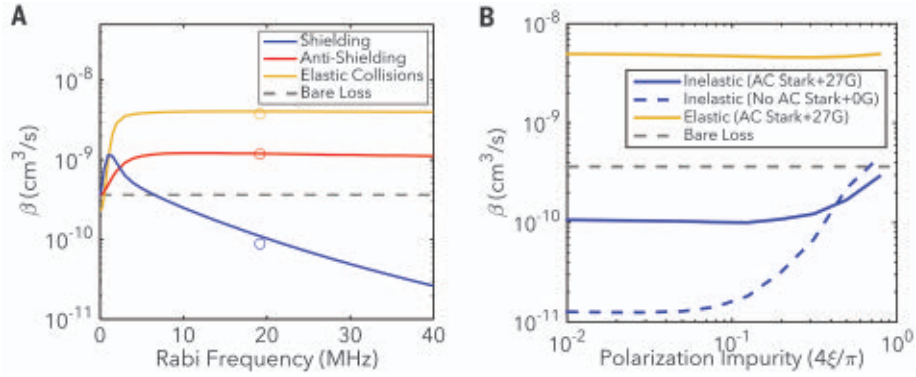



**Fig. 3. Theory calculations of shielding parameters.** (**A**) Rate coefficient versus Rabi frequency. Both shielding (blue) and antishielding (red) are shown. The elastic rate is shown in yellow. The solid lines are results without including spin; the circles include spin. (**B**) Plot of loss versus microwave ellipticity angle, ξ (*44*), for a Rabi frequency of 23 MHz. The tweezer trap's tensor ac Stark shift was the dominant factor reducing the effective degree of shielding. The elastic rate was nearly unaffected by the ac Stark shift and magnetic field.

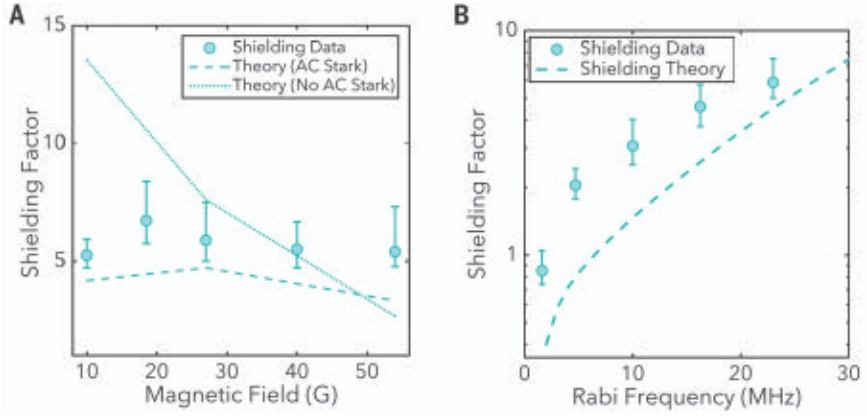



**Fig. 4. Dependence of shielding on microwave power and magnetic field.** The shielding factor is the ratio of the bare loss rate to the measured loss rate. (**A**) Shielding factor versus magnetic field at a Rabi frequency of 23 MHz. We found the effect of shielding to be robust over this range of magnetic fields, a result of the tensor ac Stark shift from the trap light. (**B**) Shielding factor versus Rabi frequency. We found the crossover point where shielding began to be around 3 MHz. The experimental data were taken at 27 G, and the theory curve did not include the effect of spin. The error bars represent the standard error of the fitted value.

frequency of the $\sigma^+$, $\sigma^-$, and $\pi$ transitions between the states $|N=1,J=1/2,F=0,m_f=0\rangle$ and $|N=0,J=1/2,F=1,m_f=\pm1,0\rangle$. Accounting for the magnetic field–dependent matrix elements, the $\sigma^+$ and $\sigma^-$ field components indicated the degree of circular polarization in the plane transverse to the axial magnetic field, and the $\pi$ component of the field was related to the tilt angle of the polarization ellipse relative to the $z$ axis. Using the measured Rabi frequencies, we then adjusted the phases of the four individual antennas to maximize the target circular field component and minimized the other two polarizations. The helical antenna array generated clean circular polarization in free space; however, the reflections from metal components in and around the vacuum chamber degraded the polarization cleanliness to a power ratio of right- to left-handed circular polarization of 100 (*44*).

To create collisional shielding, we used the $|N=1,J=1/2,F=1,m_f\rangle$ hyperfine manifold, with an applied 27-G magnetic field. The magnetic field direction was such that the upper state in the manifold $|N=1,J=1/2,F=1,m_f=-1\rangle$ was driven by the high-purity circular polarization. After merging the tweezers, we prepared the upper dressed state by switching on low-power microwaves with a frequency of a few megahertz blue detuned to $|N=1,J=1/2,F=1,m_f=-1\rangle$. Then, adiabatically, the amplitude of the microwaves was ramped to full power in ~100 μs while the detuning remained fixed. In this way, we produced near-resonant dressing where the detuning was small with respect to the Rabi frequency. Using the highest-energy $m_f$ level ensured that the upper dressed state did not cross any other levels as the microwave power was ramped up. The lifetime of the single-particle dressed state in the optical tweezer was limited to >500 ms by the phase noise of the microwave source.

The collisional lifetime of the bare ground state was the reference for our shielding performance comparison. We measured the trap frequency and temperature of the bare ground state molecules to be the same as molecules prepared in the upper dressed state, thus ensuring that the densities of microwave-dressed molecules and bare ground-state molecules were comparable. The ratio of the measured lifetimes was thus the ratio of the two-body loss rates.

At a Rabi frequency of 23 MHz, and in the upper dressed state (blue detuned from the $|N=0,F=0,m_f=0\rangle$ to $|N=1,J=1/2,F=1,m_f=-1\rangle$ transition) in a 27-G field (Fig. 2), the shielded lifetime was 64 ms [two-body rate coefficient ($\beta$) = 7.2(2.0) × $10^{-11}$ $cm^3/s$], six times longer than the bare ground-state lifetime of 10.8 ms [$\beta$ = 4.2(0.8) × $10^{-10}$ $cm^3/s$]. The ratio of the lifetimes was in agreement with a coupled-channel loss-rate calculation (Fig. 3). We found experimentally that the shielded lifetime was relatively independent of the polarization purity from 100:1 to 10:1 in power at a Rabi frequency of ~23 MHz (*44*).

Although the upper dressed state produced a repulsive shielding potential, the lower dressed state adiabatically connected to the attractive branch of the dipole-dipole interaction as the molecules approached during the collision, causing antishielding (*37*, *38*). Guided by this theory, we prepared the lower dressed state by flipping the direction of the magnetic field, which effectively swapped the handedness of the microwaves such that the lowest $m_f$ level ($|N=1,J=1/2,F=1,m_f=+1\rangle$) was now the one being driven most strongly by the circularly polarized microwaves. We prepared the lower dressed state with a microwave power ramp, with the microwaves red detuned. We measured this antishielded lifetime to be 2.7 ms [$\beta$ = 1.7(0.5) × $10^{-9}$ $cm^3/s$], faster than the bare ground state by a factor of about 4 and faster than the shielded state by a factor of 24 (see Fig. 2).

We used coupled-channel methods to calculate microwave shielding of CaF molecules. Similar to previous work (*38*), the colliding molecules were modeled as rigid rotors interacting through dipole-dipole interactions and with external magnetic and microwave fields. For details, see (*44*). In contrast to previous studies, we included the tensor ac Stark shift caused by the intense tweezer light. At short range, a fully absorbing boundary condition

was imposed that yielded universal loss in the absence of microwave dressing. Nonadiabatic transitions between dressed states led to microwave-induced loss, whereby the microwave Rabi frequency was converted into kinetic energy. Short-range losses occurred to a lesser extent because the potentials involved were mainly repulsive. The results of two sets of calculations are shown in Fig. 3. First, we assumed that the microwave polarization was perfectly circular about the magnetic field and tweezer polarization directions (Fig. 3A). Cylindrical symmetry was exploited to expedite the calculations. Second, the ellipticity of the microwave polarization was added (Fig. 3B), breaking cylindrical symmetry, which made the computations more demanding such that explicitly accounting for (hyper)fine structure became intractable. Hence, we made the approximation of treating spin implicitly by an enhanced rotational *g*-factor and tested the accuracy of this approximation in Fig. 3A. As discussed in (*44*), this approximation was reasonable for low magnetic fields and in the presence of a dominant tensor ac Stark interaction.

Previous theoretical work on microwave shielding (*37*, *38*) indicated a strong dependence on the polarization. The tensor Stark shift due to the optical tweezer light aligned the molecules, which competed with the resonant dipolar interactions that lead to shielding. The coupled-channel calculations performed here indicated that this fact limited shielding for perfectly circular polarization but reduced the sensitivity to polarization imperfections (see Fig. 3B). The $|N = 1, J = 1/2, F = 1, m_f\rangle$ hyperfine manifold, used for shielding, had a substantial tensor polarizability (*44*). The tweezer in which the collisions took place was linearly polarized along the $z$ axis, parallel to the magnetic field and perpendicular to the plane containing the polarization ellipse of the microwaves. At a tweezer trap depth of 1.8 mK for the ground state and no applied magnetic field, we observed a splitting of 10 MHz between the $m_f = 0$ and $m_f = \pm1$ states. This tensor Stark shift was the dominant limiting factor of the observed shielding process.

It has been shown previously (*37*) that the CaF($^2\Sigma$) fine structure can limit the effectiveness of microwave shielding, leading to losses enhanced by orders of magnitude compared with calculations neglecting (hyper)fine structure or to $^1\Sigma$ bialkali molecules, with much weaker hyperfine interactions. Here, we used the $|N = 0, J = 1/2, F = 0\rangle \rightarrow |N = 1, J = 1/2, F = 1\rangle$ transition, see Fig. 1A, to achieve shielding. At low magnetic fields, the electron and nuclear spins in these states are approximately coupled to a total spin singlet, which effectively eliminates the fine-structure couplings. Our calculations (Fig. 4A) predicted an enhancement of the shielding at low magnetic field but only in the absence of tensor ac Stark shifts. Including tensor ac Stark shifts, this interaction dominated the loss at low magnetic field, and the interplay between these two effects resulted in only a weak magnetic field dependence; see (*44*) for a detailed discussion. This result was in agreement with experimentally observed shielding lifetimes that were similar for 10 to 54 G (Fig. 4A).

As predicted from theory, the shielding effect showed a clear power dependence (Fig. 4B). The loss rate increased from $\beta = 7.2(2.0) \times 10^{-11}$ cm$^3$/s to $\beta = 2.1(0.5) \times 10^{-10}$ cm$^3$/s when the microwave power was reduced from 23 to 5 MHz of Rabi frequency on the $\sigma^-$ transition, keeping the polarization unchanged. Decreasing the power further to 1.5 MHz of Rabi frequency, we measured an increased loss rate of $\beta = 5.0(1.1) \times 10^{-10}$ cm$^3$/s, slightly higher than the bare ground state. The losses measured were almost entirely from long-range microwave-driven nonadiabatic transitions; therefore, the bare ground-state loss rate is not a lower bound (*37*, *38*). As the gap between the upper dressed state providing the repulsive potential and the lower dressed state decreased, the loss rate at low Rabi frequency was faster than the bare ground-state loss rate.

We demonstrated microwave shielding of inelastic collisions in three dimensions with ultracold CaF molecules in an optical tweezer trap. The relative ratios of experimentally measured two-body lifetimes agreed well with results of coupled-channel calculations and the qualitative features of shielding theory. By blue detuning the microwaves to prepare the shielded upper dressed state, we observed a factor of six suppression of inelastic loss relative to the bare ground state. By red detuning, we created an antishielded lower dressed state, leading to an enhanced loss rate. This shielding mechanism may be extended to a wide range of polar molecules prepared in a single quantum state (*37*, *38*), including polyatomic molecules (*46*, *47*). It is notable that the predicted elastic scattering rate with microwave dressing was greatly enhanced, leading to a ratio ($\gamma$) of elastic to inelastic rates of more than 50, which is more than enough for effective direct evaporative cooling. Theory also suggested that the shielding was limited by ac Stark shifts from the tweezer traps, indicating the possibility of improved shielding by using lower optical trap intensities.

## REFERENCES AND NOTES


1. L. D. Carr, D. DeMille, R. V. Krems, J. Ye, *New J. Phys.* **11**, 055049 (2009).
2. J. L. Bohn, A. M. Rey, J. Ye, *Science* **357**, 1002–1010 (2017).
3. E. S. Shuman, J. F. Barry, D. Demille, *Nature* **467**, 820–823 (2010).
4. J. F. Barry, D. J. McCarron, E. B. Norrgard, M. H. Steinecker, D. DeMille, *Nature* **512**, 286–289 (2014).
5. S. Truppe *et al.*, *Nat. Phys.* **13**, 1173–1176 (2017).
6. L. Anderegg *et al.*, *Phys. Rev. Lett.* **119**, 103201 (2017).
7. A. L. Collopy *et al.*, *Phys. Rev. Lett.* **121**, 213201 (2018).
8. K.-K. Ni *et al.*, *Science* **322**, 231–235 (2008).
9. T. Takekoshi *et al.*, *Phys. Rev. Lett.* **113**, 205301 (2014).
10. J. W. Park, S. A. Will, M. W. Zwierlein, *Phys. Rev. Lett.* **114**, 205302 (2015).
11. P. D. Gregory, J. Aldegunde, J. M. Hutson, S. L. Cornish, *Phys. Rev. A* **94**, 041403 (2016).
12. M. Guo *et al.*, *Phys. Rev. A* **96**, 052505 (2017).
13. T. M. Rvachov *et al.*, *Phys. Rev. Lett.* **119**, 143001 (2017).
14. L. R. Liu *et al.*, *Science* **360**, 900–903 (2018).
15. W. B. Cairncross *et al.*, *Phys. Rev. Lett.* **126**, 123402 (2021).
16. L. De Marco *et al.*, *Science* **363**, 853–856 (2019).
17. L. Anderegg *et al.*, *Nat. Phys.* **14**, 890–893 (2018).
18. H. J. Williams *et al.*, *Phys. Rev. Lett.* **120**, 163201 (2018).
19. D. J. McCarron, M. H. Steinecker, Y. Zhu, D. DeMille, *Phys. Rev. Lett.* **121**, 013202 (2018).
20. S. Ospelkaus *et al.*, *Phys. Rev. Lett.* **104**, 030402 (2010).
21. J. W. Park, Z. Z. Yan, H. Loh, S. A. Will, M. W. Zwierlein, *Science* **357**, 372–375 (2017).
22. F. Seeßelberg *et al.*, *Phys. Rev. Lett.* **121**, 253401 (2018).
23. C. Chou *et al.*, *Nature* **545**, 203–207 (2017).
24. G. Valtolina *et al.*, *Nature* **588**, 239–243 (2020).
25. H. Son, J. J. Park, W. Ketterle, A. O. Jamison, *Nature* **580**, 197–200 (2020).
26. X. Ye, M. Guo, M. L. González-Martínez, G. Quéméner, D. Wang, *Sci. Adv.* **4**, eaaq0083 (2018).
27. P. D. Gregory *et al.*, *Nat. Commun.* **10**, 3104 (2019).
28. M.-G. Hu *et al.*, *Science* **366**, 1111–1115 (2019).
29. L. W. Cheuk *et al.*, *Phys. Rev. Lett.* **125**, 043401 (2020).
30. S. Bali, D. Hoffmann, T. Walker, *Europhys. Lett.* **27**, 273–277 (1994).
31. K.-A. Suominen, M. J. Holland, K. Burnett, P. Julienne, *Phys. Rev. A* **51**, 1446–1457 (1995).
32. P. S. Julienne, *J. Res. Natl. Inst. Stand. Technol.* **101**, 487–503 (1996).
33. R. Napolitano, J. Weiner, P. S. Julienne, *Phys. Rev. A* **55**, 1191–1207 (1997).
34. A. V. Gorshkov *et al.*, *Phys. Rev. Lett.* **101**, 073201 (2008).
35. G. Quéméner, J. L. Bohn, *Phys. Rev. A* **93**, 012704 (2016).
36. M. L. González-Martínez, J. L. Bohn, G. Quéméner, *Phys. Rev. A* **96**, 032718 (2017).
37. T. Karman, J. M. Hutson, *Phys. Rev. Lett.* **121**, 163401 (2018).
38. T. Karman, J. M. Hutson, *Phys. Rev. A* **100**, 052704 (2019).
39. T. Karman, *Phys. Rev. A* **101**, 042702 (2020).
40. K. Matsuda *et al.*, *Science* **370**, 1324–1327 (2020).
41. Z. Z. Yan *et al.*, *Phys. Rev. Lett.* **125**, 063401 (2020).
42. L. W. Cheuk *et al.*, *Phys. Rev. Lett.* **121**, 083201 (2018).
43. L. Anderegg *et al.*, *Science* **365**, 1156–1158 (2019).
44. See supplementary materials.
45. U. Kraft, *IEEE Trans. Antenn. Propag.* **44**, 515–522 (1996).
46. I. Kozyryev, L. Baum, K. Matsuda, J. M. Doyle, *Chem. Phys. Chem.* **17**, 3641–3648 (2016).
47. D. Mitra *et al.*, *Science* **369**, 1366–1369 (2020).
48. L. Anderegg, S. Burchesky, Y. Bao, S. S. Yu, T. Karman, E. Chae, K.-K. Ni, W. Ketterle, J. M. Doyle, Data for observation of microwave shielding of ultracold molecules. Zenodo (2021); https://doi.org/10.5281/zenodo.4724359.


## ACKNOWLEDGMENTS


**Funding:** This work was supported by the US Army Research Office, US Department of Energy, and National Science Foundation (NSF). L.A. acknowledges support from the Harvard Quantum Initiative. S.B. and S.S.Y. acknowledge support from the NSF Graduate Research Fellowships Program. E.C. is supported by National Research Foundation of Korea (2021R1C1C1009450 and 2020R1A4A1018015). **Author contributions:** L.A., S.B., Y.B., S.S.Y., E.C., K.-K.N., W.K., and J.M.D. contributed to the experimental effort. T.K. performed the theoretical calculations. All authors discussed the results and contributed to the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in this paper are present in the paper or in the supplementary materials. All data presented in this paper are deposited at Zenodo (*48*).


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/779/suppl/DC1
Supplementary Text
Figs. S1 to S7
References (*49*–*52*)


9 February 2021; accepted 7 July 2021
10.1126/science.abg9502

POLYMER CHEMISTRY

# Chemically recyclable thermoplastics from reversible-deactivation polymerization of cyclic acetals

Brooks A. Abel[1]†, Rachel L. Snyder[1]†, Geoffrey W. Coates[1]*

Identifying plastics capable of chemical recycling to monomer (CRM) is the foremost challenge in creating a sustainable circular plastic economy. Polyacetals are promising candidates for CRM but lack useful tensile strengths owing to the low molecular weights produced using current uncontrolled cationic ring-opening polymerization (CROP) methods. Here, we present reversible-deactivation CROP of cyclic acetals using a commercial halomethyl ether initiator and an indium(III) bromide catalyst. Using this method, we synthesize poly(1,3-dioxolane) (PDXL), which demonstrates tensile strength comparable to some commodity polyolefins. Depolymerization of PDXL using strong acid catalysts returns monomer in near-quantitative yield and even proceeds from a commodity plastic waste mixture. Our efficient polymerization method affords a tough thermoplastic that can undergo selective depolymerization to monomer.

Plastics are among the most high-performance and cost-effective materials available today. Consumer industries depend on the versatile properties of plastics. As a result, plastic production increases by ~8% annually, with single-use packaging materials making up nearly 40% of plastic products (*1*, *2*). However, the mass manufacture and uncontrolled disposal of plastics has come at both economic and environmental costs (*1*, *3*–*7*). Currently, most collected postconsumer plastics are processed for reuse via downcycling approaches such as mechanical recycling, affording relatively small quantities of low-value materials with diminished properties (*2*, *8*, *9*). As the plastics crisis grows, global organizations and governments (*10*, *11*) are targeting more promising strategies to simultaneously combat the environmental and economic impacts of the plastics problem. These methods include upcycling, where plastic waste is used as a feedstock for value-added materials (*12*–*16*) and chemical recycling to monomer (CRM). CRM converts plastic waste directly back to monomer, enabling a circular plastics economy that both mitigates the need for continuous feedstock sourcing and could potentially eliminate the accumulation of plastic waste. To reduce our dependence on fossil fuels, prevent the accumulation of millions of tons of plastic waste each year, and turn massive economic losses into gains, the development of a circular plastics economy via CRM must be at the forefront of sustainability efforts (*17*, *18*).

Polymers derived from moderately strained heterocyclic monomers are viable candidates for CRM because of their moderate ceiling temperatures ($T_c$ < 250°C), the temperature at which the change in Gibbs free energy ($\Delta G$) = 0 for polymerizations where both the change in enthalpy ($\Delta H$) and the change in entropy ($\Delta S$) are negative (*19*). To date, polyesters (*20*–*25*), polycarbonates (*26*–*29*), and polymers derived from other heterocyclic monomers (*30*–*35*) are capable of CRM. Several of these systems exhibit noteworthy properties. Chen and co-workers reported on the synthesis of poly[*trans*-hexahydro-2(3*H*)-benzofuranone], which displayed high tensile stress at break ($\sigma_B$ = 55 MPa), high melting temperature ($T_m$) values (≥126°C), and good thermal stability [decomposition temperature ($T_d$) = 340°C] (*23*). Helms and co-workers synthesized cross-linked polymers comprising diketoenamine linkages, which were depolymerized in the presence of mixed plastic waste (*33*). Mecking and co-workers developed long-chain aliphatic polyesters with polyolefin-like mechanical properties that were synthesized on a large scale from biorenewable monomers and chemically recycled via solvolysis (*36*). Moving forward, important advancements to these foundational reports (and others) will include implementing easily accessible monomers, improving the efficiency of polymer syntheses, accessing useful material properties, and invoking simple processes for depolymerization to monomer (*9*, *12*, *37*). Designing systems that meet these criteria will enable the widespread use of chemically recyclable polymers.

Polyacetals are promising candidates for CRM because their dynamic acetal functionalities facilitate depolymerization at relatively low temperatures (<150°C) (*34*). They also possess suitable thermal and chemical stability for real-world application, as evidenced by the commercialization of polyoxymethylene (POM) as an engineering thermoplastic (*38*). Polyacetals are commonly synthesized from cyclic acetal monomers, which are readily made on a large scale from diols and formaldehyde (*39*).

1,3-Dioxolane (DXL) is commercially synthesized in a single step on a large scale from common industrial C1 and C2 feedstocks—formaldehyde and ethylene glycol, respectively—both of which also have bioderived routes (*34*, *40*). Early work on poly(1,3-dioxolane) (PDXL) suggested its potential applications (Fig. 1A) (*41*). However, consistently obtaining high molecular weight polyacetals has remained challenging because cyclic acetal polymerization using Brønsted or Lewis acid catalysts affords poor molecular weight control (Fig. 1B) (*42*). Therefore, we sought to achieve molecular weight control during the cationic ring-opening polymerization (CROP) of DXL.

In typical living and/or controlled cationic polymerizations, dormant halide-terminated chain ends generated from discrete initiators are reversibly activated by a Lewis acid catalyst. This reversible-deactivation mechanism maintains low cation concentrations, a requirement for minimizing undesired termination and irreversible chain transfer. Living cationic polymerizations were initially developed in the 1970s by Higashimura, Sawamoto, and colleagues to polymerize vinyl ether and *N*-vinylcarbazole monomers (*43*, *44*) and later were adapted to isobutylene systems by Faust and Kennedy (*45*). More recently, Aoshima and co-workers used a reversible-deactivation strategy to copolymerize vinyl ethers and cyclic acetals (*46*). In this work, we identify an initiator and Lewis acid catalyst system capable of promoting reversible deactivation of halide-terminated polyacetals to achieve both molecular weight control and high living chain-end retention (Fig. 1C) (*47*).

In reversible-deactivation CROP (RD-CROP), the active 3° oxonium and oxocarbenium species present during CROP of cyclic acetals are in equilibrium with dormant halomethyl ether species (Fig. 2A). Therefore, we selected chloromethyl methyl ether (MOMCl) as the initiator for RD-CROP to mirror the dormant halide-terminated polyacetal chain ends. To prevent undesired initiation or chain transfer by protic impurities, the sterically hindered base 2,6-di-*tert*-butyl pyridine (DTBP) was used as a proton trap (*48*). We screened a range of commercial Lewis acid catalysts of the form $MCl_n$ [M = B(III), Al(III), Ga(III), In(III), Sb(III)/Sb(V), Bi(III), Sn(IV), Zn(II), Fe(III), Ti(IV), Zr(IV), or Hf(IV)] owing to their prior use in cationic polymerizations and Friedel-Crafts reactions. In the presence of a Lewis acid and halomethyl ether, cyclic acetal polymerization can proceed via three pathways: (i) initiation by Brønsted acid impurities, (ii) monomer activation with the Lewis acid catalyst, or (iii) ionization of halomethyl ether chain ends by the Lewis acid catalyst (i.e., RD-CROP). We determined the selectivity

[1]Department of Chemistry and Chemical Biology and Joint Center for Energy Storage Research, Baker Laboratory, Cornell University, Ithaca, NY 14853, USA.
*Corresponding author. Email: coates@cornell.edu
†These authors contributed equally to this work.

of each catalyst toward RD-CROP of DXL using $MCl_n$ only (method A), $MCl_n$ + DTBP (proton trap; method B), or $MCl_n$ + DTBP + MOMCl initiator (method C). Catalysts were deemed viable for RD-CROP if polymerization did not occur using method B (i.e., both $H^+$ activation and Lewis acid monomer activation were absent) but proceeded when initiator was added to the system (method C; fig. S1). No polymerization was observed within 18 hours when the Lewis acids $BCl_3$, $AlCl_3$, $SbCl_3$, $BiCl_3$, $TiCl_4$, $ZrCl_4$, or $HfCl_4$ were used in the presence or absence of MOMCl (table S2). Polymerization occurred in both the presence and absence of MOMCl when the Lewis acids $GaCl_3$, $SbCl_5$, $SnCl_4$, or $FeCl_3$ were used (table S2), indicating that these Lewis acids directly activate monomer. $InCl_3$ and $ZnCl_2$ catalyzed the polymerization of DXL exclusively in the presence of MOMCl. We hypothesize that the increased halophilicity of soft Lewis acids such as $InCl_3$ and $ZnCl_2$ improves selectivity toward chain-end activation relative to binding oxygens in the monomer or polymer (*49*). $ZnCl_2$ demonstrated incomplete conversion, even at higher catalyst loadings over 18 hours. Meanwhile, $InCl_3$ reached maximum conversion (~85%, based on measured standard state $\Delta H°$ and $\Delta S°$ values when $[DXL]_0$ = 8 M; see below) in under 18 hours at lower catalyst loading and was selected for further optimization (table S2). Despite lower activity than In-based catalysts, Zn-based catalysts could be a more cost-effective alternative (*23*).

To increase the polymerization rate, we exchanged Cl with more-labile Br leaving groups to increase the extent of chain-end ionization (i.e., concentration of active chain ends). The $InX_3$/MOMX halide composition was varied from a [Br]:[Cl] ratio of 0:4 to 4:0 by combining the appropriate amounts of MOMCl, MOMBr, $InCl_3$, and $InBr_3$ to achieve a particular halogen stoichiometry. At [Br]:[Cl] = 0:4, polymerization reached 81% conversion in 40 min. When [Br]:[Cl] was increased to 4:0, 81% conversion was reached in just 90 s, demonstrating a more than 25-fold increase in polymerization rate (Fig. 2B and table S3). In all cases, we observed a linear dependence of number-average molecular weight as measured by gel permeation chromatography ($M_{n,GPC}$) on DXL conversion, suggesting living chain-end retention throughout the polymerization (table S3 and fig. S2).

RD-CROP with $InBr_3$/MOMBr showed excellent molecular weight control for DXL and other monocyclic acetal derivatives including 1,3-dioxepane (DXP), 1,3-dioxocane (DXC), 1,3,7-trioxocane (TXC), and the bicyclic acetal *trans*-hexahydro-1,3-benzodioxole (HBD) (Fig. 2C; figs. S3 to S7; and table S5). For each monomer derivative, $M_{n,GPC}$ increases linearly with increasing monomer-to-initiator ratio

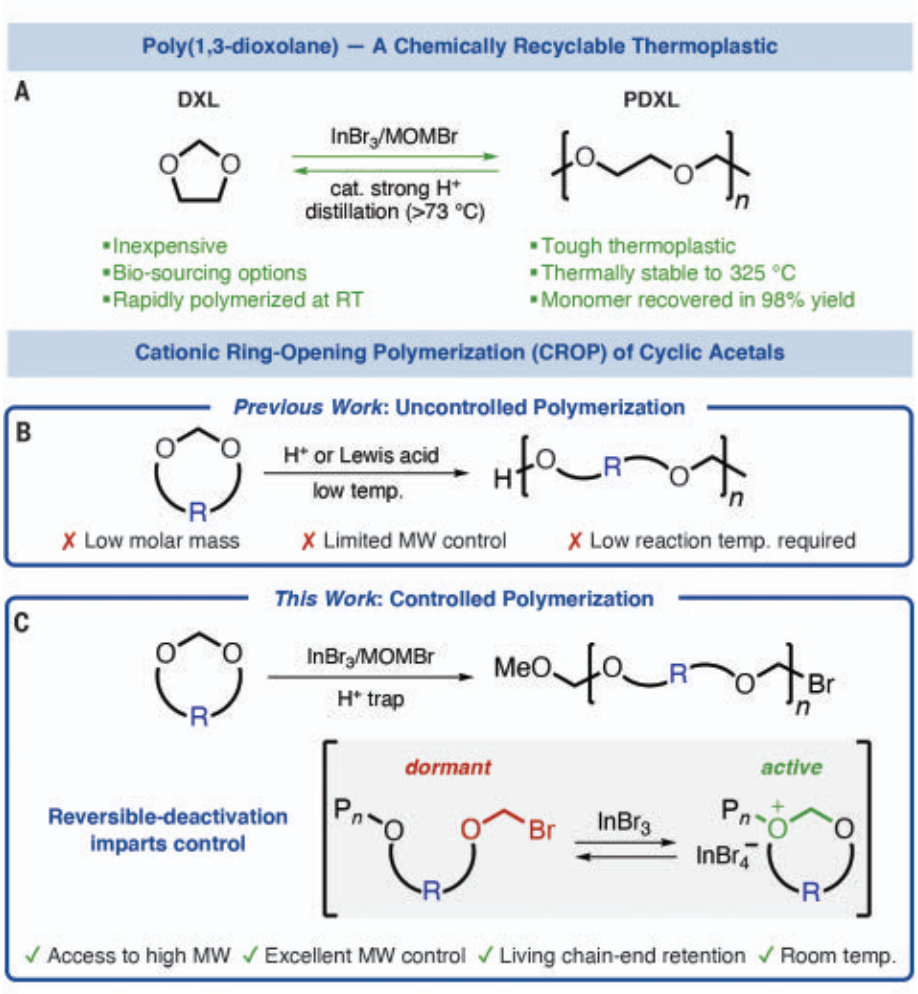



**Fig. 1. Methods of polyacetal synthesis.** (**A**) Polymerization of DXL is reversible through heating in acid. RT, room temperature. (**B**) Uncontrolled CROP of cyclic acetals using Brønsted or Lewis acid initiators affords poor molecular weight (MW) control. (**C**) RD-CROP imparts control over cyclic acetal polymerization. R, organic linker; MeO, methoxy group.

($[monomer]_0$:$[MOMBr]_0$), indicating that initiation by MOMBr is predominant. Sequential monomer additions over 24 hours during RD-CROP of DXL further confirmed living chain-end retention at room temperature (Fig. 2D). $M_{n,GPC}$ values increased in proportion to the amount of DXL added, and no change in dispersity ($Đ$) was observed during the course of the experiment (table S6 and fig. S9).

Despite excellent molecular weight control, we consistently observed $Đ$ values ranging from 1.51 to 1.74 owing to transacetalization, a known phenomenon in CROP of cyclic acetals (*50*). During transacetalization, acetal linkages in the polymer backbone react with cationic polymer chain ends because of the similar nucleophilicities of cyclic and acyclic acetals (fig. S10). We confirmed transacetalization during RD-CROP of cyclic acetals by introducing increasing amounts of an acyclic acetal, diethoxymethane (DEM), relative to [MOMBr] at a constant monomer-to-initiator ratio. Gel permeation chromatography (GPC) analyses show good agreement between theoretical molecular weights ($M_{n,th}$) calculated using equation S2 (see supplementary materials) and experimental $M_{n,GPC}$ values (table S7 and fig. S11). For all studied polyacetals, end-group analysis of low molecular weight samples was performed using matrix-assisted laser desorption/ionization–time-of-flight (MALDI-TOF) mass spectrometry. As anticipated, degenerative chain transfer to polymer via transacetalization gives a statistical 1:2:1 distribution of chain ends from the MOMBr initiator and sodium benzyloxide (NaOBn) quenching agent (figs. S10 to S14).

In previous reports, residual catalyst or acid species from incomplete polymer purification prematurely catalyzed depolymerization of polyacetals owing to the thermodynamics of polymerization and depolymerization (Fig. 2E;

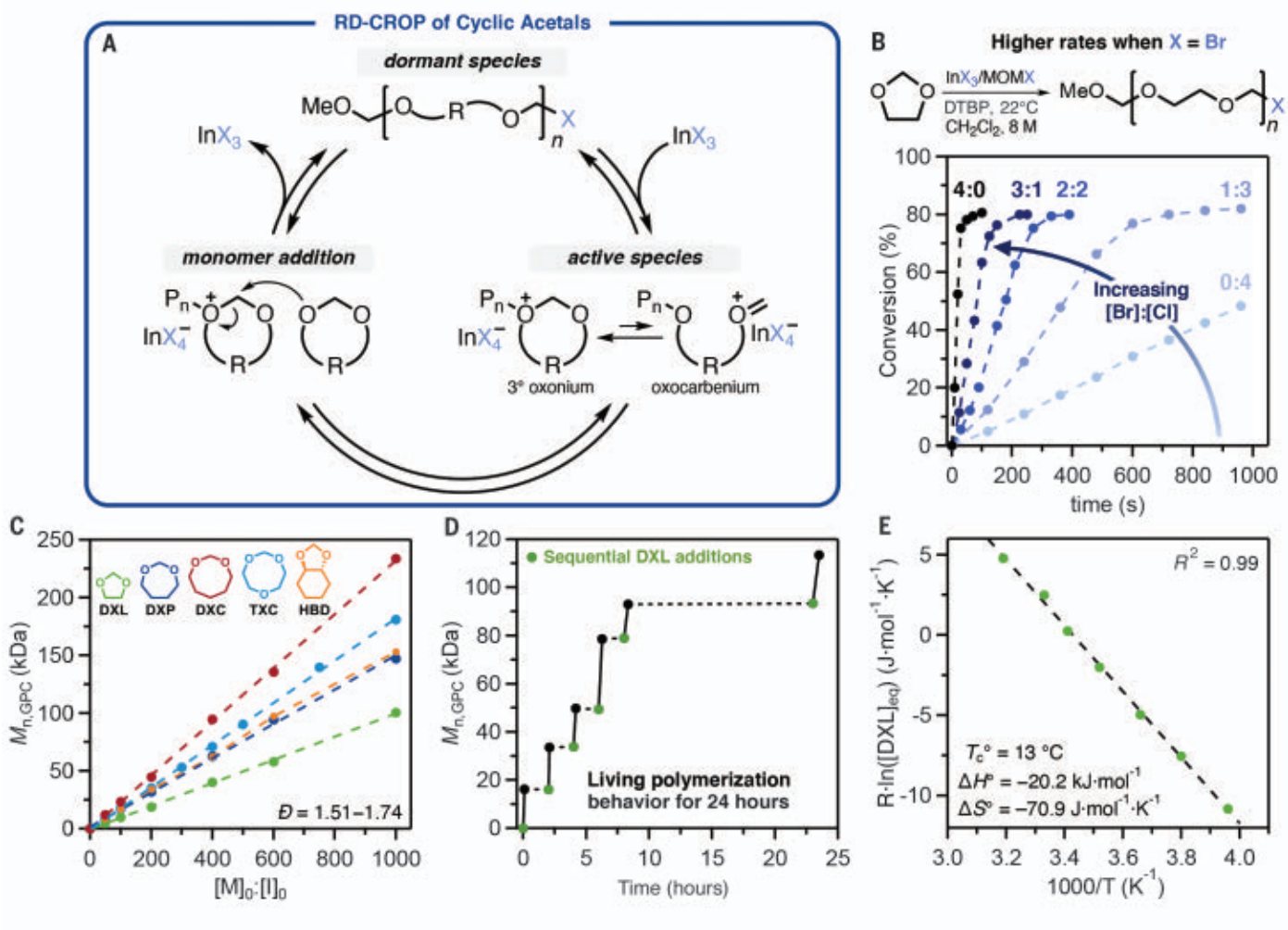



**Fig. 2. Reversible-deactivation cationic ring-opening polymerization of cyclic acetals.** (**A**) The proposed mechanism of cyclic acetal RD-CROP involves reversible chain-end activation by the $InBr_3$ catalyst after MOMBr initiation. (**B**) Reaction rate increases with $InX_3$/MOMX [Br]:[Cl] ratio in the RD-CROP system. (**C**) Molecular weight control is observed with linear dependence of $M_{n,GPC}$ on [monomer]$_0$:[MOMBr]$_0$ for monomer scope. (**D**) High living chain-end retention is observed during RD-CROP of DXL. (**E**) Van't Hoff analysis of variable temperature NMR data provides $\Delta H°$, $\Delta S°$, and $T_c°$ values for PDXL.

see below) (*51*). With an improved purification method (see supplementary materials), all five polyacetal derivatives exhibited excellent thermal stability by thermogravimetric analysis (TGA), with degradation temperatures at 5% mass loss ($T_{d,5\%}$) above 337°C (fig. S17 and table S8). Isothermal TGA data shows >99% mass remaining after 2 hours at 140°C for each polyacetal derivative (fig. S18 and table S9).

Differential scanning calorimetry (DSC) was used to measure the thermal transition temperatures for each polyacetal derivative, some of which are semicrystalline (fig. S19 and table S10). Semicrystalline PDXL exhibits a low glass transition temperature ($T_g$) of –63°C and a $T_m$ of 58°C. The $T_m$ of PDXL is comparable to thermal transitions in several commercial materials, including poly(lactic acid) ($T_g$ = 60°C), poly(ethylene terephthalate) ($T_g$ = 61°C), poly(ε-caprolactone) ($T_m$ = 60°C), and poly(ethylene oxide) ($T_m$ = 66°C) (*52*). Both thermal transitions in PDXL are invariant with molecular weight from 37.9 to 220 kDa (fig. S20 and table S11).

In general, tensile strength increases with polymer molecular weight. To date, the tensile properties of PDXL have been overlooked because prior catalyst systems typically yielded low molecular weight materials (see above) (*42*). To measure the effect of molecular weight on tensile properties, we synthesized PDXL on a 10-g scale with $M_{n,GPC}$ values ranging from 37.9 to 220 kDa (fig. S21 and table S12). Bulk PDXL was melt pressed to give colorless, opaque films (Fig. 3A). Uniaxial tensile elongation tests were performed on the PDXL samples of increasing molecular weight to identify the critical molecular weight, above which robust properties are obtained (figs. S20 to S22 and tables S13 to S19). Low molecular weight PDXL (37.9 kDa), which is comparable to the highest previously reported molecular weight (*42*), demonstrates poor mechanical properties, with a $\sigma_B$ of only 13.5 ± 1.3 MPa at 5 ± 0.3% strain ($\varepsilon_B$) (table S14 and fig. S24). At 82.3 kDa, PDXL tensile properties increase significantly to $\sigma_B$ = 33.3 ± 1.2 MPa and $\varepsilon_B$ = 640 ± 45% (table S16 and fig. S24). At 180 kDa, PDXL showed high tensile strength, with $\sigma_B$ = 40.4 ± 1.2 MPa and $\varepsilon_B$ = 720 ± 20%, revealing the impressive toughness and ductility of PDXL at high molecular weights (Fig. 3A and table S19). In fact, the tensile strength of PDXL is comparable to the two most prevalent commodity plastic materials, isotactic polypropylene ($\sigma_B$ = 26 MPa and $\varepsilon_B$ = 420%) and high-density polyethylene ($\sigma_B$ = 30.2 MPa and $\varepsilon_B$ = 900%). Like other semicrystalline polymers, PDXL undergoes considerable strain-induced crystallization, forming visible white striations that give way to fibrous filaments after fracture (fig. S25).

The stability of the films at elevated temperature and/or humidity was studied by

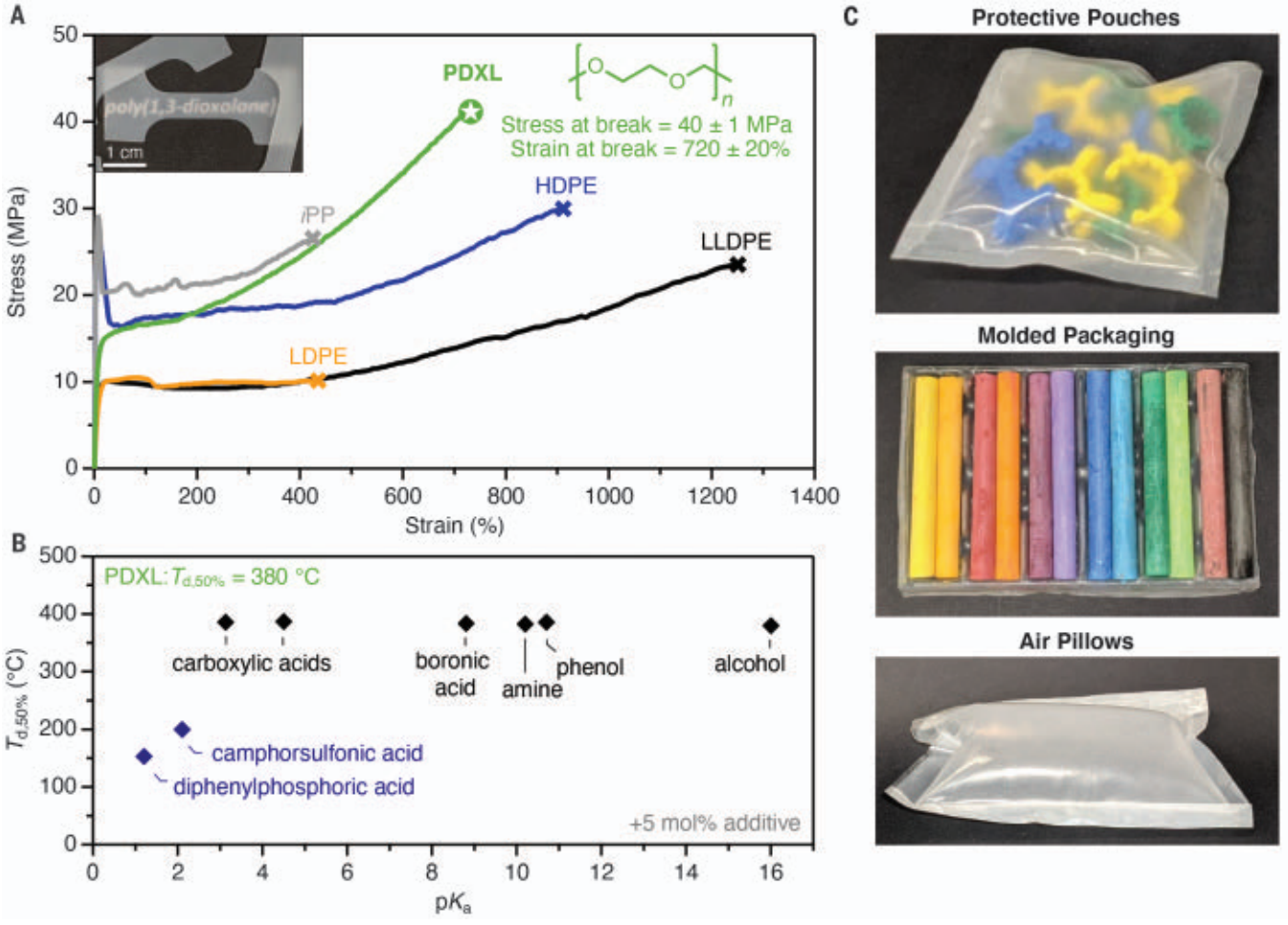



**Fig. 3. Tensile properties and thermal stability of poly(1,3-dioxolane).** (**A**) PDXL exhibits high tensile strength comparable to isotactic polypropylene (*i*PP) and high-density polyethylene (HDPE). LDPE, low-density polyethylene; LLDPE, linear low-density polyethylene. Inset: Image of colorless, semicrystalline PDXL films synthesized on multigram scales. (**B**) PDXL shows good thermal stability ($T_{d,50\%}$ ≥ 380°C) in the presence of weak acids and amines but degrades readily with strong acid additives (p$K_a$ ≤ 2.1). (**C**) PDXL prototype items were formed by melt processing of PDXL films.

placing a PDXL sample (110 kDa, 100 mg) in a 20-ml vial under either a dry atmosphere or at 100% relative humidity. After 7 days at 57°C, the polymer samples exposed to dry or humid environments showed identical proton nuclear magnetic resonance ($^1$H NMR) spectra to those of the pristine material and no decrease in $M_n$. These experiments highlight the stability of PDXL at elevated temperatures for prolonged periods of time, even at 100% relative humidity (fig. S26). Additionally, no transacetalization was observed after polymer isolation, confirming full removal of active catalytic species from the polymer samples during purification (fig. S27). High molecular weight PDXL (180 kDa) is readily soluble in $CH_2Cl_2$ but demonstrates good solvent resistance and minimal swelling toward hexanes, acetone, methyl ethyl ketone (MEK), and ethanol (EtOH) with <3% dissolution by mass after 100 days submersed in these solvents under ambient conditions (fig. S28 and table S20). PDXL showed some solubility in tetrahydrofuran, with 90 wt % remaining after 100 days. Submersion in water for extended periods of time results in partial dissolution, reaching 17% mass loss in 7 days and 64% mass loss over 100 days. Both high (182 kDa) and low (37.9 kDa) molecular weight PDXL showed identical GPC traces after 100 days in $H_2O$, indicating that PDXL is resistant to hydrolysis under neutral pH conditions over time. Although the hydrophilicity of PDXL precludes its use in some applications, many plastic products, such as air pillows, protective pouches, or molded packaging, require robust tensile strengths but remain dry during use (Fig. 3C). Detailed $^1$H NMR analysis of PDXL in $CDCl_3$ or $D_2O$ at 50°C for 48 hours showed no change in the NMR spectra, such as the appearance of any degradation by-products including formaldehyde, epoxides, alcohols, or monomer (fig. S29). Therefore, although PDXL dissolves in water, it does not hydrolytically degrade, even above the ceiling temperature, in the absence of an acid catalyst.

As with any newly synthesized polymer, PDXL will need to undergo complete stability, safety, and toxicity testing before it sees use. While we currently have no evidence of premature depolymerization or the production of by-products other than DXL monomer, it will be critical to analyze the polymer degradation across a broader range of environmental conditions to ensure both its safety and utility during application. Detailed studies to elucidate the by-products and mechanism(s) of PDXL biodegradation in both soil and marine environments are currently in progress.

Having demonstrated that PDXL has good stability and promising tensile strength, we next demonstrated that PDXL could be efficiently chemically recycled to monomer (Fig. 4A). Ceiling temperature is an important indicator of CRM-capable polymers and informs the conditions required for depolymerization. We calculated standard state $T_c^\circ$ for PDXL using an established variable temperature NMR method in which the equilibrium DXL concentration

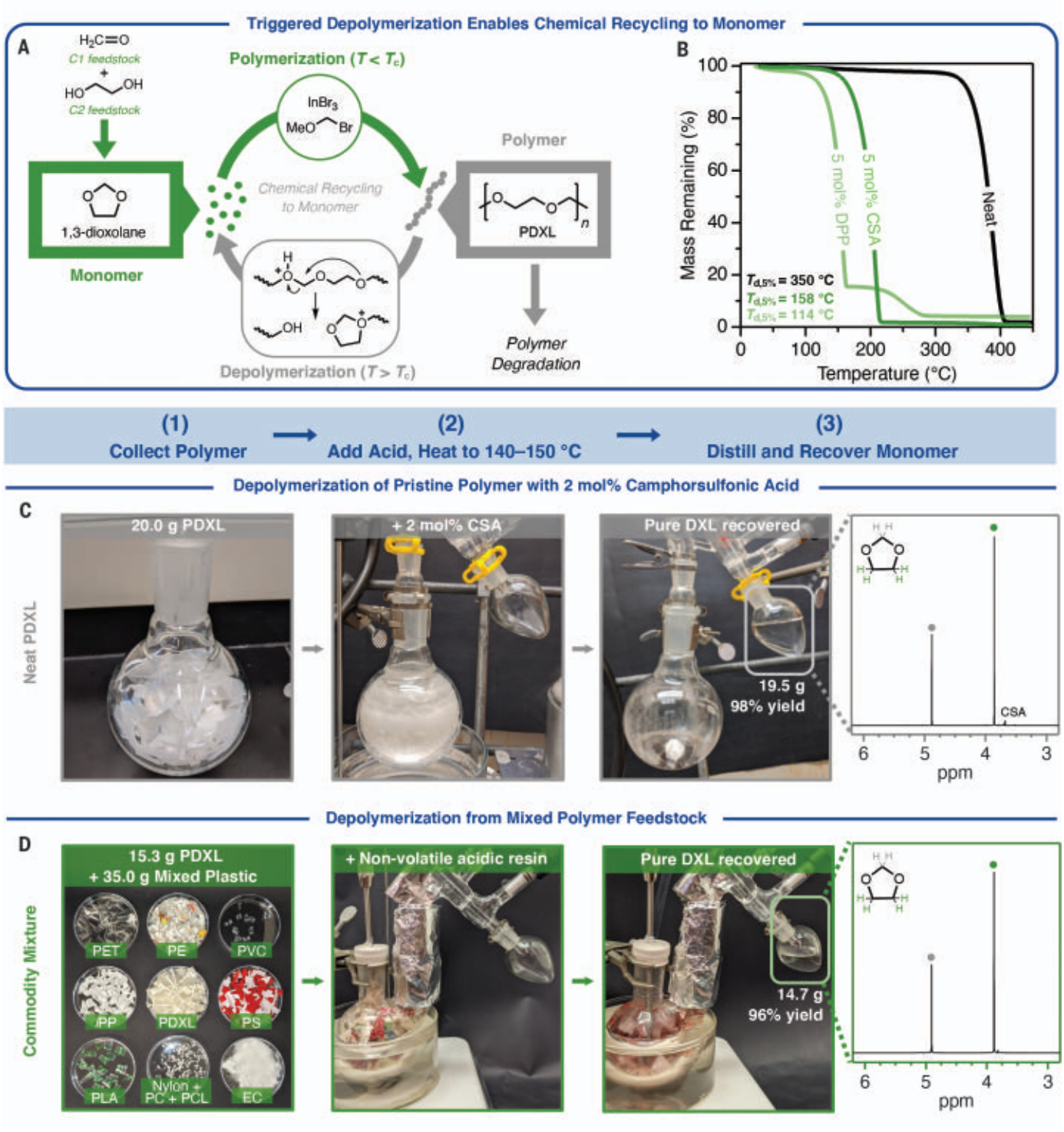



**Fig. 4. Chemical recycling to monomer of poly(1,3-dioxolane).** (**A**) DXL can be polymerized to PDXL and chemically recycled to monomer. (**B**) PDXL is thermally stable to $T_{d,5\%}$ = 353°C but readily depolymerized with 5 mol % added CSA or DPP. (**C**) PDXL was depolymerized via distillation in ambient atmosphere at 150° ± 5°C over 60 min to recover DXL monomer in near-quantitative yields. $^{1}$H NMR of the recovered monomer shows only DXL and trace amounts of CSA (ppm, parts per million), which is readily removed during monomer drying with $CaH_2$. (**D**) Using a polymer-supported acid catalyst, CRM of PDXL from a mixed commodity plastics feedstock yields exclusively DXL monomer without contamination by dyes, plasticizers, or other additives. PET, polyethylene terephthalate; PE, polyethylene; PVC, polyvinyl chloride; PS, polystyrene; PLA, polylactic acid; PC, polycarbonate; PCL, poly(ε-caprolactone); EC, ethyl cellulose.

($[DXL]_{eq}$) was measured as a function of temperature (fig. S31 and table S21) (*30*). The resultant van't Hoff plot and equations S5 to S7 (see supplementary materials) were used to determine the standard state $\Delta H°$, $\Delta S°$, and $T_c°$ values for RD-CROP of DXL (Fig. 2E and fig. S32). PDXL demonstrates $\Delta H° = -20$ kJ·mol$^{-1}$, $\Delta S° = -70.9$ J·mol$^{-1}$·K$^{-1}$, and a relatively low $T_c°$ of 13°C, which is in good agreement with previously reported values (table S22) (*53*).

Given its low $T_c$, it is essential that PDXL undergoes depolymerization only in the presence of an external catalyst to avoid unintended degradation during use. Methylene acetals are known to remain stable toward weak acids and bases (*54*). Similarly, PDXL showed excellent thermal stability in the presence of acid or base additives across a range of p$K_a$ values (where $K_a$ is the acid dissociation constant) (Fig. 3B). Additives with relatively low volatility were chosen to provide the largest temperature window possible before boiling or sublimation of the additive occurred. Polymer thermal stability was quantified using the degradation temperature at 50% mass remaining ($T_{d,50\%}$) to account for mass loss from the additive (8 to 15 wt %) (table S23). PDXL retains $T_{d,50\%}$ values near 380°C with 5 mol % loading of weaker acids, including citric acid (p$K_{a,1}$ = 2.9), *trans*-cinnamic acid (p$K_a$ = 4.5), or 4-tolylboronic acid (p$K_a$ = 8.8) (fig. S33 and table S23). Likewise, 5 mol % loadings of representative alcohols 4-*tert*-butyl phenol and 1-pentadecanol and nitrogen-containing bases octylamine, tributylamine, and 7-methyl-1,5,7-triazabicyclo[4.4.0]decene (MeTBD) did not affect PDXL degradation (figs. S34 and S35 and table S23). As expected, strong acids such as diphenylphosphoric acid (DPP; p$K_a$ = 1.2) and camphorsulfonic acid (CSA; p$K_a$ = 2.1) efficiently catalyzed PDXL degradation, affording accessible $T_{d,50\%}$ values of 153° and 200°C, respectively (Fig. 4B and table S23). Lewis acidic lithium bis(trifluoromethane)sulfonimide and transition metal titanium(IV) isopropoxide also catalyzed PDXL degradation ($T_{d,50\%}$ values of 323° and 313°C, respectively), although much less efficiently than did CSA and DPP (fig. S36 and table S23).

CSA was selected as a representative depolymerization catalyst as it is inexpensive and derived from biosourced material. The rate of acid-catalyzed degradation decreased slightly with increasing $M_{n,GPC}$. At 2 mol % CSA loading, the time to 25% mass loss at 120°C increased from 6.7 to 9.0 to 10.5 min for 27.9, 110, and 220 kDa PDXL in the melt, respectively (fig. S37 and table S24). As expected, increasing degradation rates were also observed at increasing temperatures (100° to 140°C) for a 127 kDa PDXL sample with 2 mol % added CSA (fig. S37). The rate of degradation (percent mass loss per second) was calculated from the initial slope of the isothermal TGA experiments (90 to 70% mass remaining) and used to determine an Arrhenius activation energy of 85.0 kJ/mol [coefficient of determination ($R^2$) = 0.998] for CSA-catalyzed PDXL depolymerization (fig. S38). Upon increasing the CSA loading from 5 to 10 to 15 mol %, the $T_{d,5\%}$ of PDXL (182 kDa) decreased from 157° to 108° to 80°C, respectively. At 20 or 25 mol % added CSA, PDXL $T_{d,5\%}$ values plateaued at 73° and 74°C, respectively, coinciding with the boiling point of DXL monomer (fig. S39 and table S25).

Once depolymerization is kinetically activated by the catalyst, the combined low $T_c°$ of PDXL and low volatility of DXL permit facile monomer collection via simple distillation at moderate temperatures under ambient pressure, an accessible and inexpensive procedure that could mitigate the need to transport substantial amounts of waste to specialized recycling facilities (*16*). We first doped PDXL (20.0 g; 60 to 220 kDa) with 2 mol % CSA. The catalyst was solvent-cast into the material in $CH_2Cl_2$ to ensure a homogeneous dispersion of CSA and study the efficacy of PDXL CRM (Fig. 4C). A short-path distillation was then set up under ambient atmosphere, and the PDXL/CSA mixture was lowered into an oil bath preheated to 140°C. After heating for 14 min, pure DXL began collecting in the receiving flask. The distillation continued for 78 min in total, at which point only black residue from degraded CSA remained, and 19.5 g (98% yield) of pure DXL monomer was collected with trace amounts of CSA impurities, as confirmed by $^1$H NMR spectroscopy (fig. S40 and table S26). The recovered DXL was later dried over $CaH_2$ and repolymerized using $InBr_3$/MOMBr-catalyzed RD-CROP in $CH_2Cl_2$ at room temperature to give PDXL ($M_{n,GPC}$ = 84.4 kDa) with identical tensile properties to pristine 80.7 kDa material (fig. S41 and table S27).

Plastics collected from mixed waste streams generally include a variety of dyes and/or pigments, plasticizers, stabilizers, and other impurities (*9*). A key advantage of CRM is the ability to isolate pure monomer from such complex mixtures, avoiding the need for time- and cost-intensive separation processes. To investigate the efficacy of selective PDXL depolymerization and DXL collection from a mixture of plastic waste, we performed CRM of PDXL in the presence of poly(ethylene terephthalate), high-density and low-density polyethylene, poly(vinyl chloride), isotactic polypropylene, polystyrene, poly(lactic acid), poly(ε-caprolactone), polycarbonate, nylon-6,6, ethyl cellulose, and the plasticizer bis(2-ethylhexyl) phthalate. The plastics used were either postconsumer materials containing dyes, additives, and plasticizers or were purchased from chemical suppliers (table S28). PDXL (15.3 g; 80 to 180 kDa) was then added to the mixture, and an acidic resin (Dowex-50, 5 wt %) was selected as a nonvolatile catalyst (Fig. 4D). The heterogeneous mixture was then mechanically stirred at 150° ± 5°C under ambient atmosphere. A primary fraction of DXL monomer (13.8 g) was collected during the first 30 min. A second fraction (0.931 g) was collected over an additional 30 min, giving a total DXL yield of 96%. Most notably, the two collected fractions contained >99% DXL monomer in the presence of commodity plastics containing acid- and/or heat-sensitive linkages (i.e., esters, amides, carbonates, and glycosidic bonds) and small-molecule dyes and plasticizers (figs. S42 and S43). This result demonstrates that polymer separation is not required prior to CRM of PDXL from mixed waste streams.

We have developed an RD-CROP method that affords molecular weight control and living chain-end retention for the polymerization of cyclic acetal monomers. RD-CROP enables the synthesis of high molecular weight PDXL—a thermally stable, semicrystalline thermoplastic with high tensile strength suitable for large-scale applications such as packaging products. Moving forward, we are focusing on several major goals to improve upon the properties of PDXL and other CRM-enabled polyacetals. To begin, we are investigating new monomer designs to afford polyacetals that maintain moderate ceiling temperatures but show improved hydrophobicity and higher melting points that better mimic the properties of current commodity polymers. Next, we are optimizing the polymerization conditions to further meet the criteria of green chemistry, including bulk polymerization, using new initiating and quenching agents, and modifying the reaction workup. Lastly, we will explore stabilizers to afford improved oxidative and chemical stability of PDXL. Overall, we believe polyacetal synthesis via RD-CROP of cyclic acetals will prove an important strategy in the development of the circular plastics economy.

## REFERENCES AND NOTES


1. R. Geyer, J. R. Jambeck, K. L. Law, *Sci. Adv.* **3**, e1700782 (2017).
2. Z. O. G. Schyns, M. P. Shaver, *Macromol. Rapid Commun.* **42**, 2000415 (2021).
3. J. M. Garcia, M. L. Robertson, *Science* **358**, 870–872 (2017).
4. J. Brahney, M. Hallerud, E. Heim, M. Hahnenberger, S. Sukumaran, *Science* **368**, 1257–1260 (2020).
5. E. MacArthur, *Science* **358**, 843–843 (2017).
6. S. B. Borrelle *et al*., *Science* **369**, 1515–1518 (2020).
7. K. L. Law *et al*., *Sci. Adv.* **6**, eabd0288 (2020).
8. D. J. Fortman *et al*., *ACS Sustain. Chem. Eng.* **6**, 11145–11159 (2018).
9. S. Billiet, S. R. Trenor, *ACS Macro Lett.* **9**, 1376–1390 (2020).
10. "Chemical Upcycling of Polymers," Report of the Basic Energy Sciences Roundtable on Chemical Upcycling of Polymers, Bethesda, Maryland, 30 April to 1 May 2019; https://science.osti.gov/-/media/bes/pdf/reports/2020/Chemical_Upcycling_Polymers.pdf.
11. Ellen MacArthur Foundation, "Universal circular economy policy goals" (2021); https://policy.ellenmacarthurfoundation.org/universal-policy-goals.
12. J. C. Worch, A. P. Dove, *ACS Macro Lett.* **9**, 1494–1506 (2020).

13. A. Rahimi, J. M. García, *Nat. Rev. Chem.* **1**, 0046 (2017).
14. A. Kumar *et al.*, *J. Am. Chem. Soc.* **142**, 14267–14275 (2020).
15. F. Zhang *et al.*, *Science* **370**, 437–441 (2020).
16. C. Jehanno, M. M. Pérez-Madrigal, J. Demarteau, H. Sardon, A. P. Dove, *Polym. Chem.* **10**, 172–186 (2019).
17. D. E. Fagnani *et al.*, *ACS Macro Lett.* **10**, 41–53 (2020).
18. W. W. Y. Lau *et al.*, *Science* **369**, 1455–1461 (2020).
19. H. Höcker, *J. Macromol. Sci. Part A* **30**, 595–601 (1993).
20. H. Nishida *et al.*, *Macromolecules* **44**, 12–13 (2011).
21. G. W. Fahnhorst, T. R. Hoye, *ACS Macro Lett.* **7**, 143–147 (2018).
22. L. Cederholm, P. Olsén, M. Hakkarainen, K. Odelius, *Polym. Chem.* **11**, 4883–4894 (2020).
23. J.-B. Zhu, E. M. Watson, J. Tang, E. Y.-X. Chen, *Science* **360**, 398–403 (2018).
24. X. Tang, E. Y.-X. Chen, *Chem* **5**, 284–312 (2019).
25. R. M. Cywar, J.-B. Zhu, E. Y.-X. Chen, *Polym. Chem.* **10**, 3097–3106 (2019).
26. Y. Liu, H. Zhou, J.-Z. Guo, W.-M. Ren, X.-B. Lu, *Angew. Chem. Int. Ed.* **56**, 4862–4866 (2017).
27. W. C. Ellis *et al.*, *Chem. Sci.* **5**, 4004–4011 (2014).
28. D. J. Saxon, E. A. Gormong, V. M. Shah, T. M. Reineke, *ACS Macro Lett.* **10**, 98–103 (2021).
29. G. W. Coates, Y. D. Y. L. Getzler, *Nat. Rev. Mater.* **5**, 501–516 (2020).
30. J. P. Lutz *et al.*, *J. Am. Chem. Soc.* **141**, 14544–14548 (2019).
31. S. Enthaler, *J. Appl. Polym. Sci.* **131**, 39791 (2014).
32. Y. Liu, Y. Jia, Q. Wu, J. S. Moore, *J. Am. Chem. Soc.* **141**, 17075–17080 (2019).
33. P. R. Christensen, A. M. Scheuermann, K. E. Loeffler, B. A. Helms, *Nat. Chem.* **11**, 442–448 (2019).
34. A. Hufendiek, S. Lingier, F. E. Du Prez, *Polym. Chem.* **10**, 9–33 (2018).
35. C. Shi *et al.*, *Sci. Adv.* **6**, eabc0495 (2020).
36. M. Häußler, M. Eck, D. Rothauer, S. Mecking, *Nature* **590**, 423–427 (2021).
37. M. Hong, E. Y.-X. Chen, *Green Chem.* **19**, 3692–3706 (2017).
38. C. Suemitsu, A. Kanazawa, S. Aoshima, *Polym. Chem.* **12**, 822–830 (2021).
39. R. L. Snyder *et al.*, *ACS Energy Lett.* **6**, 1886–1891 (2021).
40. M. A. Hillmyer, *Science* **358**, 868–870 (2017).
41. W. F. Gresham, Preparation of Polydioxolane, US Patent 2394910 (1946).
42. H. Qiu, Z. Yang, M. Köhler, J. Ling, F. H. Schacher, *Macromolecules* **52**, 3359–3366 (2019).
43. T. Higashimura, O. Kishiro, *Polym. J.* **9**, 87–93 (1977).
44. M. Sawamoto, J. Fujimori, T. Higashimura, *Macromolecules* **20**, 916–920 (1987).
45. R. Faust, J. P. Kennedy, *J. Polym. Sci. A Polym. Chem.* **25**, 1847–1869 (1987).
46. K. Maruyama, A. Kanazawa, S. Aoshima, *Polym. Chem.* **10**, 5304–5314 (2019).
47. R. Szymanski, P. Kubisa, S. Penczek, *Macromolecules* **16**, 1000–1008 (1983).
48. Y. C. Bae, R. Faust, *Macromol. Symp.* **132**, 11–23 (1998).
49. K. M. Osten, P. Mehrkhodavandi, *Acc. Chem. Res.* **50**, 2861–2869 (2017).
50. L. C. Reibel, C. P. Durand, E. Franta, *Can. J. Chem.* **63**, 264–269 (1984).
51. J. Fejgin, W. Sadowska, M. Tomaszewicz, S. Penczek, *Makromol. Chem.* **95**, 155–167 (1966).
52. J. Brandup, E. H. Immergut, E. A. Grulke, *Polymer Handbook* (Wiley, ed. 4, 2003).
53. P. H. Plesch, P. H. Westermann, *J. Polym. Sci. Part C Polym. Symp.* **16**, 3837–3843 (1967).
54. P. G. M. Wuts, T. W. Greene, *Greene's Protective Groups in Organic Synthesis* (Wiley-Interscience, ed. 4, 2007).

### ACKNOWLEDGMENTS

The authors thank Y. D. Y. L. Getzler for invaluable discussion. **Funding:** This work was fully supported by the Joint Center for Energy Storage Research (JCESR), an Energy Innovation Hub funded by the US Department of Energy, Office of Science, Basic Energy Sciences. This work made use of the Cornell Center for Materials Research and the NMR Facility at Cornell University, which are supported by the NSF under awards DMR-1719875 and CHE-1531632, respectively. **Author contributions:** B.A.A. and R.L.S. designed and performed all experiments. G.W.C. directed research. All authors prepared the manuscript. **Competing interests:** B.A.A., R.L.S., and G.W.C. are inventors on US provisional patent application D-9743, submitted by Cornell University, which covers synthesis of polyacetals by RD-CROP and their chemical recycling. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

### SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/783/suppl/DC1
Materials and Methods
Synthetic Procedures
Figs. S1 to S63
Tables S1 to S29
References (*55*–*59*)

12 February 2021; resubmitted 10 May 2021
Accepted 2 July 2021
10.1126/science.abh0626

BLACK HOLES

# A characteristic optical variability time scale in astrophysical accretion disks

**Colin J. Burke[1,2], Yue Shen[1,3]*, Omer Blaes[4], Charles F. Gammie[1,3,5,6], Keith Horne[7], Yan-Fei Jiang[8], Xin Liu[1,3], Ian M. McHardy[9], Christopher W. Morgan[10], Simone Scaringi[11], Qian Yang[1,3]**

**Accretion disks around supermassive black holes in active galactic nuclei produce continuum radiation at ultraviolet and optical wavelengths. Physical processes in the accretion flow lead to stochastic variability of this emission on a wide range of time scales. We measured the optical continuum variability observed in 67 active galactic nuclei and the characteristic time scale at which the variability power spectrum flattens. We found a correlation between this time scale and the black hole mass extending over the entire mass range of supermassive black holes. This time scale is consistent with the expected thermal time scale at the ultraviolet-emitting radius in standard accretion disk theory. Accreting white dwarfs lie close to this correlation, suggesting a common process for all accretion disks.**

Accretion disks are present around growing supermassive black holes (SMBHs) found in active galactic nuclei (AGNs). Standard theory of radiatively efficient accretion disks (*1*) can reproduce the broadband emission from AGNs (*2*, *3*), but the exact structure and physical processes occurring in accretion disks remain unknown. Because AGN accretion disks are too small to resolve in direct observations, constraints on their structure have been derived from gravitational microlensing (*4*, *5*) and time delay measurements of accretion disk echoes to flux variations from the innermost region around the SMBH (reverberation mapping) (*6*–*9*).

For unknown reasons, optical emission from AGN accretion disks exhibits stochastic variability (*10*, *11*). Optical light curves (the time series of fluxes tracing the variable accretion disk emission) for large samples of AGNs can be used to measure the variability characteristics of the accretion disk emission. The power spectrum density (PSD) of AGN optical variability can be approximated by a damped random walk (DRW) model (*12*–*17*), which varies smoothly between a $f^{-2}$ power law (where $f$ is the frequency) at the high-frequency end and white noise at the low-frequency end (*12*, *13*). There are deviations from the $f^{-2}$ scaling at the highest frequencies (*16*, *18*) in some individual AGNs. The transition frequency, corresponding to a characteristic damping time scale ($\tau_{\rm damping}$), is typically several hundred days for quasars (*14*, *15*, *17*), the most luminous subset of AGNs with a bolometric luminosity $L_{\rm bol} \gtrsim 10^{45}$ erg s$^{-1}$. There is no widely accepted physical interpretation for this damping time scale. There is tentative evidence that it may correlate with the mass of the SMBH and/or luminosity of the AGN (*12*, *14*, *19*), but such claims have been controversial (*17*) and the results inconsistent (*12*, *14*, *15*). The range of SMBH mass in those studies has been limited to two orders of magnitude, and the measurements of the damping time scales are susceptible to biases because of the limited observing period (*20*, *21*).

To address these limitations, we compiled optical light curves from the literature for AGNs with estimated SMBH masses. To robustly constrain the damping time scale, we excluded any light curves that did not have sufficient signal-to-noise ratio or duration (*21*). Starting from an initial set of ~400 AGNs, our selection criteria led to a final sample containing 67 AGNs that spanned the entire SMBH mass range of ~$10^4$ to $10^{10}$ solar masses ($M_\odot$). We

[1]Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [2]Center for AstroPhysical Surveys, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [3]National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [4]Department of Physics, University of California, Santa Barbara, CA 93106, USA. [5]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [6]Illinois Center for Advanced Study of the Universe, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [7]School of Physics and Astronomy, University of St. Andrews, Fife KY16 9SS, UK. [8]Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA. [9]Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK. [10]Department of Physics, United States Naval Academy, Annapolis, MD 21402, USA. [11]Department of Physics, University of Durham, Durham DH1 3LE, UK.
*Corresponding author. Email: shenyue@illinois.edu

derived $\tau_{damping}$ for each AGN by fitting DRW models to each light curve (*21*). Herein, all time scales have been converted to the rest frame of the AGN and all quoted uncertainties and scatter are 1σ unless otherwise specified.

Figure 1 shows the relation between our derived damping time scales and the SMBH masses. There was a correlation (Pearson correlation coefficient $r$ = 0.82) over the SMBH mass range of ~$10^4$ to $10^{10}$ $M_\odot$. We verified that this correlation persisted if we made different choices for the details in measuring the damping time scale or methods of SMBH mass estimation (*21*). The best-fitting model relation is

$$\tau_{damping} = 107^{+11}_{-12}\ \text{days} \left(\frac{M_{BH}}{10^8\ M_\odot}\right)^{0.38^{+0.05}_{-0.04}} \quad (1)$$

where $M_{BH}$ is the mass of the SMBH. The data have an additional 1σ intrinsic scatter of $0.09^{+0.05}_{-0.04}$ dex around the best-fitting model. This relation is sufficiently tight that inversion of Eq. 1 can predict SMBH mass given $\tau_{damping}$ with a 1σ precision of ~0.3 dex. Alternatively, fitting a linear model for $M_{BH}$ given $\tau_{damping}$ yields

$$M_{BH} = 10^{7.97^{+0.14}_{-0.14}}\ M_\odot\ (\tau_{damping}/100\ \text{days})^{2.54^{+0.34}_{-0.35}} \quad (2)$$

with an intrinsic scatter of $0.33^{+0.11}_{-0.11}$ dex in $M_{BH}$. This intrinsic scatter in the predicted SMBH mass is similar to the systematic uncertainties in SMBH mass measurements (*22*, *23*). Previous studies of AGN optical variability (*14*, *15*) have found that the damping time scale depends weakly on wavelength λ as $\tau_{damping} \propto \lambda^{0.17}$. The measured damping time scale in different bands (and at different redshifts, $z$) were scaled using this relation to a rest frame wavelength of 2500 Å (Fig. 2). Because lower-mass systems were generally at lower redshifts than higher-mass systems in our AGN sample because of observational biases, a positive wavelength dependence of $\tau_{damping}$ slightly flattened its observed mass dependence. However, the measured weak wavelength dependence of the damping time scale (*14*, *15*) means that the mass dependence (i.e., the slope of the $\tau_{damping}$ – $M_{BH}$ relation) at a fixed rest frame wavelength is <0.5.

The light curve duration of our dataset was generally not long enough to constrain the damping time scale in the most massive (>$10^9$ $M_\odot$) or distant ($z$ > 1) AGNs. By restricting our sample to AGNs with $\tau_{damping}$ shorter than one-tenth of the light curve baseline (*21*), we potentially introduced a bias by underestimating the average damping time scale for the most massive or distant AGNs. Nevertheless, this caveat does not affect the existence of a damping time scale to mass correlation (*21*).

Most radiatively efficient AGNs accrete within a narrow range of accretion rates (normalized by SMBH mass), with an average Eddington ratio [the ratio between $L_{bol}$ and the mass-dependent maximum luminosity $L_{Edd} \equiv 1.3 \times 10^{38} (M_{BH}/M_\odot)$erg $s^{-1}$] of $L_{bol}/L_{Edd}$ ~ 0.2 and a dispersion of ~0.3 dex (*24*, *25*). This is similar to the median and dispersion of Eddington ratios in our sample (*21*). We found no correlation between the model residuals and $L_{bol}/L_{Edd}$, which we interpret as being caused by the limited dynamic range and large systematic uncertainties in the measured Eddington ratios. Any dispersion in the true $L_{bol}/L_{Edd}$ (i.e., free of measurement uncertainties) can potentially contribute to the intrinsic scatter around the average relation.

To extend the $\tau_{damping}$–mass scaling relation to accretors with much smaller masses, we considered accreting white dwarfs that are noneruptive (that is, the accretion rate is approximately stable). The $\tau_{damping}$ measurements for white dwarfs are based on optical light curves and taken directly from a previous study (*26*). A $\tau_{damping}$–mass scaling with a mass slope of 0.5 is consistent with measurements of accretion disks in these white dwarfs [Fig. 1 and (*21*)]. We did not consider optical variability in accretion disks around neutron stars or stellar mass black holes because the optical accretion disk emission could be complicated by x-ray reprocessing (*27*) or overwhelmed by optical light from a companion star.

This average scaling between the damping time scale and SMBH mass can be qualitatively understood within the standard theory of accretion disks. Both the orbital time $t_{orb}$ (the time to orbit around the black hole) and the thermal time $t_{th}$ (the time scale to restore thermal equilibrium) scale with the SMBH mass and radius as follows (*1*):

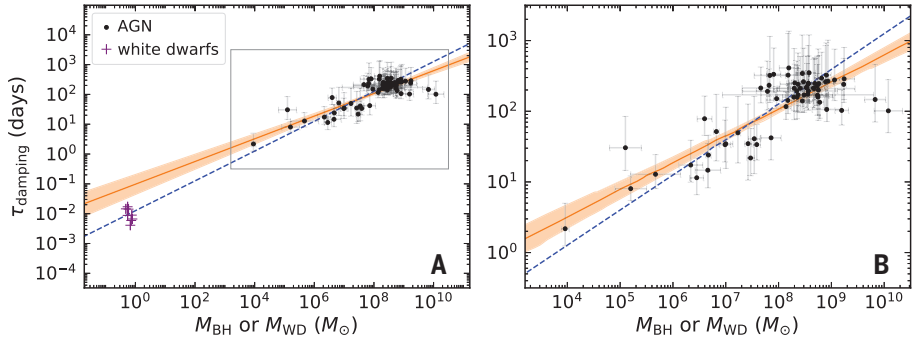



**Fig. 1. Optical variability damping time scale as a function of accretor mass.** (**A**) Rest frame damping time scale ($\tau_{damping}$) measured from AGN light curves plotted as a function of SMBH mass $M_{BH}$ for AGNs (black circles). The orange line and shaded band are the best-fitting model and 1σ uncertainty for the AGN sample, respectively. Purple crosses show equivalent measurements for white dwarfs (*26*), where $M_{WD}$ denotes the mass of the white dwarf; these do not fall in the orange band but are consistent with a model that has a fixed mass slope of 0.5 (blue dashed line). The typical uncertainties on $M_{WD}$ and the white dwarf damping time scale are 0.2 dex and 0.01 days, respectively (*26*). All error bars are 1σ. (**B**) Magnified view of the region within the gray box in (A).

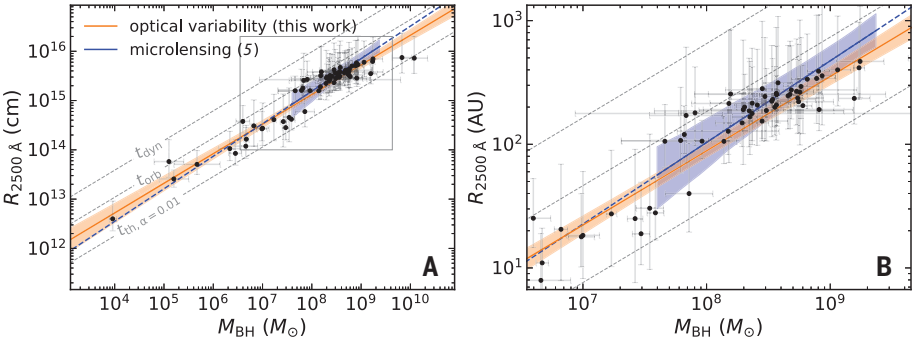



**Fig. 2. Accretion disk-emitting radius at rest frame 2500 Å as a function of SMBH mass.** (**A**) Emitting radius, computed as $R_{2500Å} \propto M_{BH}^{1/3} \alpha^{2/3} \tau_{damping}^{2/3}$, assuming that $\tau_{damping}$ is the thermal time and α = 0.05. The data points (black circles) are overplotted with the best-fitting linear model (orange line) and 1σ uncertainty (orange shaded area). The relation derived from microlensing observations (*5*) is shown in the overlapping mass range (blue solid line and 1σ shaded region). The blue dashed line indicates an extrapolation of the microlensing results to other black hole masses. The three gray dashed lines are the corresponding radius if $\tau_{damping}$ is identified as $t_{dyn}$, $t_{orb}$, or $t_{th}$ with α = 0.01, respectively. All error bars are 1σ. (**B**) Magnified view of the region within the gray box in (A) with radii converted to astronomical units (au).

$$t_{\rm orb} = 100\left(\frac{M_{\rm BH}}{10^8\,M_\odot}\right)\left(\frac{R}{100\,R_S}\right)^{3/2}\ \text{days} \quad (3)$$

$$t_{\rm th} = 1680\left(\frac{\alpha}{0.01}\right)^{-1} \times \left(\frac{M_{\rm BH}}{10^8\,M_\odot}\right)\left(\frac{R}{100\,R_S}\right)^{3/2}\ \text{days} \quad (4)$$

where $R_S = 2GM_{\rm BH}/c^2$ is the Schwarzschild radius of the black hole, $G$ is the gravitational constant, $c$ is the speed of light in vacuum, and $\alpha$ is the viscosity parameter. The relation among different time scales is: $t_{\rm th} \approx (2\pi\alpha)^{-1}t_{\rm orb} = \alpha^{-1}t_{\rm dyn} \approx (H/R)^2 t_{\rm vis}$, where $t_{\rm dyn}$ is the dynamical time (an alternative to $t_{\rm orb}$ used in the literature), $t_{\rm vis}$ is the viscous time on which matter diffuses through the accretion disk caused by viscosity, and $H/R$ is the ratio between the scale height $H$ and the radial extent $R$ of the accretion disk.

Assuming all AGNs accrete at constant Eddington ratio ($\dot{M}_{\rm BH}/M_{\rm BH} \propto L_{\rm bol}/L_{\rm Edd}$ = constant, where $\dot{M}_{\rm BH}$ is the mass accretion rate) and any dispersion in Eddington ratio leads to intrinsic scatter around the average relation, the standard theory predicts a scaling relation between the effective emitting radius (at a given rest wavelength) and the black hole mass as $R_\lambda \propto M_{\rm BH}^{2/3}$. We assumed that the radiative efficiency $\eta \propto L_{\rm bol}/\dot{M}_{\rm BH}$ is also constant. In reality, our sample contained AGNs with different accretion rates and possibly a range of black hole spins, which may lead to different values of $\eta$, introducing additional scatter around the average relation. At a given rest frame wavelength, the orbital and thermal time scales therefore scale with mass as $t_{\rm orb}$, $t_{\rm th} \propto M_{\rm BH}^{1/2}$. If the damping time scale that we measured were associated with the orbital time or the thermal time, then we would expect a slope of 0.5 in the $\tau_{\rm damping} - M_{\rm BH}$ relation, which is consistent with the observed slope within 2.5σ. A steeper wavelength dependence of $\tau_{\rm damping}$ than previously reported (*14*), as discussed above, would improve the agreement.

The physical origin of the damping time scale could be associated with the thermal time scale at the radius where variability is driven. To compare our results with AGN disk sizes measured from microlensing, we first scaled the damping time scale to rest frame 2500 Å using its measured wavelength dependence $\tau_{\rm damping} \propto \lambda^{0.17}$ (*14*). Assuming that this ultraviolet (UV)–emitting part of the disk is where variability is driven, we derived the effective UV-emitting radius as $R_{2500\text{Å}} \propto M_{\rm BH}^{1/3}\alpha^{2/3}\tau_{\rm damping}^{2/3}$ using Eq. 4.

Figure 2 shows the relation between the derived physical radius $R_{2500\text{Å}}$ that emits at rest frame 2500 Å and the SMBH mass for our sample. We found a correlation that is consistent with the prediction from the standard model, $R_{2500\text{Å}} \propto M_{\rm BH}^{2/3}$. We have assumed a fiducial viscosity parameter of $\alpha = 0.05$, which leads to $t_{\rm th} \approx 3t_{\rm orb} \approx 20t_{\rm dyn}$. This fiducial value of $\alpha$ is higher than the typical value of ~0.01 found in standard magnetohydrodynamic simulations of accretion disks (*28*) but is consistent with the range of 0.05 to 0.1 in simulations of radiation pressure–dominated AGN accretion disks (see the supplementary text). Figure 2 also shows the relation derived from microlensing measurements of accretion disk sizes for luminous quasars (*5*).

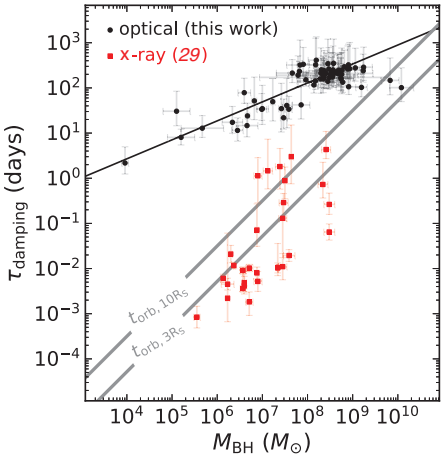



**Fig. 3. Comparison of AGN variability time scales at optical and x-ray wavelengths.** Red squares show measurements of the break time scale for x-ray variability (*29*), and black circles are our optical measurements (same as Fig. 1), both as functions of the SMBH mass. The thick gray lines indicate the orbital time scale at 10 and three times the Schwarzschild radius, which has a linear dependence on mass. The optical and x-ray data have different slopes. The correlation in the optical is tighter than the x-ray correlation. All error bars are 1σ.

The size-mass relation that we derived from the optical variability data is consistent with the constraints from microlensing in the overlapping mass range (Fig. 2) but also extends to lower masses. This suggests an association of the damping time scale with the thermal time at the effective UV-emitting radius. Because the normalization of our $\tau_{\rm damping} - M_{\rm BH}$ relation was constrained to within ~10% (1σ) and we considered the microlensing results reliable, $t_{\rm dyn}$ or $t_{\rm orb}$ was less favorable than $t_{\rm th}$ (with $\alpha = 0.05$) as the origin for $\tau_{\rm damping}$. If $\tau_{\rm damping}$ were associated with $t_{\rm th}$, then $\alpha$ must be in the range of 0.03 to 0.12 to be within ~1σ of the microlensing constraints. Both our analysis and the microlensing study assumed the same standard accretion disk model but differed in additional assumptions. For example, our variability approach assumed that the damping time scale was the thermal time scale with a fiducial viscosity parameter $\alpha = 0.05$ without needing to know the orientation of the disk; the microlensing analysis did not make this assumption on $\alpha$ but did assume a mean orientation of the disk. The correlation in Fig. 2 is tighter (Pearson correlation coefficient $r = 0.96$) than in Fig. 1 because the computation of $R_{2500\text{Å}}$ includes an explicit mass dependence, biasing the correlation strength. Nevertheless, the agreement with the microlensing results at the high-mass end and a scaling that was different from $R_{2500\text{Å}} \propto M_{\rm BH}^{1/3}$ expected from pure self-correlation led us to conclude that the correlation seen in Fig. 2 is real and not due to self-correlation.

Figure 3 compares the optical damping time scale of the accretion disk with the characteristic x-ray variability time scale at different AGN SMBH masses with the x-ray variability measurements obtained from a previous study (*29*). x-ray emission in AGNs mainly arise from a hot, optically thin but geometrically thick gas (a region referred to as the corona), much closer to the SMBH than the UV-optical–emitting part of the accretion disk, so the characteristic x-ray variability time scales are expected to be substantially shorter (e.g., Eqs. 3 and 4). The best-fitting time scale–mass relation for x-ray variability has a slope close to unity (*26*, *29*–*31*), as would be expected if the x-ray time scale traces the orbital or thermal time near the innermost stable circular orbit (ISCO), which has a radius that scales linearly with black hole mass. By contrast, the characteristic time scales measured from optical variability were several orders of magnitude longer than the x-ray time scale and had a different mass dependence. The higher scatter in the x-ray relation compared with the optical relation could have been caused by other parameters that affect the ISCO radius (such as black hole spin) or by the range of x-ray–emitting radii within the optically thin corona. The thermal time scale described by Eq. 4 does not apply to the corona, but we expect that any x-ray variability would operate on the orbital time scale in the corona region.

The correlation between the damping time scale of optical variability and the SMBH mass has implications for AGN accretion disk models. It implies an intrinsic origin for AGN optical variability, as opposed to extrinsic causes such as microlensing. The difference with x-ray variability ruled out simple reprocessing of x-ray emission into the optical variability (*27*) on similar time scales as the damping time scale, instead requiring internal accretion disk processes that either drive the optical variability themselves or modify the x-ray reprocessing.

There is no detailed physical model that can explain the observed variability characteristics of accretion disks (see the supplementary text), but the standard accretion disk model (*1*) provides qualitative agreement with the observed time scale–mass relation over 10 orders of magnitude in accretor mass (combining AGNs and

white dwarfs). The association of the characteristic variability time scale with the thermal time scale of the accretion disk explains the observed low-frequency break in the optical variability PSD. Figure 3 and fig. S7 show that the x-ray variability time scale is consistent with the orbital time at the ISCO, although the large scatter and the small number of stellar mass black holes (in fig. S7) cannot rule out other time scales (such as the viscous time $t_{vis}$) as the origin for the break in the x-ray variability PSD.

It remains unclear which processes drove these accretion disk flux variations and whether additional accretion parameters (such as accretion rate, black hole spin, etc.) were involved. The measured AGN accretion disk sizes from microlensing were larger than predicted by the standard model (*32*, *33*). The measured wavelength dependence of the damping time scale (*14*) was shallower than standard model predictions ($\tau \propto \lambda^2$), implying a more complicated mapping from the damping time scales to the physical radii as a function of wavelength. We speculate that the variability was driven in the inner part of the accretion disk, emitting at rest frame UV, which induced optical variability by rapid outward propagation, during which the damping time scale was approximately preserved (see the supplementary text). In other words, the damping time scale traces the thermal time scale at the UV-emitting part of the accretion disk, even when measured at longer (e.g., optical) wavelengths.

Regardless of the physical mechanism, the observed $\tau_{damping} - M_{BH}$ relation can be used to estimate the SMBH mass of an AGN using optical variability. The correlation parameters are sufficiently well constrained to provide mass estimates that are as accurate as reverberation mapping and single-epoch methods. The method can be applied to AGNs at the low-mass end of SMBHs, where the broadline emission is often too weak to measure a robust SMBH mass using spectral methods.

## REFERENCES AND NOTES


1. N. I. Shakura, R. A. Sunyaev, *Astron. Astrophys.* **24**, 337 (1973).
2. G. A. Shields, *Nature* **272**, 706–708 (1978).
3. W. H. Sun, M. A. Malkan, *Astrophys. J.* **346**, 68 (1989).
4. C. W. Morgan, C. S. Kochanek, N. D. Morgan, E. E. Falco, *Astrophys. J.* **712**, 1129–1136 (2010).
5. C. W. Morgan *et al.*, *Astrophys. J.* **869**, 106 (2018).
6. S. G. Sergeev, V. T. Doroshenko, Y. V. Golubinskiy, N. I. Merkulova, E. A. Sergeeva, *Astrophys. J.* **622**, 129–135 (2005).
7. E. M. Cackett, K. Horne, H. Winkler, *Mon. Not. R. Astron. Soc.* **380**, 669–682 (2007).
8. M. M. Fausnaugh *et al.*, *Astrophys. J.* **821**, 56 (2016).
9. R. Edelson *et al.*, *Astrophys. J.* **840**, 41 (2017).
10. M.-H. Ulrich, L. Maraschi, C. M. Urry, *Annu. Rev. Astron. Astrophys.* **35**, 445–502 (1997).
11. P. Padovani *et al.*, *Astron. Astrophys. Rev.* **25**, 2 (2017).
12. B. C. Kelly, J. Bechtold, A. Siemiginowska, *Astrophys. J.* **698**, 895–910 (2009).
13. S. Kozłowski *et al.*, *Astrophys. J.* **708**, 927–945 (2010).
14. C. L. MacLeod *et al.*, *Astrophys. J.* **721**, 1014–1033 (2010).
15. K. L. Suberlak, Ž. Ivezić, C. MacLeod, *Astrophys. J.* **907**, 96 (2021).
16. Y. Zu, C. S. Kochanek, S. Kozlowski, A. Udalski, *Astrophys. J.* **765**, 106 (2013).
17. T. Simm *et al.*, *Astron. Astrophys.* **585**, A129 (2016).
18. R. F. Mushotzky, R. Edelson, W. Baumgartner, P. Gandhi, *Astrophys. J.* **743**, L12 (2011).
19. S. Collier, B. M. Peterson, *Astrophys. J.* **555**, 775–785 (2001).
20. S. Kozłowski, *Astron. Astrophys.* **597**, A128 (2017).
21. Materials and methods are available as supplementary materials.
22. B. M. Peterson, *Space Sci. Rev.* **183**, 253–275 (2014).
23. Y. Shen, *Bull. Astron. Soc. India* **41**, 61 (2013).
24. J. A. Kollmeier *et al.*, *Astrophys. J.* **648**, 128–139 (2006).
25. Y. Shen *et al.*, *Astrophys. J. Suppl. Ser.* **194**, 45 (2011).
26. S. Scaringi *et al.*, *Sci. Adv.* **1**, e1500686 (2015).
27. C. Done, M. Gierliński, A. Kubota, *Astron. Astrophys. Rev.* **15**, 1–66 (2007).
28. S. A. Balbus, J. F. Hawley, *Rev. Mod. Phys.* **70**, 1–53 (1998).
29. O. González-Martín, S. Vaughan, *Astron. Astrophys.* **544**, A80 (2012).
30. I. M. McHardy, E. Koerding, C. Knigge, P. Uttley, R. P. Fender, *Nature* **444**, 730–732 (2006).
31. E. G. Körding *et al.*, *Mon. Not. R. Astron. Soc.* **380**, 301–310 (2007).
32. J. Dexter, E. Agol, *Astrophys. J.* **727**, L24 (2011).
33. M. Sun *et al.*, *Astrophys. J.* **891**, 178 (2020).
34. C. J. Burke *et al.*, Data for: A characteristic optical variability timescale in astrophysical accretion disks, version 1, Zenodo (2021); https://doi.org/10.5281/zenodo.4914484.


## ACKNOWLEDGMENTS


**Funding:** C.J.B. acknowledges support from an Illinois Graduate Survey Science Fellowship. Y.S. was supported by NSF grant AST-2009947. C.F.G. was supported by NSF grants AST-1716327 and OISE-1743747. K.H. was supported by UK STFC grant ST/R000824/1. I.M.M. was supported by UK STFC grant ST/R000638/1. C.W.M. was supported by NSF grant AST-2007680. **Author contributions:** C.J.B. led the data compilation and analysis. Y.S. designed the project and led the manuscript writing. O.B., C.F.G., and Y.-F.J. led the theoretical interpretation. X.L. and Q.Y. contributed to data compilation. I.M.M. and S.S. led the x-ray variability and white dwarf discussion. C.W.M. led the microlensing discussion. K.H. led the disk reverberation mapping discussion. All authors contributed to the scientific interpretation and manuscript writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Optical light curves of the AGN sample were taken from the publicly available sources listed in table S1 and data S1; our compilation is archived on Zenodo (*34*). White dwarf optical variability time scale measurements were taken from (*26*). x-ray variability time scale measurements were from (*26*, *29*). Our derived optical $\tau_{damping}$ measurements are provided in table S1 (for the final sample) and data S1 (for the initial sample). The full figure set for our initial AGN sample (an example is shown in fig. S5) is also available on Zenodo (*34*). The software used is publicly available from the cited references, and a Python notebook to fully reproduce our analysis is available at https://github.com/burke86/taufit/tree/master/paper .


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/789/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Table S1
References (*35*–*92*)
Data S1

10 February 2021; accepted 11 June 2021
10.1126/science.abg9933

PALEOBOTANY

# A fossil record of land plant origins from charophyte algae

**Paul K. Strother[1]* and Clinton Foster[2]**

**Molecular time trees indicating that embryophytes originated around 500 million years ago (Ma) during the Cambrian are at odds with the record of fossil plants, which first appear in the mid-Silurian almost 80 million years later. This time gap has been attributed to a missing fossil plant record, but that attribution belies the case for fossil spores. Here, we describe a Tremadocian (Early Ordovician, about 480 Ma) assemblage with elements of both Cambrian and younger embryophyte spores that provides a new level of evolutionary continuity between embryophytes and their algal ancestors. This finding suggests that the molecular phylogenetic signal retains a latent evolutionary history of the acquisition of the embryophytic developmental genome, a history that perhaps began during Ediacaran-Cambrian time but was not completed until the mid-Silurian (about 430 Ma).**

Establishing the timing of land plant (embryophyte) origins provides an important constraint for modeling evolving environmental conditions on the Earth's surface, including many aspects of the global carbon cycle (*1*), throughout geologic time. However, the origin of the embryophytes did not occur as a singularity in geologic time, it happened over a period of time as preexisting algal genes, and de novo genes were assembled into the genome that specifies embryonic development in the plant sporophyte that we see today (*2*). The algal-plant transition is also intimately linked to adaptation to living on the land surface. Bower (*3*) long ago used this fact to propose that it was the serial accumulation of morphological adaptation to the subaerial environment during the evolution of an "interpolated" sporophyte phase in the life cycle that characterized the origin of the land plants. His morphological and developmental model predicted that the plant spore preceded the evolutionary origin of the plant sporophyte, a supposition subsequently borne out by both the spore fossil record (*4*, *5*) and the morphological dynamics of meiosis and spore development in bryophytes today (*6*).

Refinement of molecular clock time trees (*7*, *8*) continues to approach an Ediacaran to

[1]Weston Observatory, Department of Earth and Environmental Sciences, Boston College, Weston, MA 02493, USA. [2]Research School of Earth Sciences, The Australian National University, Canberra ACT 2601, Australia.
*Corresponding author. Email: strother@bc.edu

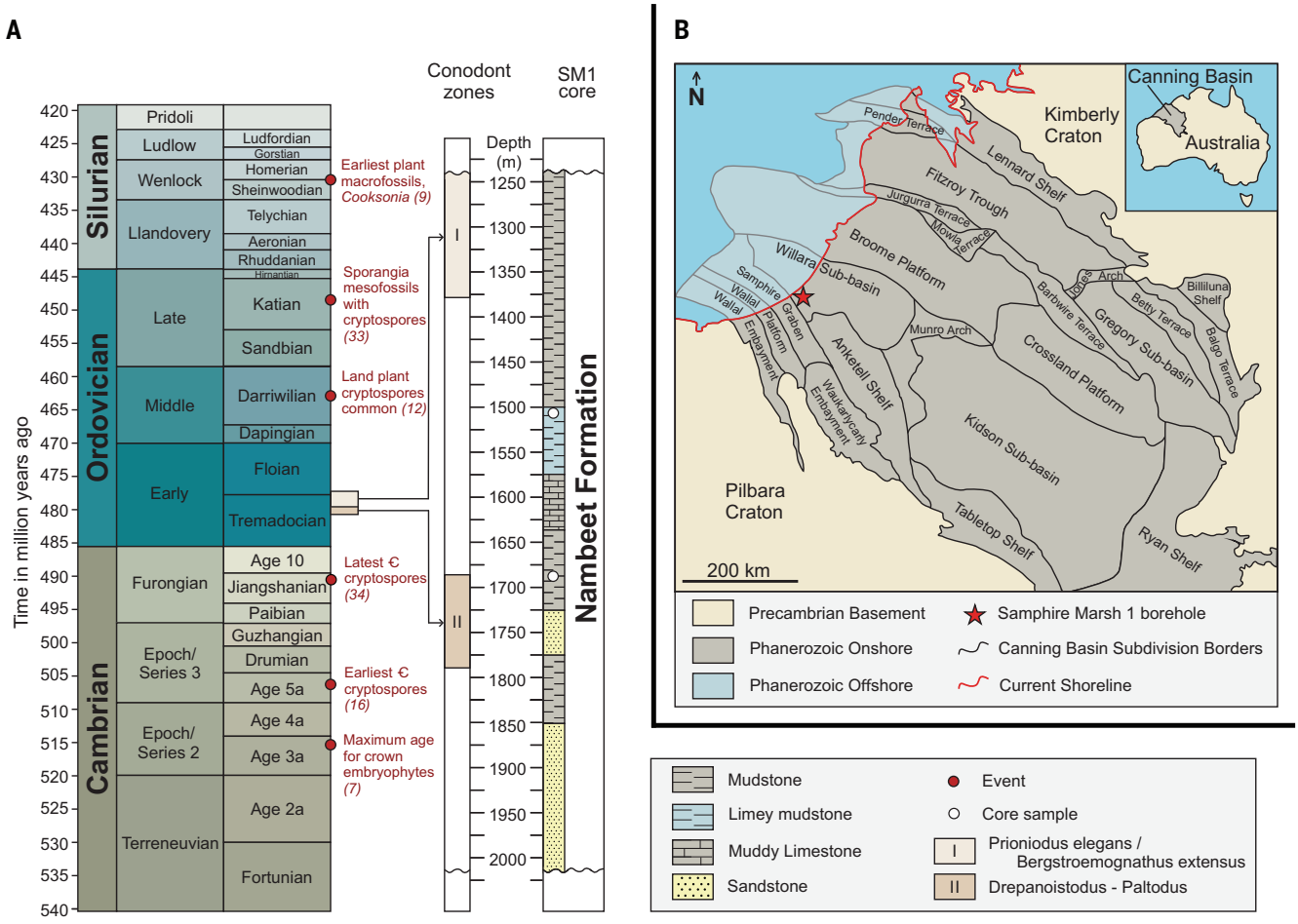



**Fig. 1. Stratigraphic placement, evolutionary context, and location of the Samphire Marsh 1 (SM1) borehole.** (**A**) Stratigraphic position of the SM1 assemblages in the context of the major landmarks in lower Paleozoic plant evolution and position within the SM1 core. Important events in plant evolution documented by cryptospores were as follows: "Earliest plant macrofossils, *Cooksonia*" (*9*), "Sporangia mesofossils with cryptospores" (*35*), "Land plant cryptospores abundant" (*12*), "Latest Cambrian cryptospores" (*20*), "Earliest Cambrian cryptospores" (*16*), and "Maximum age for crown embryophytes" (*7*). Stratigraphic columns are based on (*36*). (**B**) Map showing location of SM1 borehole and geologic structural units of the Canning Basin [according to (*37*)].

Cambrian maximum age for the origin of crown group land plants. This has led to ad hoc explanations of the much later arrival of the earliest plant mesofossil, *Cooksonia* (*9*), in the rock record; for example, the incompleteness of the rock record (*10*), the lack of readily fossilizable plant tissues (*7*), and the scarcity of lower Paleozoic terrestrial deposits (*1*). Fossil spores of Middle Ordovician (Darriwilian) age exhibit features that are shared with the spore-bearing embryophytes, including wall ultrastructure (*11*), and spores borne in either tetrahedral (*12*) or dyad (*13*) configurations. Records of older Cambrian "spore-like" remains (*14*–*16*) have been used to inform soft lower boundaries in molecular clocks (*7*, *17*), but their relevance to the question of embryophyte origins has also been dismissed because, it is argued, they lack characters unique to the embryophytes (*1*, *18*).

In fact, Cambrian cryptospores do not have the key synapomorphy that aligns with embryophyte spores: spore bodies borne in tetrahedral tetrads. Instead, they generally occur in spore packets, which include varying numbers of spore bodies that may appear to be in various stages of development. Their interpretation as charophytes, as opposed to other algae (*4*), is based on the observation that living members of some charophyte algae (e.g., *Coleochaete* Brébisson 1844) undergo irregular forms of meiosis (*19*) that are consistent with the endosporic development seen in Cambrian spore packets (*4*, *20*). In addition, sporoderm characteristics of Cambrian and Ordovician cryptospores appear to be homologous to the lamellated spore walls of some crown group liverworts (*4*, *15*, *21*).

With the exception of a single occurrence in South China (*22*), Cambrian cryptospores are known only from the Laurentian paleocontinent (*14*–*16*). Here, we record an assemblage of cryptospores that occurs in the Tremadocian (Lower Ordovician) of the Canning Basin of northern Western Australia, corresponding to an age of ~480 million years ago (Ma). These populations of cryptospores exhibit morphological variations that bear similarities to both older Cambrian cryptospores from Laurentia and younger Ordovician forms from Gondwana, filling in a temporal gap in the prior cryptospore fossil record and expanding the geographical distribution of the fossil record of the algal-embryophyte transition.

The cryptospore assemblages occur in the Samphire Marsh 1 borehole (Fig. 1A and supplementary text), which was drilled in the southern part of the Canning Basin in 1958 (Fig. 1B). Cryptospore-dominated assemblages occur in the type Nambeet Formation, a 775-m-thick sequence of gray-green glauconitic shales and siltstones interbedded with limestones, which overlies a basal fine- to coarse-grained sandstone unit (*23*). Paleogeographic reconstruction (fig. S1) places Samphire Marsh 1 in the intertidal zone of a transgressive sequence unconformably overlying Cambrian granitic basement. Nicoll (*24*) found faunas from overlying cores 4 and 5 belonging

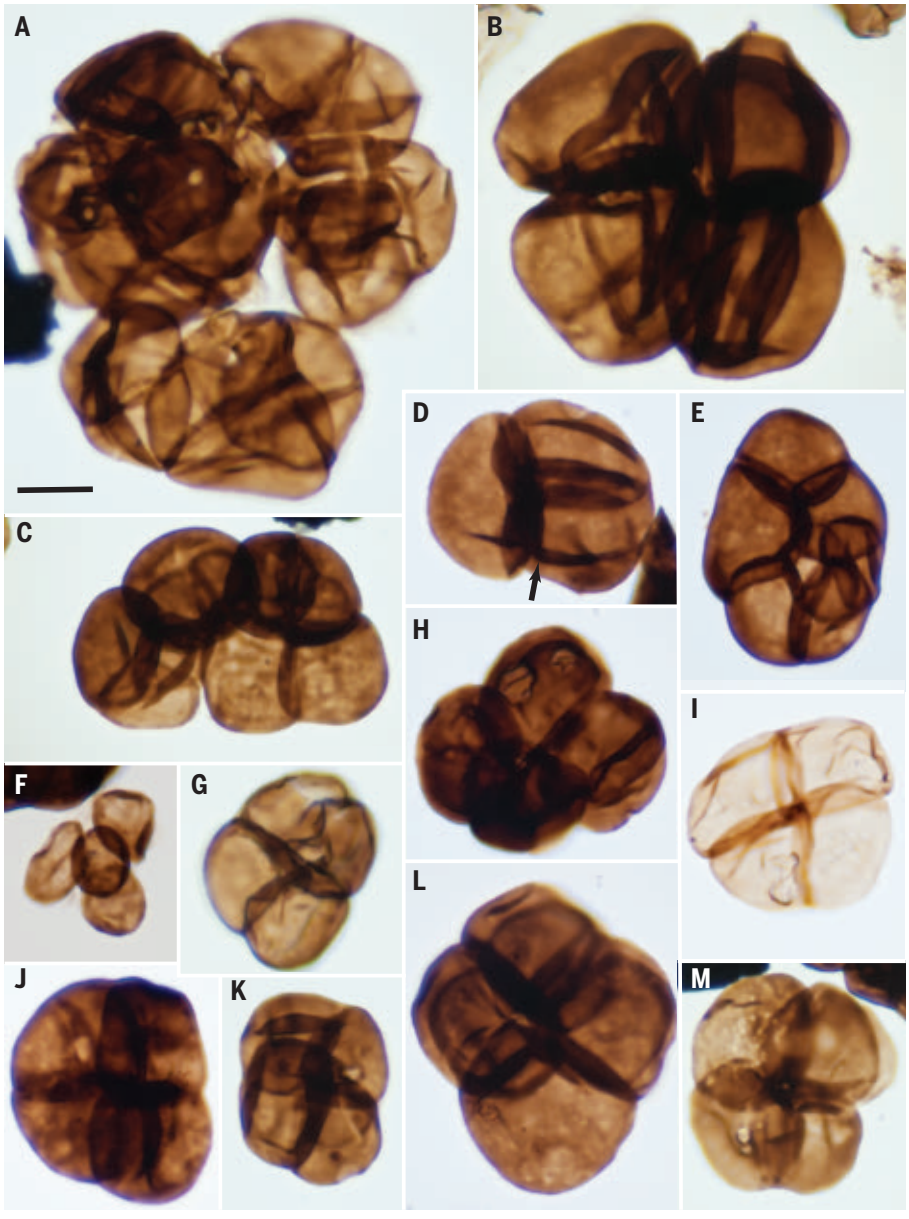



**Fig. 2. Cryptospores from the SM1 borehole: spore packets and tetrads.** (**A**) Cluster of six packets of cryptospores typical of preservation in the assemblage. (**B**) Planar tetrad of smooth-walled cryptospores. (**C**) Irregular cluster of circular cryptospores, *Spissuspora* cf. *S. laevigata*. (**D**) Cryptospore triad showing an immature dyad adpressed to a larger monad. Note the fold (arrow) that crosses both spore bodies. (**E**) Quasiplanar cryptospore polyad showing seven spore bodies. (**F**) *Rimosotetras* cf. *R. subsphaerica*, a tetrahedral tetrad of small, loosely attached isometric spore bodies. (**G**) Simple thin-walled cryptospore tetrad. (**H**) Cryptospore tetrad with highly folded wall. (**I**) Planar tetrad similar to *T. laevigatus* but with a thinner wall and smaller diameter than the type population. (**J**) Cryptospore tetrad. (**K**) Cryptospore tetrad with folded walls. (**L**) Cryptospore tetrad with a singular larger spore body. (**M**) Cryptospore tetrad with folded walls. Scale bar in (A) is 10 μm and applies to all images; sample depth interval was 5535 to 5547 ft (1687 to 1691 m).

to the *Prioniodus elegans–Bergstroemognathus extensus* Zone that are Arenig (now, Floian) in age. Assemblages from underlying cores 8 and 9 were assigned to the *Drepanoistodus–Paltodus* Zone of late Tremadocian age. On the basis of these conodont biozonations, the age of the cryptospore assemblages reported here from cores 6 and 8 is late Tremadocian (Early Ordovician), corresponding to ~480 Ma (Fig. 1A).

The assemblage is dominated by highly variable clusters of packets of spore-like microfossils (Fig. 2A), along with isolated spore polyads (Fig. 2, B to E), tetrads (Fig. 2, F to M), and dyads (Fig. 3, A to E and G to I). Isolated spore monads such as *Laevolancis divellomedium* (Chibrikova) Burgess and Richardson 1991 (fig. S2A) also occur, and, although their provenance from embryophytes is unclear, it is noteworthy that previous records are from the Late Ordovician (Hirnatian) through the late Silurian (Přídolí) and younger (*25*). All cryptospore walls in this deposit appear simple, without any evidence of added sculpturing. Surface textures vary from smooth (e.g., Figs. 2, A and B, and 3, C, E, and G) to somewhat blotchy in appearance (e.g., Figs. 2, C and J, and 3, A and H), reflecting variations in wall thickness. Overall, two general kinds of spore walls are present: those with a more or less homogeneous structure (e.g., Figs. 2, C to E, and 3, A and H) and those that show features indicating an underlying laminated sporoderm (e.g., Figs. 2, A, B, H, and L, and 3G). The wall ultrastructure of multilaminate cryptospores is well documented elsewhere (*15*, *21*, *26*), and the sporomorphs seen here are similar enough to enable us to interpret their wall structure using light microscopy. Individual laminae in these multilaminate walls may fold back on themselves, generating the appearance of irregular (Fig. 2H) to concentric thickenings or bands (Figs. 2L and 3G, arrows) when viewed under light microscopy. In other cases, linear folds in the spore wall may cross between spore-body pairs, serving to reinforce the fact that these forms are not simply random associations of unrelated cells but rather have a common developmental origin. The first wall type with thicker homogeneous structure first occurs in the older, Cambrian cryptospore *Spissuspora* Strother 2016, but homogeneous walls are more characteristic of Darriwilian and younger cryptospores (*11*).

An important feature that is used to discriminate between cryptospores and simple clusters of vegetative algal cells is the intimate nature of cell-cell contact that persists in cryptospores as evidence of developmental continuity during sporogenesis. For example, contact regions between the members of dyads tend to occupy nearly the full diameter of the spore bodies, a feature that is evident in all dyads illustrated in Fig. 3. In simple dyads, the thickened contact surface just corresponds to an upturned fold of the contact face (e.g., Fig. 3A), but in some specimens (Fig. 3C and fig. S2, C and F) a marginal thickening on the proximal spore contact face is quite pronounced, corresponding to a thin equatorial cingulum in dispersed true spores. Therefore, cell-cell contact in these forms does not appear to be random; these dyads and tetrads were formed together as end products of cytokinesis of a common generative cell, effectively a spore mother cell. Although tetrahedrally aligned spore bodies are not characteristic of the assemblage overall, a few, corresponding to *Rimosotetras* cf. *R. subsphaerica* Strother,

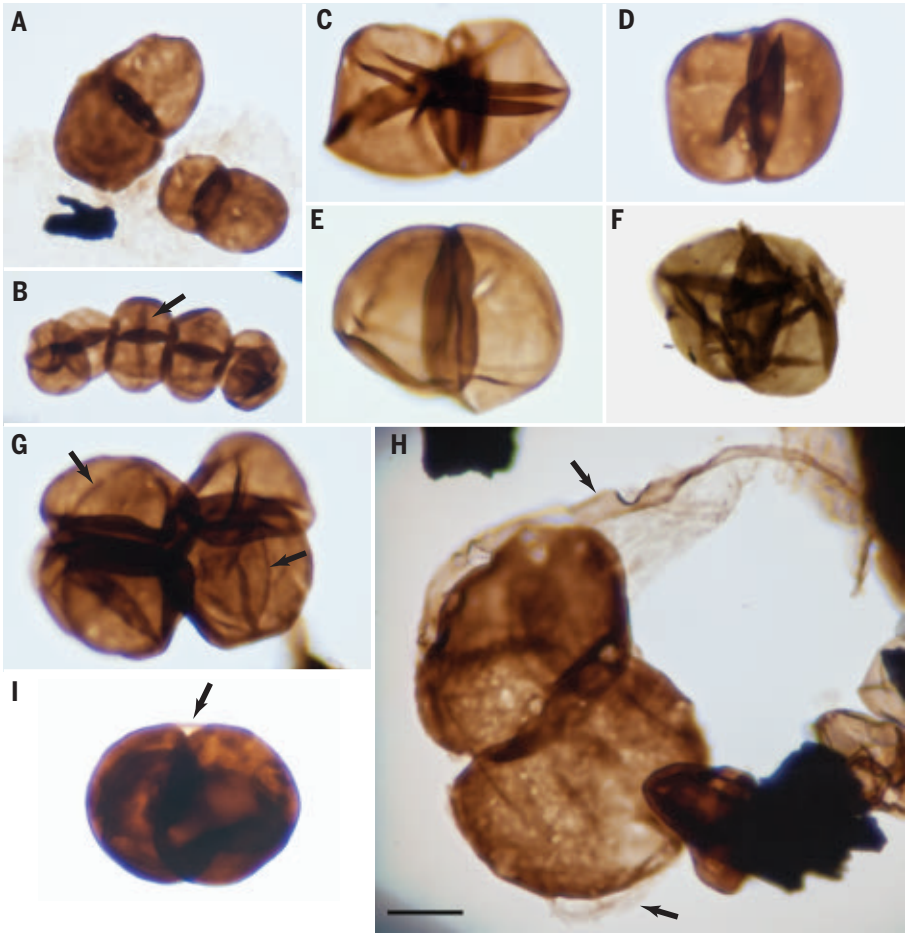



**Fig. 3. Cryptospore dyads from the SM1 borehole and the Silurian of Pennsylvania.** (**A** to **E** and **G** to **I**) Cryptospore dyads from the SM1 borehole. (A) Two simple spore dyads that are irregular in size and have blotchy walls. (B) Row of six aligned small dyads. Note the presence of a transverse thickened band (arrow). (C) *D. murusattenuata*. Note the thickened contact ring and folded walls, features that characterize this species. (D) *D. murusattenuata*, a simple smooth-walled dyad pair. (E) *D. murusattenuata*. (**F**) Cryptospore dyad from the Silurian of Pennsylvania. Shown is a *D. murusattenuata* paratype, which is Silurian (Wenlock) in age. (**G**) Planar set of two dyads with thin folded walls. This form is similar to the Cambrian species *A. geminus*. (**H**) *Abditusdyadus laevigatus* with associated partially detached membrane (arrows). (**I**) *Abditusdyadus* sp. with a very thin enclosing membrane (arrow) and thick, blotchy walls. Scale bar in (H) is 10 μm and applies to all images; sample depth interval was 5535 to 5547 ft (1687 to 1691 m).

Traverse and Vecoli 2015, have been found (Fig. 2F). Similar, although larger, loosely arranged tetrahedral tetrads also occur in the Darriwilian of Saudi Arabia (*12*, *27*). Overall, in this assemblage, most tetrads exhibit either planar (Fig. 2, G to M) or paired dyad (Fig. 3G) configurations. This latter topology is more characteristic of older, Cambrian forms, especially *Adinosporus geminus* Strother 2016. Conversely, planar tetrads composed of isometric spore bodies with simple walls (e.g., Fig. 2, G and I) appear to be comparable to *Tetraplanarsporites laevigatus* Wellman, Steemans and Miller 2015, which is known from the Upper Ordovician (*28*, *29*). Finally, many smooth-walled dyads illustrated here (Fig. 3, C to E) are not readily distinguishable from the widely distributed taxon, *Dyadospora murusattenuata* Strother & Traverse 1979, a paratype (Silurian, Wenlock) of which is illustrated in Fig. 3F for comparison.

There is a distinct lack of envelopes or enclosing membranes in the assemblage; in fact, we were able to find only two examples of this feature. Figure 3I (arrow) reveals a very thin membrane traversing the marginal gap between the two spore members of a dyad. Such a thin structure is not likely the residual spore wall of a resting cyst, especially given the robust nature of the spores themselves. Figure 3H (arrows) preserves a fragmentary diaphanous membrane that has broken away from a rather robust dyad. Envelopes surrounding spore tetrads and dyads from Silurian deposits were once speculated to be the relictual walls of an ancestral algal zygote (*30*). However, cryptospores from Lower and Middle Ordovician strata do not consistently show envelope-enclosed forms to be more prevalent than in the Silurian. Neither do Ordovician cryptospore assemblages show evidence of persistent zygospores or aplanospores that might indicate a zygnematacean affinity, adding to a perception that envelopes enclosing cryptospores of late Ordovician and Silurian age were derived from spore mother cells, not from ancestral zygospore walls. This implies that the robust enveloping walls seen in Silurian *Abditusdyadus* Wellman and Richardson 1996 and *Velatitetras* Burgess 1991 are not plesiomorphic with respect to the algal ancestry of the land plants.

A salient feature of Bower's original antithetic hypothesis was a prediction that mitotic divisions of (diploid) zygotes would first produce multicellular thalli of sporogenous cells, which only later in evolutionary time would lose their totipotency to function as fully differentiated vegetative cells. The fossil record shows that this is likely to be the case, first in the retention of spore-like packets attached in short linear chains (e.g., Fig. 3B), but more substantially in sheets of paired spore-like cells, as seen here in *Grododowon orthogonalis* Strother 2017 (Fig. 4, A and B, and fig. S2J). *G. orthogonalis*, which forms planar sheets generated by successive orthogonal cell divisions, has been described elsewhere from Middle (*31*) and Upper Ordovician (*32*) deposits. Rosettes of *G. orthogonalis* show a growth pattern that is similar to thallus growth in extant *Coleochaete* (*31*), but the basic pattern of orthogonal cell division seen here is certainly not restricted to the charophyte algae. The relation to potential vegetative tissues is unknown for *G. orthogonalis*, but the specimen illustrated in Fig. 4A retains an intriguing, ring-like fragment (arrow) that may represent an attachment feature or a remnant of a vegetative structure or covering.

It is tempting to conclude that the spores described here document a Lower Ordovician origin to crown group land plants, thus bringing molecular time trees and fossils into a closer alignment. However, that approach masks the far more intriguing possibility that the fossil record does indeed inform us of the tempo and mode of the evolutionary origins of plant development. This Lower Ordovician

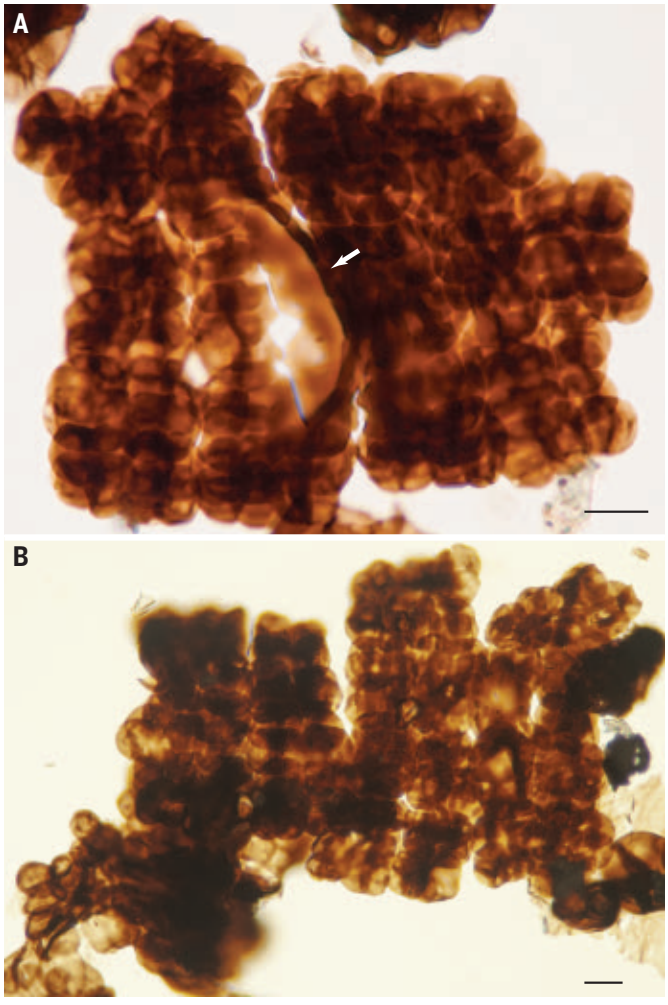



**Fig. 4. *G. orthogonalis* from the SM1 borehole.** (**A**) *G. orthogonalis* comprises planar sheets of geometrically aligned spore packets, here with an attached arcuate structure (arrow). (**B**) This larger specimen shows varying degrees of preservational quality but retains the orthogonal alignment of small spore dyads that characterize the genus. Scale bars are 10 μm; sample depth interval was 5535 to 5547 ft (1687 to 1691 m).

assemblage includes some taxa that range into the Late Ordovician and others that are known previously from middle Cambrian deposits (fig. S3). It is the intermediate nature of the assemblage that strengthens our assertion that the earlier cryptospore record is relevant to the question of the origin of land plants. These pre-Darriwilian cryptospores are direct fossil evidence of charophyte adaptation to terrestrial settings. The genomic component of that adaptation that was subsequently incorporated into the embryophyte genome might account for molecular clock dates that precede the arrival of fossilized plant axes in the rock record. This idea, that early cryptospores are a manifestation of spore development that preceded vegetative plant development, is also consistent with sedimentological (*33*) and other earth system proxies (*34*) for the effects of land plants indicating a later, Silurian origin to the crown group embryophytes.

## REFERENCES AND NOTES


1. P. Kenrick, C. H. Wellman, H. Schneider, G. D. Edgecombe, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 519–536 (2012).
2. S. K. Floyd, J. L. Bowman, *Int. J. Plant Sci.* **168**, 1–35 (2007).
3. F. O. Bower, *The Origin of a Land Flora: A Theory Based on the Facts of Alternation* (Macmillan, 1908).
4. P. K. Strother, W. A. Taylor, "The evolutionary origin of the plant spore in relation to the antithetic origin of the plant sporophyte" in *Transformative Paleobotany*, M. Krings, C. J. Harper, N. R. Cuneo, G. W. Rothwell, Eds. (Elsevier, 2018), pp. 3–20.
5. C. H. Wellman, J. Gray, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**, 717–731, discussion 731–732 (2000).
6. R. C. Brown, B. E. Lemmon, *New Phytol.* **190**, 875–881 (2011).
7. J. L. Morris *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2274–E2283 (2018).
8. Y. Nie *et al*., *Syst. Biol.* **69**, 1–16 (2020).
9. M. Libertín, J. Kvaček, J. Bek, V. Žárský, P. Štorch, *Nat. Plants* **4**, 269–271 (2018).
10. A. B. Smith, A. J. McGowan, *Paleobiology* **34**, 155–161 (2008).
11. W. A. Taylor, P. K. Strother, M. Vecoli, S. Al-Hajri, *Rev. Micropaleontol.* **60**, 281–288 (2017).
12. P. K. Strother, A. Traverse, M. Vecoli, *Rev. Palaeobot. Palynol.* **212**, 97–110 (2015).
13. K. S. Renzaglia *et al*., *Bot. J. Linn. Soc.* **179**, 658–669 (2015).
14. P. K. Strother, J. H. Beck, "Spore-like microfossils from Middle Cambrian strata: Expanding the meaning of the term cryptospore" in *Pollen and Spores*, M. M. Harley, C. M. Morton, S. Blackmore, Eds. (The Systematics Association, 2000), pp. 413–424.
15. P. K. Strother, G. D. Wood, W. A. Taylor, J. H. Beck, *Mem. Assoc. Australas. Palaeontol.* **29**, 99–113 (2004).
16. P. K. Strother, *Rev. Palaeobot. Palynol.* **227**, 28–41 (2016).
17. J. T. Clarke, R. C. M. Warnock, P. C. J. Donoghue, *New Phytol.* **192**, 266–301 (2011).
18. C. H. Wellman, "Dating the origin of land plants," in *Telling the Evolutionary Time*, P. C. J. Donoghue, M. P. Smith, Eds. (CRC Press, 2003), pp. 119–141.
19. D. Haig, *J. Phycol.* **46**, 860–867 (2010).
20. W. A. Taylor, P. K. Strother, *Rev. Palaeobot. Palynol.* **153**, 296–309 (2009).
21. W. A. Taylor, *Rev. Palaeobot. Palynol.* **156**, 7–13 (2009).
22. L. Yin, Y. Zhao, L. Bian, J. Peng, *Sci. China Earth Sci.* **56**, 703–709 (2013).
23. P. E. Playford, R. N. Cope, A. E. Cockbain, G. H. Low, D. C. Lowry, "Phanerozoic," in *Geology of Western Australia* (Geological Survey of Western Australia, memoir 2, 1975), pp. 223–433.
24. R. S. Nicoll, "Ordovician conodont distribution in selected petroleum exploration wells, Canning Basin, Western Australia," in *Australian Geological Survey Organisation Record* (AGSO, 1993), vol. 17, pp. 1–136.
25. C. H. Wellman, P. Steemans, M. Vecoli, "Palaeophytogeography of Ordovician–Silurian land plants," in *Geological Society, London, Memoirs* (Geological Society of London, 2013), vol. 38, pp. 461–476.
26. W. A. Taylor, P. K. Strother, *Rev. Palaeobot. Palynol.* **151**, 41–50 (2008).
27. P. K. Strother, S. Al-Hajri, A. Traverse, *Geology* **24**, 55–58 (1996).
28. C. H. Wellman, P. Steemans, M. A. Miller, *Rev. Palaeobot. Palynol.* **212**, 111–126 (2015).
29. M. Ghavidel-Syooki, *Rev. Palaeobot. Palynol.* **231**, 48–71 (2016).
30. A. R. Hemsley, *Biol. Rev. Camb. Philos. Soc.* **69**, 263–273 (1994).
31. P. K. Strother, W. A. Taylor, J. H. Beck, M. Vecoli, *Palynology* **41** (sup1), 57–68 (2017).
32. N. Navidi-Izad *et al*., *Palynology* **44**, 575–585 (2020).
33. N. S. Davies, M. R. Gibling, *Earth Sci. Rev.* **98**, 171–200 (2010).
34. D. Edwards, L. Cherns, J. A. Raven, *Palaeontology* **58**, 803–837 (2015).
35. C. H. Wellman, P. L. Osterloff, U. Mohiuddin, *Nature* **425**, 282–285 (2003).
36. L. S. Normore, L. M. Dent, "Petroleum source potential of the Ordovician Nambeet Formation, Canning Basin: Evidence from petroleum well Olympic 1," in *Geology of Western Australia* (Geological Survey of Western Australia, report 169, 2017); pp. 1–20.
37. A. J. Mory, "A review of mid-Carboniferous to Triassic stratigraphy, Canning Basin, Western Australia, in *Geology of Western Australia* (Geological Survey of Western Australia, report 107, 2010); pp. 1–130.


## ACKNOWLEDGMENTS

We thank L. van Maldegem (ANU) for drafting the figures. The prior transmission electron microscopy work of W. A. Taylor (University of Wisconsin, Eau Claire) has been essential in enabling the interpretation of sporoderm structure based solely on light microscopy. We thank the Executive Director, Geological Survey and Resource Strategy, Western Australian Department of Mines, Industry Regulation and Safety, for the loan of the Samphire Marsh 1 palynological slides. We also acknowledge the critical science infrastructure role of curating geological samples and data by the Western Australia (WA) government. This paper contributes to Geoscience Australia's research of the Ordovician in Barnicarndy 1, Canning Basin, by C.F. Last, we would like to acknowledge our late colleague, Gordon Wood, who first brought to our attention the potential existence of a lowermost Paleozoic spore record. **Funding:** None. **Author contributions:** C.F. made the initial discovery, including recognition and specimen acquisition, and prepared all figures. P.K.S. wrote the initial draft and was responsible for all photomicrography. Both authors were responsible for conception, review, and editing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** This report is based on material housed and curated by the WA Department of Mines, Industry Regulation and Safety, which may be accessed through the WAPIMS database https://wapims.dmp.wa.gov.au/wapims. One sample, slide S-05-77/2, is housed at the Paleobotany Laboratory at Weston Observatory of Boston College; contact P.K.S. for access. All data are available in the manuscript or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/798/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S3
Tables S1 and S2
References (*38*–*46*)

3 May 2021; accepted 17 June 2021
10.1126/science.abj2927

SUPERCONDUCTIVITY

# Multicomponent superconducting order parameter in $UTe_2$

I. M. Hayes[1]†, D. S. Wei[2,3]†, T. Metz[1], J. Zhang[4], Y. S. Eo[1], S. Ran[1,5], S. R. Saha[1,5], J. Collini[1], N. P. Butch[1,5], D. F. Agterberg[6], A. Kapitulnik[2,3,7,8]*, J. Paglione[1,5,9]*

**An unconventional superconducting state was recently discovered in uranium ditelluride ($UTe_2$), in which spin-triplet superconductivity emerges from the paramagnetic normal state of a heavy-fermion material. The coexistence of magnetic fluctuations and superconductivity, together with the crystal structure of this material, suggests that a distinctive set of symmetries, magnetic properties, and topology underlie the superconducting state. Here, we report observations of a nonzero polar Kerr effect and of two transitions in the specific heat upon entering the superconducting state, which together suggest that the superconductivity in $UTe_2$ is characterized by a two-component order parameter that breaks time-reversal symmetry. These data place constraints on the symmetries of the order parameter and inform the discussion on the presence of topological superconductivity in $UTe_2$.**

Unconventional superconductors can host topologically protected edge states provided that the right set of symmetries is broken at the superconducting transition temperature, $T_c$. The superconducting state of $UTe_2$ has attracted immense attention because several observations, including a temperature-independent nuclear magnetic resonance Knight shift (*1*), an anomalously large upper critical field ($H_{c2}$) (*1*, *2*), reentrant superconductivity at high fields (*3*), chiral behavior imaged by scanning tunneling microscopy (*4*), and a point-node gap structure (*5*), all point to an odd-parity, spin-triplet pairing state. However, the key question of whether time-reversal symmetry is broken remains open. A prior attempt to measure time-reversal symmetry breaking (TRSB) in $UTe_2$ by using muon spin relaxation was unsuccessful because of the presence of dynamic local magnetic fields (*6*). Furthermore, TRSB in $UTe_2$ seems unlikely as the irreducible point group ($D_{2h}$) representations of the orthorhombic crystal symmetry of $UTe_2$ are all one-dimensional (*7*). For a superconducting order parameter to break time-reversal symmetry, it needs to have two components with a relative phase. Owing to the presence of strong spin-orbit coupling in $UTe_2$, which can be inferred from the strong anisotropy in its magnetic susceptibility (*1*), this means that the two components must belong to different irreducible representations of $D_{2h}$. This would necessarily result in multiple superconducting transitions, a rare phenomenon exhibited by only three other systems: $UPt_3$, Th-doped $UBe_{13}$, and $PrOs_4Sb_{12}$ (*8*–*10*). In this study, we propose a multicomponent order parameter that is experimentally supported by measurements of TRSB in $UTe_2$, as well as two distinct phase transitions in specific heat measurements. Together, these experiments allow us to strongly constrain the symmetry classification of the order parameters to two candidates.

To test for possible TRSB in the superconducting state of $UTe_2$, we performed high-resolution polar Kerr effects (PKE) measurements using a zero-area Sagnac interferometer (*11*) (ZASI). In general, the generation of a Kerr effect arises from the unequal reflection (in both polarization and phase) of left and right circularly polarized light from a given material, resulting in reflected light relative to the incident light that is phase-shifted by a Kerr angle $\theta_K$. The Kerr effect is not sensitive to Meissner effects, which normally prevent the measurement of global magnetic effects, and is therefore an optimal probe of TRSB in a superconducting system. At the same time, probing the system at frequencies ($\omega$) much larger than the superconducting gap energy ($\Delta$) will reduce a typical ferromagnetic-like signal of order ~1 rad by a factor of $(\Delta/\hbar\omega)^2 \sim 10^{-7}$, where $\hbar$ is the reduced Planck constant, yielding a typical theoretically predicted signal of about 0.1 to 1 μrad (*12*–*18*). However, owing to the high degree of common-mode rejection of the ZASI

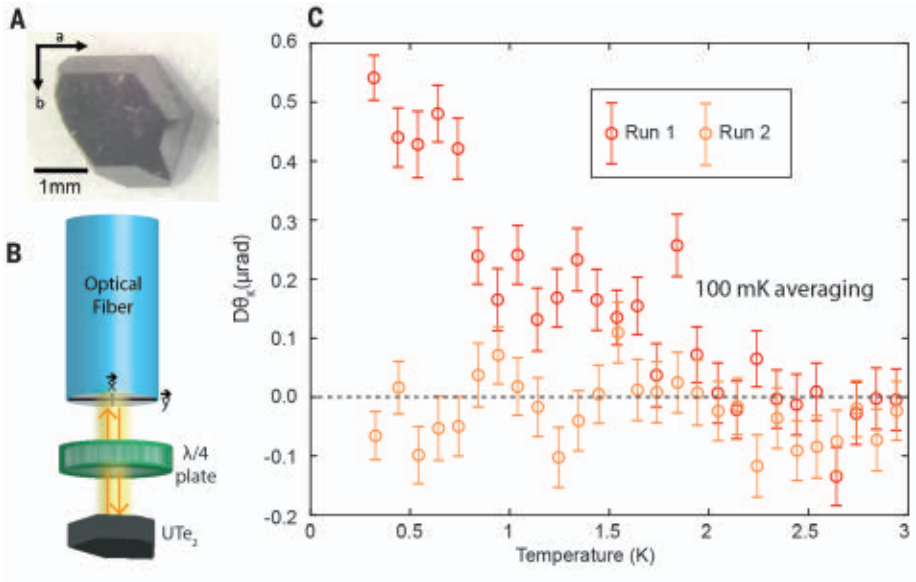



**Fig. 1. Polar Kerr angle evolution across the superconducting transition temperature in $UTe_2$.** (**A**) Optical image of the $UTe_2$ single crystal used in this work. (**B**) Schematic of the Sagnac interferometer used to measure the polar Kerr angle. The two orthogonal axes of the fiber compose the arms of the interferometer. Light from one axis is converted to circularly polarized light at the quarter-wave plate, reflected off the sample, converted back to linearly polarized light at the quarter-wave plate, and then transmitted into the axis orthogonal to the one from which is originated (focusing lenses are omitted for clarity). The light is reflected off the *a*-*b* plane. (**C**) Kerr angle plotted as a function of increasing temperature, after the $UTe_2$ single crystal was cooled through the superconducting transition temperature (1.6 K) with zero applied magnetic field. Error bars represent statistical error of hundreds of data points averaged together over 100-mK range bins (*28*). Two separate runs are shown. Run 1 shows no change in Kerr angle as the sample is cooled through $T_c$. Run 2 shows an increase in the Kerr angle around $T_c$, saturating at ~500 nrad.

[1]Department of Physics, Quantum Materials Center, University of Maryland, College Park, MD 20742, USA. [2]Geballe Laboratory for Advanced Materials, Stanford University, Stanford, CA 94305, USA. [3]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA. [4]State Key Laboratory of Surface Physics, Department of Physics, Fudan University, Shanghai 200433, China. [5]NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. [6]Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA. [7]Department of Physics, Stanford University, Stanford, CA 94305, USA. [8]Stanford Institute for Materials and Energy Sciences (SIMES), SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA. [9]The Canadian Institute for Advanced Research, Toronto, Ontario, Canada.
***Corresponding author. Email: paglione@umd.edu (J.P.); aharonk@stanford.edu (A.K.)**
†These authors contributed equally to this work.

for any reciprocal effects (e.g., linear birefringence, optical activity, etc.), we can detect these small signals.

The design and operation of our ZASI interferometer are detailed in (*11*, *19*), and the basic operation is as follows: Polar Kerr measurements are performed with 1550-nm wavelength light (20-μW incident power) that is polarized and then directed into a two-axes polarization-maintaining optical fiber that threads down into a $^3$He cryostat until it reaches our $UTe_2$ sample, which is mounted onto a copper stage thermally anchored to the cold finger of the cryostat. There, a beam along each axis is reflected off the $UTe_2$ crystal face (incident on the *a*-*b* plane of the crystal) and launched back up the opposite axis of the fiber. The two reflected beams, which have both traveled along identical paths (passing through the same optical plates, lenses, and fiber axes), compose the arms of the Sagnac interferometer. Any relative phase shift between the two beams must arise from encountering the sample and are revealed upon interference with one another to produce a signal from which the Kerr angle rotation can be extracted. The $UTe_2$ single crystal used in this study and a basic schematic of the setup are shown in Fig. 1, A and B. Previously, this technique has been used to confirm TRSB in $Sr_2RuO_4$ (*20*), with a low-temperature saturation value of the Kerr effect of ~0.1 μrad. The heavy-fermion uranium-based superconductors superconductors $UPt_3$ (*21*) and $URu_2Si_2$ (*22*), and the filled-skutterdite $PrOs_4Sb_{12}$ (*23*), gave a larger signal of ~0.4 to 0.7 μrad, which is expected owing to their strong spin-orbit interaction. Crucially, testing the apparatus with reciprocal reflecting media such as simple conventional superconductors, gold mirrors, and the spin-singlet d-wave heavy-fermion compound $CeCoIn_5$ (*24*), have yielded an expected null result.

To begin, we report the results of polar Kerr measurements performed at low temperatures on a single crystal of $UTe_2$. The sample was first cooled below the $T_c$ of $UTe_2$ (~1.6 K) in ambient magnetic field ($H_{ext}$ <0.3Oe), and the Kerr angle was subsequently measured as the sample was warmed above $T_c$. We found a small (~400 nrad at 300 mK), field-trainable Kerr effect that onsets near $T_c$ ~ 1.6 K, which is consistent with a TRSB superconducting order parameter; Fig. 1C shows two runs performed identically in this manner. Whereas Run 1 shows a signal emerging around $T_c$, and saturating at ~500 nrad, Run 2 shows no discernible signal. This indicates that without an applied field, domains are formed in the sample that can orient in opposite directions and give a finite signal or no signal at all, with an average signal of zero and a standard deviation dependent on the domain-to-beam size (10 μm) (*11*). Similar results were found for a second $UTe_2$ crystal, from a separate growth batch (fig. S1A). The detection of a finite positive Kerr signal indicates that a spontaneously large domain forms upon cooling the sample, owing to TRSB in $UTe_2$.

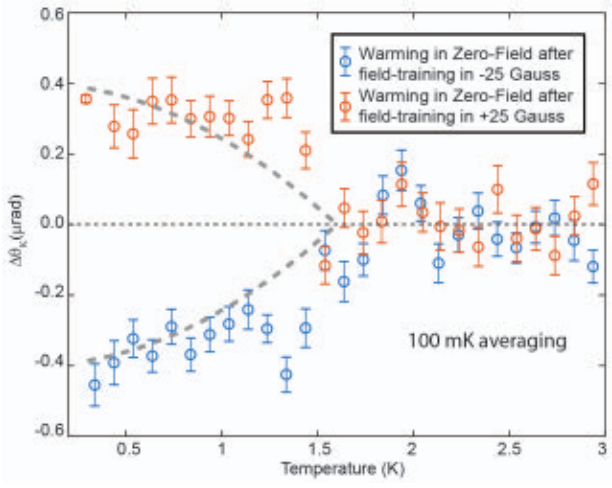



**Fig. 2. Magnetic field training of the Kerr effect.** Kerr angle for two different runs in which the sample is warmed up past $T_c$ after being cooled in an applied field. For a positive (negative) applied field of +25 G (−25 G), a positive (negative) Kerr signal emerges at $T_c$ and saturates around ~400 nrad. Dashed lines are guides to the eye to indicate where the onset of the Kerr signal begins. The actual temperature dependence of the Kerr signal is expected to be more complicated owing to coupling of the superconducting order parameter to the magnetic fluctuations.

To orient all of the domains in one direction, the sample was cooled through $T_c$ in a small applied field of +25 G. Experimentally, the magnitude of this training field has been found to be on the order of the lower superconducting critical field, $H_{c1}$ (*21*). Once the sample reaches base temperature (~300 mK), the external field is removed, and the Kerr angle is measured as the sample is warmed slowly up past $T_c$. Figure 2 shows a positive finite Kerr value develop around $T_c$ in this zero-field measurement. The sign of $\theta_K$ is reversed with a negative training field (−25 G), indicating that the broken time-reversal symmetry shares the same symmetry as that of a magnetic moment. This is because trainability with field implies a linear coupling between the field and broken time-reversal symmetric order parameter.

The magnitude of the trained signal should indicate the maximum signal size when all of the domains are aligned in the same direction. Depending on the size of the domains as the sample is cooled through $T_c$, the signal can be very small, or close to the full signal that can be achieved with field-training field [see, e.g., data on $UPt_3$ in (*21*)]. This may depend on the size of the sample, its purity and thus ability to pin domains, and the size of the probing beam. Additionally, small differences between signals in different runs may arise extrinsically owing to a change in background noise (from optical and electronic components and equipment) between runs. The measured Kerr signal was found to be independent of the magnitude of a small, applied training field (fig. S1B), indicating that the signal does not arise from magnetic vortices in the sample.

As discussed above, a TRSB order parameter in $UTe_2$ can only be built out of two components belonging to two different irreducible representations (*25*). This makes $UTe_2$ different from $LaNiGa_2$, which also hosts TRSB superconductivity that develops out of a paramagnetic state (*26*). In that material, which also has a $D_{2h}$ point group, a multidimensional representation can be obtained by including the spin degree of freedom, an option that is precluded in $UTe_2$ owing to spin-orbit coupling. It is exceptionally rare for a system to support superconducting transitions in two symmetry channels, which might caution against our interpretation. However, we find direct evidence for the existence of two superconducting order parameters in the specific heat of $UTe_2$. Normally, superconducting states are identified by resistivity or magnetization measurements, but neighboring superconducting states would both show zero resistance and diamagnetism. For this reason, specific heat measurements have played a central role in identifying all previous examples of superconductors with multicomponent order parameters (*8*–*10*).

The specific heat at zero field was first measured by means of the small-pulse method with $\Delta T$ = 0.5 to 1% (*27*). Four single crystals were measured from two growth batches (*28*). Figure 3 shows $C_p/T$ near the superconducting transition for these four samples. In each case, there is a shoulder-like feature at a temperature about 75 to 100 mK above the peak in $C_p/T$. This feature is quite sharp and divides the jump in the specific heat into two local maxima in the derivative, $d(C_p/T)/dT$, representing two thermodynamic anomalies. The existence of two transitions seems to be stable to whatever perturbations are responsible for the notable difference in $T_c$ between sample S4 and samples S1, S2, and S3. An additional three samples from a third growth batch also show two transitions (fig. S4). The consistent splitting of $T_c$ across seven samples despite differences in growth conditions and

absolute value of $T_c$ supports our inference that this splitting is intrinsic to $UTe_2$, reflecting the presence of a multicomponent order parameter, and is not an artifact of inclusions or intergrowths in these crystals. Furthermore, the inclusion scenario would not help explain the observed Kerr signal, because the order parameter must have two components locally for a Kerr effect to exist. This is an essential point; prior observations of split superconducting transitions were only accepted slowly by the community because of the possibility that the two observed transitions were the result of inhomogeneity in the crystals. However, both the specific heat data and the polar Kerr data point to a two-component order parameter in $UTe_2$, which makes the conclusion compelling in the present case. This is especially true given that these two measurements probe the material on different scales: a 10-μm spot on the surface in the case of the Kerr effect and the bulk of the crystal in the case of specific heat. Finally, recent work has shown that there are two well-separated transitions under pressure (*29*, *30*), which supports the idea that there are two nearly degenerate symmetry-breaking possibilities for a superconducting state in $UTe_2$. Both studies show that the splitting diminishes rapidly with decreasing pressure, with one study confirming the existence of splitting at ambient pressure (*30*), suggesting that the small splitting evident in our heat capacity experiments may be sensitive to sample quality and difficult to discern in other published work (*2*, *31*).

Beyond demonstrating that the order parameter in $UTe_2$ is two-component and TRSB, our data provide strong constraints on the particular irreducible representations to which the two components of the order parameter belong. Our observation that the TRSB in $UTe_2$ can be trained by a magnetic field along the crystallographic $c$ axis requires the presence of a term $\sim iH_c(\psi_1\psi_2^* - \psi_1^*\psi_2)$ in the free energy, where $H_c$ is the $c$-axis component of the magnetic field and $\psi_1$, $\psi_2$ are the two components of the superconducting order parameter. Symmetry requires that this term exists for only four possibilities (*28*):

(i) $\psi_1 \in B_{3u}$ and $\psi_2 \in B_{2u}$
(ii) $\psi_1 \in B_{1u}$ and $\psi_2 \in A_u$
(iii) $\psi_1 \in B_{3g}$ and $\psi_2 \in B_{2g}$
(iv) $\psi_1 \in B_{1g}$ and $\psi_2 \in A_{1g}$

using the notation for irreducible representations adopted in (*5*). However, because $UTe_2$ is known to be a spin-triplet superconductor, we can narrow the possibilities to the first two: $B_{3u}$ and $B_{2u}$, or $B_{1u}$ and $A_u$.

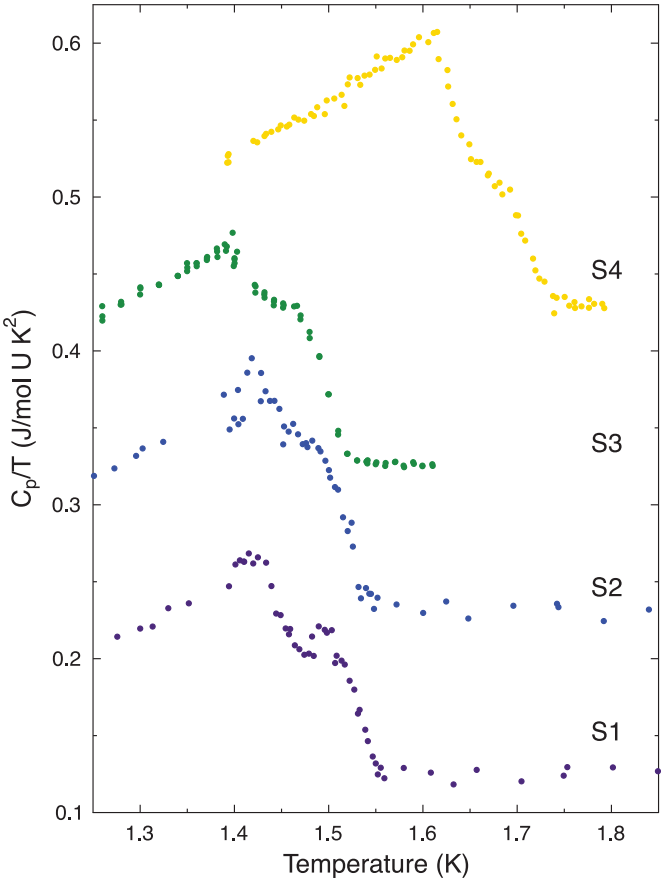



**Fig. 3. Superconducting transitions of $UTe_2$ in specific heat.** The specific heat over temperature per mole of uranium is plotted versus temperature for four samples of $UTe_2$. The $y$ axis is accurate for sample S1, whereas the curves for the other three samples have been offset in increments of 100 mJ/mol K$^2$. Each sample shows two anomalies, separated by ~80 mK, indicating the presence of two superconducting transitions. Samples S1 to S3 come from one growth batch, whereas S4 comes from another [see fig. S4 (*28*) for further details and data on three additional samples].

This picture can be checked by studying the two superconducting transitions as a function of magnetic field. The symmetry considerations that suggest that TRSB can be trained with a field applied along the $c$ axis imply that the lower-temperature transition should broaden and vanish with increasing field applied along that direction. This follows from the linear coupling between the two order parameters created by the term $\sim iH_c(\psi_1\psi_2^* - \psi_1^*\psi_2)$ when a $c$-axis field is present. Symmetry considerations preclude the existence of terms like these for fields along the $a$ and $b$ axes. Therefore, we expect the two transitions to remain distinct when a magnetic field is applied along those axes but not when a field is applied along the $c$ axis.

The field dependence of the split $T_c$ transition was measured on samples S1 and S2 by a large-pulse method (*28*) using oriented fields up to 5 T in a vector magnet (Fig. 4). The crystal orientation was determined by measuring the field angle–dependent $T_c$ for a field of 2 T. For fields oriented along the $a$ and $b$ axes, there are two discernable features at all fields. For fields along the $c$ axis, the two transitions broaden in field so that the splitting is no longer discernable above ~2 T, consistent with what we expected on the trainability of the TRSB with a $c$-axis field. Both the trainability of the Kerr signal and the magnetic field dependence of the specific heat point to the same symmetry classification of the order parameters, which confirms that the two phenomena are connected and intrinsic to $UTe_2$.

We thus arrive at a consistent picture of a superconducting state characterized by two order parameters that belong either to $B_{3u}$ and $B_{2u}$, or $B_{1u}$ and $A_u$, and that have a relative phase of $\pi/2$, leading to a TRSB state. Next, we turn to the question of the appearance of nodes. Thermal conductivity and penetration depth measurements suggest point node excitations (*5*). These have been interpreted as originating from the symmetry-required points nodes of a time-reversal invariant odd-parity state (*5*). Naïvely, TRSB gaps these point nodes, leading to a fully gapped state that appears inconsistent with these experiments (*5*). However, topological arguments allow for the possibility of Weyl point nodes (*32*–*34*), which we consider in more detail below. Weyl nodes are topologically protected by an integer Chern number, $Z$, and have associated surface Fermi arc states (*32*–*34*). For an odd-parity superconducting state in a Kramer's doubly degenerate pseudo-spin band, the single-particle gaps in the quasi-particle spectrum for the two pseudo-spin species are in general different and are given by (*25*)

$$E_{+/-} = \sqrt{(\epsilon(k) - \mu)^2 + |\boldsymbol{d}(k)|^2 \pm |\boldsymbol{q}(k)|^2}$$

Here the gap function is $\Delta(k) = [\boldsymbol{d}(k) \cdot \boldsymbol{\sigma}](i\sigma_y)$ (with Pauli matrices acting in pseudo-spin

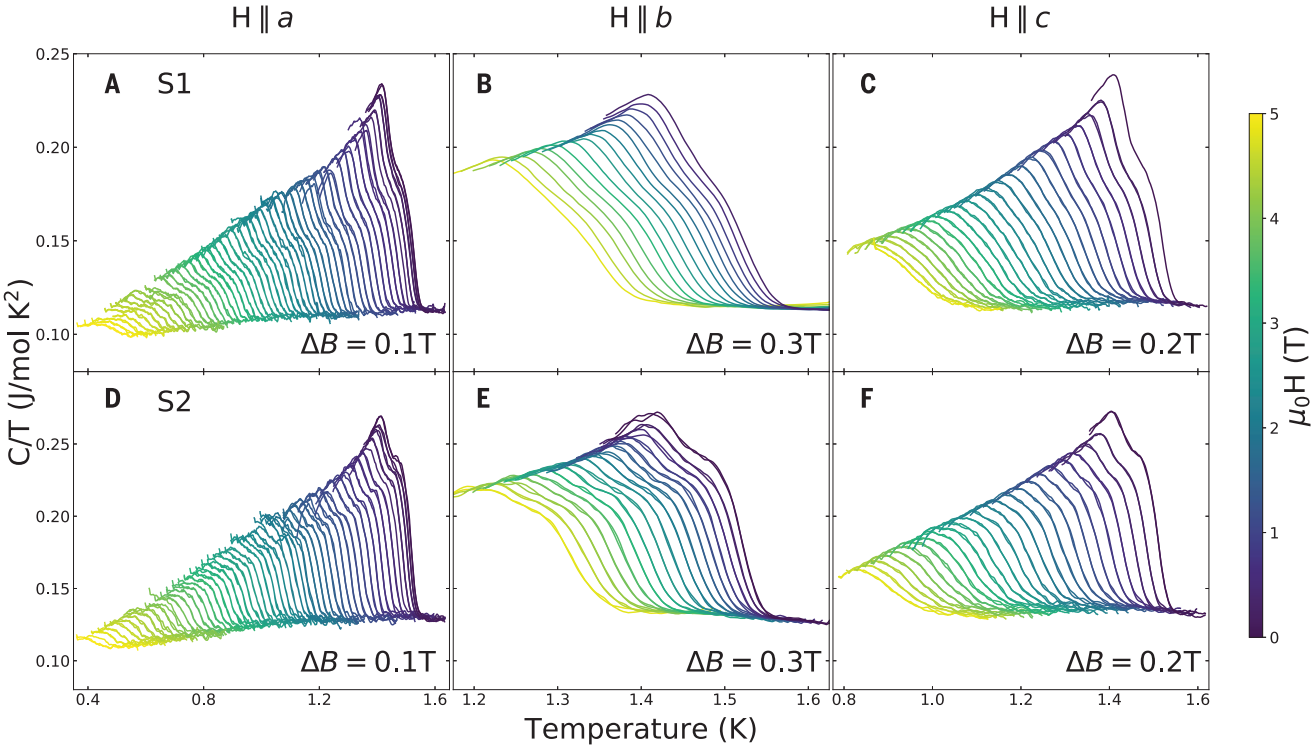



**Fig. 4. Magnetic field evolution of the split superconducting transition of $UTe_2$.** (**A** to **F**) For samples S1 and S2, specific heat was measured by a long-pulse method (see text for details) at every 100 or 300 mT along each of the crystallographic axes. Each panel corresponds to one field orientation for one of the samples and shows $C_p/T(T)$ curves gathered at each magnetic field. The curves have not been offset. When the field is oriented along the crystallographic *c* axis, the two transitions are indistinguishable above ~2 T, consistent with the linear field coupling to the product of the order parameters implied by Kerr data. However, for the other field orientations, two features are visible up to much higher magnetic fields.

space), and $\boldsymbol{q}(k) = \boldsymbol{d}(k) \times \boldsymbol{d}^*(k)$ denotes the nonunitary part that naturally arises when time-reversal symmetry is broken. In the two possibilities discussed above, the gap function is given by $\boldsymbol{d} = \boldsymbol{d}_1 + i\boldsymbol{d}_2$, where $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are both real. This gap function then gives rise to the single-particle gaps

$$\begin{aligned}|\Delta_\pm|^2 &= |\boldsymbol{d}(k)|^2 \pm |\boldsymbol{q}(k)|^2 \\ &= |\boldsymbol{d}_1|^2 + |\boldsymbol{d}_2|^2 \pm 2|\boldsymbol{d}_1 \times \boldsymbol{d}_2| \\ &= |\boldsymbol{d}_1|^2 + |\boldsymbol{d}_2|^2 \pm 2 \sin\theta |\boldsymbol{d}_1||\boldsymbol{d}_2|\end{aligned}$$

where θ is the angle between $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$. Nodes can only appear in $\Delta_-$, and for this to occur, two conditions must be met: (i) $\boldsymbol{d}_1 \cdot \boldsymbol{d}_2 = 0$ ($\sin\theta = 1$) and (ii) $|\boldsymbol{d}_1| = |\boldsymbol{d}_2|$. In general, these two conditions will be satisfied on a line in momentum space. If this line intersects the Fermi surface, there will be a Weyl point. If it does not, the superconductor will be fully gapped. Consequently, for the two gap structures discussed above, Weyl nodes can occur at arbitrary momenta on the Fermi surface.

Although the above argument reveals that Weyl points can generically occur, it does not guarantee that they exist in $UTe_2$. Surprisingly, it can be shown that Weyl points are expected for the $B_{2u} + iB_{3u}$ state. Given the current lack of detailed understanding of the electronic structure that gives rise to superconductivity, it is not possible to precisely identify the momentum dependence of the gap structure. However, symmetry places constraints on this. The symmetry-dictated form of the of the corresponding gap functions are $\boldsymbol{d}_{B2u} = f_{z,2}(\boldsymbol{k})\hat{x} + f_{u,2}(\boldsymbol{k})\hat{y} + f_{x,2}(\boldsymbol{k})\hat{z}$ and $\boldsymbol{d}_{B3u} = f_{u,3}(\boldsymbol{k})\hat{x} + f_{z,3}(\boldsymbol{k})\hat{y} + f_{y,3}(\boldsymbol{k})\hat{z}$, where the unknown functions $f_{x,i}, f_{y,i}, f_{z,i}, f_{u,i}$ share the same symmetry properties as $k_x, k_y, k_z, k_xk_yk_z$. To find Weyl points for such a gap, it is helpful to use insight for the Weyl semimetal $MoTe_2$, in which it was found that Weyl points appear in mirror planes (*35*). In particular, consider $k_x = 0$, then $f_{x,i} = f_{u,i} = 0$ by symmetry. This immediately implies $\boldsymbol{d}_{B2u} \cdot \boldsymbol{d}_{B3u} = 0$ in this mirror plane. Furthermore, the nodal condition $|\boldsymbol{d}_{B2u}| = |\boldsymbol{d}_{B3u}|$ implies that $\tilde{g}_z \equiv {f_{z,2}}^2 - {f_{z,3}}^2 = {f_{y,3}}^2$. It is also possible to carry out a similar analysis for the mirror plane $k_y = 0$, for which $\boldsymbol{d}_{B2u} \cdot \boldsymbol{d}_{B3u} = 0$ is also satisfied. In this case, $|\boldsymbol{d}_{B2u}| = |\boldsymbol{d}_{B3u}|$ implies $-\tilde{g}_z = {f_{x,2}}^2$. The relative minus sign in the two expressions $\tilde{g}_z = {f_{y,3}}^2$ ($k_x = 0$) and $-\tilde{g}_z = {f_{x,2}}^2$ ($k_y = 0$) ensures that the Weyl point will exist, provided that the cross section of the Fermi surface in both the $k_x$ and $k_y = 0$ planes has a circular topology (to see this, note that $\tilde{g}_z = 0$ for $k_z = 0$, $f_{x,i} = 0$, $k_x = 0$, and $f_{y,i} = 0$ for $k_y = 0$ so that one of these two expressions must be satisfied somewhere on a closed Fermi surface encircling the origin in the $k_z - k_x$ or $k_z - k_y$ planes). Recent angle-resolved photoemission experiments suggest that such a Fermi surface exists (*36*), revealing a likely *f*-electron–derived Fermi surface surrounding the *Z* point in the Brillouin zone. Consequently, this state will give rise to at least four Weyl points in either the $k_x = 0$ or the $k_y = 0$ planes. A similar analysis for the $A_u + iB_{1u}$ state (*28*) shows that, in this case, Weyl nodes are not guaranteed but are likely.

The above considerations, together with prior evidence for point nodes in $UTe_2$ from thermal conductivity and penetration-depth measurements (*5*), support the conclusion that $UTe_2$ is a Weyl superconductor. The Weyl points in the superconducting gap would give rise to surface Fermi arc states, providing a natural explanation for the observation of chiral surface states (*4*). These findings suggest the possibility of using $UTe_2$ for topological quantum computing, as well as for the exploration of superconducting analogs to phenomena in Weyl semimetals, including Fermi arcs and unusual Hall effects (*37*).

## REFERENCES AND NOTES

1. S. Ran *et al.*, *Science* **365**, 684–687 (2019).
2. D. Aoki *et al.*, *J. Phys. Soc. Jpn.* **88**, 043702 (2019).
3. S. Ran *et al.*, *Nat. Phys.* **15**, 1250–1254 (2019).
4. L. Jiao *et al.*, *Nature* **579**, 523–527 (2020).

5. T. Metz *et al*., *Phys. Rev. B* **100**, 220504 (2019).
6. S. Sundar *et al*., *Phys. Rev. B* **100**, 140502 (2019).
7. V. Hutanu *et al*., *Acta Crystallogr. B* **76**, 137 (2020).
8. R. Joynt, L. Taillefer, *Rev. Mod. Phys.* **74**, 235–294 (2002).
9. H. R. Ott, H. Rudigier, Z. Fisk, J. L. Smith, *Phys. Rev. B Condens. Matter* **31**, 1651–1653 (1985).
10. M. B. Maple *et al*., *J. Phys. Soc. Jpn.* **71**, 23–28 (2002).
11. J. Xia, P. T. Beyersdorf, M. M. Fejer, A. Kapitulnik, *Appl. Phys. Lett.* **89**, 062508 (2006).
12. J. Goryo, *Phys. Rev. B Condens. Matter Mater. Phys.* **78**, 060501 (2008).
13. R. M. Lutchyn, P. Nagornykh, V. M. Yakovenko, *Phys. Rev. B Condens. Matter Mater. Phys.* **80**, 104508 (2009).
14. E. Taylor, C. Kallin, *Phys. Rev. Lett.* **108**, 157001 (2012).
15. P. M. R. Brydon, D. S. L. Abergel, D. F. Agterberg, V. M. Yakovenko, *Phys. Rev. X* **9**, 031025 (2019).
16. E. J. König, A. Levchenko, *Phys. Rev. Lett.* **118**, 027001 (2017).
17. K. I. Wysokiński, J. F. Annett, B. L. Györffy, *J. Supercond. Nov. Magn.* **26**, 1909–1913 (2013).
18. M. Gradhand, K. I. Wysokinski, J. F. Annett, B. L. Györffy, *Phys. Rev. B Condens. Matter Mater. Phys.* **88**, 094504 (2013).
19. A. Kapitulnik, J. Xia, E. Schemm, A. Palevski, *New J. Phys.* **11**, 055060 (2009).
20. J. Xia, Y. Maeno, P. T. Beyersdorf, M. M. Fejer, A. Kapitulnik, *Phys. Rev. Lett.* **97**, 167002 (2006).
21. E. R. Schemm, W. J. Gannon, C. M. Wishne, W. P. Halperin, A. Kapitulnik, *Science* **345**, 190–193 (2014).
22. E. R. Schemm *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **91**, 140506 (2015).
23. E. M. Levenson-Falk, E. R. Schemm, Y. Aoki, M. B. Maple, A. Kapitulnik, *Phys. Rev. Lett.* **120**, 187004 (2018).
24. E. R. Schemm, E. M. Levenson-Falk, A. Kapitulnik, *Phys. C Supercond. Its Appl.* **535**, 13–19 (2017).
25. M. Sigrist, K. Ueda, *Rev. Mod. Phys.* **63**, 239–311 (1991).
26. A. D. Hillier, J. Quintanilla, B. Mazidian, J. F. Annett, R. Cywinski, *Phys. Rev. Lett.* **109**, 097001 (2012).
27. R. Bachmann *et al*., *Rev. Sci. Instrum.* **43**, 205–214 (1972).
28. See supplementary materials.
29. D. Braithwaite *et al*., *Commun. Phys.* **2**, 147 (2019).
30. S. M. Thomas *et al*., *Sci. Adv.* **6**, eabc8709 (2020).
31. L. P. Cairns, C. R. Stevens, C. D. O'Neill, A. Huxley, *J. Phys. Condens. Matter* **32**, 415602 (2020).
32. V. Kozii, J. W. F. Venderbos, L. Fu, *Sci. Adv.* **2**, e1601835 (2016).
33. K. Shiozaki, M. Sato, *Phys. Rev. B Condens. Matter Mater. Phys.* **90**, 165114 (2014).
34. T. Bzdušek, M. Sigrist, *Phys. Rev. B* **96**, 155105 (2017).
35. Z. Wang *et al*., *Phys. Rev. Lett.* **117**, 1–5 (2016).
36. L. Miao *et al*., *Phys. Rev. Lett.* **124**, 076401 (2020).
37. B. Yan, C. Felser, *Annu. Rev. Condens. Matter Phys.* **8**, 337–354 (2017).
38. I. M. Hayes *et al*., Replication Data for, "Multi-component Superconducting Order Parameter in $UTe_2$," Zenodo (2021); https://doi.org/10.5281/zenodo.4924576.



## ACKNOWLEDGMENTS

We thank E. Schemm, S. Tomarken, P. Brydon, T. Shishidou, and M. Weinert for useful discussions. **Funding:** Work at Stanford University was supported by the Department of Energy, Office of Basic Energy Sciences, under contract no. DE-AC02-76SF00515 and the Gordon and Betty Moore Foundation through Emergent Phenomena in Quantum Systems (EPiQS) Initiative Grant no. GBMF4529. Research at the University of Maryland was supported by the Air Force Office of Scientific Research Award no. FA9550-14-1-0332 (support of T.M.), the Department of Energy Award no. DE-SC-0019154 (specific heat experiments), the National Science Foundation Division of Materials Research Award DMR-1905891 (J.C.), the Gordon and Betty Moore Foundation's EPiQS Initiative through grant no. GBMF9071 (materials synthesis), NIST, and the Maryland Quantum Materials Center. S.R.S acknowledges support from the National Institute of Standards and Technology Cooperative Agreement 70NANB17H301. D.S.W. acknowledges support from the Karel Urbanek Fellowship in Applied Physics at Stanford University. **Author contributions:** D.S.W., I.M.H., A.K., and J.P. conceived and designed the experiments. S.R., S.R.S., and N.P.B. synthesized the $UTe_2$. S.R., S.R.S., and J.C. helped characterize the samples. D.S.W. and J.Z. performed the polar Kerr effect measurements. I.M.H., T.M., and Y.S.E. performed the specific heat measurements. D.F.A. provided the theoretical analysis. I.M.H., D.S.W., T.M., S.R., N.P.B., D.F.A., J.P., and A.K. analyzed the data and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data presented in this manuscript are available at Zenodo (*38*).

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/797/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Table S1
References (*39*–*42*)

5 February 2020; resubmitted 30 June 2020
Accepted 30 June 2021
Published online 15 July 2021
10.1126/science.abb0272


2D MATERIALS

# Boridene: Two-dimensional $Mo_{4/3}B_{2-x}$ with ordered metal vacancies obtained by chemical exfoliation


**Jie Zhou[1]*, Justinas Palisaitis[2], Joseph Halim[1], Martin Dahlqvist[1], Quanzheng Tao[1], Ingemar Persson[2], Lars Hultman[2], Per O. Å. Persson[2], Johanna Rosen[1]***



**Extensive research has been invested in two-dimensional (2D) materials, typically synthesized by exfoliation of van der Waals solids. One exception is MXenes, derived from the etching of constituent layers in transition metal carbides and nitrides. We report the experimental realization of boridene in the form of single-layer 2D molybdenum boride sheets with ordered metal vacancies, $Mo_{4/3}B_{2-x}T_z$ (where $T_z$ is fluorine, oxygen, or hydroxide surface terminations), produced by selective etching of aluminum and yttrium or scandium atoms from 3D in-plane chemically ordered $(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ in aqueous hydrofluoric acid. The discovery of a 2D transition metal boride suggests a wealth of future 2D materials that can be obtained through the chemical exfoliation of laminated compounds.**


Two-dimensional (2D) materials, such as graphene (*1*), transition metal chalcogenides (TMCs) (*2*), hexagonal boron nitride (BN) (*3*), and many hydroxides (*4*), have attracted extensive interest because of their intrinsic high surface area–to–volume ratio, electronic structure, and physicochemical properties (*2*, *5*, *6*). Further enrichment of the 2D materials family is sought after for the exploration of effects from quantum confinement and low dimensionality (*3*). Most 2D materials are produced by exfoliation of van der Waals solids (*5*, *7*), such as graphite, $MoS_2$, or phosphorous, cleaving laminated crystals into 2D counterparts. More recently, selective etching (also referred to as chemical exfoliation) has emerged as an alternative route for the preparation of 2D materials, such as the family of MXenes (*8*). MXenes are produced by selective removal of the A-layers, typically Al, in laminated $M_{n+1}AX_n$ (MAX) phases (where M indicates an early transition metal, A indicates an A-group element, and X indicates C and/or N) and related materials (*9*–*11*) and are commonly described by the generic formula $M_{n+1}X_nT_z$, where $T_z$ denotes surface terminations, typically -O, -OH, -F, and Cl (*8*, *12*, *13*).

The family of MXenes has been restricted to carbides and/or nitrides (*14*, *15*). Inspired by MXenes, a few attempts have been made to derive boron-based 2D materials, referred to as MBenes, through the selective etching of ternary, atomically laminated MAB phases ($M_{n+1}AlB_{2n}$ and $M_4AlB_4$, where n = 1 to 3) using HCl, HF, NaOH, and/or LiF-HCl etchants (*16*–*18*). However, the synthesis of individual single-layer sheets has, to date, not been realized. Instead, phase transformation, partial etching of Al, and dissolution of the 3D compounds are common obstacles. Dealloying the indium layer from $Ti_2InB_2$ yields a bulk layered TiB structure (*19*), short of a 2D material. Experimental realization of 2D hydrogen boride sheets (HB sheets) was recently reported (*20*), produced by ion-exchange reaction between protons and magnesium cations in magnesium diboride ($MgB_2$). The amorphous sheets, however, lack long-range order (*21*) and are not stable at ambient conditions. The growth of 2D boron sheets (borophene) on Ag (111) has been shown under ultrahigh vacuum conditions (*22*). Theoretical calculations on hypothetical 2D $Mo_2B_2$, $Fe_2B_2$, and $Mn_2B_2$ suggest high potential for catalysis, energy storage, and magnetism-based applications (*23*, *24*).

The discovery of in-plane chemically ordered MAX phase quaternaries, coined *i*-MAX (*25*–*27*), has enabled the synthesis of 2D MXenes with in-plane chemical or vacancy ordering, such as $Mo_{4/3}C$ and $W_{1.33}C$, showing promise for energy storage and catalysis (*25*, *26*, *28*). We theoretically predicted 15 ternary laminated boride


[1]Materials Design, Department of Physics, Chemistry and Biology (IFM), Linköping University, SE-581 83 Linköping, Sweden. [2]Thin Film Physics, Department of Physics, Chemistry and Biology (IFM), Linköping University, SE-581 83 Linköping, Sweden.
*Corresponding author. Email: jie.zhou@liu.se (J.Z.); johanna.rosen@liu.se (J.R.)

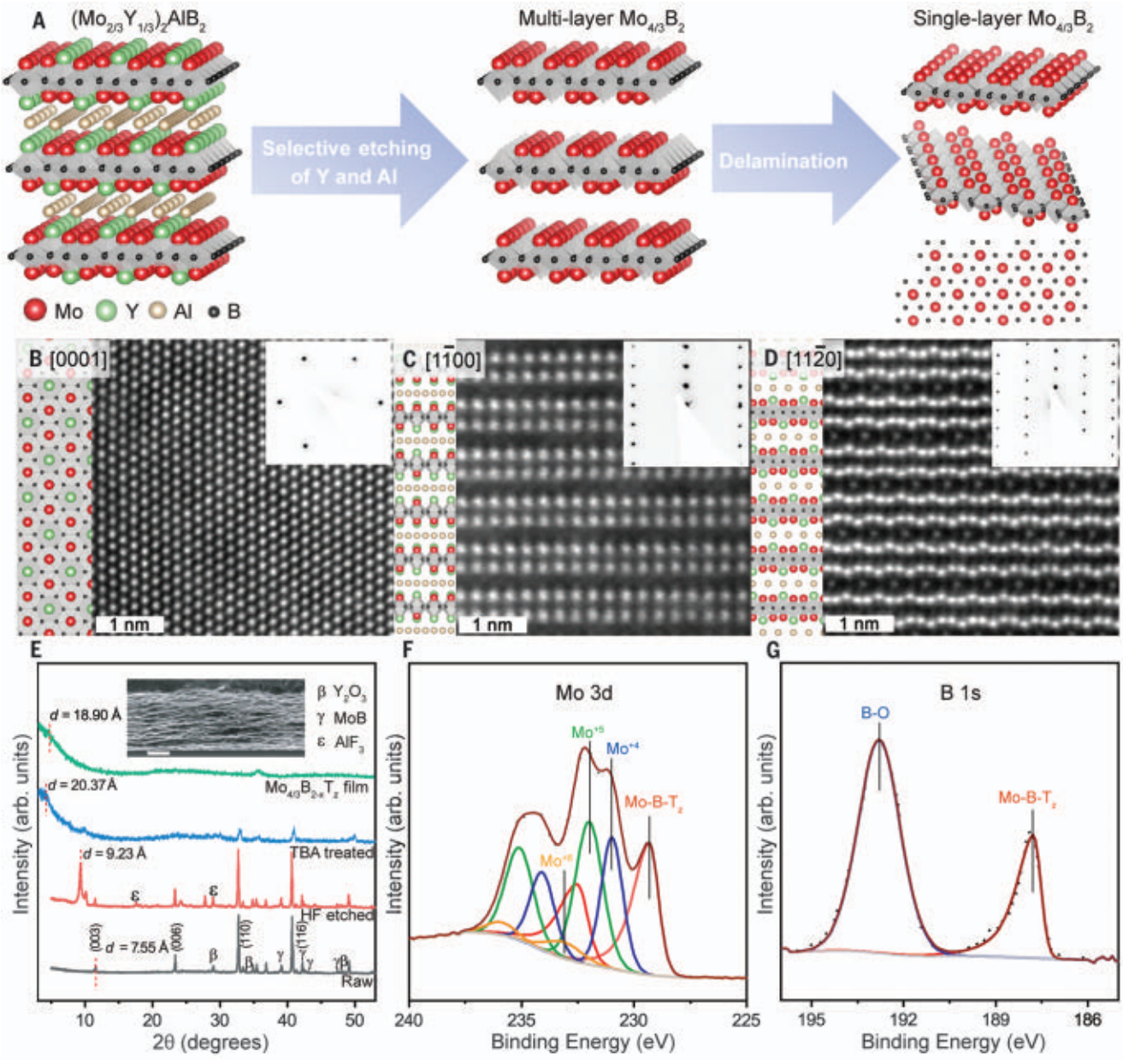



**Fig. 1. Synthesis and characterization of 3D $(Mo_{2/3}Y_{1/3})_2AlB_2$ and its 2D derivative $Mo_{4/3}B_{2-x}T_z$ (boridene).** (**A**) Schematic of the transformation process from 3D *i*-MAB to 2D boridene, with the schematic atomic structure of $(Mo_{2/3}Y_{1/3})_2AlB_2$ before etching (left), after etching (middle), and after delamination (right). The in-plane ordering of Mo and Y in $(Mo_{2/3}Y_{1/3})_2AlB_2$ gives rise to ordered vacancies upon the removal of Y after etching. (**B** to **D**) In-plane chemical ordering of the $(Mo_{2/3}Y_{1/3})_2AlB_2$ *i*-MAB phase evident from STEM images along the [0001] (B), [1$\bar{1}$00] (C), and [11$\bar{2}$0] (D) zone axes, with corresponding SAED patterns. The schematics to the left of each image represent the atomic arrangement in space group $R\bar{3}m$ (no. 166). (**E**) XRD pattern of $(Mo_{2/3}Y_{1/3})_2AlB_2$ before (black) and after (red) HF etching and after TBAOH intercalation (blue) and delamination (green). (Inset) SEM image showing the cross section of a $Mo_{4/3}B_{2-x}T_z$ film (scale bar, 2 μm). arb. units, arbitrary units. (**F** and **G**) High-resolution XPS spectra of a filtered film of boridene with peak fittings for the Mo 3d (F) and B 1s (G) regions.

phases with in-plane chemical ordering, of the general formula $(M'_{2/3}M''_{1/3})_2AlB_2$ (*29*), which we called *i*-MAB phases. The predictions were experimentally verified for $(Mo_{2/3}Sc_{1/3})_2AlB_2$ and $(Mo_{2/3}Y_{1/3})_2AlB_2$ (*29*).

We report 2D transition metal borides, $Mo_{4/3}B_{2-x}T_z$ (where $T_z$ denotes surface terminations -F, -O, or -OH), derived by selective etching of laminated *i*-MAB phases in concentrated aqueous hydrofluoric (HF) acid. We propose to name these phases boridene or MBene depending on the context in relation to graphene or MXene, respectively. Our etching applied to prepare boridene yields a 2D material with ordered vacancies on the metal sites, as inherited from the parent bulk phase.

$(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ compounds were prepared by a solid-state reaction of elemental powders (details are presented in the materials and methods section of the supplementary materials). Scanning electron microscopy (SEM) shows a layered crystal structure with chemical composition $(Mo_{2/3}Y_{1/3})_2AlB_2$ and with the atomic molar ratio of Mo:Y:Al confirmed by energy dispersive x-ray (EDX) analysis to be 1.3:0.7:1.0 (figs. S1 and S2). A schematic of the target material is shown in Fig. 1A (left panel). Scanning transmission electron microscopy (STEM) images of the $(Mo_{2/3}Y_{1/3})_2AlB_2$ *i*-MAB phase along the [0001], [1$\bar{1}$00], and [11$\bar{2}$0]

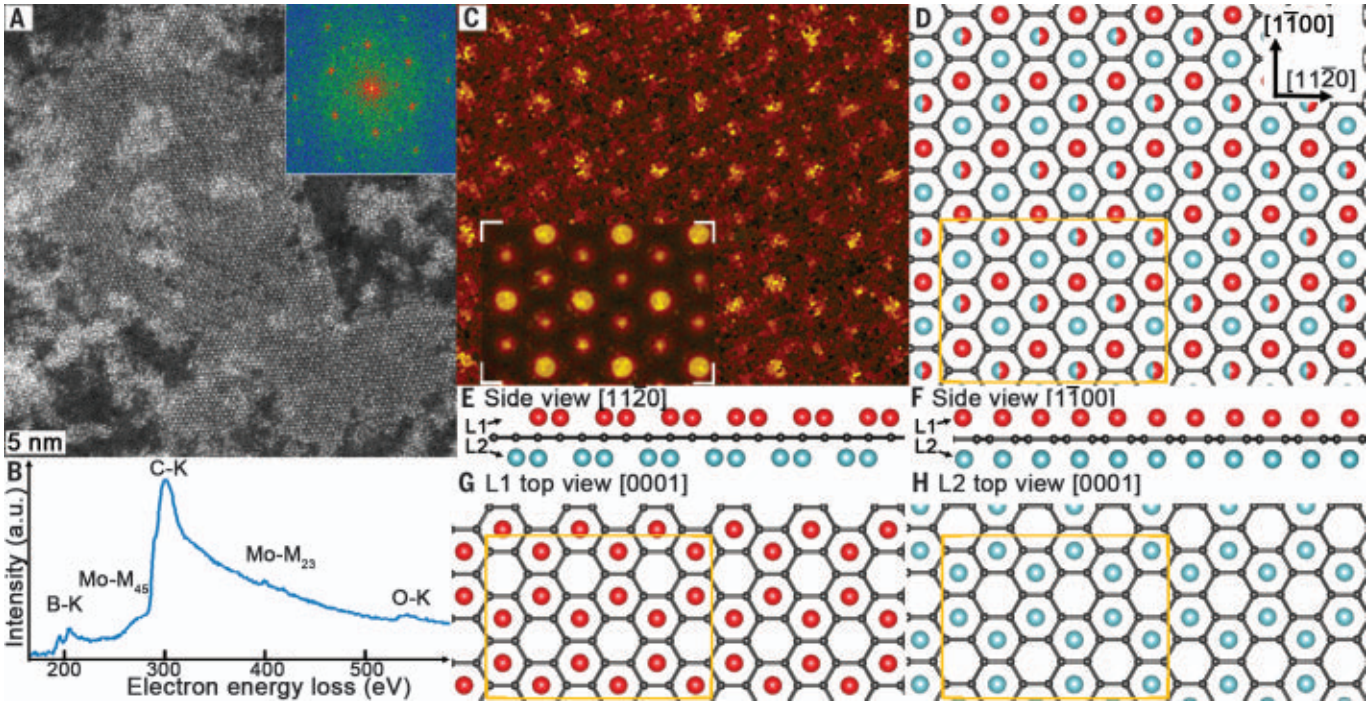



**Fig. 2. Structural characterization of the monolayer boridene (2D $Mo_{4/3}B_{2-x}T_z$ sheets).** (**A** and **B**) STEM image of single-layer $Mo_{4/3}B_{2-x}T_z$ sheet (dispersed on amorphous carbon support) (A) with corresponding FFT (inset) and core-loss EELS spectrum (B). a.u., arbitrary units. (**C**) Higher-magnification STEM image together with simulated STEM image (inset) based on an ideal (theoretical) sheet. (**D**) Top view of the ideal atomic structure of the single-layer $Mo_{4/3}B_2$ sheet from theoretically simulated parent phase, Mo, represented by red (top layer) and blue (bottom layer) atoms. (**E** and **F**) Side views of the ideal atomic structure of the single-layer sheet along the [11$\bar{2}$0] (E) and [1$\bar{1}$00] (F) zone axes. (**G** and **H**) Top views of the ideal atomic structure if only considering atoms of layer 1 (L1) and layer 2 (L2), respectively, with L1 and L2 depicted in (E) and (F).

zone axes are shown in Fig. 1, B to D, respectively. In these micrographs, the Mo and Y atoms appear brightest, the Al atoms are less bright, and B is too light to be detected under applied imaging conditions. The schematics to the left of each micrograph indicate the expected atomic arrangements based on the hexagonal $(Mo_{2/3}Y_{1/3})_2AlB_2$ in space group $R\bar{3}m$ (no. 166). From the top view in Fig. 1B, a close-packed (hexagonal) structure is evident. Along the [1$\bar{1}$00] zone axis shown in Fig. 1C, Mo and Y are partially overlapped and thus display an average contrast. However, when viewed from the [11$\bar{2}$0] zone axis (Fig. 1D), chemical ordering of Mo and Y in the *M* layer is revealed, with the Y atoms extending from the Mo layer toward the Al layer, consistent with theoretical predictions (*29*). Furthermore, the varied contrast shown within the Al layer is an indication of the formation of a Kagomé lattice, originating from the A-layer relaxation as a result of the extending Y atoms. The in-plane ordered *i*-MAB structure (*29*) is also revealed by the selected-area electron diffraction (SAED), shown in the insets of Fig. 1, B to D. The measured x-ray diffraction (XRD) pattern of a sample with nominal composition of $(Mo_{2/3}Y_{1/3})_2AlB_2$ is shown in Fig. 1E (black curve). The corresponding Rietveld refinement and refined parameters are shown in fig. S3 and table S1, respectively. The phase composition determined by the refinement is 80 wt % $(Mo_{2/3}Y_{1/3})_2AlB_2$, 15 wt % MoB, and 5 wt % $Y_2O_3$.

Chemical exfoliation (selective etching) of both $(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ was implemented using a concentrated aqueous HF solution. A schematic of the 3D-to-2D conversion is shown in Fig. 1A. In addition to removal of the Al layers, the in-plane ordering of Mo and Y in $(Mo_{2/3}Y_{1/3})_2AlB_2$ gives rise to ordered metal vacancies upon the additional removal of Y upon etching (see also schematic in fig. S4). This is concluded based on the following: When the $(Mo_{2/3}Y_{1/3})_2AlB_2$ powder was immersed in 40 wt % HF aqueous solution and stirred in an oil bath at 33° to 35°C, bubbles formed immediately, which indicates an exothermal reaction between the 3D boride and the HF. This resembles HF etching of MAX phases for MXene derivation, where it is assumed that evolved gas is $H_2$ (*8*). Compared with most MAX phases (*8*, *14*), the *i*-MAB phases herein are more reactive when etched in a HF solution, hence a reduced time is required (210 min in the present case). Further details of the etching process are given in the materials and methods section, and the optimization of the etching process is described in the supplementary text (figs. S5 and S6). The corresponding XRD result of the etched powder is shown in Fig. 1E (red curve), where the peak intensities originating from the $(Mo_{2/3}Y_{1/3})_2AlB_2$ precursor have substantially decreased after the HF treatment. The (000*l*) peak downshifts to a lower angle of $2\theta = 9.35°$, corresponding to an enlarged *d*-spacing value (*d*) of 9.23 Å, from the original 7.55 Å of $(Mo_{2/3}Y_{1/3})_2AlB_2$. A schematic showing the change of the *d* value during the etching process is provided in fig. S7. This additional low-angle (000*l*) peak is commonly observed in MXenes produced by selective etching of a MAX phase in HF (*8*, *14*). Moreover, a small amount of an etching by-product, $AlF_3$, is observed. The binary MoB and $Y_2O_3$ impurities present in the raw precursor sample did not react with HF, so the relative intensity of their peaks remains.

Intercalation of the etched multilayer crystals is performed using a dilute tetrabutylammonium hydroxide (TBAOH) aqueous solution and mild shaking. The *d* value is then further increased to 20.37 Å, as shown in Fig. 1E (blue curve). After the TBAOH treatment, the etched multilayer crystals delaminate spontaneously in water through mild shaking (by hand), and a colloidal suspension of the delaminated 2D sheets with a concentration of 4 mg/ml can

easily be reached. A film (inset in Fig. 1E) is then produced by vacuum filtration of the colloidal suspension through a nanoporous polypropylene membrane, and the corresponding XRD result is shown in Fig. 1E (topmost scan), displaying a broad (000$l$) peak with an angle of 2θ = 4.74°, corresponding to a $d$ value of 18.90 Å. A signal of the (000$l$) peak implies that the film is made of stacked 2D sheets only (fig. S8). EDX analysis of the sample before and after etching is presented in fig. S2, showing the presence of Mo, B, O, and F in the 2D material after etching and the effective removal of both Al and Y from the 3D precursor. The surface and cross-sectional SEM images and corresponding EDX analysis of the filtered film are provided in figs. S9 and S10, which confirm the complete conversion from 3D $(Mo_{2/3}Y_{1/3})_2AlB_2$ to 2D $Mo_{4/3}B_{2-x}$ during the etching process. The prevalence of concerned species is corroborated by x-ray photoelectron spectroscopy (XPS) analysis, presented in Fig. 1, F and G, and in figs. S11 and S12. The delaminated 2D material is therefore referred to as $Mo_{4/3}B_{2-x}T_z$.

An overview STEM image from a typical portion of an ~50-nm single layer $Mo_{4/3}B_{2-x}T_z$ boridene flake is shown in Fig. 2A. The corresponding fast Fourier transform (FFT) reveals a hexagonal symmetry, as inherited from the parent phase (Fig. 1B). Electron energy-loss spectroscopy (EELS) of the same flake shows the presence of Mo, B, O, and C (Fig. 2B), where the carbon signal originates from the amorphous carbon support. Notably, EELS confirms the absence of Y (fig. S13), whereas the presence of F at low concentrations cannot be ruled out but is below the detection limit because of low signal to noise (as a result of the comparably thick carbon support). The molar ratio of Mo:B:O is approximated to 24:28:48, which gives the formula $Mo_{4/3}B_{1.55}O_{2.66}$ for the 2D sheet. Based on the removal of Al and Y upon etching of the $(Mo_{2/3}Y_{1/3})_2AlB_2$ precursor established above, the ideal final 2D product without surface terminations would be $Mo_{4/3}B_2$ (see schematic in Fig. 1A). In addition to the suggested ordered vacancies formed upon removal of Y, the resulting formula also indicates a deficiency in the occupation of the B sites in the 2D sheets.

The atomically resolved structure (top view) of the boridene is shown in Fig. 2C, where two different column intensities can be seen. A corresponding schematic of the sheet structure, obtained from an ideal single 2D sheet after selective etching of Al and Y from $(Mo_{2/3}Y_{1/3})_2AlB_2$, is shown in Fig. 2D. The sheet consists of a top and bottom layer of Mo, separated by a layer of B. To differentiate the two Mo layers in the $Mo_{4/3}B_2$ sheet, the layers are marked as red [layer 1 (L1)] and blue [layer 2 (L2)], respectively. The atoms divided in red or blue represent the sheet columns containing two Mo atoms in the projected column, both in the top (red) and bottom (blue) layer. The sheet viewed from the [0001], [11$\bar{2}$0], and [1$\bar{1}$00] zone axes is shown in Fig. 2, D to F, and a top view of the separate L1 and L2 is illustrated in Fig. 2, G and H. Both L1 and L2 exhibit the same honeycomb structure with ordered Y vacancies. The layers are, however, mutually shifted in the ideal $Mo_{4/3}B_2$ sheet, which causes overlap between Mo atoms in every third column (indicated by a blue or red atom in Fig. 2D). Furthermore, because of this shift, the vacancies in the L1 and L2 layers overlap with Mo atoms in the corresponding layer (as shown in Fig. 2D and indicated in Fig. 2G). To verify this ideal sheet structure, a simulated STEM image is included as an inset in Fig. 2C. The match between experimental and proposed structures further supports formation of $Mo_{4/3}B_{2-x}T_z$ with ordered Y vacancies.

The average lateral size of the 2D boridene flakes is 50 nm, which is smaller than the size of typical MXene materials (*8*), for which a 3D grain size of >10 μm (of the precursor) is available. The parent 3D-boride phases we prepared, however have a smaller grain size (~1 μm), thus putting a practical limit to the boridene flake size. Furthermore, the ordered vacancies are accompanied by other defects, primarily point defects (Fig. 2A), which are known to affect the material performance. Defects may decrease the elasticity slightly but may also introduce additional electrochemically active sites (*30*, *31*). Related edge morphology analysis of a $Mo_{4/3}B_{2-x}T_z$ sheet is provided in fig. S14.

XPS measurements of the $(Mo_{2/3}Y_{1/3})_2AlB_2$ precursor and the filtered $Mo_{4/3}B_{2-x}T_z$ film help identify the chemical states [binding energies (BEs)] of those constituting elements. The XPS spectrum of the Mo 3d region (Fig. 1F) for $Mo_{4/3}B_{2-x}T_z$ film is fitted by four peaks. The first corresponds to Mo-B-$T_z$ (BE = 229.6 eV) species belonging to the $Mo_{4/3}B_{2-x}T_z$ sheets, whereas the other three correspond to oxidized states of $Mo^{+4}$, $Mo^{+5}$, and $Mo^{+6}$ originating from surface oxidation of the film after exposure to ambient atmosphere during drying. Figure 1G presents the XPS spectrum of the B 1s region for the film, where the spectrum is fitted by two species: Mo-B-$T_z$ (BE = 187.8 eV) belonging to $Mo_{4/3}B_{2-x}T_z$ sheets and the other species corresponding to $B_2O_3$ surface oxide. The binding energies for the Mo-B-$T_z$ species in the Mo 3d region show a shift to a higher value compared with that for the 3D compound $(Mo_{2/3}Y_{1/3})_2AlB_2$, whereas the Mo-B-$T_z$ species in the B 1s region shows a shift to a lower BE value (fig. S11). From the XPS measurements and analysis, we observed that, along with the complete removal of Al atoms when etching $(Mo_{2/3}Y_{1/3})_2AlB_2$ (fig. S11D), essentially all Y atoms are removed (fig. S11B). Furthermore, from the XPS spectra of O 1s and F 1s of the $Mo_{4/3}B_{2-x}T_z$ film (fig. S12, A and B), surface terminations $T_z$ were identified to be a mixture of -O, -OH, and -F. However, quantification of the elemental ratios in the 2D and 3D compounds is hindered by the presence of surface oxides (figs. S11 and S12 and tables S2 to S7). From the combined XPS and EELS analysis, we propose a first estimation of the boridene, $Mo_{4/3}B_{2-x}T_z$, composition, with x being up to ~0.5 and z being in the range of 2 to 3.

On the basis of the aforementioned results, we conclude that the relatively weakly bonded Al and Y atoms in $(Mo_{2/3}Y_{1/3})_2AlB_2$ were selectively etched when immersed in 40 wt % HF solution and that the surface of the remaining 2D Mo-B layer was functionalized by O, OH, and F surface terminations. The chemical reactions taking place during the etching process can be described as

$$(Mo_{2/3}Y_{1/3})_2AlB_2 + 5HF = AlF_3 + 2/3YF_3 + 5/2H_2 + Mo_{4/3}B_{2-x} \quad (1)$$

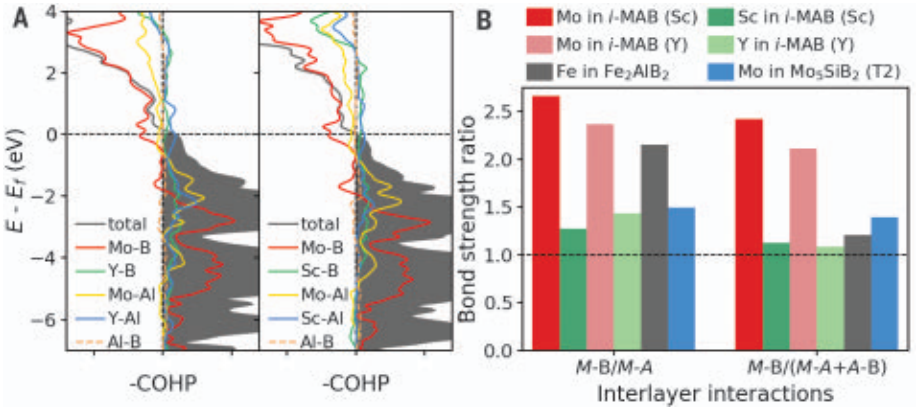



**Fig. 3. Theoretical bond analysis and interlayer interaction analysis for selected laminated borides.** (**A**) Calculated COHP curves for $(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ *i*-MAB phases. The filled regions represent occupied states, and the Fermi level is set to 0 eV. (**B**) Bonding strength ratios for various interactions in $(Mo_{2/3}Y_{1/3})_2AlB_2$, $(Mo_{2/3}Sc_{1/3})_2AlB_2$, $Fe_2AlB_2$, and $Mo_5SiB_2$.

$$Mo_{4/3}B_{2-x} + 2H_2O = Mo_{4/3}B_{2-x}O_2 + 2H_2 \quad (2)$$

$$Mo_{4/3}B_{2-x} + 2H_2O = Mo_{4/3}B_{2-x}(OH)_2 + H_2 \quad (3)$$

$$Mo_{4/3}B_{2-x} + 2HF = Mo_{4/3}B_{2-x}F_2 + H_2 \quad (4)$$

Derivation of 2D $Mo_{4/3}B_{2-x}T_z$ sheets can also be obtained from selective etching of the Al and Sc atoms from the parent 3D $(Mo_{2/3}Sc_{1/3})_2AlB_2$ *i*-MAB phase (fig. S15). XRD analysis (fig. S16) shows that the peak intensities originating from the $(Mo_{2/3}Sc_{1/3})_2AlB_2$ precursor have substantially decreased after HF treatment, and the (00l) peak downshifts to a lower angle of 2θ = 7.78° from the original 11.69° of $(Mo_{2/3}Sc_{1/3})_2AlB_2$. Moreover, a small amount of the etching by-product $ScF_3$ is observed. The binary $Mo_3Al_8$ impurities in the 3D sample are found to be dissolved in the HF solution, whereas the binary MoB impurities did not react with HF. After the TBAOH treatment, the etched multilayer crystals delaminate spontaneously in water through mild shaking, and a colloidal suspension of the delaminated 2D sheets is obtained (fig. S17). The EDX results of the sample before and after etching indicate complete removal of both Al and Sc, with the sheets being composed of Mo, B, and O, respectively (fig. S17). The initial optimization of the etching process and the characterization of the filtered film derived from $(Mo_{2/3}Sc_{1/3})_2AlB_2$ are provided in figs. S18 to S20.

To explain why it is possible to selectively etch Al, Y, and Sc from $(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ *i*-MAB phases, we compare the bond strength and interlayer interaction between the atomic layers. The chemical bonding was quantitatively analyzed on the basis of a crystal orbital Hamilton population (COHP) analysis. Details are provided in the supplementary materials. Figure 3A shows that the COHP curve for the total average interactions in $(Mo_{2/3}Y_{1/3})_2AlB_2$ and $(Mo_{2/3}Sc_{1/3})_2AlB_2$ is composed of occupied bonding states. Decomposing these curves into their partial COHP (pCOHP) shows that the Mo-B interaction also has populated antibonding states that could be improved if the Fermi level $E_f$ can be shifted to lower energy (all pCOHP curves are illustrated in figs. S21 and S22). The bond strength was quantified by integrating the occupied states of the pCOHP curves (IpCOHP), as illustrated in figs. S23 and S24. The strongest individual bonds in $(Mo_{2/3}Y_{1/3})_2AlB_2$ are the in-plane B-B interactions followed by in-plane Al-Al and out-of-plane Mo-B interactions. However, taking bonding coordination into consideration reveals that Mo-B interaction is dominating (fig. S23B). In an attempt to compare the interaction between atomic layers, we considered IpCOHP ratios for *M*-B to *M*-*A* bonds (left-hand bars in Fig. 3B). For $(Mo_{2/3}Y_{1/3})_2AlB_2$, the Mo-B interaction is 2.7 times as strong as that of Mo-Al, whereas the Y-B interaction is only 1.3 times as strong as that of Y-Al. Similar numbers were found for $(Mo_{2/3}Sc_{1/3})_2AlB_2$. This indicates that chemical exfoliation to remove Al is likely to also affect Y and Sc.

The selective etching of *i*-MAB phases is in contrast to failed attempts for traditional layered MAB phases such as $Fe_2AlB_2$ and $Mo_5SiB_2$. Our bond analysis for $Fe_2AlB_2$ gives a ratio of 2.1 for Fe-B/Fe-Al, whereas $Mo_5SiB_2$ reveals a ratio of 1.5 for Mo-B/Mo-Al (left-hand bars in Fig. 3B and figs. S25 to S28). However, when comparing IpCOHP ratios between different crystal structures, it is essential to include all contributing interactions. In a previous attempt to predict exfoliation (*32*), important contributions such as the Al-B interactions in $Fe_2AlB_2$ were neglected, even though that bond is quite strong (fig. S25). The right-hand bars in Fig. 3B represent *M*-B to *M*-*A*+*A*-B IpCOHP ratios (other ratios shown in fig. S29). For the two *i*-MAB phases, we found that Mo bonds substantially more strongly to B than to Al, whereas similar bond strengths are found for Y and Sc bonding to B or Al. For both $Fe_2AlB_2$ and $Mo_5SiB_2$, we found that *M* bonds as strongly to B as *M* to *A* combined with *A* to B. As a first approximation, this lack of diverging interlayer interaction may explain why selective etching of Al (and Y or Sc) from *i*-MAB phases is successful, whereas attempts made for etching Al in $Fe_2AlB_2$ and Si in $Mo_5SiB_2$ fails.

Boridene (2D molybdenum boride sheets) has been demonstrated in the form of free-standing sheets larger than 50 nm. It is realized by selectively etching the Y and Al atoms from an in-plane chemically ordered quaternary *i*-MAB phase $(Mo_{2/3}Y_{1/3})_2AlB_2$ compound in a HF solution. The boridene formula is $Mo_{4/3}B_{2-x}T_z$, with ordered vacancies on the metal sites. Compared with the parent phase, the sheets maybe slightly deficient in B, with x up to ~0.5. The surface terminations $T_z$ were identified to be a mixture of O, OH, and F, with z in the range of 2 to 3. The 2D material can be selectively prepared in multilayer form, as delaminated single-layer sheets in colloidal suspension or as an additive-free filtered film. The 2D $Mo_{4/3}B_{2-x}T_z$ sheets can be prepared by our top-down approach, realizing a suspension of high concentration where the sheets are stable and processable in an aqueous environment. We also showed that the same method can realize corresponding boridene by selective etching of Al and Sc from the *i*-MAB phase $(Mo_{2/3}Sc_{1/3})_2AlB_2$. Present and referenced theoretical predictions show corresponding promise for a wealth of boridenes (or MBenes, as also named).

## REFERENCES AND NOTES


1. K. S. Novoselov *et al*., *Science* **306**, 666–669 (2004).
2. Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, M. S. Strano, *Nat. Nanotechnol.* **7**, 699–712 (2012).
3. S. Z. Butler *et al*., *ACS Nano* **7**, 2898–2926 (2013).
4. R. Ma, T. Sasaki, *Adv. Mater.* **22**, 5082–5104 (2010).
5. J. N. Coleman *et al*., *Science* **331**, 568–571 (2011).
6. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162 (2009).
7. L. Li *et al*., *Nat. Nanotechnol.* **9**, 372–377 (2014).
8. M. Naguib *et al*., *Adv. Mater.* **23**, 4248–4253 (2011).
9. M. W. Barsoum, *Prog. Solid State Chem.* **28**, 201–281 (2000).
10. M. Sokol, V. Natu, S. Kota, M. W. Barsoum, *Trends Chem.* **1**, 210–223 (2019).
11. J. Zhou *et al*., *Angew. Chem. Int. Ed.* **55**, 5008–5013 (2016).
12. M. Li *et al*., *J. Am. Chem. Soc.* **141**, 4730–4737 (2019).
13. A. VahidMohammadi, J. Rosen, Y. Gogotsi, *Science* **372**, eabf1581 (2021).
14. M. Naguib *et al*., *ACS Nano* **6**, 1322–1331 (2012).
15. P. Urbankowski *et al*., *Nanoscale* **8**, 11385–11391 (2016).
16. L. T. Alameda, C. F. Holder, J. L. Fenton, R. E. Schaak, *Chem. Mater.* **29**, 8953–8957 (2017).
17. L. T. Alameda *et al*., *J. Am. Chem. Soc.* **141**, 10852–10861 (2019).
18. K. Kim, C. Chen, D. Nishio-Hamane, M. Okubo, A. Yamada, *Chem. Commun.* **55**, 9295–9298 (2019).
19. J. Wang *et al*., *Nat. Commun.* **10**, 2284 (2019).
20. H. Nishino *et al*., *J. Am. Chem. Soc.* **139**, 13761–13769 (2017).
21. R. Kawamura *et al*., *Nat. Commun.* **10**, 4880 (2019).
22. A. J. Mannix *et al*., *Science* **350**, 1513–1516 (2015).
23. Z. Jiang, P. Wang, X. Jiang, J. Zhao, *Nanoscale Horiz.* **3**, 335–341 (2018).
24. X. Guo *et al*., *Adv. Funct. Mater.* **31**, 2008056 (2021).
25. Q. Tao *et al*., *Nat. Commun.* **8**, 14949 (2017).
26. R. Meshkian *et al*., *Adv. Mater.* **30**, e1706409 (2018).
27. M. Dahlqvist *et al*., *Sci. Adv.* **3**, e1700642 (2017).
28. B. Ahmed, A. El Ghazaly, J. Rosen, *Adv. Funct. Mater.* **30**, 2000894 (2020).
29. M. Dahlqvist *et al*., *J. Am. Chem. Soc.* **142**, 18583–18591 (2020).
30. J. Xie *et al*., *Adv. Mater.* **25**, 5807–5813 (2013).
31. J. Bao *et al*., *Angew. Chem. Int. Ed.* **54**, 7399–7404 (2015).
32. M. Khazaei *et al*., *Nanoscale* **11**, 11305–11314 (2019).


## ACKNOWLEDGMENTS


**Funding:** J.R. acknowledges support from the Knut and Alice Wallenberg (KAW) Foundation for a fellowship/scholar grant and project funding (KAW 2020.0033) and from the Swedish Foundation for Strategic Research (SSF) for project funding (EM16-0004). The KAW Foundation is also acknowledged for support provided to the Linköping Electron Microscopy Laboratory. J.Z. acknowledges support from the National Natural Science Foundation of China (grant no. 21805295). P.O.Å.P. acknowledges SSF through the Research Infrastructure Fellow program no. RIF 14-0074. Support from the Swedish Government Strategic Research Area in Materials Science on Functional Materials at Linköping University (faculty grant SFO-Mat-LiU no. 2009 00971) is also acknowledged. The calculations were carried out using supercomputer resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre (NSC) and the PDC Center for High Performance Computing, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. **Author contributions:** J.R. and J.Z. initiated the study. Q.T. and J.Z. performed materials synthesis and materials analysis (EDX and XRD, including Rietveld refinement) under the supervision of J.R. J.H. performed the XPS measurements and analysis. M.D. performed the density functional theory analysis under the supervision of J.R. J.P. performed the STEM-SAED-EELS analysis, and I.P. performed the STEM image simulation under the supervision of P.O.Å.P. J.Z. wrote the manuscript with contributions from the other authors. All coauthors read and commented on successive drafts of the manuscript. **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.sciencemag.org/content/373/6556/801/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S29
Tables S1 to S7
References (*33*–*54*)


8 November 2020; accepted 6 July 2021
10.1126/science.abf6239

PALEONTOLOGY

# Lifetime mobility of an Arctic woolly mammoth

Matthew J. Wooller[1,2]*†, Clement Bataille[3,4]*†, Patrick Druckenmiller[5,6], Gregory M. Erickson[7], Pamela Groves[8], Norma Haubenstock[1], Timothy Howe[1], Johanna Irrgeher[9], Daniel Mann[8], Katherine Moon[10,11], Ben A. Potter[12], Thomas Prohaska[9], Jeffrey Rasic[13], Joshua Reuther[5], Beth Shapiro[10,11], Karen J. Spaleta[1], Amy D. Willis[14]

**Little is known about woolly mammoth (*Mammuthus primigenius*) mobility and range. Here we use high temporal resolution sequential analyses of strontium isotope ratios along an entire 1.7-meter-long tusk to reconstruct the movements of an Arctic woolly mammoth that lived 17,100 years ago, during the last ice age. We use an isotope-guided random walk approach to compare the tusk's strontium and oxygen isotope profiles to isotopic maps. Our modeling reveals patterns of movement across a geographically extensive range during the animal's ~28-year life span that varied with life stages. Maintenance of this level of mobility by megafaunal species such as mammoth would have been increasingly difficult as the ice age ended and the environment changed at high latitudes.**

The extent to which environmental changes and human activity altered woolly mammoth (*Mammuthus primigenius*) populations is key to understanding the causes of megafaunal mass extinctions after the last ice age (*1*–*4*). Because mammoths are extinct, however, we know very little about their natural history, including the size of their home ranges or lifetime movement (*2*, *4*). Movement patterns of extant elephantids (*5*) and Arctic herbivores such as caribou (*6*) include regular movement across lifetime ranges, and it has been assumed that Arctic woolly mammoths exhibited similar behaviors.

Sequential isotopic analyses of elephantid tusks have been valuable for reconstructing life history records, including mobility (*7*, *8*). As tusks grow, they continually incorporate ingested strontium (Sr), and the incremental record of strontium isotope ratios ($^{87}Sr/^{86}Sr$) in tusks and teeth can be used to investigate proboscidean movements (*9*, *10*). The $^{87}Sr/^{86}Sr$ ratios in soils and plants show strong and predictable geographic variations that primarily reflect the underlying bedrock geology and that change little at millennial time scales (*10*). As animals ingest local strontium, the bioavailable $^{87}Sr/^{86}Sr$ patterns on the landscape are reflected in their tissues and can be used to trace an animal's movement (*11*, *12*). Strontium isotope analyses can be augmented by stable oxygen isotope analyses to aid in determining the provenance of organisms (*8*, *12*). To date, studies of elephantid movement have either used bulk digestion of relatively large samples of tusks that average the isotopic variability associated with movements at low temporal resolution or have focused on incremental records of shorter duration found in teeth, short mandibular tusks, adolescent tusks, or tusk segments (*7*–*9*, *13*).

The unglaciated portion of northern Alaska has a long and well-preserved mammoth paleontological record, with woolly mammoths persisting in mainland Alaska until ~13,000 calibrated radiocarbon years before present (*4*, *14*) and more-recent survival on islands (*1*). Our aim is to use sequential isotopic analyses of one of these well-preserved remains, a tusk, to examine the life history of a woolly mammoth that was alive during the last ice age, when conditions likely favored woolly mammoth adaptations (*15*). We selected a tusk [University of Alaska Museum Earth Science (UAMES) collection catalog number 29496] from an animal that died above the Arctic Circle ~17,100 calibrated years before present (*14*). The well-preserved remains of our study specimen include both tusks (total length: ~2.4 m), fragments of the skull, and a complete mandible with well-preserved teeth (*14*). Genetic analyses of the specimen showed that it has a single copy of the X chromosome and is therefore male (*14*),

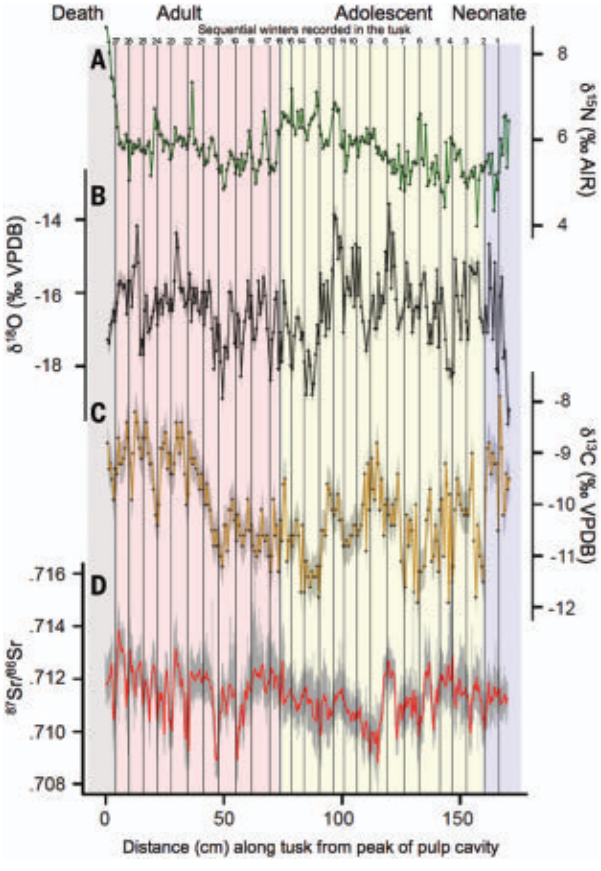



**Fig. 1. Sequential isotopic analyses along an entire ~1.7-m-long transect of a mammoth tusk from Arctic Alaska.** (**A**) Stable nitrogen, (**B**) oxygen, (**C**) carbon, and (**D**) strontium isotope values ($\delta^{15}N$, $\delta^{18}O$, $\delta^{13}C_{carbonate}$, and $^{87}Sr/^{86}Sr$ values, respectively). Vertical lines represent annual markers [peak winter (*14*)], and color shading corresponds to four main life periods (neonate, adolescent, adult, and end of life) (*14*). VPDB, Vienna Pee Dee Belemnite international standard; AIR, atmospheric nitrogen international standard.

[1]Alaska Stable Isotope Facility, University of Alaska Fairbanks, Fairbanks, AK, USA. [2]Department of Marine Biology, University of Alaska Fairbanks, Fairbanks, AK, USA. [3]Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, ON, Canada. [4]Department of Biology, University of Ottawa, Ottawa, ON, Canada. [5]University of Alaska Museum of the North, Fairbanks, AK, USA. [6]Department of Geosciences, University of Alaska Fairbanks, Fairbanks, AK, USA. [7]Department of Biological Science, Florida State University, Tallahassee, FL, USA. [8]Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK, USA. [9]Department of General, Analytical and Physical Chemistry, Montanuniversität Leoben, Leoben, Austria. [10]Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA, USA. [11]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA. [12]Arctic Studies Center, Liaocheng University, Liaocheng City, Shandong Province, China. [13]National Park Service, Fairbanks, AK, USA. [14]Department of Biostatistics, University of Washington, Seattle, WA, USA.
***Corresponding author. Email: mjwooller@alaska.edu (M.J.W.); cbataill@uottawa.ca (C.B.)**
†These authors contributed equally to this work.

and analyses of its complete mitochondrial genome placed it within mitochondrial clade I (*14*). Macroscopic and microscopic examination of growth layers exposed on the interior surface of the bisected tusk, along with x-ray and isotope data along the entire tusk ($^{87}Sr/^{86}Sr$, $\delta^{18}O$, $\delta^{13}C$, and $\delta^{15}N$ values) (Fig. 1), established a minimum age at death of 28 years.

The ultrahigh-resolution $^{87}Sr/^{86}Sr$ (~340,000 individual $^{87}Sr/^{86}Sr$ ratio measurements) and lower-resolution $\delta^{18}O$ variations [i.e., about weekly temporal resolution (*14*)] along the entire tusk (Fig. 1) were used to infer the geographic range and movements of the mammoth during its main life stages (*14*). Our approach involved comparing the $^{87}Sr/^{86}Sr$ data from the tusk, averaged at about weekly resolution, along with the $\delta^{18}O$ values, to a set of predictive isotope maps for Alaska and northwestern Canada. We then used an isotopically informed Markov chain Monte Carlo approach to identify the most probable routes and most frequently visited areas (Fig. 2) (*14*) and reconstruct the mammoth's movement history (*14*).

The isotope data indicate four main life stages: neonate, juvenile, adult, and end of life (the last ~1.5 years) (*14*). Data from the first ~10 cm from the tusk tip showed minimal $^{87}Sr/^{86}Sr$ variation (Fig. 1), suggesting that the young mammoth mostly occupied a range in the lower Yukon River basin in interior Alaska (*14*). As a juvenile (2 to 16 years, represented by the next ~75 cm of tusk), the mammoth used a larger range spanning some of the lowlands of interior Alaska between the Alaska and Brooks ranges (Fig. 2) (*14*). The mammoth undertook regular north-south movements within this large core area (Fig. 2) (*14*) as well as several long-distance movements, sometimes reaching the eastern end of the Brooks Range and the northern Seward Peninsula in the west (Fig. 2) (*14*). These juvenile-age movements probably represent the movements of a herd (*16*–*18*). Living *Loxodonta* and *Elephas* species both form stable matriarchal social units comprising related females together with their juvenile male and female offspring (*16*–*18*).

With increasing maturity, our study mammoth broadened his range (Fig. 2). After ~16 years, a distinctive transition occurred involving higher variance in $^{87}Sr/^{86}Sr$ along with other isotopic changes (Fig. 1) (*14*). This implied change in the animal's range probably reflects a transition to reproductive maturity accompanied by long-distance travel between interior Alaska and the North Slope of the Brooks Range (Fig. 2). These movements were probably in response to seasonal changes in resource availability. Today, male individuals of *Elephas maximus* and *Loxodonta africana* tend to be more mobile than females, and they typically leave matriarch-led herds to lead solitary lives or form all-male groups (*16*–*18*). It is noteworthy that some of the mammoth's areas of frequent use (Fig. 2) (*14*) mirror those of extant caribou (*6*), suggesting the use of these areas by Arctic herbivores over many millennia. Holarctic mammoth distributions seem to favor habitats in mountainous settings (*19*). Notably, some of the areas most frequently visited by our mammoth bull are proximate to locations of higher densities of mammoth remains and some of the earliest sites of human occupation in Alaska (Fig. 2).

The isotopic proxies for movements ($^{87}Sr/^{86}Sr$ ratios and $\delta^{18}O$ values), habitat, and diet [$\delta^{13}C$ and $\delta^{15}N$ values (*14*)] from the final ~10 cm of the tusk, representing the last ~1.5 years of our mammoth's life, reveal that it occupied a reduced range entirely restricted to the region north of the Brooks Range (Fig. 2), where it most likely died from starvation (*14*). Evidence for starvation includes a substantial increase in $\delta^{15}N$ values and a corresponding decrease in $\delta^{13}C$ values (Fig. 1) (*14*). Death seems to have occurred in late winter or spring, after it was restricted to a small area surrounding the location of its death (Fig. 2). A winter or spring season of mortality is indicated by declining $\delta^{18}O$ values immediately before its death. Winter conditions can include a scarcity of food and low temperatures, requiring greater investments in thermoregulation by arctic mammals, and this seasonality of death is consistent with previously analyzed mammoths from Chukotka and Wrangel Island (*7*).

The last persisting woolly mammoth populations were geographically constrained on isolated and relatively small islands (*1*, *7*), with no option to maintain large-scale movement patterns. Their extinctions were probably due to stochastic environmental changes and inbreeding that contributed to their eventual demise (*1*, *7*). Our study mammoth, alternatively, moved widely across unglaciated parts of Alaska. If woolly mammoth populations on mainland Beringia maintained a similar degree of mobility during the transition from the last ice age to the warmer and wetter Holocene, this behavior could have imparted additional ecophysiological stress, in particular as the forb- and/or graminoid-dominated ecosystems diminished and forests and peatlands expanded (*4*, *20*). This may have compounded their vulnerabilities to other stressors, including predation from humans and large carnivores (*21*), potentially

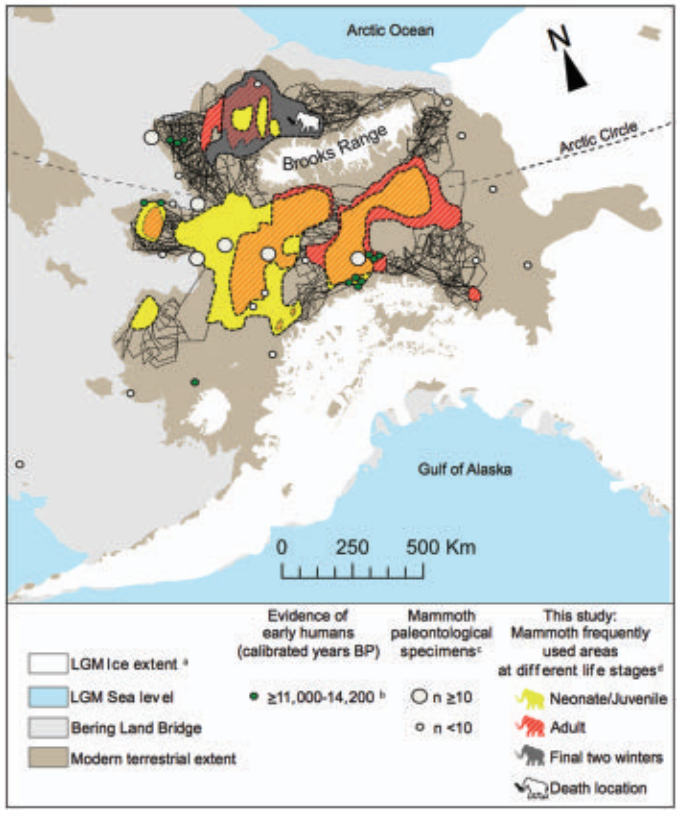



**Fig. 2. Summary life history of this study's woolly mammoth within the geographic, climatic, and human dispersal context of Beringia during the Late Quaternary.** The core areas correspond to those that were visited most frequently by the individual during each life stage (colored polygons) (*14*) (orange areas signify areas of overlap between neonate/juvenile and adult frequently used areas). The black dashed lines between the core areas represent the routes produced by the best walks (*14*). The white mammoth symbol indicates the area where the specimen was found (i.e., death location). Also shown are locations where the remains of other mammoths have been found and where evidence of early humans has been reported. Superscript letters indicate sources: a, (*22*); b, (*23*–*25*); c, compiled from Arctos and Neotoma databases (*14*); and d, this study. LGM, Last Glacial Maximum; BP, before present.

contributing to their eventual extinction from mainland Beringia.

## REFERENCES AND NOTES


1. R. W. Graham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9310–9314 (2016).
2. E. D. Lorenzen *et al.*, *Nature* **479**, 359–364 (2011).
3. G. M. MacDonald *et al.*, *Nat. Commun.* **3**, 893 (2012).
4. D. H. Mann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14301–14306 (2015).
5. E. C. Mills *et al.*, *PLOS ONE* **13**, e0199387 (2018).
6. A. P. Baltensperger, K. Joly, *Mov. Ecol.* **7**, 18 (2019).
7. J. J. El Adli, D. C. Fisher, S. L. Vartanyan, A. N. Tikhonov, *Quat. Int.* **445**, 135–145 (2017).
8. D. C. Fisher, D. L. Fox, *Annu. Rev. Earth Planet. Sci.* **46**, 229–260 (2006).
9. K. A. Hoppe, P. L. Koch, R. W. Carlson, S. D. Webb, *Geology* **27**, 439–442 (1999).
10. C. P. Bataille, B. E. Crowley, M. J. Wooller, G. J. Bowen, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **555**, 109849 (2020).
11. S. R. Brennan *et al.*, *Sci. Adv.* **1**, e1400124 (2015).
12. J. Funck, C. Bataille, J. Rasic, M. Wooller, *J. Quat. Sci.* **36**, 76–90 (2021).
13. K. T. Uno *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **559**, 109962 (2020).
14. See supplementary materials.
15. V. J. Lynch *et al.*, *Cell Rep.* **12**, 217–228 (2015).
16. L. A. Taylor *et al.*, *J. Anim. Ecol.* **89**, 57–67 (2020).
17. C. J. Moss, H. Croze, P. C. Lee, Eds., *The Amboseli Elephants: A Long-Term Perspective on a Long-Lived Mammal* (Univ. of Chicago Press, 2011).
18. L. Laursen, M. Bekoff, *Mamm. Species* **92**, 1–8 (1978).
19. R. D. Kahlke, *Quat. Int.* **379**, 147–154 (2015).
20. M. J. Wooller *et al.*, *R. Soc. Open Sci.* **5**, 180145 (2018).
21. B. Van Valkenburgh, M. W. Hayward, W. J. Ripple, C. Meloro, V. L. Roth, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 862–867 (2016).
22. A. S. Dalton *et al.*, *Quat. Sci. Rev.* **234**, 106223 (2020).
23. C. E. Holmes, in *From the Yenisei to the Yukon: Interpreting Lithic Assemblage Variability in Late Pleistocene/Early Holocene Beringia*, T. Goebel, I. Buvit, Eds. (Texas A&M Univ. Press, 2011), pp. 179–191.
24. B. A. Potter, *Environ. Archaeol.* **12**, 3–23 (2007).
25. F. B. Lanoë, J. D. Reuther, C. E. Holmes, *J. Archaeol. Method Theory* **25**, 818–838 (2018).


## ACKNOWLEDGMENTS


We dedicate this paper to the memory of Dr. Sean Brennan. We thank D. Patterson and M. Grif at TVC Radiology for taking the x-rays of our study tusk. The discovery and excavation of the Kik mammoth was funded and supported by the Arctic Field Office of the Bureau of Land Management. We thank M. Kunz, B. Gaglioti, M. Yazwinsky, and C. Adkins for field operations and assistance in recovering the specimen. Our thanks go to D. Pomraning and his co-workers at the Geophysical Institute for help sectioning the tusk we studied. We thank T. Cade (ISOMASS) and T. Oare (Thermo Fisher) for technical assistance and K. Joly for comments on our manuscript. **Funding:** This study was supported by the University of Alaska's Faculty Initiative Fund, the M. J. Murdock Charitable Trust, and the National Science Foundation [MCT SR-10 201811010, NSF DBI MRI 1625573 (M.J.W.); NSF BMMB EAGER 1937050 (G.M.E.)] and by the National Sciences and Engineering Research Council [Discovery Grant RGPIN-2019-05709 (C.B.)]. **Author contributions:** Conceptualization: M.J.W., C.B., P.D., and A.D.W. Methodology: M.J.W., C.B., A.D.W., G.M.E., B.S., K.J.S., J.I., N.H., T.H., T.P., P.D., and K.M. Formal analysis: K.M., B.S., C.B., A.D.W., T.H., N.H., M.J.W., and K.J.S. Investigation: C.B., M.J.W., and A.D.W. Resources: M.J.W. and P.D. Writing – original draft: M.J.W., P.D., C.B., A.D.W., K.J.S., J.Ra., G.M.E., and B.S. Writing – review and editing: All authors. Visualization: M.J.W., C.B., A.D.W., and K.J.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** DNA sequences are deposited in GenBank, and all other data are available in the main text or supplementary materials. All code associated with our spatial modeling will be made freely accessible on the journal website [and at GitHub (https://github.com/statdivlab/KikWalk)].


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/806/suppl/DC1
Materials and Methods
Figs. S1 to S21
Tables S1 to S7
References (*26*–*90*)

MDAR Reproducibility Checklist
Movie S1
Data S1 to S3

10 December 2020; accepted 12 July 2021
10.1126/science.abg1134

METABOLISM

# Daily energy expenditure through the human life course

**Herman Pontzer[1,2]*†, Yosuke Yamada[3,4]*, Hiroyuki Sagayama[5]*, Philip N. Ainslie[6], Lene F. Andersen[7], Liam J. Anderson[6,8], Lenore Arab[9], Issaad Baddou[10], Kweku Bedu-Addo[11], Ellen E. Blaak[12], Stephane Blanc[13,14], Alberto G. Bonomi[15], Carlijn V. C. Bouten[12], Pascal Bovet[16], Maciej S. Buchowski[17], Nancy F. Butte[18], Stefan G. Camps[12], Graeme L. Close[6], Jamie A. Cooper[13], Richard Cooper[19], Sai Krupa Das[20], Lara R. Dugas[19], Ulf Ekelund[21], Sonja Entringer[22,23], Terrence Forrester[24], Barry W. Fudge[25], Annelies H Goris[12], Michael Gurven[26], Catherine Hambly[27], Asmaa El Hamdouchi[10], Marjije B. Hoos[12], Sumei Hu[28], Noorjehan Joonas[29], Annemiek M. Joosen[12], Peter Katzmarzyk[30], Kitty P. Kempen[12], Misaka Kimura[3], William E. Kraus[31], Robert F. Kushner[32], Estelle V. Lambert[33], William R. Leonard[34], Nader Lessan[35], Corby Martin[30], Anine C. Medin[7,36], Erwin P. Meijer[12], James C. Morehen[37,6], James P. Morton[6], Marian L. Neuhouser[38], Teresa A. Nicklas[18], Robert M. Ojiambo[39,40], Kirsi H. Pietiläinen[41], Yannis P. Pitsiladis[42], Jacob Plange-Rhule[43]‡, Guy Plasqui[44], Ross L. Prentice[38], Roberto A. Rabinovich[45], Susan B. Racette[46], David A. Raichlen[47], Eric Ravussin[30], Rebecca M. Reynolds[48], Susan B. Roberts[20], Albertine J. Schuit[49], Anders M. Sjödin[50], Eric Stice[51], Samuel S. Urlacher[52], Giulio Valenti[12,15], Ludo M. Van Etten[12], Edgar A. Van Mil[53], Jonathan C. K. Wells[54], George Wilson[6], Brian M. Wood[55,56], Jack Yanovski[57], Tsukasa Yoshida[4], Xueying Zhang[27,28], Alexia J. Murphy-Alford[58], Cornelia Loechl[58], Amy H. Luke[59]*, Jennifer Rood[30]*, Dale A. Schoeller[60]*, Klaas R. Westerterp[61]*, William W. Wong[18]*, John R. Speakman[62,27,28,63]*†, IAEA DLW Database Consortium§**

**Total daily energy expenditure ("total expenditure") reflects daily energy needs and is a critical variable in human health and physiology, but its trajectory over the life course is poorly studied. We analyzed a large, diverse database of total expenditure measured by the doubly labeled water method for males and females aged 8 days to 95 years. Total expenditure increased with fat-free mass in a power-law manner, with four distinct life stages. Fat-free mass–adjusted expenditure accelerates rapidly in neonates to ~50% above adult values at ~1 year; declines slowly to adult levels by ~20 years; remains stable in adulthood (20 to 60 years), even during pregnancy; then declines in older adults. These changes shed light on human development and aging and should help shape nutrition and health strategies across the life span.**

All of life's essential tasks, from development and reproduction to maintenance and movement, require energy. Total daily energy expenditure (total expenditure; megajoules per day) is thus central to understanding both daily nutritional requirements and the body's investment among activities. Yet, we know surprisingly little about total expenditure in humans or how it changes over the life span. Most large ($n$ > 1000 subjects) analyses of human energy expenditure have been limited to basal expenditure—the metabolic rate at rest (*1*), which accounts for only a portion (usually ~50 to 70%) of total expenditure—or have estimated total expenditure from basal expenditure and daily physical activity (*2*–*5*). Doubly labeled water studies provide measurements of total expenditure in free-living subjects but have been limited in sample size ($n$ < 600 subjects), geographic and socioeconomic diversity, and/or age (*6*–*9*).

Body composition, size, and physical activity change over the life course, often in concert, making it difficult to parse the determinants of energy expenditure. Total and basal expenditures increase with age as children grow and mature (*10*, *11*), but the relative effects of increasing physical activity and age-related changes in tissue-specific metabolic rates are unclear (*12*–*16*). Similarly, the decline in total expenditure beginning in older adults corresponds with declines in fat-free mass and physical activity but may also reflect age-related reductions in organ metabolism (*9*, *17*–*19*).

We investigated the effects of age, body composition, and sex on total expenditure using a large ($n$ = 6421 subjects; 64% female), diverse ($n$ = 29 countries) database of doubly labeled

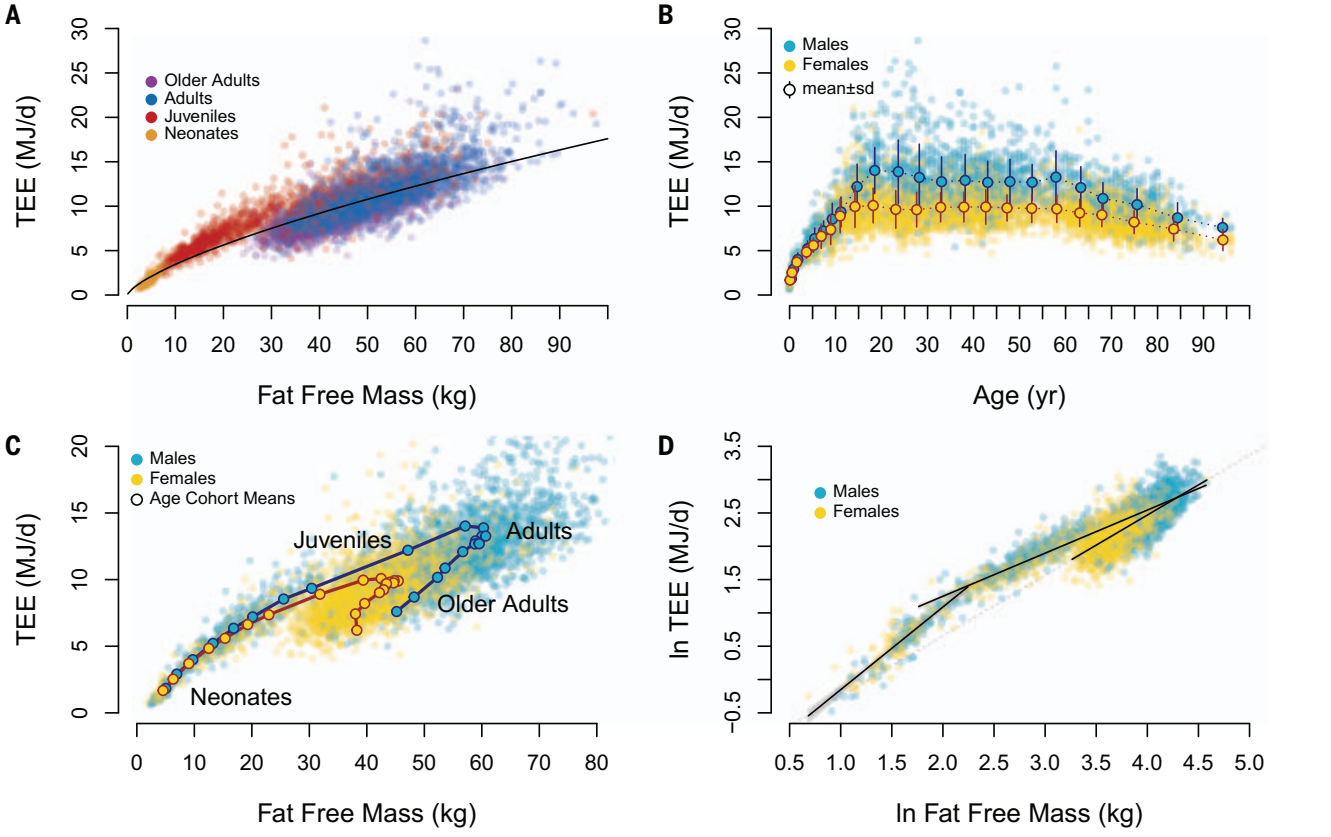



**Fig. 1. Total energy expenditure (TEE) through the human life course.** (**A**) TEE increases with fat-free mass (FFM) in a power-law manner, but age groups cluster about the trend line differently. The black line indicates TEE = $0.677\text{FFM}^{0.708}$. Coefficient of determination ($R^2$) = 0.83; $P < 0.0001$ (table S2). (**B**) Total expenditure rises in childhood, is stable through adulthood, and declines in older adults. Means ± SD for age-sex cohorts are shown. (**C**) Age-sex cohort means show a distinct progression of total expenditure and fat-free mass over the life course. (**D**) Neonates, juveniles, and adults exhibit distinct relationships between fat-free mass and expenditure. The dashed line, extrapolated from the regression for adults, approximates the regression used to calculate adjusted total expenditure.

[1]Department of Evolutionary Anthropology, Duke University, Durham, NC, USA. [2]Duke Global Health Institute, Duke University, Durham, NC, USA. [3]Institute for Active Health, Kyoto University of Advanced Science, Kyoto, Japan. [4]National Institute of Health and Nutrition, National Institutes of Biomedical Innovation, Health and Nutrition, Tokyo, Japan. [5]Faculty of Health and Sport Sciences, University of Tsukuba, Ibaraki, Japan. [6]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK. [7]Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, 0317 Oslo, Norway. [8]Crewe Alexandra Football Club, Crewe, UK. [9]David Geffen School of Medicine, University of California, Los Angeles. [10]Unité Mixte de Recherche en Nutrition et Alimentation, CNESTEN–Université Ibn Tofail URAC39, Regional Designated Center of Nutrition Associated with AFRA/IAEA, Rabat, Morocco. [11]Department of Physiology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. [12]Maastricht University, Maastricht, Netherlands. [13]Department of Nutritional Sciences, University of Wisconsin, Madison, WI, USA. [14]Institut Pluridisciplinaire Hubert Curien, CNRS Université de Strasbourg, UMR7178, France. [15]Phillips Research, Eindoven, Netherlands. [16]Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland. [17]Division of Gastroenterology, Hepatology, and Nutritiion, Department of Medicine, Vanderbilt University, Nashville, TN, USA [18]Department of Pediatrics, Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, TX, USA. [19]Department of Public Health Sciences, Parkinson School of Health Sciences and Public Health, Loyola University, Maywood, IL, USA. [20]Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, USA. [21]Department of Sport Medicine, Norwegian School of Sport Sciences, Oslo, Norway. [22]Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH), Institute of Medical Psychology, Berlin, Germany. [23]School of Medicine, University of California Irvine, Irvine, CA, USA. [24]Solutions for Developing Countries, University of the West Indies, Mona, Kingston, Jamaica. [25]Department of Biomedical and Life Sciences, University of Glasgow, Glasgow, UK. [26]Department of Anthropology, University of California Santa Barbara, Santa Barbara, CA, USA. [27]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK. [28]State Key Laboratory of Molecular developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [29]Central Health Laboratory, Ministry of Health and Wellness, Candos, Mauritius. [30]Pennington Biomedical Research Center, Baton Rouge, LA, USA. [31]Department of Medicine, Duke University, Durham, NC, USA. [32]Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [33]Health through Physical Activity, Lifestyle and Sport Research Centre (HPALS), Division of Exercise Science and Sports Medicine (ESSM), FIMS International Collaborating Centre of Sports Medicine, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. [34]Department of Anthropology, Northwestern University, Evanston, IL, USA. [35]Imperial College London Diabetes Centre, Abu Dhabi, United Arab Emirates and Imperial College London, London, UK. [36]Department of Nutrition and Public Health, Faculty of Health and Sport Sciences, University of Agder, 4630 Kristiansand, Norway. [37]The FA Group, Burton-Upon-Trent, Staffordshire, UK. [38]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center and School of Public Health, University of Washington, Seattle, WA, USA. [39]Kenya School of Medicine, Moi University, Eldoret, Kenya. [40]Rwanda Division of Basic Sciences, University of Global Health Equity, Rwanda. [41]Helsinki University Central Hospital, Helsinki, Finland. [42]School of Sport and Service Management, University of Brighton, Eastbourne, UK. [43]Department of Physiology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. [44]Department of Nutrition and Movement Sciences, Maastricht University, Maastricht, Netherlands. [45]The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK. [46]Program in Physical Therapy and Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA. [47]Biological Sciences and Anthropology, University of Southern California, CA, USA. [48]Centre for Cardiovascular Sciences, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK. [49]School of Social and Behavioral Sciences, University of Tilburg, Tilburg, Netherlands. [50]Department of Nutrition, Exercise and Sports, Copenhagen University, Copenhagen, Denmark. [51]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford CA, USA. [52]Department of Anthropology, Baylor University, Waco, TX, USA. [53]Maastricht University, Maastricht and Lifestyle Medicine Center for Children, Jeroen Bosch Hospital, Hertogenbosch, Netherlands. [54]Population, Policy, and Practice Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, London, UK. [55]Department of Anthropology, University of California Los Angeles, Los Angeles, CA, USA. [56]Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [57]Growth and Obesity, Division of Intramural Research, NIH, Bethesda, MD, USA. [58]Nutritional and Health Related Environmental Studies Section, Division of Human Health, International Atomic Energy Agency, Vienna, Austria. [59]Division of Epidemiology, Department of Public Health Sciences, Loyola University School of Medicine, Maywood, IL, USA. [60]Biotech Center and Nutritional Sciences University of Wisconsin, Madison, WI, USA. [61]Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University Medical Centre, Maastricht, Netherlands. [62]Center for Energy Metabolism and Reproduction, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. [63]CAS Center of Excellence in Animal Evolution and Genetics, Kunming, China.
***Corresponding author. Email: herman.pontzer@duke.edu (H.P.); j.speakman@abdn.ac.uk (J.R.S.); yyamada831@gmail.com (Y.Y.); sagayama.hiroyuki.ka@u.tsukuba.ac.jp (H.S.); aluke@luc.edu (A.H.L.); jennifer.rood@pbrc.edu (J.R.); dschoell@nutrisci.wisc.edu (D.A.S.); k.westerterp@maastrichtuniversity.nl (K.R.W.); wwong@bcm.edu (W.W.W.)**
†These authors contributed equally to this work. ‡Deceased. §IAEA DLW Database Consortium members are listed in the supplementary materials.

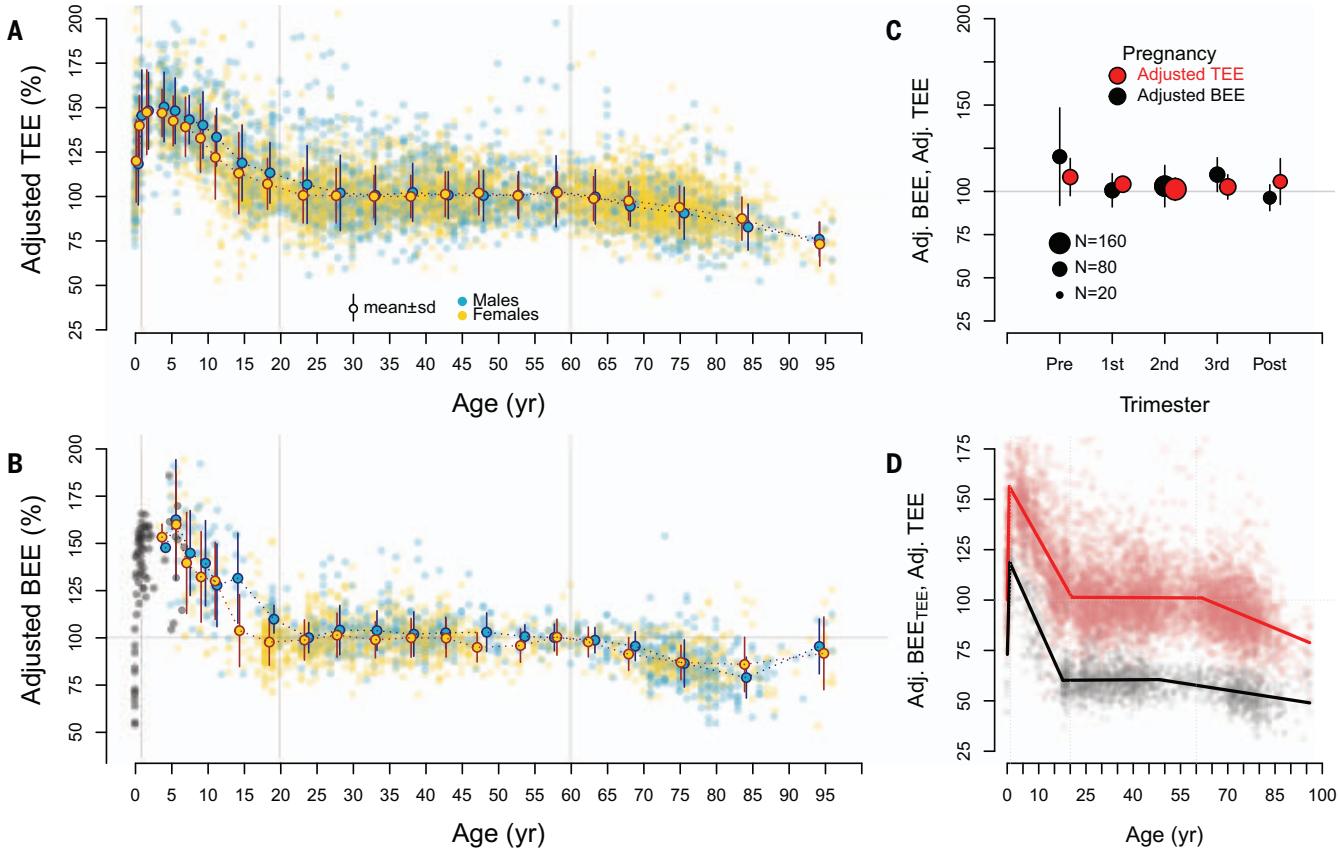



**Fig. 2. Fat-free mass– and fat mass–adjusted expenditures over the life course.** Individual subjects and age-sex cohort mean ± SD are shown. For both (**A**) total expenditure (adjusted TEE) and (**B**) basal expenditure (adjusted BEE), adjusted expenditures begin near adult levels (~100%) but quickly climb to ~150% in the first year. Adjusted expenditures decline to adult levels at ~20 years of age then decline again in older adults. Basal expenditures for infants and children not in the DLW Database are shown in gray. (**C**) Pregnant mothers exhibit adjusted total and basal expenditures similar to those of nonreproducing adults (Pre, before pregnancy; Post, 27 weeks postpartum). (**D**) Segmented regression analysis of adjusted total (red) and adjusted basal expenditure (black) (calculated as a portion of total, Adj. $BEE_{TEE}$) indicates a peak at ~1 year of age, adult levels at ~20 years of age, and decline at ~60 years of age.

water measurements for subjects aged 8 days to 95 years (*20*), calculating total expenditure from isotopic measurements by using a single, validated equation for all subjects (*21*). Basal expenditure, measured with indirect calorimetry, was available for $n$ = 2008 subjects, and we augmented the dataset with additional published measures of basal expenditure in neonates and doubly labeled water–mesaured total expenditure in pregnant and postpartum women (supplementary materials, materials and methods, and table S1).

We found that both total and basal expenditure increased with fat-free mass in a power-law manner (Fig. 1, figs. S1 and S2, and table S1), requiring us to adjust for body size to isolate potential effects of age, sex, and other factors. Because of the power-law relation with size, the ratio of energy expenditure/mass does not adequately control for body size because the ratio trends lower for larger individuals (fig. S1). Instead, we used regression analysis to control for body size (*22*). We used a general linear model with log-transformed values of energy expenditure (total or basal), fat-free mass, and fat mass in adults 20 to 60 years (table S2) to calculate residual expenditures for each subject. We converted these residuals to "adjusted" expenditures for clarity in discussing age-related changes: 100% indicates an expenditure that matches the expected value given the subject's fat-free mass and fat mass, 120% indicates an expenditure 20% above expected, and so on. Using this approach, we also calculated the portion of adjusted total expenditure attributed to basal expenditure (Fig. 2D and materials and methods). Segmented regression analysis (materials and methods) revealed four distinct phases of adjusted total and basal expenditure over the life span.

The first phase is of neonates, up to 1 year of age. Neonates in the first month of life had size-adjusted energy expenditures similar to that of adults, with adjusted total expenditure of 99.0 ± 17.2% ($n$ = 35 subjects) and adjusted basal expenditure of 78.1 ± 15.0% ($n$ = 34 subjects) (Fig. 2). Both measures increased rapidly in the first year. In segmented regression analysis, adjusted total expenditure rose 84.7 ± 7.2% per year from birth to a break point at 0.7 years of age [95% confidence interval (CI): 0.6, 0.8]; a similar rise and break point were evident in adjusted basal expenditure (table S4). For subjects between 9 and 15 months of age, adjusted total and basal expenditures were nearly ~50% elevated compared with that of adults (Fig. 2).

The second phase is of juveniles, 1 to 20 years of age. Total and basal expenditure continued to increase with age throughout childhood and adolescence along with fat-free mass (Fig. 1), but size-adjusted expenditures steadily declined. Adjusted total expenditure declined at a rate of −2.8 ± 0.1% per year from 147.8 ± 22.6% for subjects 1 to 2 years of age to 102.7 ± 18.1% for subjects 20 to 25 years of age (tables S2

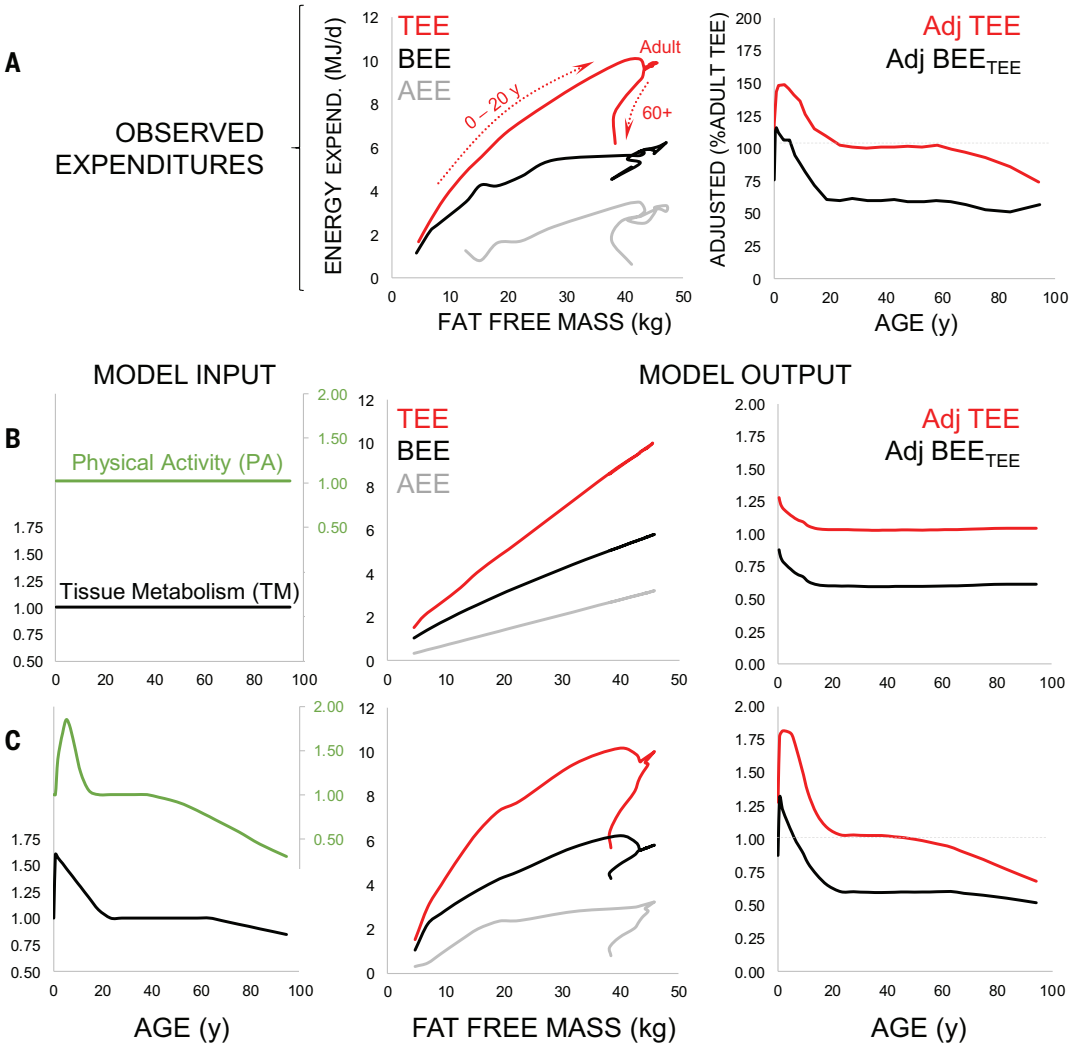



**Fig. 3. Modeling the contribution of physical activity and tissue-specific metabolism to daily expenditures.** (**A**) Observed total expenditure (TEE; red), basal expenditure (BEE; black), and activity expenditure (AEE; gray) (table S1) show age-related variation with respect to fat-free mass (Fig. 1C) that is also evident in adjusted values (Fig. 2D and table S3). (**B**) These age effects do not emerge in models that assume constant physical activity (PA; green) and tissue-specific metabolic rate (TM; black) across the life course. (**C**) When physical activity and tissue-specific metabolism follow the life course trajectories evident from accelerometry and adjusted basal expenditure, respectively, model output is similar to observed expenditures.

and S4). Segmented regression analysis identified a break point in adjusted total expenditure at 20.5 years (95% CI: 19.8, 21.2), after which it plateaued at adult levels (Fig. 2); a similar decline and break point were evident in adjusted basal expenditure (Fig. 2 and table S4). No pubertal increases in adjusted total or basal expenditure were evident among subjects 10 to 15 years of age (Fig. 2 and table S3). In multivariate regression for subjects 1 to 20 years of age, males had a higher total expenditure and adjusted total expenditure (tables S2 and S3), but sex had no detectable effect on the rate of decline in adjusted total expenditure with age (sex:age interaction, $P = 0.30$).

The third phase is adulthood, from 20 to 60 years of age. Total and basal expenditure and fat-free mass were all stable from ages 20 to 60 years (Figs. 1 and 2 and tables S1 and S2). Sex had no effect on total expenditure in multivariate models with fat-free mass and fat mass, nor in analyses of adjusted total expenditure (tables S2 and S4). Adjusted total and basal expenditures were stable even during pregnancy; the elevation in unadjusted expenditures matched those expected from the gain in mothers' fat-free mass and fat mass (Fig. 2C). Segmented regression analysis identified a break point at 63.0 years of age (95% CI: 60.1, 65.9), after which adjusted total expenditure begins to decline. This break point was somewhat earlier for adjusted basal expenditure (46.5, 95% CI: 40.6, 52.4), but the relatively small number of basal measures for 45 to 65 years of age (Fig. 2D) reduces our precision in determining this break point.

The fourth phase is of older adults, >60 years of age. At ~60 years of age, total and basal expenditure begin to decline, along with fat-free mass and fat mass (Fig. 1, fig. S3, and table S1). Declines in expenditure are not only a function of reduced fat-free mass and fat mass, however. Adjusted total expenditure declined by −0.7 ± 0.1% per year, and adjusted basal expendiure fell at a similar rate (Fig. 2, fig. S3, supplementary text S1, and table S4). For subjects 90+ years of age, adjusted total expenditure was ~26% below that of middle-aged adults.

Our analyses provide empirical measures and predictive equations for total and basal expenditure from infancy to old age (tables S1 and S2) and bring to light major metabolic changes across the life course. To begin, we can infer fetal metabolic rates from maternal measures during pregnancy: If body size–adjusted expenditures were elevated in the fetus, then adjusted expenditures for pregnant mothers—particularly late in pregnancy, when the fetus accounts for a substantial portion of a mother's weight—would be likewise elevated. Instead, the stability of adjusted total and basal expenditures at ~100% during pregnancy (Fig. 2B) indicates that the growing fetus maintains a fat-free mass– and fat mass–adjusted metabolic rate similar to that of adults, which is consistent with adjusted expenditures of neonates (both ~100%) (Fig. 2) in the first

weeks after birth. Total and basal expenditures, both absolute and size-adjusted values, then accelerate rapidly over the first year. This early period of metabolic acceleration corresponds to a critical period in early development in which growth often falters in nutritionally stressed populations (*23*). Increasing energy demands could be a contributing factor.

After rapid acceleration in total and basal expenditure during the first year, adjusted expenditures progressively decline thereafter, reaching adult levels at ~20 years of age. Elevated adjusted expenditures in this life stage may reflect the metabolic demands of growth and development. Adult expenditures, adjusted for body size and composition, are remarkably stable, even during pregnancy and postpartum. Declining metabolic rates in older adults could increase the risk of weight gain. However, neither fat mass nor percentage increased in this period (fig. S3), which is consistent with the hypothesis that energy intake is coupled to expenditure (*24*).

Following previous studies (*15*, *16*, *19*, *25*, *26*), we calculated the effect of organ size on basal expenditure over the life span (materials and methods). Organs with a high tissue-specific metabolic rate, particularly the brain and liver, account for a greater proportion of fat-free mass in young individuals. Thus, organ-based basal expenditure, estimated from organ size and tissue-specific metabolic rate, follows a power-law relationship with fat-free mass that is roughly consistent with observed basal expenditures (materials and methods, and fig. S6). Still, observed basal expenditure exceeded organ-based estimates by ~30% in early life (1 to 20 years of age) and was ~20% lower than organ-based estimates in subjects over 60 years of age (fig. S6), which is consistent with studies indicating that tissue-specific metabolic rates are elevated in juveniles (*15*, *16*) and reduced in older adults (*19*, *25*, *26*).

We investigated the contributions of daily physical activity and changes in tissue-specific metabolic rate to total and basal expenditure using a simple model with two components: activity and basal expenditure (Fig. 3 and materials and methods). Activity expenditure was modeled as a function of physical activity and body mass, assuming that activity costs are proportional to weight, and could either remain constant over the life span or follow the trajectory of daily physical activity measured with accelerometry, peaking at 5 to 10 years of age and declining thereafter (Fig. 3) (*12*, *17*, *18*). Similarly, basal expenditure was modeled as a power function of fat-free mass (consistent with organ-based basal expenditure estimates) (materials and methods) multiplied by a "tissue-specific metabolism" term, which could either remain constant at adult levels across the life span or follow the trajectory observed in adjusted basal expenditure (Fig. 2). For each scenario, total expenditure was modeled as the sum of activity and basal expenditure (materials and methods).

Models that hold physical activity or tissue-specific metabolic rates constant over the life span do not reproduce the observed patterns of age-related change in absolute or adjusted measures of total or basal expenditure (Fig. 3). Only when age-related changes in physical activity and tissue-specific metabolism are included does model output match observed expenditures, indicating that variation in both physical activity and tissue-specific metabolism contribute to total expenditure and its components across the life span. Elevated tissue-specific metabolism in early life may be related to growth or development (*15*, *16*). Conversely, reduced expenditures in later life may reflect a decline in organ-level metabolism (*25*–*27*).

Metabolic models of life history commonly assume continuity in tissue-specific metabolism over the life course, with metabolic rates increasing in a stable, power-law manner (*28*, *29*). Measures of humans here challenge this view, with deviations from the power-law relationships for total and basal expenditure in childhood and old age (Figs. 1 and 2). These changes present a potential target for investigating the kinetics of disease, drug activity, and healing, processes that are intimately related to metabolic rate. Further, interindividual variation in expenditure is considerable even when controlling for fat-free mass, fat mass, sex, and age (Figs. 1 and 2 and table S2). Elucidating the processes underlying metabolic changes across the life course and variation among individuals may help reveal the roles of metabolic variation in health and disease.

## REFERENCES AND NOTES


1. C. J. Henry, *Public Health Nutr.* **8**, 1133–1152 (2005).
2. Food and Agriculture Organization of the United Nations (FAO), *Food Nutr. Bull.* **26**, 166 (2005).
3. K. R. Westerterp, J. O. de Boer, W. H. M. Saris, P. F. M. Schoffelen, F. ten Hoor, *Int. J. Sports Med.* **05**, S74–S75 (1984).
4. P. D. Klein *et al*., *Hum. Nutr. Clin. Nutr.* **38**, 95–106 (1984).
5. J. R. Speakman, *Doubly Labelled Water: Theory and Practice* (Chapman and Hall, 1997).
6. A. E. Black, W. A. Coward, T. J. Cole, A. M. Prentice, *Eur. J. Clin. Nutr.* **50**, 72–92 (1996).
7. L. R. Dugas *et al*., *Am. J. Clin. Nutr.* **93**, 427–441 (2011).
8. H. Pontzer *et al*., *Curr. Biol.* **26**, 410–417 (2016).
9. J. R. Speakman, K. R. Westerterp, *Am. J. Clin. Nutr.* **92**, 826–834 (2010).
10. N. F. Butte, *Am. J. Clin. Nutr.* **72** (suppl.), 1246S–1252S (2000).
11. H. L. Cheng, M. Amatoury, K. Steinbeck, *Am. J. Clin. Nutr.* **104**, 1061–1074 (2016).
12. D. L. Wolff-Hughes, D. R. Bassett, E. C. Fitzhugh, *PLOS ONE* **9**, e115915 (2014).
13. E. A. Schmutz *et al*., *Int. J. Behav. Nutr. Phys. Act.* **15**, 35 (2018).
14. J. A. Hnatiuk, K. E. Lamb, N. D. Ridgers, J. Salmon, K. D. Hesketh, *Int. J. Behav. Nutr. Phys. Act.* **16**, 42 (2019).
15. A. Hsu *et al*., *Am. J. Clin. Nutr.* **77**, 1506–1511 (2003).
16. Z. Wang *et al*., *Am. J. Hum. Biol.* **22**, 476–483 (2010).
17. D. L. Wolff-Hughes, E. C. Fitzhugh, D. R. Bassett, J. R. Churilla, *J. Phys. Act. Health* **12**, 447–453 (2015).
18. Y. Aoyagi, S. Park, S. Cho, R. J. Shephard, *Prev. Med. Rep.* **11**, 180–186 (2018).
19. D. Gallagher, A. Allen, Z. Wang, S. B. Heymsfield, N. Krasnow, *Ann. N. Y. Acad. Sci.* **904**, 449–455 (2000).
20. J. R. Speakman *et al*., *Ann. Nutr. Metab.* **75**, 114–118 (2019).
21. J. R. Speakman *et al*., *Cell Rep. Med.* **2**, 100203 (2021).
22. D. B. Allison, F. Paultre, M. I. Goran, E. T. Poehlman, S. B. Heymsfield, *Int. J. Obes. Relat. Metab. Disord.* **19**, 644–652 (1995).
23. H. Alderman, D. Headey, *PLOS ONE* **13**, e0195904 (2018).
24. J. E. Blundell *et al*., *Physiol. Behav.* **219**, 112846 (2020).
25. Z. Wang *et al*., *Am. J. Clin. Nutr.* **92**, 1369–1377 (2010).
26. Z. Wang, S. Heshka, S. B. Heymsfield, W. Shen, D. Gallagher, *Am. J. Clin. Nutr.* **81**, 799–806 (2005).
27. Y. Yamada *et al*., *J. Gerontol. A Biol. Sci. Med. Sci.* **65**, 510–516 (2010).
28. G. B. West, J. H. Brown, B. J. Enquist, *Nature* **413**, 628–631 (2001).
29. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, *Ecology* **85**, 1771–1789 (2004).


## ACKNOWLEDGMENTS


We are grateful to the International Atomic Energy Agency (IAEA), Taiyo Nippon Sanso, and SERCON for their support and to T. Oono for his tremendous efforts at fundraising on our behalf. **Funding:** The IAEA Doubly Labeled Water (DLW) Database is generously supported by the IAEA, Taiyo Nippon Sanso, and SERCON. The authors also gratefully acknowledge funding from the US National Science Foundation (BCS-1824466) awarded to H.P. The funders played no role in the content of this manuscript. **Author contributions:** H.P., Y.Y., H.S., A.H.L., J.R., D.A.S., H.S., K.R.W., W.W.W., and J.R.S. wrote the paper. H.P., Y.Y., H.S., P.N.A., L.F.A., L.J.A., L.A., I.B., K.B.-A., E.E.B., S.B., A.G.B., C.V.C.B., P.B., M.S.B., N.F.B., S.G.C., G.L.C., J.A.C., R.C., S.K.D., L.R.D., U.E., S.E., T.F., B.W.F., A.H.G., M.G., C.H., A.E.H., M.B.H., S.H., N.J., A.M.J., P.K., K.P.K., M.K., W.E.K., R.F.K., E.V.L., W.R.L., N.L., C.M., A.C.M., E.P.M., J.C.M., J.P.M., M.L.N., T.A.N., R.M.O., K.H.P., Y.P.P., J.P.-R., G.P., R.L.P., R.A.R., S.B.R., D.A.R., E.R., R.N.R., S.B.R., A.J.S., A.M.S., E.S., S.S.U., G.V., L.M.V.E., E.A.V.M., J.C.K.W., G.W., B.M.W, J.Y., T.Y., X.Y., A.H.L., J.R., D.A.S., K.R.W., W.W.W., and J.R.S. contributed data. P.N.A., L.F.A., L.J.A., L.A., I.B., K.B.-A., E.E.B., S.B., A.G.B., C.V.C.B., P.B., M.S.B., N.F.B., S.G.C., G.L.C., J.A.C., R.C., S.K.D., L.R.D., U.E., S.E., T.F., B.W.F., A.H.G., M.G., C.H., A.E.H., M.B.H., S.H., N.J., A.M.J., P.K., K.P.K., M.K., W.E.K., R.F.K., E.V.L., W.R.L., N.L., C.M., A.C.M., E.P.M., J.C.M., J.P.M., M.L.N., T.A.N., R.M.O., K.H.P., Y.P.P., J.P.-R., G.P., R.L.P., R.A.R., S.B.R., D.A.R., E.R., R.N.R., S.B.R., A.J.S., A.M.S., E.S., S.S.U., G.V., L.M.V.E., E.A.V.M., J.C.K.W., G.W., B.M.W, J.Y., T.Y., and X.Y. read and commented on the manuscript. J.R.S., C.L., A.H.L., A.J. M-A., H.P., J.R., D.A.S., H.S., K.R.W., W.W.W., and Y.Y. assembled and managed the database. H.P., Y.Y., H.S., A.H.L., J.R., D.A.S., H.S., K.R.W., W.W.W., and J.R.S. performed and discussed the analysis. **Competing interests:** The authors have no conflicts of interest to declare. **Data availability:** All data used in these analyses is freely available via the IAEA DLW Database, which can be found at https://doubly-labelled-water-database.iaea.org/home and www.dlwdatabase.org.


## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/808/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S4
IAEA DLW Database Consortium Collaborators List
References (*30*–*54*)
MDAR Reproducibility Checklist

27 August 2020; accepted 21 June 2021
10.1126/science.abe5017

MICROBIOTA

# High-fat diet–induced colonocyte dysfunction escalates microbiota-derived trimethylamine *N*-oxide

Woongjae Yoo[1], Jacob K. Zieba[1], Nora J. Foegeding[1], Teresa P. Torres[1], Catherine D. Shelton[1], Nicolas G. Shealy[1], Austin J. Byndloss[2], Stephanie A. Cevallos[2], Erik Gertz[3,6], Connor R. Tiffany[2], Julia D. Thomas[1], Yael Litvak[2,4], Henry Nguyen[2], Erin E. Olsan[2,5], Brian J. Bennett[3,6], Jeffrey C. Rathmell[1,7,8,9], Amy S. Major[1,8,10], Andreas J. Bäumler[2]*, Mariana X. Byndloss[1,7,8,9]*

**A Western-style, high-fat diet promotes cardiovascular disease, in part because it is rich in choline, which is converted to trimethylamine (TMA) by the gut microbiota. However, whether diet-induced changes in intestinal physiology can alter the metabolic capacity of the microbiota remains unknown. Using a mouse model of diet-induced obesity, we show that chronic exposure to a high-fat diet escalates *Escherichia coli* choline catabolism by altering intestinal epithelial physiology. A high-fat diet impaired the bioenergetics of mitochondria in the colonic epithelium to increase the luminal bioavailability of oxygen and nitrate, thereby intensifying respiration-dependent choline catabolism of *E. coli*. In turn, *E. coli* choline catabolism increased levels of circulating trimethlamine *N*-oxide, which is a potentially harmful metabolite generated by gut microbiota.**

A Western-style, high-fat diet is often associated with cardiovascular disease, and one explanation for this is that members of the gut microbiota catabolize dietary choline into trimethylamine (TMA) (*1*), which is absorbed in the intestine and oxidized in the liver to trimethylamine *N*-oxide (TMAO), a metabolite that promotes atherosclerosis (*2*). A critical but unexplored aspect of the TMAO pathway is how the interaction between diet-impaired host physiology and microbial communities affects TMA production. Gene clusters responsible for TMA production are commonly found in obligately anaerobic *Clostridia* (phylum *Firmicutes*) and facultatively anaerobic *Enterobacteriaceae* (phylum *Proteobacteria*) (*1*, *3*), but only the latter taxon features a substantial increase in abundance in the feces of individuals on a high-fat diet (*4*–*7*). In addition to altering the microbiota composition, a high-fat diet also changes host physiology because saturated fatty acids impair mitochondrial bioenergetics by inducing hydrogen peroxide production in the mitochondria (*8*, *9*). Notably, in the colon, high mitochondrial oxygen consumption is vital for maintaining epithelial hypoxia (*10*), which preserves anaerobiosis to drive dominance of obligately anaerobic bacteria (*11*–*13*) while suppressing growth of facultative anaerobic *Enterobacteriaceae* (*14*). Therefore, we wanted to determine whether diet-impaired mitochondrial bioenergetics would escalate microbial choline catabolism by increasing the abundance of *Enterobacteriaceae*, which was modeled with *E. coli*.

To address this question, we first investigated whether diet-impaired mitochondrial bioenergetics were responsible for the increased *Enterobacteriaceae* abundance observed in individuals on a high-fat diet (*4*–*7*). We used mice from The Jackson Laboratory (C57BL/6J) that do not carry endogenous *Enterobacteriaceae* (*15*), which provided experimental control over this taxon. Inoculation of mice reared on a low- or high-fat diet (10% or 60% fat, respectively) with a single dose of *E. coli* Nissle 1917 (family *Enterobacteriaceae*), a commensal human isolate marketed as a probiotic (*16*), resulted in significantly higher fecal *E. coli* carriage in the latter group (Fig. 1A). This mirrors an increase in abundance of *Enterobacteriaceae* in the human fecal microbiota, driven by a high-fat diet (*6*, *7*).

Increased abundance of *Enterobacteriaceae* in fecal microbiota has been linked to mucosal inflammation (*17*) and shifts in epithelial metabolism (*14*); as a result, we investigated diet-induced mucosal responses in mice. Weight gain induced by a high-fat diet (fig. S1A) was associated with low-grade mucosal inflammation characterized by a reduction in colon length (fig. S1B), epithelial metaplasia (fig. S1C), an increased number of mitotic figures in the colonic epithelium (fig. S1, D and E), and reduced numbers of goblet cells (fig. S1, F and G), which was consistent with previous observations (*18*). A high-fat diet triggered similar responses in germ-free (Swiss Webster) mice (fig. S1, H to K), suggesting that the generation of low-grade mucosal inflammation was microbiota independent. In line with previous reports on saturated fatty acid–mediated impairment of mitochondrial bioenergetics (*8*, *9*), a high-fat diet was associated with reduced mitochondrial activity in the epithelium, as indicated by diminished expression of mitochondrial markers in mRNA isolated from colonic epithelial cells (Fig. 1B), as well as reduced levels of adenosine triphosphate (ATP) (Fig. 1C) and pyruvate dehydrogenase in the colonic epithelium (Fig. 1D). A high-fat diet is rich in saturated fatty acids, which have been implicated in diet-impaired mitochondrial bioenergetics (*8*, *9*). Consistent with the role of saturated fatty acids in reducing mitochondrial bioenergetics in the epithelium, palmitate treatment diminished expression of mitochondrial genes (fig. S2A) and reduced mitochondrial ATP production (fig. S2, B and C) in a human colonic epithelial (Caco-2) cancer cell line in vitro.

To investigate whether mitochondrial bioenergetics impaired by a high-fat diet were linked to increased epithelial oxygenation in the colon, we visualized epithelial hypoxia with the exogenous hypoxic marker pimonidazole, which is reduced under hypoxic conditions to hydroxylamine intermediates that irreversibly bind to nucleophilic groups in proteins or DNA (*19*, *20*). Pimonidazole staining revealed that for mice on a low-fat diet, the colonic epithelial surface was hypoxic, but hypoxia was eliminated in mice receiving a high-fat diet (Fig. 1, E and F). Consistent with the idea that saturated fatty acids act directly on the epithelium (fig. S2), prolonged high-fat intake also eliminated epithelial hypoxia (Fig. 1, G and H), diminished expression of mitochondrial markers in mRNA isolated from colonic epithelial cells (Fig. 1I), reduced ATP levels (Fig. 1J), and decreased levels of pyruvate dehydrogenase in the colonic epithelium (Fig. 1K) in germ-free mice. Taken together, this suggests that the mechanism triggering changes in epithelial oxygenation is microbiota independent.

Mice from The Jackson Laboratory remain *Enterobacteriaceae*-free because the vendor screens against the presence of this taxon using pathogen-free procedures (*15*). Because the niche of *Enterobacteriaceae* remains vacant in mice from The Jackson Laboratory, microbiota assembly can be completed by the designed engraftment of *E. coli* strains engineered to probe the contribution of predestined metabolic pathways to bacterial growth. We used this approach for precision editing of microbiota to determine the bioavailability of oxygen,

[1]Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [2]Department of Medical Microbiology and Immunology, School of Medicine, University of California at Davis, Davis, CA 95616, USA. [3]Agriculture Research Service (ARS-USDA), University of California at Davis, Davis, CA 95616, USA. [4]Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus Givat-Ram, Jerusalem 9190401, Israel. [5]Department of Biological Sciences, California State University, Sacramento, CA 95819, USA. [6]Department of Nutrition, University of California at Davis, Davis, CA 95616, USA. [7]Vanderbilt Institute for Infection, Immunology, and Inflammation, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [8]Vanderbilt Center for Immunobiology, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [9]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [10]Department of Medicine, Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA.
*Corresponding author. Email: mariana.x.byndloss@vumc.org (M.X.B.); jbaumler@ucdavis.edu (A.J.B.)

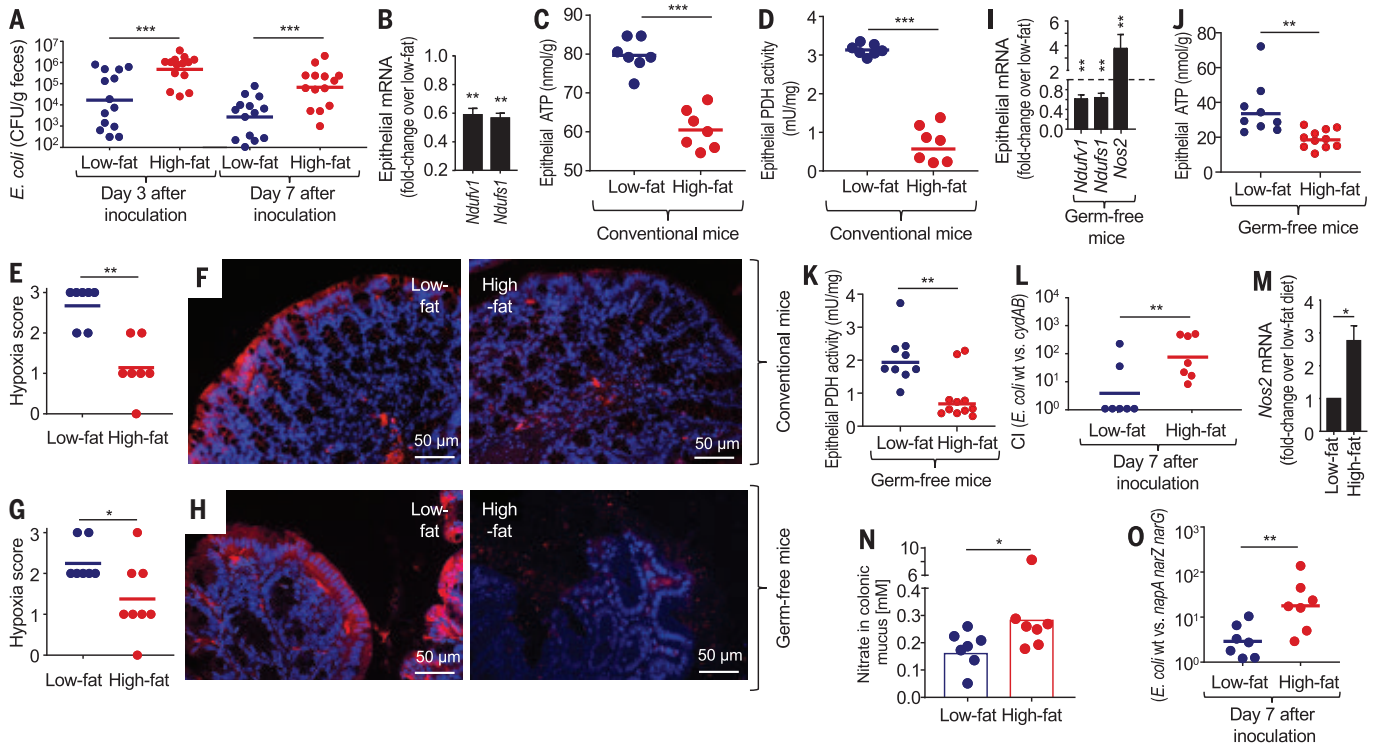



**Fig. 1. A high-fat diet changes epithelial physiology to increase the luminal availability of host-derived respiratory electron acceptors.** Mice were reared and maintained on a low- or high-fat diet. (**A**) Mice were inoculated with *E. coli* strain Nissle 1917, and colony-forming units (CFU) of *E. coli* Nissle 1917 in feces were determined at the indicated time points. (**B** to **K** and **M**) Preparations of colonic epithelial cells were used to isolate RNA and prepare cell lysates. (B) Fold change in mice on the high-fat diet compared with low-fat diet controls in epithelial transcripts was determined by quantitative real-time polymerase chain reaction (PCR) for genes encoding nicotinamide adenine dinucleotide and hydrogen (NADH):ubiquinone oxidoreductase core subunit V1 (*Ndufv1*) and NADH:ubiquinone oxidoreductase core subunit S1 (*Ndufs1*) ($n$ = 6 biological replicates). (C) Cytosolic concentrations of ATP. (D) Cytosolic concentrations of pyruvate dehydrogenase (PDH) activity. (E to H) Mice were injected with pimonidazole 1 hour before euthanasia. Binding of pimonidazole was detected with hypoxyprobe-1 primary antibody and a Cy-3 conjugated goat anti-mouse secondary antibody (red fluorescence) in the sections of proximal colon that were counterstained with DAPI (4′,6-diamidino-2-phenylindole) nuclear stain (blue fluorescence). (E) Pimonidazole staining was quantified by scoring blinded sections of proximal colon from conventional mice. (F) Representative images of colonic sections from conventional mice are shown. (G) Pimonidazole staining was quantified by scoring blinded sections of proximal colon from germ-free mice. (H) Representative images of colonic sections from germ-free mice are shown. (I) Epithelial transcripts of the indicated genes determined by quantitative real-time PCR in samples from germ-free mice ($n$ = 6 biological replicates). (J) Cytosolic concentrations of ATP in germ-free mice. (K) Cytosolic concentrations of PDH activity in germ-free mice. (**L**) Mice were inoculated with a 1:1 mixture of *E. coli* strain Nissle 1917 (wt); an isogenic *cydAB* mutant and CFU were determined at the indicated time point to calculate the competitive index (CI). (M) Fold change in epithelial *Nos2* transcripts was determined by quantitative real-time PCR. Bars represent geometric means ± geometric error ($n$ = 6). (**N**) Nitrate concentrations were determined in colonic mucus. (**O**) Mice were inoculated with a 1:1 mixture of *E. coli* strain Nissle 1917 (wt) and an isogenic *napA narG narZ* mutant and CFU were determined at the indicated time point to calculate the CI. (A, C to E, G, I, K, and L) Each dot represents data from one animal (biological replicate). *$P$ < 0.05; ** $P$ < 0.01; ***$P$ < 0.001 using an unpaired two-tailed Student's $t$ test [(A) to (D) and (I) to (O)] or a one-tailed Mann-Whitney test [(E) and (G)].

a factor linked to growth of *Enterobacteriaceae* (*14*, *21*–*23*). This was accomplished by comparing the fitness of the aerobic respiration-proficient *E. coli* strain Nissle 1917 with a genetically identical (isogenic) strain lacking cytochrome *bd*-II oxidase (*cydAB* mutant), an enzyme required for aerobic respiration under microaerophilic conditions (*14*). The use of these indicator strains to measure the bioavailability of oxygen has been validated in mouse models of antibiotic treatment and chemically induced colitis (*14*, *23*). Increased recovery of an aerobic respiration–proficient wild type (*E. coli* Nissle 1917) over the aerobic respiration–deficient mutant (*cydAB* mutant) supported the idea that a high-fat diet increased the bioavailability of oxygen in the intestinal lumen (Fig. 1L and fig. S1L).

Reduced mitochondrial activity in colonic epithelial cells is associated with increased *Nos2* expression (*14*), which was observed in mRNA isolated from the colonic epithelial cells of mice on a high-fat diet (Fig. 1, I and M). The *Nos2* gene encodes inducible nitric oxide synthase (iNOS), an enzyme that generates nitric oxide, which is converted into nitrate in the intestinal mucous layer (*17*). Consistent with this chain of events, a high-fat diet increased the nitrate concentration in the colonic mucus layer of both conventional (Fig. 1N) and germ-free (fig. S1M) mice. To investigate whether a high-fat diet–induced increase in luminal nitrate availability provided a nitrate respiration–mediated fitness advantage for facultative anaerobic *Enterobacteriaceae*, we compared growth of wild type *E. coli* Nissle 1917 and an isogenic mutant lacking nitrate reductase activity encoded by the *napFDAGHBC*, *narGHJI*, and *narZYWV* operons (*napA narG narZ* mutant) (fig. S3A). Inoculation of conventional mice with a 1:1 mixture of both strains resulted in increased recovery of the wild type (*E. coli* Nissle 1917) over a nitrate respiration-deficient mutant (*napA narG narZ* mutant), suggesting that nitrate respiration provided a growth advantage for *E. coli* in mice on a high-fat diet (Fig. 1O and fig. S1N). Collectively, these data indicated that the elevated abundance of

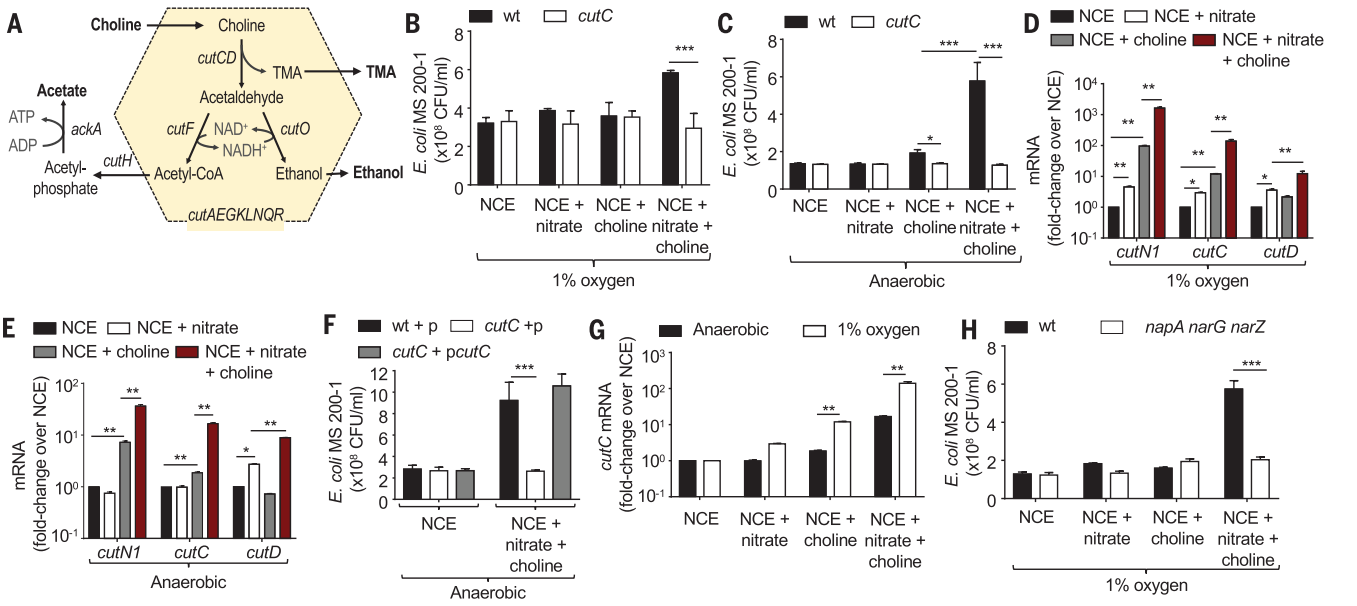



**Fig. 2. *E. coli* choline catabolism requires nitrate respiration in vitro.** (**A**) Schematic of choline catabolism encoded by the *cut* gene cluster of *E. coli* MS 200-1. (**B** and **C**) In vitro growth of *E. coli* MS 200-1 (wt) and an isogenic *cutC* mutant in no-carbon essential (NCE) medium supplemented with the indicated nutrients in a hypoxia chamber with 1% oxygen (B) or in an anaerobic chamber (C). (D and C) Expression of the indicated genes was determined by quantitative real-time PCR in RNA isolated from *E. coli* MS 200-1, which was grown in the indicated media in a hypoxia chamber with 1% oxygen (**D**) or in an anaerobic chamber (**E**). (**F**) In vitro growth in an anaerobic chamber of *E. coli* MS 200-1 carrying a cloning vector (wt+p), a *cutC* mutant carrying a cloning vector (*cutC*+p), and a *cutC* mutant complemented with the cloned *cutC* gene (*cutC*+p*cutC*) in NCE medium supplemented with the indicated nutrients. (**G**) Expression of *cutC* was determined by quantitative real-time PCR in RNA isolated from *E. coli* MS 200-1 grown under the indicated conditions. (**H**) In vitro growth of *E. coli* MS 200-1 (wt) and an isogenic *napA narG narZ* mutant in NCE medium supplemented with the indicated nutrients, in a hypoxia chamber with 1% oxygen. (B to H) Bars represent geometric means ± geometric error. $n = 4$ biological replicates (average of triplicate technical replicate per biological replicate). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ using an unpaired two-tailed Student's $t$ test [(B), (C), (G), and (H)] or a one-way analysis of variance (ANOVA) followed by Tukey's HSD test [(D) to (F)].

*E. coli* induced by a high-fat diet was driven by an increased bioavailability of host-derived respiratory electron acceptors, such as oxygen and nitrate, which fueled *E. coli* proliferation.

The *cut* gene cluster is present in various *Enterobacteriaceae* (*1*, *3*) and encodes a microcompartment thought to protect the bacterial cell from acetaldehyde, a toxic intermediate generated by choline TMA-lyase (Fig. 2A). Orthologous microcompartments are used for the breakdown of ethanolamine and 1,2-propanediol by some *Enterobacteriaceae*, but catabolism of these substrates in the mouse intestine requires the presence of respiratory electron acceptors such as tetrathionate, nitrate, or oxygen (*24*, *25*). Therefore, we wanted to investigate whether a high-fat diet–induced increase in the availability of respiratory electron acceptors would escalate the production of TMA in *E. coli* microcompartments. To this end, we used *E. coli* strain MS 200-1, which encodes the *cut* gene cluster involved in converting choline into TMA, acetate, and ethanol (*26*) (Fig. 2A). *E. coli* strain MS 200-1 was stably engrafted in mice from The Jackson Laboratory (C57BL6/J), resulting in fecal shedding at levels similar to those observed for a murine *E. coli* isolate (Mt1b1) or fecal shedding of endogenous *Enterobacteriaceae* from Charles River (C57BL6/NCrl) mice (fig. S4A). A high-fat diet increased fecal carriage of *E. coli* strain MS 200-1 (fig. S4B) by promoting growth with oxygen (fig. S4C) and nitrate (fig. S4D and fig. S3B) as electron acceptors. The *cutC* gene, encoding choline TMA-lyase, provided no growth benefit (Fig. 2B) or only a modest growth benefit (Fig. 2C) when *E. coli* strain MS 200-1 was cultured under conditions that mimicked the gut environment [i.e., growth in minimal medium supplemented with choline under microaerophilic (1% oxygen) or under anaerobic conditions, respectively]. Further, the presence of nitrate markedly enhanced *cutC*-dependent growth on choline as a carbon source (Fig. 2, B and C), which was associated with elevated expression of the *cut* genes (Fig. 2, D and E). Nitrate respiration-dependent growth of the *cutC* mutant on choline could be restored by introducing the cloned *cutC* gene on a plasmid (Fig. 2F and fig. S4E). The induction of *cut* gene expression by nitrate was further enhanced when *E. coli* was cultured under microaerobic conditions, compared with being cultured anaerobically (Fig. 2G and fig. S4, F and G). However, a nitrate respiration–deficient mutant (*E. coli* MS 200-1 *napA narG narZ* mutant) was unable to grow in minimal medium supplemented with choline and nitrate (Fig. 2H and fig. S4H), suggesting that growth on choline was dependent on nitrate respiration. These observations predicted that a high-fat diet–induced increase in the luminal concentration of host-derived nitrate would escalate choline catabolism by *E. coli* strain MS 200-1 in the digestive tract.

To test this idea in vivo, mouse diets were supplemented with 1% choline, because the high-fat diet used in previous experiments (Fig. 1) contained only trace amounts of this nutrient (<0.1%). Mice were reared on a choline-supplemented low-fat (10% fat) diet or on a choline-supplemented high-fat (60% fat) diet and then inoculated with *E. coli* strain MS 200-1. Rearing animals on a choline-supplemented high-fat diet was associated with increased weight gain (fig. S5A), reduced colon length (fig. S5B), low-grade mucosal inflammation (fig. S5C), an increased number of mitotic figures in the colonic epithelium (fig. S5D), loss of epithelial hypoxia (fig. S5, D and E), and elevated fecal shedding of *E. coli* (Fig. 3A). The choline-supplemented high-fat diet–induced proliferation of *E. coli* was respiration dependent, as it was no longer observed when the ability to respire nitrate or oxygen under microaerobic conditions was genetically ablated (Fig. 3, A to C). The presence of 1% choline in the diet provided a *cutC*-dependent growth advantage for *E. coli* in mice on a high-fat diet but not in animals on a low-fat diet (Fig. 3D), suggesting that changes in the gut environment induced

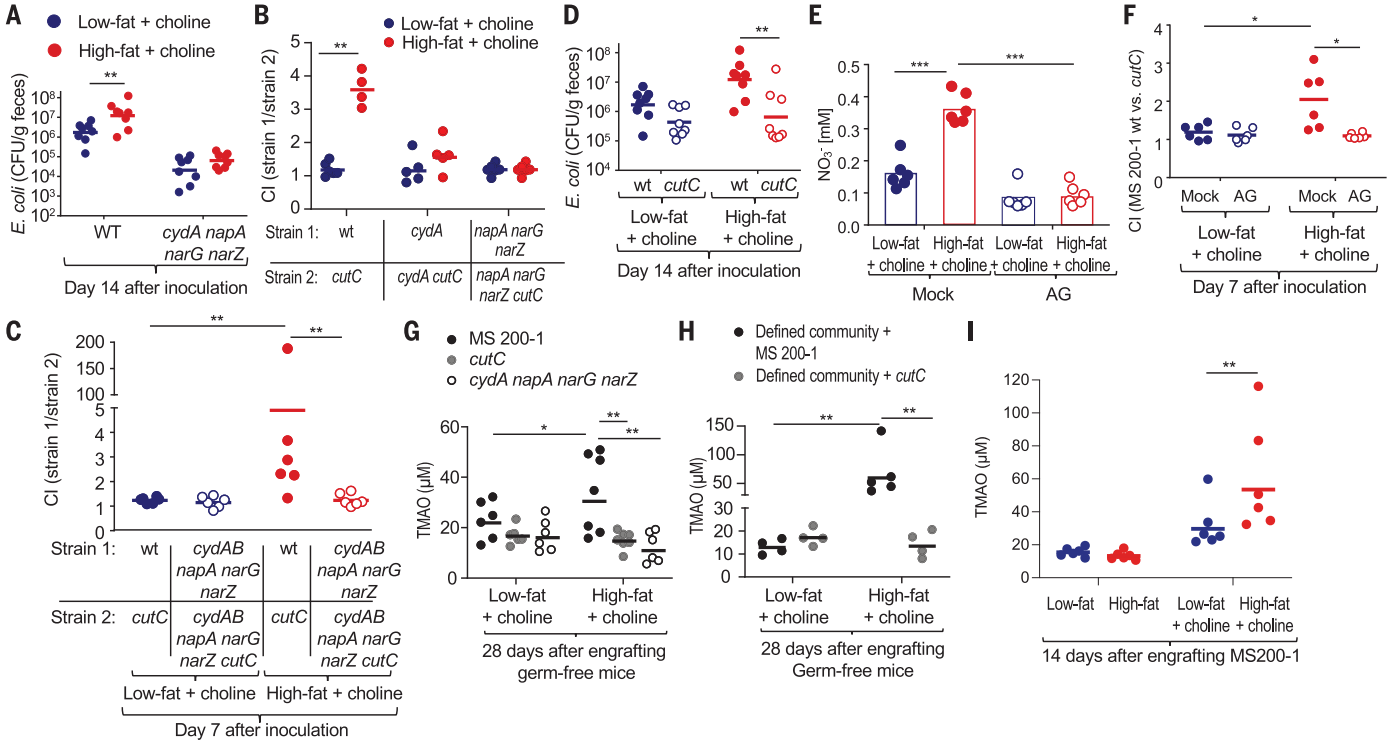



**Fig. 3. Host-derived electron acceptors license *E. coli* choline catabolism in vivo.** Mice were reared and maintained on a low-fat diet supplemented with 1% choline or a high-fat diet supplemented with 1% choline unless indicated otherwise. (**A**) Mice (C57BL/6J) were inoculated with *E. coli* strain MS 200-1 (wt) or an isogenic *cydA napA narG narZ* mutant, and CFU of *E. coli* in the feces were determined. (**B** and **C**) Mice (C57BL/6J) were inoculated with the indicated *E. coli* MS 200-1 strain mixtures, and the CI in the feces was determined 14 (B) or 7 (C) days later . (**D**) Mice (C57BL/6J) were inoculated with *E. coli* strain MS 200-1 (wt) or an isogenic *cutC* mutant, and CFU of *E. coli* in the feces were determined. (**E** and **F**) Mice (C57BL/6J) were mock-treated or received drinking water supplemented with the iNOS-inhibitor aminoguanidine (AG). (E) Nitrate concentrations were determined in colonic mucus. (F) Mice were inoculated with a 1:1 mixture of *E. coli* strain MS 200-1 (wt); an isogenic *cutC* mutant and CFU of each *E. coli* strain in the feces were determined to calculate the CI. (**G**) Germ-free (Swiss Webster) mice were mono-associated with the indicated *E. coli* strains, and TMAO levels in the plasma were determined 28 days later. (**H**) Germ-free (Swiss Webster) mice were engrafted with a defined microbial community containing either *E. coli* strain MS 200-1 or an isogenic *cutC* mutant. TMAO levels in the plasma were determined 28 days later. (**I**) Conventional (C57BL/6J) mice were engrafted with *E. coli* strain MS 200-1 and maintained on the indicated diet. TMAO levels in the plasma were determined 14 days later. (A to I) Each dot represents data from one animal (biological replicate). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ using an unpaired two-tailed Student's *t* test [(A) to (F), (H), and (I)] or a one-way ANOVA followed by Tukey's HSD test (G).

by a high-fat diet were necessary for *E. coli* to catabolize choline. To determine whether a high-fat diet–induced increase in the luminal nitrate concentration was required for the *cutC*-mediated growth advantage, mice were treated with aminoguanidine hydrochloride (AG), an inhibitor of the host enzyme iNOS. The choline-supplemented high-fat diet increased the nitrate concentration in the mucus layer, which was abrogated in mice receiving the AG treatment (Fig. 3E). The AG treatment eliminated the growth advantage conferred upon *E. coli* by *cutC*-mediated choline utilization (Fig. 3F), suggesting that the high-fat diet enhanced choline catabolism by elevating the availability of host-derived nitrate.

To investigate whether *E. coli* choline catabolism would alter circulating TMAO levels, germ-free mice were mono-associated with the *E. coli* MS 200-1 wild type, an isogenic *cutC* mutant, or an isogenic *cydA napA narG narZ* mutant (fig. S5G). TMAO plasma levels were significantly increased in mice mono-associated with the *E. coli* MS 200-1 wild type compared with the other groups, which supported our hypothesis that a high-fat diet escalates *E. coli* choline catabolism in vivo by supporting bacterial respiration (Fig. 3G). Next, we engrafted germ-free mice with a defined microbial community composed of *cutC*-negative bacteria (*Bacteroides thetaiotaomicron*, *Bacteroides caecimuris*, *Limosilactobacillus reuteri*, *Clostridium innocuum*, *Clostridioides mangenotti*, *Clostridium cochlearium*, *and Clostridium sporogenes*) (fig. S6) and either the *E. coli* MS 200-1 wild type or a *cutC* mutant (fig. S5H). TMAO plasma levels were significantly increased in mice engrafted with the defined microbial community containing a *cutC*-positive bacterium (*E. coli* MS 200-1 wild type) compared with the other groups (Fig. 3H), thus further corroborating the idea that a high-fat diet escalates *E. coli* choline catabolism in vivo. Since TMA producers are naturally present in conventional mouse microbiota (*27*), we wanted to determine whether engrafting mice with a single TMA-producing *E. coli* strain would measurably increase TMAO levels in the plasma during a high-fat diet. Compared with low-fat diet controls, inoculation of mice on the choline-supplemented high-fat diet with *E. coli* MS 200-1 resulted in an increased abundance of the *cutC* gene in the microbiota (*1*) (fig. S5I), an increased abundance of *E. coli* (fig. S5J), and a higher TMAO plasma level (Fig. 3I and fig. S5K). This increase in TMAO plasma levels was dependent on *E. coli* respiration, because it was no longer observed when mice were engrafted with an isogenic *cydA* mutant or an isogenic *napA narG narZ* mutant (fig. S5K). Collectively, these data suggested that an increased availability of host-derived electron acceptors promoted *E. coli* choline catabolism, so that this species was able to singlehandedly elevate levels of TMAO in the plasma of mice on the high-fat diet.

Finally, we wanted to determine whether treatment with drugs that reduce the availability of host-derived respiratory electron acceptors would lower circulating TMAO levels in mice

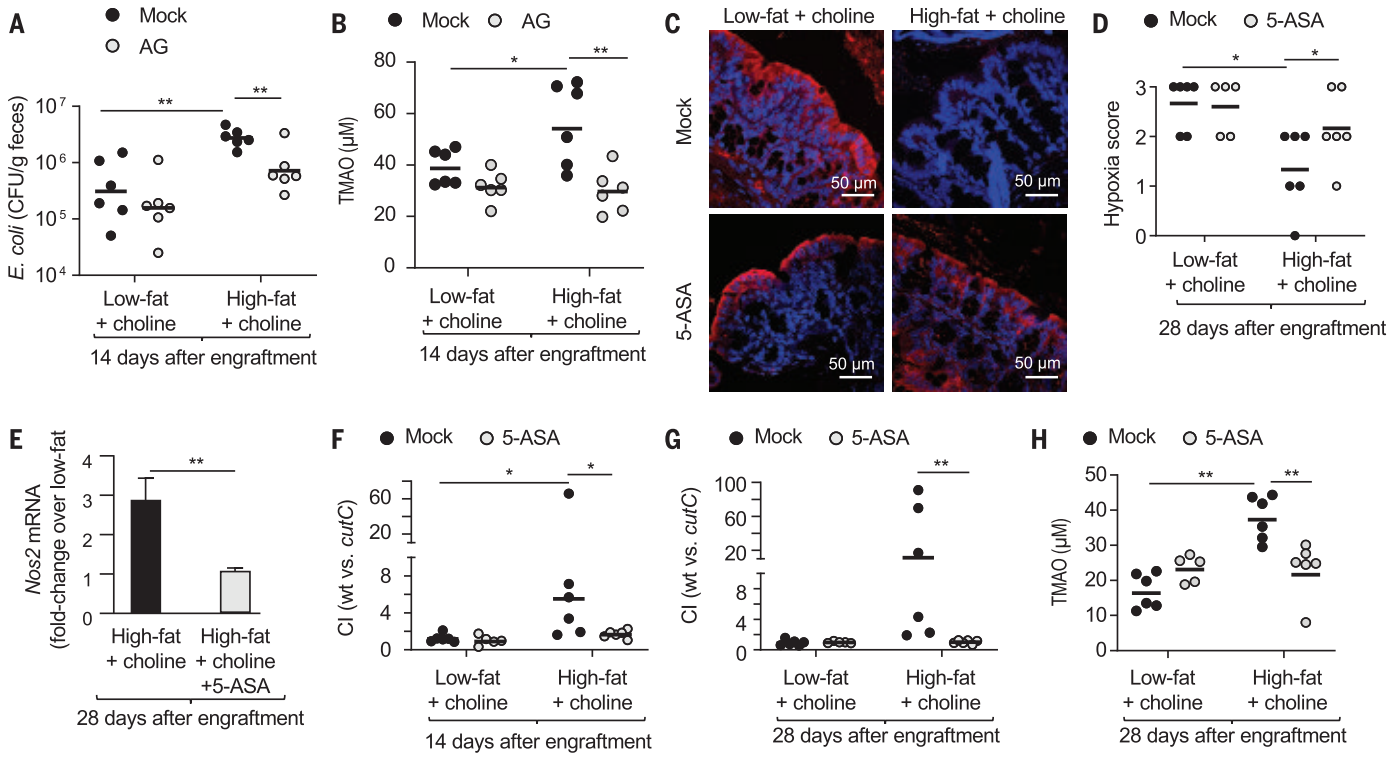



**Fig. 4. Restoring normal epithelial physiology blunts an *E. coli*–induced increase in circulating TMAO levels.** (**A** and **B**) Mice (C57BL/6J) maintained on a low- or high-fat diet supplemented with 1% choline were engrafted with *E. coli* MS 200-1 and received drinking water supplemented with AG or no supplementation (mock). (A) CFU of *E. coli* in the feces were determined. (B) TMAO levels in the plasma were determined. (**C** to **H**) Mice (C57BL/6J) maintained on a low- or high-fat diet supplemented with 1% choline were engrafted with a mixture of *E. coli* MS 200-1 wild type (wt) and *cutC* mutant, and received chow supplemented with 5-ASA or no 5-ASA supplementation (mock). (C and D) Mice were injected with pimonidazole one hour before euthanasia. Binding of pimonidazole was detected using hypoxyprobe-1 primary antibody and a Cy-3 conjugated goat anti-mouse secondary antibody (red fluorescence) in the sections of proximal colon that were counterstained with DAPI nuclear stain (blue fluorescence). (C) Representative images of colonic sections from conventional mice are shown. (D) Pimonidazole staining was quantified by scoring blinded sections of proximal colon from conventional mice. (E) Fold change in epithelial *Nos2* transcripts was determined by quantitative real-time PCR. Bars represent geometric means ± geometric error. (F and G) The CI was determined 14 (F) and 28 (G) days after engraftment. (H) TMAO levels in the plasma were determined. (A, B, D, and F to H) Each dot represents data from one animal (biological replicate).*$P$ < 0.05; **$P$ < 0.01 using an unpaired two-tailed Student's $t$ test [(A), (B), and (E) to (H)] or a one-tailed Mann-Whitney test (D).

engrafted with *E. coli* MS 200-1. Inhibition of host iNOS activity by AG treatment blunted growth of *E. coli* MS 200-1 in mice on the choline-supplemented high-fat diet (Fig. 4A) and reduced circulating TMAO levels (Fig. 4B). Next, we targeted peroxisome proliferator–activated receptor gamma (PPAR-γ), a nuclear receptor that maintains epithelial hypoxia and suppresses *Nos2* expression in the colonic epithelium (*14*). To this end, mice were treated with 5-aminosalicylic acid (5-ASA), a PPAR-γ agonist (*28*) that activates mitochondrial bioenergetics specifically in the intestinal epithelium (*29*). In mice on a choline-supplemented high-fat diet, 5-ASA treatment restored epithelial hypoxia (Fig. 4, C and D), and diminished *Nos2* expression in mRNA isolated from colonic epithelial cells (Fig. 4E). Furthermore, 5-ASA treatment of mice on a choline-supplemented high-fat diet abrogated the fitness advantage conferred by *cutC*-mediated choline catabolism in *E. coli* MS 200-1 (Fig. 4, F and G). The 5-ASA treatment blunted the increase in circulating TMAO in mice on the choline-supplemented high-fat diet (Fig. 4H).

Taken together, our data suggest that high-fat diet–induced low-grade mucosal inflammation is associated with diet-impaired mitochondrial bioenergetics in the colonic epithelium, thereby eliminating epithelial hypoxia. The consequent increase in the availability of respiratory electron acceptors, such as oxygen and nitrate, drives the proliferation of *E. coli* and other *Enterobacteriaceae*, explaining some of the changes in the composition of microbiota observed in individuals on a high-fat diet (*4*–*7*). Our results also suggest that *E. coli* strains carrying the *cut* operon catabolize choline only when high-fat diet–mediated changes in host physiology increase the luminal concentration of host-derived nitrate and oxygen. The emerging picture suggests that epithelial hypoxia ensures the microbiota remains beneficial by limiting the availability of respiratory electron acceptors, which limits *E. coli* choline catabolism. However, a prolonged high-fat diet impairs this host control mechanism, thereby elevating *Enterobacteriaceae* abundance while simultaneously escalating *E. coli* choline catabolism to elevate circulating TMAO levels. Dose response meta-analysis implicates this metabolite in increasing the relative risk for all-cause mortality in patients by 7.6% per each 10-μmol/liter increment of TMAO (*30*).

## REFERENCES AND NOTES


1. A. Martínez-del Campo *et al.*, *mBio* **6**, e00042–e15 (2015).
2. Z. Wang *et al.*, *Nature* **472**, 57–63 (2011).
3. Y. Zhu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4268–4273 (2014).
4. S. Devkota *et al.*, *Nature* **487**, 104–108 (2012).
5. N. Fei, L. Zhao, *ISME J.* **7**, 880–884 (2013).
6. M. Martinez-Medina *et al.*, *Gut* **63**, 116–124 (2014).
7. M. Anitha *et al.*, *Cell. Mol. Gastroenterol. Hepatol.* **2**, 328–339 (2016).
8. P. A. Kakimoto, F. K. Tamaki, A. R. Cardoso, S. R. Marana, A. J. Kowaltowski, *Redox Biol.* **4**, 375–380 (2015).
9. A. R. Cardoso, P. A. Kakimoto, A. J. Kowaltowski, *PLOS ONE* **8**, e77088 (2013).
10. C. J. Kelly *et al.*, *Cell Host Microbe* **17**, 662–671 (2015).
11. Y. Litvak, M. X. Byndloss, A. J. Bäumler, Colonocyte metabolism shapes the gut microbiota. *Science* **362**, eaat9076 (2018).

12. Y. Litvak, A. J. Bäumler, *Immunity* **51**, 214–224 (2019).
13. C. R. Tiffany, A. J. Bäumler, *Am. J. Physiol. Gastrointest. Liver Physiol.* **317**, G602–G608 (2019).
14. M. X. Byndloss *et al*., *Science* **357**, 570–575 (2017).
15. E. M. Velazquez *et al*., *Nat. Microbiol.* **4**, 1057–1064 (2019).
16. A. Nissle, *DMW-Deutsche Medizinische Wochenschrift* **51**, 1809–1813 (1925).
17. S. E. Winter *et al*., *Science* **339**, 708–711 (2013).
18. M. Gulhane *et al*., *Sci. Rep.* **6**, 28990 (2016).
19. N. Terada, N. Ohno, S. Saitoh, S. Ohno, *Histochem. Cell Biol.* **128**, 253–261 (2007).
20. S. Kizaka-Kondoh, H. Konse-Nagasawa, *Cancer Sci.* **100**, 1366–1373 (2009).
21. F. Rivera-Chávez *et al*., *Cell Host Microbe* **19**, 443–454 (2016).
22. C. A. Lopez *et al*., *Science* **353**, 1249–1253 (2016).
23. S. A. Cevallos *et al*., *mBio* **10**, e02244–e19 (2019).
24. P. Thiennimitr *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17480–17485 (2011).
25. F. Faber *et al*., *PLOS Pathog.* **13**, e1006129 (2017).
26. S. Craciun, E. P. Balskus, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21307–21312 (2012).
27. Z. Wang *et al*., *Cell* **163**, 1585–1595 (2015).
28. C. Rousseaux *et al*., *Carcinogenesis* **34**, 2580–2586 (2013).
29. S. A. Cevallos *et al*., *mBio* **12**, e03227–e20 (2021).
30. G. G. Schiattarella *et al*., *Eur. Heart J.* **38**, 2948–2956 (2017).

### ACKNOWLEDGMENTS

We thank E. Balkus for providing *E. coli* strain MS 200-1 and A. Goodman for providing *Bacteroides thetaiotaomicron* strain VPI-5482. **Funding:** W.Y. was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) by the Ministry of Education 2020R1A6A3A03037326. Y.L. was supported by Vaadia-BARD Postdoctoral Fellowship FI-505-2014. E.E.O. was supported by Public Health Service Grant TR001861. C.S. was supported by the Dorothy Beryl and Theodore Roe Austin Pathology Research Fund and T32AI112541. N.J.F. was supported by T32DK007673-07. N.G.S. was supported by T32ES007028-46. E.G. and B.J.B. were supported by the U.S. Department of Agriculture (USDA) Project 2032-51530-025-00D. Work in A.J.B.'s laboratory was supported by USDA/NIFA award 2015-67015-22930; by Crohn's and Colitis Foundation of America Senior Investigator Award 650976; and by Public Health Service Grants AI044170, AI096528, AI112445, AI112949, AI146432, and AI153069. Work in M.X.B.'s laboratory was supported by V Scholar grant V2020-013 from The V Foundation for Cancer Research, Vanderbilt Digestive Disease Pilot and Feasibility grant P30 058404, ACS Institutional Research Grant IRG-19-139-59, VICC GI SPORE grant P50CA236733, United States-Israel Binational Science Foundation grant 2019136, and Vanderbilt Institute for Clinical and Translational Research Grant VR53102 and VR54267. **Author contributions:** W.Y., A.J.B., and M.X.B. conceptualized this study. W.Y., J.K.Z., N.J.F., T.P.T., C.D.S., N.G.S., A.J.B., S.A.C., E.G., C.R.T., J.D.T., Y.L., H.N., E.E.O., A.S.M., and M.X.B. performed the experiments. W.Y., N.J.F., C.D.S., E.G., B.J.B., J.C.R., A.S.M., and M.X.B. analyzed the data. A.J.B. and M.X.B. wrote the manuscript and all authors reviewed it. A.J.B. and M.X.B. coordinated and acquired funding for the work. **Competing interests:** J.C.R. is a founder, scientific advisory board member, and stockholder of Sitryx Therapeutics as well as a member of the scientific advisory boards of Caribou Biosciences and Nirogy Therapeutics. J.C.R. has also consulted for Merck and Pfizer within the past 3 years and has received research support from Incyte Corp., Calithera Biosciences, and Tempest Therapeutics. All other authors declare no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials.

### SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/813/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S9
References (*31*–*40*)
MDAR Reproducibility Checklist

26 November 2019; accepted 29 June 2021
10.1126/science.aba3683

CORONAVIRUS

# Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants

**Meng Yuan[1]†, Deli Huang[2]†, Chang-Chun D. Lee[1]†, Nicholas C. Wu[3,4]†, Abigail M. Jackson[1], Xueyong Zhu[1], Hejun Liu[1], Linghang Peng[2], Marit J. van Gils[5], Rogier W. Sanders[5,6], Dennis R. Burton[2,7], S. Momsen Reincke[8,9], Harald Prüss[8,9], Jakob Kreye[8,9], David Nemazee[2], Andrew B. Ward[1], Ian A. Wilson[1,10]***

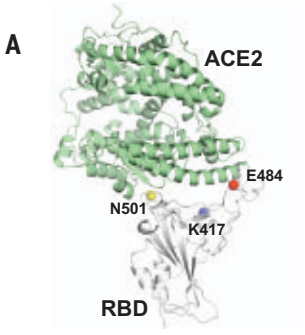
**Neutralizing antibodies (nAbs) elicited against the receptor binding site (RBS) of the spike protein of wild-type severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) are generally less effective against recent variants of concern. RBS residues Glu[484], Lys[417], and Asn[501] are mutated in variants first described in South Africa (B.1.351) and Brazil (P.1). We analyzed their effects on angiotensin-converting enzyme 2 binding, as well as the effects of two of these mutations (K417N and E484K) on nAbs isolated from COVID-19 patients. Binding and neutralization of the two most frequently elicited antibody families (IGHV3-53/3-66 and IGHV1-2), which can both bind the RBS in alternative binding modes, are abrogated by K417N, E484K, or both. These effects can be structurally explained by their extensive interactions with RBS nAbs. However, nAbs to the more conserved, cross-neutralizing CR3022 and S309 sites were largely unaffected. The results have implications for next-generation vaccines and antibody therapies.**

The COVID-19 pandemic has already lasted for more than a year, but new infections are still escalating throughout the world. Although several different COVID-19 vaccines have been deployed globally, a major concern is the emergence of antigenically distinct severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern (VOCs). In particular, the B.1.1.7 lineage that arose in the UK (*1*) and quickly became dominant, the B.1.351 (also known as 501Y.V2) lineage in South Africa (*2*), the B.1.1.28 lineage (and its descendant B.1.1.28.1, also known as P.1/501Y.V3) in Brazil (*3*), and B.1.232/B.1.427/B.1.429 (also called CAL.20C and CAL.20A) in the United States (*4*) have raised serious questions about the nature, extent, and consequences of antigenic drift in SARS-CoV-2. In the receptor binding site (RBS) of the spike (S) protein receptor-binding domain (RBD), the B.1.1.7 lineage has acquired an Asn[501] → Tyr (N501Y) mutation, the B.1.351 and P.1 lineages share this mutation along with Lys[417] → Asn/Thr (K417N/T) and Glu[484] → Lys (E484K), whereas the California variants have a Leu[452] → Arg mutation that is also present in the B.1.617 variant, which was first isolated in India, with Glu[484] → Gln (*5*). E484K has also been detected in a few B.1.1.7 genomes (*1*) (Fig. 1A). We therefore investigated the structural and functional consequences of such mutations on neutralizing antibodies (nAbs) isolated from COVID-19 convalescent patients, as well as their effect on angiotensin-converting enzyme 2 (ACE2) receptor binding.

N501Y was previously reported to enhance binding to the human ACE2 receptor (*6*, *7*). Here, we quantified binding of K417N, E484K, N501Y, and double and triple combinations in the RBD to ACE2 by biolayer interferometry (Fig. 1A and fig. S1). N501Y indeed increased RBD binding to ACE2 relative to wild-type RBD (binding affinity $K_D$ = 3.3 nM versus 7.0 nM), whereas K417N substantially reduced ACE2 binding (41.6 nM). E484K slightly reduced binding (11.3 nM). N501Y could rescue binding of K417N (9.0 nM), and the triple mutant K417N/E484K/N501Y (as in B.1.351) had similar binding (6.5 nM) to the wild type (Fig. 1A and fig. S1). Consistently, K417N/T mutations are associated with N501Y in naturally circulating SARS-CoV-2. Among 585,054 SARS-CoV-2 genome sequences in the GISAID database (5 March 2021) (*8*), about 95% of K417N/T mutations occur with N501Y, despite N501Y being present in only 21% of all analyzed sequences. In contrast, only 36% of E484K mutations occur with N501Y.

We and others have shown that most SARS-CoV-2 nAbs that target the RBD and their epitopes can be classified into different sites and subsites (*9*–*12*). Certain IGHV genes are highly

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA. [3]Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [4]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [5]Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centers, Location AMC, University of Amsterdam, Amsterdam, Netherlands. [6]Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY 10021, USA. [7]Ragon Institute of MGH, Harvard, and MIT, Cambridge, MA 02139, USA. [8]German Center for Neurodegenerative Diseases (DZNE) Berlin, Berlin, Germany. [9]Department of Neurology and Experimental Neurology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität Berlin, and Berlin Institute of Health, Berlin, Germany. [10]Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.
***Corresponding author. Email: wilson@scripps.edu**
†These authors contributed equally to this work.




| hACE2 vs. RBD | Binding $K_D$ (nM) |
|---|---|
| wild type | 7.0 ± 0.1 |
| K417N | 41.6 ± 0.5 |
| E484K | 11.3 ± 0.2 |
| N501Y | 3.3 ± 0.1 |
| K417N / N501Y | 9.0 ± 0.1 |
| E484K / N501Y | 3.0 ± 0.1 |
| K417N / E484K / N501Y | 6.5 ± 0.1 |

C

| Antibody | VH germline | Epitope [categorized in (9)] | Binding class [categorized in (10)] | Neutralization effect ($IC_{50}$, mutant vs. WT) | | Effect on binding affinity (mutant vs. WT) | | Ref. |
|---|---|---|---|---|---|---|---|---|
| | | | | K417N | E484K | K417N | E484K | |
| CC12.1 | VH3-53 | RBS-A | Class 1 | × | - | × | - | (11, 13) |
| CC12.3 | VH3-53 | RBS-A | Class 1 | × | - | × | - | (11, 13) |
| COVA2-04 | VH3-53 | RBS-A | Class 1 | × | - | × | - | (20, 75) |
| COVA2-07 | VH3-53 | N.S. | N.S. | × | - | × | - | (20) |
| REGN10933 | VH3-11 | RBS-B | N.C. | + | + | + | + | (28) |
| COVA2-39 | VH3-53 | RBS-B | Class 2 | - | × | - | × | (20, 75) |
| C144 | VH3-53 | RBS-B | Class 2 | - | × | - | × | (10, 15) |
| C121 | VH1-2 | RBS-B | Class 2 | - | × | - | × | (10, 15) |
| CV05-163 | VH1-2 | RBS-B | Class 2 | - | × | + | × | (17) |
| CV07-250 | VH1-18 | RBS-B | N.C. | - | - | - | ++ | (17) |
| CV07-270 | VH3-11 | RBS-C | N.C. | - | × | - | × | (17) |
| REGN10987 | VH3-30 | RBS-D | Class 3 | - | - | - | - | (28) |
| COVA2-15 | VH3-23 | RBS-D | N.C. | - | ++ | - | ++ | (20) |
| COVA1-16 | VH1-46 | CR3022 | Class 4 | - | - | - | - | (20, 80) |
| C135 | VH3-30 | S309 | Class 3 | - | - | - | - | (10, 15) |
| CV38-142 | VH5-51 | S309 | Class 3 | - | - | - | - | (20, 81) |
| REGN10933+ REGN10987 | VH3-11 VH3-30 | RBS-B, RBS-D | N.C. Class 3 | - | - | - | - | (28) |
| CC6.29 | VH7-4-1 | N.S. | N.S. | - | ++ | - | - | (11) |

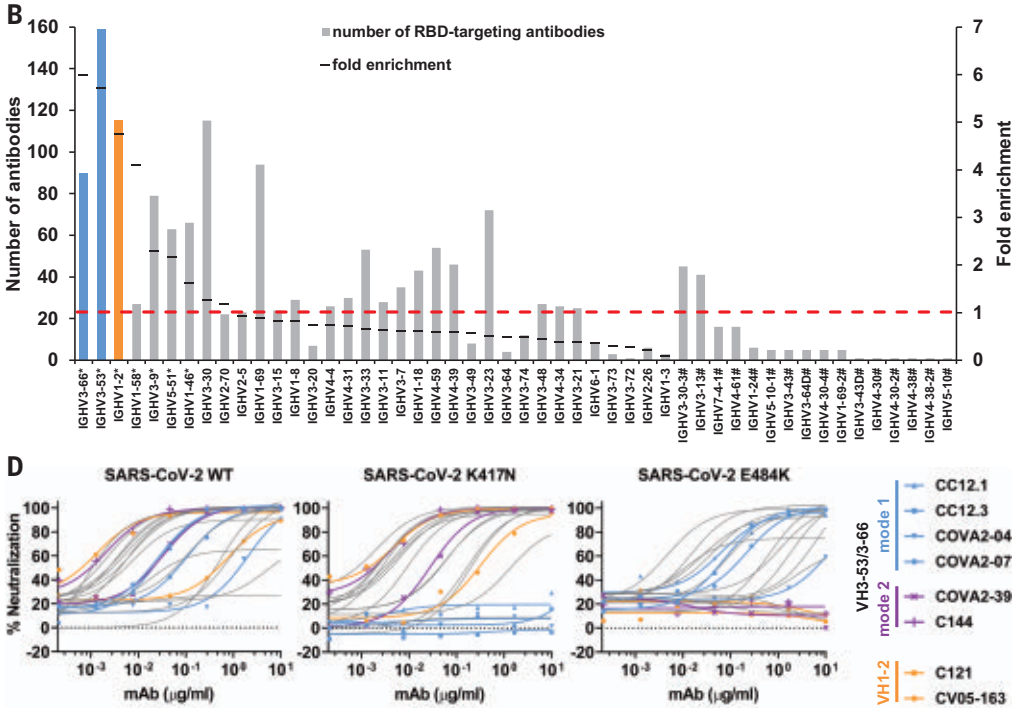



**Fig. 1. Emergent SARS-CoV-2 variants escape two major classes of neutralizing antibodies.** (**A**) Emergent mutations (spheres) in the RBS of B.1.351 and P.1 lineages are mapped onto a structure of SARS-CoV-2 RBD (white) in complex with ACE2 (green) (PDB 6M0J) (*90*). Binding affinities of Fc-tagged human ACE2 against SARS-CoV-2 RBD wild type and mutants were assayed by biolayer interferometry (BLI) experiments. Detailed sensorgrams are shown in fig. S1. (**B**) Distribution of IGHV gene usage. Numbers of RBD-targeting antibodies encoded by each IGHV gene are shown. The frequently used IGHV3-53 and IGHV3-66 genes are highlighted in blue, and IGHV1-2 in orange. The IGHV gene usage in 1593 SARS-CoV-2 RBD-targeting antibodies (*11*, *14*–*44*) relative to healthy individuals (baseline) (*76*) (fold enrichment) is shown as black line segments. Hashtags (#) denote IGHV gene frequencies in healthy individuals that were not reported in (*76*). Asterisks denote IGHV genes that are significantly enriched over the baseline repertoire (*76*) ($P < 0.05$, one-sample proportion test with Bonferroni correction). A fold enrichment of 1 (red dashed line) represents no difference over baseline. (**C**) Effects of single mutations on the neutralization activity and binding affinity of each neutralizing antibody. "–" denotes an increase in half-maximal inhibitory concentration ($IC_{50}$) or $K_D$ by a factor of <10; "+", by a factor of 10 to 100; and "++", by a factor of >100. Results in red with "×" indicate that no neutralization activity or binding was detected at the highest amount of immunoglobulin G used. N.C., not categorized in the original studies; N.S., no structure available. (**D**) Neutralization of pseudotyped SARS-CoV-2 virus and variants carrying K417N or E484K mutations. We tested a panel of 17 neutralizing antibodies, including four mode-1 IGHV3-53 antibodies (blue), two mode-2 IGHV3-53 antibodies (purple), and two IGHV1-2 antibodies (orange). The discrepancy between CV05-163 neutralizing SARS-CoV-2 pseudotyped virus ($IC_{50}$ = 0.47 μg/ml) and authentic virus ($IC_{50}$ = 0.02 μg/ml) reported in our previous study (*17*) is possibly due to different systems (pseudovirus versus authentic virus) and host cells (HeLa cells versus Vero E6 cells) used in these experiments.

enriched in the antibody response to SARS-CoV-2 infection, with IGHV3-53 (*11*, *13*–*16*) and IGHV3-66, which differ by only one conservative substitution (V12I), and IGHV1-2 (*11*, *17*, *18*) being the most enriched IGHV genes used among 1593 RBD-targeting antibodies from 32 studies (*11*, *14*–*44*) (Fig. 1B). We investigated the effects of the prevalent SARS-CoV-2 mutations on neutralization by these major classes of antibodies found in multiple donors, and the consequences for current vaccines and therapeutics.

K417N and E484K in VOCs B.1.351 and P.1 have been reported to decrease the neutralizing activity of sera as well as neutralizing monoclonal antibodies (mAbs) isolated from COVID-19 convalescent plasma and vaccinated individuals (*45*–*62*). Relative to wild-type SARS-CoV-2, B.1.351 is more resistant to neutralization by convalescent plasma and mRNA vaccine sera by a factor of ~8 to 14, and P.1 is more resistant by a factor of ~2.6 to 5 (*63*–*68*). Some variants are able to escape neutralization by some nAbs (e.g., LY-CoV555, 910-30, COVOX-384, S2H58, C671, etc.), whereas others retain activity (e.g., 1-57, 2-7, mAb-222, S309, S2E12, COV2-2196, C669, etc.) (*50*, *57*, *63*, *66*, *69*, *70*). Here, we selected a representative panel of 17 human nAbs isolated from COVID-19 patients or humanized mice to study the escape mechanism. These nAbs cover all known neutralizing sites on the RBD and include those encoded by V genes that are the most frequently used and also significantly enriched (Fig. 1B). We tested the activity of a panel of nAbs against wild-type (Wuhan strain) SARS-CoV-2 pseudovirus and single mutants K417N and E484K (Fig. 1C). Binding and neutralization of four and five antibodies out of the 17 tested were abolished by K417N and E484K, respectively. Strikingly, binding and neutralization by all six highly potent IGHV3-53 antibodies (*71*) that we tested were abrogated by either K417N (RBS-A/class 1) or E484K (RBS-B/class 2) (Fig. 1, C and D, and fig. S2). In addition, binding and neutralization of IGHV1-2 antibodies was severely reduced for the E484K mutation (Fig. 1, C and D, and fig. S2).

We next examined 54 SARS-CoV-2 RBD-targeting human antibodies with available structures. The antibody epitopes on the RBD can be classified into six sites—four RBS subsites, RBS-A, B, C, and D; CR3022 site; and S309 site (fig. S3) (*72*)—that are generally related to the four classes assigned in (*10*) (Fig. 1C). Of 23 IGHV3-53/3-66 antibodies, 21 target RBS-A (Fig. 2A). All IGHV1-2 antibodies with known structures bind to the RBS-B epitope. A large fraction of antibodies in these two main families make contact with $Lys^{417}$, $Glu^{484}$, or $Asn^{501}$ in their epitopes (Fig. 2A) (*73*). Almost all RBS-A antibodies interact extensively with $Lys^{417}$ and $Asn^{501}$, whereas most RBS-B and RBS-C antibodies contact $Glu^{484}$, and most RBS-C antibodies interact with $Leu^{452}$. We also examined the buried surface area (BSA) of $Lys^{417}$, $Glu^{484}$, and $Asn^{501}$ upon interaction with these RBD-targeting antibodies (fig. S3C). The extensive BSA confirmed why mutations at positions 417 and 484 affect binding and neutralization. Antibodies targeting RBS-D, or the cross-neutralizing S309 and CR3022 sites, are minimally or not involved in interactions with these four RBD mutations (Fig. 2A and fig. S3C).

IGHV3-53/3-66 RBD antibodies can adopt two different binding modes (*9*, *10*), which we refer here to as binding modes 1 and 2 (*74*), with distinct epitopes and approach angles (Fig. 2B and fig. S4). All IGHV3-53/3-66 RBD antibodies to date with binding mode 1 have a short, heavy-chain complementarity-determining region 3 (CDR H3) of <15 amino acids (Kabat numbering) and bind RBS-A (*13*, *16*, *32*), whereas those with binding mode 2 have a longer CDR H3 (≥15 amino acids) and target RBS-B (*9*, *10*, *75*). These dual binding modes enhance recognition of this antibody family for the SARS-CoV-2 RBD, although most IGHV3-53/3-66 RBD antibodies adopt binding mode 1 (Fig. 2 and fig. S4). $Lys^{417}$ is a key epitope residue for antibodies with IGHV3-53/3-66 binding mode 1 (Fig. 2B and fig. S4). IGHV3-53 germline residues $V_H$ $Tyr^{33}$ and $Tyr^{52}$ make hydrophobic interactions with the aliphatic moiety of $Lys^{417}$, and its ε-amino group interacts with CDR H3 through a salt bridge ($Asp^{97}$ or $Glu^{97}$), hydrogen bond (H-bond), or cation-π interaction ($Phe^{99}$) (Fig. 2B). K417N/T would diminish such interactions and therefore affect antibody binding and neutralization, providing

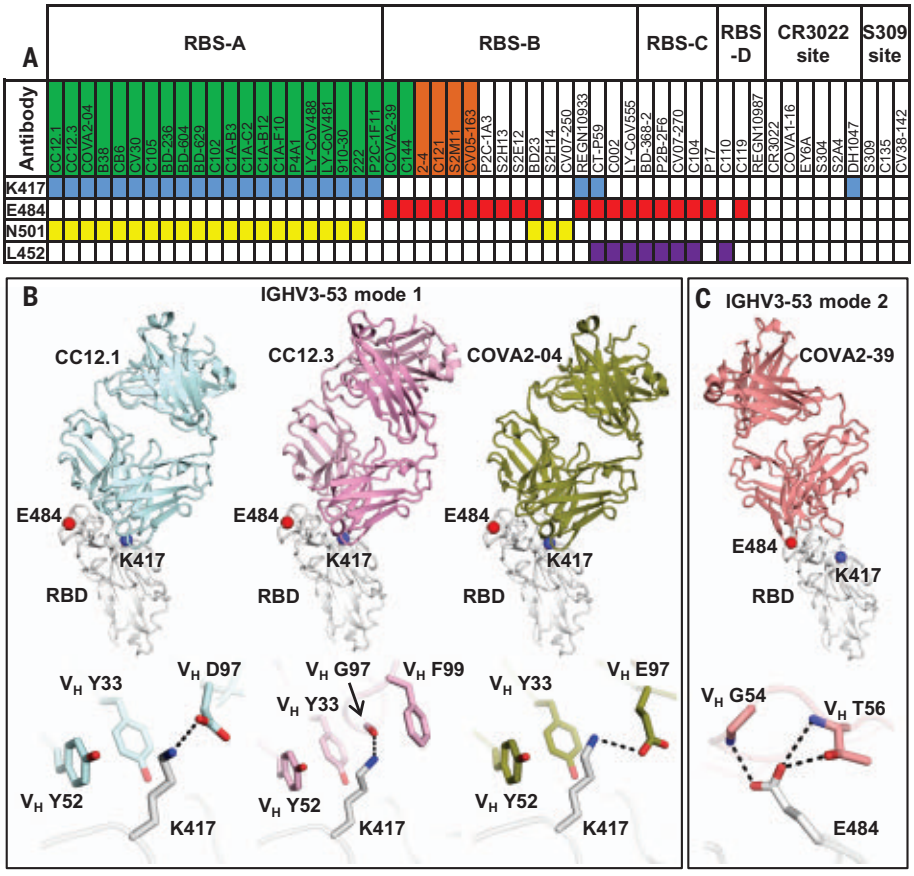



**Fig. 2. Antibody binding to the SARS-CoV-2 RBS.** (**A**) Antibodies making contact with RBD residues $Lys^{417}$, $Glu^{484}$, and $Asn^{501}$ are represented by blue, red, and yellow boxes, respectively (cutoff distance = 4 Å). Antibodies encoded by the most frequently elicited IGHV3-53/3-66 and IGHV1-2 in convalescent patients are shown in green and orange boxes, respectively. Antibodies are ordered by epitopes originally classified in (*9*) with an additional epitope, RBS-D, that maps to a region in the RBS above or slightly overlapping with the S309 site. Details of the epitope classifications are shown in fig. S3A. Structures of RBD-targeting antibodies that were isolated from patients are analyzed (*91*). (**B** and **C**) Residues that are mutated in recently circulating variants are integral to the binding sites of IGHV3-53 antibodies. Representative structures are shown for IGHV3-53 binding mode 1 [CC12.1 (PDB 6XC3), CC12.3 (PDB 6XC4) (*13*), and COVA2-04 (PDB 7JMO) (*75*)] (B) and binding mode 2 [COVA2-39 (PDB 7JMP) (*75*)] (C). The SARS-CoV-2 RBD is in white and Fabs in different colors. Residues $Lys^{417}$ and $Glu^{484}$ are represented by blue and red spheres, respectively. Hydrogen bonds and salt bridges are represented by black dashed lines.

a structural explanation for K417N escape in IGHV3-53/3-66 antibodies with binding mode 1 (Fig. 1, C and D, Fig. 2B, and fig. S2). In contrast, IGHV3-53 antibodies with binding mode 2 do not interact with RBD-Lys$^{417}$ (fig. S4) but with Glu$^{484}$ through H-bonds with CDR H2 (Fig. 2C). Consistently, binding and neutralization of IGHV3-53 antibodies with binding mode 2 (fig. S4) are abolished by E484K but not K417N (Fig. 1, C and D, and fig. S2). Interestingly, unlike most IGHV3-53 antibodies that are sensitive to K417N/T or E484K, a recently discovered IGHV3-53-encoded mAb-222, which binds RBS, retains activity against P.1 and B.1.351. The mAb-222 light chain could largely restore the neutralization potency of other IGHV3-53 antibodies, which suggests that light-chain interactions can compensate for loss of binding of K417N/T by the heavy chain (*63*). However, this antibody may represent only a small portion of IGHV3-55/3-66 antibodies that can neutralize VOCs.

Among the IGHV genes used in RBD antibodies, IGHV1-2 is also highly enriched over the baseline frequency in the antibody repertoire of healthy individuals (*76*) and is second only to IGHV3-53/3-66 (Fig. 1B). We compared three structures of IGHV1-2 antibodies, namely 2-4 (*27*), S2M11 (*30*), and C121 (*10*), that target RBS-B. Despite being encoded by different IGK(L)V genes, 2-4 (IGLV2-8), S2M11 (IGKV3-20), and C121 (IGLV2-23) share a nearly identical binding mode and epitope (Fig. 3A). Structural analysis reveals that the $V_H$ Gly$^{26}$-Tyr-Thr-Phe-Thr-Gly-(Tyr)-Tyr$^{33}$, Trp$^{50}$-(Ile)-Asn/Ser-(Pro)-X-Ser-X-Gly-Thr$^{57}$, and Thr$^{73}$-Ser-(Ile)-Ser/Thr$^{76}$ motifs are important for RBD binding (fig. S5, A to D). Although only a small part of the epitope interacts with the light chains of 2-4, S2M11, and C121, $V_L$ residues 32 and 91 (and also residue 30 in some antibodies) play an important role in forming a hydrophobic pocket together with $V_H$ residues for binding RBD-Phe$^{486}$, which is another key binding residue in such classes of antibodies (*9*) (fig. S5, E to I). Three other IGHV1-2 antibodies, 2-43, 2-15, and H4, also bind in a similar mode (*77*), further highlighting the structural convergence of IGHV1-2 antibodies in targeting the same RBD epitope. All IGHV1-2 antibodies to date form extensive interactions with Glu$^{484}$ (Fig. 3A and fig. S3C). In particular, germline-encoded $V_H$ Tyr$^{33}$, Asn$^{52}$ (somatically mutated to Ser$^{52}$ in C121), and Ser$^{54}$ are involved in polar interactions with the RBD-Glu$^{484}$ side chain that would be altered by substitution with Lys (Fig. 3A) and thereby diminish binding and neutralization of IGHV1-2 antibodies against E484K (Fig. 1, C and D, and fig. S2).

We previously isolated another potent IGHV1-2 antibody, CV05-163, targeting the SARS-CoV-2 RBD (fig. S6) from a COVID-19 patient (*17*). CV05-163 likely represents a shared antibody response for IGHV1-2 RBD antibodies across patients (fig. S7). Negative-stain electron microscopy (nsEM) of CV05-163 in complex with the SARS-CoV-2 S trimer illustrates that it can bind in various stoichiometries, including molar ratios of 1:1, 2:1, and 3:1 (Fab to S protein trimer), and can accommodate RBDs in both up and down conformations (fig. S8). We also determined a crystal structure of Fab CV05-163 with SARS-CoV-2 RBD and Fab CR3022 to 2.25 Å resolution (Fig. 3B, figs. S9 to S11, and tables S1 and S2) and found that it does indeed bind RBS-B (Fig. 3) and makes extensive interactions with Glu$^{484}$ through H-bonds ($V_H$ Trp$^{50}$ and $V_H$ Asn$^{58}$) and a salt bridge ($V_L$ Arg$^{91}$) (Fig. 3B); these interactions explain why CV05-163 binding and neutralization were diminished with E484K (Fig. 1, C and D, and fig. S2). However, CV05-163 is rotated 90° (Fig. 3B) relative to IGHV1-2 antibodies 2-4, S2M11, and C121 (Fig. 3A). Thus, IGHV1-2 antibodies, akin to IGHV3-53/66 (*75*), can engage the RBD in two different binding modes, both of which are susceptible to escape by E484K but not by K417N (Fig. 1, C and D).

A further group of antibodies target the back side of the RBS ridge (RBS-C) (*9*). To date, five nAbs isolated from COVID-19 patients are known to bind RBS-C: CV07-270 (*17*), BD-368-2 (*38*), P2B-2F6 (*18*), C104 (*10*), and P17 (*78*). These RBS-C nAbs also interact with Glu$^{484}$ (Fig. 4A), mainly through an arginine in CDR H3, which suggests that E484K may have an adverse impact on RBS-C antibodies. Indeed, binding and neutralization by CV07-270 was abrogated by E484K (Fig. 1C and fig. S2). Intriguingly, these five RBS-C antibodies are encoded by five different IGHV genes (*79*), but they all target a similar epitope with similar angles of approach. In addition, neutralization by REGN10933, a potent antibody used for therapeutic treatment, was reduced to a lesser extent by K417N and E484K (Fig. 1C) (*28*). REGN10933 binds at a slightly different angle from RBS-A antibodies and other RBS-B antibodies. Lys$^{417}$ then interacts with CDRs H1 and H3 of REGN10933, whereas Glu$^{484}$ contacts CDR H2 (fig. S12). Overall, our results demonstrate that RBS mutations K417N and E484K can either abolish or extensively reduce the binding and neutralization of several major classes of SARS-CoV-2 RBD antibodies.

Two other non-RBS sites that are distant from Lys$^{417}$ and Glu$^{484}$ have been repeatedly shown to be neutralizing sites on the SARS-CoV-2 RBD, namely the CR3022 cryptic site and the S309 proteoglycan site (*9*) (Fig. 1C and Fig. 4B). Antibodies from COVID-19 patients can neutralize SARS-CoV-2 by targeting

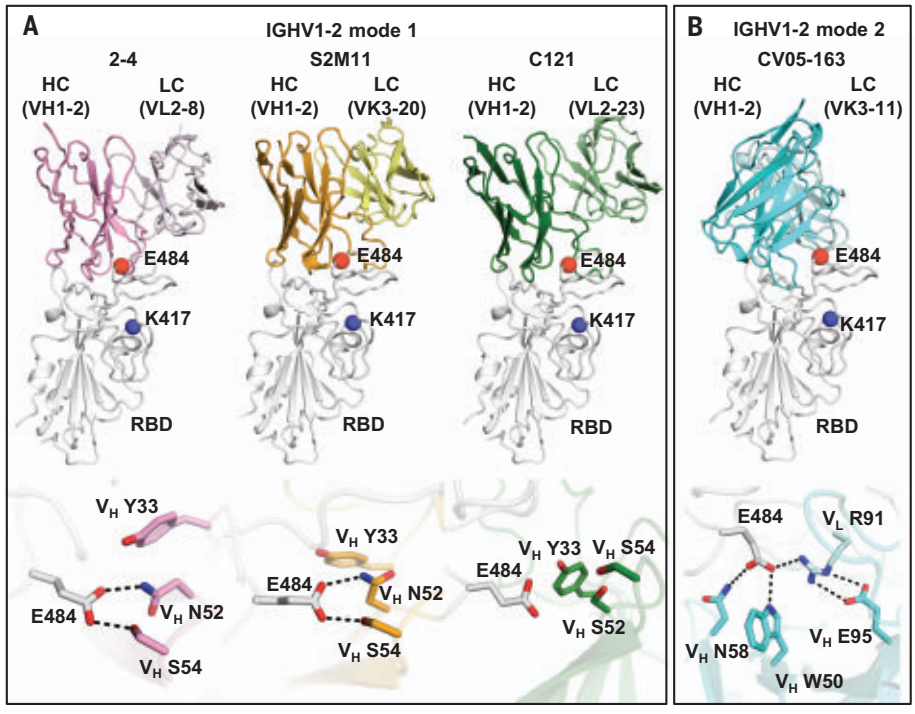



**Fig. 3. Glu$^{484}$ is critical for RBD recognition of IGHV1-2 antibodies.** (**A**) Binding mode 1; (**B**) binding mode 2. Heavy and light chains of antibody 2-4 (PDB 6XEY) (*27*) are shown in pink and light pink, respectively; S2M11 (PDB 7K43) (*30*) in orange and yellow; C121 (PDB 7K8X) (*10*) in dark and light green; and CV05-163 in cyan and light cyan. The RBD is shown in white. Glu$^{484}$ and Lys$^{417}$ are highlighted as red and blue spheres, respectively. Hydrogen bonds are represented by dashed lines; hydrogen bonds are not shown for C121 because of limited resolution (3.9 Å).

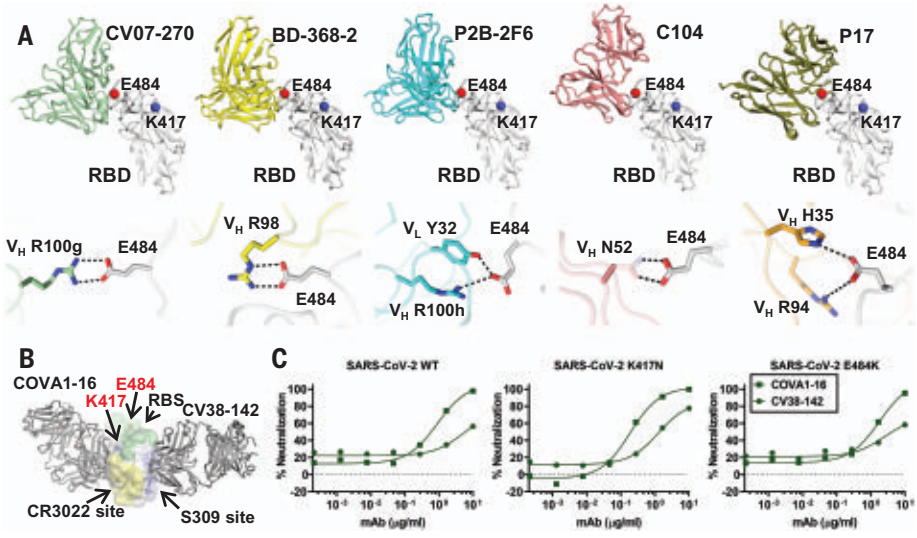



**Fig. 4. Antibodies targeting other major antigenic sites are differentially affected by mutations in recent variants.** (**A**) Interactions between RBS-C antibodies and SARS-CoV-2 RBD. The RBD is shown in white with $Glu^{484}$ and $Lys^{417}$ represented as red and blue spheres, respectively. The various antibodies illustrated are in different colors. Only the variable domains are shown for clarity. Hydrogen bonds and salt bridges to $Glu^{484}$ are represented by dashed lines. Published structures [PDB 6XKP (*17*), 7CHF (*38*), 7BWJ (*18*), 7K8U (*10*), and 7CWN (*78*)] are used to depict structures of SARS-CoV-2 RBD with CV07-270, BD-368-2, P2B-2F6, C104, and P17, respectively. The electron density for the full side chain of $V_H$ $Asn^{52}$ was not well resolved in the 3.8-Å structure of C104 in complex with SARS-CoV-2 S. The full side chain is modeled here and shown as transparent sticks to illustrate a possible interaction with $Glu^{484}$. (**B**) Cross-neutralizing antibodies to the RBD are not affected by $Glu^{484}$ and $Lys^{417}$ mutations. COVA1-16 targets the CR3022 cryptic site (yellow) (*80*) and CV38-142 targets the S309 proteoglycan site (blue) (*81*) to the RBD. Glycans at the $Asn^{343}$ glycosylation site are represented by sticks. The RBS surface is shown in green. $Glu^{484}$ and $Lys^{417}$ are highlighted as red and blue spheres, respectively. (**C**) Neutralization of CV38-142 and COVA1-16 against SARS-CoV-2 wild type, K417N, or E484K pseudoviruses.

the CR3022 site, including COVA1-16 (*80*), S304, S2A4 (*31*), and DH1047 (*41*). Recently, we isolated antibody CV38-142 that targets the S309 site (*81*). Antibodies targeting these two epitopes are often cross-reactive with other sarbecoviruses, as these sites are more evolutionarily conserved than the RBS. To test the effect of the K417N and E484K mutations on nAbs that target the S309 and CR3022 sites, we assessed binding and neutralization by CV38-142 and COVA1-16 to SARS-CoV-2. Both mutations have minimal effect on these antibodies (Fig. 1C and Fig. 4C).

The most potent neutralizing antibodies to SARS-CoV-2 generally tend to target the RBS (table S3), as they directly compete with receptor binding. Such RBS antibodies often interact with $Lys^{417}$, $Glu^{484}$, or $Asn^{501}$ and are therefore sensitive to RBS mutations at these positions in the VOCs. On the other hand, antibodies targeting the CR3022 and S309 sites are often less potent but are less affected by the VOCs, as their epitopes do not contain mutated residues. In fact, recent studies have shown that sera from convalescent or vaccinated individuals can retain neutralization activity, albeit reduced, against the mutated variants (*48*, *50*, *82*), which is possibly due to antibodies targeting other epitopes including the CR3022 and S309 sites. Thus, the CR3022 and S309 sites are promising targets to avoid interference by SARS-CoV-2 mutations observed to date.

As SARS-CoV-2 continues to circulate in humans and increasing numbers of COVID-19 vaccines are administered, herd immunity to SARS-CoV-2 should be approached locally and globally. However, as with other RNA viruses, such as influenza and HIV (*83*), further antigenic drift is anticipated in SARS-CoV-2. Within-host antigenic drift has also been observed in an immunosuppressed COVID-19 patient who had low titers of neutralizing antibodies that allowed the emergence of N501Y and E484K mutations (*84*). Whereas antibody responses elicited by the wild-type lineage that initiated the COVID-19 pandemic have been well characterized in natural infection (*11*, *14*–*18*, *20*, *23*, *26*, *27*, *29*) and vaccination (*50*, *85*–*87*), data on the immune response to VOCs are now only starting to emerge (*88*) and will clarify the similarity and differences in the antibodies elicited. Because SARS-CoV-2 is likely to become endemic (*89*), the findings here and in other recent studies can be used to fast-track the development of more broadly effective vaccines and therapeutics.

## REFERENCES AND NOTES


1. M. Chand *et al.*, *Investigation of Novel SARS-CoV-2 Variant: Variant of Concern 202012/01. Technical Briefing 5*. Public Health England (2020); https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959426/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5.pdf.
2. H. Tegally *et al.*, *Nature* **592**, 438–443 (2021).
3. N. R. Faria *et al.*, "Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings" (2021); https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586.
4. V. Tchesnokova *et al.*, bioRxiv 432189 [preprint]. 11 March 2021.
5. P. D. Yadav *et al.*, *Clin. Infect. Dis.* ciab411 (2021).
6. X. Zhu *et al.*, *PLOS Biol.* **19**, e3001237 (2021).
7. T. N. Starr *et al.*, *Cell* **182**, 1295–1310.e20 (2020).
8. Y. Shu, J. McCauley, *Euro Surveill.* **22**, 30494 (2017).
9. M. Yuan, H. Liu, N. C. Wu, I. A. Wilson, *Biochem. Biophys. Res. Commun.* **538**, 192–203 (2021).
10. C. O. Barnes *et al.*, *Nature* **588**, 682–687 (2020).
11. T. F. Rogers *et al.*, *Science* **369**, 956–963 (2020).
12. W. Dejnirattisai *et al.*, *Cell* **184**, 2183–2200.e22 (2021).
13. M. Yuan *et al.*, *Science* **369**, 1119–1123 (2020).
14. Y. Cao *et al.*, *Cell* **182**, 73–84.e16 (2020).
15. D. F. Robbiani *et al.*, *Nature* **584**, 437–442 (2020).
16. C. O. Barnes *et al.*, *Cell* **182**, 828–842.e16 (2020).
17. J. Kreye *et al.*, *Cell* **183**, 1058–1069.e19 (2020).
18. B. Ju *et al.*, *Nature* **584**, 115–119 (2020).
19. D. Pinto *et al.*, *Nature* **583**, 290–295 (2020).
20. P. J. M. Brouwer *et al.*, *Science* **369**, 643–650 (2020).
21. Y. Wu *et al.*, *Science* **368**, 1274–1278 (2020).
22. X. Chi *et al.*, *Science* **369**, 650–655 (2020).
23. E. Seydoux *et al.*, *Immunity* **53**, 98–105.e5 (2020).
24. R. Shi *et al.*, *Nature* **584**, 120–124 (2020).
25. X. Han *et al.*, *Front. Immunol.* **12**, 653189 (2021).
26. S. J. Zost *et al.*., *Nat. Med.* **26**, 1422–1427 (2020).
27. L. Liu *et al.*, *Nature* **584**, 450–456 (2020).
28. J. Hansen *et al.*, *Science* **369**, 1010–1014 (2020).
29. C. Kreer *et al.*, *Cell* **182**, 843–854.e12 (2020).
30. M. A. Tortorici *et al.*, *Science* **370**, 950–957 (2020).
31. L. Piccoli *et al.*, *Cell* **183**, 1024–1042.e21 (2020).
32. S. A. Clark *et al.*, *Cell* **184**, 2605–2617.e18 (2021).
33. M. Mor *et al.*, *PLOS Pathog.* **17**, e1009165 (2021).
34. R. Babb *et al.* (Regeneron Pharmaceuticals Inc.), U.S. Patent 10787501 (2020); https://uspto.report/patent/grant/10,787,501.
35. M. Yuan *et al.*, *Science* **368**, 630–633 (2020).
36. N. K. Hurlburt *et al.*, *Nat. Commun.* **11**, 5413 (2020).
37. T. Noy-Porat *et al.*, *Nat. Commun.* **11**, 4303 (2020).
38. S. Du *et al.*, *Cell* **183**, 1013–1023.e13 (2020).
39. Y. Zhou *et al.*, *Cell Rep.* **34**, 108699 (2021).
40. A. R. Shiakolas *et al.*, *Cell Rep. Med.* **2**, 100313 (2021).
41. D. Li *et al.*, *Cell* 10.1016/j.cell.2021.06.021 (2021).
42. B. B. Banach *et al.*, bioRxiv 424987 [preprint]. 3 January 2021.
43. G. Bullen *et al.*, *Front. Immunol.* **12**, 678570 (2021).
44. J. Wan *et al.*, *Cell Rep.* **32**, 107918 (2020).
45. Y. Weisblum *et al.*, *eLife* **9**, e61312 (2020).
46. A. J. Greaney *et al.*, *Cell Host Microbe* **29**, 44–57.e9 (2021).
47. E. Andreano *et al.*, bioRxiv 424451 [preprint]. 28 December 2020.
48. C. K. Wibmer *et al.*, *Nat. Med.* **27**, 622–625 (2021).
49. A. J. Greaney *et al.*, *Cell Host Microbe* **29**, 463–476.e6 (2021).
50. Z. Wang *et al.*, *Nature* **592**, 616–622 (2021).
51. L. Stamatatos *et al.*, *Science* eabg9175 (2021).
52. L. Wang *et al.*, *Science* **373**, eabh1766 (2021).
53. X. Xie *et al.*, *Nat. Med.* **27**, 620–621 (2021).
54. D. A. Collier *et al.*, *Nature* **593**, 136–141 (2021).
55. W. F. Garcia-Beltran *et al.*, *Cell* **184**, 2372–2383.e9 (2021).
56. V. Shinde *et al.*, *N. Engl. J. Med.* **384**, 1899–1909 (2021).
57. R. E. Chen *et al.*, *Nat. Med.* **27**, 717–726 (2021).
58. X. Shen *et al.*, *Cell Host Microbe* **29**, 529–539.e3 (2021).
59. S. Cele *et al.*, *Nature* **593**, 142–146 (2021).
60. V. V. Edara *et al.*, *Cell Host Microbe* **29**, 516–521.e3 (2021).
61. J. Hu *et al.*, *Cell Mol. Immunol.* **18**, 1061–1063 (2021).
62. R. Wang *et al.*, *Immunity* **54**, 1611–1621.e5 (2021).
63. W. Dejnirattisai *et al.*, *Cell* **184**, 2939–2954.e9 (2021).
64. M. Hoffmann *et al.*, *Cell* **184**, 2384–2393.e12 (2021).
65. D. Planas *et al.*, *Nat. Med.* **27**, 917–924 (2021).
66. P. Wang *et al.*, *Nature* **593**, 130–135 (2021).
67. P. Wang *et al.*, *Cell Host Microbe* **29**, 747–751.e4 (2021).

68. D. Zhou *et al*., . *Cell* **184**, 2348–2361.e6 (2021).
69. G. Cerutti *et al*., *Structure* **29**, 655–663.e4 (2021).
70. T. N. Starr *et al*., *Nature* 10.1038/s41586-021-03807-6 (2021).
71. Binding and neutralization of IGHV3-66 antibodies were also abolished by K417N, as shown by (*62*), where IGHV3-66 antibody CB6, which binds to RBD in IGHV3-53/3-66 binding mode 1 (fig. S4), was not able to bind or neutralize K417N and B.1.351 or P.1.
72. The epitopes have been assigned on their interaction with a single RBD. Quaternary epitopes are not considered in these assignments.
73. Although the antibodies structurally characterized to date do not represent all of the antibodies in a polyclonal response, they do represent major families of antibodies that have been found in the sera of convalescent SARS-CoV-2 patients.
74. We originally referred these IGHV3-53/3-66 binding modes as "A" and "B" in (*9*) and (*75*). To distinguish from "RBS-A" and "RBS-B", these IGHV3-53/3-66 binding modes are referred to as "IGHV3-53/3-66 binding mode 1" and "2" in this study. Likewise, IGHV1-2 antibodies approach the RBD in two binding modes, which are referred as "IGHV1-2 binding mode 1" and "2". The definitions of the binding modes in these two germline-encoded antibodies differ from "class 1" and "class 2" in (*10*), which were defined as ACE2-blocking antibodies that bind to "up-RBD only" and to "up/down RBD", respectively.
75. N. C. Wu *et al*., *Cell Rep.* **33**, 108274 (2020).
76. S. D. Boyd *et al*., *J. Immunol.* **184**, 6986–6992 (2010).
77. M. Rapp *et al*., *Cell Rep.* **35**, 108950 (2021).
78. H. Yao *et al*., *Cell Res.* **31**, 25–36 (2021).
79. RBS-C antibodies CV07-270, BD-368-2, P2B-2F6, C104, and P17 are encoded by IGHV3-11, IGHV3-23, IGHV4-38-2, IGHV4-34, and IGHV3-30, respectively.
80. H. Liu *et al*., *Immunity* **53**, 1272–1280.e5 (2020).
81. H. Liu *et al*., *Cell Host Microbe* **29**, 806–818.e6 (2021).
82. K. Wu *et al*., bioRxiv 427948 [preprint]. 25 January 2021.
83. G. B. Karlsson Hedestam *et al*., *Nat. Rev. Microbiol.* **6**, 143–155 (2008).
84. B. Choi *et al*., *N. Engl. J. Med.* **383**, 2291–2293 (2020).
85. K. Wu *et al*., *N. Engl. J. Med.* **384**, 1468–1470 (2021).
86. Y. Liu *et al*., *N. Engl. J. Med.* **384**, 1466–1468 (2021).
87. K. R. W. Emary *et al*., *Lancet* **397**, 1351–1362 (2021).
88. T. Moyo-Gwete *et al*., *N. Engl. J. Med.* NEJMc2104192 (2021).
89. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* **368**, 860–868 (2020).
90. J. Lan *et al*., *Nature* **581**, 215–220 (2020).
91. Structures used for the analysis: CC12.1 (PDB: 6XC3), CC12.3 (6XC4), COVA2-04 (7JMO), B38 (7BZ5), CB6 (7C01), CV30 (6XE1), C105 (6XCN), BD-236 (7CHB), BD-604 (7CH4), BD-629 (7CH5), C102 (7K8M), C1A-B3 (7KFW), C1A-C2 (7KFX), C1A-B12 (7KFV), C1A-F10 (7KFY), P4A1 (7CJF), P2C-1F11 (7CDI), LY-CoV481 (7KMI), LY-CoV488 (7KM8), 910-30 (7KS9), 222 (7NX6), S2H14 (7JX3), COVA2-39 (7JMP), C144 (7K90), BD23 (7BYR), 2-4 (6XEY), CV07-250 (6XKQ), REGN10933 (6XDG), C121 (7K8X), C002 (7K8S), P2C-1A3 (7CDJ), S2E12 (7K4N), S2M11 (7K43), S2H13 (7JV2), CT-P59 (7CM4), LY-CoV555 (7L3N), BD-368-2 (7CHH), P2B-2F6 (7BWJ), CV07-270 (6XKP), C104 (7K8U), P17 (7CWO), C110 (7K8V), C119 (7K8W), REGN10987 (6XDG), CR3022 (6W41), COVA1-16 (7JMW), EY6A (6ZDG), S304 (7JW0), S2A4 (7JVA), DH1047 (7LD1), S309 (6WPS), C135 (7K8Z), and CV38-142 (7LM8). The paratope residues of the C104 3.8-Å structure that were truncated due to weak electron density were modeled as full side chains before performing calculations.

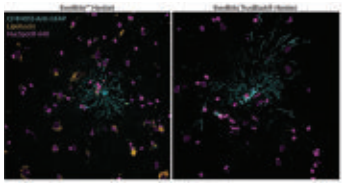

## ACKNOWLEDGMENTS

We thank H. Tien for technical support with the crystallization robot, J. Matteson and Y. Hua for their contributions to mammalian cell culture, W. Yu for insect cell culture, and R. Stanfield for assistance in data collection. **Funding:** Supported by Bill and Melinda Gates Foundation grants OPP1170236 and INV-004923 INV (I.A.W., A.B.W., and D.R.B.); NIH grants R00 AI139445 (N.C.W.), NIH P01 AI110657 (I.A.W., A.B.W., and R.W.S.), R01 AI132317 (D.N. and D.H.), and R01 AI142945 (L.P.); German Research Foundation grants PR 1274/3-1 and PR 1274/5-1, Helmholtz Association grants HIL-A03 and SO-097, and German Federal Ministry of Education and Research Connect-Generate grant 01GM1908D (H.P.); and a Vici fellowship from the Netherlands Organisation for Scientific Research (NWO) (R.W.S.). 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. Extraordinary facility operations were supported in part by the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on the response to COVID-19, with funding provided by the Coronavirus CARES Act. **Author contributions:** M.Y., D.H., C.-C.D.L., N.C.W., and I.A.W. conceived and designed the study; M.Y., C.-C.D.L., N.C.W., and H.L. expressed and purified the proteins for crystallization; S.M.R., H.P., and J.K. provided CV05-163 and other antibody clones and sequences; M.J.v.G., R.W.S., and D.R.B. provided plasmids for some of the antibodies reported in (*11*, *20*), respectively; M.Y. and X.Z. performed the crystallization and x-ray data collection and determined and refined the x-ray structures; D.H., L.P., and D.N. performed the neutralization assays; M.Y. and C.-C.D.L. carried out the binding assays; A.M.J. and A.B.W. provided nsEM data and performed reconstructions; M.Y., C.-C.D.L., N.C.W., and I.A.W. wrote the paper; and all authors reviewed and/or edited the paper. **Competing interests:** Related to this work, the German Center for Neurodegenerative Diseases (DZNE) and Charité–Universitätsmedizin Berlin previously filed a patent application that included anti-SARS-CoV-2 antibody CV05-163 first reported in (*17*). **Data and materials availability:** The x-ray coordinates and structure factors have been deposited to the RCSB Protein Data Bank under accession code 7LOP. The EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD-23466 (one bound), EMD-23467 (two bound), and EMD-23468 (three bound). This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6556/818/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S3
References (*92*–*110*)
MDAR Reproducibility Checklist

17 February 2021; accepted 12 May 2021
Published online 20 May 2021
10.1126/science.abh1139



### Mounting Medium

EverBrite TrueBlack Hardset is the first mounting medium that offers strong antifade protection and quenches lipofuscin autofluorescence commonly found in human and aged animal tissues. The mounting medium cures as it hardens and has a refractive index that is well-matched to coverglass. EverBrite TrueBlack Hardset is available without nuclear stain, with DAPI (4',6-diamidino-2-phenylindole), or with far-red NucSpot nuclear counterstain. This product adds to Biotium's growing line of unique TrueBlack reagents designed for optimal quenching of fluorescence background from lipofuscin and other sources.

**Biotium**
For info: +1-800-304-5357
www.biotium.com

### 3D Liver Spheroid Kit

Leveraging our 3D cell culture technology as well as our expertise in liver biology and absorption, distribution, metabolism, excretion, and toxicity (ADME-Tox) studies. Corning has developed an all-in-one solution for the development of 3D liver models that better recapitulate in vivo behavior. The Corning HepGo assay-ready 3D liver spheroid kit delivers a more physiologically relevant model than traditional 2D models, for reliable, accurate drug safety screening and drug efficacy testing. The kit is a plug-and-play solution that includes 96 preplated primary human hepatocyte (PHH) spheroids in a 96-well spheroid microplate (one uniform spheroid per well) and companion media. It provides easy access to 3D PHH liver spheroids, which may improve accuracy and efficiency in various culture applications, including hepatotoxicity assays. With its assay-ready platform, the HepGo kit allows both imaging-based and biochemical assays to be performed on the same plate—no transfer required. This streamlines the process, reducing time and cost as well as the potential for error, ensuring reliable, accurate, reproducible data.

**Corning Life Sciences**
For info: +1-800-492-1110
www.corning.com

### Cryopreservation Media

STEM-CELLBANKER is a ready-to-use, chemically defined, animal-free freezing medium produced under Good Manufacturing Practice conditions. It is optimized for embryonic stem cell and induced pluripotent stem cell storage, as well as being suitable for cryopreservation of other fragile cell types. Containing only European- or U.S. Pharmacopoeia–grade ingredients, STEM-CELLBANKER is the optimal choice for storage of cells developed for cell therapy applications. It is also available as a dimethyl sulfoxide–free formulation. STEM-CELLBANKER significantly increases cell viability while maintaining cell pluripotency, normal karyotype, and proliferation ability after freeze–thaw. STEM-CELLBANKER is ready to use and requires no special devices, such as a controlled-rate freezer, to achieve consistently high viabilities following resuscitation from cryopreservation, even over extended long-term storage.

**AMS Biotechnology**
For info: +1-617-945-5033
www.amsbio.com/stem-cell-cryopreservation

### gDNA Removal Columns

Porvair Sciences has launched Chromatrap gDNA Removal Columns that quickly remove unwanted genomic DNA (gDNA) from RNA samples. Removal of gDNA from RNA samples is particularly important to scientists undertaking real-time PCR, RNA-seq, and microarray-based applications where high purity and accurate measurement of RNA is vital. Employing a simple load-and-spin method, the new columns efficiently capture gDNA on the membrane, allowing efficient flow-through of DNA-free cell lysates for downstream RNA purification. Unlike traditional gDNA removal methods, Chromatrap gDNA removal columns require no time-consuming enzymatic digestion that can also impact RNA purity. The captured gDNA may also be eluted for downstream applications such as PCR, Southern blotting, cloning, and sequencing.

**Chromatrap**
For info: +44-(0)-1978-661144
www.chromatrap.com/gdna-removal-columns

### Breath Biopsy Panel for Respiratory Diseases

Owlstone Medical has introduced the Respiratory Diseases Research Use Only (RUO) Panel. The breath biomarker panel is intended to support research to distinguish between different types of chronic inflammatory airway diseases including asthma, chronic obstructive pulmonary disease (COPD), and idiopathic pulmonary fibrosis, and to facilitate better therapeutic decision making and monitoring. Owlstone Medical's RUO Panel consists of a set of biologically relevant volatile organic compounds that have been linked to inflammatory respiratory diseases, and can be easily collected on breath to enable direct characterization of disease biology in the lungs by noninvasively sampling the airways. The panel will be used by pharmaceutical and academic/clinical research clients to support therapeutic development through characterizing disease endotypes and monitoring response to therapy, and to better understand the onset, development, and exacerbation of disease.

**Owlstone Medical**
For info: +1-984-244-7811
www.owlstonemedical.com

### Radiation-Resistant Lenses for CCTV Cameras

Resolve Optics reports on growing demand for radiation-resistant fixed-focus and zoom lenses for closed-circuit TV (CCTV) cameras used for visual inspection in nuclear power stations. After the widely reported accidents at Chernobyl and (more recently) the Fukushima Daiichi nuclear power plant, the need for CCTV camera systems that can operate even when subject to high radiation became evident. Drawing upon its expertise in radiation-resistant lenses, Resolve Optics has established itself as a world leader in the supply of key camera lenses used in applications including the inspection of fuel elements, pressure vessels, and nuclear cores, as well as the inspection of loading and offloading of nuclear materials. All optical elements in Resolve Optics radiation-resistant lens designs for CCTV cameras are made using cerium oxide–doped glass or synthetic silica, enabling our fixed-focus and zoom camera lenses to withstand radiation doses up to 1 billion rads and temperatures up to 100°C without significant discoloration or performance degradation. These specialist lenses also provide high image resolution and minimum geometric distortion from 400 nm to 750 nm.

**Resolve Optics**
For info: +44-(0)-1494-777100
www.resolveoptics.com/nuclear

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

Department of Pharmacology | Yale University

## Assistant or Associate Professor of Pharmacology

**The Department of Pharmacology at Yale University School of Medicine invites applications for Assistant or Associate Professor positions in the tenure track.** Applicants must have an M.D. and/or Ph.D. or equivalent degree and have demonstrated excellent qualifications in research and education. Areas of interest include studies of molecular function such as biochemistry, proteomics, chemical biology, structural biology (including electron microscopy), and computational biology in different areas of drug discovery, cell signaling, metabolism, genetic diseases, rare diseases, cancer and epigenetic mechanisms.

Responsibilities include establishing a vigorous and independently funded research program while supervising and mentoring students with diverse backgrounds plus contributing to the graduate and medical school educational missions. We seek individuals with strong records of independent creative accomplishments, who will interact productively with colleagues within the Department and across Yale School of Medicine and the University.

Pharmacology at Yale University has a rich history of pre-eminence in drug discovery, indeed the first cancer chemotherapy clinical trials were conducted here. Today, the research in our Department encompasses a broad range of fields, with researchers aiming to both understand and treat human disease, with particular strengths in drug discovery, integrative cell signaling, structural biology, and neuroscience. For more details, see **https://medicine.yale.edu/pharm/.**

Please submit a letter describing qualifications, along with a CV, a two-page summary of current and proposed research, and three letters of reference to the following Interfolio website: **apply.interfolio.com/90286.**

Informal inquiries may be submitted electronically to **Professor Joseph Schlessinger, Chair of Pharmacology, c/o Nathan Kucera (nathan.kucera@yale.edu).**

Consideration of applications will begin **November 1st, 2021.**

*Yale University is an Equal Opportunity/Affirmative Action Employer. We seek candidates who embrace and reflect diversity in the broadest sense. Yale values diversity among its students, staff, and faculty and strongly welcomes applications from women, persons with disabilities, protected veterans, and underrepresented minorities.*

The EGL Charitable Foundation
invites you to apply to the

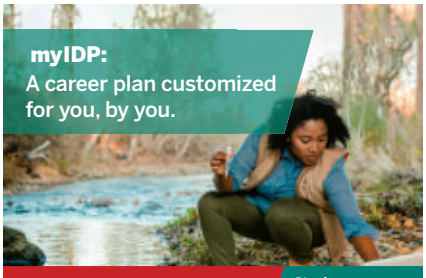
## Gruss Lipper Postdoctoral Fellowship Program

Eligibility:

- Israeli citizenship
- Candidates must have completed PhD and/or MD/PhD degrees in Biomedical Science at an accredited Israeli University/Medical School, or be in their final year of studies
- Candidates must have been awarded a postdoctoral position by a U.S. host research institution

Details available at: www.eglcf.org
Application deadline: October 4, 2021






There's only one Science

**Features in myIDP include:**

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers

In partnership with:

UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND
University of Massachusetts Medical School
FASEB



ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Confused about your next career move?

Download Free Career Advice Booklets!

ScienceCareers.org/booklets

MICHIGAN MEDICINE
UNIVERSITY OF MICHIGAN

**The Geriatrics Center and Biology of Aging Group at the University of Michigan are seeking applications for Assistant or Associate Professor positions.** Successful candidates will join a highly-interactive, well-funded, creative faculty, in an environment that includes the Paul F. Glenn Center for Aging Research, NIA Intervention Testing Program, the UM Cardiovascular Aging Center, Claude Pepper Center, the Ann Arbor VA GRECC, and a long-standing NIA-funded training grant. These programs, together with substantial start-up funding, provide valuable infrastructure and resources for new and established researchers.

Minimum qualifications are a PhD or MD, several years of highly productive postdoctoral research, and a clear interest in problems relevant to aging. Established scientists would need an international reputation as a leader in aging research, a strong track record of extramural funding, and a reputation as an outstanding mentor. Successful candidates will be housed in modern lab space, and will receive a primary appointment in a basic science or research-oriented clinical department as appropriate. **Inquiries can be directed to Rich Miller (millerr@umich.edu).** Applicants should submit a cover letter, a CV, a brief description of research interests, a synopsis of current and previous research support, and contact information for 3 – 5 referees.

*By* **Brian Mustanski**

# Keep quiet about homophobia or open up?

I pride myself on coming from a place of "yes." So it was uncharacteristic that, when my department head asked me to share my experiences of homophobia at a recent virtual diversity town hall for faculty, my first reaction was to decline. He did not know what had happened to me just the week before. I was out for a run when an SUV pulled up next to me. A young man rolled down his window, hung his head out, yelled "faggot"* at me, and laughed as the SUV drove away. I said nothing. I wish I had shouted, "This kind of bullying is the reason 20% of gay teens attempt suicide" in the hope that it might help him understand the implications of his actions. But in that moment, I wasn't Dr. Mustanski, leader of an LGBTQ health research institute. I was just the same Brian who had been called "fag" countless times—and had learned in such situations it was safer to keep quiet.

In my work, it's a different story. I have dedicated my career to research advancing the health and well-being of the LGBTQ community, including documenting the physical and mental health effects of bullying and victimization. I talk passionately about my research to policymakers and journalists, and in public forums. Still, I rarely discuss my own experiences. I've always thought, "Who wants to hear about me when I can share the voices of thousands of research participants?"

Our faculty meetings are typically dedicated to announcing new grants and departmental policies, not telling harrowing personal stories. So, I asked my department head, "Why would I ever share such experiences at a faculty meeting?" He offered that it might make me seem more approachable. The comment both stung and resonated. Like many minority individuals, I put up walls because that feels safer than to risk looking vulnerable, which can be a lure for harassment and abuse. Why would I take off this armor?

**"I put up walls because that feels safer than to risk looking vulnerable."**

But in the end I decided to prepare an essay to read, hoping that opening up and showing my vulnerability could break us out of academia's typical dispassionate and intellectualized discussions about discrimination and diversity. I described how the slur was yelled at me. I told a story of a customs agent at the airport refusing to process me and my then-husband together; instead of demanding the agent accept our civil union or asking to speak to his boss, I remained silent, too afraid of what could happen if I complained and too exhausted from 8 hours on a plane and a lifetime of homophobia to protest. I shared how I was advised to hide being gay during graduate school interviews because being open could risk my chance of admission, being harassed at academic conferences, being told not to mention the words "gay" or "bisexual" in grant applications, and having interviewers on a faculty search committee literally walk out of the room when they heard about the LGBTQ focus of my work. Sharing these stories only took about 10 minutes, but it felt like a grueling hour.

When I finished, supportive and appreciative text messages, emails, and Zoom chats started to pour in, and continued for days. One particularly stuck with me: "Your essay was deeply moving, and I appreciate that you were willing to be vulnerable and give voice to what many can't say."

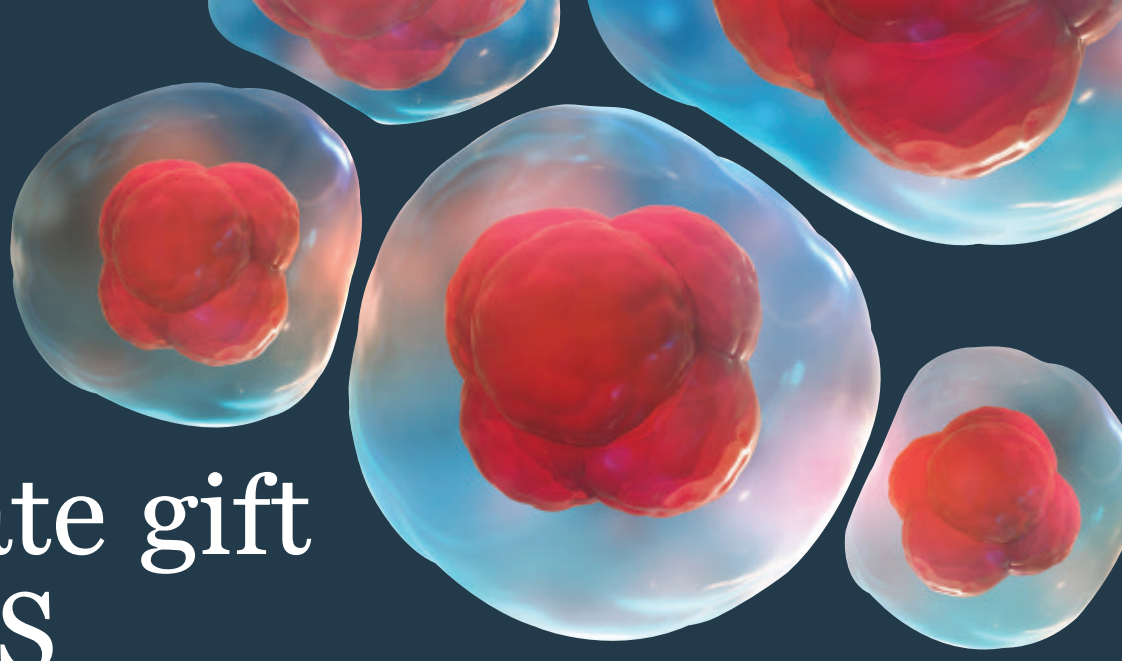
In the months since that meeting, I have had mixed feelings. I believe it is unfair to ask marginalized people to take on the burden of educating others about the -isms that plague academia and beyond. Yet I saw the motivating power of sharing vulnerability with the right audience, and I have noticed other faculty and staff feeling newly empowered to share their personal experiences. Will such inspiration be enough to advance the hard work of making structural changes to increase equity and inclusion? I honestly don't know. But I do know that, although the young man who called me a "faggot" probably won't read this essay, the scientific community is in my—and our—sphere of influence. And by opening up, perhaps I can help spur change. ■

Brian Mustanski is a professor of medical social sciences and director of the Institute for Sexual and Gender Minority Health and Wellbeing at Northwestern University in Chicago. Send your career story to SciCareerEditor@aaas.org.

*Science's policy is to not print slurs in full unless it is crucial to the story.

ILLUSTRATION: ROBERT NEUBECKER

1848 SOCIETY
AAAS



# An estate gift to AAAS

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

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

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

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

**—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member**

---

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

**Legal Name:** American Association for the Advancement of Science
**Federal Tax ID Number:** 53-0196568
**Address:** 1200 New York Avenue, NW, Washington, DC 20005

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

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

cut here

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

1848 SOCIETY
AAAS

**PLEASE CONTACT ME AT:**

Name: _____

Address: _____

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

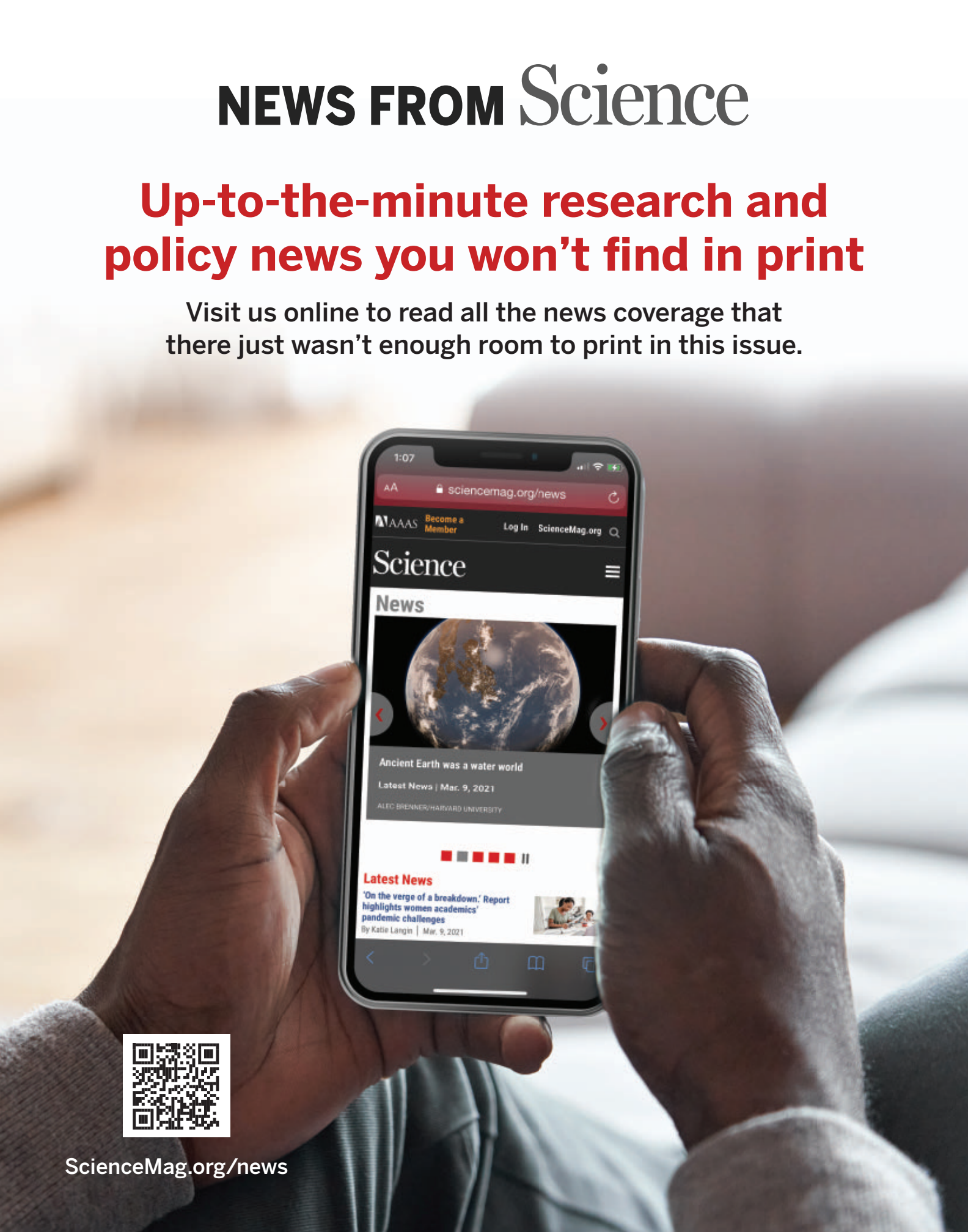
Email: _____ Phone: _____

**RETURN THIS FORM TO:**
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

# 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.




ScienceMag.org/news