Will rising ocean acidity really mess with fish's minds? p. 560

Rethinking hominoid ancestors p. 587

Nocturnal migrant flies high in the daytime p. 646

Science

$15
7 MAY 2021
sciencemag.org

AAAS



MAKING A MALE

Uncommon chromosomal arrangements in creeping voles p. 592



Explore what's possible with innovative research tools

Overcome experimental limitations and gain the freedom to pursue your next discovery with our complete research solution. From leading-edge cell analyzers, sorters, multiomics instrumentation and informatics to advanced reagents, we're committed to providing the critical tools you need to propel your research forward.

So, go beyond your research limitations and explore with confidence.
Discover the difference.

bdbiosciences.com/explore

BD

For Research Use Only. Not for use in diagnostic or therapeutic procedures.
BD and the BD Logo are trademarks of Becton, Dickinson and Company.
© 2021 BD. All rights reserved. BD-26642 (v1.0) 0221

GRC Connects

A Cutting-Edge Virtual Experience

Gordon Research Conferences

Our virtual initiative, *GRC Connects*, provides scientific communities with short and informative opportunities to connect and collaborate with their GRC colleagues until we can resume in-person conferences, furthering GRC's commitment to find new and creative ways to advance the frontiers of science.

Event formats include: panel discussions, networking and mentoring sessions, GRC Power Hours™ and scientific talks!

These free events are open to all members of the scientific community and scientists of all career stages are encouraged to attend.

TESTIMONIALS FROM ATTENDEES:

"I love that these events are FREE! These presentations excited me, and I can't wait to return to an in-person GRC."

"This event was excellent. It was my first, but will not be my last!"

"Very well run and productive! No technical glitches like many other virtual experiences."

Visit **www.grc.org** today to register for one of our upcoming *Connects* events and apply for one of the 17 GRCs and GRSs scheduled to take place, in person, in October and November in New England and California. We look forward to connecting with and seeing you soon!



UPCOMING EVENT TOPICS INCLUDE:

Multi-Drug Efflux Systems

Nanoscale Science and Engineering for Agriculture and Food Sciences

Mechanisms of Epilepsy and Neuronal Synchronization

Preclinical Form and Formulation for Drug Discovery

Natural Products and Bioactive Compounds

Environmental Nanotechnology

Plant Metabolic Engineering





The heart of the matter.

The NEBNext® Ultra™ II workflow lies at the heart of NEB's portfolio for next gen sequencing library preparation. With specially formulated master mixes and simplified workflows, high quality libraries can be generated with low inputs and reduced hands-on time.

As sequencing technologies improve and applications expand, the need for compatibility with ever-decreasing input amounts and sub-optimal sample quality grows. Scientists must balance reliability and performance with faster turnaround, higher throughput and automation compatibility.

NEBNext Ultra II modules and kits for Illumina® are the perfect combination of reagents, optimized formulations and simplified workflows, enabling you to create DNA or RNA libraries of highest quality and yield, even when starting from extremely low input amounts.

The Ultra II workflow is central to many of our NEBNext products, including:

- Ultra II DNA & FS DNA Library Prep
- Enzymatic Methyl-seq
- Ultra II RNA & Directional RNA Library Prep
- Single Cell/Low Input RNA Library Prep
- Module products for each step in the workflow

The Ultra II workflow is available in convenient kit formats or as separate modules – it is easily scalable and automated on a range of liquid handling instruments.

End Repair/ dA-Tailing	Adaptor Ligation	Clean Up/ Size Selection	PCR Enrichment	Clean Up
Module NEB #E7546	Module NEB #E7595	Component of NEB #E7103	Module NEB #M0544	Component of NEB #E7103
NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB #E7645)				
NEBNext Ultra II Library Prep with Sample Purification Beads (NEB #E7103)				

The NEBNext Ultra II workflow has been cited in thousands of publications, as well as a growing number of preprints and protocols related to COVID-19. Citation information and extensive performance data for each product is available on neb.com.

To learn more about why NEBNext is the choice for you, visit **NEBNext.com**.

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. Illumina® is a registered trademark of Illumina, Inc.
© Copyright 2020, New England Biolabs, Inc.; all rights reserved.

NEW ENGLAND BioLabs® Inc.

be INSPIRED
drive DISCOVERY
stay GENUINE

CONTENTS

7 MAY 2021 • VOLUME 372 • ISSUE 6542



NEWS

IN BRIEF

548 News at a glance

IN DEPTH

552 **Is India's coronavirus death 'paradox' vanishing?**
Limited evidence suggests in 2020 the country had relatively low COVID-19 mortality
By J. Cohen

554 **Brazil and Russia face off over vaccine contamination charge**
Does Sputnik V contain replication-competent viruses?
By S. Moutinho and M. Wadman

554 **NASA set to announce Earth system observatory**
Climate-monitoring satellites mark resurgence for agency's earth science division *By P. Voosen*

556 **Fatal attraction to light at night pummels insects**
Fearing artificial lights add to an "insect apocalypse," researchers seek solutions
By E. Pennisi

557 **'Campfires' may drive heating of solar atmosphere**
Observations suggest small flares are corona's mystery heat source *By D. Clery*

558 **USDA now only partially inspects some animal labs**
Internal documents reveal agency moved to "focused" inspections to save work
By D. Grimm

FEATURES

560 **Sea of doubts**
Dozens of papers linking high carbon dioxide to unsettling changes in fish behavior fall under suspicion
By M. Enserink

INSIGHTS

POLICY FORUM

566 **Mapping out a future for ungulate migrations**
Limited mapping of migrations hampers conservation
By M. J. Kauffman et al.



PERSPECTIVES

570 **Macroscale entanglement and measurement**
The entangled motion of macroscopic vibrating membranes can be measured precisely *By H.-K. Lau and A. A. Clerk*
REPORTS pp. 622 & 625

571 **Rapid antigen testing in COVID-19 responses**
SARS-CoV-2 transmission was reduced with measures centered on rapid antigen testing
By M. García-Fiñana and I. E. Buchan
REPORT p. 635

573 **Reversible fusion-fission fibers**
Reversible assembly of graphene oxide fibers creates a pathway for practical applications
By R. Cruz-Silva and A. L. Elías
REPORT p. 614

574 **Illuminating the first bacteria**
A new analysis aims to uncover the root of the bacterial tree of life
By Laura A. Katz
RESEARCH ARTICLE p. 588

575 **Making sense of dinosaurs and birds**
Advances in imaging and statistics illuminate dinosaur sensory biology and behavior *By L. M. Witmer*
RESEARCH ARTICLE p. 601; REPORT p. 610

577 **Carbohydrates, insulin, and obesity**
Insulin plays a role in body fat regulation independent of dietary carbohydrates
By J. R. Speakman and K. D. Hall

PHOTOS: (TOP TO BOTTOM) FREDRIK JUTFELT; LAWRENCE WITMER/OHIO UNIVERSITY

RECENT PHD IN GENOMICS, PROTEOMICS AND SYSTEMS BIOLOGY APPROACHES

Gain international recognition, win 30,000 USD, and have your thesis essay published in Science.

The Science & SciLifeLab Prize for Young Scientists is a global prize aimed at rewarding scientists at an early stage of their careers, awarded in four categories.

Research in this category focuses on genomics, proteomics, integrative omics and systems biology approaches, including computational, to facilitate comprehensive understanding of living cells, organisms and species,

Entrants for the 2021 prize must have received their PhD after January 1, 2019.

APPLY FOR THE *SCIENCE* & SCILIFELAB PRIZE FOR YOUNG SCIENTISTS

APPLY BEFORE JULY 15, 2021

SCIENCEPRIZE.SCILIFELAB.SE

Science & SciLifeLab PRIZE For Young Scientists

Science AAAS

SciLifeLab

Knut och Alice Wallenbergs Stiftelse

BOOKS *ET AL.*

579 More than the message
A new guide offers advice for navigating barriers to successful science communication *By J. Wai*

580 The people at the dawn of civilization
Sumerian language and culture take center stage in a new anthropological analysis *By A. Robinson*

LETTERS

581 Biosecurity for humanitarian aid
By M. P. van den Burg et al.

582 Inclusion through part-time science
By E. Theusch

582 Integrate U.S. science and diplomacy
By A. Banerjee et al.

RESEARCH

IN BRIEF

584 From *Science* and other journals

REVIEW

587 Primate evolution
Fossil apes and human evolution
S. Almécija et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABB4363
VIDEO

RESEARCH ARTICLES

588 Bacterial phylogeny
A rooted phylogeny resolves early bacterial evolution *G. A. Coleman* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABE0511
PERSPECTIVE p. 574

589 Genetics
Environmental robustness of the global yeast genetic interaction network *M. Costanzo* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF8424

590 Paleogenomics
Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments *B. Vernot* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF1667



591 Neurodevelopment
The coding and long noncoding single-cell atlas of the developing human fetal striatum
V. D. Bocchi et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF5759

592 Sex determination
Sex chromosome transformation and the origin of a male-specific X chromosome in the creeping vole *M. B. Couger* et al.

601 Paleontology
The early origin of a birdlike inner ear and the evolution of dinosaurian movement and vocalization *M. Hanson* et al.
PERSPECTIVE p. 575; REPORT p. 610

REPORTS

610 Paleontology
Evolution of vision and hearing modalities in theropod dinosaurs *J. N. Choiniere* et al.
PERSPECTIVE p. 575; RESEARCH ARTICLE p. 601

614 Materials science
Reversible fusion and fission of graphene oxide–based fibers *D. Chang* et al.
PERSPECTIVE p. 573

618 Solar cells
Interfacial toughening with self-assembled monolayers enhances perovskite solar cell reliability *Z. Dai* et al.

Quantum systems
622 Direct observation of deterministic macroscopic entanglement *S. Kotler* et al.
PODCAST

625 Quantum mechanics–free subsystem with mechanical oscillators
L. Mercier de Lépinay et al.
PERSPECTIVE p. 570

630 Ferroelectrics
Reversible oxygen migration and phase transitions in hafnia-based ferroelectric devices *P. Nukala* et al.

635 Coronavirus
The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia
M. Pavelka et al.
PERSPECTIVE p. 571

642 Coronavirus
X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease
S. Günther et al.

646 Migration
Extreme altitudes during diurnal flights in a nocturnal songbird migrant
S. Sjöberg et al.

DEPARTMENTS

545 Editorial
Opening the path to biotech
By Sangeeta Bhatia, Nancy Hopkins, Susan Hockfield

547 Editorial
The frontier is not endless for all
By H. Holden Thorp

654 Working Life
From professor to patient X
By Anne Crecelius

ON THE COVER

A young creeping vole (*Microtus oregoni*) moves cautiously through the forest in Olympic National Park, United States. In this species, the standard mammalian sex determination system has been rearranged, such that there is no standard Y chromosome, yet Y chromosome genes are present in both sexes. The creeping vole genome reveals rare divergence from the relatively inflexible mammalian sex determination system. See page 592.
Photo: David Moskowitz



Science Staff 546
New Products 649
Science Careers 650

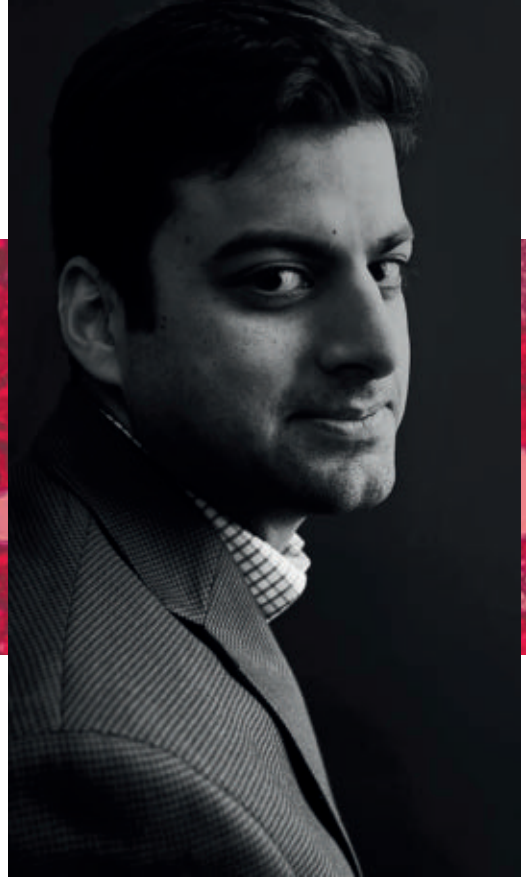
IMAGE: COURTESY OF THE PENN MUSEUM/RECONSTRUCTION OF HEAD AND HEADDRESS OF LADY PU-ABI FROM UR BY KATHARINE WOOLLEY, CA. 1930

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.

HUMAN VACCINES PROJECT
DECODING THE IMMUNE SYSTEM

MICHELSON
MEDICAL RESEARCH
FOUNDATION

MICHELSON PRIZES:
NEXT GENERATION GRANTS



Dr. Ansuman Satpathy,
Michelson Prize Winner 2018
Stanford University

The Michelson Prizes: Next Generation Grants support young investigators applying disruptive research concepts to significantly advance the development of vaccines and immunotherapies for major global diseases.

The 2021 Michelson Prizes will be awarded for research proposals in two focus areas:

- Human Immunology and Vaccine Research
- Climate Change and Human Immunology

Researchers from a wide array of disciplines are encouraged to apply.

Are you 35 or younger?
Apply now for the 2021 Michelson Prizes!

www.michelsonprizes.smapply.org

"The Michelson Prize was instrumental in my early career. It allowed me to pursue high-risk, high-impact ideas."

APPLY TODAY

APPLICATIONS DUE:
JUNE 18, 2021

GRANT AWARD:

$150,000

EDITORIAL

Opening the path to biotech

In 1999, the Massachusetts Institute of Technology (MIT) released a study that documented how women faculty in its School of Science were afforded fewer resources and opportunities than men—a discrepancy it attributed to unconscious biases that had marginalized women faculty "even in the light of obvious good will." The report inspired policy changes at universities across the country that have made faculty resources more equitable. But a study released last month by MIT members (including the authors of this editorial) of the Boston Biotech Working Group (BBWG) now documents a similar problem at the interface of academia and industry: Fewer women than men faculty at MIT move their research discoveries into companies, and fewer serve as scientific advisers or on boards of directors. This disparity holds back women faculty and denies the full promise of innovation to the universities they work for, the biotech industry, and society at large.

In 2019, the BBWG brought together a diverse group of biotech and biopharma participants to investigate the situation. The results are sobering, if not surprising. In contrast to their male peers, women in the biological sciences at MIT rarely start biotech companies or sit on their boards. This is true despite an increase in women faculty with qualifications equal to those of their male colleagues. A 2018 study from Stanford University, led by Ann Arvin, described a similar finding at that institution. The BBWG study focused on only 7 of the 14 departments in MIT's Schools of Science and Engineering, but even so it arrived at a remarkable conclusion: Had the women in those departments participated in founding companies at the same rate as the men, they could have started roughly 40 additional companies helping to predict, prevent, and treat disease.

> "...women... could have started roughly 40 additional companies..."

Just as sobering are the stories of women faculty being passed over as cofounders and scientific advisers, and left out of networking events, leaving them on the sidelines of an emerging ecosystem that is critical both to advancing their discoveries and to propelling the careers of their trainees. Some describe feeling "invisible" in spite of their expertise, and others have been advised to take along male students or post-docs when pitching for capital to be taken more seriously.

What can be done? A first step is to acknowledge the importance of network effects. Many social and professional groups remain inaccessible to women. Such networks mediate how money is raised, how teams are assembled, and how people learn about the commercialization process. Women in the biological sciences with ideas ripe for the market clearly don't enjoy equal access to these networks—in 2019 only 2.7% of US venture capital dollars, for example, supported women-founded companies. As concerning, the lack of network access is even greater for women (and men) of color.

To address these imbalances, the BBWG launched several programs that could be easily replicated at other universities. Its Future Founders Initiative convenes boot camps where aspiring women academic entrepreneurs can get candid advice from experienced founders. Its "VC Pledge" invites venture capital groups to commit to achieving, within 2 years, a goal of at least 25% women on the boards of companies that they influence. Its Accelerator Fellowships offer tenured women faculty the chance to spend time at venture capital firms, watching ideas get vetted and deals get done. The BBWG Data Group has created a framework to quantify faculty participation in the entrepreneurial ecosystem and potentially reveal points of intervention. For example, MIT and Stanford data show that women and men faculty in some engineering departments are starting companies at similar rates, suggesting that adjusting departmental "microclimates" may be helpful.

The MIT study complements other efforts in academia and the broader ecosystem. Washington University's Equalize is a national pitch competition that pairs academic women inventors with experienced mentors to turn ideas into start-ups. The Mass General Brigham Innovation office sponsors board director candidate workshops and an Innovation Academy for prospective founders. To increase access to capital, angel groups and accelerator funds, such as MassNextGen and AIM-HI, and networks of women venture capital investors seek to increase women's participation. And Equality Can't Wait, founded by philanthropists Melinda Gates and MacKenzie Scott, aims to expand women's influence.

These programs could address another challenge: underrepresentation of minority faculty in tech transfer. The discoveries women and minority researchers are making today have great potential as a force for good in the world—but reaching that potential is only possible if paths to real-world applications are open to everybody.

–Sangeeta Bhatia, Nancy Hopkins, Susan Hockfield

Sangeeta Bhatia is the Wilson Professor of Engineering and director of the Marble Center for Cancer Nanomedicine at MIT, Cambridge, MA, USA. sbhatia@mit.edu

Nancy Hopkins is Amgen Inc. Professor of Biology emerita at MIT, Cambridge, MA, USA. nhopkins@mit.edu

Susan Hockfield is President Emerita, professor of Neuroscience, and a member of the Koch Institute for Integrative Cancer research at MIT, Cambridge, MA, USA. hockfield@mit.edu

10.1126/science.abj2642

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

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Kelly Servick, Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERN** Sofia Moutinho **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** Donisha Adams, Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **MULTIMEDIA MANAGING PRODUCER** Joel Goldberg **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **WEB DESIGNER** Jennie Pajerowski **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **VIDEO PRODUCER** Meagan Cantwell **SENIOR PODCAST PRODUCER** Sarah Crespi

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman, Meron Kebede **FINANCIAL ANALYST** Alexander Lee **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SENIOR SALES COORDINATOR** Shirley Young **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Jacob Hedrick, 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** Tori Velasquez, Mike Romano, Ashley Hylton **DIGITAL MARKETING SPECIALIST** Asleigh Rojanavongse **SENIOR DESIGNER** Kim Huynh

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

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

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

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

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

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

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

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

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

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

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

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

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

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

AAAS BOARD OF DIRECTORS
CHAIR Claire M. Fraser
PRESIDENT Susan G. Amara
PRESIDENT-ELECT Gilda A. Barabino
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER Sudip Parikh
BOARD Cynthia M. Beall, Rosina M. Bierbaum, Ann Bostrom, Janine Austin Clayton, Laura H. Greene, Kaye Husbands Fealing, Maria M. Klawe, Robert B. Millard, William D. Provine

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

Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
Sebastian Amigorena, *Institut 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*
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.*
Bradley Bernstein, *Mass. General Hospital*
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*
Nick Chater, *U. of Warwick*
Ling-Ling Chen, *SIBCB, CAS*
M. Keith Chen, *UCLA*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
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*
Jennifer Elisseeff, *Johns Hopkins U.*
Andrea Encalada, *U. San Francisco de Quito*
Nader Engheta, *U. of Penn.*
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*
Elizabeth Grove, *U. of Chicago*
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*
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*
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)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci., Res. & 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*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Institut 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 Centre*
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.*
Kimani Toussaint, *Brown U.*
Wim van der Putten, *Netherlands Inst. of Ecology*
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)**
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*

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.

EDITORIAL

The frontier is not endless for all



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

Recent weeks have seen numerous calls for more investment in research and development (R&D) in the United States. This is understandable with a new administration that is friendlier to science and with The Endless Frontier Act—a measure that could double the budget of the National Science Foundation in 5 years—under consideration in Congress. Proponents of the bill are heralding its potential to enhance America's competitiveness: A large part of the new money would go for "use-inspired" basic research aimed at economic growth. Although the new money for science would be long overdue, and there are provisions in the bill to try to extend its geographical benefit, care must be taken to ensure that funds are distributed more equitably than in the past. If science in the United States is truly to be an endless frontier, the benefits must extend equitably to all.

No one has had a better front row seat for efforts in the United States to boost competitiveness than Deborah Wince-Smith, the president and chief executive officer of the Council on Competitiveness, an eclectic group of leaders from business, academia, organized labor, and national laboratories. The council is one of a very small number of bodies that bring together a wide political spectrum on common interests. It has succeeded for 35 years because of the passion and deep knowledge that Wince-Smith brings to her leadership. I asked her how things had changed in the past 35 years. "When the council was formed, people thought what corporate America does is good for America," she said, "But we've learned during this period with the demise of many of our manufacturing centers, that individual Americans in regions throughout our country have not been part of the prosperity game that's at the heart of who we are as a nation."

We know that simply getting more research funding to these regions is not the whole answer. Of the top 30 universities in R&D spending, 11 are responsible for $11.3 billion in funding in the thriving coastal cities of New York, Los Angeles, San Francisco, Boston, Seattle, and San Diego. But the other 19 members of this group, which account for $20 billion in spending, include cities and states in regions that have not benefited from the economic resurgence of the coastal hubs. There are plenty of great and well-funded scientists in these areas, but nowhere near an equal share of fast-growing and innovative companies.

In addition to poor geographic distribution, the number of women and people of color who are participating in the American innovation economy is dismal. In another editorial in this issue, Sangeeta Bhatia, Nancy Hopkins, and Susan Hockfield note that only 2.7% of venture capital goes to women-founded companies—and the statistics for people of color are even worse.

Over the years, American science policy has been shaped by two canonical reports: *Science: The Endless Frontier* (1945) and *Rising Above the Gathering Storm* (2005). Both were patriotic jeremiads proclaiming that American leadership in science and technology would lead to American strength in economic and foreign policy. Neither dealt explicitly with systemic sexism and racism in science or the poor geographic distribution of opportunity. *Rising Above the Gathering Storm* addressed failures in science education. But it did not address how the prevention of women and people of color from earning science degrees and advancing in their fields affected American competitiveness.

> "'We can't say we are a great nation when... so many... are not participating...'"

And that is what sets apart a recent report by the Council on Competitiveness, *Competing in the Next Economy*, which explicitly calls for widening the innovation economy in the United States to include more people and places. "We need to think of this in a new way and not be crippled by the old language of industrial policy that is really hurting us," Wince-Smith said. "We can't say we are a great nation when we have so many parts of our country and so many citizens that are not participating in the opportunities."

In other words, although China and Europe are formidable scientific competitors of the United States, achieving true competitiveness as a nation starts at home. As the United States plans for another welcome surge of research funding, it must work ever harder to expand the reach of this investment.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

10.1126/science.abj2583

3.1 MILLION

Hectares of land in Indonesia burned in 2019, an area the size of Maryland and twice the Indonesian government's estimate. In 2020, Indonesia deported the lead author of the new estimate after disagreeing with his preliminary one (accepted for review by *Earth System Science Data*).

IN BRIEF

Edited by **Jeffrey Brainard**

CLINICAL TRIALS

In a first, FDA enforces trial data posting

In an unprecedented enforcement action, the U.S. Food and Drug Administration (FDA) last week informed a clinical trial sponsor that it had violated the law that requires results from human studies of medical treatments and devices be posted to the federal repository ClinicalTrials.gov. The action was leveled against the drug company Acceleron Pharma for failing to provide data from a completed trial of its experimental drug to treat kidney cancer. Acceleron ceased the drug's development in 2017, following the trial's disappointing results, according to a company spokesperson. FDA gave the company 30 more days to supply the results or face fines; the spokesperson says it will comply. Congress created the reporting law in 2007 after pharma companies suppressed data revealing lucrative drugs to be unsafe or ineffective. In 2017, FDA and the National Institutes of Health, which oversees the law for the researchers it funds, clarified its rules. But thousands of trial sponsors have continued to ignore the law, investigations in *Science* and elsewhere have found.

Senate grills adviser nominee

LEADERSHIP | Supporters of Eric Lander, President Joe Biden's choice to lead the White House's science office, are hoping his apology last week for "understating" the contributions of two female scientists to the discovery of the CRISPR gene-editing technology will clear the way for his Senate confirmation. The molecular biologist mounted an aggressive defense of his character and suitability for the job during a 2-hour hearing that included pointed questions from both Democrats and Republicans about his treatment of women. "It is clear that you wanted to make sure your lab got the patents and the credit for the CRISPR technology, and in the process you may have marginalized the contributions of" Jennifer Doudna and Emmanuelle Charpentier, said Senator Cynthia Lummis (R–WY). "I felt terrible," Lander replied. "I made a mistake." The lawmakers also scrutinized his meetings with sex offender and science philanthropist Jeffrey Epstein when Lander was director of the Broad Institute. Lander said the two met briefly twice in 2012 but that the institute neither requested nor received funds from Epstein.

U.S. curbs some greenhouse gases

CLIMATE POLICY | The U.S. Congress and the Biden administration are moving to clamp down on releases of two major planet-warming gases. On 29 April, the Senate voted 52 to 42 to cancel a decision by former President Donald Trump that had overturned the first ever federal limits, imposed in 2016, on methane leaking from oil and gas wells. The House of Representatives is expected to join the Senate in approving the cancellation. Then on 3 May, the Environmental Protection Agency (EPA) announced plans to start deep cuts in emissions of hydrofluorocarbons (HFCs), chemicals used in refrigeration and air conditioning. The new rules, which set HFC limits for 2022 and 2023, are part of EPA's effort to implement a law passed last year by Congress that calls for slashing HFC emissions by 85% over 15 years.

Japan to fight theft of ideas

INTELLECTUAL PROPERTY | To reduce risks of intellectual property (IP) theft, a Japanese governmental advisory committee on 27 April approved a draft policy that will require university and national laboratory scientists to disclose all international sources of outside income and to confirm they are not cooperating with organizations identified as security risks by the Japanese government. Those failing to comply with the rules, to be finalized this year, could be ineligible for government grants for up to 5 years. The new initiatives, which echo measures in the United States, were pushed by politicians and officials concerned that Japan has been slow to bolster protection of its advanced technology, even though there have been no high-profile allegations of theft. The same concerns have led to plans for closer scrutiny of international student visa applications. The policy does not mention particular countries, but Japanese officials have been wary of possible IP theft by China, which sends more than 120,000 students to Japan annually.

Mars helicopter scouts for rover

PLANETARY SCIENCE | After four modest flights, NASA last week extended the mission of its small helicopter on Mars, Ingenuity (below). Its next, monthlong assignment: to scout sites for the Perseverance rover to sample.



PHOTO: NASA/JPL-CALTECH/ASU



Snorkelers swim near mangroves in Mexico's Casa Cenote, one of the sinkholes researchers studied.

PHOTO: ALEX MUSTARD/NPL/MINDEN PICTURES

ENVIRONMENT

Yucatán Peninsula sinkholes are vaults of stored carbon

Fed by mangrove roots, the turquoise sinkholes on the Yucatán Peninsula known as cenotes have accumulated some of the highest levels of soil organic carbon found on Earth, a study reports. Scientists measured samples of peat collected by divers under mangroves at the edges of three cenotes near Tulum, Mexico, the team reports on 5 May in *Biology Letters*. Mangroves were already known to be among the most productive ecosystems; using radiocarbon aging, the researchers estimate that these cenotes have been storing their carbon for more than 3000 years. If preserved, the Yucatán Peninsula's more than 2000 sinkholes could provide an economic boon to their owners, most of whom are Mayan communities, who could sell the carbon-storage capabilities as offsets for carbon losses from land development elsewhere, the team writes. But the survival of the mangroves and their cenotes is jeopardized by rising sea level, poorly planned tourism, groundwater extraction, and water pollution.

After instrument checks, the rover will explore rocks over a 2-kilometer stretch that could include volcanic deposits and mudstone remnants of an ancient lakebed. Ingenuity's new flights are meant to provide warnings of steep terrain inaccessible to Perseverance, and to show how rotorcraft could scout for future missions. But because of the time Ingenuity needs to recharge its battery, Perseverance eventually will move faster than the helicopter. Once the helicopter is left in the dust, its communications link on the rover will be severed, ending its lookout jaunts.

Congo's Ebola outbreak ends

PUBLIC HEALTH | The latest outbreak of Ebola in the Democratic Republic of the Congo (DRC) is officially over. The first cases were identified in North Kivu province in early February, but prompt control measures, including more than 1000 vaccinations, kept the spread to just 12 people, with six fatalities. On 3 May, the World Health Organization declared the outbreak over, after the official 42-day waiting period had elapsed since the last diagnosed person tested negative for the virus. The DRC has contained 12 outbreaks of the hemorrhagic disease since 1976, nine of them since 2007. North Kivu was the center of a much larger outbreak that killed more than 2200 people between 2018 and 2020. Another outbreak that also started in February, in Guinea, has so far caused at least 18 recorded cases and nine deaths. More than 5000 people in the West African country have been vaccinated as part of control efforts.

Ant name honors gender diversity

TAXONOMY | Scientists have chosen a Latin name for a newly discovered ant species that is one of the first meant to honor gender diversity and transgender people. The miniature trap jaw ant was dubbed *Strumigenys ayersthey*, honoring artist Jeremy Ayers. Typically, animal species named after people have Latin spellings ending in "-ae" when the person identifies as a woman and "-i" when they identify as a man. In this case, however, the researchers chose "they," a pronoun used by some nonbinary people who identify as neither male nor female. The ant was discovered in 2018 in a forest preserve in Ecuador, the authors report this week in *ZooKeys*.

Bird counts resume, cautiously

ORNITHOLOGY | More than 2000 volunteer citizen-scientists are again identifying and counting North American birds after a long-running annual survey was suspended for 2020 because of the COVID-19 pandemic. Results from the North American Breeding Bird Survey help land and resource managers understand the status and trends of more than 500 bird species. Data collected in past years have informed understanding of the effects of pollution



HUMAN ORIGINS

Africa's oldest known grave held a boy

Some 78,000 years ago, a community in what today is East Africa laid to rest a small boy, about 3 years old. They dug a grave, curled his body, and may have rested his head on a pillow. In 2013, a team that included a Kenyan archaeologist and skilled local diggers unearthed the remains of the boy (illustration, above)—whom they dubbed Mtoto, meaning "boy" in Swahili—in a cave near Kenya's coast. Complex behaviors, indicated by jewelry and the use of ochre pigment, appeared in Africa, the birthplace of our species, as long as 150,000 years ago, but Mtoto represents the oldest evidence on the continent of people buying their dead, the researchers announce on 6 May in *Nature*. The body's careful preparation adds to evidence that ancient people may have conducted special funerary rites for children, who are overrepresented in the ancient archaeological record.

NEWSMAKERS

Biden science team fills out

ENERGY President Joe Biden last week said he will nominate University of Oregon chemist Geraldine Richmond to be undersecretary for science at the Department of Energy (DOE). The position's duties include coordinating R&D programs across the agency's 17 national laboratories. An expert on the molecular characteristics of water surfaces, Richmond served on the National Science Board and chaired DOE's Basic Energy Sciences Advisory Committee. In 2016, she received the National Medal of Science. She is a past president of AAAS (which publishes *Science*) and co-founded the Committee on the Advancement of Women Chemists, which works to encourage women scientists.

NASA After breezing through his U.S. Senate confirmation last week, Bill Nelson, a former astronaut and U.S. senator from Florida, was sworn in as NASA administrator on 3 May. Nelson will face immediate challenges, including overseeing a $2.9 billion contract NASA awarded last month to SpaceX to develop its Starship spacecraft for use as the agency's human lunar lander. The two other bidders for the contract, Blue Origin and Dynetics, have protested the decision on technical grounds. NASA has suspended the contract pending review by Congress's Government Accountability Office.

and land development on North America's bird population, which has declined by nearly 3 billion or 29% since 1970, a 2019 paper in *Science* reported. This year's survey, coordinated jointly by agencies in Canada, Mexico, and the United States, began in April and ends in July. The volunteers working outdoors have little social interaction, but organizers have encouraged them to take standard precautions against the COVID-19 virus.

U.S. moves to build bomb triggers

NUCLEAR WEAPONS | The U.S. Department of Energy (DOE) last week took a step toward upgrading its nuclear arsenal by approving a conceptual design for manufacturing new plutonium triggers to detonate thermonuclear bombs. The "pits" would replace aging ones made decades ago. The Los Alamos National Laboratory will produce up to 30 pits annually starting by 2026, under the plan from DOE's National Nuclear Security Administration (NNSA). Critics have questioned whether the work is needed, citing estimates that existing pits might function as designed for decades more. The work also faces opposition from environmental groups. Operational and safety problems have delayed Los Alamos from executing long-standing plans to produce new pits, and Congress's Government Accountability Office last year questioned whether the project can meet its timetable and contain costs. DOE's Savannah River Site in South Carolina may build additional pits, but NNSA has yet to release a conceptual design for producing those.

IBM unveils denser chip

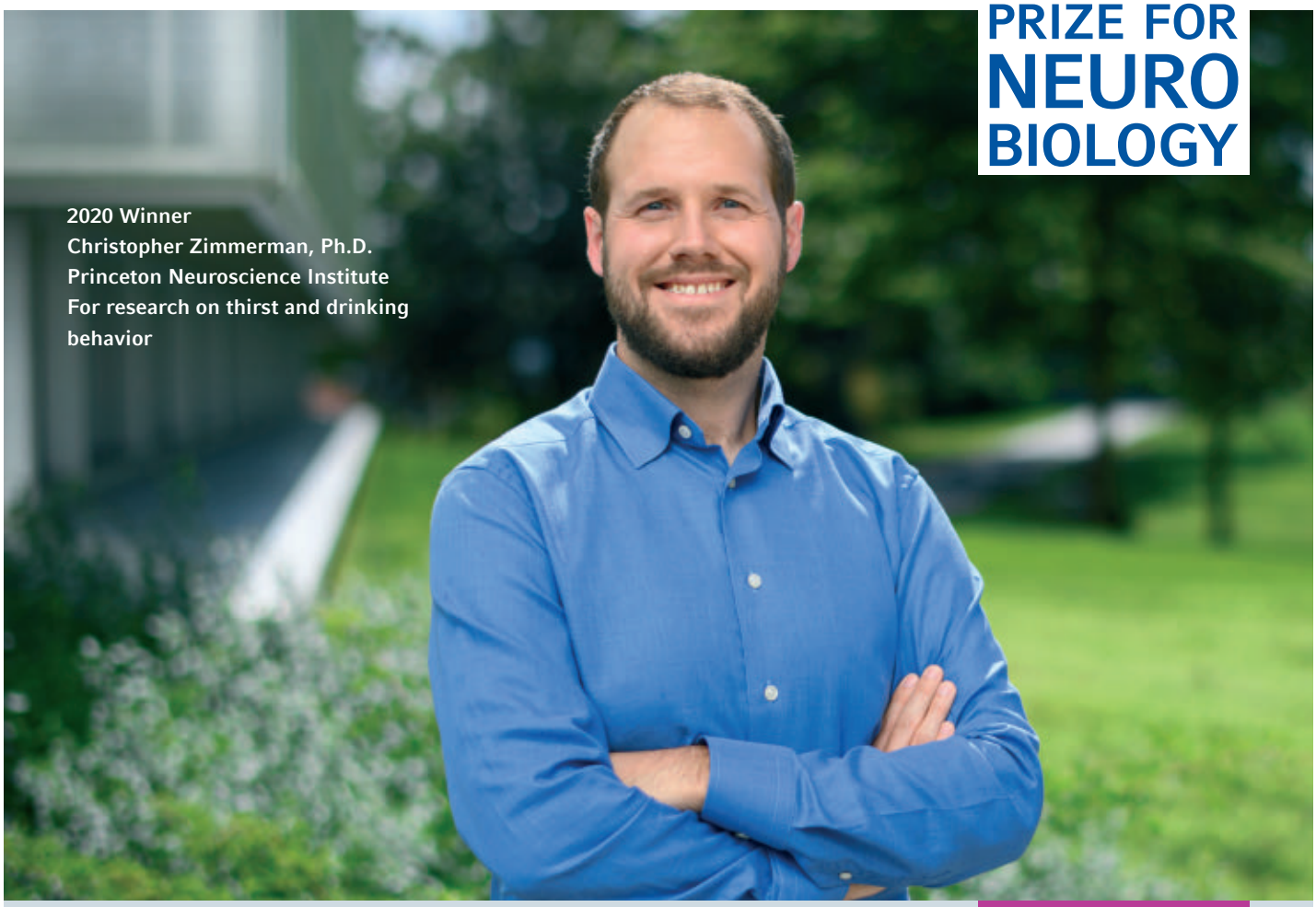
COMPUTING | IBM this week announced a prototype computer chip that packs 50 billion transistors in an area the size of a fingernail—a new record for density and computing power. The design meets an industry standard called 2-nanometer, which is based on a composite of metrics. For 5 decades, semiconductor engineers have steadily improved computing power by inventing new techniques to carve ever smaller features on chips, allowing them to pack more and more transistors closer together. IBM's new chips will be the first to rely exclusively on a technique known as extreme ultraviolet lithography. The company projects the chips will improve computing performance by 45%, or use 75% less power, than the current generation of 7-nanometer chips. After production begins in 2024, the new chips may power everything from cellphones that go 4 days between charges to more powerful supercomputers used in research.

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

PHOTO: JORGE GONZÁLEZ/ELENA SANTOS

eppendorf & Science

PRIZE FOR NEUROBIOLOGY



2020 Winner
Christopher Zimmerman, Ph.D.
Princeton Neuroscience Institute
For research on thirst and drinking behavior

Application Deadline
June 15, 2021

Now It's Your Turn!

Eppendorf & Science Prize for Neurobiology
The annual Eppendorf & Science Prize for Neurobiology is an international prize which honors young scientists for their outstanding contributions to neurobiological research based on methods of molecular and cell biology. The winner and finalists are selected by a committee of independent scientists, chaired by *Science*'s Senior Editor, Dr. Peter Stern. If you are 35 years of age or younger and doing great research, now is the time to apply for this prize.

As the Grand Prize Winner, you could be next to receive
> Prize money of US$25,000
> Publication of your work in *Science*
> Full support to attend the Prize Ceremony held in conjunction with the Annual Meeting of the Society for Neuroscience in the USA
> 10-year AAAS membership and online subscription to *Science*
> Complimentary products worth US$1,000 from Eppendorf
> An invitation to visit Eppendorf in Hamburg, Germany

It's easy to apply! Write a 1,000-word essay and tell the world about your work. Learn more at:

eppendorf.com/prize

eppendorf Science AAAS

AAAS® and Science® are registered trademarks of the American Association for the Advancement of Science, USA. Eppendorf® and the Eppendorf Brand Design are registered trademarks of Eppendorf AG, Germany. All rights reserved, including graphics and images. Copyright © 2021 by Eppendorf AG. Photography: Sameer A. Khan.



IN DEPTH

A team of researchers in rural Maharashtra state visits houses to track the spread of COVID-19 in India.

PHOTO: RAJA SENGUPTA

COVID-19

Is India's coronavirus death 'paradox' vanishing?

Limited evidence suggests in 2020 the country had relatively low COVID-19 mortality

By **Jon Cohen**, *in Vadu, India*

At a tiny rural hospital about 1 hour's drive northeast of Pune, India, in early April, workers loaded an SUV with coolers, syringes, vials, thermometers, and electronic tablets. They drove 20 minutes to the village of Karandi, slowing to pass caravans of migrant sugarcane cutters in ox carts. They spent more than an hour taking blood samples at a cluster of houses shared by three generations of one family. Later, the team would scour the blood for antibodies that indicate past run-ins with COVID-19.

Girish Dayma, who helps oversee this research program run by a satellite of King Edward Memorial (KEM) Hospital in Pune, says the team's surveys show that up to 40% of these villagers have antibodies for SARS-CoV-2. "It was thought that the rural area was not much affected," Dayma says. "The data are very much important to convince the policymakers that we need interventions in rural areas."

Studies like KEM's are also crucial to determining whether, as some researchers believe, India's horrific death toll is actually lower than expected from the rate of infections. Good data are scarce. Last week, hundreds of Indian researchers appealed to the government to release what it has and collect more. "[O]ur inability to adequately manage the spread of infections has, to a large extent, resulted from epidemiological data not being systematically collected and released in a timely manner," they wrote.

The current surge in COVID-19 cases has humbled those who thought the country had bested the disease. In early February, with cases dropping below 10,000 per day, restrictions were dropped, political leaders staged massive rallies, and masks became rare in many crowded locales. But the devastating surge starting in late March gave the lie to the suggestion that India might be approaching herd immunity; 10,000 cases hit Pune alone the day the KEM team visited Karandi. A few weeks later, India topped 400,000 cases in a single day.

Debate has swirled over whether new variants or waning immunity are at work, just how many people have become infected, and—most contentious—how many have died. Official figures suggest that, compared with other countries, India has recorded relatively few deaths given its count of COVID-19 cases. "We have been trying to find explanations for the low number of deaths in India since last year," says a signatory of the appeal, microbiologist Gagandeep Kang from the Christian Medical College, Vellore.

"The 'Indian paradox' really is quite puzzling," says Prabhat Jha, an epidemiologist at the University of Toronto. Explanations include underestimates of deaths, demographic effects, and environmental factors like abundant vitamin D from the Indian climate. But now, with hospitals struggling to find enough oxygen for their COVID-19 patients, crematoria overwhelmed, and media reports of intentional undercounting of deaths to make the current deluge look less dire, the seeming paradox may be disappearing.

In India's first wave, which ran from June through November 2020, cases never went above 100,000 per day. Hospitals struggled—the KEM intensive care unit in Pune for a time relied on raincoats instead of proper gowns—but few reached capacity with severely ill patients.

Even then, it was hard to nail down the magnitude of infections and death. "We rely on reporting of positive cases, which obviously leaves big gaps because a large percentage of people are asymptomatic, and a

Travel for this story was supported by the Pulitzer Center; *Science*'s COVID-19 reporting is also supported by the Heising-Simons Foundation.

lot of people don't have access to testing," says Soumya Swaminathan, chief scientist at the World Health Organization and a native of India. For mortality, she notes that only 20% of death certificates list a cause.

The notion of an Indian paradox surfaced as early as April 2020 and remains largely speculative despite frequent references by the health minister. One convincing study looked at 450,000 people who sought COVID-19 tests between June and the end of 2020 in 12 of the most populous Indian cities, including New Delhi, Mumbai, Pune, Kolkata, and Chennai. Led by Jha, it found that seropositivity over time jumped from about 17.8% to 41.4%, implying a huge increase in cases. Yet even after factoring in 30% underreporting of COVID-19 deaths—the worldwide average—the team calculated about 41 deaths from COVID-19 per 100,000 population, they reported in March on medRxiv. That mortality rate is less than half the corresponding U.S. figure.

Other studies, however, suggest the demographics of the outbreak could explain the anomaly. One thorough study looked at reported COVID-19 cases and deaths last spring and summer in two southern Indian states, Andhra Pradesh and Tamil Nadu, that are home to about 10% of the country's population. The researchers reported that older adults—the group at greatest risk of dying—accounted for relatively few of India's infections (*Science*, 6 November 2020, p. 691). Only 17.9% of the deaths in the study were in people age 75 or older, compared with 58.1% in that age bracket in the United States.

One reason is that India's population skews young. In 2011, the most recent census year, 45% of the population was 19 years or younger, and only 4.8% was 65 or older. And infection rates in the old were unusually low, perhaps because those who survive to old age in India are often wealthier and better able to socially distance, the researchers argue.

The results don't mean COVID-19 is any less deadly in India, notes the paper's first author, Ramanan Laxminarayan, an economist and epidemiologist who founded the Center for Disease Dynamics, Economics & Policy in Washington, D.C., and New Delhi. Unsurprisingly, increasing age was accompanied by a steady climb in the COVID-19 death rate, peaking at 16.6% in those 85 and older. "If you have 65% of your population in an age group where mortality rates are extremely low, then obviously, you're going to see an overall case fatality rate that's extremely low," he says. He calls claims of an Indian paradox "nonsense."

Other factors also help explain India's seemingly low death rates, Laxminarayan says. In the first wave, infections spread disproportionately in the urban poor, many of whom had to show up for work even during lockdowns, he says. Compared with wealthier city dwellers and those in rural villages, the urban poor are younger and have less obesity—characteristics linked with lower likelihood of severe COVID-19.

The states where the team worked have reliable death numbers because they started disease surveillance early, the researchers write. But elsewhere in the country, Laxminarayan suspects far more people have died than reported, noting that cases have been vastly underestimated. A study from the Indian Council of Medical Research found antibodies in 7.1% of nearly 29,000 people in 21 of India's 36 states and union territories. Published on 27 January in *The Lancet Global Health*, the findings imply that when the study finished collecting data in mid-August 2020, India had nearly 75 million cases—about 30 times higher than the case count then. "By that token, is it really unreasonable to think that deaths are underreported by a factor of four or five?" Laxminarayan asks.

Yet those who believe India's death rate is unusually low point to several factors. One, Jha says, is household structure. As with the family in Karandi, three generations in a home is a norm in many places. India's relatively small older population means young people are the most likely to bring COVID-19 into a household, and they tend to have lower levels of virus and more asymptomatic infections. Jha notes that reports suggest between 70% to 90% of infected people in India don't develop symptoms. As a result, older people tend to be exposed to lower doses of virus, which their immune systems may be more likely to control.

Some scientists have suggested genetics might also play a role. Anurag Agrawal, who heads the Council of Scientific & Industrial Research's Institute of Genomics and Integrative Biology, the leading contributor of a consortium that sequences SARS-CoV-2 in India, says genetic and environmental factors might be linked. Indians who live in the United States or the United Kingdom, he says, suffer just as much from severe COVID-19 as other people there. His team has its own "very controversial" theory, which it has yet to publish because the lead author fell ill with COVID-19. Some hotly debated studies have found lower rates of COVID-19 hospitalization in smokers, Agrawal notes. He points out that high COVID-19 death rates tend to occur in countries with the best air quality. His team contends that smokers and the many Indians who live with bad air pollution might overexpress a variation of an enzyme, CY1P1A1, that "detoxifies" the lungs and destroys the virus through a previously described phenomenon, "xenobiotic metabolism."

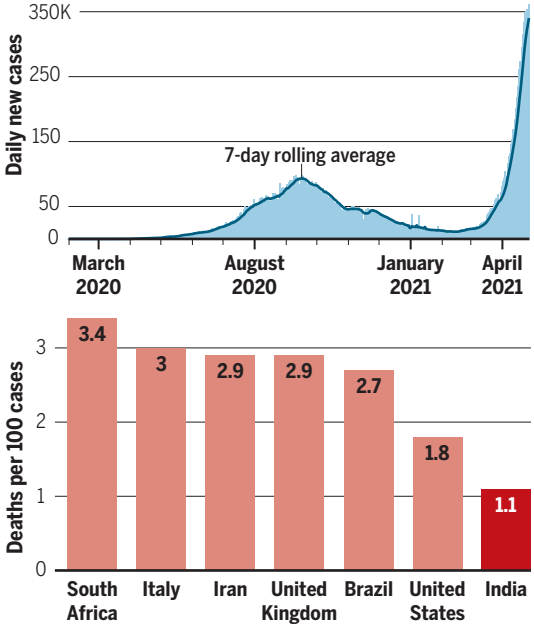
Jha and others are skeptical. "There's very little association with particulate matter and COVID infection cases or deaths in our analysis," he says.

The mortality pattern may shift during the current surge. This time, the virus appears to be causing serious illness in younger people more frequently and walloping wealthier populations. And Swaminathan notes that unlike in India's first wave, when hospitals never filled to capacity, "People are dying unnecessarily because health systems can't cope." But Jha says those trends are not dispelling the paradox. Recent data from Maharashtra suggest mortality rates of confirmed cases haven't changed much—deaths have surged catastrophically, but so have cases overall.

Only more and better data will resolve whether India is benefiting from a paradox and, if so, whether it will hold. Agrawal, who is in New Delhi, says India is in a wait-and-see mode. "It's just crazy here these days," he says. If patterns from other countries play out in India, he predicts the wave will begin to die down in mid-May. "Until then, we need to hold on." ■

A puzzle or a mirage

Even as millions have fallen ill in India, researchers have struggled to explain why COVID-19 mortality rates are lower there than in other countries.



CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) OUR WORLD IN DATA COVID-19 DATA REPOSITORY VIA JOHNS HOPKINS CENTER FOR SYSTEMS SCIENCE AND ENGINEERING; JOHNS HOPKINS CORONAVIRUS RESOURCE CENTER

COVID-19

Brazil and Russia face off over vaccine contamination charge

Does Sputnik V contain replication-competent viruses?

By **Sofia Moutinho** *and* **Meredith Wadman**

A high-stakes international fight broke out last week over the Russian COVID-19 vaccine known as Sputnik V, after Brazil declined to authorize its import. Although the Brazilian Health Regulatory Agency (Anvisa) cited a litany of reasons, including being denied access to the vaccine's quality control center and some production sites in Russia, its most surprising allegation was that the vaccine's second dose contains adenoviruses capable of replication, a potential danger to vaccine recipients.

The confrontation escalated when the vaccine's official Twitter account accused Brazil of bowing to "political" pressure and said "Sputnik V is undertaking a legal defamation proceeding" against the regulators. Anvisa officials fired back with a press conference, displaying documentation from the vaccine's maker and video clips of a meeting with its representatives. Although some observers initially took Anvisa's side, others soon wondered whether it had misinterpreted information from the vaccine's developer, the Moscow-based Gamaleya National Center of Epidemiology and Microbiology.

> **"Sputnik should have responded ... instead of starting a fight."**
> **Ricardo Gazzinelli**, Brazilian Immunology Society

Many outside the fray also felt Anvisa indicated originally it had directly tested for replicating viruses, but the agency later clarified it had relied on Gamaleya's reports. "The data we evaluated shows the presence of replicating virus," Gustavo Mendes, general manager of medicines and biological products at Anvisa, said at the press conference. Gamaleya issued a detailed rebuttal, declaring that Anvisa's allegations "have no scientific grounds and cannot be treated seriously." It added that a four-stage purification process prevents contamination.

The furor comes as Brazil, which has one of the highest burdens of COVID-19 in the world, has vaccinated only about 15% of its people with a first dose. Governors from 14 of its 26 states want to import 30 million doses of Sputnik V. Although it has not yet received the green light from the World Health Organization or the European Medicines Agency, more than 60 countries have authorized its use, and it is seen as an important weapon against the global pandemic. "We need this vaccine. It's cheap. It's effective. It's easy to store and transport," says Hildegund Ertl of the Wistar Institute, who studies adenovirus-based vaccines.

Several COVID-19 vaccines use adenoviruses to deliver the gene for the spike protein of SARS-CoV-2. Sputnik V relies on two, Ad26 and Ad5, in sequential doses. Both adenoviruses are stripped of *E1*, a gene that allows them to replicate, then mass-produced by cultured human cells with a stand-in copy of *E1*. But Ad5 is known to reacquire the gene from them on rare occasions.

A source of common colds, adenoviruses typically cause mild symptoms, but are occasionally lethal, and immunocompromised people could be at particular risk. Vaccinemakers therefore test for replicating adenoviruses. Anvisa said that although the standard worldwide is zero tolerance, Gamaleya set a much higher limit. Its documents displayed by Anvisa stated the tested vaccine batches had "less than 100" replication-capable particles per dose.

Some scientists suspect that wording merely reflects the sensitivity of Gamaleya's test or an arbitrary limit it established, not evidence that the vaccine contains active adenoviruses. Anvisa rejects that explanation.

The 14 governors sent new quality control documents from Gamaleya to Anvisa late last week and asked immunologist Amilcar Tanuri, at the Federal University of Rio de Janeiro's main campus, to independently review them. He tells *Science* that the additional information on the tested vaccine batches makes it clear that they had no replicating viruses.

Ricardo Gazzinelli, president of the Brazilian Immunology Society, calls the impasse "unnecessary turmoil." "Sputnik should have responded to [Anvisa's concerns] instead of starting a fight." Anvisa officials say that if Gamaleya clarifies the issues raised by the regulatory agency, the import ban can still be reversed. "Sputnik V is not out of the question" for use in Brazil, Mendes says. ■



REMOTE SENSING

NASA set to announce Earth system observatory

Climate-monitoring satellites mark resurgence for agency's earth science division

By **Paul Voosen**

NASA is about to announce its next generation of Earth-observing satellites. As soon as this month, it will lay out preliminary plans for a multibillion-dollar set of missions that will launch later this decade. This "Earth system observatory," as NASA calls it, will offer insights into two long-standing wild cards of climate change—clouds and aerosols—while providing new details about the temperatures and chemistry of the planet's changing surface. The satellite fleets also mark a revival for NASA's earth science, which has languished over the past decade compared with exploration of Mars and other planets.

Although officials have been planning the missions for several years, the Biden administration is accelerating them as part of its focus on addressing climate change. "Earth system science is poised to make an enormous difference in our ability to mitigate, adapt to, and plan for changes we're seeing," says Karen St. Germain, director of NASA's earth science division. "The pace we're going to have to do that is much higher in the decade in front of us than the decade behind us."

Agency spokespeople declined to provide details about the missions because they have not yet been formally approved. But at a workshop last month, Charles Webb, an associate director for flight programs at NASA, said four missions would go ahead, launching as soon as 2028—an acceleration of plans under the Trump administration, when only two missions were scheduled to begin development. "It

PHOTO: SANTIAGO BORJA



Planned NASA satellites will measure the vertical motions that lead to towering storm clouds.

became pretty clear the greatest science return is having all of these in operation close to each other," Webb said. The missions lack official names, but go by the shorthand of ACCP (Aerosol, Cloud, Convection, and Precipitation), which covers two missions; Surface Biology and Geology (SBG), and Mass Change, which would measure tiny variations in gravity indicating changes in ice and water.

The administration's proposed 15% bump for NASA's earth science budget for next year, to $2.3 billion, would help fund the accelerated program. The increase would also be welcome news for NASA's earth science researchers after 2 decades operating under administrations leery of climate. Jeremy Werdell of NASA's Goddard Space Flight Center recalls multiple attempts by the Trump administration to kill Pace, a $900 million ocean-monitoring satellite for which he is principal investigator. "You'd see the chart with all the upcoming missions, and you'd see yours isn't there." (The mission survived and is due to launch in 2023.) However, even President Joe Biden's proposed investment would leave NASA's inflation-adjusted spending on earth science below the levels 20 years ago, says Waleed Abdalati, director of the Cooperative Institute for Research in Environmental Sciences at the University of Colorado, Boulder, and former NASA chief scientist. "We're well behind where we were."

The ACCP satellites, with a total cost of up to $1.6 billion, would replace CloudSat and CALIPSO, probes launched in 2006 that have provided clarity on how clouds and aerosols can either slow global warming, by reflecting sunlight, or speed it by trapping heat. CloudSat, able to spot previously invisible light rain and snowfall, provided the first global picture of total precipitation. The pair also showed that certain aerosols, when mixed with clouds, can suppress rainfall and extend clouds' lifetimes and their ability to block light. "We learned yes, this is an effect we can observe," says Tristan L'Ecuyer, a climate scientist at the University of Wisconsin, Madison—and one that human-caused pollution, like soot, may amplify.

But the satellite duo is aging: CloudSat has lost two of the four reaction wheels it uses to stabilize its orbit, and CALIPSO is out of fuel and onto its backup laser. The three or more satellites of the ACCP missions will pick up where these two left off, with more precise measurements that can reach closer to the planet's surface to quantify the response of low clouds to warming. They will be armed with two primary instruments: an advanced laser capable of identifying individual aerosol types, such as smoke from an volcanic eruption or dust blown off the Sahara; and several bands of radar, including a Doppler instrument that can detect the vertical convective motions that drive major storms. "We're still trying to understand why hurricanes intensify," L'Ecuyer says.

SBG, the chemistry and temperature mission, which would cost up to $650 million, will include a satellite with a high-resolution hyperspectral imaging spectrometer that can subdivide reflected light into more than 400 wavelength channels across the visible and into the infrared. It serves as a molecular mapping system, sensitive to the spectral signatures of specific gases in the air column below or compounds at the surface.

By measuring the intensity of green chlorophyll or detecting the signatures of excess salts or fungus, for example, researchers can evaluate the health of crops and forests. They will be able to prospect for minerals in remote regions, map and identify different coral and algae species, and track plumes of greenhouse gases. "Spectroscopy truly is the future of planetary imaging," says Greg Asner, an ecologist at Arizona State University, Tempe, who has tested airborne prototypes of the spectrometer. (A smaller scale version will be launched and mounted on the space station next year.)

Another satellite in SBG would host a thermal radiometer, capable of capturing the heat coming off Earth's surface. Such an instrument is already mounted on the space station for the Ecostress mission, where it watches for the high leaf temperatures that result when plants run short of water. Ecostress data also reveal the heat of wildfires and volcanoes, track warming ocean currents, and pinpoint urban heat islands. SBG would cover the whole globe, not just the midlatitudes as Ecostress does, while capturing three times as much land with each path. "It's a really rich future," says Simon Hook, Ecostress's principal investigator at NASA's Jet Propulsion Laboratory.

Less is certain about Mass Change, beyond that the mission—likely a partnership with European space agencies—will continue the measurements made by the Gravity Recovery and Climate Experiment Follow-On (GRACE-FO), which calculates tiny shifts in Earth's gravity from the fluctuating distance between two satellites flying in tandem. GRACE-FO has allowed scientists to measure ice loss in Greenland and Antarctica and monitor water loss in soil and underground aquifers due to drought. Mass Change will include a tandem of satellites similar to GRACE-FO, which flies in an orbit around the poles. Scientists also hope the mission will fly a second set in a different orbit to reduce distortions in the gravity signal it collects.

With NASA planning to turn powerful new eyes on Earth, NASA's earth scientists see a brighter future for their work, if not for the beleaguered planet. "We're in a time of challenge, there's no doubt about that," St. Germain says. "But it's also a time of unprecedented opportunity." ■



Streetlights lure mayflies by the millions to their deaths, as seen on this bridge in Navarra, Spain. Researchers are testing ways to keep the insects closer to the water.

CONSERVATION ECOLOGY

Fatal attraction to light at night pummels insects

Fearing artificial lights add to an "insect apocalypse," researchers seek solutions

By **Elizabeth Pennisi**

Each summer, on bridges across the world, mayfly massacres occur. First, warm weather prompts the transformation of the insects' aquatic larvae. Within hours, the short-lived, flying adults pop out of streams, rivers, and lakes, eager to mate and lay eggs by the millions.

But bridges illuminated with artificial light can lure the newly emerged adults away from the water to a futile death before breeding. Others, fooled by the sheen of reflective pavement, drop their eggs on the bridge road instead of the water. Because mayflies control the growth of algae and are food for fish, the fate of these humble insects may reverberate through ecosystems, says Ádám Egri, a biological physicist at the Centre for Ecological Research in Budapest, Hungary, who is working to save endangered mayflies there.

Mayflies aren't alone in their fatal attraction to what researchers refer to as ALAN: artificial light at night. Studies from around the globe are finding worrisome impacts on insect mating and abundance, says Stéphanie Vaz, an entomologist at the Federal University of Rio de Janeiro's main campus. In the past year, researchers have published the first experimental and regional studies of the problem, and in March, *Insect Conservation and Diversity* devoted a special issue to the topic.

Some researchers think brighter nights may be a factor in recently documented insect declines (*Science*, 12 May 2017, p. 576), says Stephen Ferguson, a physiological ecologist at the College of Wooster. With insect numbers dropping by 80% in some places and 40% of insect species headed for extinction by some estimates, "Some researchers have started to make more noise about the 'insect apocalypse,'" Ferguson says. "ALAN is almost certainly one of the drivers."



In Grand Teton National Park, a new system of dimmer, reddish lights attracts fewer insects—and lets visitors see the stars.

Even as they begin to raise the alarm, scientists are pointing to simple solutions. Egri, for example, has found that mounting bright lights low on the sides of bridges keeps the mayflies close to the water. But researchers are "still at the very beginning of the story of global, ecologically friendly artificial lighting," he says.

Many insects and other animals are drawn to light because they depend on the Moon or Sun for navigation, Ferguson says. And light at night is increasing by up to 40% per year, according to ALAN researcher Franz Hölker at the Free University of Berlin, who calculated this estimate using satellite, energy use, and other data. Cities are using more light-emitting diodes, whose blue light is brighter than the yellow glow of sodium vapor streetlights.

Even dark areas are no longer very dark. "Protected areas are not able to buffer these light intensities as we thought," Vaz says. On Moon-less nights, artificial sky glow now exceeds the combined light of stars and other natural sources on 22% of the globe's total land, with biodiversity hot spots dispropor-

PHOTOS: (TOP TO BOTTOM) EDUARDO BLANCO/MINDEN PICTURES; BARBER LAB/HUNTER COLE

tionately affected, Brett Seymoure, a behavioral ecologist at Washington University in St. Louis, and his colleagues report in the preprint elibrary SSRN.

Given the many other factors also hurting insects, such as habitat degradation and climate change, linking light to species' declines is challenging. "It is a very understudied field," Hölker says. But scattered studies suggest the impact may be powerful. He and others have calculated that Germany's 9 million streetlights attract about 1 billion insects a night, many of which die or are killed by bats and other predators. Researchers have estimated that at least one-third of the insects swarming around artificial lights die of exhaustion or are eaten by predators.

One recent study underscores the magnitude of the effect. On the night of 27 July 2019, the glow of Las Vegas lights lured massive numbers of migrating grasshoppers into the air above the city, according to a 31 March paper in *Biology Letters*. The clouds of grasshoppers were visible on weather radar; by estimating numbers of insects seen on radar before, during, and after the swarm, Elske Tielens, an ecologist at the University of Oklahoma, Oklahoma City, and her colleagues calculated that at its peak, the swarm weighed 30.2 tons and contained 48 million grasshoppers.

There were "more grasshoppers in the air on that single July night than human visitors to Las Vegas in a whole year," Tielens says. "This is probably happening on smaller scales in many places, and with many more insects," Ferguson adds.

In the Netherlands, a consortium of universities, nonprofit organizations, industry, and government is exploring light's effects on local ecosystems through the Light on Nature project. It set up long-term experiments in seven sets of plots in dark areas. The researchers lit up some plots with lights of different colors and monitored bat and insect communities. Between 2012 and 2016, moth numbers remained steady in dark plots but decreased 14% in lighted areas, Roy van Grunsven, an entomologist at Dutch Butterfly Conservation, and colleagues reported in June 2020 in *Current Biology*.

"This study represents the only published experimental evidence to date" about ALAN's long-term effects, says Daniel Boyes, an entomologist at the UK Centre for Ecology and Hydrology in Wallingford. "The bottom line is that moths are being bombarded with unnatural night conditions that their sensory systems are not adapted for," Seymoure adds.

Most of the research on artificial light so far has taken place in temperate climates. But Vaz's modeling studies point to light pollution as a possible cause for a decline in firefly diversity in Brazil's Atlantic Forest. And Jessica Deichmann, an applied ecologist at the Smithsonian Conservation Biology Institute, documented what happens when electric lights were first turned on in a remote tropical forest in Peru. "I've witnessed firsthand the truly massive storm clouds of insects drawn to lights when they are first installed, and this sight is hard to forget," she says. Most of the insects, particularly flying ants and flies, die of exhaustion or are eaten.

She worries the nightly tolls will curtail pollination and other ecosystem services provided by these species. So, like more and more ALAN researchers, she is seeking solutions. Her team set up experimental plots in the forest lit by lights of different colors and discovered amber lights attracted 60% fewer insects than white light.

> **"Substantial areas of our planet [must] remain dark forever."**
>
> **Jessica Deichmann**, Smithsonian Conservation Biology Institute

But what's good for some flying insects may be bad for others, as Tufts University graduate student Avalon Owens described in January at a virtual meeting of the Society for Integrative and Comparative Biology. Owens evaluated how fireflies and other flying insects reacted to red, blue, and amber light in Kellettville, Pennsylvania, a rural area with little light pollution and so many *Photinus carolinus* fireflies that the town hosts an annual firefly festival. Observing fireflies in the wild, "I found red light is 'best,' and amber is 'worst' for interfering with courtship," she says.

In the lab, she found that in amber light, "females go almost completely dark," leaving males no way to find them, she and her colleagues reported in the special issue.

Egri and his colleagues, too, tested the impact of color, hanging beacons of different hues low on a bridge, then photographing and counting mayflies. Blue lights, being even brighter than the yellowish road lights, kept more insects close to the water. For two springs now, blue beacons installed on the Tahitótfalu bridge in northern Hungary have shone for 3 hours past sunset, while lights on the roadway are dimmed. This seems to work, Egri says. "No mayflies left the river."

Elsewhere, dimmer, redder lights are being tested, including at a visitor center in Grand Teton National Park. But Egri says his own effort and others "are still too little." Deichmann agrees that more ambitious measures are needed. For the sake of insects and ecosystems, "It is absolutely essential to ensure substantial areas of our planet remain dark forever." ■

HELIOPHYSICS

'Campfires' may drive heating of solar atmosphere

Observations suggest small flares are corona's mystery heat source

By **Daniel Clery**

Researchers are getting closer to solving an enduring mystery: why the Sun's wispy atmosphere is nearly 200 times hotter than its surface. Results from new instruments that see the Sun in sharper spatial and temporal resolution, presented last week at the European Geosciences Union (EGU) meeting, suggest tiny, flickering flares, dubbed "campfires," could be sufficient to heat the atmosphere.

"It's still one of the major unsolved problems in solar physics," says Sarah Matthews a solar physicist at University College London. "Campfires are a tantalizing prospect."

Temperatures ought to decline as one moves out from the Sun's core, where its fusion fires burn. However, in the 1930s scientists discovered that the solar atmosphere, or corona, seethes at more than 1 million degrees Celsius, far hotter than the 5500°C of the surface. Flares have been a prime suspect for providing the missing heat. Researchers think these eruptions of energy and particles occur when magnetic field lines, looping up into the corona, become twisted and suddenly snap into a configuration of lower stress, a process known as magnetic reconnection that releases huge amounts of energy.

Solar telescopes on Earth and in space reveal large magnetic loops 10 times the diameter of Earth flaring several times a day when the Sun is magnetically active. At those rates, the flares cannot provide enough heat to maintain the corona's temperature. So researchers in the 1960s proposed that the Sun is constantly erupting in "nanoflares" too small to be seen.

Enter Solar Orbiter, launched by the European Space Agency just over 1 year ago. In May 2020, while still maneuvering toward its operational orbit, the spacecraft ob-

served the corona from a vantage halfway between the Sun and Earth using its Extreme Ultraviolet Imager (EUI). At extreme ultraviolet wavelengths, hot gases deep in the corona shine brightly, revealing a welter of brief brightenings the team called "campfires": almost 1500 of them during a 4-minute observation, ranging in size from 400 to 4000 kilometers long. "We didn't expect them to stand out so clearly," says EUI Principal Investigator David Berghmans, a solar physicist at the Royal Observatory of Belgium. The team was especially surprised to find these nanoflares going off on the Sun's so-called "quiet" regions, where magnetic activity is low and big flares rare. Seeing them there, Berghmans says, is "like a palm tree at the North Pole."

The question solar physicists now want answered is: "Are [campfires] enough to power the quiet Sun on average?" says Solar Orbiter project scientist Daniel Müller at the European Space Agency's research center in the Netherlands. The EUI alone can't settle the question, but, Müller says, "Indications are that they could."

Part of that confidence comes from computer models, such as one presented at EGU by a team led by Yajie Chen of Peking University and Hardi Peter of the Max Planck Institute for Solar System Research. Their 3D model extends from below the Sun's turbulent surface up into the corona and across an expanse of the quiet Sun similar to that observed by EUI. The "brightenings" it produced were remarkably similar to the larger campfires in size and duration, the team reported. "Everything we checked matched," Peter says. The model also suggested the nanoflares could sustain the million-degree temperatures of the quiet parts of the corona.

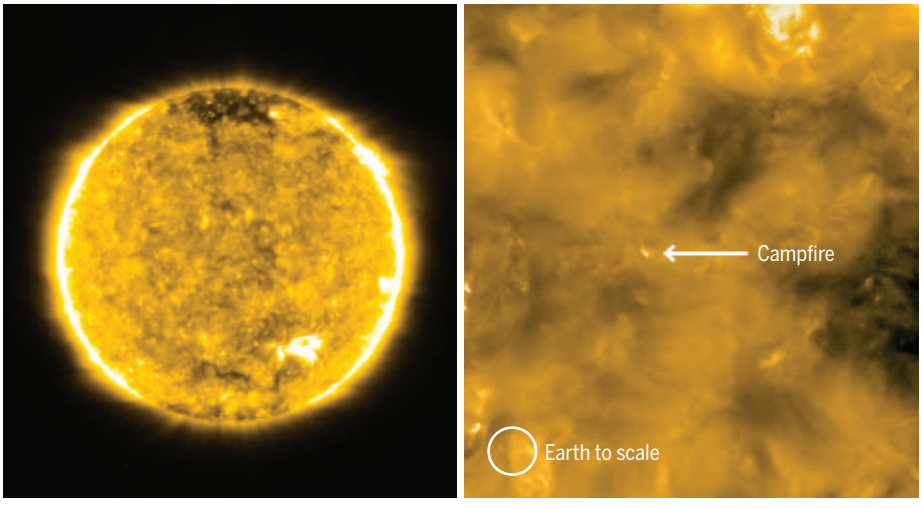
Radio observations of the Sun support that picture. At EGU, researchers from India's Tata Institute of Fundamental Research reported data from the Murchison Widefield Array, a radio telescope in Western Australia made up of 4100 spiderlike wire antennas spread across several kilometers. During a 70-minute observation of quiet regions on the Sun, the team recorded 20,000 bright radio flashes, each lasting 1 second or less, with estimated energies matching theoretical predictions for nanoflares. "It jumped out of the data," says Tata's Surajit Mondal.

Mondal adds that it's difficult to estimate the energy of a flare from its radio signal alone. But, he says, "If there is this number of flares, it can maintain coronal heating."

The mystery of the corona's heating isn't solved yet—other processes besides the small flares could contribute. The main rival is waves, carrying energy from the churning plasma below the Sun's surface up into the corona. But researchers are struggling to explain how the waves would break and deposit their energy into the diffuse coronal plasma—and whether they arise independently of the flares. "Flare processes also drive wave modes, they are intertwined," Matthews says.

Solar physicists will have their best chance yet to crack this problem in the next few years, as Solar Orbiter teams up with NASA's Parker Solar Probe, launched in 2018, and the Daniel K. Inouye Solar Telescope, currently being commissioned on a mountaintop in Hawaii. Solar Orbiter will be fully in place in 2022, swooping down every 6 months to just one-quarter of the Sun-Earth distance.

Never before has a telescope stared directly at the Sun from such a close vantage. Peter says: "It will be an exciting year." ■



Ultraviolet images from Solar Orbiter revealed "campfires" (arrow), small flares that could power coronal heating.

IMAGES: SOLAR ORBITER/EUI TEAM/ESA & NASA; CSL; IAS; MPS; PMOD/WRC; ROB; UCL/MSSL

ANIMAL RESEARCH

USDA now only partially inspects some animal labs

Internal documents reveal agency moved to "focused" inspections to save work

By **David Grimm**

In February 2019, the U.S. Department of Agriculture (USDA) made a significant—and apparently secret—change to how it oversees laboratory animal welfare, *Science* has learned. Instead of fully inspecting all of the nearly 1100 facilities that house monkeys, rabbits, and other creatures used in biomedical research, it mandated partial "focused" inspections for labs accredited by a private organization of veterinarians and scientists called AAALAC International.

Such partial inspections violate the Animal Welfare Act, argues Katherine Meyer, director of Harvard University's Animal Law & Policy Clinic, which discovered the change after law student Brett Richey combed through more than 1000 pages of internal USDA documents. The federal law states that USDA must enforce "minimum requirements for handling, housing, [and] feeding" of research animals, as well as adequate veterinary care, Meyer notes. "How do you ensure that labs are in compliance with those standards if USDA is doing incomplete inspections?"

USDA counters that it still inspects all lab animal facilities, as required by law. The agency "is not using AAALAC inspections. [It] is conducting focused inspections of research facilities because facilities that are AAALAC accredited generally have better compliance records, and we can expend less resources on said facilities," a spokesperson wrote in an email to *Science*. In deciding how and when to inspect facilities, he said, "AAALAC accreditation is one factor that is considered, along with the facility's history of noncompliances based on our previous inspections."

USDA's approach is legal, says Larry Carbone, former associate director of the University of California, San Francisco's



A U.S. Department of Agriculture inspector examines ferret cages. A new policy mandates that inspectors do lighter inspections at certain lab animal facilities.

PHOTO: U.S. DEPARTMENT OF AGRICULTURE

animal care facility and an expert on animal welfare policy. There are also other safeguards for animal welfare, he notes. Still, he says, the agency's secrecy about the policy is "troubling."

Harvard's animal law clinic obtained the USDA documents—including PowerPoint slides, FAQs, and email exchanges—via a public records request, in the course of a separate effort to force the agency to update its welfare standards for research monkeys. In one document, USDA, citing concerns about workload, says it has "made it mandatory … for inspectors to perform focused inspection at AAALAC-accredited research facilities unless the research facility requested a full inspection." And, in a redacted line that *Science* has since uncovered, it adds, "This focused inspection counts as the facility's annual inspection."

Other USDA documents suggest what a "focused inspection" entails. Instead of taking a full look at an institution's records, animals, and the facility itself (such as air conditioning units and surgery rooms) every year, an inspector only needs to look at one of those aspects or a sampling of them, according to a USDA PowerPoint. "An inspector could just look at a sampling of paperwork—and not a single animal," Meyer argues.

The Federation of American Societies for Experimental Biology (FASEB) and other organizations that promote animal research have argued that USDA should consider AAALAC accreditation when assessing whether a facility is at risk for animal welfare violations. "AAALAC is an indicator that a facility goes above and beyond," says Naomi Charalambakis, FASEB's senior science policy analyst.

But AAALAC accreditation should supplement, rather than supplant, USDA inspections, she says. FASEB has not recommended "that USDA should allow limited inspections just because a facility is AAALAC accredited," adds Jennifer Zeitzer, the organization's director of public affairs.

The internal USDA documents also imply the new policy is confidential. The guidelines are "for official use, internal only," reads one FAQ. "There will be no stakeholder announcement," and the details will not be included in USDA's official inspection guidelines, it continues. The agency also made no mention of the changes at a conference last month for heads of animal care facilities and others responsible for lab animal welfare.

"The apparent sneakiness on USDA's part is striking," Carbone says. He says he's also "disturbed" that the agency—as internal documents indicate—has told its inspectors not to ask a facility for proof of its AAALAC accreditation, or to probe whether it's on AAALAC probation. "I'd for sure think they'd want to know *why* a place was put on AAALAC probation."

Meyer argues that AAALAC itself is problematic. She and others in the animal advocacy community have long contended that the organization is not an appropriate substitute for USDA because it only inspects institutions every 3 years, is made up of many of the same people that run animal research facilities, and—unlike USDA—schedules inspections in advance rather than showing up unannounced. (AAALAC did not respond to repeated requests for comment.)

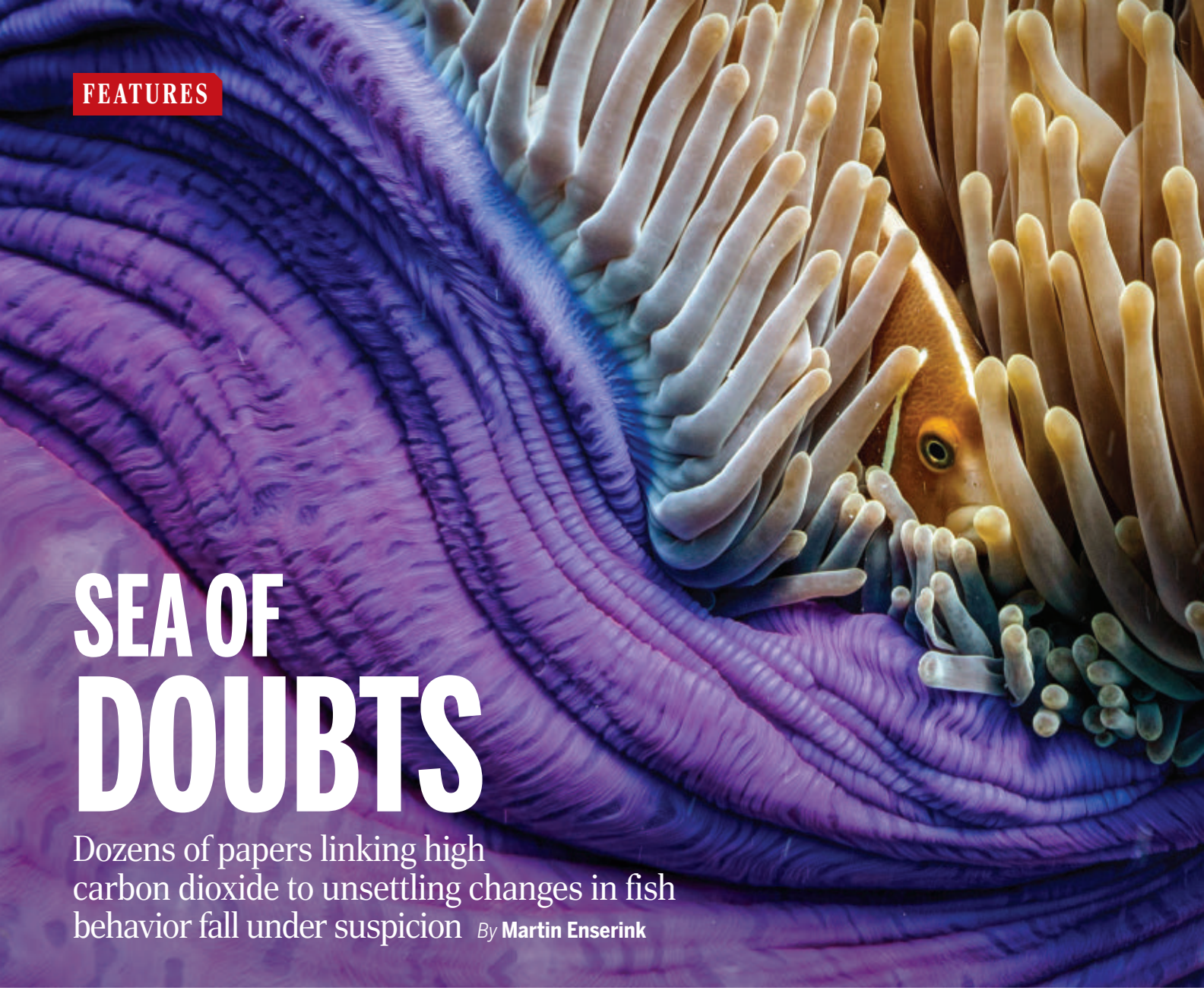
Over the past 5 years, USDA reports have cited several AAALAC-accredited facilities for critical violations of the Animal Welfare Act, including pain relievers not given in a timely manner and unmonitored animals bleeding to death inside their cages. Because AAALAC inspections are confidential—even to USDA—the agency could be giving such labs a minimal inspection and a "clean bill of health" under the new policy, Meyer says.

Whether that is actually happening is unclear. USDA has inspected 322 research facilities since its new guidelines took effect, an internal document reveals. Of the 151 that the agency believed were AAALAC accredited, 91 got focused inspections. (Some facilities requested full inspections, the document states, "because they were concerned about the appearance of a 'focused inspection.'") But inspectors are not to note in their report what aspect of a facility they looked at, internal documents indicate. So an outsider would not be able to tell whether USDA missed potentially critical animal welfare violations because, for example, it didn't look at animals that year.

Some animal care experts support a heavier reliance on AAALAC. It "would benefit all involved," says Troy Hallman, director of Yale University's Office of Animal Research Support. In his experience, he says, four AAALAC inspectors come for 3 days, versus one USDA inspector for 2 days; AAALAC also looks at mice, rodents, and other creatures not covered by USDA. "'Problematic' labs would be identified just as easily by AAALAC as USDA," Hallman says.

Yet USDA's own inspections are vital, says Allyson Bennett, senior editor at Speaking of Research, which advocates for animal studies. "USDA external oversight is critical to ensuring open public view of animal research, she says. "AAALAC can be wonderful, but it's a private body with relatively few mechanisms for transparency."

Bennett adds, however, that it's unrealistic to expect USDA to conduct thorough annual inspections without more funding, given the agency's limited resources and the approximately 800,000 animals it oversees. "It's a big job," she says. "If we want USDA to do more, then it needs to be able to do more." ■



SEA OF DOUBTS

Dozens of papers linking high carbon dioxide to unsettling changes in fish behavior fall under suspicion

By **Martin Enserink**

When Philip Munday discussed his research on ocean acidification with more than 70 colleagues and students in a December 2020 Zoom meeting, he wasn't just giving a confident overview of a decade's worth of science. Munday, a marine ecologist at James Cook University (JCU), Townsville, was speaking to defend his scientific legacy.

Munday has co-authored more than 250 papers and drawn scores of aspiring scientists to Townsville, a mecca of marine biology on Australia's northeastern coast. He is best known for pioneering work on the effects of the oceans' changing chemistry on fish, part of it carried out with Danielle Dixson, a U.S. biologist who obtained her Ph.D. under Munday's supervision in 2012 and has since become a successful lab head at the University of Delaware (UD), Lewes.

In 2009, Munday and Dixson began to publish evidence that ocean acidification—a knock-on effect of the rising carbon dioxide (CO_2) level in Earth's atmosphere—has a range of striking effects on fish behavior, such as making them bolder and steering them toward chemicals produced by their predators. As one journalist covering the research put it, "Ocean acidification can mess with a fish's mind." The findings, included in a 2014 report from the Intergovernmental Panel on Climate Change (IPCC), could ultimately have "profound consequences for marine diversity" and fisheries, Munday and Dixson warned.

But their work has come under attack. In January 2020, a group of seven young scientists, led by fish physiologist Timothy Clark of Deakin University in Geelong, Australia, published a *Nature* paper reporting that in a massive, 3-year study, they didn't see these dramatic effects of acidification on fish behavior at all.

The paper has proved so polarizing in the field, "It's like Republicans and Democrats," says co-author Dominique Roche of Carleton University in Ottawa, Canada. Some scientists hailed it as a stellar example of research replication that cast doubt on extraordinary claims that should have received closer scrutiny from the start. "It is by far the best environmental science paper I have read for a long time," declared ecotoxicologist John Sumpter of Brunel University London.

Others have criticized the paper as needlessly aggressive. Although Clark and his colleagues didn't use science's F-word, fab-

PHOTO: FREDRIK JUTFELT



Orange clownfish are among the tropical species studied in 22 papers now facing scrutiny.

rication, they did say "methodological or analytical weaknesses" might have led to irreproducible results. And many in the research community knew the seven authors take a strong interest in sloppy science and fraud—they had blown the whistle on a 2016 *Science* paper by another former Ph.D. student of Munday's that was subsequently deemed fraudulent and retracted—and felt the *Nature* paper hinted at malfeasance. The seven were an "odd little bro-pocket" whose "whole point is to harm other scientists," marine ecologist John Bruno of the University of North Carolina, Chapel Hill—who hasn't collaborated with Dixson and Munday—tweeted in October 2020. "The cruelty is the driving force of the work."

What few researchers know is that in August 2020, Clark and three others in the group took another, far bigger step: They asked three funders that together spent millions on Dixson's and Munday's work—the Australian Research Council (ARC), the U.S. National Science Foundation (NSF), and the U.S. National Institutes of Health (NIH)—to investigate possible fraud in 22 papers.

The request, which they shared with a *Science* reporter, rests on what they say is evidence of manipulation in publicly available raw data files for two papers, one published in *Science*, the other in *Nature Climate Change*, combined with remarkably large and "statistically impossible" effects from CO_2 reported in many of the other papers. They also provided testimony from former members of the Dixson and Munday labs, some of whom monitored Dixson's activities and concluded she made up data.

ARC and NSF declined to discuss the case with *Science*, but said they generally refer such cases to the research institutions—in this case JCU; the Georgia Institute of Technology, where Dixson worked between 2011 and 2015; and UD. NIH said it refers cases to the U.S. Office of Research Integrity, which does not comment on cases.

Munday calls the allegations of fraud "abhorrent" and "slanderous," and a JCU spokesperson says the university has dismissed the allegations after a preliminary investigation. (Munday retired from JCU in April and has moved to Tasmania, but emphasizes there is no connection between that timing and the allegations.) UD says it cannot comment on personnel matters; a Georgia Tech spokesperson declined to comment except to say the institute "takes all allegations of research misconduct seriously." Dixson denies making up data as well. "I fully stand by all the data I've collected, I stand by the papers that we've published," she told *Science* in a February interview. "The data was collected with integrity. I mean, I preach that to my students."

But multiple scientists and data experts unconnected to the Clark group who reviewed the case at *Science*'s request flagged a host of problems in the two data sets, and one of them found what he says are serious irregularities in the data for additional papers co-authored by Munday.

The fight, between two groups united by their passion for fish, isn't just about data and the future of the oceans. It highlights issues in the sociology, psychology, and politics of science, including pressure on researchers to publish in top-tier journals, the journals' thirst for eye-catching and alarming findings, and the risks involved in whistleblowing.

Members of the Clark group say they will soon publicize the alleged data problems on PubPeer, a website for discussion of published work. And they say they thought long and hard about whether to discuss their concerns with a reporter while investigations may be ongoing. "In my experience, whistleblowers, myself as well as others, are shamed for talking to the media before an investigation has concluded misconduct," says Josefin Sundin of the Swedish University of Agricultural Sciences, the last author on the *Nature* replication paper. "But why is that? If an investigation even takes place, it can drag on for a very long time. If you know that data have been fabricated, why is it considered the right thing to do to stay silent about it for months and even years?"

TOWNSVILLE MAY BE one of the world's best places to go if you want to become a marine biologist. JCU's website boasts "a unique tropical learning environment with research stations, state-of-the-art laboratories and the Great Barrier Reef right on our doorstep." The university is home to the ARC Centre of Excellence for Coral Reef Studies, where Munday had his lab. For field studies, scientists fly to Lizard Island, a granite rock on the reef that's legendary among marine biologists, thanks to the Australian Museum's well-run research station, otherworldly diving, and beach barbecues. ("Man, I miss that island," Dixson tweeted last year.)

Munday's lab has been one of the engines of JCU's success. Originally focused on competition in reef fish and their ability to switch sex, Munday shifted his attention to ocean acidification in the late 2000s. Dixson, who arrived at JCU in 2007, embraced the topic.

Acidification, which results as rising CO_2 levels cause more of the gas to dissolve in the ocean, poses serious threats to ocean life, weakening corals and other organisms with carbonate shells or skeletons. But a 2009 paper in the *Proceedings of the National Academy of Sciences* (*PNAS*) on which Munday and Dixson were the first and second author, respectively, reported another troubling effect. When the pH of the water in fish tanks was lowered from 8.15, the current level in ocean water, to 7.8, the level expected by the end of the 21st century, larvae of the orange clownfish were less attracted by the chemical cues from a healthy reef—but more attracted to cues from grass and a pungent swamp tree whose smell normally repels them. That could cause clownfish to lose their ability to find suitable homes on the reef, the authors concluded.

Based on more lab experiments, Munday, Dixson, and other researchers later reported that high CO_2 levels mess with fish minds in other ways as well: They become disoriented, hyperactive, and venture farther from shelter, for instance, while their vision and hearing deteriorate. For many of those studies,





Fredrik Jutfelt and Timothy Clark (right) teamed up with other marine biologists to replicate key results in the challenged papers, without success.

the scientists measured fish's preferences by placing them in a flume, an apparatus that forces them to make a choice (see graphic, p. 564). Water from two different sources flows into the flume, side by side, and researchers measure how much time the fish spend in water from either source.

Munday and Dixson often found unusually large effects from ocean acidification. In the *PNAS* paper, for example, the time orange clownfish spent on the foul-smelling side of the flume went from 0% to 80%. In a 2010 study in *Ecology Letters*, clownfish larvae reared in normal ocean water completely avoided chemical cues of two predator species, the small rockcod and the dottyback, but in more acidic water they spent 100% of their time around those predators' scents—a "fatal attraction," the authors said. A 2013 paper in *Marine Biology* reported that coral trout, an economically important species, became 90 times more active at a high CO_2 level.

Dixson also used the flume for studies not related to ocean acidification. The *Science* paper the whistleblowers have challenged, published in 2014 while she was at the lab of Georgia Tech marine ecologist Mark Hay, showed that fish and coral larvae collected on Fiji's coast are attracted by the chemical cues from healthy, protected reefs, but repelled by water from overfished reefs dominated by seaweeds. (It was, again, bad news, suggesting degraded reefs will have trouble recovering on their own because they're unable to entice coral and fish larvae.)

Not long after that paper was published, *Science* received a "technical comment" from JCU reef ecologist Andrew Baird, who noted several problems, including the fact that the water flow quoted for Dixson's flume was much faster than any coral larvae have been reported to swim, meaning larvae would be washed out of the back of the flume. *Science*'s review process deemed the comment "as low priority for publication," says Deputy Editor for Research Sacha Vignieri. (*Science*'s news and editorial departments operate independently of each other.) Baird published it as a preprint instead, but it drew little attention.

The Fiji paper was Dixson's second in *Science*—in 2012 she and Hay had reported some corals secrete chemical signals that "recruit" fish to trim toxic seaweeds—and it cemented her status as a rising star. In 2015, she left Georgia Tech to start her own lab in Delaware.

CLARK, a fish physiologist, came to Townsville in 2011 to take a research job at the Australian Institute of Marine Science, a government laboratory on a cape 50 kilometers east of the city. Looking for new things to study, he says he started to read Dixson's and Munday's ocean acidification papers—and was struck by the large effect sizes. "I thought they were some of the most phenomenal findings in the whole discipline of biology," he says. He set out to Lizard Island to repeat the work with predator cues, thinking he could unravel the physiology behind the phenomenon.

But he didn't get the same results at all. Placed in the flume, fish would start to explore their surroundings, but they rarely had the strong preference for one side or the other that Dixson and Munday reported, and amping up the CO_2 did not make a difference. Some fish were "terrified," and didn't move at all, says Fredrik Jutfelt of the Norwegian University of Science and Technology, who joined Clark for a season on Lizard Island in 2014, along with Sundin and several other scientists. "They're taken out of their environment and placed in a highly unnatural situation," Jutfelt says.

Clark was eager to publish the findings, but says, "We could think of quite a few methodological reasons that could be used to downplay our findings." He and Sundin organized a second season on Lizard Island to re-create other studies. The researchers videotaped every experiment and used automated tracking software to monitor fish behavior as a way to rule out bias, which Munday and Dixson had not done.

The work was still in progress when, in June 2016, *Science* published a paper by marine ecologist Oona Lönnstedt, another successful JCU alum. After obtaining her Ph.D. in Townsville—where Munday was one of her three supervisors—Lönnstedt had returned to her native Sweden in 2014 and become a postdoc at Uppsala University (UU) to explore a new environmental threat: the tiny fragments of plastic waste known as microplastics that pollute many aquatic environments. In the 2016 *Science* paper, she and UU limnologist Peter Eklöv reported that perch larvae from the Baltic Sea had a penchant for swallowing microplastics instead of real food, which reduced their growth and—just like elevated CO_2—disturbed their response to chemical cues, making them more likely to end up as fish food themselves.

Sundin, also at UU at the time, says she immediately thought the paper was "a total fantasy." She and Jutfelt had spent time with Lönnstedt at a Baltic island research station, but had never seen a study on the scale described in the paper; many

PHOTOS: (LEFT TO RIGHT) IDA JUTFELT; FREDRIK JUTFELT

other details didn't add up to them either (*Science*, 24 March 2017, p. 1254). They decided to report their suspicions, with support from Clark, Roche, and other researchers they knew from Lizard Island.

The group learned some painful lessons. After a preliminary inquiry, a UU panel dismissed the request for an investigation in a terse report and berated the team for failing to discuss its concerns with Lönnstedt and Eklöv in a "normal scholarly discussion." Lönnstedt said the group was simply jealous. The accusers spent many months gathering additional documentation, at the expense of their own research. In April 2017, Sweden's Central Ethical Review Board concluded there had indeed been "scientific dishonesty" in the research, and *Science* retracted the paper; 8 months later, a full UU investigation concluded the data had been fabricated. (Eklöv blamed Lönnstedt; Lönnstedt maintained her innocence.)

The brazenness of the apparent deception shocked Jutfelt. "It really triggered my skepticism about science massively," he says. "Before that paper, I could not understand how anyone could fabricate data. It was inconceivable to me." Now, he began to wonder how many other papers might be a total fantasy. The experience also taught the group that, if they were ever to blow the whistle again, they would have to bring a stronger case right from the start, Clark says.

On the other side of the globe, the group's accusations had Munday's attention. "It seems that Clark and Jutfelt are trying to make a career out of criticizing other people's work. I can only assume they don't have enough good ideas of their own to fill in their time," he wrote to Lönnstedt in a June 2016 email that she used in her defense to the ethics board. "Recently, I found out they have been 'secretly' doing work on the behavioural effects of high CO_2 on coral reef fishes, presumably because they want to be critical of some aspects of our work."

Pomacentrus moluccensis

Control	39	16	80	5
Control	0	0	90	6
Control	9	2	100	4
Control	85	18	100	7
Control	137	4	70	6
Control	64	12	80	8
Control	163	16	100	6
Control	135	2	100	8
Control	0	0	90	7
Control	6	4	80	5
CO2	0	18	0	0
CO2	0	18	0	0
CO2	63	16	0	3
CO2	45	16	0	1
CO2	59	16	20	0
CO2	0	16	10	1

Dascyllus aruanus

Control	131	16	100	7
Control	33	0	100	5
Control	9	2	90	7
Control	27	18	80	5
Control	74	4	90	6
Control	34	12	90	7
Control	84	16	80	5
Control	95	2	100	8
Control	184	0	80	6
Control	6	4	90	5
CO2	21	14	0	1
CO2	0	15	30	0
CO2	0	16	0	2
CO2	33	14	0	0
CO2	18	14	20	0
CO2	0	18	0	2

Timothy Clark and colleagues say duplications in raw data files—such as these in activity measurements for two fish species in a 2014 paper—suggest fraud. Philip Munday says the block of 10 duplicated numbers is the result of a human error that he will correct.

IN JANUARY 2020, *Nature* published the Clark team's findings: Elevated CO_2 levels in water had a "negligible" effect on fish's attraction to chemical cues from predators, their activity levels, and "lateralization"—their tendency to favor their left or right side in some behaviors. Based on a statistical procedure called a bootstrapping simulation, the team reported that Munday's and Dixson's data on chemical signal preference had a "0 out of 10,000" chance of being real. They left it to the reader to decide what to think about this.

As Clark had predicted, the ensuing debate focused on differences in methods. (It often does when replications fail, says psychologist Brian Nosek of the University of Virginia, a pioneer of the replication movement.) In a rebuttal published in *Nature* in October 2020, Munday, Dixson, and 11 other coauthors argued that the replication effort differed in at least 16 "crucial" points from the original studies, including the species used, the CO_2 measurement methods, and the fact that some of the work was done during a 2016 heat wave on Lizard Island, which could dampen CO_2's effects. "They altered methods in critical ways that reduce the likelihood of detecting effects," Munday said in his December 2020 Zoom seminar.

Clark acknowledges real differences in his team's approach, sometimes necessary because the original papers were vague or the methodology as described didn't work, he says. Other points were easy to counter, he says. The team did use several of the same species as Munday and Dixson, for example, and the heat wave did not affect the temperature in their fish tanks. And, he notes, "Massive effects in many species should not suddenly disappear because of small changes or improvements in the methodology."

Sumpter agrees. The Munday and Dixson defense strategy was "I'll overwhelm them with Christ-knows-how-many different reasons. That doesn't seem to me the way to go about this," he says. But the rebuttal convinced Bruno of the "bro-pocket" tweet. "I am shocked that *Nature* published a putative 'repeatability' study that didn't even come close to mimicking the original science. Lame," he tweeted. "I'd be amazed if Clark et al wasn't retracted."

Others defended Munday and Dixson more diplomatically. Four "grandfathers in the field," as Bruno calls them, criticized the replication in a paper in *Biogeosciences*. One author, Hans-Otto Pörtner of the Alfred Wegener Institute in Bremerhaven, Germany, says his own work had been on the receiving end of criticism from the "youngish group" in the past. "Building a career on judging what other people did is not right," says Pörtner, who co-chairs one of IPCC's three working groups. "If such a controversy gets outside of the community, it's harmful because the whole community loses credibility."

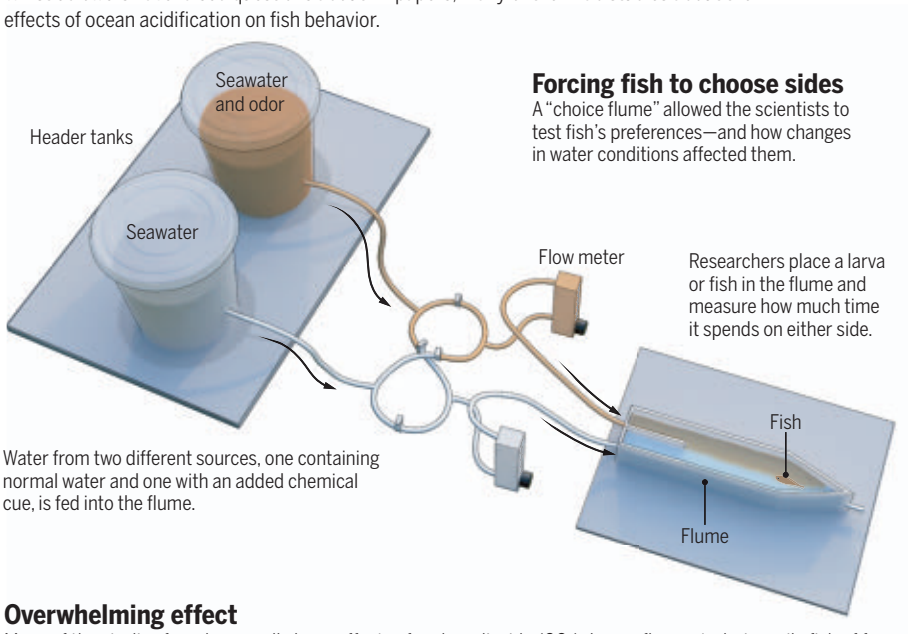
JUTFELT SAYS the team soon had reason to go even further. After the publication in *Nature*, they were approached by several former members of the Munday and Dixson labs eager to discuss their experiences. He and Clark also contacted former students themselves and collected written statements, eight of which are included in the request for the investigation they sent to the three funding agencies in August 2020. (All but one of former students and colleagues were willing to have their identities revealed to the funders, although several told *Science* they worried about retaliation.)

The statements about Munday's lab don't contain concrete fraud accusations, but those about Dixson do. One former colleague became "slowly more suspicious," especially about the fluming studies, and began to monitor them covertly. The testimony contains text messages and photos of lab notebooks that appear to show Dixson did not spend enough time in the lab for one particular study and could not possibly have produced the data she reported.

IMAGE: MUNDAY ET AL., *NATURE CLIMATE CHANGE*, **4**, 487 (2014)

Oceans in the lab

Whistleblowers have raised questions about 22 papers, many of them lab studies about the effects of ocean acidification on fish behavior.

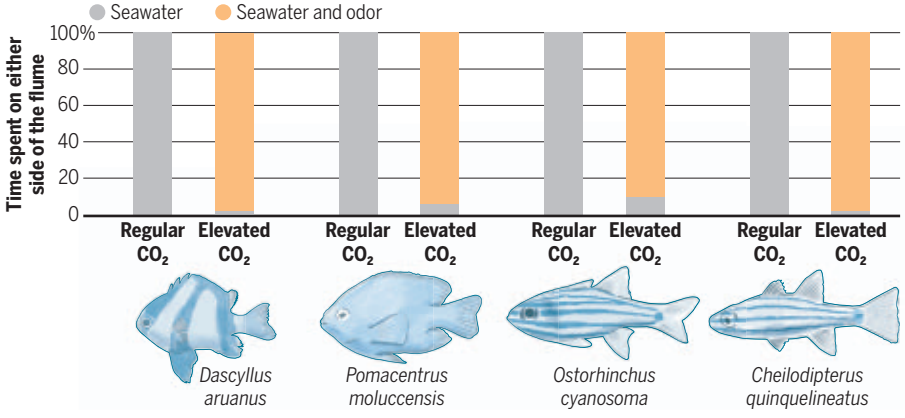


Forcing fish to choose sides

A "choice flume" allowed the scientists to test fish's preferences—and how changes in water conditions affected them.

Overwhelming effect

Many of the studies found unusually large effects of carbon dioxide (CO_2). In one flume study, juvenile fish of four different species were strongly repelled by the odor of a predator at normal CO_2 levels, spending all of their time in odor-free water. At an elevated CO_2 level, all four spent more than 90% of their time in the presence of the predator's odor.



Speaking to *Science*, this researcher said they left UD after asking officials there to investigate Dixson's work. (A UD spokesperson says the university does not comment on personnel matters.) Two other former lab members say they, too, suspected Dixson of making up data.

One of the former lab members also started to examine raw data files for papers authored by Munday and Dixson, looking for irregularities. For most papers, the files were not available—many journals ask scientists to post their raw data, but don't enforce that policy—but in two available files, they found several problems. Dixson's 2014 *Science* paper—on which Munday was not an author—appears to have dozens of duplicated sets of numbers. For example, one column of 20 numbers appears at least six times in the Excel file—suggesting that the same experiment in 20 individuals in each of six species produced exactly the same results every time. In the past, scientists committing fraud are known to have duplicated data as an easy way to make them up.

Chris Hartgerink, a Dutch statistician who has a Ph.D. on detecting fraud in data sets and now heads a company named Liberate Science, confirmed the duplications in a review done at *Science*'s request and says they "would warrant an expression of concern at the least." He would not immediately label them a sign of fraud, however. "It seems like something went deeply wrong in this research project, and that can be due to malfeasance, systematic negligence, or a wide range of other distorting factors," Hartgerink says. Another "data detective" consulted for this story, Nicholas Brown, says the file provides "essentially indisputable evidence" that the data were made up. (Brown, who obtained a Ph.D. from the University of Groningen in the Netherlands in 2019, is retired and lives in Spain.)

Dixson, in the February interview, said she did not know about the allegations. Although she denies making up data, "There hypothetically could be an error in there," she said, perhaps because of mistakes in transcribing the data; "I don't know. I'm human." Hay, the paper's last author, says he'll be "able to discuss what I know and my impressions once investigations are finalized," but did not specify who is investigating. Vignieri, the *Science* editor, says she had not heard about the problems, but that universities, not the journal, would normally do an investigation. She agrees the sample sizes and the reported effects were large, but says "the two often go together and neither can be taken as worrying indicators, by default."

Clark and colleagues also found problems in the data for the 2014 paper in *Nature Climate Change*, which showed fish behavior is altered near natural CO_2 seeps off the coast of Papua New Guinea. (Munday was the first of five authors on the study, Dixson the third.) That data set also contained several blocks of identical measurements, although far fewer than in the *Science* paper. Ecologist Nicholas DiRienzo of the University of Arizona, who was consulted for this story, confirmed the duplications—and found additional ones that he calls "another strong indicator of fabrication."

Munday says Dixson has recently provided him with one original data sheet for the study, which shows she made a mistake transcribing the measurements into the Excel file, explaining the largest set of duplications. "This is a simple human error, not fraud," he says. Many other data points are similar because the methodology could yield only a limited combination of numbers, he says. Munday says he has sent *Nature Climate Change* an author correction but says the mistake does not affect the paper's conclusions.

Brown, who decided to delve more deeply into the case on his own, identified problems of a different nature in two more Munday papers that had not been flagged as suspicious by the Clark team and on which Dixson was not an author. At about 20 places in a very large data file for another 2014 paper in *Nature Climate Change*, the raw data do not add up to total scores that appear a few columns farther to the right. And in a 2016 paper in *Conservation Physiology*, fractions that together should add up to exactly one often do not; instead the sum varies from 0.15 to 1.8.

Munday concedes that both data sets have problems as well, which he says are

CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) MUNDAY *ET AL.*, *NATURE CLIMATE CHANGE*, **4**, 487 (2014)



Many of the disputed studies were done at Lizard Island (far left), a marine biology mecca on the Great Barrier Reef.

PHOTO: FREDRIK JUTFELT

due to their first authors hand copying data into the Excel files. He says the files will be corrected and both journals notified. But Brown says the anomalies strongly suggest fabrication. No sensible scientist would calculate results manually and then enter the raw data and the totals—thousands of numbers in one case—into a spreadsheet, he says.

To him, the problems identified in the data sets also cast suspicions on the "ludicrous effect sizes" in many of the 22 papers flagged by the whistleblowers. "Suppose you're going to the house of somebody you think may have been handling stolen televisions, and you found 22 brand new televisions in his basement, and three had serial numbers that corresponded to ones that have been stolen from shops," Brown says. "Are you going to say, 'Yeah, we'll assume you've got the purchase receipts for the other 19?'" Hartgerink is more cautious. "I did find patterns of recurring issues across the 22 papers, which may prove problematic and warrant further investigation," he wrote in his report for *Science*, "but it is not possible to establish intent, motivation, or origin of these issues."

JCU says it has dismissed the case brought by the Clark group following the advice of an external investigator. The problem in the data on fish living near the CO_2 seeps constituted only a "minor breach" of JCU's Research Code that Munday has offered an explanation for, a spokesperson says.

Clark and his colleagues say they are not surprised. Years ago, neither UU nor JCU responded adequately to the concerns about Lönnstedt's work, they say. "Universities have a big conflict of interest," Sundin says, echoing the experience of many other whistleblowers. "Thankfully, Sweden had a Central Ethical Review Board," Clark says. "Australia has no such board."

The group now worries that if Dixson is found to have fabricated data, Munday will escape responsibility despite being her supervisor during part of her work. That would be reminiscent of the microplastics case, in which Eklöv bore part of the responsibility, according to the second review panel, but was not found guilty of fabrication and kept his job while Lönnstedt lost hers. The outcome would be doubly sad because Lönnstedt and Dixson served as role models in a field where women are underrepresented, says Sandra Binning of the University of Montreal, a co-author of the *Nature* replication: "That's not lost on me."

> **"I stand by the papers that we've published. ... The data was collected with integrity. I mean, I preach that to my students."**
>
> **Danielle Dixson**, University of Delaware, Lewes

SINCE MUNDAY AND DIXSON began to publish on ocean acidification in 2009, their work has spawned a minifield that drew in dozens of other researchers. Many—even members of the Clark group, in earlier work—have reported seeing changes in how fish behave, although rarely as dramatic as those Munday and Dixson claimed. Researchers also tried to unravel the physiological and neural underpinnings of the behavioral changes. In the rebuttal in *Nature*, Munday noted that 85 papers, with more than 180 co-authors from more than 90 institutions, have by now reported effects from elevated CO_2. As Dixson asked in a tweet criticizing the replication effort, "Can one study claim to overturn 11 years of research?"

Jutfelt says the reality is more complex. Of those 85 papers, 43 were co-authored by Munday, he notes. More important, unconscious and conscious bias may have played a role, he says. Many studies that did report an effect, including his own, weren't blinded. Researchers knew which fish were exposed to high CO_2 levels, and they knew what fish were expected to do under those circumstances. Some of his own past studies, including one that found strong effects of CO_2 in sticklebacks in Sweden, were "poorly designed," Jutfelt admits. "I've become much more aware of all of this. I think I have become a better scientist," he says, pointing to the automated analyses in the replication effort.

Sumpter is also not surprised so many published studies went in the same direction. He says his own field, ecotoxicology, is rife with small studies with large effect sizes; they're easy to publish. "But if you claim that chemical X doesn't seem to do very much at relevant concentrations, it's quite likely that a journal editor will just bat it right back to you, saying, without using quite these words: not really exciting."

Still, the reported effects of CO_2 on fish behavior and ecology have seemed to fade as the years passed. A recent meta-analysis of 95 papers by Jeff Clements of Fisheries and Oceans Canada, published as a preprint with Jutfelt, Sundin, and Clark, showed the field is experiencing a strong "decline effect," a phenomenon where, after dramatic initial findings, reported effects become smaller and smaller.

Munday, in his December seminar, acknowledged that later studies had identified "mitigating factors" that dampened the alarming impacts seen in his lab's early experiments. "The effects that we would predict now we have this additional knowledge from a decade of research," he said, "would be much less than what we would have predicted when the very first studies were published." Time may tell whether those effects exist in the first place. ■

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

INSIGHTS



POLICY FORUM

CONSERVATION

Mapping out a future for ungulate migrations

Limited mapping of migrations hampers conservation

By **Matthew J. Kauffman, Francesca Cagnacci, Simon Chamaillé-Jammes, Mark Hebblewhite, J. Grant C. Hopcraft, Jerod A. Merkle, Thomas Mueller, Atle Mysterud, Wibke Peters, Christiane Roettger, Alethea Steingisser, James E. Meacham, Kasahun Abera, Jan Adamczewski, Ellen O. Aikens, Hattie Bartlam-Brooks, Emily Bennitt, Joel Berger, Charlotte Boyd, Steeve D. Côté, Lucie Debeffe, Andrea S. Dekrout, Nandintsetseg Dejid, Emiliano Donadio, Luthando Dziba, William F. Fagan, Claude Fischer, Stefano Focardi, John M. Fryxell, Richard W. S. Fynn, Chris Geremia, Benito A. González, Anne Gunn, Elie Gurarie, Marco Heurich, Jodi Hilty, Mark Hurley, Aran Johnson, Kyle Joly, Petra Kaczensky, Corinne J. Kendall, Pavel Kochkarev, Leonid Kolpaschikov, Rafał Kowalczyk, Frank van Langevelde, Binbin V. Li, Alex L. Lobora, Anne Loison, Tinaapi H. Madiri, David Mallon, Pascal Marchand, Rodrigo A. Medellin, Erling Meisingset, Evelyn Merrill, Arthur D. Middleton, Kevin L. Monteith, Malik Morjan, Thomas A. Morrison, Steffen Mumme, Robin Naidoo, Andres Novaro, Joseph O. Ogutu, Kirk A. Olson, Alfred Oteng-Yeboah, Ramiro J. A. Ovejero, Norman Owen-Smith, Antti Paasivaara, Craig Packer, Danila Panchenko, Luca Pedrotti, Andrew J. Plumptre, Christer M. Rolandsen, Sonia Said, Albert Salemgareyev, Aleksandr Savchenko, Piotr Savchenko, Hall Sawyer, Moses Selebatso, Matthew Skroch, Erling Solberg, Jared A. Stabach, Olav Strand, Michael J. Suitor, Yasuyuki Tachiki, Anne Trainor, Arnold Tshipa, Munir Z. Virani, Carly Vynne, Stephanie Ward, George Wittemyer, Wenjing Xu, Steffen Zuther**

Migration of ungulates (hooved mammals) is a fundamental ecological process that promotes abundant herds, whose effects cascade up and down terrestrial food webs. Migratory ungulates provide the prey base that maintains large carnivore and scavenger populations and underpins terrestrial biodiversity (fig. S1). When ungulates move in large aggregations, their hooves, feces, and urine create conditions that facilitate distinct biotic communities. The migrations of ungulates have sustained humans for thousands of years, forming tight cultural links among Indigenous people and local communities. Yet ungulate migrations are disappearing at an alarming rate (*1*). Efforts by wildlife managers and conservationists are thwarted by a singular challenge: Most ungulate migrations have never been mapped in sufficient detail to guide effective conservation. Without a strategic and collaborative effort, many of the world's great migrations will continue to be truncated, severed, or lost in the coming decades. Fortunately, a combination of animal tracking datasets, historical records, and local and In-

PHOTO: THEO ALLOFS/MINDEN PICTURES



Over 1,300,000 blue wildebeest migrate in a 600-km circuit every year between the Serengeti plains in Tanzania and the Masai Mara in Kenya searching for food and water. Such seasonal movements are becoming more difficult as the human footprint expands.

digenous knowledge can form the basis for a global atlas of migrations, designed to support conservation action and policy at local, national, and international levels.

NEW TECHNOLOGY, NEW DISCOVERIES

New technologies have enabled precise mapping of long-distance migrations and are revealing that the movements of ungulates across the globe are more diverse and behaviorally complex than previously recognized. When animal tracks are overlaid on dynamic maps of seasonal resources, they reveal diverse migration patterns, from long-distance movements across climatic gradients, to shorter elevational movements to access alpine habitats. Recent discoveries span several continents (see the figure). In 2014, a zebra migration was discovered that stretches 500 km across Namibia and Botswana, a new record for the species (*2*). On the Mongolian Steppe, gazelle were found to explore an area roughly the size of Hungary (100,000 km^2) over their lifetimes (*3*). In Ethiopia's Gambella National Park, researchers discovered that white-eared kob migrate in an 860-km circuit connecting to the Boma-Bandingilo migration in South Sudan, extending the species' known migratory range (*4*).

Detailed movement data are also leading to new ecological discoveries. A key finding suggests that in some species, migratory behavior is a type of animal culture that must be learned and transmitted between generations. In North America, bighorn sheep and moose failed to migrate when first translocated into new landscapes. Over multiple generations, however, individual populations gained knowledge to move seasonally and find forage at broader scales, and they became more migratory (*5*). This reliance on culture carries a stark warning for conservation—namely, that the persistence of a migration corridor may depend on the survival of individuals that possess the knowledge to travel along it.

MIGRATIONS IN DANGER

Historical accounts describe numerous migrations that have been lost. The millions of Cape springbok that once traversed the Karoo landscape of South Africa were eradicated by fencing, disease, and hunting at the end of the 19th century. In Kenya's Kajiado County, migrations of wildebeest, zebra, and Thomson's gazelle have collapsed owing to competition with livestock and massive land conversion. When the tens of millions of bison that roamed North America were slaughtered, the abundance necessary to promote migration was also lost. The Yellowstone population is one of just a few bison herds that still migrate.

Barriers have long restricted the free movement of migratory herds. Today, nomadic movements of Mongolian gazelles and khulan are constrained by railroads and border fences (*3*). In Botswana, veterinary fences built in the 1950s caused the death of hundreds of thousands of wildebeest. In Russia's Kola Peninsula, the construction of a railroad divided the wild reindeer population and eliminated the longest of the region's migrations. In Europe, red deer must now navigate a landscape shaped by millennia of human infrastructure and habitat fragmentation, exploiting seasonal cycles of forage only where their routes are not disrupted (see the figure). Recent estimates project that 25 million km of new roads will be built across the globe by 2050 (*6*), which will constrain and sever seasonal migrations even further.

Climate change poses an additional threat. Many ungulates time their migrations to exploit patterns of plant green-up and other key weather events. Droughts are becoming commonplace, making it more difficult for animals to move in synchrony with green-up and access the best forage (*7*). In southern Africa, zebra, blue wildebeest, and African elephant migrations are driven by water availability, which is changing as rainfall patterns shift, leading to population declines of some species (*8*). In the Arctic, barren-ground caribou (fig. S2) have shifted migration and calving dates by up to 0.5 days per year over the last three decades in response to climate change (*9*). Preliminary tracking data indicate that wild reindeer in the Taimyr-Evenk population of Russia are showing similar effects in response to Arctic warming of 4°C over the past 40 years. An extreme example has been observed in endangered Dolphin and Union caribou in the Canadian Arctic Archipelago, which experience mass mortality when they break through thin ice caused by delayed fall freeze-up (*10*). Migratory ungulates are thus doubly challenged by changes in climate that alter snow, forage, and water distribution, combined with barriers that prevent them from adapting their movement tactics as conditions change.

Author affiliations are listed in the supplementary materials.
Email: mkauffm1@uwyo.edu

EXISTING CONSERVATION AND POLICY

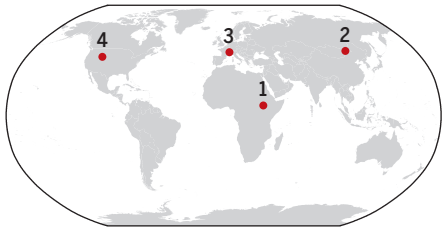
The Convention on Migratory Species (CMS) advanced conservation by formally recognizing animal migrations in 1979. Yet, tracking data were not available in sufficient detail to map ungulate migrations until recently, preventing Parties to the Convention from developing concrete policies to protect migrations. New technologies and methods now allow corridors to be mapped in detail from tracking data (*11*). When migration paths are overlaid on landscapes within a geographic information system, barriers and other threats can be identified, pointing towards effective conservation solutions.

Globally, policies to protect ungulate migrations typically require migration maps or animal tracking data. In 2008, the Path of the Pronghorn in Wyoming (USA) became the first federally protected corridor, when a map based on telemetered animals was included in the Bridger-Teton National Forest Plan, which required that forest managers consider impacts to the corridor (see the figure). In 2018, the Tanzania Wildlife Conservation Regulations were signed, establishing a means to designate wildlife corridors between the country's protected areas. In 2011, Angola, Botswana, Namibia, Zambia, and Zimbabwe signed an international treaty creating the Kavango-Zambezi Transfrontier Conservation Area (KAZA). The KAZA aims to facilitate a network of protected areas linked by dispersal corridors and migratory routes delineated from tracking data of species like blue wildebeest, African buffalo, zebra, and elephant. In Kazakhstan, the Yrgyz-Torgai-Zhylanshyk "ecological corridor"—the first of its kind for the country—was created to allow migration of the Betpakdala saiga population between protected areas. Ungulate tracking data have also facilitated the expansion of protected areas. In Mongolia, the movements of khulan beyond the border of Great Gobi B Strictly Protected Area during the extreme winter of 2009–2010 prompted a doubling of the protected area in 2019.

For classic back-and-forth migrations, corridors delineated from empirical tracking data are capable of identifying even more targeted conservation solutions. In Wyoming (USA), for example, GPS mapping of a high-use mule deer corridor (*11*) allowed researchers to identify a 400-m bottleneck used by ~5000 deer (see the figure). When this land was threatened with subdivision, The Conservation Fund raised funding to purchase and protect it as a Wildlife Habitat Management Area. Enumerating the threats to this corridor also prompted the State of Wyoming to "designate" migration corridors and manage them for "no net loss of function." In the first quarter of 2019, the US ad-

Ungulate migrations around the world

Animal tracking studies are being conducted around the world, facilitating discovery of previously unknown movements and making it possible to map migrations and identify threats with precision.

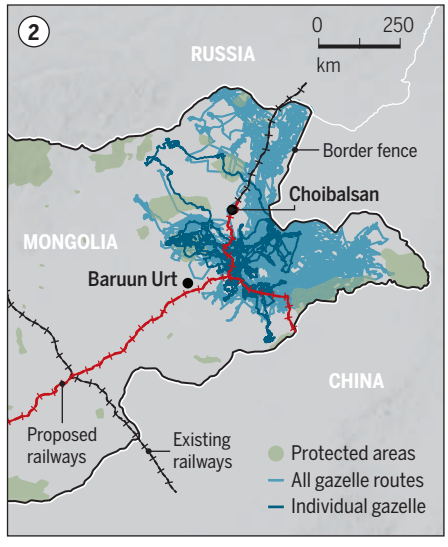


White-eared kob



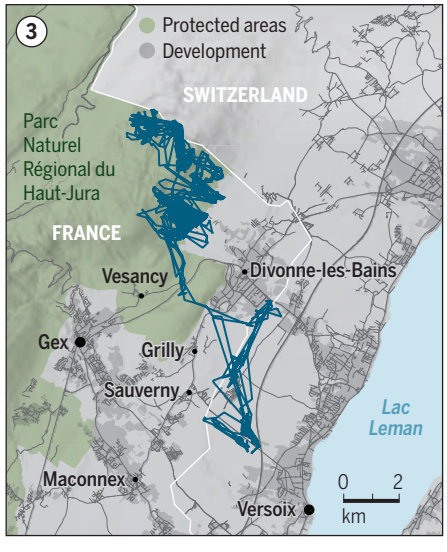
In 2018, white-eared kob were discovered to make an 860-km migratory circuit between Ethiopia and South Sudan, extending the known migratory range. The herds traverse working landscapes consisting of oil and gas concessions, hotspots of armed conflict, and commercial agricultural developments in the Boma-Gambella landscape.

Mongolian gazelle



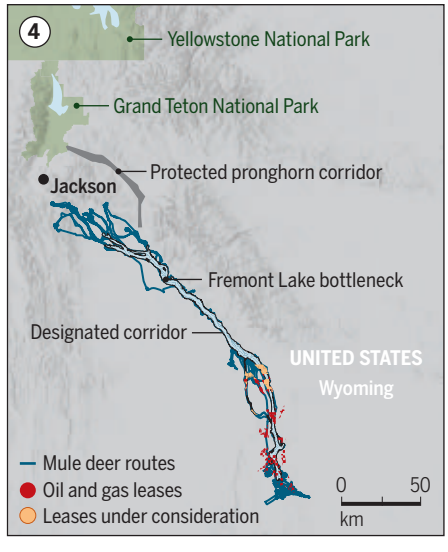
On the Eastern Steppe, the nomadic movements of Mongolian gazelle are an order of magnitude larger than the region's protected areas despite being bound by impermeable border fences with Russia and China. A proposed railroad threatens to further constrain their wide-ranging movements.

Red deer



Ungulates living in human-dominated landscapes may try to maintain migration, but their movements are often constrained and truncated by human settlements and associated infrastructure. Red deer at the Swiss-French border navigate a semi-urban landscape using remaining forest patches interrupted by highways and fences.

Mule deer



In 2014, a 242-km mule deer migration was discovered in Wyoming (U.S.) and analyzed to delineate the high-use corridor. The corridor was formally "designated" by state officials in 2016, prompting protection of the Fremont Lake bottleneck and deferral of oil and gas leases that would have put the migration at risk.

GRAPHIC: K. FRANKLIN/*SCIENCE*

ministration deferred 5674 ha of oil and gas leases on federal land because drilling was deemed too risky to the functioning of the corridor.

Lost migrations also need to be inventoried and mapped by leveraging Indigenous, local, and expert knowledge. Although there are only a few examples where an ungulate migration has been restored, mapping historical migrations provides an important baseline to evaluate changes and potential future conservation opportunities. In some South American rangelands, for example, conservationists are hopeful that ranch abandonment will create opportunities to restore lost migrations of guanacos and globally threatened huemul deer.

BETTER POLICIES THROUGH BETTER MAPS

The next decade is likely to witness a massive global effort to combat the twin crises of biodiversity loss and climate change, and maps of the world's ungulate migrations can inform these efforts. We therefore propose a global atlas of ungulate migration, built from tracking studies and traditional knowledge. We envision a digital archive that translates migration data into actionable migration maps that are standardized, in a central database, and publicly available. Together with new guidelines for conserving ecological connectivity (*12*), this global atlas will provide the migration maps that planners and policy-makers need to develop corridor and connectivity plans at national and regional scales, review impacts of potential developments, and consider new policies to conserve migrations.

Most of the world's governments have signed on to the Convention on Biological Diversity (CBD; 193 parties) and the CMS (132 parties). To adhere to these conventions, governments need reliable information on where biodiversity and ecological processes like migrations occur. Under the CBD, the Conference of the Parties will meet in 2021 to establish the Post-2020 Global Biodiversity Framework, which many entities hope will target at least 30% of Earth's surface for some form of protection by 2030. Signatories to the CBD are encouraged to update National Biodiversity Strategy and Action Plans, nearly 90% of which lack plans for maintaining connectivity or conserving migrations (*13*). A global migration atlas can help define the areas to be targeted by these new efforts.

Internationally funded development projects are accelerating across Asia and Africa (*6*), driven in part by China's Belt and Road Initiative. Long-distance migrations are highly susceptible to cumulative impacts, such as those that span hundreds of kilometers or cross international boundaries. Although international financing institutions are developing new policies that aim to reduce impacts to wildlife connectivity, adequate migration maps are lacking. A detailed atlas would allow these institutions to fund smart and proactive actions, such as project relocation or redesign, to reduce impacts to existing migrations across international boundaries.

Data on protected areas, Key Biodiversity Areas, and the distribution of threatened species are provided to the public and investors through the Integrated Biodiversity Assessment Tool (www.ibat-alliance.org). These data allow projects to identify environmental concerns early, avoid impacts to biodiversity, and reduce financial risks. Key Biodiversity Areas can be used to identify migratory stopover or bottleneck sites, but standardized and publicly available maps of migration pathways are also required to inform policies to maintain connectivity.

Mapping the world's ungulate migrations will also make clear that numerous stakeholders work the lands that underpin these migrations or benefit from the herds directly. The cultural traditions and identity of the Inuit and Tlicho, for example, are based on livelihoods that depend on migratory caribou (*14*). Whether for mule deer in Wyoming, caribou in the Arctic, or gazelle across the Mongolian steppe, migration mapping highlights that such journeys traverse habitats that have been stewarded for generations by local communities and private landowners. Expansive migrations across working landscapes thus demonstrate the need for sustainable conservation grounded in a social-ecological systems approach that considers the costs and benefits to people whose lands are essential for sustaining connectivity (*15*).

Developing a global atlas of ungulate migrations will require unprecedented collaboration to assemble the required knowledge, data, and analytical tools. Fortunately, tracking datasets are growing each year, and existing data can be used to map many known migrations. Unmapped migrations can be targeted for new field studies, and historical knowledge can be digitally archived. To coordinate this effort, we have created the Global Initiative for Ungulate Migration (GIUM). Under the auspices of CMS, and in partnership with the Convention's secretariat, the GIUM brings together scientists, conservationists, and wildlife managers worldwide to create a collaborative knowledge base, develop a global atlas, and catalyze new conservation actions and policies.

Many efforts, from regional regulations to national legislation and international treaties, seek to sustain the world's biodiversity. A new global atlas on ungulate migration will help target action under all existing frameworks. First, migration maps will be derived from empirical data, using established analytical methods, and the resulting maps will be peer-reviewed and centrally curated at the CMS (https://www.cms.int/gium). The recent collaborative effort by the US Geological Survey and state wildlife agencies to map ungulate migrations of the western US is a timely example of this approach (see supplementary materials). Second, empirical migration maps depict the actual movements animals make on an annual basis. Thus, the footprint of migration maps can be readily integrated into existing spatial conservation planning. For example, the recent International Union for Conservation of Nature Connectivity Guidelines (*12*) establish criteria for "ecological corridors," some of which could be delineated by mapped migration corridors. Finally, the GIUM itself is composed of both scientists and conservationists working together to compile disparate migration data in one place and explore the best means to incorporate migration maps into existing conservation frameworks.

Like much of the world's biodiversity, migrations need to be mapped and archived before they are lost. Such an ambitious effort could guide the locations of new roads, fences, and other infrastructure and identify the habitats that should remain undeveloped to safeguard the seasonal movement of ungulates, the ecosystems they sustain, and the diverse benefits they provide to humanity. ■

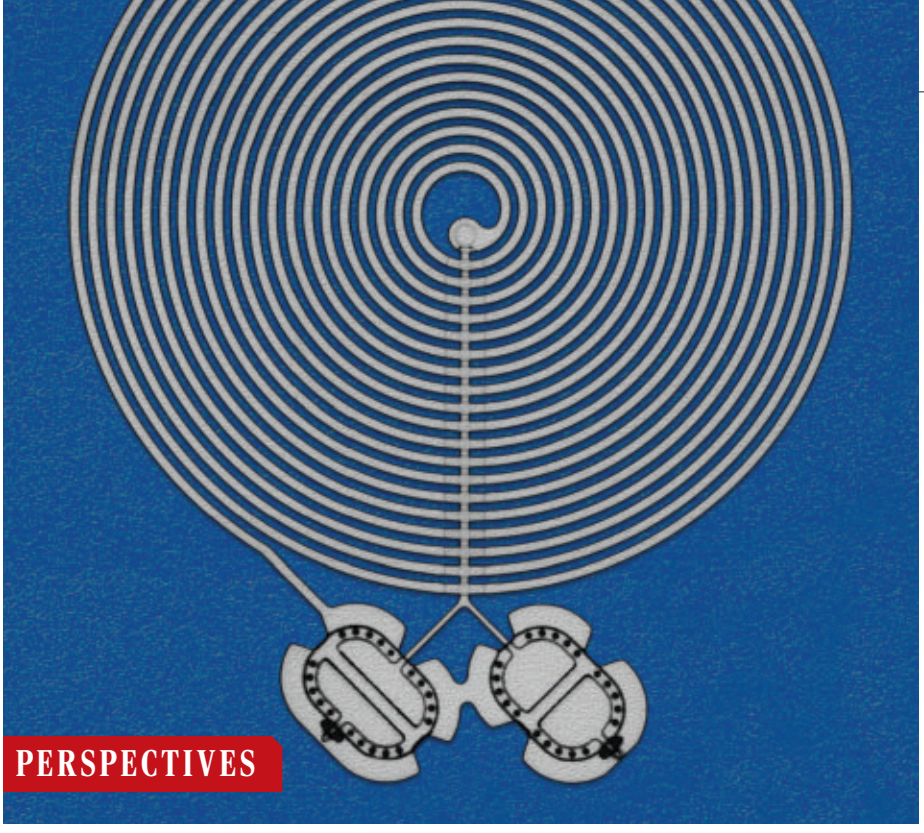
REFERENCES AND NOTES

1. G. Harris, S. Thirgood, J. G. C. Hopcraft, J. P. Cromsight, J. Berger, *Endanger. Species Res.* **7**, 55 (2009).
2. R. Naidoo *et al.*, *Oryx* **50**, 138 (2016).
3. D. Nandintsetseg *et al.*, *J. Appl. Ecol.* **56**, 1916 (2019).
4. A. Kassahun, Movement Ecology and Home Ranges of White eared Kob (*Kobus kob leucotis*) in the Boma-Gambella Transboundary Landscape Between the Republic of South Sudan and Ethiopia, thesis, Addis Ababa University, Ethiopia (2019).
5. B. R. Jesmer *et al.*, *Science* **361**, 1023 (2018).
6. W. F. Laurance *et al.*, *Nature* **513**, 229 (2014).
7. E. O. Aikens *et al.*, *Glob. Change Biol.* **26**, 4215 (2020).
8. N. Owen-Smith, J. O. Ogutu, in *Wildlife Conservation in a Changing Climate*, J. F. Brodie, E. S. Post, D. F. Doak, Eds. (Univ. of Chicago Press, 2012), pp. 153–178.
9. S. C. Davidson *et al.*, *Science* **370**, 712 (2020).
10. K. G. Poole, A. Gunn, B. R. Patterson, M. Dumond, *Arctic* **63**, 414 (2010).
11. M. J. Kauffman *et al.*, *Wild Migrations: Atlas of Wyoming's Ungulates* (Oregon State Univ. Press, 2018).
12. J. Hilty *et al.*, *Guidelines for Conserving Connectivity Through Ecological Networks and Corridors*, Best Practice Protected Area Guidelines Series No. 30 (International Union for Conservation of Nature, Gland, Switzerland, 2020).
13. E. Dinerstein *et al.*, *BioScience* **67**, 534 (2017).
14. A. Cunsolo *et al.*, *Am. Imago* **77**, 31 (2020).
15. C. Kremen, A. M. Merenlender, *Science* **362**, eaau6020 (2018).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/566/suppl/DC1

10.1126/science.abf0998



PERSPECTIVES

A spiral inductor and two capacitors (each with a vibrating membrane) create a microwave cavity.

QUANTUM SYSTEMS

Macroscale entanglement and measurement

The entangled motion of macroscopic vibrating membranes can be measured precisely

By **Hoi-Kwan Lau**[1] *and* **Aashish A. Clerk**[2]

Quantum mechanics governs both fundamental particles and large objects, but for the latter, a myriad of different factors conspire to mask its effects and render deviations from a purely classical description all but invisible. However, careful and controlled experiments can reveal purely quantum mechanical phenomena of large objects. On pages 622 and 625 of this issue, Kotler *et al.* (*1*) and Mercier de Lépinay *et al.* (*2*), respectively, report experimental examples of the direct observation of two effects of quantum mechanics in macroscopic objects that cannot be seen classically. Kotler *et al.* report direct evidence of quantum entanglement of macroscopic objects (vibrating drumhead membranes), and Mercier de Lépinay *et al.* detect and even evade back-action in quantum-mechanical measurements of an analogous system.

[1]Department of Physics, Simon Fraser University, Burnaby, BC, Canada. [2]Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL, USA. Email: aaclerk@uchicago.edu

A number of physical platforms have been developed to exploit quantum effects in systems much larger than a single atom or molecule. Among the largest are quantum optomechanical systems (*3*), in which the motion of a macroscopic mechanical resonator (such as a vibrating cantilever or membrane) is controlled with photons in an electromagnetic resonator. Both Kotler *et al.* and Mercier de Lépinay *et al.* used a superconducting aluminum resonant inductor–capacitor circuit with a microwave resonant frequency (a few gigahertz), and a thin suspended aluminum membrane with an extremely low-dissipation, drumhead-like vibrational mode in the megahertz range (see the image). These membranes have diameters of ~10 μm and masses of ~100 pg (~10^{12} atoms).

The membrane forms one plate of the capacitor in the circuit, so its motion directly modulates the resonant frequency of the circuit (see the figure, top). The number of microwave photons in the circuit acts as a force on the mechanical system. By driving the circuit with appropriate microwave voltages, the mechanical motion could be prepared, manipulated, and read out with quantum-level precision. Crucially, in each experiment, the circuit is coupled to two distinct mechanical drumhead resonators (realized as two vibrating capacitors) so that quantum correlations between different systems could be explored.

Kotler *et al.* drove the circuit with tailored microwave pulses that strongly correlate the motion of the two vibrating drumheads at the quantum level, creating a quantum-entangled state of two macroscopic objects. Entanglement is a strictly quantum effect in which distinct objects exhibit correlations that, in a certain sense, are stronger than what is allowed classically (see the figure, top). In various nascent quantum technologies, such as computation and sensing, entanglement enables quantum advantages over purely classical devices.

To verify entanglement, the authors used the circuit to measure correlations in the motion of the two vibrating drumheads. Although strong evidence of macroscopic entanglement has been previously reported in related systems (*4*, *5*), the results of Kotler *et al.* represent an advance on several fronts. The entanglement was generated in a deterministic (nonrandom) manner, and all of the correlations needed to verify entanglement were directly measured rather than inferred. Further, the measurement efficiency was substantially higher than in previous microwave-frequency studies, so that entanglement was apparent even without the subtraction of amplifier noise from the measured data.

Another purely quantum phenomena is back-action, the concept that measurement of an object must disturb it in some way. According to Heisenberg's uncertainty principle, precise observation of an object's position necessarily imparts a random "kick" that disturbs its momentum. For an oscillator whose position is being continuously monitored, measurement precision is often expressed in terms of the standard quantum limit (SQL) (*6*). However, back-action need not limit force measurement, and a variety of "quantum back-action–evasion" strategies have been developed.

A particularly powerful one is the construction of so-called quantum mechanics–free subsystem (QMFS) (*7*, *8*). A single oscillator is effectively encoded into two distinct oscillators, so that the measurement back-action can be moved into collective degrees of freedom that are dynamically uncoupled from the quantities being measured (see the figure, bottom). This general strategy allows the measurement of a narrow-band (classical) force without any fundamental quantum limit.

PHOTO: KOTLER *ET AL.* (*1*)

The experiment by Mercier de Lépinay *et al.* explicitly demonstrated this idea using an optomechanical scheme first analyzed in (*9*). The unusual QMFS dynamics of the two mechanical membranes is realized by driving the circuit with four distinct microwave control tones, each at a different slightly detuned frequency. The net result is an effective single harmonic oscillator encoded in collective variables of the two physical vibrating membranes that can be measured without any degradation from quantum back-action. Indeed, they achieved a measurement precision better than the SQL. Their technique allows for a complete measurement of the effective mechanical motion without a back-action limit in both the sine and cosine quadratures of the motion. Previous work protected only one quadrature of the mechanical motion from back-action (*10*–*12*), could not measure both protected collective quadratures (*13*), or coupled a mechanical resonator to an atomic ensemble (*14*).

Mercier de Lépinay *et al.* also used their technique to generate entanglement. By slightly breaking the conditions needed for a perfect QMFS measurement, the circuit realizes an effective autonomous measurement-plus-feedback operation that correlates and ultimately entangles the motions of the two vibrating drumheads. This approach both generated and stabilized an entangled state as long as the circuit is energized. This method is complementary to that of Kotler *et al.*, which does not stabilize an entangled state but has the advantage of being a true entangling quantum gate that preserves information in the initial mechanical state.

Beyond demonstrating direct evidence of quantum entanglement and measurement beyond the conventional limits imposed by the quantum back-action for macroscopic objects, the advanced techniques developed by both groups could have a broader impact. The pulsed unitary entangling operation by Kotler *et al.* could be used as logic gates for quantum computation with continuous variables, and the collective measurement process by Mercier de Lépinay *et al.* could be combined with entanglement to enable new kinds of enhanced measurements. The refined microwave optomechanical devices of both groups could be used to faithfully convert quantum information between different physical platforms (*15*). Apart from practical applications, these experiments address how far into the macroscopic realm experiments can push the observation of distinctly quantum phenomena. ■

REFERENCES AND NOTES

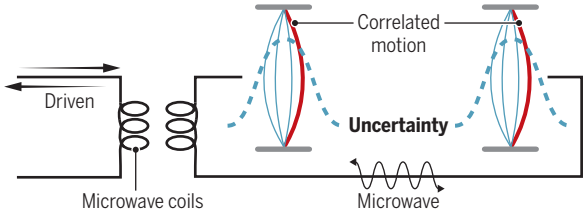
1. S. Kotler *et al.*, *Science* **372**, 622 (2021).
2. L. Mercier de Lépinay, C. F. Ockeloen-Korppi, M. J. Woolley, M. A. Sillanpää, *Science* **372**, 625 (2021).
3. M. Aspelmeyer, T. J. Kippenberg, F. Marquardt, *Rev. Mod. Phys.* **86**, 1391 (2014).
4. C. F. Ockeloen-Korppi *et al.*, *Nature* **556**, 478 (2018).
5. R. Riedinger *et al.*, *Nature* **556**, 473 (2018).
6. V. B. Braginsky, F. Ya Khalili, K. S. Thorne, *Quantum Measurement* (Cambridge Univ. Press, 1992).
7. M. Tsang, C. M. Caves, *Phys. Rev. X* **2**, 031016 (2012).
8. E. S. Polzik, K. Hammerer, *Ann. Phys.* **527**, A15 (2015).
9. M. J. Woolley, A. A. Clerk, *Phys. Rev. A* **87**, 063846 (2013).
10. J. Suh *et al.*, *Science* **344**, 1262 (2014).
11. F. Lecocq, J. B. Clark, R. W. Simmonds, J. Aumentado, J. D. Teufel, *Phys. Rev. X* **5**, 041037 (2015).
12. I. Shomroni, L. Qiu, D. Malz, A. Nunnenkamp, T. J. Kippenberg, *Nat. Commun.* **10**, 2086 (2019).
13. C. F. Ockeloen-Korppi *et al.*, *Phys. Rev. Lett.* **117**, 140401 (2016).
14. C. B. Møller *et al.*, *Nature* **547**, 191 (2017).
15. A. A. Clerk, K. W. Lehnert, P. Bertet, J. R. Petta, Y. Nakamura, *Nat. Phys.* **16**, 257 (2020).

10.1126/science.abh3419

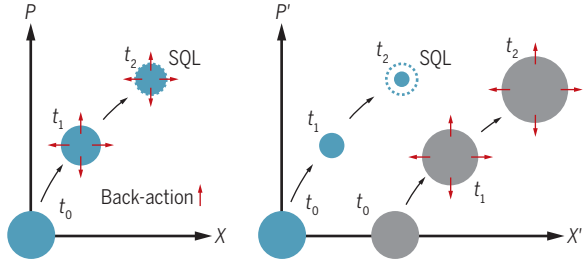
Quantum effects writ large

Two studies use microwave-driven membranes to demonstrate quantum-mechanical effects of macroscopic objects.

Entangled positions
Kotler *et al.* showed that unlike the position *X* and momentum *P* of unentangled oscillators (with fundamental uncertainty shown in blue), entangled oscillators had strongly correlated motion (red).



Beating back-action limits
(Right) Uncertainty in the oscillator's motion (shaded area) is reduced by measurement over times *t* but could not beat the standard quantum limit (SQL, dashed circle). (Left) Mercier de Lépinay *et al.* moved quantum back-action into a second, unmeasured effective oscillator whose uncertainty grows (gray shaded area) so that the measurement precision of both quadratures *X′* and *P′* (blue) beat the SQL.



GRAPHIC: C. BICKEL/*SCIENCE*

CORONAVIRUS

Rapid antigen testing in COVID-19 responses

SARS-CoV-2 transmission was reduced with measures centered on rapid antigen testing

By **Marta García-Fiñana** *and* **Iain E. Buchan**

The value of rapid antigen testing of people (with or without COVID-19 symptoms) to reduce transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been discussed extensively (*1*–*5*) but remains a topic of policy debates (*6*, *7*). Lateral flow devices (LFDs) to test for SARS-CoV-2 antigen are inexpensive, provide results in minutes, and are highly specific (*2*–*4*), and although less sensitive than reverse transcriptase polymerase chain reaction (RT-PCR) tests to detect viral RNA, they detect most cases with high viral load (*2*, *3*, *8*), which are likely the most infectious (*8*, *9*). Successful mass testing relies on public trust, the social and organizational factors that support uptake, contact tracing, and adherence to quarantine. On page 635 of this issue, Pavelka *et al.* (*10*) report the substantial reduction in transmission that population-wide rapid antigen testing had, in combination with other measures, in Slovakia.

Slovakia ran mass testing interventions from the last week of October to the second week of November 2020, with 65% of the target populations taking rapid antigen tests. Testing started in the four counties with the highest rates of infection, continued with national mass testing, then was followed up with more testing in high-prevalence areas. Nasopharyngeal swabs for the LFDs were taken by clinical staff, not self-administered. Sample quality and test accuracy are higher with tests taken by health professionals (*3*). Although the specific impact of Slovakia's mass testing could not be disentangled from

Institute of Population Health, University of Liverpool, Waterhouse Building, Liverpool L69 3BX, UK.
Email: buchan@liverpool.ac.uk; martaf@liverpool.ac.uk

the contribution of other concurrent control measures (including closure of secondary schools and restrictions on hospitality and indoor leisure activities), statistical modelling by Pavelka *et al.* estimated a 70% reduction in the prevalence of COVID-19 cases compared with unmitigated growth.

The UK piloted mass testing in Liverpool in November 2020 after the city experienced the highest COVID-19 prevalence in the country. Slovakia applied more pressure on its citizens to get tested than did Liverpool, by requiring anyone not participating in mass testing to quarantine. The Liverpool testing uptake was consequently lower than Slovakia's, involving 25% of the population in 4 weeks. Liverpool's public health service valued the testing as an additional control measure, but impacts were limited by lack of support for those in socioeconomically deprived areas facing income loss from quarantine after a positive test (*2*): Test positivity rates were highest and testing uptake lowest in the most deprived areas (*2*, *11*). Similar socioeconomic barriers were reported for test uptake among care home staff (*12*). This highlights the importance of addressing public perceptions of testing and support for low-income workers to quarantine when implementing mass testing.

The predictive value of testing varies with the population prevalence of infection and phase of the epidemic curve (*7*). As the prevalence of SARS-CoV-2 infections decreases, the proportion of false-positive test results increases, whereas the number of false-negative test results decreases. For example, with 99.9% specificity (proportion of noninfections that the test rejects) and 80% sensitivity (proportion of infections that the test detects), the positive predictive value (proportion of people with a positive test result who are infected) is 89% when the prevalence is 1%, and it drops to 44% at 0.1% prevalence (55 in 100 positive test results are false). In absolute terms, however, if testing 100,000 people, these scenarios would result in 99 false positives (out of 899 positive results) and 100 false positives (out of 180 positive results) for 1% and 0.1% prevalence, respectively (see the figure). Confirmatory RT-PCR tests after a positive LFD test result was recently reintroduced by Public Health England because of both the low positive predictive values of testing at low prevalence of infection and the utility of reusing PCR samples for viral genetic sequencing in variant surveillance (*13*).

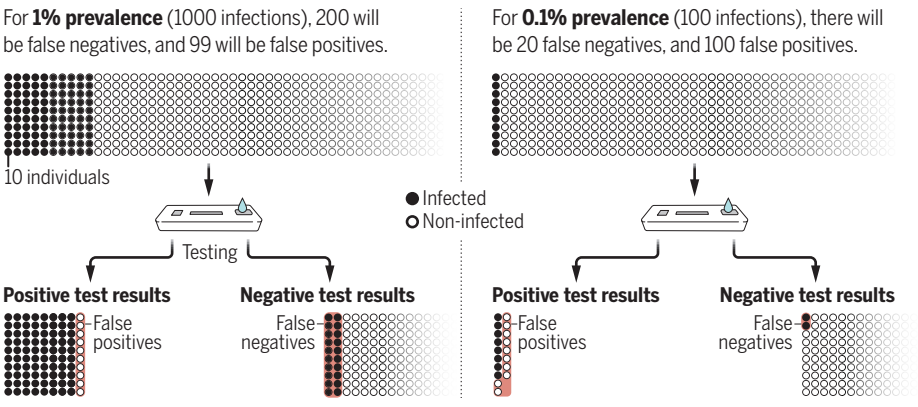
The pilot in Slovakia was conducted while the prevalence was still high (3.9% in areas with the highest rate of infection). Rapid antigen testing was used as an additional tool to identify a substantial proportion of asymptomatic SARS-CoV-2–infected individuals, who were required to quarantine. Additionally, those who did not agree to take part in testing were required to quarantine, thus reducing the chance of transmission among those who were permitted to mix. At higher prevalence, more SARS-CoV-2 infections can be identified, but the proportion of false-negative tests is also higher, so the reliance on other control measures is greater. No matter what the prevalence, mass testing regimes can only properly be considered amid other health protection measures.

By the end of the mass testing program in Slovakia, rapid antigen tests had identified more than 50,000 people without COVID-19 symptoms who were likely contagious with SARS-CoV-2. UK mass testing pilots in Liverpool and also in Wales that started at a similar time as the pilot in Slovakia, but with fewer pressures to take part, identified more than 4000 asymptomatic cases in the Cheshire and Merseyside region around Liverpool (*14*) and more than 700 in Wales (*15*). Although the testing technology was equivalent across Slovakia, England, and Wales, the interventions were different, spanning a variety of population prevalence, phases of the epidemic curve, surges of new variants, periods of lockdown, periods of reopening of large-scale social mixing, and targeting of testing. For example, the Liverpool project shifted in public messaging from "Let's All Get Tested" to "Test Before You Go" to "Testing Our Front Line" (for anyone having to leave home to go to work in lockdown).

In places with low SARS-CoV-2 prevalence, mindful of the cumulative harms from COVID-19 restrictions, the emphasis is on restarting social and economic activities while minimizing infections. As research continues to clarify the impact of vaccines on SARS-CoV-2 transmission, there is a need to use rapid antigen testing as a part of comprehensive public health measures that reduce the risk of the virus escaping vaccine or natural immunity through avoidable transmission—for example, testing to secure workplaces and large events as societies reopen after lockdowns. Successful implementation, however, depends on public participation in testing and adequate support to quarantine. ■

Predictive value of testing changes with prevalence

When testing 100,000 individuals with a lateral flow device with 80% sensitivity and 99.9% specificity, the proportion of false-positive and false-negative test results will vary according to the prevalence of infection.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. Z. Kmietowicz, *BMJ* **372**, n81 (2021). 10.1136/bmj.n81
2. I. Buchan *et al.*, Liverpool COVID-19 community testing pilot. Interim evaluation report. 2020 (University of Liverpool, 2020); www.gov.uk/government/publications/liverpool-covid-19-community-testing-pilot-interim-evaluation-report-summary.
3. T. Peto *et al.*, *medRxiv* 10.1101/2021.01.13.21249563 (2021).
4. A. Crozier, S. Rajan, I. Buchan, M. McKee, *BMJ* **372**, 208 (2021).
5. M. J. Mina, T. E. Peto, M. García-Fiñana, M. G. Semple, I. E. Buchan, *Lancet* **397**, 1425 (2021).
6. L. Y. W. Lee *et al.*, *medRxiv* 10.1101/2021.03.31.21254687 (2021).
7. R. W. Peeling, P. Olliaro, *Lancet* 10.1016/S1473-3099(21)00152-3 (2021).
8. L. Y. W. Lee *et al.*, *medRxiv* 10.1101/2021.03.31.21254687 (2021).
9. M. Marks *et al.*, *Lancet Infect. Dis.* (2021). 10.1016/S1473-3099(20)30985-3
10. M. Pavelka *et al.*, *Science* 372, 635 (2021).
11. M. A. Green *et al.*, *medRxiv* 10.1101/2021.02.10.21251256 (2021).
12. J. Tulloch *et al.*, *SSRN* 10.2139/ssrn.3822257 (2021).
13. S. Hopkins, Gov.UK 30 March 2021); https://publichealthmatters.blog.gov.uk/2021/03/30/covid-19-reintroducing-confirmatory-pcr-testing.
14. NHS Cheshire and Merseyside, Combined Intelligence for Population Health Action (2021): www.cipha.nhs.uk.
15. K. Nnoaham, Evaluation of the lateral flow device testing pilot for COVID-19 in Merthyr Tydfil and the lower Cynon Valley (2021); https://cwmtafmorgannwg.wales/Docs/Publications/FINAL_V2_Whole%20Area%20Testing%20Evaluation%20Full%20Report%2020210325.pdf.

ACKNOWLEDGMENTS

I.E.B. and M.G.-F. received grant funding from the UK Department of Health and Social Care to evaluate the Liverpool community testing pilot. I.E.B. reports fees from AstraZeneca as chief data scientist adviser through Liverpool University and a senior investigator grant from the National Institute for Health Research (NIHR) outside the submitted work.

10.1126/science.abi6680

TEXTILES

Reversible fusion-fission fibers

Reversible assembly of graphene oxide fibers creates a pathway for practical applications

By **Rodolfo Cruz-Silva**[1] *and* **Ana Laura Elías**[2]

Materials scientists have long used highly coveted biomimetic techniques in the search for synthetic structural materials that resemble muscles and other natural fibers (*1*). Among these techniques, self-assembly processes inspired by cell fusion are gaining notoriety (*2*). Within biological systems, fission occurs when one entity can separate into two or more parts, and the opposite process, fusion, occurs when two or more parts merge into one object (*3*). These processes can be naturally triggered by stimuli present in the environment—such as light, temperature, or humidity—but are ultimately controlled by the organism's own metabolism. On page 614 of this issue, Chang *et al.* (*4*) demonstrate the assembly of wet-spun graphene oxide (GO) microfibers through a reversible solvent-triggered process, which is regulated by the individual fiber's chemistry and morphology, mimicking biological fusion.

The fusion process of Chang *et al.* results in hierarchical assemblies that involve thousands of individual GO fibers (see the figure). When the fiber assemblies are immersed in a suitable solvent, the process can be reversed, resembling fission. Notably, the transition between the individual fibers and the complex yarn can be repeated without damaging the primary fiber structure, ultimately preserving the GO flakes and their arrangement. The authors also used this technique to prepare crosslinked nets and functional yarns that can capture and release foreign particles. This fusion-fission behavior has not been found in other ceramic or polymeric fibers, and it grants a new functionality to GO fibers that can be used to manufacture highly complex architectures with many applications.

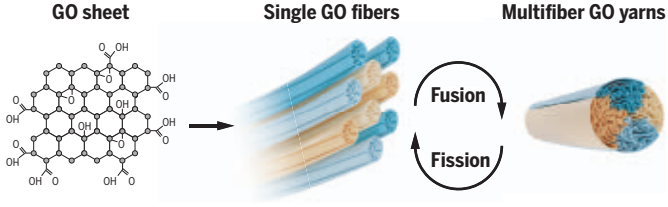
Humans have long relied on fibers to make clothes and various structural materials such as ropes and nets—from the primitive fibers used tens of thousands of years ago (*5*) to the modern functional fibers made using nanomaterials (*1*). In the 20th century, advanced fibers played a role in developing a myriad of technologies, including those used in aeronautics, electronics, and space exploration. Regardless of the era, fibers have been assembled hierarchically to form threads, yarns, ropes, and fabrics with different levels of complexity. As Chang *et al.* demonstrate, applying fusion and fission concepts previously studied in materials science through soft structures such as micelles and vacuoles is of paramount importance when making hierarchical fibers. Because modern fiber assembly techniques require complex machinery and a high-energy input, simplified fiber assembly processes are highly attractive, especially those with minimum energy requirements that can reduce the environmental footprint. Therefore, the development of a simplified and reversible fiber assembly process tackles one of the major challenges that human-made fibers have faced.

GO has attracted a lot of attention as a two-dimensional building block of complex architectures because of its water dispersibility, sub-nanometer-thin nature, synthesis scalability, and chemical reactivity. In the past decade, possibly inspired by the commercial success of carbon fibers, the wet-spinning technique has been used extensively for the integration of GO into fibers with micro- and nanometric diameters (*6*). Because GO fibers can be converted into electrically conductive fibers through reduction processes, they hold an enormous potential for multiple applications as sensors, electronic components, smart textiles, actuators, thermally conductive materials, and high-performance fibers, among others (*7*). Nevertheless, the industrial application of GO fibers has proven difficult. Developing processes to transform GO fibers into hierarchical assemblies like nonwoven fabrics, clothes, ropes, and nets (*8*), and possibly combining GO fibers with functional biological materials such as proteins (*9*), is necessary to achieve advanced functional nanocomposites useful for a broad range of applications.

Low-energy, stimulus-based assembly of GO-based architectures stands as an attractive field with many possible applications where reversibility is key. Textile and high-performance multifilament fibers account for the immediate applications of this technology. A potential use for the fibers may lie in the controlled release and capture of foreign materials such as particles and organic compounds. This requires developing a fundamental understanding of the role played by the individual GO sheets and their nanoscale assembly in the fusion and fission cycles. Even though biological fiber assemblies exhibit a much higher level of complexity and involve multiple components, the reversible assembly of GO fibers mimics nature and holds the refreshing potential to move the field forward. Specifically, GO fiber–based architectures featuring self-healing (*9*), self-sensing, and self-powered actuation should be vigorously pursued to finally bring GO fibers into practical applications. ■

Graphene oxide yarns

Graphene oxide (GO) fibers are assembled through wet-spinning techniques from GO sheets. When multiple GO fibers are immersed in a suitable solvent, they assemble into a hierarchical yarn. This assembly can be reversed, mimicking biological fusion-fission cycles.



GRAPHIC: N. DESAI/*SCIENCE* BASED ON CHANG *ET AL.* (*4*)

[1]Research Initiative for Supramaterials and Aqua Global Innovation Center, Shinshu University, Nagano, Japan. [2]Department of Physics, Binghamton University, Binghamton, NY 13902, USA. Email: alelias@binghamton.edu

REFERENCES AND NOTES

1. Z. F. Liu *et al.*, *Science* **349**, 400 (2015).
2. E. Rideau, R. Dimova, P. Schwille, F. R. Wurm, K. Landfester, *Chem. Soc. Rev.* **47**, 8572 (2018).
3. B. Alberts, *Molecular Biology of the Cell* (Garland Science, 2008).
4. D. Chang *et al.*, *Science* **372**, 614 (2021).
5. E. Kvavadze *et al.*, *Science* **325**, 1359 (2009).
6. Z. Xu, C. Gao, *Acc. Chem. Res.* 47, 1267 (2014).
7. G. Xin *et al.*, *Science* **349**, 1083 (2015).
8. R. Cruz-Silva *et al.*, *ACS Nano* **8**, 5959 (2014).
9. M. C. Demirel, M. Vural, M. Terrones, *Adv. Funct. Mater.* **28**, 1704990 (2018).

ACKNOWLEDGMENTS

The authors have a US patent (no. 9,284,193) for making graphene oxide films and fibers.

10.1126/science.abh2283

EVOLUTION

Illuminating the first bacteria

A new analysis aims to uncover the root of the bacterial tree of life

By **Laura A. Katz**

The ability to sequence genes and, more recently, whole genomes has transformed our understanding of the tree of life by elucidating the tremendous diversity of microorganisms and by placing plants, animals, and fungi as branches nested among microbial lineages (*1*–*3*). The resulting evolutionary tree divides life into three domains: the exclusively microbial Bacteria and Archaea, and Eukarya, organisms whose cells contain nuclei (including ciliates, amoebae, and animals). Yet, the ordering of the earliest branching events on the tree and the nature of now-extinct ancestors remains unclear. On page 588 of this issue, Coleman *et al.* (*4*) provide a new estimate of the root of the bacterial tree of life, that is, the ancestor from which all bacterial species are derived. Knowledge of the root of the bacterial tree is important because it defines the evolutionary starting point for the tremendous diversity of Bacteria and offers glimpses into the nature of the first bacterial cells.

Although genome sequencing has given biologists a wealth of data for estimating phylogenies (i.e., evolutionary trees), it has also revealed the importance of nonvertical inheritance because genes, and sometimes whole genomes, have been transferred across species boundaries (*2*, *3*, *5*). This lateral gene transfer (LGT; also called horizontal gene transfer) has substantially altered bacterial evolution, for example, through the spread of antibiotic resistance and the acquisition of metabolic pathways (*2*). In some lineages, such as the bacterial clade Thermatogales, a large number of LGTs have generated a chimeric lineage: An estimated 20% of the genes in the *Thermotoga* genus have been acquired from Archaea rather than through vertical transmission during binary division (*5*). Eukaryotes represent another chimera, because they arose through the merger of an archaeon and a bacterium, the latter eventually became a mitochondrion (*6*, *7*). Events such as these make estimation of the root of a tree challenging because transferred genes reflect different histories than vertically inherited genes and thus can confound inferences.

Coleman *et al.* use a method developed by their group called amalgamated likelihood estimation (ALE) to estimate the position of the root of the bacterial tree of life while attempting to account for gene duplications and losses as well as LGTs. To determine the root of the bacterial tree, Coleman *et al.* first construct an unrooted phylogeny based on 62 single-copy genes sampled from 265 bacterial genomes and then use ALE to evaluate various positions of the root on this tree through analyses of the evolutionary history of 11,000 gene families sampled from across the same lineages.

Coleman *et al.* estimate that two-thirds of the transmissions among gene families analyzed are vertical and that the root of the bacterial tree of life likely lies between "Gracilicutes"—predominantly Gram-negative bacteria, including Proteobacteria, Acidobacteria, and Spirochaeta—and "Terrabacteria," which includes Gram-positive bacteria, Cyanobacteria, Firmicutes, and "CPR" (candidate phyla radiation); taxonomic names in quotes indicate uncertainty (see the figure). The placement of two clades—"DST" (Deinococcota, Synergistota, and Thermotogota) and Fusobacteriota (an anaerobic clade of Gram-negative bacteria)—that are the recipients of a disproportionate number of LGTs is unresolved. This is exemplary of the uncertainty introduced by LGT in reconstructing a bacterial tree of life.

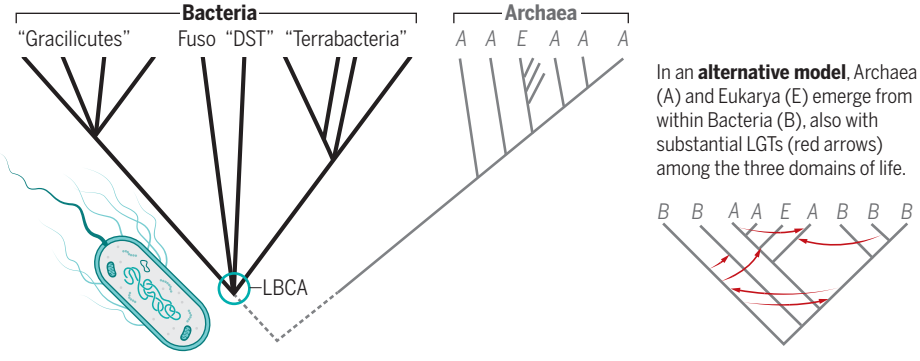
Knowing the root of the bacterial tree of life allows both polarization of the evolution of bacterial traits and reconstruction of the last bacterial common ancestor (LBCA), at least for those aspects that have evolved through vertical transmission. Coleman *et al.* infer that LBCA was a free-living rod-shaped cell surrounded by a double membrane and with a flagellum. The genome of this ancestor of all Bacteria encoded key components of informational processing (e.g., DNA transcription and replication); metabolic pathways, including CO_2 fixation and the Krebs cycle; and elements of the CRISPR-Cas9 adaptive immune system. Overall, these inferences agree with findings of others on the physiological and morphological features of LBCA (*8*, *9*).

Did they get it right? It is difficult to know because reconstructing such ancient events [the earliest fossil bacteria are estimated to have existed around 3 billion to 3.4 billion years ago (*10*)] presents a challenge even in the absence of LGT. Results are dependent on the parameters used in the mathematical models, and although Coleman *et al.* extensively evaluate these parameters through simulations, assumptions about the relative rates of gene duplication and LGT have been questioned (*11*). To further complicate inferences, Coleman *et al.* note the importance of taxon sampling because changing representative species within major bacterial clades affects the estimate of the position of the root. This suggests that future analyses including different, and perhaps newly discovered, lineages may alter both estimates of the root of the bacterial tree of life and inferences about LBCA.

Another confounding factor is the strength of the assumption that Bacteria are monophyletic (that Bacteria descended from a common ancestor that existed more recently

The root of the bacterial tree of life

Coleman *et al.* infer that the root of the bacterial tree of life lies between "Gracilicutes" and "Terrabacteria," which enables reconstruction of the last bacterial common ancestor (LBCA). Unresolved are two clades, "DST" (Deinococcota, Synergistota, and Thermotogota) and Fusobacteriota (Fuso), that were the recipients of a disproportionately large number of lateral gene transfers (LGTs). Names in quotes represent provisional taxonomic terms.



GRAPHIC: N. CARY/*SCIENCE*

Department of Biological Sciences, Smith College, Northampton, MA, USA. Email: lkatz@smith.edu

than their shared common ancestor with Archaea). This requires either the independent evolution of Bacteria and Archaea at the origin of life (*12*) or a complex pattern of diversification and extinction events to give rise to reciprocally monophyletic Bacteria and Archaea (i.e., the evolution of two distinct lineages of microorganisms). Coleman *et al.* state that they are agnostic as to the monophyly of Bacteria, pointing out the limitations of previous studies that rely on an archaeal outgroup for rooting the bacterial tree of life. Instead, their model excludes Archaea and the nested eukaryotes (*6*, *7*) in their estimates of the bacterial root. Alternative models suggest that Archaea descended from LBCA (*13*, *14*), in which case Archaea would have to be included to determine the root of the bacterial tree of life (see the figure).

Whereas some will contest the proposed position of the root of the bacterial tree of life, others challenge the concept of a tree as a model for the diversification of life on Earth (*14*, *15*). Although Coleman *et al.* may have accurately captured the vertical portion of the bacterial tree of life, the tree that they root is missing the history of biological innovations and ecological adaptations that are derived from LGTs. To portray these lateral events, some have suggested that the tree of life be replaced with a circle, a web, or a network depicting the complex history of inheritance (*2*, *3*, *5*). It has been suggested that "to save the trees, one might define organisms as more than the sums of their genes and imagine organismal lineages to have a sort of emergent reality" (*2*). In other words, there is room for alternative methods and innovation to encompass both vertical and lateral inheritance, focusing on the evolutionary history of phenotypic features such as metabolism, morphology, and life history. ■

REFERENCES AND NOTES

1. L. A. Hug *et al.*, *Nat. Microbiol.* **1**, 16048 (2016).
2. W. F. Doolittle, *Science* **284**, 2124 (1999).
3. T. Dagan, W. Martin, *Genome Biol.* **7**, 118 (2006).
4. G. A. Coleman *et al.*, *Science* **372**, eabe0511 (2021).
5. J. P. Gogarten, J. P. Townsend, *Nat. Rev. Microbiol.* **3**, 679 (2005).
6. T. A. Williams *et al.*, *Nature* **504**, 231 (2013).
7. A. B. Collens, L. A. Katz, *J. Hered.* **112**, 140 (2021).
8. W. F. Martin, *Front. Microbiol.* **11**, 817 (2020).
9. M. Miyata *et al.*, *Genes Cells* **25**, 6 (2020).
10. E. J. Javaux, *Nature* **572**, 451 (2019).
11. T. J. Treangen, E. P. C. Rocha, *PLOS Genet.* **7**, e1001284 (2011).
12. F. L. Sousa *et al.*, *Philo. Trans. R. Soc. London Ser. B* **368**, 1 (2013).
13. T. Cavalier-Smith, *Int. J. Syst. Evol. Microbiol.* **52**, 7 (2002).
14. J. A. Lake *et al.*, *Philos. Trans. R. Soc. London B Biol. Sci.* **364**, 2177 (2009).
15. S. Gribaldo, C. Brochier, *Res. Microbiol.* **160**, 513 (2009).

ACKNOWLEDGMENTS

I thank A. Cote-L'Heureux and several other colleagues for thought-provoking conversations on this topic. L.A.K. is supported by grants from the National Science Foundation (OCE-1924570, DEB-1651908, and DEB-1541511) and National Institutes of Health (R15HG010409).

10.1126/science.abh2814

PALEONTOLOGY

Making sense of dinosaurs and birds

Advances in imaging and statistics illuminate dinosaur sensory biology and behavior

By **Lawrence M. Witmer**

Birds are dinosaurs. A couple of decades of revealing finds and careful phylogenetic work have shown that birds are nested within the group of small, mostly predatory, and running dinosaurs that includes dromaeosaurs and troodontids, among others (*1*). Now, scientists are using that phylogenetic tapestry to trace the evolution of traits such as sensory biology and behavior. Two studies in this issue pair cutting-edge imaging with sophisticated statistical analyses to explore the evolution of the hearing apparatus and other sensory systems. On page 601, Hanson *et al.* (*2*) show that the inner ear of dinosaurs, as well as that of other archosaurs (the group that includes crocodilians, dinosaurs, and birds), provides clues to locomotion, hearing, and the evolution of vocalization. On page 610, Choiniere *et al.* (*3*) examine the inner ear and also aspects of the visual system to demonstrate the notably owl-like nocturnal adaptations of one group of birdlike dinosaurs.

Until recently, the advances presented by these teams of authors were unthinkable, in that many aspects of internal anatomy and certainly their connection to habits like parental care and daily activity patterns had been out of reach. Two innovations that may fairly be called revolutionary made these discoveries possible. The first revolution was x-ray computed tomography (CT) scanning, which allowed scientists to peer through flesh, bone, and, in the case of fossils, rock to visualize and model internal anatomical structures (*4*). Early on, CT scanners at hospitals were the main option, but the widespread deployment of micro- and nano-CT scanners over the past several years has permitted the collection of numerous high-resolution datasets. Moreover, the "open-data movement" (*5*) has resulted in the sharing of open-access data collected by laboratories worldwide. Thus, Hanson *et al.* and Choiniere *et al.* were able to analyze the anatomy of inner ears from 124 and 91 extinct and extant species, respectively—an increase by perhaps an order of magnitude from what was possible 10 to 15 years ago. The second revolution is the widespread use of sophisticated statistical software that allows such large datasets to be queried and probed with unprecedented detail and clarity. Both studies implemented R (*6*), the open-source, seemingly infinitely tweakable software that has taken by storm disciplines ranging from science to economics and has contributed to the widespread coding movement. The combination of open-access data and open-source analytical tools means that both studies reach a high standard for replicability.

Anatomically, Hanson *et al.* and Choiniere *et al.* both focused on the inner ear, a tiny organ consisting of a labyrinth of canals that is buried deep within the skull. In the past, that position made the inner ear inaccessible to study, but that same inaccessibility protected it during the process of fossilization, which allows the inner ear labyrinth to be digitally extracted relatively easily with CT scanning. And that is indeed helpful because the inner ear is two sensory organs in one: Three delicate, mutually perpendicular semicircular canals (SSCs) sit atop the inner ear, and hanging below is the cochlea. The SSCs are involved with the sense of equilibrium and balance and have been targeted as a source of information on posture and locomotion. The cochlea is the hearing organ. Although the actual sensory tissues and neural wiring are long gone in the fossils, the shape and orientation of the bony canals allow some broad-brush assessments of function, and the SSCs have been used to infer locomotor shifts in groups as diverse as primates (*7*) and plesiosaurs (*8*). Likewise, the length of the bony cochlea is a good proxy for the sensory basilar papilla and has been used to study auditory evolution in prior studies of birds and dinosaurs (*9*).

The Hanson *et al.* study is a landmark, showing that the two parts of the inner ear are different modules that follow somewhat separate evolutionary paths, and both have stories to tell about behavior. The SSCs can be used to discriminate different locomotor

Department of Biomedical Sciences, Heritage College of Osteopathic Medicine, Ohio University, Athens, OH 45701, USA. Email: witmerl@ohio.edu



Scientists use x-ray computed tomography scanning to peer inside fossils, such as the early bird *Archaeopteryx* (top) and a birdlike troodontid theropod dinosaur (bottom). Visualizing the inner ear's semicircular canals and cochlea provides information on locomotion and hearing.

grades. It may not be surprising that species at opposite ends of the spectrum, like quadrupeds (e.g., early archosaurs and extant crocodilians) and highly aerobatic extant birds, can be differentiated easily, but the middle ground is exciting because this is where the transition to birds takes place. This multivariate "shape cloud" includes the small, birdlike nonavian theropods and early birds such as the iconic *Archaeopteryx*, as well as some flightless birds (e.g., ostriches and penguins) and "reluctant fliers" (e.g., chickens and cranes) that have apparently reverted their SSC structure. Notably, some very birdlike theropods, such as troodontids, plot very closely to early "simple fliers" like *Archaeopteryx* (see the photos), providing further evidence that multiple groups of birdlike theropods were experimenting with aerodynamic flight surfaces (feathers) and that birds simply represent the group that survived long enough to hone and perfect flight. By contrast, the long cochlea of extant birds evolved very early and is found in almost all archosaurs but is absent in other reptiles like turtles and lizards. Hanson *et al.* discovered that extant archosaurs (crocodilians and birds) differ from other reptiles in that not only do archosaurs have parental care, but archosaur babies, unlike turtle or lizard babies, also chirp to capture their parents' attention. Thus, the authors perceptively suggest that juvenile vocalization to elicit parental care coevolved with the elongation of the cochlea in archosaurs.

Choiniere *et al.* also found the cochlea to be informative about behavior in small birdlike theropods, many of which had the elongated cochlea suggestive of sensitive hearing. But two groups—again troodontids, but especially an unusual group called alvarezsaurids—had exceptionally long, curved cochleae. In some alvarezsaurids, such as *Shuvuuia*, the cochlea approaches—in length and structure—that of today's barn owl, perhaps the most notable avian nocturnal specialist. To further test the idea that *Shuvuuia* and other alvarezsaurids might have been nocturnal, the team turned to the visual system. Fortunately, dinosaurs (including birds) and most other reptiles had bony eye rings that previous researchers (*10*) had shown provided links to daily activity patterns, whereby the eye rings of nocturnal species tend to have wider openings to admit more light than the rings of species active during the day. Choiniere *et al.* carefully reconstructed the bony eye rings in two alvarezsaurids, and, sure enough, they plotted with extant nocturnal birds and reptiles. Thus, two semi-independent systems—the cochlea and eye ring—point to alvarezsaurids being nighttime specialists, perhaps providing new clues to the enigmatic habits of these peculiar, short-armed species (*11*).

Documenting life's history with fossil discoveries remains the core of paleontology, but these two studies typify a new wave of paleontologists, armed not with a pick and shovel but with a CT scanner and R code. Likewise, the basic phylogenetic work of who is related to whom continues, but now these phylogenies are being used to explain evolution's narrative. Teeth and limbs will always provide clues for reconstructing evolutionary histories, but relatively new anatomical players like the inner ear and bony eye rings are opening new windows into the past. And through such windows, we now can hear baby dinosaurs calling to their parents, watch a small troodontid clumsily fly to a low branch, and maybe catch a glimpse of the moon shadow of a passing *Shuvuuia*. ■

PHOTO: LAWRENCE WITMER/OHIO UNIVERSITY

REFERENCES AND NOTES

1. A. Cau, *Boll. Soc. Paleontol. Ital.* **57**, 1 (2018).
2. M. Hanson *et al.*, *Science* **372**, 601 (2021).
3. J. N. Choiniere *et al.*, *Science* **372**, 610 (2021).
4. L. M. Witmer et al., in *Anatomical Imaging: Towards a New Morphology* (Springer, 2008), pp. 67–87.
5. T. G. Davies *et al.*, *Proc. Biol. Sci.* **284**, 20170194 (2017).
6. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, 2020).
7. F. Spoor *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10808 (2007).
8. J. M. Neenan *et al.*, *Curr. Biol.* **27**, 3852 (2017).
9. S. A. Walsh, P. M. Barrett, A. C. Milner, G. Manley, L. M. Witmer, *Proc. Biol. Sci.* **276**, 1355 (2009).
10. L. Schmitz, R. Motani, *Science* **332**, 705 (2011).
11. P. Senter, *Paleobiology* **31**, 373 (2005).

10.1126/science.abi5697

HYPOTHESIS

Carbohydrates, insulin, and obesity

Insulin plays a role in body fat regulation independent of dietary carbohydrates

By **John R. Speakman**[1,2,3,4] *and* **Kevin D. Hall**[5]

The primary cause of common human obesity remains uncertain. There are several plausible explanations, including the popular "carbohydrate-insulin" model (CIM), which suggests that body-fat gain results from consumption of carbohydrates that stimulate postprandial insulin, which promotes energy storage and further intake in a vicious cycle. The theoretical basis of the CIM has been refuted by several recent experiments. We suggest that although insulin plays an important role in body fat regulation, the CIM fails because it focuses on the direct action of insulin on adipose tissue after the consumption of a meal containing carbohydrates. Rather, we propose that the role of insulin in obesity may be better understood by considering its pleiotropic action on multiple organs that is driven by factors mostly independent of carbohydrate intake. Reconsidering the role of insulin may improve our understanding of the causes of obesity and its treatment.

The CIM puts the adipocyte at center stage by highlighting the role of insulin in promoting fat storage and inhibiting its release (*1*). Carbohydrate consumption stimulates insulin secretion, which partitions circulating fuels (such as triglycerides) toward storage in adipose tissue. This is postulated to reduce the energy available for metabolically active tissues such as skeletal muscle. Deprived of fuel, these nonadipose tissues experience a state of cellular "internal starvation" that motivates the individual to respond as they would to actual starvation—by seeking and consuming more food and reducing metabolic rate to conserve energy. Therefore, according to the CIM, excess energy consumption is the result of adipose tissue fat storage due to carbohydrate-driven postprandial insulin. The corresponding "obesity solution" is to replace carbohydrates with dietary fat, which does not stimulate postprandial insulin secretion: the so-called low-carb, high-fat or "ketogenic" diet.

Conceptually, testing the CIM should be simple: Randomize people to consume different diets varying widely in carbohydrates and fat and then measure obesity prevalence in each diet group. But such an experiment, if prolonged, raises ethical concerns due to potential harm to health, and if short, raises questions about whether the absence of effects was due to the short duration. Moreover, ensuring adherence to the assigned diets is not currently possible in people living outside a laboratory setting. A good (but imperfect) solution is to perform the studies on experimental animals, with the caveat that improved experimental control comes at the expense of failing to completely capture the complexity of the situation in humans.

In this context, a large dietary manipulation study exposed mice to 29 different diets to address the impact of diet composition on body fatness (*2*). In 16 of these diets, the macronutrient manipulations allowed a direct test of the predictions of the CIM because protein was held constant (in 8 diets at 10% and 8 diets at 25%) while fat and carbohydrate varied reciprocally between 10 and 80%. The carbohydrates were a mix of corn starch, maltodextrin, and sucrose (typically regarded as refined "high glycemic index" carbohydrates that induce high blood glucose and insulin responses). The mice were exposed to the diets for 12 weeks, which is roughly equivalent to 9 years in humans. The CIM prediction is that as dietary carbohydrate increased, so would postprandial insulin, and the mice would develop obesity and eat more total calories. However, the opposite happened. Mice feeding on diets with a high proportion of carbohydrates ate fewer calories and gained both less body weight and body fat despite higher postprandial insulin. Although this appears to refute the CIM, uncritical extrapolation from mice to humans is problematic.

Although controlled feeding studies lasting multiple years in humans are not feasible, the CIM can be tested by examining its predictions using shorter-term experiments. For example, if individuals are exposed to diets with very different proportions of fat and carbohydrate, the CIM predicts that a high-carbohydrate diet will result in greater postprandial blood insulin concentrations that drive fat accumulation in adipose tissue, thereby increasing hunger and energy intake compared to a low-carbohydrate diet. This prediction was tested in a month-long inpatient metabolic ward study where 20 adults were randomized to receive a diet composed of ~10% carbohydrate, ~75% fat or a diet with ~10% fat, 75% carbohydrate and instructed to eat as much or as little as they wanted (*3*). After 2 weeks, participants switched to the alternate diet. In accordance with CIM predictions, the high-carbohydrate diet resulted in much higher concentrations of circulating postprandial insulin and therefore should have partitioned more energy to body fat storage, thereby increasing hunger and energy intake compared to the low-carbohydrate diet. However, ~700 kcal/day less food was consumed on the high-carbohydrate diet, and participants reported both diets to be equally satisfying and pleasant with no differences in hunger or fullness. Furthermore, despite substantially higher daily insulin secretion, only the high-carbohydrate diet resulted in significant body fat loss.

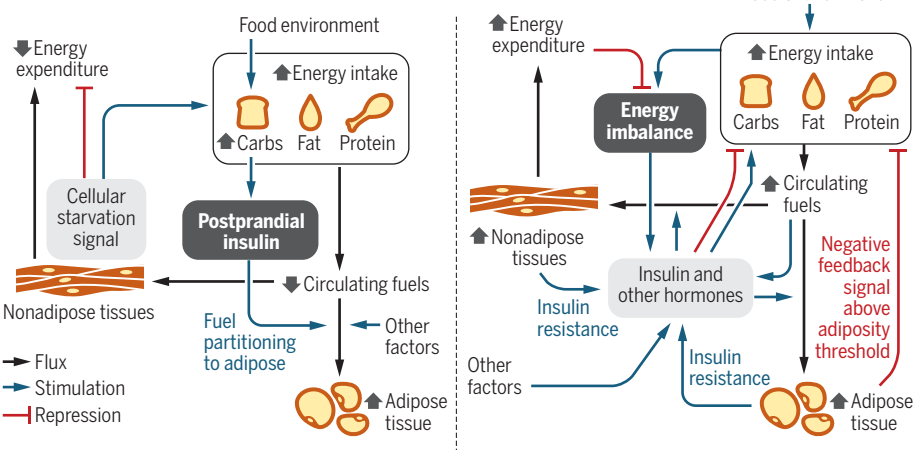
Although this was a relatively short-term experiment, another study found significantly increased satiety after 10 to 15 weeks of consuming a high-carbohydrate diet compared with a low carbohydrate diet (*4*). Furthermore, a 1-year study in freely living individuals randomized to consume low-carbohydrate versus high-carbohydrate diets found no sustained differences in objective measurements of energy intake (*5*). As a result, long-term average weight loss was almost identical, and individual differences in postprandial insulin secretion did not predict who lost most weight on each diet (*6*). Therefore, the energy intake predictions of the CIM failed to materialize in both short-term and long-term studies.

The CIM also predicts that decreased insulin secretion during a low-carbohydrate diet should increase body fat loss by mobilizing fat trapped in adipose tissue, thereby restoring the fuel supply to metabolically active tissues, compared with an isocaloric diet with higher carbohydrate. Alleviation of cellular "internal starvation" in these tissues should therefore increase energy expenditure. Two studies admitted participants with overweight or obesity to metabolic wards with strict control over food

[1]Center for Energy Metabolism and Reproduction, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. [2]State Key Lab of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [3]Institute of Biological and Environmental Sciences, University of Aberdeen, Scotland, UK. [4]CAS Centre of Excellence in Animal Evolution and Genetics, Kunming, China. [5]National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, USA. Email: j.speakman@abdn.ac.uk; kevinh@niddk.nih.gov

intake for 1 to 2 months and showed that although carbohydrate restriction led to substantial decreases in daily postprandial insulin secretion, both studies found less body fat loss with carbohydrate restriction compared with isocaloric high-carbohydrate diets (*7*, *8*). In one study, the reduced carbohydrate diet led to a significant decrease in energy expenditure (*7*). In the other study, a low-carbohydrate ketogenic diet led to a small, transient increase in energy expenditure that dissipated within a few weeks (*8*).

Two models of body weight regulation



Carbohydrate insulin model
A food environment promoting increased carbohydrate intake stimulates postprandial insulin, which partitions circulating ingested fuels into adipose tissue. This reduces the flux of these fuels into nonadipose tissues, leading to a cellular starvation signal with two consequences: reduced energy expenditure and further stimulation of intake.

Energy balance model
Energy imbalance is driven by increased energy intake caused by the obesogenic food environment. Insulin facilitates the uptake of circulating fuels and provides a negative feedback signal to the brain, which regulates intake in combination with other hormones and signals from adipose tissue when adiposity rises above a critical level.

In contrast to these inpatient controlled feeding studies, two outpatient studies reported increased energy expenditure during low-carbohydrate diets in people who were weight-stable following a period of weight loss (*9*, *10*). However, these results were likely due to a miscalculation of energy expenditure (*11*). Indeed, the reported energy expenditure increases in people on low-carbohydrate diets were not consistent with differences in body weight or measured components of energy expenditure such as resting metabolic rate, physical activity, or skeletal muscle work efficiency. Supporting the absence of effects on energy expenditure, in the mouse study of 29 diets (*2*) there were no effects of variable carbohydrate content on energy expenditure or physical activity.

Supporters of the CIM have criticized the human and mouse experiments that failed to confirm CIM predictions. The test diets in the mouse study were claimed to be inadequate. The human experiments were argued to be too short. We propose that these data should not be ignored, but rather should inform an improved model in agreement with the data (see the figure). Although the "carbohydrate" half of the CIM is disputed by recent experimental tests, this does not discount an important role for insulin in regulating body fat. Genetic manipulation of insulin secretion in mice (*12*) or pharmacologic inhibition of insulin secretion in humans (*13*) can lead to reduced body fat in the absence of diabetes. People with diabetes often experience weight loss prior to diagnosis, and treatment that increases endogenous insulin secretion or exogenous insulin therapy often results in weight regain. Nevertheless, hyperinsulinemia is not associated with meaningful differences in adiposity, and hyperinsulinemia does not necessarily result in increased weight or reliably predict future weight changes (*14*). Furthermore, genetic polymorphisms derived from genome-wide association studies for body mass do not identify targets linked to insulin action in adipose tissue as important causal variants for obesity. Therefore, the extent to which susceptibility to obesity is explained by differences in insulin secretion or insulin action is uncertain, but direct action of carbohydrate-driven postprandial insulin on adipose tissue is unlikely to be the dominant driver of common obesity, as proposed by the CIM.

Postprandial insulin is not the most important factor regulating adipose uptake and storage of fat, which can occur without increasing circulating insulin above basal concentrations (*15*). Basal insulin may be more important because adipose tissue release of fat is exquisitely sensitive to changes in insulin around basal levels, but the effect of insulin quickly saturates in the postprandial range and therefore may be relatively insensitive to dietary carbohydrate. Furthermore, reduction of dietary fat decreases basal insulin to a similar degree as isocaloric reduction in carbohydrate (*7*), indicating that basal insulin concentrations respond to the imbalance between energy intake and expenditure as much as diet composition per se.

Insulin has pleiotropic effects on multiple organs, and its role in body fat regulation is best understood as part of a dynamic network of factors controlling and mediating the effects of energy imbalance. For example, insulin provides a negative feedback signal to the brain that combines with signals from adipose tissue when body fat rises above a critical threshold concentration and serves to regulate energy intake. Adipose tissue and ectopic fat deposition in nonadipose tissues can also drive insulin resistance, thereby affecting circulating insulin concentrations independent of dietary carbohydrate. Therefore, the mechanisms underlying the effects of insulin on adiposity are more complex than proposed by the CIM. Failure of the CIM should not be taken to mean that low-carbohydrate, high-fat diets cannot be beneficial for weight loss. However, direct modulation of the carbohydrate-insulin axis in adipose tissue is unlikely to be the primary mechanism underpinning body fat loss in individuals successfully engaged in such diets. A new model of the role of insulin in obesity is required that is commensurate with data refuting key aspects of the CIM. ■

REFERENCES AND NOTES

1. D. S. Ludwig, C. B. Ebbeling, *JAMA Intern. Med.* **178**, 1098 (2018).
2. S. Hu *et al.*, *Mol. Metab.* **32**, 27 (2020).
3. K. D. Hall *et al.*, *Nat. Med.* **27**, 344 (2021).
4. K. J. Shimy *et al.*, *J. Endocr. Soc.* **4**, bvaa062 (2020).
5. J. Guo, J. L. Robinson, C. D. Gardner, K. D. Hall, *Obesity* (Silver Spring) **27**, 420 (2019).
6. C. D. Gardner *et al.*, *JAMA* **319**, 667 (2018).
7. K. D. Hall *et al.*, *Cell Metab.* **22**, 427 (2015).
8. K. D. Hall *et al.*, *Am. J. Clin. Nutr.* **104**, 324 (2016).
9. C. B. Ebbeling *et al.*, *BMJ* **363**, k4583 (2018).
10. C. B. Ebbeling *et al.*, *JAMA* **307**, 2627 (2012).
11. K. D. Hall, J. Guo, J. R. Speakman, *Int. J. Obes.* **43**, 2350 (2019).
12. N. M. Templeman, S. Skovsø, M. M. Page, G. E. Lim, J. D. Johnson, *J. Endocrinol.* **232**, R173 (2017).
13. Z. Huang, W. Wang, L. Huang, L. Guo, C. Chen, *Obesity* (Silver Spring) **28**, 2098 (2020).
14. D. Tricò, A. Natali, S. Arslanian, A. Mari, E. Ferrannini, *JCI Insight* **3**, e124912 (2018).
15. A. C. Carpentier, *Am. J. Physiol. Endocrinol. Metab.* **320**, E653 (2021).

ACKNOWLEDGMENTS

J.R.S. is supported by the 1000 Talents Program, a professorial fellowship from the Chinese Academy of Sciences President's International Fellowship Initiative (PIFI), and a Wolfson Research Merit Award from the UK Royal Society. K.D.H. is supported by the Intramural Research Program of the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

10.1126/science.aav0448

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*



Scientist Emily Grossman addresses climate activists during a demonstration on 22 November 2019.

PHOTO: MARK KERRISON/ALAMY STOCK PHOTO

BOOKS *et al.*

SCIENCE AND SOCIETY

More than the message

A new guide offers advice for navigating barriers to successful science communication

By **Jonathan Wai**

"If it disagrees with experiment, it's wrong," argued Richard Feynman in 1964 (*1*). "It doesn't make any difference how beautiful your guess is, it doesn't make any difference how smart you are, who made the guess, or what his name is." While this may indeed be true of a scientific theory, the effective communication of that theory to a broader audience is a bit more complicated. Building on such predecessors as *The Craft of Science Writing* and *Escape from the Ivory Tower*, in *Getting to the Heart of Science Communication*, Faith Kearns reminds readers that how a scientist presents his or her research can drastically affect how people outside of academia think, and especially how they feel, about a given topic (*2*, *3*).



Getting to the Heart of Science Communication
Faith Kearns
Island Press, 2021.
280 pp.

The book offers a view from the front lines of science communication, profiling practitioners who explain their journeys and share stories of relationship building and community engagement. Framing herself as a scientist turned science communicator, Kearns describe her vision for the future of the field, one in which relational communication is fundamental.

In Part I, Kearns describes the evolution of science communication, highlighting a 1951 *Science* article by William Hewitt (*4*) that established a road map for the future of scientific communication. Hewitt argued for the creation of a field of science communicators who would be charged with organizing the rapidly expanding body of research to establish productive communication channels between scientists and the public. Kearns describes this vision as prescient while noting that contemporary discussions about science communication often overprioritize finding ways to encourage or incentivize senior scientists to communicate about their work.

Turning her attention to those who do pursue scientific communication as a career, Kearns notes that many such individuals are not in tenured positions and suggests that more infrastructure is needed to support this growing community. Institutions of higher education claim to value the role of public scholarship and engagement, she argues, and if that is the case, they should defend science communicators when they come under fire, just as they reap the rewards when communication efforts are successful.

In the book's second section, Kearns focuses on relational tools of communication, including how to listen, work through conflict, and understand trauma. She notes that in situations where emotions, conflict, and power are salient to communication, communicating effectively requires the cultivation of relationships and the use of tools such as listening and empathizing.

Kearns argues that, until recently, scientists had mostly taken an agnostic approach to the identity of the communicator. She explains why this is worth rethinking, demonstrating how factors such as race, gender, sexuality, age, ability, class, and power can influence who asks questions, what questions they ask, and who benefits from the information being disseminated.

The book closes with Kearns's vision of what science communication can be, including how we can ensure that it is equitable, inclusive, and just and that communicators emphasize self-care and collective care. She encourages those who engage in this work to prioritize their emotional and physical selves and to incorporate practices such as establishing boundaries, taking time for reflection, and making space for joy.

Kearns correctly highlights that there remains a disconnect between doing science well and communicating science well. However, I would have liked to have seen more engagement with the realities that can undermine effective communication, including the tenure incentive structure and time constraints experienced by academic scientists.

It is possible that, as the science communication community grows, a plurality of communication efforts will need to be valued and incentivized, but it is worth making the effort to do so now. If these challenges can be addressed, and future scientists learn how to effectively communicate their work, it is more than the lay community that stands to benefit. Scientists themselves will gain valuable insights by engaging in conversations with the people and communities that they hope to help. ■

The reviewer is at the Department of Education Reform and Department of Psychology, University of Arkansas, Fayetteville, AR 72701, USA. Email: jwai@uark.edu

REFERENCES AND NOTES

1. Project Tuva: Richard Feynman's Messenger Lecture Series (2009); www.microsoft.com/en-us/research/project/tuva-richard-feynman/.
2. S. Carpenter, Ed., *The Craft of Science Writing: Selections from The Open Notebook* (The Open Notebook, 2020).
3. N. Baron, *Escape from the Ivory Tower: A Guide to Making Your Science Matter* (Island Press, 2010).
4. W. F. Hewitt Jr., *Science* **114**, 134 (1951).

10.1126/science.abi5203

ANTHROPOLOGY

The people at the dawn of civilization

Sumerian language and culture take center stage in a new anthropological analysis

By **Andrew Robinson**

"A narrative history of the early Near East is virtually impossible," confessed archaeologist Paul Collins, the curator for the ancient Near East at the Ashmolean Museum, Oxford, in his 2016 study, *Mountains and Lowlands: Ancient Iran and Mesopotamia*. I was reminded of this comment while reading his stimulating new book, *The Sumerians*, the latest contribution to the Lost Civilizations series, which eschews narrative in favor of separate chapters on, for example, "The first cities" and "The first writing." Even the book's opening chronology refers to "Sumerian" only as a descriptor of language and script, not as a people, state, dynasty, or empire.

As Collins frankly concludes, "the Sumerians were never actually lost, and in fact could never be lost, since they may never have existed, at least not as a distinct ethno-linguistic population. What was lost and rediscovered was the Sumerian language as preserved in the cuneiform script of Mesopotamia." Documented on clay tablets in Uruk in the late fourth millennium BCE, and first used at Ur to write Sumerian around 2800 BCE, if not before, cuneiform was later used to write Akkadian, a Semitic language (unlike Sumerian, which was an isolate) spoken in Assyria and Babylonia, along with many other languages of the Near East, until it finally disappeared as a script soon after the first century CE.

The Sumerians, for all their doubtful status as a formal society, have a remarkable list of achievements to their credit. Besides being the world's earliest attested civilization in the fourth millennium BCE, they invented cuneiform—the world's earliest writing—and the sexagesimal system of mathematics. Their cities, such as Uruk and Ur, were the headquarters of the world's earliest city-states, with bureaucracies, legal codes, divisions of labor, and a money economy.

Sumerian art, literature, and theology had a profound influence on culture and religion, long after their language died, around 2000 BCE, serving as the prototype of Akkadian, Hurrian, Canaanite, Hittite, and eventually biblical literature. The Akkadian epic of Gilgamesh was based on a legendary Sumerian king of Uruk, and the Hebrew patriarch Abraham hailed from Ur, according to the Bible. In the 20th century CE, notes Collins, the sculptor Henry Moore ranked Sumerian sculpture—appealingly illustrated throughout the book—among the greatest in world history for its "richness of feeling for life and its wonder and mystery." Meanwhile, science fiction writers, including astronomer Carl Sagan, speculated that Sumerian civilization might provide evidence of extraterrestrial contact.

A fine example of the complexity of Sumerian history is the famous "Great Death Pit" in the royal tombs at Ur, dating to around 2400 BCE, that were excavated between 1922 and 1934. One of the site's best-preserved tombs revealed a queen, beside whom was found a lapis lazuli cylinder seal inscribed with three cuneiform signs giving her name. On the assumption that she was Sumerian, the first two signs were originally read as *shub* and *ad*, and the third as *nin*, meaning "queen": Queen Shub-ad. But later analysis suggested that the first two signs make more sense if read in Akkadian as *pu* and *abi* or, more correctly, as "Pu-abum," the Akkadian for "word of the Father," while the third sign is read as *eresh* (Akkadian for "queen"): Queen Pu-abi. The most celebrated Sumerian queen "turns out to be Akkadian!" observes Collins.

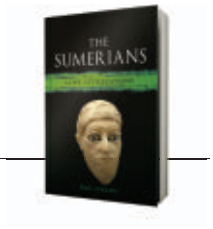
By the second millennium BCE, when Sumerian was no longer spoken in the street or in the court, it continued to inspire Akkadian-speaking rulers. At least nine Sumerian narrative poems based on the heroic kings of Uruk were composed by royal scribes. Among the best known is "Enmerkar and the Lord of Aratta," in which King Enmerkar sends a messenger across seven mountains to Aratta, demanding that the unnamed king of this region supply lowland Uruk with skilled workers and the region's precious metals and stones. The message proves too long for his messenger to memorize, so Enmerkar writes it on a clay tablet, thus inventing writing. Although the lord of Aratta cannot read the tablet, he realizes that he has been outwitted by a cleverer king—a suitably enigmatic comment on writing's still-mysterious origins from a civilization made vivid by Collins's clear and expert text. ■



A bust crafted by archaeologist Katharine Woolley in 1928 showcases an elaborate headdress that belonged to Queen Shub-ad/Pu-abi.

IMAGE: COURTESY OF THE PENN MUSEUM/RECONSTRUCTION OF HEAD AND HEADDRESS OF LADY PU-ABI FROM UR BY KATHARINE WOOLLEY, CA. 1930

The Sumerians
Paul Collins
Reaktion Books, 2021.
208 pp.



The reviewer is the author of *Writing and Script: A Very Short Introduction* (Oxford University Press, 2009). Email: andrew@andrew-robinson.org

10.1126/science.abf9619



LETTERS

During natural disasters such as the eruption of La Soufrière volcano on the island of St. Vincent, biosecurity screening of imported goods can fall by the wayside.

PHOTO: ROBERTSON S. HENRY/REUTERS

Edited by **Jennifer Sills**

Biosecurity for humanitarian aid

Since 9 April, the La Soufrière volcano on St. Vincent in the Caribbean has been highly active with multiple explosive eruptions. More than 20,000 people were forced to evacuate, and continued geologic activity threatens island-wide water and food supplies (*1*). Although globalization can facilitate rapid delivery of essential supplies to regions affected by natural disasters like this one, foreign aid brings with it the potential for environmental contaminants such as invasive alien species. Given that human health and welfare are dependent on the condition and stability of local ecosystems (*2*), we urge governments and aiding parties responding to catastrophes to recognize and mitigate potential risks.

Invasive alien species are among the greatest threats to global biodiversity and ecosystem health (*3*), often causing multimillion-dollar impacts (*4*), especially on islands (*5*). In countries recovering from disaster, invasive alien species may prove impossible to eradicate (*6*) and can create long-term ecological and financial costs in addition to already serious humanitarian challenges (*7*). Dominica, for example, faces impacts from several vertebrate invasive alien species, including *Iguana iguana* and *Osteopilus septentrionalis*, that arrived with post-hurricane aid shipments in 2017 (*8*).

Some countries, such as New Zealand, have the resources to implement effective bioscreening measures (*9*). Other countries have insufficient protocols or lack the funding, administrative capacity, or political will to manage them effectively. Regardless of a country's usual procedures, in disaster scenarios where rapidly providing aid supplies to those in need is prioritized, biosecurity-screening capacity in affected regions may be reduced or nonexistent. Ideally, aiding parties would make biosecurity screening an integral part of their process to prevent unintended long-term consequences of aid delivery. However, relief comes from both governmental and nongovernmental sources, from locations near and far, with a range of organizational structures and financial resources. Moreover, the type and location of the disaster can affect the functionality of ports and other entry points, necessitating tailored strategies for each event.

To address the need for robust biosecurity during disasters, especially among island nations, scientists and policy-makers should work to establish regional strategies supported by legislation, including who should provide information to aiding parties, shipping logistics, and processes to ensure biosecurity upon arrival. Global financial assistance would help support development, training, and implementation. Given the frequency of natural disasters, implementing biosecurity and invasive alien species surveillance programs, in addition to prioritizing rapid aid to people in need, are critically important to protect both humans and wildlife from the spread and impacts of invasive alien species.

Matthijs P. van den Burg[1], Jennifer C. Daltry[2,3], Baptiste Angin[4], Erik Boman[5], Jeanelle L. K. Brisbane[1,6], Katrina Collins[7], Jane E. Haakonsson[1,8], Arica Hill[9], Julia A. Horrocks[10], Farah Mukhida[11], FitzGerald Providence[12], Karl Questel[13], Naitram Ramnanan[14], Sophia Steele[15], Isabel M. Vique Bosquet[1,3], Charles R. Knapp[1,16]*

[1]International Union for Conservation of Nature Species Survival Commission Iguana Specialist Group, Gland, Switzerland. [2]Global Wildlife Conservation, Austin, TX 78746, USA. [3]Fauna & Flora International, Cambridge, UK. [4]Ardops Environment, Les Abymes, Guadeloupe. [5]St. Eustatius National Parks Foundation, Caribbean Netherlands. [6]WildDominique, Roseau, Commonwealth of Dominica. [7]Union Island Environmental Attackers, Union Island, St. Vincent & the Grenadines. [8]Department of Environment, Cayman Islands government, Grand Cayman, Cayman Islands. [9]Environmental Awareness Group, St. John's, Antigua. [10]University of the West Indies, Cave Hill Campus, Barbados. [11]Anguilla National Trust, Anguilla, British West Indies. [12]St. Vincent & the Grenadines Forestry Department, Kingstown, St. Vincent & the Grenadines. [13]Agence Territoriale de L'Environnement de Saint-Barthélemy, Gustavia, Saint Barthélemy. [14]CABI, St. Augustine,

Trinidad and Tobago. [15]Fauna & Flora International Corp., St. John's, Antigua. [16]Daniel P. Haerther Center for Conservation and Research, John G. Shedd Aquarium, Chicago, IL 60605, USA.
*Corresponding author.
Email: cknapp@sheddaquarium.org

REFERENCES AND NOTES

1. "Volcanic eruption leaves 'entire population' of Saint Vincent without clean water," *UN News* (2021).
2. S. Whitmee *et al*., *The Lancet* **386**, 1973 (2015).
3. F.W. Allendorf *et al*., *Trends Ecol. Evol.* **16**, 613 (2001).
4. C. Diagne *et al*., *Nature* 10.1038/s41586-021-03405-6 (2021).
5. J. Russel *et al*., *Environ. Conserv.* **44**, 359 (2017).
6. C. R. Knapp *et al*., *Anim. Conserv.* doi.org/10.1111/acv.12660 (2020).
7. P. Pyšek *et al*., *Biol. Rev*. **95**, 1511 (2020).
8. M. P. Van den Burg, J. L. K. Brisbane, C. R. Knapp, *Biol. Invasions* **22**, 195 (2020).
9. D. Simberloff, *Ecol. Eng*. **65**, 112 (2014).

10.1126/science.abj0449

Inclusion through part-time science

Over the past year, the science community has been transformed by alternative work arrangements. Building on this foundation to increase flexible job offerings in the post-pandemic future would promote workforce diversity, including people who sometimes struggle to balance work with other life priorities.

For instance, pursuing a productive scientific career while raising young children presents a daunting challenge. A disproportionate burden is typically placed on mothers due to unequal housework and child care allocation (*1*). In the United States, insufficient parental leave exacerbates the problem, usually requiring full-time workers to find alternative child care arrangements within 3 months of childbirth, relinquishing valuable bonding time with their infant and often introducing additional financial costs. Even with good parental leave benefits, time off could derail a principal investigator's career by leading to gaps in funding and mentoring, or it could perpetuate inequality if women make use of it and men do not. As a result, many new mothers and, to a lesser degree, new fathers leave the full-time science, technology, engineering, and mathematics workforce (*2*). The underrepresentation of women with children in the postdoctoral workforce and the dearth of women pursuing principal investigator positions could be improved by offering more part-time options in addition to affordable child care (*3*).

Increased availability of part-time early career scientist positions could also make balancing career and other responsibilities more feasible for other caregivers, people with health limitations, and anyone else who requires more flexibility. Currently, tenure-track faculty positions and part-time positions are mutually exclusive at some institutions. Furthermore, many postdoctoral fellowships [such as NIH F32 (*4*)] and career development awards [such as NIH K99/R00 (*5*)] that can help accelerate junior scientists' careers are only awarded to full-time applicants. Part-time researchers are uncommon and are often perceived to be less dedicated (*6*) than their full-time counterparts, putting them at a disadvantage when judged by their peers for funding and advancement. More part-time opportunities in the scientific workforce would enrich scientific culture, adding valuable new perspectives and enhancing diversity. The next time you are recruiting for an open position, consider explicitly welcoming part-time applicants—the quality of the resulting applicant pool may surprise you.

Elizabeth Theusch
Department of Pediatrics, University of California San Francisco, Oakland, CA 94609, USA. Email: elizabeth.theusch@ucsf.edu

REFERENCES AND NOTES

1. L. Schiebinger, S. K. Gilmartin, *Academe* **96**, 39 (2010).
2. E. A. Cech, M. Blair-Loy, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4182 (2019).
3. E. D. Martinez *et al*., *EMBO Rep.* **8**, 977 (2007).
4. Ruth L. Kirschstein National Research Service Award (NRSA) Individual Postdoctoral Fellowship (Parent F32) (NIH, 2020); https://grants.nih.gov/grants/guide/pa-files/PA-21-048.html.
5. NIH Pathway to Independence Award (Parent K99/R00 Independent Clinical Trial Not Allowed) (NIH, 2020); https://grants.nih.gov/grants/guide/pa-files/PA-20-188.html.
6. L. S. Malisheski, "Part-time science in perspective," *Science* 10.1126/science.caredit.a0700177 (2007).

10.1126/science.abi9023

Integrate US science and diplomacy

As evidenced by the COVID-19 pandemic, the ramifications of not adequately integrating science practitioners in the ranks of US diplomatic leadership can be severe. Since President Biden's inauguration, the new administration has rightly placed science as a key discipline to drive the trajectory of both US domestic and foreign policy (*1*). To implement this goal, the Biden administration should appoint more scientists and engineers to the senior ranks of the State Department.

Achieving broad human progress goals, as articulated in many of the UN's Sustainable Development Goals (*2*), will require that diplomacy and policy-making be informed by science. If the United States follows through with its plans to center science in decision-making, it can provide vital leadership in addressing emerging global challenges. Some of the most effective international agreements were forged through science-based policy crafting. The Antarctic Treaty System of 1959 (*3*) enables research at the South Pole. The Outer Space Treaty of 1967 (*4*) facilitates the work at the International Space Station. The Montreal Protocol of 1987 (*5*) stymied the depletion of the ozone layer. The success of these treaties demonstrates that international cooperation informed by science is already a model for future progress.

Harnessing the talents of the United States' leading scientists and engineers is key to the advancement of its foreign policy and the resolution of national security priorities. To accomplish this goal, the Biden-Harris administration should create a new position of undersecretary for scientific affairs and reestablish the science affairs officer within the State Department. Scientists and engineers should also be installed in the Foreign Service and at the National Security Council. As the Biden-Harris transition proclaimed: The people have chosen science (*6*). It is the responsibility of our diplomatic community and our partners in science and technology to make good on that choice.

Amrita Banerjee[1,2], Lyndsey Gray[3,2], W. Robert Pearson[4], Benjamin L. Schmitt[5,4]*, Katherine Shield[6,2], Giovanni Zanalda[4]
[1]AAAS Congressional Science and Engineering Fellowship, Washington, DC 20004, USA. [2]National Science Policy Network, San Francisco, CA 94104, USA. [3]Colorado State University, Fort Collins, CO 80523, USA. [4]Duke University Center for International and Global Studies, Durham, NC 27708, USA. [5]Harvard-Smithsonian Center for Astrophysics, Harvard University, Cambridge, MA 02138, USA. [6]University of California, Berkeley, Berkeley, CA 94720, USA.
*Corresponding Author: benjamin.schmitt@cfa.harvard.edu

REFERENCES AND NOTES

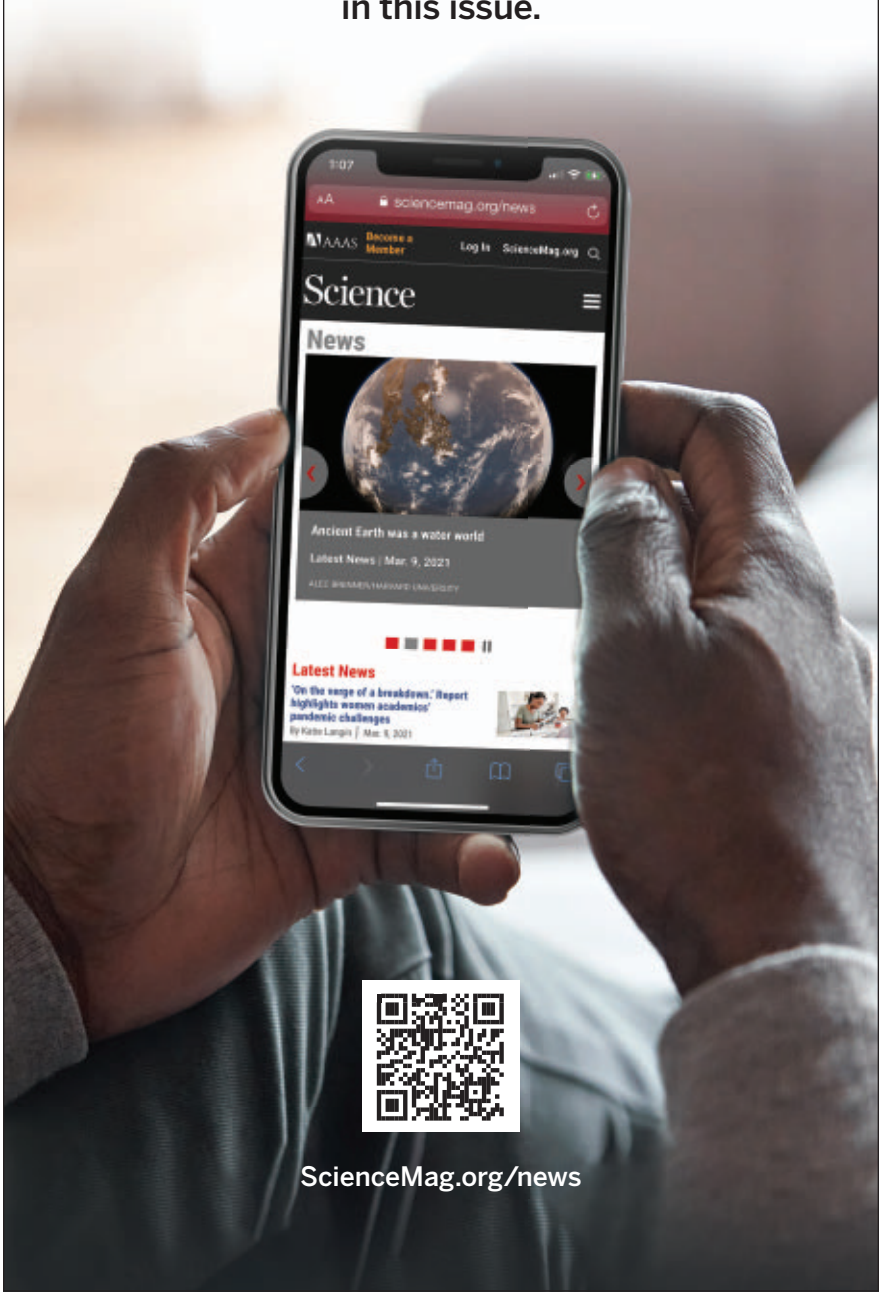
1. J. Daley, "Biden elevates science in week one actions," *Scientific American* (2021).
2. UN Department of Economic and Social Affairs, "The 17 Goals" (https://sdgs.un.org/goals).
3. Scientific Committee on Antarctic Research, "The Antarctic Treaty System" (www.scar.org/policy/antarctic-treaty-system/).
4. UN Office for Outer Space Affairs, "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies" (www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/introouterspacetreaty.html).
5. US Department of State, "The Montreal Protocol on Substances that Deplete the Ozone Layer" (www.state.gov/key-topics-office-of-environmental-quality-and-transboundary-issues/the-montreal-protocol-on-substances-that-deplete-the-ozone-layer/).
6. S. Hilgartner, J. B. Hurlbut, S. Jasanoff, *Science* **371**, 893 (2021).

10.1126/science.abi8644

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

Open-Design Upright Microscope

- Optional motorized or fixed XY stage, or motorized translator
- Open-design microscope with motorized focus
- Quickly configurable based on experimental needs
- Optimized *In vivo* and *In vitro* experimentation on one setup
- Uses standard Olympus objectives
- Free Multi-Link™ software coordinates movement with MPC-200
- OCC or DIC transmitted light (LED)
- Epi-fluorescent imaging



BOB™

The Sutter BOB – designed to eliminate the conventional microscope frame – is a versatile, open-design upright microscope platform ideal for slice electrophysiology, widefield fluorescent imaging, two-photon imaging, photostimulation and new techniques just being developed!

SUTTER INSTRUMENT

PHONE: +1.415.883.0128
FAX: +1.415.883.0572
EMAIL: INFO@SUTTER.COM
WWW.SUTTER.COM

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



MIGRATION

High fliers

Migrating from hemisphere to hemisphere is a global strategy for many bird species. Despite allowing birds to track productivity, these long-distance movements bring them in contact with inhospitable regions such as deserts and oceans. Sjöberg *et al*. used geolocators to monitor flight in great reed warblers (*Acrocephalus arundinaceus*) and found that when over these types of regions, this normally nocturnal migrating species flew both day and night. During the day, the birds increased the altitudes at which they flew, rising to more than 5000 meters. Such behavior may allow them to avoid heat stress or other daytime threats during migration. —SNV *Science*, this issue p. 646

Great reed warblers, like this one shown here singing at sunset, fly at extreme altitudes during the daytime as they migrate nonstop across the Sahara desert.

SOLAR CELLS

Tougher solar cell interfaces

The low formation energies of the active layers in perovskite solar cells lead to low-toughness materials that are compliant and soft, which limits their interface stability and long-term reliability. Dai *et al*. show that treatment with iodine-terminated self-assembled monolayers that react with surface hydroxyl groups (which ultimately creates unwanted charge traps and voids) leads to a 50% increase of adhesion toughness between the electron transport layer and a mixed-composition perovskite thin film. The projected point at which 80% of the operating efficiency in perovskite solar cells was still retained increased from ~700 to 4000 hours for 1-sun exposure with continuous maximum power point tracking. —PDS

Science, this issue p. 618

QUANTUM SYSTEMS

Quantum entanglement goes large

Quantum entanglement occurs when two separate entities become strongly linked in a way that cannot be explained by classical physics; it is a powerful resource in quantum communication protocols and advanced technologies that aim to exploit the enhanced capabilities of quantum systems. To date, entanglement has generally been limited to microscopic quantum units such as pairs or multiples of single ions, atoms, photons, and so on. Kotler *et al*. and Mercier de Lépinay *et al*. demonstrate the ability to extend quantum entanglement to massive macroscopic systems (see the Perspective by Lau and Clerk). Entanglement of two mechanical oscillators on such a large length and mass scale is expected to find widespread use in both applications and fundamental physics to probe the boundary between the classical and quantum worlds. —ISO

Science, this issue p. 622, p. 625; see also p. 570

CORONAVIRUS

A large-scale screen to target SARS-CoV-2

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genome is initially expressed as two large polyproteins. Its main protease, M^{pro}, is essential to yield functional viral proteins, making it a key drug target. Günther *et al*. used x-ray crystallography to screen more than 5000 compounds that are either approved drugs or drugs in clinical trials. The screen identified 37 compounds that bind to M^{pro}. High-resolution structures showed that most compounds bind at the active site but also revealed two allosteric sites where binding of

PHOTO: BLICKWINKEL/ALAMY STOCK PHOTO

a drug causes conformational changes that affect the active site. In cell-based assays, seven compounds had antiviral activity without toxicity. The most potent, calpeptin, binds covalently in the active site, whereas the second most potent, pelitinib, binds at an allosteric site. —VV

Science, this issue p. 642

PALEONTOLOGY

Revealing behavioral secrets in extinct species

Extinct species had complex behaviors, just like modern species, but fossils generally reveal little of these details. New approaches that allow for the study of structures that relate directly to behavior are greatly improving our understanding of the lifestyles of extinct animals (see the Perspective by Witmer). Hanson *et al*. looked at three-dimensional scans of archosauromorph inner ears and found clear patterns relating these bones to complex movement, including flight. Choiniere *et al*. looked at inner ears and scleral eye rings and found a clear emergence of patterns relating to nocturnality in early theropod evolution. Together, these papers reveal behavioral complexity and evolutionary patterns in these groups. —SNV

Science, this issue p. 601, p. 610; see also p. 575



An endocast of the *Byronosaurus jaffei* inner ear contains features indicative of complex behaviors such as flight and high-frequency vocalizations.

CREDITS (FROM LEFT): HANSON *ET AL.*; RAKAN ZAHAWI

CANCER IMAGING

Through a shrimp eye brightly

A camera for intraoperative imaging of tumors could improve surgical outcomes, but some imaging technologies have been difficult to translate to clinical practice. Blair *et al*. designed an imaging system based on the eye of the mantis shrimp. This system detected multiple near-infrared fluorescent signals simultaneously and was tested in a mouse model of human prostate cancer. In support of clinical feasibility, the authors show that the sensor can detect fluorescently labeled sentinel lymph nodes in patients with breast cancer who are undergoing surgical resection. This bioinspired imaging sensor could offer a flexible tool for image-guided surgical removal of tumors. —OMS

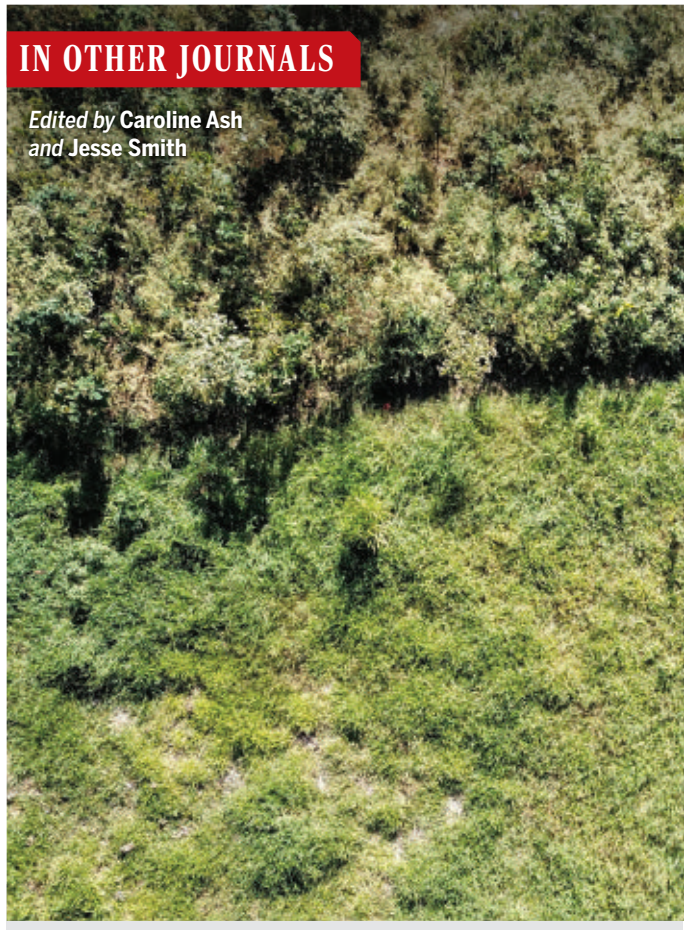
Sci. Transl. Med. **13**, eaaw7067 (2021).

ARCHAEOLOGY

Human impact on landscapes in early Africa

Although fire has been used by many cultures to improve the availability of subsistence resources, the antiquity of this practice is poorly understood. Thompson *et al*. investigated the northern basin of Lake Malawi in southern-central Africa using archaeological, geomorphological, and palaeoecological data and reconstructed the environmental context of the Middle Stone Age hunters and gatherers who lived in this region around 92,000 years ago. These peoples used fire to clear forests and thin understory growth. The signature of these fires differs from those created by natural phenomena and signals an abrupt transformation of regional ecology when compared with earlier periods. This record of anthropogenic burning is much earlier than similar practices in prehistory. —MSA

Sci. Adv. 10.1126/sciadv.abf9776 (2021).



IN OTHER JOURNALS

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

RESTORATION ECOLOGY

Fill-up for the forest

The restoration of degraded and disturbed habitats has become a major focus of applied ecology because of the benefits it brings to ecosystem functioning and biodiversity. Cole and Zahawi show how nutrients from agricultural waste can contribute to the recovery of tropical forest on postagricultural land. In an experiment in Costa Rica, a 0.5-meter-deep layer of coffee pulp applied to the land raised nutrient levels and inhibited the growth of ground cover, especially of pasture grasses. This input promoted a succession of forest trees and shrubs, which reached substantially greater biomass and canopy height after 2 years than vegetation in a control plot. —AMS *Ecol. Solut. Evid.* **2**, e12054 (2021).

Forest restoration in Costa Rica is accelerated (top section of image) by the application of coffee-processing waste as a deep mulch.

SUSTAINABILITY

Pollinator-focused solar arrays

Although the power generated from the sun is "green," the deployment of large receiving arrays is often less so. At their worst, these are large swaths of solar panels that sit above dry and dead fields. Large-scale solar production can be improved if panels are combined with native or other flowering plants. Graham *et al*. show that such "agrivoltaic" combinations provide a variety of microclimates with different light, temperature, and moisture



X-ray image of a female mole rat revealing skeletal changes as she transitions into a queen

DEVELOPMENT

Fragile queen mole rats

Honey bees are social animals, with populations consisting of castes of siblings with defined divisions of labor and matching phenotypes. Drones, queens, and workers have specific jobs within the hive for reproduction, rearing the brood, or gathering pollen and nectar. The Damaraland mole rat (*Fukomys damarensis*) is similarly social. Groups of mole rats consist of a single breeding pair and their offspring, which divide the labor of locating food, burrowing, nest building, and caring for offspring. Knowing that, similar to honey bees, mole rats show phenotypic differences depending on caste, Johnston *et al*. investigated the molecular basis for these transformations. If a female mole rat is experimentally induced to become a breeding "queen," her vertebral column lengthens. The authors found that transitions in breeding status were accompanied by skeletal gene-regulatory changes that endowed major skeletal remodeling. In addition to a lengthening of the lumbar vertebrae, gene expression analysis indicated increased bone resorption and changes in estrogen levels in queens that reduced long bone growth and femur density. These morphological costs are thus a trade-off between fertility and bone fragility. —BAP *eLife* **10**, e65760 (2021).

levels and thus plant composition. The authors found that 65 species of native pollinators visited flowers in full-sun and partially shaded plots within a solar farm. Moreover, the partially shaded plots extended bloom time, which benefited late-season foragers in water-compromised regions and seasons. Combining solar panels with plants can thus make a planet-friendly energy source even more friendly. —SNV

Sci. Rep. **11**, 7452 (2021).

CELL BIOLOGY

Signals of stress

Senescence is a permanent arrest to cell cycling induced by stress or damage. It is reinforced in tissues by the senescence-associated secretory phenotype (SASP) in which secretion of various mediators maintains the proliferative arrest. If senescent cells (called senolysis) could be therapeutically targeted to remove them from tissues, function might be restored to organs compromised by age-associated pathology. Some senolytic drugs have already entered clinical trials, but biomarkers are needed to measure their efficacy and selectivity. Wiley *et al*. show that senescent cells synthesize oxylipins, which are bioactive lipids that reinforce senescence associated with SASP. A specific oxylipin is released upon senolysis, and so detection of this in extracellular compartments could be the biomarker that is needed. —GKA

Cell Metab. 10.1016/j.cmet.2021.03.008 (2021).

EDUCATION AND CULTURE

Policy changing culture

The introduction of pension plans in cultures where adult daughters (matrilocal) or sons (patrilocal) traditionally live with their aging parents led parents, now less dependent upon a child for future support, to invest less in that child's education. Bau shows that a program in Indonesia led to differentially reduced educational investments in females, but not males, from matrilocal ethnic groups and that the women were less likely to practice matrilocality as adults. Similar effects were found for a program in Ghana that led to differential impacts on sons, but not daughters, in patrilocal groups. —BW

Amer. Econ. Rev. www.aeaweb.org/articles?id=10.1257/aer.20190098 (2021).

MICROPLASTICS

Plastic dust in the wind

Plastic pollution is ubiquitous throughout the environment, the endmember being microplastics generated from the breakdown of larger, more primary pieces. Brahney *et al*. investigated the sources of atmospheric microplastics using a combination of observations and atmospheric transport models. They found that the majority of microplastics in the air are derived from the fragmentation of plastics on roadways, with much smaller fractions coming from oceans and agricultural soils. This finding underscores the importance of controlling primary plastic pollution. —HJS

Proc. Natl. Acad. Sci. U.S.A. **118**, e2020719118 (2021).

IMAGING

Extending the reach of LIDAR

Light detection and ranging (LIDAR) is an invaluable tool used in remote sensing to map out the detailed three-dimensional features of terrain. Based on time-of-flight measurements of emitted, reflected, and detected single photons, the geographic scenery can be computationally reconstructed with high resolution. Photon scattering, high background noise, and losses in the atmosphere typically limit the range of the technique to a few tens of kilometers. Li *et al*. developed high-efficiency optical components and an efficient noise-suppression method to extend the range of LIDAR to more than 200 kilometers. Such a long reach could be used to complement traditional imaging techniques for remote-sensing applications with low-Earth-orbit satellites. —ISO

Optica **8**, 344 (2021).

JOHNSTON ET AL., ELIFE 10, E65760 (2021)

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

HYPOTHESIS

Rethinking insulin in obesity

The causes of obesity remain unclear, and several models have been proposed to help prevent or reverse weight gain. The carbohydrate-insulin model centers on the role of insulin in driving adiposity and limiting satiety in response to consumed carbohydrates. It has thus been proposed that low-carbohydrate, high-fat diets (a ketogenic diet) could be used to limit the postprandial effects of insulin. In a Perspective, Speakman and Hall discuss the accumulating evidence that this role of insulin may not be the key factor in regulating weight gain. Instead, they propose that the basal level of insulin in response to the balance of energy intake and expenditure, rather than diet composition per se, is more important. —GKA

Science, this issue p. 577

PRIMATE EVOLUTION

A distinctive ancestor

There has been much focus on the evolution of primates and especially where and how humans diverged in this process. It has often been suggested that the last common ancestor between humans and other apes, especially our closest relative, the chimpanzee, was ape- or chimp-like. Almécija *et al*. review this area and conclude that the morphology of fossil apes was varied and that it is likely that the last shared ape ancestor had its own set of traits, different from those of modern humans and modern apes, both of which have been undergoing separate suites of selection pressures. —SNV

Science, this issue p. 587

BACTERIAL PHYLOGENY

Reconstructing ancestral bacteria

The origin of the eubacteria and phylogenetic relationships between subgroups have been difficult to resolve. Applying a phylogenetic analysis and recent computational methods to the expanded diversity of bacterial sequences from metagenomic analyses, Coleman *et al*. infer the root of the eubacterial tree (see the Perspective by Katz). The root was determined without using the Archaea as an outgroup, to avoid the possibility of a false result due to long branch attraction. This method places the eubacterial root in the neighborhood of Fusobacteriota. Using this information, the authors reconstructed the eubacterial ancestor, identifying that this organism likely had a double-membrane cell envelope, flagellum-mediated motility, antiphage defense mechanisms, and diverse metabolic pathways. —LMZ

Science, this issue p. 588; see also p. 574

GENETICS

Environmental impacts on gene networks

A phenotype can be affected by genes interacting with other genes, the environment, or both other genes and the environment (a differential interaction). To better understand how these interactions function in yeast, Costanzo *et al*. mapped gene-gene interactions using single- and double-mutant deletions and temperature-sensitive alleles under 14 environmental conditions. Many deleted or temperature-sensitive nonessential genes affected yeast fitness both positively and negatively under at least one of the environmental conditions tested. In these cases, up to 24% of yeast genes were affected. A minority of these differential interactions point to previously unknown genetic connections across functional networks, informing on how genetic architecture responds to environmental variation. —LMZ

Science, this issue p. 589

PALEOGENOMICS

The value of dirty DNA

Environmental DNA can identify the presence of species, even from the distant past. Surveying three cave sites in western Europe and southern Siberia, Vernot *et al*. identified nuclear DNA and confirmed that it is from the close relatives of anatomically modern humans—Neanderthal and Denisovan individuals. A phylogenetic analysis and modeling show that the DNA in sediment samples from several layers corresponds to previously studied skeletal remains. These results demonstrate that environmental data can be applied to study the population genetics of the extinct Neanderthal and Denisovan lineages, identifying a turnover of Neanderthal populations ~100,000 years ago. —LMZ

Science, this issue p. 590

NEURODEVELOPMENT

Development of the human striatum revealed

Deep in the brain, the striatum receives and coordinates inputs from other parts of the brain. Bocchi *et al*. surveyed molecular features as the striatum develops in the human brain. Single-cell surveys of long intergenic noncoding RNAs revealed a progenitor for medium spiny neurons and provide insight into evolutionary divergence of this critical part of the brain. —PJH

Science, this issue p. 591

SEX DETERMINATION

Mystery solved?

Chromosomal sex determination arises when an autosomal locus acquires a sex-determining function. In some taxa, this process occurs often. The XY system in mammals, however, has been evolutionarily stable across a wide array of species. Fifty years ago, a variation on this norm was described in the creeping vole (*Microtus oregoni*), but the details have remained mostly unknown. Couger *et al*. sequenced the sex chromosomes in this species and found that the Y chromosome has been lost, the male-determining chromosome is a second X that is largely homologous to the female X, and both the maternally inherited and male-specific sex chromosomes carry vestiges of the ancestral Y. —SNV

Science, this issue p. 592

FERROELECTRICS

A role for vacancies

Hafnia-based materials are of interest because of their potential use in microelectronic components. Hafnia-oxide is a ferroelectric material, but whether the polarization switching comes from the polar crystal phases or the migration of oxygen vacancies has remained an open question. Nukala *et al*. attempted to resolve this controversy by conducting electron microscopy during the operation of a hafnium zirconium oxide capacitor. The authors found that vacancy migration is intertwined with the ferroelectric switching, which has implications for the use of these materials in a range of microelectronic applications. —BG

Science, this issue p. 630

MATERIALS SCIENCE

Reversible fiber fusion and fission

Materials that can cycle between states are of interest for actuators, soft robotics, or recoverable membranes for separations. Chang *et al*. show that a collection of graphene oxide fibers can fuse into a single stronger fiber

upon immersion in a solvent, extraction, and drying under tension (see the Perspective by Cruz-Silva and Elías). The geometrical deformation of the fibers during drying and swelling plays an important role in the reversible cycles, with a large volume change between the dried and swelled fibers. Moreover, fibers made from polymers, glass, metal, or silk can be given these abilities when coated with a micron-sized layer of graphene oxide. —MSL

Science, this issue p. 614; see also p. 573

CORONAVIRUS

The Slovakian test case

Toward the end of 2020, Slovakia decided that it would test and then isolate positive severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) cases among its entire population of ~5.5 million, and more than 50,000 positive cases were found during a rapid antigen testing campaign. Pavelka *et al.* analyzed the data and found that in 41 counties before and after the two rounds of testing, infection prevalence declined by about 80% (see the Perspective by García-Fiñana and Buchan). They also used the data to test a microsimulation model for one county. Quarantine of the whole household after a positive test was essential to achieving a large reduction in prevalence. Since Autumn 2020, transmission in Slovakia has rebounded, despite other interventions, because high-intensity testing was not sustainable. —CA

Science, this issue p. 635; see also p. 571

INFLAMMATION

DUBbing new inflammasome inhibitors

Inflammasome assembly and activation leading to mature interleukin-1β (IL-1β) release is dysregulated in a wide range of inflammatory diseases. Optimal activation of the NLRP3 inflammasome, a protein complex important for IL-1β release, requires the activity of BRISC, a deubiquitinating enzyme (DUB) complex composed of four protein subunits including the metalloprotease BRCC3. Ren *et al*. demonstrate that the compound thiolutin, a zinc chelator that inhibits BRCC3, can potently inhibit NLRP3 deubiquitination and inflammasome activation. Thiolutin was effective at inhibiting NLRP3 activation and preventing IL-1β production in multiple mouse models of inflammatory disease, including a model of diet-induced nonalcoholic fatty liver disease. Holomycin, a derivative of thiolutin with reduced toxicity, was also effective at inhibiting NLRP3, paving the way for the development of agents that selectively target deubiquitination of NLRP3 to regulate its activity. —CO

Sci. Immunol. **6**, eabe2933 (2021).

NEUROSCIENCE

Stabilized by protein interactions

Neurons maintain stable excitability despite prolonged changes in synaptic input by adjusting the strength of connections between neurons. Heavner *et al*. found that in mice, associations between proteins in a postsynaptic protein interaction network were differentially altered by prolonged increases compared with prolonged decreases in activity. Some of these changes did not occur in mice lacking either *Homer1* or *Shank3B*, genes that are mutated in some patients with autism spectrum disorder. Thus, these scaffolding proteins may serve as structural hubs for synaptic plasticity. —LKF

Sci. Signal. **14**, eabd7325 (2021).

REVIEW SUMMARY

PRIMATE EVOLUTION

Fossil apes and human evolution

Sergio Almécija*, Ashley S. Hammond, Nathan E. Thompson, Kelsey D. Pugh, Salvador Moyà-Solà, David M. Alba

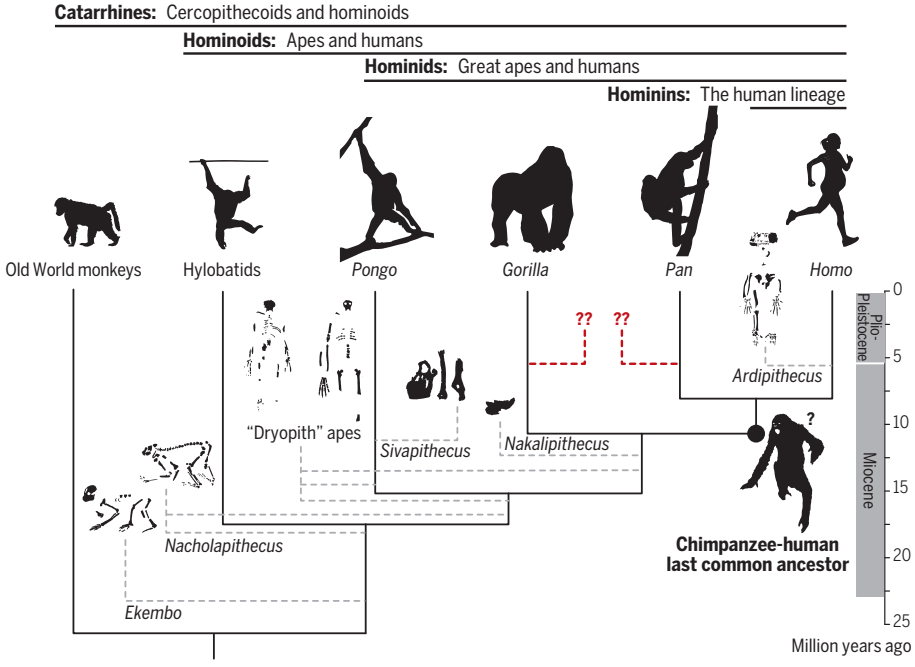
BACKGROUND: Ever since the writings of Darwin and Huxley, humans' place in nature relative to apes (nonhuman hominoids) and the geographic origins of the human lineage (hominins) have been heavily debated. Humans diverged from apes [specifically, the chimpanzee lineage (*Pan*)] at some point between ~9.3 million and ~6.5 million years ago (Ma), and habitual bipedalism evolved early in hominins (accompanied by enhanced manipulation and, later on, cognition). To understand the selective pressures surrounding hominin origins, it is necessary to reconstruct the morphology, behavior, and environment of the *Pan-Homo* last common ancestor (LCA). "Top-down" approaches have relied on living apes (especially chimpanzees) to reconstruct hominin origins. However, "bottom-up" perspectives from the fossil record suggest that modern hominoids represent a decimated and biased sample of a larger ancient radiation and present alternative possibilities for the morphology and geography of the *Pan-Homo* LCA. Reconciling these two views remains at the core of the human origins problem.

ADVANCES: There is no consensus on the phylogenetic positions of the diverse and widely distributed Miocene apes. Besides their fragmentary record, disagreements are due to the complexity of interpreting fossil morphologies that present mosaics of primitive and derived features, likely because of parallel evolution (i.e., homoplasy). This has led some authors to exclude known Miocene apes from the modern hominoid radiation. However, most researchers identify some fossil apes as either stem or crown members of the hominid clade [i.e., preceding the divergence between orangutans (pongines) and African great apes and humans (hominines), or as a part of the modern great ape radiation]. European Miocene apes have prominently figured in discussions about the geographic origin of hominines. "Kenyapith" apes dispersed from Africa into Eurasia ~16 to 14 Ma, and some of them likely gave rise to the European "dryopith" apes and the Asian pongines before 12.5 Ma. Some authors interpret dryopiths as stem hominines and support their back-to-Africa dispersal in the latest Miocene, subsequently evolving into modern African apes and hominins. Others interpret dryopiths as broadly ancestral to hominids or an evolutionary dead end.

Increased habitat fragmentation during the late Miocene in Africa might explain the evolution of African ape knuckle walking and hominin bipedalism from an orthograde arboreal ancestor. Bipedalism might have allowed humans to escape the great ape "specialization trap"—an adaptive feedback loop between diet, specialized arboreal locomotion, cognition, and life history. However, understanding the different selection pressures that underlie knuckle walking and bipedalism is hindered by locomotor uncertainties about the *Pan-Homo* LCA and its Miocene forebears. In turn, the functional interpretation of Miocene ape mosaic morphologies is challenging because it depends on the relevance of primitive features. Furthermore, adaptive complexes can be co-opted to perform new functions during evolution. For instance, features that are functionally related to quadrupedalism or orthogrady can be misinterpreted as bipedal adaptations. Miocene apes show that the orthograde body plan, which predates below-branch suspension, is likely an adaptation for vertical climbing that was subsequently co-opted for other orthograde behaviors, including habitual bipedalism.

OUTLOOK: Future research efforts on hominin origins should focus on (i) fieldwork in unexplored areas where Miocene apes have yet to be found, (ii) methodological advances in morphology-based phylogenetics and paleoproteomics to retrieve molecular data beyond ancient DNA limits, and (iii) modeling driven by experimental data that integrates morphological and biomechanical information, to test locomotor inferences for extinct taxa. It is also imperative to stop assigning a starring role to each new fossil discovery to fit evolutionary scenarios that are not based on testable hypotheses.

Early hominins likely originated in Africa from a Miocene LCA that does not match any living ape (e.g., it might not have been adapted specifically for suspension or knuckle walking). Despite phylogenetic uncertainties, fossil apes remain essential to reconstruct the "starting point" from which humans and chimpanzees evolved. ■



The evolutionary history of apes and humans is largely incomplete. Whereas the phylogenetic relationships among living species can be retrieved using genetic data, the position of most extinct species remains contentious. Surprisingly, complete-enough fossils that can be attributed to the gorilla and chimpanzee lineages remain to be discovered. Assuming different positions of available fossil apes (or ignoring them owing to uncertainty) markedly affects reconstructions of key ancestral nodes, such as that of the chimpanzee-human LCA.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: salmecija@amnh.org (S.A.)
Cite this article as S. Almécija *et al.*, *Science* 372, eabb4363 (2021). DOI: 10.1126/science.abb4363

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb4363

REVIEW

PRIMATE EVOLUTION

Fossil apes and human evolution

Sergio Almécija[1,2,3]*, Ashley S. Hammond[1,2], Nathan E. Thompson[4], Kelsey D. Pugh[1,2], Salvador Moyà-Solà[3,5,6], David M. Alba[3]

Humans diverged from apes (chimpanzees, specifically) toward the end of the Miocene ~9.3 million to 6.5 million years ago. Understanding the origins of the human lineage (hominins) requires reconstructing the morphology, behavior, and environment of the chimpanzee-human last common ancestor. Modern hominoids (that is, humans and apes) share multiple features (for example, an orthograde body plan facilitating upright positional behaviors). However, the fossil record indicates that living hominoids constitute narrow representatives of an ancient radiation of more widely distributed, diverse species, none of which exhibit the entire suite of locomotor adaptations present in the extant relatives. Hence, some modern ape similarities might have evolved in parallel in response to similar selection pressures. Current evidence suggests that hominins originated in Africa from Miocene ape ancestors unlike any living species.

In 1871, Darwin (*1*) speculated that humans originated in Africa based on the anatomical similarities with African apes (gorillas and chimpanzees) identified by Huxley (*2*). However, Darwin urged caution until more fossils became available—the European *Dryopithecus* was the only recognized fossil ape at the time (*3*). After 150 years of continuous discoveries, essential information about human origins remains elusive owing to debates surrounding the interpretation of fossil apes (Figs. 1 and 2).

Genomic data indicate that humans and chimpanzees are sister lineages ("hominins" and "panins," respectively; Box 1) that diverged from a last common ancestor (LCA) toward the end of the Miocene, at some point between ~9.3 million and ~6.5 million years ago (Ma) (*4*, *5*). All extant hominoids (apes and humans) are characterized by the lack of an external tail, high joint mobility (e.g., elbow, wrist, hip), and the possession of an "orthograde" (upright) body plan, as opposed to the more primitive, "pronograde" body plan of other anthropoids and most other mammals (Fig. 2). These body plans are associated with two different types of positional (postural and locomotor) behaviors: pronograde behaviors, taking place on nearly horizontal supports with the trunk held roughly horizontally; and orthograde (or "antipronograde") behaviors, with the torso positioned vertically (*6*, *7*). Extant ape features also include enhanced joint mobility, long forelimbs relative to hindlimbs, and (except gorillas) long hands with high-to-very-high finger curvature (*8*–*10*). The orthograde body plan is generally interpreted as a suspensory adaptation (*11*, *12*), or as an adaptation for vertical climbing subsequently co-opted for suspension (*13*).

Based on similarities between chimpanzees and gorillas, a prevalent evolutionary model argues that African apes represent "living fossils" and that knuckle-walking chimpanzees closely reflect the morphology and behavior of the *Pan-Homo* LCA—the "starting point" of human evolution (*14*, *15*). This working paradigm also postulates that modern African apes occupy the same habitats as their ancestors (*16*) (Fig. 1). This assumption is based on a classical scenario that situates hominin origins in East Africa, owing to environmental changes after the rifting of East African Rift Valley during the Miocene (*17*). For some, a chimpanzee-like *Pan-Homo* LCA could also imply that all extant ape locomotor adaptations were inherited from a modern ape-like ancestor (*18*). However, the fossil record denotes a more complex picture: Miocene apes often display mosaic morphologies, and even those interpreted as crown hominoids do not exhibit all the features present in living apes (*19*) (Fig. 3).

The *Pan*-like LCA model builds on the "East Side Story" of hominin origins (*17*), a seriously challenged scenario. First, it is grounded in the living-ape geographic distribution, which may not match that at the time of the *Pan-Homo* split (Fig. 1). Second, the model relies on an outdated account of the fossil record (from the 1980s), when the earliest known hominin (*Australopithecus afarensis*) was recorded in East Africa, and no possible fossil gorillas and chimpanzees were known (*17*). Subsequent fossil discoveries are incompatible with such a narrative: *Australopithecus* remains from Chad indicate that early hominins were living ~2500 km west of the East African Rift ~3.5 Ma (*20*). Furthermore, if *Sahelanthropus* is a hominin, it would push back the human lineage presence in north-central Africa to ~7 Ma (*21*). Moreover, continued fieldwork efforts in less explored areas have shown that hominoids lived across Afro-Arabia during the Miocene (*22*–*25*). In addition, remains of putative hominines have been found in East Africa (*26*, *27*), perhaps even in Europe (*28*, *29*). Finally, paleoenvironmental reconstructions for late Miocene apes and hominins suggest that the *Pan-Homo* LCA inhabited woodlands, not tropical rainforests (*30*–*33*).

Current debates about the transition from an ape into a bipedal hominin are centered on the morphological and locomotor reconstruction of the *Pan-Homo* LCA, as well as its paleobiogeography. Discrepancies are caused by conflicting evolutionary signals among living and fossil hominoids, indicating rampant "homoplasy" (independent evolution causing "false homology"), and are further complicated by the highly incomplete and fragmentary nature of the hominoid fossil record. This review argues that, despite the limitations, the information provided by fossil apes is essential to inform evolutionary scenarios of human origins.

Evidence as to humans' place in nature

Humans' inner primate

Since Linnaeus established modern taxonomy in 1758 (*34*) and until the 1960s, morphological similarity was the main basis for classifying organisms. Linnaeus included modern humans (*Homo sapiens*) within the order Primates, but it was not until 1863 that Huxley provided the first systematic review of differences and similarities between humans and apes (*2*). Imagining himself as a "scientific Saturnian," Huxley stated that, "The structural differences between Man and the Man-like apes certainly justify our regarding him as constituting a family apart from them; though, inasmuch as he differs less from them than they do from other families of the same order, there can be no justification for placing him in a distinct order" [(*2*), p. 104]. Huxley's work was motivated by widespread claims (e.g., Cuvier, Owen) that humans' "uniqueness" warranted their placement in a separate order. Darwin concurred with Huxley that humans should be classified in their own family within primates (*1*).

We now know that most "human features" are primitive traits inherited from primate (e.g., trichromatic stereoscopic vision, manual grasping) or earlier (e.g., five digits) ancestors (*35*). Even humans' distinctively large brains and delayed maturation are framed within a primate trend of increased encephalization and slower life history compared with other

[1]Division of Anthropology, American Museum of Natural History (AMNH), New York, NY 10024, USA. [2]New York Consortium in Evolutionary Primatology at AMNH, New York, NY 10024, USA. [3]Institut Català de Paleontologia Miquel Crusafont (ICP), Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Barcelona, Spain. [4]Department of Anatomy, New York Institute of Technology (NYIT) College of Osteopathic Medicine, Old Westbury, NY 11568, USA. [5]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain. [6]Unitat d'Antropologia Biològica, Departament de Biologia Animal, Biologia Vegetal i Ecologia, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Barcelona, Spain.
*Corresponding author. Email: salmecija@amnh.org

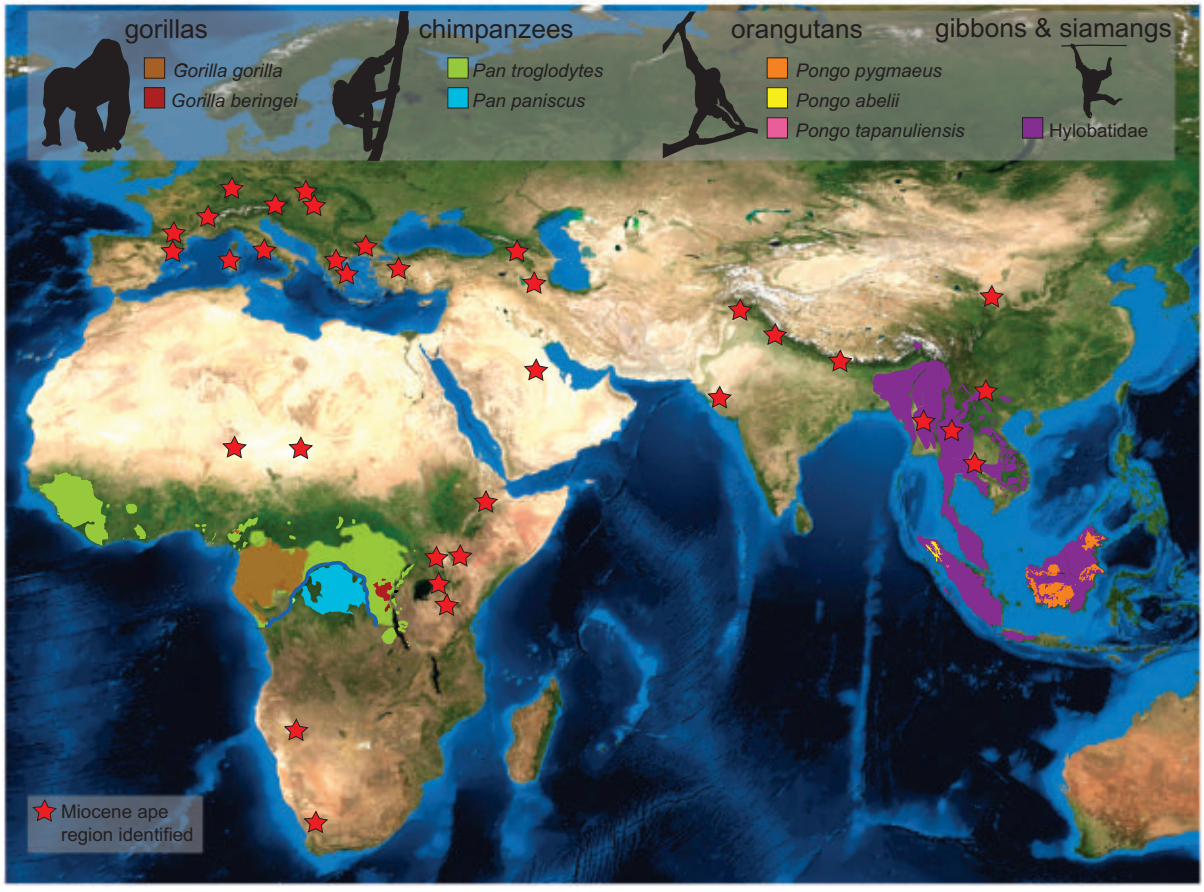


Fig. 1. Extant and fossil ape distribution. Extant apes live in (or nearby) densely forested areas around the equator in Africa and Southeast Asia. Except for the recently recognized tapanuli orangutan (which may represent a subspecies of the Sumatran orangutan), each of the three extant great ape genera presently has two geographically separated species. The Congo River (highlighted in dark blue) acts as the current barrier between common chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*). Red stars indicate regions with Miocene sediments (spanning ~23 to 5.3 Ma) where fossil apes have been uncovered. (Some regions may contain more than one site; contiguous regions are indicated with different stars if they extend over more than one political zone.) It is possible that modern great ape habitats do not represent the ancestral environments where the great ape and human clade evolved. Paleontologically, the vast majority of Africa, west of the Rift Valley, remains highly unexplored. Extant ape ranges were taken from the International Union for Conservation of Nature (IUCN Red List). Background image sources: Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS user community.

mammals (*35*, *36*). Some differences in brain size may partly reflect a neocortex enlargement related to enhanced visual and grasping abilities (*37*). Like extant great apes, humans display larger body size, larger relative brain size, a slower life-history profile, and more elaborate cognitive abilities than other primates (hylobatids included) (*36*). However, modern humans are extreme outliers in terms of delayed maturation, encephalization, advanced cognition, and manual dexterity, ultimately leading to symbolic language and technology (*38*).

Anatomically, only two adaptive complexes represent synapomorphies present in all hominins: the loss of the canine honing complex and features related to habitual bipedalism (*33*, *39*). Most anthropoids possess large and sexually dimorphic canines coupled with body size differences between males and females, reflecting levels of agonistic behavior and sociosexual structure (*40*). The fossil record indicates that there was a reduction in canine height, leading to the loss of the honing complex in early hominins (*41*). Habitual bipedalism is reflected in several traits across the body (e.g., foramen magnum position and orientation; pelvic, lower-back, and lower-limb morphology), present (or inferred) in the earliest hominins (*21*, *33*, *42*).

Darwin linked the origin of bipedalism with an adaptive complex related to freeing the hands from locomotion to use and make tools (replacing large canines), leading to a reciprocal feedback loop involving brain size, cognition, culture, and, eventually, civilization (*1*). Multiple variants in the order of these events have been advocated, with the freeing of the hands alternatively linked to tools (*43*), food acquisition and carrying (*15*), or provisioning within a monogamous social structure (*44*), to name a few. There is general agreement that canine reduction (including social structure changes), enhanced manipulative capabilities, and bipedalism were interrelated during human evolution. However, determining the order of events and their causality requires reconstructing the ape-human LCA from which hominins originated. Darwin also speculated that humans and modern African ape ancestors originated in Africa (*1*), based on the anatomical similarities identified by Huxley and his own observations that many living mammals are closely related to extinct species of the same region. However, given the limited ape fossil record at that time, he concluded that it was "useless to speculate on this subject" [(*1*), p. 199]. Using the French *Dryopithecus* to calibrate his "clock," Darwin concluded that humans likely diverged as early as the Eocene and warned against "the error of supposing that the early progenitor of the whole Simian stock, including man, was identical with, or even closely resembled, any

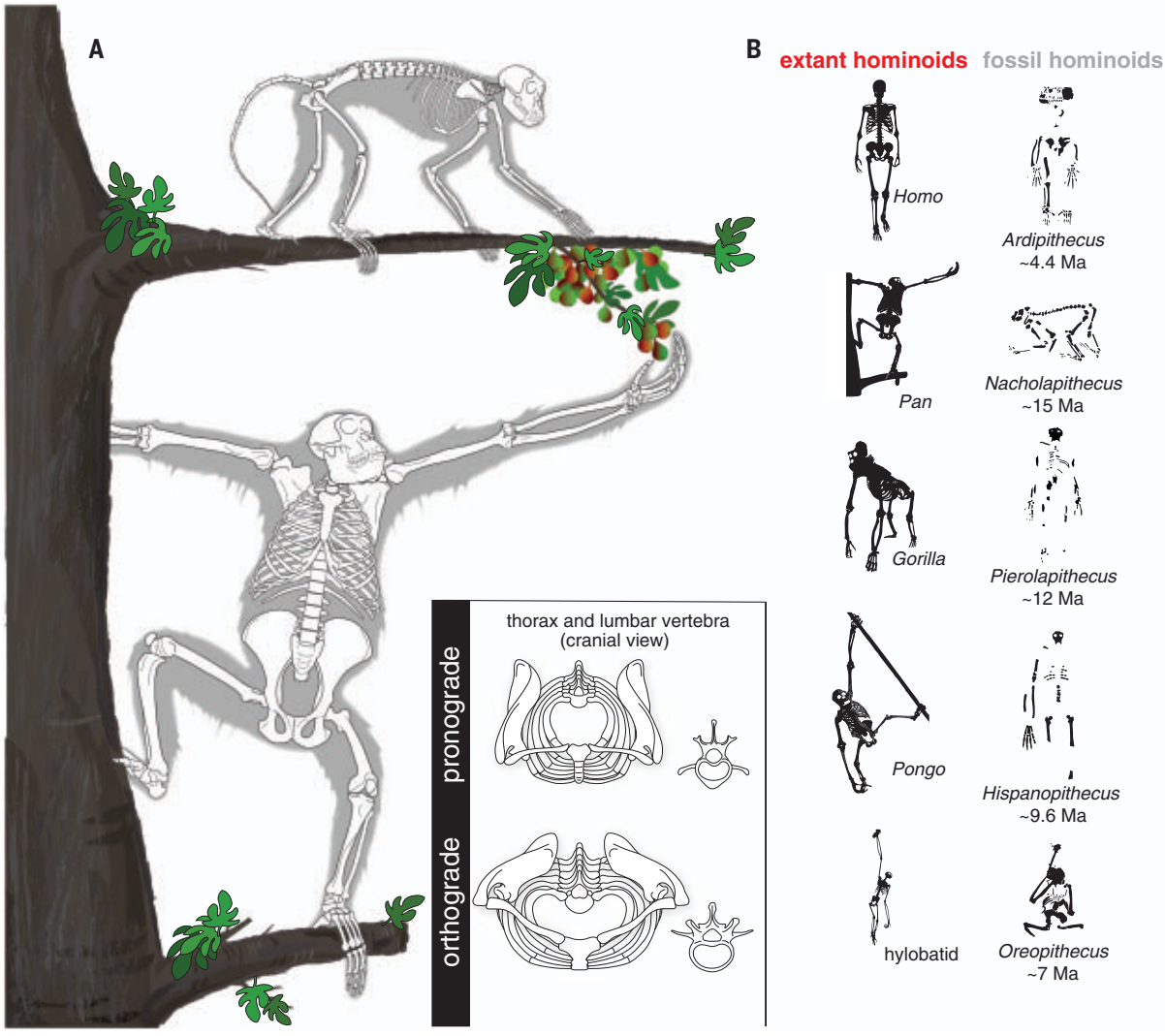


Fig. 2. Pronograde versus orthograde body plan. (**A**) Macaque (above) and chimpanzee (below) in typical postures, showing general differences between pronograde and orthograde body plan characteristics. In comparison to a pronograde monkey, the modern hominoid orthograde body plan is characterized by the lack of an external tail (the coccyx being its vestigial remnant), a ribcage that is mediolaterally broad and dorsoventrally shallow, dorsally placed scapulae that are cranially elevated and oriented, a shorter lower back, and long iliac blades. Modern hominoids have higher ranges of joint mobility, such as the full elbow extension shown here, facilitated by a short ulnar olecranon process. The inset further shows differences in lumbar vertebral anatomy, including more dorsally situated and oriented transverse processes in orthograde hominoids. (**B**) Representatives of each extant hominoid lineage (left column) show different postural variations associated with an orthograde body plan. The orthograde body plan facilitates bipedal walking in modern humans and different combinations of arboreal climbing and below-branch suspension in apes. Knuckle walking in highly terrestrial African apes is seen as a compromise positional behavior superimposed onto an orthograde ape with long forelimbs relative to the hindlimbs. Associated skeletons of fossil hominoids (right column) show that an orthograde body can be disassociated from specific adaptions for suspension (e.g., *Pierolapithecus* exhibits shorter and less curved digits than *Hispanopithecus*). Other fossil apes exhibit primitive "monkey-like" pronograde body plans with somewhat more modern ape-like forelimbs (e.g., *Nacholapithecus*). Approximate age in millions of years ago is given to representative fossils of each extinct genus: *Ardipithecus* (ARA-VP-6/500), *Nacholapithecus* (KNM-BG35250), *Pierolapithecus* (IPS21350), *Hispanopithecus* (IPS18800), and *Oreopithecus* (IGF 11778). Silhouettes of extant and fossil skeletons are shown at about the same scale.

existing ape or monkey" [(*1*), p. 199]. These ideas inaugurated a century of discussions about human's place in nature.

Reaching the "extant" consensus

Until the 1950s, the geographic origin of hominins was disputed between Africa, Asia, and Europe. After the publication of Darwin's *On the Origin of Species* (*45*), Haeckel predicted that the "missing link" (dubbed "*Pithecanthropus*," the "ape-man") would be found in Asia (*46*). This idea led to Dubois' 1891 discovery of *Homo erectus* in Indonesia (*47*). In 1925, Dart published the discovery of *Australopithecus africanus*, "the man-ape from South Africa" (*48*). However, the scientific community still focused on Europe because of the Piltdown "fossils," until they were exposed as a hoax (*49*). Asia remained a "mother continent" contender owing to the "man-like ape" *Ramapithecus*, discovered in the Indian Siwaliks (*50*).

During this time, the relationships of humans to other primates were highly contentious. Most authors advocated an ancient divergence of humans from apes (*51*, *52*) or favored a closer relationship to the great apes than to the lesser apes (*53*, *54*). A few proposed that

Box 1. Simplified taxonomy of extant primates. The adjectives "lesser" and "great" refer to the smaller size of the former relative to great apes and human group, not to old evolutionary notions based on the Scala Naturae. Given that some apes are more closely related to humans than to other apes, the word "ape" is a gradistic term used here informally to refer to all nonhominin hominoids. Finally, the taxonomic convention used (the most common), does not reflect that panins and hominins are monophyletic [although some do; e.g., (*169*)].

Order Primates
- Suborder Strepsirrhini (non-tarsier "prosimians": lemurs, galagos and lorises)
- Suborder Haplorrhini (tarsiers and simians)
  - Infraorder Tarsiiformes (tarsiers)
  - Infraorder Simiiformes (or Anthropoidea: simians or anthropoids)
    - Parvorder Platyrrhini (New World monkeys)
    - Parvorder Catarrhini (Old World simians)
      - Superfamily Cercopithecoidea (Old World monkeys)
      - Superfamily Hominoidea (apes and humans)
        - Family Hylobatidae ("lesser apes": gibbons and siamangs)
        - Family Hominidae ("great apes" and humans)
          - Subfamily Ponginae (the orangutan lineage)
            - Genus *Pongo* (orangutans)
          - Subfamily Homininae (the African ape and human lineage)
            - Tribe Gorillini (the gorilla lineage)
              - Genus *Gorilla* (gorillas)
            - Tribe Panini (the chimpanzee lineage)
              - Genus *Pan* (common chimpanzees and bonobos)
            - Tribe Hominini (the human lineage)
              - Genus *Homo* (humans)

humans were more closely related to one or both of the African apes (*55*, *56*), although these views were not widely accepted (*57*). These alternative phylogenetic hypotheses heavily affected reconstructions of the LCA. Some (e.g., Schultz, Straus) advocated for a "generalized" ape ancestor (*52*), whereas others relied on extant hominoid models. Notably, Keith developed a scenario in which a "hylobatian" brachiating stage preceded an African ape-like creature: a knuckle-walking "troglodytian" phase immediately preceding bipedalism (*11*). Focused on Keith's "hylobatian" stage, Morton proposed that the "vertically suspended posture" of a small-bodied hylobatid-like ancestor caused the erect posture of human bipedalism (*12*). Gregory, another prominent "brachiationist," supported similar views (*53*). Morton argued that knuckle walking did not represent an intermediate stage preceding bipedalism but rather a reversion toward quadrupedalism in large-bodied apes specialized for brachiation. At that time, "brachiation" was used for any locomotion in which the body was suspended by the hands. Now, it refers to the pendulum-like arm-swinging locomotion of hylobatids (*6*).

By the 1960s, the Leakeys' discoveries in Tanzania [e.g., *Paranthropus boisei* (*58*), *Homo habilis* (*59*)] reinforced the relevance of Africa in human evolution, which became firmly established as the "mother continent" with the *A. afarensis* discoveries during the 1970s (*60*, *61*). LCA models still centered on the available fossil apes (mostly represented by jaw fragments and isolated teeth) accumulated after decades of paleontological fieldwork in Africa and Eurasia. In 1965, Simons and Pilbeam (*62*) revised and organized available Miocene apes in three genera: *Dryopithecus*, *Gigantopithecus*, and *Ramapithecus*. The genus *Sivapithecus* was included in *Dryopithecus*, considered the ancestor of African apes, whereas *Ramapithecus* was considered ancestral to humans based on its short face (and inferred small canines) (*63*). Leakey (*64*) and others agreed with Simons and Pilbeam that humans belong to their own family (Hominidae, or "hominids"), whereas great apes would belong to a distinct family (Pongidae, or "pongids"). He also agreed that *Ramapithecus* was an Asian early human ancestor. However, Leakey proposed reserving the genus *Sivapithecus* for the "Asian dryopithecines" and claimed that the human lineage could be traced back to, at least, the middle Miocene of Africa with *Kenyapithecus wickeri* (~14 Ma).

Two major "revolutions" in the study of evolutionary relationships started in the 1960s. First, a series of studies jump-started the field of molecular anthropology: Blood protein comparisons by Zuckerkandl *et al.* (*65*) and Goodman (*66*) found that some great apes—gorillas and chimpanzees—were more closely related to humans than to orangutans. Sarich and Wilson developed an "immunological molecular clock" and concluded that African apes and humans share a common ancestor as recent as ~5 Ma (*67*). These results led to decades-long debates regarding the African ape–human split. For example, Washburn resurrected extant African apes as ancestral models in human evolution, proposing knuckle walking as the precursor of terrestrial bipedalism (*68*). By contrast, paleontologists argued that the molecular clock was inaccurate because of the much older age of the purported human ancestors *Kenyapithecus* and *Ramapithecus* (*69*). Second, Hennigian cladistics ("phylogenetic systematics"), which only recognizes "synapomorphies" (shared derived features) as informative for reconstructing phylogeny (*70*), became slowly implemented in anthropology by the mid-1970s (*71*).

In the 1970s and 1980s, the relationships among gorillas, chimpanzees, and humans were still disputed. Chromosomal comparisons (*72*), DNA hybridization (*73*), and hemoglobin sequencing (*74*) supported a closer relationship between chimpanzees and humans, whereas morphology-based cladistics recovered gorilla-chimpanzee as monophyletic (*75*). In the late 1980s, the first single-locus DNA sequencing studies (*76*), followed in the 1990s with multiple loci analyses, finally resolved the "trichotomy" (*77*). Current genomic evidence indicates that humans are more closely related to chimpanzees (*5*), having diverged at some time between ~9.3 and ~6.5 Ma (*4*). Ever since "the molecular revolution," the perceived relevance of fossil apes in human evolution has been in jeopardy.

African apes as time machines?

Extant African apes have been considered ancestral models since Keith's "troglodytian" stage in the 1920s (*11*), and especially since the 1960s, with updated hypotheses inspired by the "molecular revolution" (*68*, *78*) and field discoveries on chimpanzee behavior by Goodall (*79*). Leakey played a central role in promoting Goodall's pioneering research (subsequently fostering Fossey's research in gorillas and Galdikas's research in orangutans). Now, a prominent paradigm proposes that chimpanzees represent "living fossils" that closely depict the *Pan-Homo* LCA (*14*, *16*). This model combines molecular data with the anachronistic view that *Gorilla* and *Pan* are morphologically similar (*75*). Under these assumptions, knuckle walking, once used to defend African ape monophyly (*80*), is used to argue that African apes are morphologically "conservative" and only display size-related differences (*14*). This model contends that gorillas are allometrically enlarged chimps and that chimpanzees [or bonobos (*78*)] constitute a suitable model for the *Pan-Homo* LCA, perhaps even the hominine or hominid LCAs (*14*). This narrative also incorporates the paleobiogeographic assumption

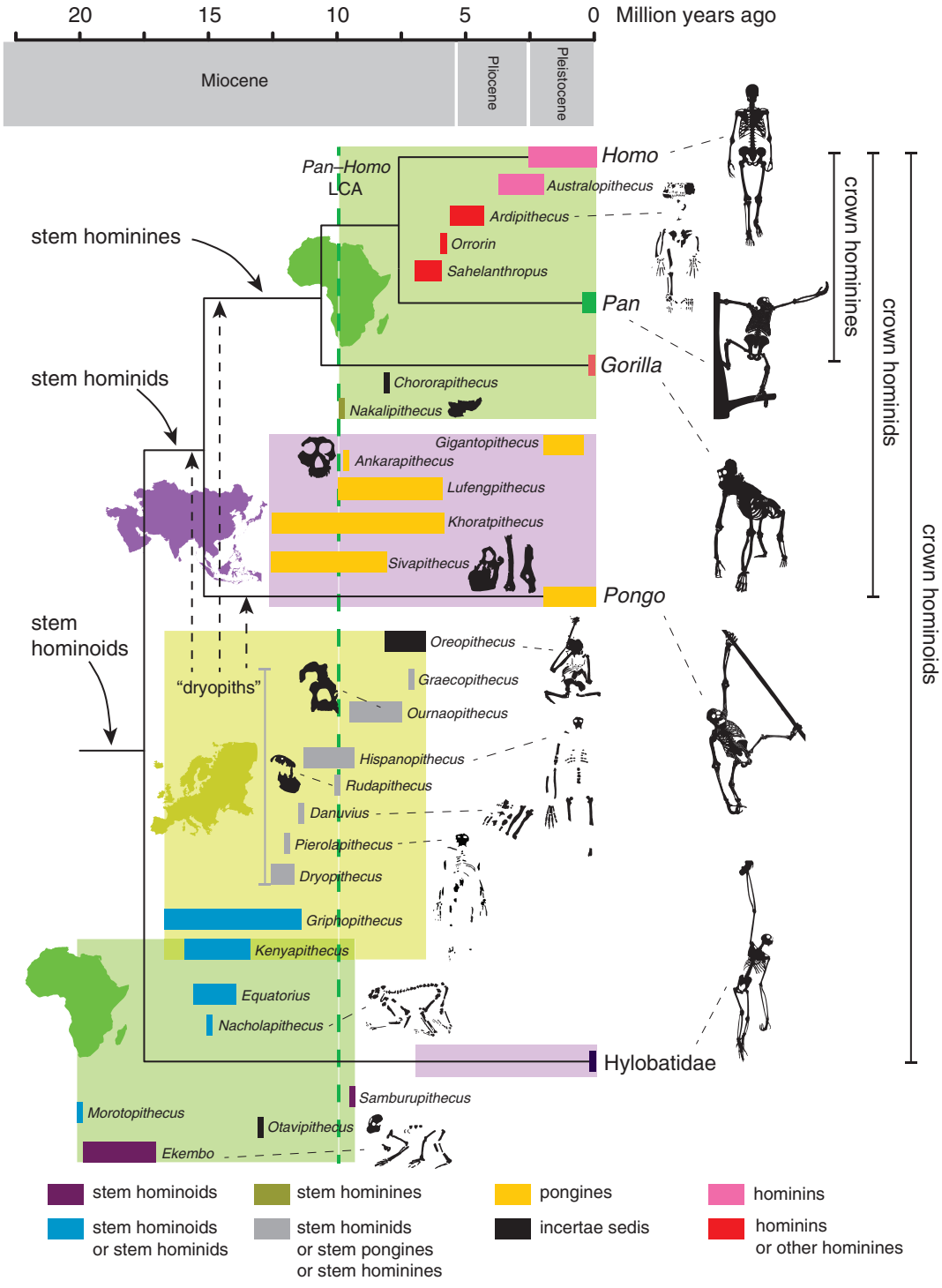


Fig. 3. Phylogenetic relationships among living hominoids and chronostratigraphic ranges of fossil hominoids. A time-calibrated phylogenetic tree of living hominoids is depicted next to the spatiotemporal ranges of the fossil hominoids mentioned in the text. Fossil taxa are color coded based on possible phylogenetic hypotheses. The vertical green dashed line indicates that there is a continuity in the African fossil ape record. However, currently, it is sparse between ~14 and 10 Ma. Robust and lasting phylogenetic inferences of apes are difficult, in part, because of the fragmentary nature of the fossil record and probable high levels of homoplasy. Many Miocene ape taxa are represented only by fragmentary dentognathic fossils, and the utility of mandibles and molars for inferring phylogeny in apes has been questioned. Another area of uncertainty relates to the position of many early and middle Miocene African apes relative to the crown hominoid node. The discovery or recognition of more complete early Miocene fossil hylobatids would help resolve their position and, thus, what really defines the great ape and human family. Splitting times are based on the molecular clock estimates of Springer *et al*. (*168*) (hominoids and hominids) and Moorjani *et al*. (*4*), which are more updated for hominines and *Pan-Homo*. Silhouettes are not to scale. Shaded boxes represent geographic distributions (green is Africa, gold is Europe, and purple is Asia).

that African apes likely occupy the same habitats as their ancestors: Without new selection pressures, there was no need for evolution.

If hominins originated from a chimpanzee-like LCA, human bipedalism must have evolved from knuckle walking (*15*), a functional compromise enabling terrestrial travel while retaining climbing adaptations (*80*). Under this view, bipedal hominins originated from an ancestor that was already terrestrial while traveling. These conclusions are logical from a "top-down" perspective, based on the evidence provided by extant hominoids and early hominins. However, a fully informed theory of hominin origins must also apply a "bottom-up" approach (*81*, *82*), from the perspective of extinct apes preceding the *Pan-Homo* split. It is also essential to clarify whether chimpanzees represent a good ancestral model for the *Pan-Homo* LCA. Unfortunately, the view from the bottom is blurry.

The tangled branches of ape evolution

The fossil ape dilemma: Homoplasy and mosaic evolution

With more than 50 hominoid genera and a broad geographic distribution (Fig. 1), the Miocene has been dubbed "the real planet of the apes" (*83*). Besides their fragmentary nature, a persistent challenge is understanding the phylogenetic relationships among fossil apes, which exhibit mosaics of primitive and

derived features with no modern analogs. The Asian Miocene ape *Sivapithecus* best exemplifies this complexity. Discoveries during the 1970s and 1980s, including a facial skeleton (*84*), clarified that *Ramapithecus* is a junior synonym of *Sivapithecus*, which is likely related to orangutans (*85*). However, two *Sivapithecus* humeri show a primitive (pronograde-related) morphology, calling into question the close phylogenetic link with *Pongo* that had been inferred from facial similarities (*86*).

The root of this "*Sivapithecus* dilemma" (*18*) is identifying where "phylogenetic signal" is best captured in hominoids: the postcranium or the cranium? The former implies that a *Pongo*-like face evolved independently twice; the latter entails that some postcranial similarities among living apes evolved more than once. Both hypotheses highlight the phylogenetic noise that homoplasy introduces in phylogenetic inference. Indeed, several studies have found that homoplasy similarly affects both anatomical areas (*87*). The conclusion that *Sivapithecus* is not a pongine relies on the assumption that suspensory adaptations and other orthograde-related features present in living hominoids were inherited from their LCA (*18*). However, this is contradicted by differences among living apes [e.g., forelimb and hand anatomy, degree of limb elongation, hip abduction capability (*8*, *9*, *19*, *80*, *88*–*91*)]. These studies concluded that apparent similarities could represent independently evolved biomechanical solutions to similar locomotor selection pressures. For instance, hand length "similarities" among living apes result from different combinations of metacarpal and/or phalangeal elongation in each extant genus (*9*).

Parallel evolution—homoplasy among closely related taxa due to shared genetic and developmental pathways—could explain some postcranial similarities related to suspensory behaviors among extant apes (*80*). Compared with convergences among distantly related taxa, parallelisms are more subtle and difficult to detect and they readily evolve when similar selection pressures appear. Within extant primates, suspensory adaptions evolved independently in atelines and between hylobatids and great apes (*8*, *80*, *88*, *91*, *92*). When the hominoid fossil record is added, independent evolution of suspensory adaptations has been inferred, too, for orangutans, chimpanzees, and some extinct lineages (*9*, *89*, *93*, *94*). Knuckle walking has also been proposed to have different origins in gorillas and chimpanzees (*80*, *93*, *95*). As for suspension, the preexistence of an orthograde body plan, vertical climbing, and general arboreal heritage could have facilitated the independent evolution of knuckle walking to circumvent similar biomechanical demands during terrestrial quadrupedalism while preserving a powerful grasping hand suitable for arboreal locomotion (*9*).

The possibility of parallelisms indicates that ancestral nodes in the hominoid evolutionary tree, including the *Pan-Homo* LCA, cannot be readily inferred without incorporating fossils. In addition, fossils from "known" evolutionary lineages are commonly used to calibrate molecular clocks despite being subject to considerable uncertainty (*4*). Even worse, relatively complete fossil apes undisputedly assigned to early members of the gorilla and chimpanzee lineages remain to be found.

Counting crowns: The case of the European Miocene apes

Sivapithecus and other fossil Asian great apes (e.g., *Khoratpithecus*, *Ankarapithecus*, *Lufengpithecus*) are generally considered pongines (Fig. 3) based on derived craniodental traits shared with *Pongo* (*94*, *96*–*98*), although alternative views exist, particularly for *Lufengpithecus* (*99*). By contrast, the phylogenetic positions of apes from the African early (e.g., *Ekembo*, *Morotopithecus*) and middle Miocene (*Kenyapithecus, Nacholapithecus*, *Equatorius*) remain very controversial. Like *Sivapithecus*, they exhibit only some modern hominoid features superimposed onto a primitive-looking pronograde ("monkey-like") body plan (Fig. 2). Some authors interpret this mosaicism as indicating that most Miocene apes do not belong within the crown hominoid radiation and, thus, are irrelevant to reconstructions of the *Pan-Homo* LCA (*14*). This is likely the case for early Miocene African taxa. However, the vertebrae of *Morotopithecus* [~20 Ma (*100*) or ~17 Ma (*101*)] display orthogrady-related features absent from other stem hominoids, indicating either a closer relationship with crown hominoids or an independent evolution of orthogrady (*102*). In turn, *Kenyapithecus* and *Nacholapithecus* are commonly regarded as preceding the pongine-hominine split owing to the possession of some modern hominid craniodental synapomorphies combined with a more primitive postcranium than that of living great apes (*94*, *103*). This raises the question: Can some Miocene apes belong to the crown hominid clade despite lacking many of the features shared by extant great apes?

The large-bodied apes from the middle-to-late Miocene of Europe are at the center of discussions about great ape and human evolution (*19*, *28*, *94*, *104*, *105*). Named after *Dryopithecus* (*3*), they are generally distinguished as a subfamily (Dryopithecinae) (*94*) or tribe (Dryopithecini) (*28*). However, it is unclear if they constitute a monophyletic group or a paraphyletic assemblage of stem and crown hominoids (*94*). Thus, we refer to them informally as "dryopiths." These apes are dentally conservative, but each genus exhibits different cranial and postcranial morphology. The dryopith fossil record includes the oldest skeletons that consistently exhibit postcranial features of living hominoids (orthograde body plan and/or long and more curved digits). *Dryopithecus* (~12 to 11 Ma) is known from craniodental remains and isolated postcranials that are too scarce to reconstruct its overall anatomy (*106*). By contrast, *Pierolapithecus* (~12 Ma) is represented by a cranium with an associated partial skeleton (*19*). Cranially a great ape, its rib, clavicle, lumbar, and wrist morphologies are unambiguous evidence of an orthograde body plan. Yet, unlike chimpanzees and orangutans (but similar to gorillas), *Pierolapithecus* lacks specialized below-branch suspensory adaptations [see discussion in (*10*)]. The recently described *Danuvius* (~11.6 Ma, Germany), and the slightly younger (~10 to 9 Ma) *Hispanopithecus* (Spain) (*105*) and *Rudapithecus* (Hungary) (*28*) represent the oldest record of specialized below-branch suspensory adaptations (e.g., long and strongly curved phalanges; Fig. 2). *Danuvius* has also been argued to show adaptations to habitual bipedalism (but see below).

The different mosaic morphology exhibited by each dryopith genus is a major challenge for deciphering their phylogenetic relationships (Fig. 3). Current competing phylogenetic hypotheses consider dryopiths as stem hominoids (*107*, *108*), stem hominids (*94*, *96*, *109*), or crown hominids closer to either pongines (*105*), hominines (*28*), or even hominins (*29*, *110*). However, recent phylogenetic analyses of apes recovered dryopiths as stem hominids (*97*, *109*), perhaps except *Ouranopithecus* (~9 to 8 Ma) and *Graecopithecus* (~7 Ma) (*97*). *Ouranopithecus* has been interpreted by some as a stem hominine, or even as a crown member more closely related to the gorilla or human lineages (*110*). *Graecopithecus* has also been advocated as a hominin (*29*), although the fragmentary available material hinders evaluation of this hypothesis. Such contrasting views about dryopiths stem from their incomplete and fragmentary fossil record coupled with pervasive homoplasy. However, because these factors are equal for all researchers, their different conclusions must also relate to analytical differences (e.g., taxonomy, sampling, polymorphic and continuous trait treatment). The root of the conflict is the remarkable differences in subjective definition and scoring of complex morphologies (e.g., "incipient supraorbital torus").

Paleobiogeography of the African ape and human clade

One hundred fifty years after Darwin speculated that modern African ape and human ancestors originated in Africa, possible hominins have been found as far back as the latest Miocene of Africa (*21*, *33*, *111*): *Sahelanthropus* (~7 Ma), *Orrorin* (~6 Ma), and *Ardipithecus kadabba* (~5.8 to 5.2 Ma). However, others question the feasibility of identifying the earliest hominins among the diverse Miocene apes

(*96*, *112*). Puzzlingly, despite some claims based on scarce remains (*113*–*115*), ancient representatives of the gorilla and chimpanzee lineages remain elusive. Some apes from the African late Miocene—*Chororapithecus* (*26*), *Nakalipithecus* (*27*), and *Samburupithecus* (*116*)—have been interpreted as hominines, but the available fragmentary remains preclude a conclusive assessment. Furthermore, *Samburupithecus* is likely a late-occurring stem hominoid (*97*, *117*).

During the middle Miocene (~16.5 to 14 Ma), apes are first found "out of Africa." These are the genera *Kenyapithecus* (Turkey) and *Griphopithecus* (Turkey and central Europe). We informally refer to them as the "kenyapiths" because there is no consensus on their relationships (*28*, *94*, *118*). Kenyapiths indicate that putative stem hominids are first recorded in Eurasia and Africa before the earliest record of both European dryopiths and Asian pongines at ~12.5 Ma (*94*). Paleobiogeographical and paleontological data suggest that kenyapiths dispersed from Africa into Eurasia as one of the multiple catarrhine intercontinental dispersal events occurred during the Miocene (e.g., hylobatids, pliopithecoids) (*83*, *94*). Although some competing evolutionary scenarios agree that kenyapiths gave rise to dryopiths in Europe, the phylogenetic and geographic origin of hominines remains contentious (*28*, *94*).

If dryopiths are stem hominids, they could either be close to the crown group or constitute an evolutionary dead end, an independent "experiment" not directly related to either pongines or hominines. Alternatively, dryopiths might be crown hominids more closely related to one of these groups. If dryopiths are hominines, this implies that the latter could have originated in Europe and subsequently dispersed "back to Africa" during the late Miocene (*28*, *29*, *83*). This would coincide with vegetation structure changes caused by a trend of increased cooling and seasonality (*32*) that ultimately drove European apes to extinction [or back to Africa (*28*)]. In this scenario, hominines and pongines would be vicariant groups that originally evolved in Europe and Asia, respectively, from early kenyapith ancestors. Given the suspensory specializations of late Miocene dryopiths (*Hispanopithecus* and *Rudapithecus*), if modern African apes originated from these forms, this scenario implies that the hominine ancestor could have been more reliant on suspension than living chimpanzees or gorillas. The claim that hominines originated outside of Africa may be justified by cladistic analyses recovering dryopiths as stem hominines but may not be based on the lack of late Miocene great apes in Africa because fossils from this critical time period have been discovered (~13 to 7 Ma) (Fig. 3). Both molecular and paleontological evidence (e.g., *Sivapithecus*) situate the pongine-hominine divergence within the middle Miocene. Hence, the debate cannot be settled without more conclusively resolving the phylogenetic relationships of middle Miocene dryopiths.

An alternative scenario proposes a vicariant divergence for hominines and pongines from kenyapith ancestors but favors the origin of hominines in Africa (*94*, *119*). It argues for a second vicariant event between European dryopiths and Asian pongines soon after the kenyapith dispersal into Eurasia. Cladistically, dryopiths would be pongines but would share none of the currently recognized pongine autapomorphies, evolved after the second vicariant event. This scenario is difficult to test, but it would be consistent with the apparent absence of clear pongine synapomorphies in *Lufengpithecus* (*99*) and the more derived nasoalveolar morphology of *Nacholapithecus* (*103*) compared with some dryopiths (*106*). However, it would imply even higher levels of homoplasy, including the independent acquisition of an orthograde body plan in Africa and Eurasia from pronograde kenyapith ancestors.

A third possibility is that none of the taxa discussed above are closely related to the African ape and human clade (*107*). Under this view, bona fide extinct nonhominin hominines have yet to be found in largely unexplored regions of Africa, explaining the virtual lack of a gorilla and chimpanzee fossil record. According to Pilbeam, paleoanthropologists could be "like the drunk looking for his keys under the lamppost where it was light rather than where he had dropped them, working with what we had rather than asking whether or not that was adequate" [(*108*), pp. 155–156]. Africa is a huge continent, and most paleontological discoveries are concentrated in a small portion of it. The greatest challenge is finding hominoid-bearing Mio-Pliocene sites outside East and South Africa, even though we know they exist (*20*–*22*). Besides insufficient sampling effort, this is hindered by numerous impediments to fieldwork in most of Africa, including geopolitical conflicts, restricted land use development, lack of suitable outcrops (due to extensive vegetation cover), and taphonomic factors [tropical forests do not favor fossil preservation (*120*)].

A Miocene view of (Miocene) hominin origins

Evolution in motion

The decades-long feud regarding arboreality and bipedalism in *A. afarensis* exemplifies the complexity of inferring function from anatomy. "Totalist" functional morphologists rely on a species' "total morphological pattern" (*121*) to infer its locomotor repertoire. Totalists see a bipedal early hominin with some ape-like retentions (e.g., curved fingers) pointing to continued use of the trees and consider that certain not-yet-human-like features (e.g., hip) indicate a different type of bipedalism (*122*). Instead, "directionalists"—for whom functional inferences are only possible for derived traits evolved for a specific function—focus exclusively on bipedal adaptations (*123*). Totalist and directionalist interpretations of the fossil record differ in the "adaptive significance" attributed to primitive features, which result in different behavioral reconstructions. Two other related factors further complicate locomotor inferences in extinct species: First, different positional behaviors have similar mechanical demands [e.g., bipedalism, quadrupedalism and some types of climbing (*39*)]. Second, preexisting morphofunctional complexes originally selected to fulfill a particular function (adaptations) can be subsequently co-opted for a new role (exaptations).

The mosaic nature of hominoid morphological evolution makes the functional reconstruction of fossil apes especially challenging, as recently exemplified by *Danuvius* (*104*): It was described as possessing long and curved fingers, a long and flexible vertebral column, hip and knee joints indicative of extended postures, and an ankle configuration aligning the foot perpendicular to the long axis of the tibia. Such a combination of features was functionally interpreted as indicating below-branch suspension combined with above-branch bipedalism. However, a critique to the original study concluded that the morphological affinities of *Danuvius* with modern great apes support a positional repertoire that includes orthogrady and suspension, but not bipedalism (*124*). Part of the "problem" with the original interpretation is that it infers a derived locomotor behavior—bipedalism—from primitive features that are also functionally related to quadrupedalism. For instance, the inferred "long-back" morphology of *Danuvius* is characteristic of most quadrupedal monkeys and other Miocene apes (*125*), denoting the lack of trunk specialization seen in extant great apes. The *Danuvius* femoral head joint, being (primitively) posterosuperiorly expanded (*126*), is consistent with flexed quadrupedal hip postures that are not used during human-like bipedalism. In addition, the distal tibia configuration of *Danuvius* is shared with *Ekembo* and cercopithecoids (*104*), thus being likely plesiomorphic and not unique to bipeds. When the primitive and derived features of *Danuvius* are considered, a totalist would argue that it combined high degrees of plesiomorphic quadrupedal locomotion with novel (suspensory) behaviors, whereas a directionalist would downplay the primitive features in favor of the newly derived adaptive traits (i.e., suspension).

The late Miocene *Oreopithecus* (~7 Ma, Italy) is another example of conflicting phylogenetic and functional signals. Phylogenetic interpretations of *Oreopithecus* include cercopithecoid, stem hominoid, and hominid (even hominin) status (*127*). However, current phylogenetic analyses suggest that *Oreopithecus* could represent

a late-occurring stem hominoid (*97*, *128*), with postcranial adaptations to alternative types of orthogrady, such as forelimb-dominated behaviors (*129*) and terrestrial bipedalism (*130*). Even if not directly related to hominins (or modern hominoids), the locomotor adaptations of *Oreopithecus*, and other Miocene apes, are worthy of further research to understand the selection pressures that led to the (independent) emergence of modern hominoid positional behaviors.

To distinguish true locomotor adaptations from exaptations, current research efforts focus on plastic "ecophenotypic" traits, potentially denoting how fossil hominoids were actually moving. Bone is a living tissue, and growth is expected to occur in predictable ways that reflect loading patterns throughout life (*131*). Thus, cross-sectional and trabecular bone properties and their links to behavior are widely investigated (*132*, *133*). Yet, experimental studies indicate that internal bone morphology does not necessarily match stereotypical loading patterns (*134*). Ample evidence suggests that irregular loading, even in low magnitude, can be more osteogenically potent than stereotypical loading (*135*). This may bias interpretations of individual fossils with a species-atypical loading pattern during life (e.g., because of an injury). Bone (re)modeling also does not consistently occur in response to changes in loading pattern: It can occur in ways that detract from, rather than enhance, function (*136*) and may manifest differentially across the skeleton (*137*). Incongruence also exists between actual bone performance and expectations based on aspects of internal morphology (*138*). Finally, there is a strong genetic component to the responsiveness of bone (re)modeling to loading (*136*), which is largely unknown for most species. The confidence with which internal bone structures can be used to retrodict behavior in fossil species remains a work in progress.

Before bipedalism

Competing hypotheses about the locomotor behavior immediately preceding hominin bipedalism include terrestrial knuckle walking (*15*), palmigrade quadrupedalism (*93*), and different types of arboreal (orthograde) behaviors such as climbing and suspension (*7*), vertical climbing (*139*), or arboreal bipedalism and suspension (*104*, *140*). Miocene great apes can enlighten this question by helping to identify the polarity of evolutionary change preceding the *Pan-Homo* divergence (*81*, *82*). For instance, if *Pierolapithecus* is interpreted as an orthograde ape without specific suspensory adaptations but retaining quadrupedal adaptations [see alternatives in (*10*)], then the orthograde body plan and ulnocarpal contact loss could be interpreted as an adaptation to vertical climbing, subsequently co-opted for suspension (*19*). Similarly, habitual bipedalism might have directly evolved from other orthograde behaviors without an intermediate stage of advanced suspension or specialized knuckle walking. Hence, *Pierolapithecus* complements previous hypotheses that biomechanical aspects of the lower limb during quadrupedalism and vertical climbing could be functionally "preadaptive" for bipedalism (*39*, *139*).

A holistic view indicates that the *Pan-Homo* LCA was a Miocene ape with extant great ape–like cognitive abilities, likely possessing a complex social structure and tool traditions (*36*, *38*, *141*). This ape would exhibit some degree of body size and canine sexual dimorphism (with large honing male canines) (*15*), indicating a polygynous sociosexual system (*40*). Based on Miocene apes and earliest hominins, it is also likely that the *Pan-Homo* LCA was orthograde and proficient at vertical climbing [see alternative interpretation based on *Ardipithecus* (*33*, *93*)], but not necessarily adapted specifically for below-branch suspension or knuckle walking (*9*, *33*). Chimpanzees seem to retain the *Pan-Homo* LCA plesiomorphic condition in some regards [e.g., brain and body size (*38*), vertebral counts (*125*), foot morphology (*142*)]. However, in others [e.g., interlimb (*93*), hand (*9*), pelvis (*143*) length proportions; femur morphology (*89*)], early hominins are more similar to generalized Miocene apes. These results further reinforce the idea that functional aspects of other locomotor types were co-opted for bipedalism during hominin origins.

The "East Side Story" scenario links the divergence of chimpanzees and humans to the rifting of East Africa, which would have triggered a vicariant speciation event from the ancestral *Pan-Homo* LCA (*17*). Chimpanzees would have remained "frozen in time" in their ancestral tropical forest environment, whereas humans would be the descendants of the group "left behind" on the east side of the Rift. Major climate and landscape changes would have then forced the earliest hominins to adapt to more open (grassland savanna) environments by acquiring bipedalism—and the rest is history. Several decades after the proposal of this scenario, where do we stand?

The landscape of East Africa has dramatically changed during the past 10 million years because of tectonic events leading to specific climatic conditions and associated changes in vegetation structure, from mixed tropical forest to more heterogeneous and arid environments than elsewhere in tropical Africa (*144*, *145*). The trend of progressive aridification did not culminate in the predominance of savanna environments until ~2.0 Ma—roughly coinciding with hominin brain size increase and the appearance of *H. erectus*—and was punctuated by alternating episodes of extreme humidity and aridity, resulting in a fluctuating extension of forests through time (*144*, *145*). Despite ongoing discussions about early hominin paleoenvironments (woodland with forest patches versus wooded savanna) (*146*), evidence from Miocene apes (*30*, *31*) supports that the *Pan-Homo* LCA inhabited some kind of woodland. Therefore, it has been suggested that the *Pan-Homo* LCA was probably more omnivorous than chimpanzees (ripe fruit specialists) and likely fed both in trees and on the ground (*33*), in agreement with isotopic analyses for *Ardipithecus ramidus* (*41*).

Bipedalism would have emerged because of the selection pressures created by the progressive fragmentation of forested habitats and the need for terrestrial travel from one feeding patch to the next. Data on extant ape positional behaviors (Fig. 4) suggest that hominin terrestrial bipedalism originated as a posture rather than a means of travel on the ground (*147*) or in trees (*140*). Rose (*39*) proposed a long process of increasing commitment to bipedality in the transition to more complex open habitats throughout the Plio-Pleistocene, and Potts (*148*) argued that key stages in hominin evolution may relate to adaptive responses to cope with highly variable environments. The fossil and archaeological records provide a new twist to the order of evolutionary events in early hominin evolution. The remains of *Orrorin* and *Ar. ramidus* indicate that habitual terrestrial bipedalism, enhanced precision grasping, and loss of canine honing evolved at the dawn of the human lineage well before brain enlargement (*9*, *33*, *89*, *93*). It was not until later in time [maybe starting with *Australopithecus* (*149*) and continuing with *Homo*], that some preexisting hand attributes were co-opted for purposive and systematic stone toolmaking in more encephalized hominins with more advanced cognitive abilities (*38*, *150*).

The specialization trap

That hominins continuously evolved since the *Pan-Homo* LCA is universally accepted, but the possibility that all living hominoids (including chimpanzees) experienced their own evolutionary histories is sometimes disregarded. Potts (*151*) suggested that the greater cognitive abilities of great apes originated to continue exploiting fruit supplies from densely forested environments in front of strong environmental variability. Coupled with locomotor adaptations (e.g., vertical climbing, suspension) enabling an efficient navigation through the canopy, this "cognitive trap" would consist of an adaptive feedback loop between diet, locomotion, cognition, and life history. Although hominids originated approximately during the "Mid-Miocene Climatic Optimum" (~17 to 15 Ma), their subsequent radiation from ~14 Ma onward paralleled a trend of climatic "deterioration" during the rest of the Miocene (*152*). Great apes might have initially thrived by

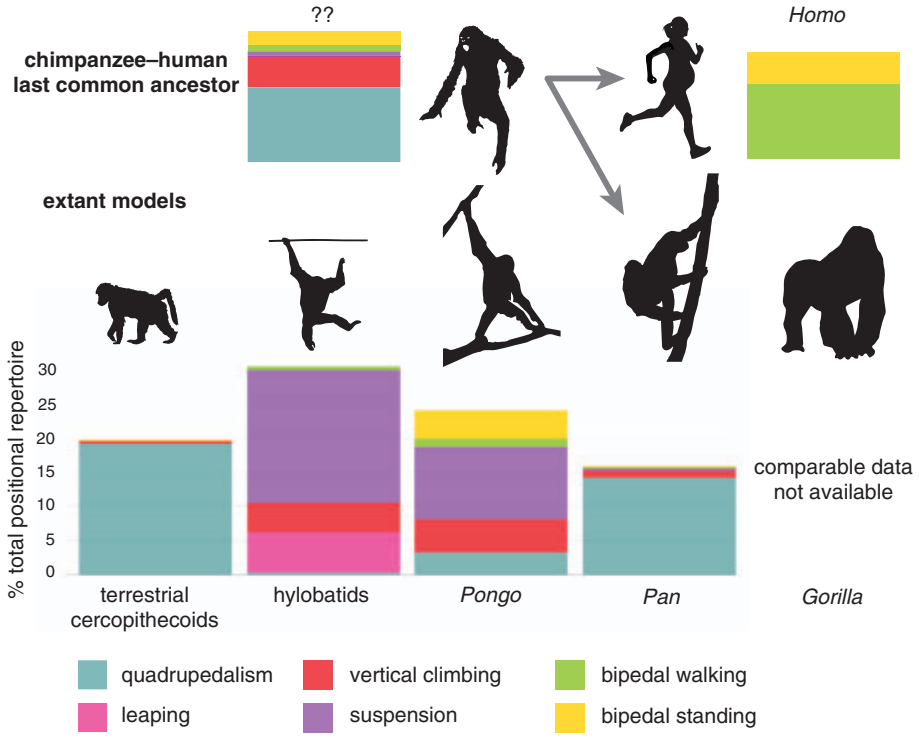


Fig. 4. The positional repertoire preceding human bipedalism. Although one particular behavior can dominate the locomotor repertoire of a given species, the full positional repertoire (postural and locomotor behaviors) of living primates is diverse, complex, and not fully understood. For example, some locomotor behaviors are not totally comparable (e.g., monkey quadrupedalism versus African ape knuckle walking). Furthermore, comprehensive data are not yet available for some extant hominoids (e.g., *Gorilla*). Bipedalism did not appear de novo in hominins; it existed as a posture or locomotion within a broader Miocene ape positional repertoire. The combined evidence of Miocene apes and early hominins indicate that the locomotor repertoire of the *Pan-Homo* LCA likely included a combination of positional behaviors not represented among living primates. Over time, bipedal behaviors became the predominant activity within the repertoire of early hominins (and knuckle walking in the chimpanzee lineage). Locomotor behaviors (plus bipedal standing) in each taxon represent percentages of total positional behavior repertoire. (The full repertoire is not shown; hence, these do not add to 100%.) Data were taken from (*92*). Quadrupedalism includes Hunt's categories "quadrupedal walk" and "quadrupedal run," suspension includes "suspensory," "brachiate," "clamber," and "transfer." The locomotor repertoire compositions of the LCA and modern humans (*Homo*) are conjectural, for illustrative purposes.

evolving particular adaptations to more efficiently exploit their habitats, thereby occupying new adaptive peaks without abandoning the same area of the adaptive landscape broadly occupied by earlier stem hominoids. Nevertheless, this evolutionary strategy would become unsustainable once a particular paleoenvironmental threshold was surpassed. This could explain the fate of European dryopiths, which survived for some time under suboptimal conditions (despite the progressive trend of cooling and increased seasonality) until they vanished (*94*).

The dietary, locomotor, and cognitive specializations of late Miocene great apes would have hindered their shift into new adaptive peaks suitable for the more open environments toward the latest Miocene (*153*). The Miocene planet of the apes gave way to the time of the more generalist Old World monkeys, enabling their survival in a wider variety of seasonal habitats (*30*, *92*, *154*). The same specialization trap can explain the delayed retreat of pongines (and hylobatids) to southeastern Asia throughout the Plio-Pleistocene. The highly specialized orangutans remain extant, but not for long because their habitat continues to shrink. African apes could have partially overcome the specialization trap by evolving (perhaps in parallel) semiterrestrial adaptations—knuckle walking. Gorillas also expanded their dietary range (more folivorous) and enlarged their body size. Contrary to the view that gorillas are "enlarged" chimpanzees, morphometric analyses indicate that gorillas underwent their own evolutionary history, resulting in different ontogenetic trajectories (*155*, *156*) and postcranial differences that cannot be explained by size-scaling effects (*9*, *143*). Why, when, and how many times knuckle walking evolved is more difficult to explain than the origin of hominin bipedalism. Habitat fragmentation coupled with a higher reliance on arboreal feeding might be invoked (i.e., knuckle walking serves both terrestrial and arboreal locomotion). This idea is difficult to reconcile with the premise that continuous-canopy forests covered the tropical belt of central and western Africa since the Miocene, unless gorillas and chimpanzees evolved in less densely forested habitats (*30*, *31*, *114*) and retreated to tropical forests when outcompeted by hominins and/or cercopithecoids. Ironically, the same specializations that allowed great apes to survive despite major environmental challenges since the late Miocene might ultimately doom them to extinction.

Hominins might have escaped the great-ape specialization trap by evolving novel and more radical adaptations: bipedalism (another specialized orthograde locomotion), concomitant freeing of the hands, and subsequent enhanced manual dexterity, brain configuration, sociosexual behavior, and culturally mediated technology. Human evolution also reflects the progressive adaptation (biological first, cultural later) to ever-changing environments (*39*, *148*). Some essential changes (upright posture, enhanced cognition) are just the continuation of a trend started in Miocene hominoids (*19*, *36*, *151*). While escaping from the great ape specialization trap, humans might have fallen into another evolutionary cul-de-sac, with current human activities and overpopulation leading the biosphere to a point beyond return (*157*). Will humans escape their own specialization trap?

Conclusions and perspectives

Fossils uniquely inform deep-time evolutionary studies, which is essential to plan for the future (*158*). However, we must be aware of the many existing limitations and the gaps in our knowledge. For example, we need more fossils because we are likely missing vastly more than what we have. More fieldwork is necessary to find fossil apes close to the gorilla or chimpanzee lineages, and it is essential to extend such efforts to unexplored or undersampled areas (Fig. 1). It is also essential to continue developing tools of phylogenetic inference. Bayesian approaches are promising, but uncertainty remains about their applicability to morphological data (*159*). Improvements in the treatment of continuous characters and recent methodological advances for analyzing three-dimensional geometric morphometric data within a cladistic framework (in combination with traditional characters) are promising for reconstructing fossil hominoid phylogeny (*160*). The oldest (recently retrieved) ancient DNA is ~1 Ma (*161*). Paleoproteomics could be a complementary solution because it has enabled sampling further back in time up to ~2 Ma, recently confirming

the pongine status of *Gigantopithecus* (*162*). Future technological advances in paleoproteomics could potentially help to answer key questions by retrieving paleoproteomes from Miocene apes.

Locomotor reconstructions of the *Pan-Homo* LCA and other fossil hominoids are seriously hampered by the lack of current analogs. Washburn spotted the fundamental limitation: "It is not possible to bring the past into the laboratory. No one can see a walking *Australopithecus*" [(*163*), p. 67]. Such inferences rely on morphofunctional assumptions of bone, joint, or muscle function, but experimentally derived biomechanical data are required to test these assumptions and provide reliable inferences from fossils. Technological advances now facilitate noninvasive kinematic data collection from animals in their natural environments (*164*). In turn, experimental and morphological information should be integrated to better predict the locomotion of fossil hominoids. Forward dynamic simulations offer a powerful pathway for predicting de novo movements in fossil species while iterating possible effects of morphology and soft tissue (*165*).

Humans are storytellers: Theories of human evolution often resemble "anthropogenic narratives" that borrow the structure of a hero's journey to explain essential aspects such as the origins of erect posture, the freeing of the hands, or brain enlargement (*166*). Intriguingly, such narratives have not drastically changed since Darwin (*166*). We must be aware of confirmation biases and ad hoc interpretations by researchers aiming to confer their new fossil the starring role within a preexisting narrative. Evolutionary scenarios are appealing because they provide plausible explanations based on current knowledge, but unless grounded in testable hypotheses, they are no more than "just-so stories" (*167*).

Many uncertainties persist about fossil apes, and the day in which the paleobiology of extinct species can be undisputedly reconstructed is still far away. However, current disagreements regarding ape and human evolution would be much more informed if, together with early hominins and living apes, Miocene apes were also included in the equation. This approach will allow us to better discern primitive and derived traits, the common from the specific, or the unique. This is the role of fossil apes in human evolution.

REFERENCES AND NOTES

1. C. Darwin, *The Descent of Man, and Selection in Relation to Sex (Vol. I)* (John Murray, 1871).
2. T. H. Huxley, *Evidence As To Man's Place in Nature* (Williams and Norgate, 1863).
3. E. Lartet, Note sur un grand singe fossile qui se rattache au groupe des singes superieurs. *C. R. Hebd. Seances Acad. Sci.* **43**, 219–223 (1856).
4. P. Moorjani, C. E. G. Amorim, P. F. Arndt, M. Przeworski, Variation in the molecular clock of primates. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10607–10612 (2016). doi: 10.1073/pnas.1600374113; pmid: 27601674
5. J. Prado-Martinez *et al*., Great ape genetic diversity and population history. *Nature* **499**, 471–475 (2013). doi: 10.1038/nature12228; pmid: 23823723
6. K. D. Hunt *et al*., Standardized descriptions of primate locomotor and postural modes. *Primates* **37**, 363–387 (1996). doi: 10.1007/BF02381373
7. J. Stern, Before bipedality. *Yearb. Phys. Anthropol.* **19**, 59–68 (1975).
8. S. G. Larson, Parallel evolution in the hominoid trunk and forelimb. *Evol. Anthropol.* **6**, 87–99 (1998). doi: 10.1002/(SICI)1520-6505(1998)6:3<87::AID-EVAN3>3.0.CO;2-T
9. S. Almécija, J. B. Smaers, W. L. Jungers, The evolution of human and ape hand proportions. *Nat. Commun.* **6**, 7717 (2015). doi: 10.1038/ncomms8717; pmid: 26171589
10. M. Nakatsukasa, S. Almécija, D. R. Begun, in *The Evolution of the Primate Hand: Anatomical, Developmental, Functional, and Paleontological Evidence*, L. T. Kivell, P. Lemelin, G. B. Richmond, D. Schmitt, Eds. (Springer, 2016), pp. 485–514.
11. A. Keith, Hunterian lectures on man's posture: Its evolution and disorders. Lecture II. The evolution of the orthograde spine. *BMJ* **1**, 499–502 (1923). doi: 10.1136/bmj.1.3247.499; pmid: 20771062
12. D. J. Morton, Evolution of man's erect posture (preliminary report). *J. Morphol.* **43**, 147–179 (1926). doi: 10.1002/jmor.1050430108
13. M. Cartmill, in *Functional Vertebrate Morphology*, M. Hildebrand, D. Bramble, K. Liem, D. Wake, Eds. (Belknap Press, 1985), pp. 73–88.
14. D. R. Pilbeam, D. E. Lieberman, in *Chimpanzees and Human Evolution*, M. N. Muller, R. W. Wrangham, D. R. Pilbeam, Eds. (The Belknap Press of Harvard Univ. Press, 2017), pp. 22–141.
15. B. G. Richmond, D. R. Begun, D. S. Strait, Origin of human bipedalism: The knuckle-walking hypothesis revisited. *Am. J. Phys. Anthropol.* **44** (suppl. 33), 70–105 (2001). doi: 10.1002/ajpa.10019; pmid: 11786992
16. R. Wrangham, D. Pilbeam, in *All Apes Great and Small*, B. M. F. Galdikas, N. E. Briggs, L. K. Sheeran, G. L. Shapiro, J. Goodall, Eds. (Kluwer Academic/Plenum Publishers, 2001), pp. 5–17.
17. Y. Coppens, East side story: The origin of humankind. *Sci. Am.* **270**, 88–95 (1994). doi: 10.1038/scientificamerican0594-88; pmid: 8197447
18. D. Pilbeam, N. Young, Hominoid evolution: Synthesizing disparate data. *C. R. Palevol* **3**, 305–321 (2004). doi: 10.1016/j.crpv.2004.01.006
19. S. Moyà-Solà, M. Köhler, D. M. Alba, I. Casanovas-Vilar, J. Galindo, *Pierolapithecus catalaunicus*, a new Middle Miocene great ape from Spain. *Science* **306**, 1339–1344 (2004). doi: 10.1126/science.1103094; pmid: 15550663
20. M. Brunet *et al*., The first australopithecine 2,500 kilometres west of the Rift Valley (Chad). *Nature* **378**, 273–275 (1995). doi: 10.1038/378273a0; pmid: 7477344
21. M. Brunet *et al*., A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**, 145–151 (2002). doi: 10.1038/nature00879; pmid: 12110880
22. G. C. Conroy, M. Pickford, B. Senut, J. Van Couvering, P. Mein, *Otavipithecus namibiensis*, first Miocene hominoid from southern Africa. *Nature* **356**, 144–148 (1992). doi: 10.1038/356144a0; pmid: 1545864
23. M. Pickford, Y. Coppens, B. Senut, J. Morales, J. Braga, Late Miocene hominoid from Niger. *C. R. Palevol* **8**, 413–425 (2009). doi: 10.1016/j.crpv.2008.11.003
24. B. Senut, M. Pickford, D. Wessels, Panafrican distribution of Lower Miocene Hominoidea. *C. R. Acad. Sci. Paris* **325**, 741–746 (1997).
25. P. J. Andrews, L. Martin, The phyletic position of the Ad Dabtiyah hominoid. *Bull. Br. Mus. Geol.* **41**, 383–393 (1987).
26. G. Suwa, R. T. Kono, S. Katoh, B. Asfaw, Y. Beyene, A new species of great ape from the late Miocene epoch in Ethiopia. *Nature* **448**, 921–924 (2007). doi: 10.1038/nature06113; pmid: 17713533
27. Y. Kunimatsu *et al*., A new Late Miocene great ape from Kenya and its implications for the origins of African great apes and humans. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19220–19225 (2007). doi: 10.1073/pnas.0706190104; pmid: 18024593
28. D. R. Begun, M. C. Nargolwalla, L. Kordos, European Miocene hominids and the origin of the African ape and human clade. *Evol. Anthropol.* **21**, 10–23 (2012). doi: 10.1002/evan.20329; pmid: 22307721
29. J. Fuss, N. Spassov, D. R. Begun, M. Böhme, Potential hominin affinities of *Graecopithecus* from the Late Miocene of Europe. *PLOS ONE* **12**, e0177127 (2017). doi: 10.1371/journal.pone.0177127; pmid: 28531170
30. P. Andrews, *An Ape's View of Human Evolution* (Cambridge Univ. Press, 2016).
31. P. Andrews, Last common ancestor of apes and humans: Morphology and environment. *Folia Primatol.* **91**, 122–148 (2020). doi: 10.1159/000501557; pmid: 31533109
32. T. E. Cerling *et al*., Comment on the paleoenvironment of *Ardipithecus ramidus*. *Science* **328**, 1105 (2010). doi: 10.1126/science.1185274; pmid: 20508112
33. T. D. White *et al*., *Ardipithecus ramidus* and the paleobiology of early hominids. *Science* **326**, 75–86 (2009). doi: 10.1126/science.1175802; pmid: 19810190
34. C. Linnaeus, *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis. Synonymis, Locis* (Laurentius Salvius, Holmiae, 1758), vol. Tomus I. Editio Decima, reformata.
35. R. D. Martin, *Primate Origins and Evolution: A Phylogenetic Reconstruction* (Chapman and Hall London, 1990).
36. D. M. Alba, Cognitive inferences in fossil apes (Primates, Hominoidea): Does encephalization reflect intelligence? *J. Anthropol. Sci.* **88**, 11–48 (2010). pmid: 20834049
37. M. Cartmill, Rethinking primate origins. *Science* **184**, 436–443 (1974). doi: 10.1126/science.184.4135.436; pmid: 4819676
38. S. Almécija, C. C. Sherwood, in *The Nervous Systems of Non-Human Primates*, vol. 3 of *Evolution of Nervous Systems*, J. Kaas, Ed. (Elsevier, ed. 2, 2017), pp. 299–315.
39. M. Rose, "The process of bipedalization in hominids" in *Origine(s) de la bipédie chez les hominidés*, Y. Coppens, B. Senut, Eds. (CNRS, Paris, 1991), pp. 37–48.
40. J. M. Plavcan, C. P. van Schaik, P. M. Kappeler, Competition, coalitions and canine size in primates. *J. Hum. Evol.* **28**, 245–276 (1995). doi: 10.1006/jhev.1995.1019
41. G. Suwa *et al*., Paleobiological implications of the *Ardipithecus ramidus* dentition. *Science* **326**, 94–99 (2009). doi: 10.1126/science.1175824; pmid: 19810195
42. M. Pickford, B. Senut, D. Gommery, J. Treil, Bipedalism in *Orrorin tugenensis* revealed by its femora. *C. R. Palevol* **1**, 191–203 (2002). doi: 10.1016/S1631-0683(02)00028-3
43. S. L. Washburn, Tools and human evolution. *Sci. Am.* **203**, 63–75 (1960). doi: 10.1038/scientificamerican0960-62; pmid: 13843002
44. C. O. Lovejoy, The origin of man. *Science* **211**, 341–350 (1981). doi: 10.1126/science.211.4480.341; pmid: 17748254
45. C. Darwin, *On the Origin of Species. Or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
46. E. Haeckel, *Natürliche Schöpfungsgeschichte* (Georg Reimer, 1868).
47. E. Dubois, On *Pithecanthropus erectus*: A transitional form between man and the apes. *J. Anthropol. Inst. G. B. Irel.* **25**, 240–255 (1896). doi: 10.2307/2842246
48. R. A. Dart, *Australopithecus africanus*: The man-ape of South Africa. *Nature* **115**, 195–199 (1925). doi: 10.1038/115195a0
49. J. S. Weiner, K. P. Oakley, W. E. Le Gros Clark, The solution of the Piltdown problem. *Bull. Br. Mus. Geol.* **2**, 139–146 (1953).
50. G. E. Lewis, Preliminary notice of new man-like apes from India. *Am. J. Sci.* **s5-27**, 161–181 (1934). doi: 10.2475/ajs.s5-27.159.161
51. H. F. Osborn, The discovery of Tertiary man. *Science* **71**, 1–7 (1930). doi: 10.1126/science.71.1827.1; pmid: 17760122
52. W. L. Straus Jr., The riddle of man's ancestry. *Q. Rev. Biol.* **24**, 200–223 (1949). doi: 10.1086/397067; pmid: 18138211
53. W. K. Gregory, How near is the relationship of man to the chimpanzee-gorilla stock? *Q. Rev. Biol.* **2**, 549–560 (1927). doi: 10.1086/394289
54. A. H. Schultz, The skeleton of the trunk and limbs of higher primates. *Hum. Biol.* **2**, 303–438 (1930).
55. G. Elliot Smith, *The Evolution of Man: Essays* (Oxford Univ. Press, 1924).
56. H. Weinert, *Ursprung der Menschheit: Über den engeren Anschluss des Menschengeschlechts an die Menschenaffen* (Ferdinand Enke, 1932).
57. A. H. Schultz, Characters common to higher primates and characters specific for man. *Q. Rev. Biol.* **11**, 259–283 (1936). doi: 10.1086/394508
58. L. S. B. Leakey, A new fossil skull from Olduvai. *Nature* **184**, 491–493 (1959). doi: 10.1038/184491a0
59. L. S. B. Leakey, P. V. Tobias, J. R. Napier, A new species of the genus *Homo* from Olduvai Gorge. *Nature* **202**, 7–9 (1964). doi: 10.1038/202007a0; pmid: 14166722

60. D. C. Johanson, M. Taieb, Plio—Pleistocene hominid discoveries in Hadar, Ethiopia. *Nature* **260**, 293–297 (1976). doi: 10.1038/260293a0; pmid: 815823
61. M. D. Leakey, R. L. Hay, Pliocene footprints in the Laetoli Beds at Laetoli, northern Tanzania. *Nature* **278**, 317–323 (1979). doi: 10.1038/278317a0
62. E. L. Simons, D. R. Pilbeam, Preliminary revision of the Dryopithecinae (Pongidae, Anthropoidea). *Folia Primatol.* **3**, 81–152 (1965). doi: 10.1159/000155026; pmid: 5320325
63. E. L. Simons, The phyletic position of *Ramapithecus*. *Postilla* **57**, 1–9 (1961).
64. L. S. B. Leakey, An early Miocene member of Hominidae. *Nature* **213**, 155–163 (1967). doi: 10.1038/213155a0; pmid: 6030571
65. E. Zuckerkandl, R. T. Jones, L. Pauling, A comparison of animal hemoglobins by tryptic peptide pattern analysis. *Proc. Natl. Acad. Sci. U.S.A.* **46**, 1349–1360 (1960). doi: 10.1073/pnas.46.10.1349; pmid: 16590757
66. M. Goodman, Immunochemistry of the primates and primate evolution. *Ann. N. Y. Acad. Sci.* **102**, 219–234 (1962). doi: 10.1111/j.1749-6632.1962.tb13641.x; pmid: 13949097
67. V. M. Sarich, A. C. Wilson, Immunological time scale for hominid evolution. *Science* **158**, 1200–1203 (1967). doi: 10.1126/science.158.3805.1200; pmid: 4964406
68. S. L. Washburn, Behaviour and the origin of man. *Proc. R. Anthropol. Inst. G. B. Irel.* **1967**, 21–27 (1967).
69. L. S. B. Leakey, The relationship of African apes, man and old world monkeys. *Proc. Natl. Acad. Sci. U.S.A.* **67**, 746–748 (1970). doi: 10.1073/pnas.67.2.746; pmid: 5002096
70. W. Hennig, *Phylogenetic Systematics* (Univ. Illinois Press, 1966).
71. E. Delson, N. Eldredge, I. Tattersall, Reconstruction of hominid phylogeny: A testable framework based on cladistic analysis. *J. Hum. Evol.* **6**, 263–278 (1977). doi: 10.1016/S0047-2484(77)80051-1
72. J. J. Yunis, O. Prakash, The origin of man: A chromosomal pictorial legacy. *Science* **215**, 1525–1530 (1982). doi: 10.1126/science.7063861; pmid: 7063861
73. C. G. Sibley, J. E. Ahlquist, The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridization. *J. Mol. Evol.* **20**, 2–15 (1984). doi: 10.1007/BF02101980; pmid: 6429338
74. M. Goodman, G. Braunitzer, A. Stangl, B. Schrank, Evidence on human origins from haemoglobins of African apes. *Nature* **303**, 546–548 (1983). doi: 10.1038/303546a0; pmid: 6406908
75. P. Andrews, L. Martin, Cladistic relationships of extant and fossil hominoids. *J. Hum. Evol.* **16**, 101–118 (1987). doi: 10.1016/0047-2484(87)90062-5
76. M. M. Miyamoto, J. L. Slightom, M. Goodman, Phylogenetic relations of humans and African apes from DNA sequences in the psi eta-globin region. *Science* **238**, 369–373 (1987). doi: 10.1126/science.3116671; pmid: 3116671
77. M. Ruvolo, Molecular phylogeny of the hominoids: Inferences from multiple independent DNA sequence data sets. *Mol. Biol. Evol.* **14**, 248–265 (1997). doi: 10.1093/oxfordjournals.molbev.a025761; pmid: 9066793
78. A. L. Zihlman, J. E. Cronin, D. L. Cramer, V. M. Sarich, Pygmy chimpanzee as a possible prototype for the common ancestor of humans, chimpanzees and gorillas. *Nature* **275**, 744–746 (1978). doi: 10.1038/275744a0; pmid: 703839
79. J. Goodall, Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature* **201**, 1264–1266 (1964). doi: 10.1038/2011264a0; pmid: 14151401
80. R. Tuttle, in *Phylogeny of the Primates*, W. P. Luckett, F. S. Szalay, Eds. (Springer, 1975), pp. 447–480.
81. P. Andrews, T. Harrison, in *Interpreting the Past: Essays on Human, Primate, and Mammal Evolution*, D. E. Lieberman, R. J. Smith, J. Kelley, Eds. (Brill Academic, 2005).
82. K. P. McNulty, Apes and tricksters: The evolution and diversification of humans' closest relatives. *Evolution (N. Y.)* **3**, 322–332 (2010). doi: 10.1007/s12052-010-0251-z
83. D. R. Begun, *The Real Planet of the Apes: A New Story of Human Origins* (Princeton Univ. Press, 2016).
84. D. Pilbeam, New hominoid skull material from the Miocene of Pakistan. *Nature* **295**, 232–234 (1982). doi: 10.1038/295232a0; pmid: 6799831
85. P. Andrews, J. E. Cronin, The relationships of *Sivapithecus* and *Ramapithecus* and the evolution of the orang-utan. *Nature* **297**, 541–546 (1982). doi: 10.1038/297541a0; pmid: 7045678
86. D. Pilbeam, M. D. Rose, J. C. Barry, S. M. I. Shah, New *Sivapithecus* humeri from Pakistan and the relationship of *Sivapithecus* and *Pongo*. *Nature* **348**, 237–239 (1990). doi: 10.1038/348237a0; pmid: 2234091
87. B. A. Williams, Comparing levels of homoplasy in the primate skeleton. *J. Hum. Evol.* **52**, 480–489 (2007). doi: 10.1016/j.jhevol.2006.11.011; pmid: 17391731
88. G. E. Erikson, Brachiation in New World monkeys and in anthropoid apes. *Symp. Zool. Soc. London* **10**, 135–163 (1963).
89. S. Almécija *et al*., The femur of *Orrorin tugenensis* exhibits morphometric affinities with both Miocene apes and later hominins. *Nat. Commun.* **4**, 2888 (2013). doi: 10.1038/ncomms3888; pmid: 24301078
90. A. S. Hammond, J. M. Plavcan, C. V. Ward, A validated method for modeling anthropoid hip abduction in silico. *Am. J. Phys. Anthropol.* **160**, 529–548 (2016). doi: 10.1002/ajpa.22990; pmid: 27088216
91. P. Andrews, C. P. Goves, Gibbons and brachiation. *Gibbon and Siamang* **4**, 167–218 (1976).
92. K. D. Hunt, Why are there apes? Evidence for the co-evolution of ape and monkey ecomorphology. *J. Anat.* **228**, 630–685 (2016). doi: 10.1111/joa.12454; pmid: 27004976
93. C. O. Lovejoy, G. Suwa, S. W. Simpson, J. H. Matternes, T. D. White, The great divides: *Ardipithecus ramidus* reveals the postcrania of our last common ancestors with African apes. *Science* **326**, 100–106 (2009). doi: 10.1126/science.1175833; pmid: 19810199
94. D. M. Alba, Fossil apes from the Vallès-Penedès Basin. *Evol. Anthropol.* **21**, 254–269 (2012). doi: 10.1002/evan.21312; pmid: 23280922
95. M. Dainton, G. A. Macho, Did knuckle walking evolve twice? *J. Hum. Evol.* **36**, 171–194 (1999). doi: 10.1006/jhev.1998.0265; pmid: 10068065
96. T. Harrison, Anthropology. Apes among the tangled branches of human origins. *Science* **327**, 532–534 (2010). doi: 10.1126/science.1184703; pmid: 20110491
97. K. D. Pugh, "The phylogenetic relationships of Middle-Late Miocene apes: Implications for early human evolution," thesis, The Graduate Center, City University of New York (2020).
98. Y. Chaimanee, V. Lazzari, K. Chaivanich, J.-J. Jaeger, First maxilla of a late Miocene hominid from Thailand and the evolution of pongine derived characters. *J. Hum. Evol.* **134**, 102636 (2019). doi: 10.1016/j.jhevol.2019.06.007; pmid: 31446972
99. J. Kelley, F. Gao, Juvenile hominoid cranium from the late Miocene of southern China and hominoid diversity in Asia. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6882–6885 (2012). doi: 10.1073/pnas.1201330109; pmid: 22511723
100. D. L. Gebo *et al*., A hominoid genus from the early Miocene of Uganda. *Science* **276**, 401–404 (1997). doi: 10.1126/science.276.5311.401; pmid: 9103195
101. M. Pickford, P. Mein, Early Middle Miocene mammals from Moroto II, Uganda. *Beiträge der Paläontologie* **30**, 361–386 (2006).
102. L. MacLatchy, The oldest ape. *Evol. Anthropol.* **13**, 90–103 (2004). doi: 10.1002/evan.10133
103. M. Nakatsukasa, Y. Kunimatsu, *Nacholapithecus* and its importance for understanding hominoid evolution. *Evol. Anthropol.* **18**, 103–119 (2009). doi: 10.1002/evan.20208
104. M. Böhme *et al*., A new Miocene ape and locomotion in the ancestor of great apes and humans. *Nature* **575**, 489–493 (2019). doi: 10.1038/s41586-019-1731-0; pmid: 31695194
105. S. Moyà-Solà, M. Köhler, A *Dryopithecus* skeleton and the origins of great-ape locomotion. *Nature* **379**, 156–159 (1996). doi: 10.1038/379156a0; pmid: 8538764
106. S. Moyà-Solà *et al*., First partial face and upper dentition of the Middle Miocene hominoid *Dryopithecus fontani* from Abocador de Can Mata (Vallès-Penedès Basin, Catalonia, NE Spain): Taxonomic and phylogenetic implications. *Am. J. Phys. Anthropol.* **139**, 126–145 (2009). doi: 10.1002/ajpa.20891; pmid: 19278017
107. B. R. Benefit, M. L. McCrossin, Miocene hominoids and hominid origins. *Annu. Rev. Anthropol.* **24**, 237–256 (1995). doi: 10.1146/annurev.an.24.100195.001321
108. D. Pilbeam, Genetic and morphological records of the Hominoidea and hominid origins: A synthesis. *Mol. Phylogenet. Evol.* **5**, 155–168 (1996). doi: 10.1006/mpev.1996.0010; pmid: 8673283
109. D. M. Alba *et al*., Miocene small-bodied ape from Eurasia sheds light on hominoid evolution. *Science* **350**, aab2625 (2015). doi: 10.1126/science.aab2625; pmid: 26516285
110. L. de Bonis, G. Bouvrain, D. Geraads, G. Koufost, New hominid skull material from the late Miocene of Macedonia in northern Greece. *Nature* **345**, 712–714 (1990). doi: 10.1038/345712a0; pmid: 2193230
111. B. Senut *et al*., First hominid from the Miocene (Lukeino Formation, Kenya). *C. R. Acad. Sci. Paris* **332**, 137–144 (2001). doi: 10.1016/S1251-8050(01)01529-4
112. B. Wood, T. Harrison, The evolutionary context of the first hominins. *Nature* **470**, 347–352 (2011). doi: 10.1038/nature09709; pmid: 21331035
113. M. Pickford, B. Senut, Hominoid teeth with chimpanzee-and gorilla-like features from the Miocene of Kenya: Implications for the chronology of ape-human divergence and biogeography of Miocene hominoids. *Anthropol. Sci.* **113**, 95–102 (2005). doi: 10.1537/ase.04S014
114. S. McBrearty, N. G. Jablonski, First fossil chimpanzee. *Nature* **437**, 105–108 (2005). doi: 10.1038/nature04008; pmid: 16136135
115. J. DeSilva, E. Shoreman, L. MacLatchy, A fossil hominoid proximal femur from Kikorongo Crater, southwestern Uganda. *J. Hum. Evol.* **50**, 687–695 (2006). doi: 10.1016/j.jhevol.2006.01.008; pmid: 16620913
116. H. Ishida, M. Pickford, H. Nakaya, Y. Nakano, Fossil anthropoids from Nachola and Samburu hills, Samburu district, Kenya. *Afr. Study Monogr.* **2**, 73–85 (1984).
117. D. R. Begun, in *Phylogeny of the Neogene Hominoid Primates of Eurasia*, vol. 2 of *Hominoid Evolution and Climatic Change in Europe*, L. de Bonis, G. D. Koufos, P. Andrews, Eds. (Cambridge Univ. Press, 2001), pp. 231–253.
118. P. Andrews, J. Kelley, Middle Miocene dispersals of apes. *Folia Primatol.* **78**, 328–343 (2007). doi: 10.1159/000105148; pmid: 17855786
119. S. M. Solà, M. Köhler, Recent discoveries of *Dryopithecus* shed new light on evolution of great apes. *Nature* **365**, 543–545 (1993). doi: 10.1038/365543a0; pmid: 8413607
120. S. M. Cote, Origins of the African hominoids: An assessment of the palaeobiogeographical evidence. *C. R. Palevol* **3**, 323–340 (2004). doi: 10.1016/j.crpv.2004.03.006
121. W. E. Le Gros Clark, *The Fossil Evidence for Human Evolution* (Univ. Chicago Press, 1964).
122. J. T. J. Stern, Climbing to the top: A personal memoir of *Australopithecus afarensis*. *Evol. Anthropol.* **9**, 113–133 (2000). doi: 10.1002/1520-6505(2000)9:3<113::AID-EVAN2>3.0.CO;2-W
123. C. V. Ward, Interpreting the posture and locomotion of *Australopithecus afarensis*: Where do we stand? *Am. J. Phys. Anthropol.* **119** (suppl. 35), 185–215 (2002). doi: 10.1002/ajpa.10185; pmid: 12653313
124. M. Böhme, N. Spassov, J. M. DeSilva, D. R. Begun, Reply to: Reevaluating bipedalism in *Danuvius*. *Nature* **586**, E4–E5 (2020). doi: 10.1038/s41586-020-2736-4; pmid: 32999478
125. N. E. Thompson, S. Almécija, The evolution of vertebral formulae in Hominoidea. *J. Hum. Evol.* **110**, 18–36 (2017). doi: 10.1016/j.jhevol.2017.05.012; pmid: 28778460
126. C. V. Ward, A. Walker, M. F. Teaford, I. Odhiambo, Partial skeleton of *Proconsul nyanzae* from Mfangano Island, Kenya. *Am. J. Phys. Anthropol.* **90**, 77–111 (1993). doi: 10.1002/ajpa.1330900106; pmid: 8470757
127. E. Delson, An anthropoid enigma: Historical introduction to the study of *Oreopithecus bambolii*. *J. Hum. Evol.* **15**, 523–531 (1986). doi: 10.1016/S0047-2484(86)80071-9
128. I. Nengo *et al*., New infant cranium from the African Miocene sheds light on ape evolution. *Nature* **548**, 169–174 (2017). doi: 10.1038/nature23456; pmid: 28796200
129. A. S. Hammond *et al*., Insights into the lower torso in late Miocene hominoid *Oreopithecus bambolii*. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 278–284 (2020). doi: 10.1073/pnas.1911896116; pmid: 31871170
130. M. Köhler, S. Moyà-Solà, Ape-like or hominid-like? The positional behavior of Oreopithecus bambolii reconsidered. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11747–11750 (1997). doi: 10.1073/pnas.94.21.11747; pmid: 9326682
131. J. D. Currey, *Bones: Structure and Mechanics* (Princeton Univ. Press, 2002).
132. C. Ruff, B. Holt, E. Trinkaus, Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *Am. J. Phys. Anthropol.* **129**, 484–498 (2006). doi: 10.1002/ajpa.20371; pmid: 16425178
133. T. L. Kivell, A review of trabecular bone functional adaptation: What have we learned from trabecular analyses in extant hominoids and what can we apply to fossils? *J. Anat.* **228**, 569–594 (2016). doi: 10.1111/joa.12446; pmid: 26879841
134. B. Demes *et al*., Patterns of strain in the macaque ulna during functional activity. *Am. J. Phys. Anthropol.* **106**, 87–100 (1998). doi: 10.1002/(SICI)1096-8644(199805)106:1<87::AID-AJPA6>3.0.CO;2-A; pmid: 9590526

135. L. E. Lanyon, The success and failure of the adaptive response to functional load-bearing in averting bone fracture. *Bone* **13** (suppl. 2), S17–S21 (1992). doi: 10.1016/8756-3282(92)90191-X; pmid: 1627409
136. I. J. Wallace, S. Judex, B. Demes, Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone* **72**, 1–8 (2015). doi: 10.1016/j.bone.2014.11.013; pmid: 25460574
137. J. P. P. Saers, Y. Cazorla-Bak, C. N. Shaw, J. T. Stock, T. M. Ryan, Trabecular bone structural variation throughout the human lower limb. *J. Hum. Evol.* **97**, 97–108 (2016). doi: 10.1016/j.jhevol.2016.05.012; pmid: 27457548
138. Z. Wood *et al*., Are we crying Wolff? 3D printed replicas of trabecular bone structure demonstrate higher stiffness and strength during off-axis loading. *Bone* **127**, 635–645 (2019). doi: 10.1016/j.bone.2019.08.002; pmid: 31390534
139. J. G. Fleagle *et al*., Climbing: A biomechanical link with brachiation and with bipedalism. *Symp. Zool. Soc. London* **48**, 359–375 (1981).
140. S. K. S. Thorpe, R. L. Holder, R. H. Crompton, Origin of human bipedalism as an adaptation for locomotion on flexible branches. *Science* **316**, 1328–1331 (2007). doi: 10.1126/science.1140799; pmid: 17540902
141. M. A. Panger, A. S. Brooks, B. G. Richmond, B. Wood, Older than the Oldowan? Rethinking the emergence of hominin tool use. *Evol. Anthropol.* **11**, 235–245 (2003). doi: 10.1002/evan.10094
142. E. J. McNutt, B. Zipfel, J. M. DeSilva, The evolution of the human foot. *Evol. Anthropol.* **27**, 197–217 (2018). doi: 10.1002/evan.21713; pmid: 30242943
143. A. S. Hammond, S. Almécija, Lower ilium evolution in apes and hominins. *Anat. Rec.* **300**, 828–844 (2017). doi: 10.1002/ar.23545; pmid: 28406561
144. M. A. Maslin *et al*., East African climate pulses and early human evolution. *Quat. Sci. Rev.* **101**, 1–17 (2014). doi: 10.1016/j.quascirev.2014.06.012
145. I. Fer, B. Tietjen, F. Jeltsch, M. H. Trauth, Modelling vegetation change during Late Cenozoic uplift of the East African plateaus. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **467**, 120–130 (2017). doi: 10.1016/j.palaeo.2016.04.007
146. M. Domínguez-Rodrigo, Is the "savanna hypothesis" a dead concept for explaining the emergence of the earliest hominins? *Curr. Anthropol.* **55**, 59–81 (2014). doi: 10.1086/674530
147. K. D. Hunt, The evolution of human bipedality: Ecology and functional morphology. *J. Hum. Evol.* **26**, 183–202 (1994). doi: 10.1006/jhev.1994.1011
148. R. Potts, Environmental hypotheses of hominin evolution. *Am. J. Phys. Anthropol.* **107** (suppl. 27), 93–136 (1998). doi: 10.1002/(SICI)1096-8644(1998)107:27+<93::AID-AJPA5>3.0.CO;2-X; pmid: 9881524
149. S. Harmand *et al*., 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature* **521**, 310–315 (2015). doi: 10.1038/nature14464; pmid: 25993961
150. D. M. Alba, S. Moyà-Solà, M. Köhler, Morphological affinities of the *Australopithecus afarensis* hand on the basis of manual proportions and relative thumb length. *J. Hum. Evol.* **44**, 225–254 (2003). doi: 10.1016/S0047-2484(02)00207-5; pmid: 12662944
151. R. Potts, Paleoenvironmental basis of cognitive evolution in great apes. *Am. J. Primatol.* **62**, 209–228 (2004). doi: 10.1002/ajp.20016; pmid: 15027093
152. J. Zachos, M. Pagani, L. Sloan, E. Thomas, K. Billups, Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001). doi: 10.1126/science.1059412; pmid: 11326091
153. F. Kaya *et al*., The rise and fall of the Old World savannah fauna and the origins of the African savannah biome. *Nat. Ecol. Evol.* **2**, 241–246 (2018). doi: 10.1038/s41559-017-0414-1; pmid: 29292396
154. N. G. Jablonski, M. J. Whitfort, N. Roberts-Smith, X. Qinqi, The influence of life history and diet on the distribution of catarrhine primates during the Pleistocene in eastern Asia. *J. Hum. Evol.* **39**, 131–157 (2000). doi: 10.1006/jhev.2000.0405; pmid: 10968926
155. P. Mitteroecker, P. Gunz, M. Bernhard, K. Schaefer, F. L. Bookstein, Comparison of cranial ontogenetic trajectories among great apes and humans. *J. Hum. Evol.* **46**, 679–697 (2004). doi: 10.1016/j.jhevol.2004.03.006; pmid: 15183670
156. S. E. Inouye, Ontogeny and allometry of African ape manual rays. *J. Hum. Evol.* **23**, 107–138 (1992). doi: 10.1016/0047-2484(92)90103-G
157. A. D. Barnosky *et al*., Approaching a state shift in Earth's biosphere. *Nature* **486**, 52–58 (2012). doi: 10.1038/nature11018; pmid: 22678279
158. A. D. Barnosky *et al*., Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science* **355**, eaah4787 (2017). doi: 10.1126/science.aah4787; pmid: 28183912
159. P. A. Goloboff, A. Torres Galvis, J. S. Arias Becerra, Parsimony and model-based phylogenetic methods for morphological data: Comments on O'Reilly *et al*. *Palaeontology* **61**, 625–630 (2018). doi: 10.1111/pala.12353
160. S. A. Catalano, M. D. Ercoli, F. J. Prevosti, The more, the better: The use of multiple landmark configurations to solve the phylogenetic relationships in musteloids. *Syst. Biol.* **64**, 294–306 (2015). doi: 10.1093/sysbio/syu107; pmid: 25516268
161. T. van der Valk *et al*., Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* **591**, 265–269 (2021). doi: 10.1038/s41586-021-03224-9; pmid: 33597750
162. F. Welker *et al*., Enamel proteome shows that *Gigantopithecus* was an early diverging pongine. *Nature* **576**, 262–265 (2019). doi: 10.1038/s41586-019-1728-8; pmid: 31723270
163. S. Washburn, Human evolution: Science or game. *Yearb. Phys. Anthropol.* **17**, 67–70 (1973).
164. W. I. Sellers, E. Hirasaki, Markerless 3D motion capture for animal locomotion studies. *Biol. Open* **3**, 656–668 (2014). doi: 10.1242/bio.20148086; pmid: 24972869
165. J. A. Nyakatura *et al*., Reverse-engineering the locomotion of a stem amniote. *Nature* **565**, 351–355 (2019). doi: 10.1038/s41586-018-0851-2; pmid: 30651613
166. M. Landau, *Narratives of Human Evolution* (Yale Univ. Press, 1993).
167. R. J. Smith, Explanations for adaptations, just-so stories, and limitations on evidence in evolutionary biology. *Evol. Anthropol.* **25**, 276–287 (2016). doi: 10.1002/evan.21495; pmid: 28004894
168. M. S. Springer *et al*., Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. *PLOS ONE* **7**, e49521 (2012). doi: 10.1371/journal.pone.0049521; pmid: 23166696
169. M. Goodman *et al*., Primate evolution at the DNA level and a classification of hominoids. *J. Mol. Evol.* **30**, 260–266 (1990). doi: 10.1007/BF02099995; pmid: 2109087

ACKNOWLEDGMENTS

We thank the many colleagues that motivated and shaped this review through their own work on the "Miocene ape-hominin transition." In particular, we highlight the decades-long work of P. Andrews, D. Begun, B. Benefit, T. Harrison, B. Jungers, J. Kelley, O. Lovejoy, L. MacLatchy, M. Nakatsukasa, M. McCrossin, M. Pickford, D. Pilbeam, B. Senut, J. Stern, C. Ward, and T. White. E. Delson and S. Catalano provided constructive criticisms on an earlier version of the manuscript. K. Younkin assisted by making Fig. 2. **Funding:** This research has been funded by the Agencia Estatal de Investigación (CGL2016-76431-P and CGL2017-82654-P, AEI/FEDER EU) and the Generalitat de Catalunya (CERCA Programme and consolidated research groups 2017 SGR 86 and 2017 SGR 116 GRC). **Competing interests:** The authors declare no competing interests.

10.1126/science.abb4363

RESEARCH ARTICLE SUMMARY

BACTERIAL PHYLOGENY

A rooted phylogeny resolves early bacterial evolution

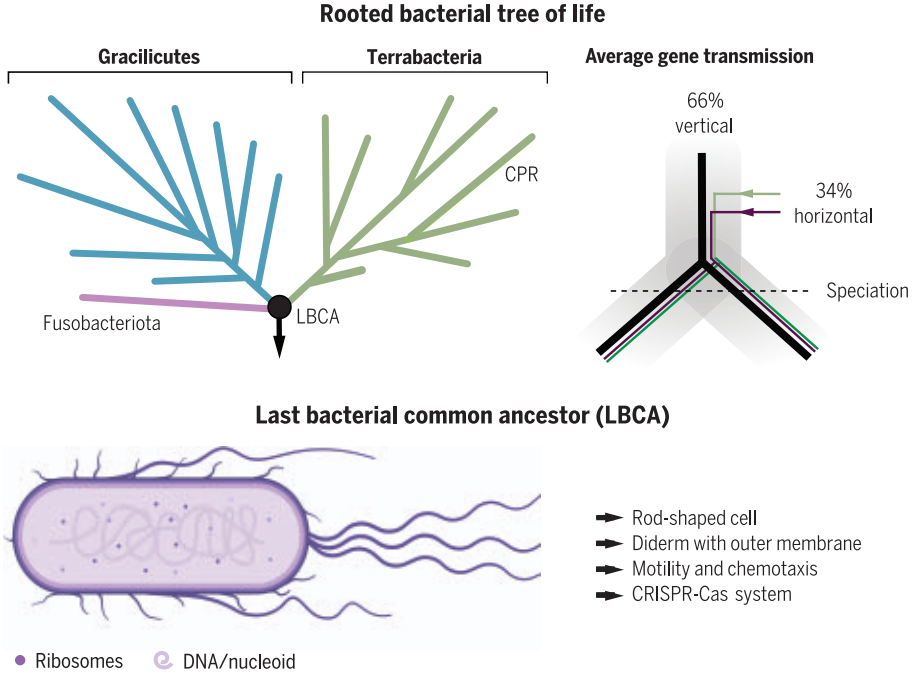
Gareth A. Coleman†, Adrián A. Davín†, Tara A. Mahendrarajah, Lénárd L. Szánthó, Anja Spang, Philip Hugenholtz‡*, Gergely J. Szöllősi‡*, Tom A. Williams‡*

INTRODUCTION: Bacteria are the most diverse and abundant cellular organisms on Earth, and in recent years environmental genomics has revealed the existence of an enormous diversity of previously unknown lineages. Despite the abundance of genomic sequence data, the root of the bacterial tree and the nature of the most recent common ancestor of Bacteria have remained elusive. The problem is that even with the help of new data, tracing billions of years of bacterial evolution back to the root has remained challenging because standard phylogenetic models do not account for the full range of evolutionary processes that shape bacterial genomes. In particular, standard models treat horizontal gene transfer as an impediment to the reconstruction of the tree of life that must be removed from analyses. But if horizontal gene transfer is modeled appropriately, it can provide information about the deep past that is unavailable to standard methods.

RATIONALE: We reconstructed and rooted the bacterial tree by applying a hierarchical phylogenomic approach that explicitly uses information from gene duplications and losses within a genome as well as gene transfers between genomes. This approach allowed us to root the tree without including an archaeal outgroup. Outgroup-free rooting is a promising approach for Bacteria, both because the position of the universal root is uncertain and because the long branch separating Bacteria from Archaea has the potential to distort the reconstruction of within-Bacteria relationships. Outgroup-free gene tree-species tree reconciliation allowed us to quantitatively model both the vertical and horizontal components of bacterial evolution and integrate information from 11,272 gene families to resolve the root of the bacterial tree. Notably, these analyses also provided estimates of the gene content of the last bacterial common ancestor.

RESULTS: Our analyses place the root between two major bacterial clades, the Gracilicutes and Terrabacteria. We found no support for a root between the Candidate Phyla Radiation (CPR), a lineage comprising putative symbionts and parasites with small genomes, and all other Bacteria. Instead, the CPR was inferred to be a member of the Terrabacteria and formed a sister lineage to the Chloroflexota and Dormibacterota. This suggests that the CPR evolved by reductive genome evolution from free-living ancestors. Gene families inferred to have been present at the root indicate that the last bacterial common ancestor was already a complex double-membraned cell capable of motility and chemotaxis that possessed a CRISPR-Cas system. Although ~92% of gene families have experienced horizontal transfers during their history, tracing their evolution along the most likely rooted tree revealed that about two-thirds of gene transmissions have been vertical. Thus, bacterial evolution has a major vertical component, despite a profound impact of horizontal gene transfer through time. Horizontal gene flows can also provide insight into the temporal sequence of events during bacterial diversification, because donor lineages must be at least as old as recipients. Analysis of gene transfers in our dataset suggests that the diversification of the Firmicutes, CPR, Acidobacteriota, and Proteobacteria is the oldest among extant bacterial phyla.

CONCLUSION: The vertical and horizontal components of genome evolution provide complementary sources of information about bacterial phylogeny. The vertical component provides a robust framework for interpreting species diversity and allows us to reconstruct ancestral states, while the horizontal component helps to root the vertical tree and orient it in time. The inferred Gracilicutes-Terrabacteria root will be useful for investigating the tempo and mode of bacterial diversification, metabolic innovation, and changes in cell architecture such as the evolutionary transitions between double (diderm) and single (monoderm) membranes. Future development of methods that harness the complementarity of vertical and horizontal gene transmission will continue to further our understanding of life on Earth. ■



A rooted phylogeny of Bacteria. The reconciliation of bacterial gene phylogenies places the root between the major clades of Gracilicutes (including Proteobacteria and Bacteroidota) and Terrabacteria (including Firmicutes and Cyanobacteria). On the basis of this hypothesis, ancestral genome reconstruction predicts that the last bacterial common ancestor (LBCA) was a complex, double-membraned cell and that, on average, two-thirds of gene transmissions have been vertically inherited along this rooted tree.

The list of author affiliations is available in the full article online.
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
*Corresponding author. Email: p.hugenholtz@uq.edu.au (P.H.); ssolo@elte.hu (G.J.Sz.); tom.a.williams@bristol.ac.uk (T.A.W.)
Cite this article as G. A. Coleman *et al.*, *Science* 372, eabe0511 (2021). DOI: 10.1126/science.abe0511

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abe0511

RESEARCH ARTICLE

BACTERIAL PHYLOGENY

A rooted phylogeny resolves early bacterial evolution

Gareth A. Coleman[1]†, Adrián A. Davín[2]†, Tara A. Mahendrarajah[3], Lénárd L. Szánthó[4,5], Anja Spang[3,6], Philip Hugenholtz[2]‡*, Gergely J. Szöllősi[4,5,7]‡*, Tom A. Williams[1]‡*

A rooted bacterial tree is necessary to understand early evolution, but the position of the root is contested. Here, we model the evolution of 11,272 gene families to identify the root, extent of horizontal gene transfer (HGT), and the nature of the last bacterial common ancestor (LBCA). Our analyses root the tree between the major clades Terrabacteria and Gracilicutes and suggest that LBCA was a free-living flagellated, rod-shaped double-membraned organism. Contrary to recent proposals, our analyses reject a basal placement of the Candidate Phyla Radiation, which instead branches sister to Chloroflexota within Terrabacteria. While most gene families (92%) have evidence of HGT, overall, two-thirds of gene transmissions have been vertical, suggesting that a rooted tree provides a meaningful frame of reference for interpreting bacterial evolution.

A species tree captures the relationships among organisms but requires a root to provide the direction of evolution. Rooting deep radiations (*1*) is among the greatest challenges in phylogenetics, and there is no consensus on the root of the bacterial tree. On the basis of evidence (*2*–*5*) that the root of the tree of life lies between Bacteria and Archaea, early analyses with an archaeal outgroup placed the bacterial root near Aquificales and Thermotogales (*6*, *7*) or Planctomycetes (*8*). Alternative approaches, including analyses of gene flows and the evolution of multimeric protein complexes as well as other complex characters (*9*), have instead suggested roots within the monoderm (single-membrane) Bacteria (*10*) or between Chloroflexi and all other cellular life (*9*). The development of techniques for sequencing microbes directly from environmental samples, without the need for laboratory cultivation, has greatly expanded the genomic representation of natural prokaryotic diversity (*11*–*14*). Recent phylogenomic analyses of expanded sets of diverse bacteria have placed the root between one of the recently identified groups, the Candidate Phyla Radiation [CPR, also known as Patescibacteria (*15*, *16*)] and all other Bacteria (*11*, *16*, *17*). The CPR is characterized by small cells and genomes that appear to have predominantly symbiotic or parasitic lifestyles, but much remains to be learned about their ecology and physiology (*15*, *17*–*19*). If correct, the early divergence of the CPR has important implications for our understanding of the earliest period of cellular evolution. Along with evidence that the root of the archaeal domain lies between the reduced and predominantly host-associated DPANN superphylum (originally named after Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota) and all other Archaea (*1*, *20*), the CPR root implies that streamlined, metabolically minimalist prokaryotes have coexisted with the more familiar, self-sufficient lineages throughout the history of cellular life (*19*, *21*).

Improved taxon sampling can help to resolve evolutionary relationships (*22*, *23*), and the quantity and diversity of genome sequence data now available presents an opportunity to address long-standing questions about the origins and diversification of Bacteria. However, deep phylogenetic divergences are difficult to resolve, both because the phylogenetic signal for such relationships is overwritten by new changes over time, and also because the process of sequence evolution is more complex than the best-fitting models currently available. In particular, variation in nucleotide or amino acid composition across the sites of the alignment and the branches of the tree can induce long branch attraction (LBA) artifacts in which deep-branching, fast-evolving, poorly sampled or compositionally biased lineages group together irrespective of their evolutionary history (*24*). These issues are widely appreciated (*11*) but are challenging to address adequately, particularly when sequences from thousands of taxa (*11*, *13*, *14*, *16*, *17*) are used to estimate trees of global prokaryotic diversity, which precludes the use of the best-fitting phylogenetic methods available.

[1]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK. [2]Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia. [3]Department of Marine Microbiology and Biogeochemistry, NIOZ, Royal Netherlands Institute for Sea Research, 1790 AB Den Burg, Netherlands. [4]Department of Biological Physics, Eötvös Loránd University, 1117 Budapest, Hungary. [5]MTA-ELTE "Lendület" Evolutionary Genomics Research Group, 1117 Budapest, Hungary. [6]Department of Cell- and Molecular Biology, Uppsala University, SE-75123 Uppsala, Sweden. [7]Institute of Evolution, Centre for Ecological Research, 1121 Budapest, Hungary.
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
***Corresponding author. Email: p.hugenholtz@uq.edu.au (P.H.); ssolo@elte.hu (G.J.Sz.); tom.a.williams@bristol.ac.uk (T.A.W.)**

Archaeal outgroup rooting does not unambiguously establish the root of the bacterial tree

The standard approach to rooting is to include an outgroup in the analysis, and all published phylogenies in which CPR forms a sister lineage to the rest of the Bacteria (*11*, *16*, *17*) have made use of an archaeal outgroup. Outgroup rooting on the bacterial tree, however, has three serious limitations. First, interpretation of the results requires the assumption that the root of the tree of life lies between Bacteria and Archaea. While this is certainly the consensus view, the available evidence is limited and difficult to interpret (*2*–*5*, *25*), and alternative hypotheses in which the universal root is placed within Bacteria have been proposed on the basis of indels (*26*, *27*) or the analysis of slow-evolving characters (*9*). Second, the long branch leading to the archaeal outgroup has the potential to distort within-Bacteria relationships because of LBA. Third, joint analyses of Archaea and Bacteria use a smaller number of genes that are widely conserved and have evolved vertically since the divergence of the two lineages, and sequence alignment is more difficult owing to the low sequence identity between homologs of the two domains.

We evaluated the performance of outgroup rooting on a bacterial tree using 265 Bacteria (see below) and 149 Archaea from a shared subset of 29 phylogenetic markers (table S1). Using this archaeal outgroup, the maximum likelihood (ML) phylogeny under the best-fitting model (LG+C60+R8+F, which accounts for site heterogeneity in the substitution process) placed the bacterial root between a clade comprising Cyanobacteria, Margulisbacteria, CPR, Chloroflexota, and Dormibacterota on one side of the root and all other taxa on the other (fig. S1). However, bootstrap support for this root, and indeed many other deep branches in both the bacterial and archaeal subtrees, was low (50 to 80%). We therefore used approximately unbiased (AU) tests (*28*) to determine whether a range of published alternative rooting hypotheses (table S2) could be rejected, given the model and data. The AU test asks whether the optimal trees that are consistent with these other hypotheses have a significantly worse likelihood score than the ML tree. In this case, the likelihoods of all tested trees were statistically indistinguishable (AU test, $P > 0.05$, table S2), indicating that outgroup rooting cannot resolve the bacterial root on this alignment.

An alternative to outgroup rooting for deep microbial phylogeny

Given the limitations of using a remote archaeal outgroup to establish the root of the bacterial tree, we explored outgroup-free rooting using gene tree-species tree reconciliation (*1*, *29*–*31*). We recently applied this approach to root the

archaeal tree (*1*), and similar approaches have been used to investigate the root of eukaryotes (*32*, *33*) and to map and characterize whole-genome duplications in plants (*34*). Gene tree-species tree reconciliation methods work by adding a layer to the standard framework for inferring trees from molecular data. This additional step models the way in which gene trees can differ from each other and the overarching rooted species tree. Substitution models [such as LG (*35*)] describe how the constituent sequences of a gene family evolve along a gene tree via a series of amino acid substitutions that allow us to infer the most likely gene tree. Reconciliation models describe how a gene tree evolves along the rooted species tree, beginning with gene birth (origination) and followed by a combination of vertical descent and events such as gene duplications, transfers, and losses (this series of events is known as a DTL reconciliation). Combining the substitution-based modeling of sequences along the gene tree with the reconciliation-based modeling of gene trees along a rooted species tree allows us to infer the most likely rooted species tree from the constituent gene families. In other words, reconciliation methods aggregate phylogenetic signal across gene families and, because the likelihood of reconciliations depends on the position of the root, can be used to test the support for competing root positions (*1*, *29*), providing a genome-wide (and gene transfer–aware) extension of the classical approach used to root the tree of life on the basis of ancient gene duplications (*3*, *4*).

Our method, amalgamated likelihood estimation (ALE), improves on earlier approaches by explicitly accounting for uncertainty in the gene tree topologies and in the events leading to those topologies while simultaneously estimating rates of gene duplication, transfer, and loss directly from the data (*31*). Simulations suggest that root inferences under ALE are robust to variation in taxon sampling and the proportion of extinct lineages (fig. S2), that the method finds the correct root even under high levels of gene transfer (*1*, *29*), and that the numbers of D, T, and L events are accurately recovered from the data (figs. S3 to S8). These results suggest that ALE is appropriate for the problem at hand (*36*).

Rooting Bacteria without an outgroup

To obtain an unrooted species tree for ALE analysis, we selected a focal dataset of 265 genomes representative of bacterial diversity according to the Genome Taxonomy Database (GTDB) (*13*). We inferred the tree from a concatenation of 62 conserved single-copy markers (table S1) using the LG+C60+R8+F model in IQ-Tree 1.6.10 (Fig. 1), which was chosen as the best-fitting model using the Bayesian information criterion (BIC) (*37*). This yielded highly congruent trees when removing 20 to 80% of the most compositionally heterogeneous sites from the alignment (fig. S9), suggesting that the key features of the topology are not composition-driven LBA artifacts. One exception was the position of the Fusobacteriota, which was recovered as a sister lineage to a clade comprising Deinococcota, Synergistota, and Thermotogota (DST) when 20% of the most heterogeneous sites were removed (fig. S9A) but was recovered as a single lineage between Terrabacteria plus DST and Gracilicutes in all other trees.

We used ALE to test the support for 62 root positions (tables S3 and S4) on the unrooted topology by reconciling gene trees for 11,272 homologous gene families [inferred using MCL (*38*)] from the 265 bacterial genomes. Note that this method does not assume that the root lies between Bacteria and Archaea. In addition to testing root positions corresponding to published hypotheses, we exhaustively tested all inner nodes of the tree above the phylum level. The ALE analysis rejected all of the root positions tested (AU test, $P < 0.05$) except for three adjacent branches, lying between the two major clades of Gracilicutes (comprising most diderm lineages) and Terrabacteria (comprising monoderm and atypical diderm lineages) (Fig. 1); the difference between the three root positions was the position of the Fusobacteriota in relation to these two major clades (Fig. 1B). Alternative roots were rejected with increasing confidence as distance from the optimal root region increased (Fig. 1C and table S3).

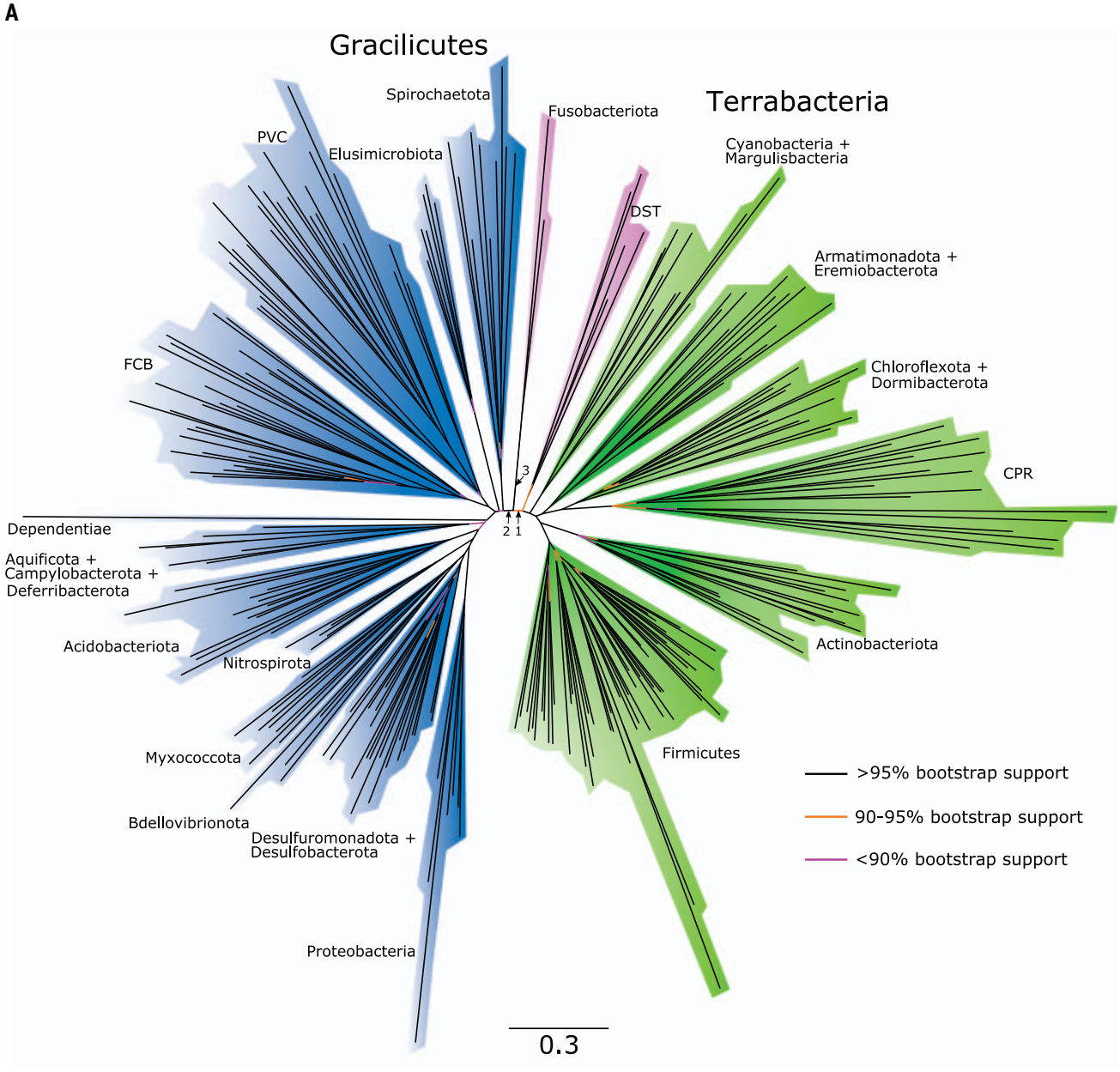
We tested the robustness of the inferred root region by (i) excluding gene families with extreme duplication, transfer, or loss rates; (ii) repeating the analysis using gene families constructed with an assignment to families in the Clusters of Orthologous Genes (COG) (*39*) ontology; and (iii) repeating the analysis on a secondary independent sampling of the tree, in which CPR makes up 40% of the genomes (*11*) (figs. S10 to S13 and table S5). These analyses consistently recovered the root between the Gracilicutes and Terrabacteria, regardless of the position of the Fusobacteriota. A Gracilicutes-Terrabacteria root was previously reported (*40*, *41*), but these studies did not include the CPR, which has recently been suggested to represent the earliest diverging bacterial lineage (*11*, *16*). Our outgroup-free analysis consistently recovered CPR nested within the Terrabacteria, as a sister clade to Chloroflexota and Dormibacterota, even with CPR representing more than 40% of the taxa included. This finding implies that the CPR evolved by genome reduction from a free-living ancestor, a scenario that has been proposed previously (*21*).

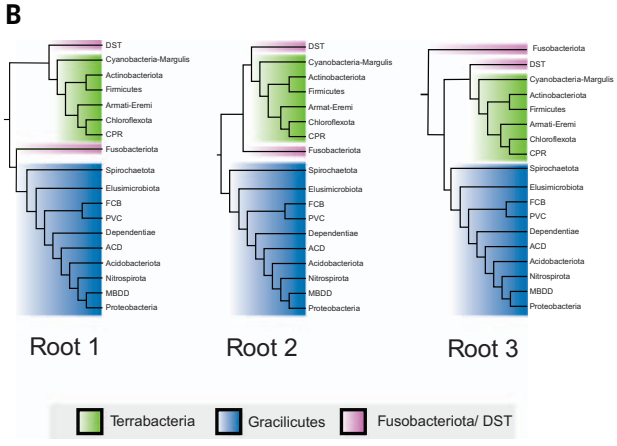
Transfers contain information about the relative timing of divergences, because for each transfer, the donor must be at least as old as the recipient (*42*, *43*). To establish the relative ages of the crown groups of different phyla, we used high-confidence relative age constraints recovered in at least 95 of 100 bootstrap replicates common to the focal and secondary datasets (*36*). Simulations suggest that this approach accurately recovers relative clade ages (98.4% accuracy on a simulated dataset the same size as the focal dataset, fig. S14). Our analysis (Fig. 2) predicts that the Firmicutes crown group is the oldest among extant bacterial phyla (median rank: 2 ± 1.43 SD) followed by the crown groups of the CPR (median rank: 3 ± 2), Proteobacteria (median rank: 3 ± 1.59), and Acidobacteriota (median rank: 3 ± 1.56), suggesting that these lineages were the earliest to diversify within Bacteria. The crown groups of lineages predominantly associated with animal hosts, Spirochaetota (median rank: 10 ± 0.85) and Elusimicrobiota (median rank: 11 ± 0.62), appear to be the youngest among extant phyla.

Is bacterial evolution treelike?

How much of bacterial evolution can be explained by the concept of a rooted species tree? Horizontal gene transfer (HGT) is frequent in prokaryotes, and published analyses indicate that most or all prokaryotic gene families have experienced HGT during their history (*1*, *44*). This implies that there is no single tree that fully describes the evolution of all bacterial genes or genomes (*45*, *46*). Extensive HGT is existentially challenging for concatenation, because it greatly curtails the number of genes that evolve on a single underlying tree (*47*). Phylogenetic networks (*46*, *48*) were the first methods to explicitly acknowledge non-vertical evolution, but they can be difficult to interpret biologically. Gene tree-species tree reconciliation unites tree and network-based approaches by modeling both the horizontal components of genome evolution (a fully reticulated network allowing all possible transfers) and the vertical trace (a common rooted species tree). This framework enables us to quantify the contributions of vertical and horizontal processes to bacterial evolutionary history.

Our analyses (Fig. 3) reveal that most bacterial gene families present in two or more species (9678 of 10,518 MCL families, or 92%) have experienced at least one gene transfer during their evolution; only very small families have escaped transfer entirely on the time scales considered here (fig. S15). Consistent with previous analyses (*1*, *49*), transfer rates vary across gene functional categories, with genes encoding proteins involved in defense mechanisms (such as antibiotic biosynthesis) and in the production of secondary metabolites being the most frequently transferred, and those coding for translational and cell cycle proteins the least (Fig. 3B). Despite this accumulation of HGT, most gene families evolve vertically the majority of the time, with


A
Gracilicutes
Terrabacteria
Spirochaetota
Fusobacteriota
PVC
Elusimicrobiota
Cyanobacteria +
Margulisbacteria
DST
Armatimonadota +
Eremiobacterota
FCB
Chloroflexota +
Dormibacterota
CPR
3
2 1
Dependentiae
Aquificota +
Campylobacterota +
Deferribacterota
Acidobacteriota
Nitrospirota
Actinobacteriota
Myxococcota
Firmicutes
Bdellovibrionota
Desulfuromonadota +
Desulfobacterota
Proteobacteria
>95% bootstrap support
90-95% bootstrap support
<90% bootstrap support
0.3


B
Root 1
Root 2
Root 3
Terrabacteria
Gracilicutes
Fusobacteriota/ DST

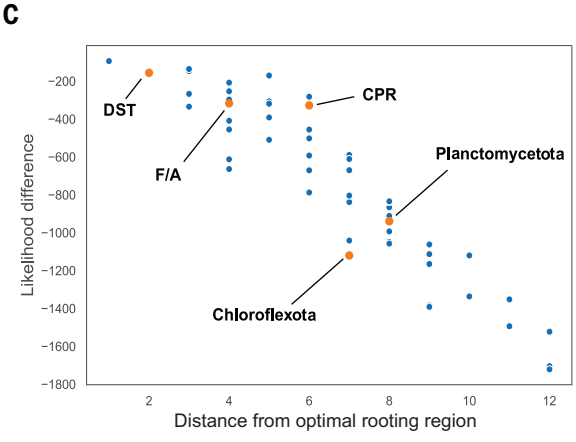

C
DST
F/A
CPR
Planctomycetota
Chloroflexota
Likelihood difference
Distance from optimal rooting region

Fig. 1. A rooted phylogeny of Bacteria. (**A**) We used gene tree-species tree reconciliation to infer the root of the bacterial tree. The unrooted maximum likelihood phylogeny was inferred from a concatenation of 62 marker genes under the best-fitting model, LG+C60+R8+F, which accounts for site heterogeneity in the substitution process and uses a mixture of eight substitution rates estimated from the data to model across-site evolutionary rate variation. Branches are colored according to bootstrap support value. The root falls between two major clades of Bacteria, the Gracilicutes and the Terrabacteria, on one of three statistically equivalent adjacent branches indicated by arrows, shown as rooted trees in (**B**). All alternative roots tested were rejected (tables S3 and S4), with likelihoods decreasing as a function of distance from the root region, as shown in (**C**). Previously proposed root positions, including the CPR root, are highlighted in red. FCB are the Fibrobacterota, Chlorobia, Bacteroidota, and related lineages; PVC are the Planctomycetota, Verrucomicrobiota, Chlamydiota, and related lineages; DST are the Deinococcota, Synergistota, and Thermotogota; ACD are Aquificota, Campylobacterota, and Deferribacterota; F/A are Firmicutes and Actinobacteriota; MBDD are Myxococcota, Bdellovibrionota, Desulfomonadota, and Desulfobacterota. Scale bar, 0.3 amino acid substitutions per site.

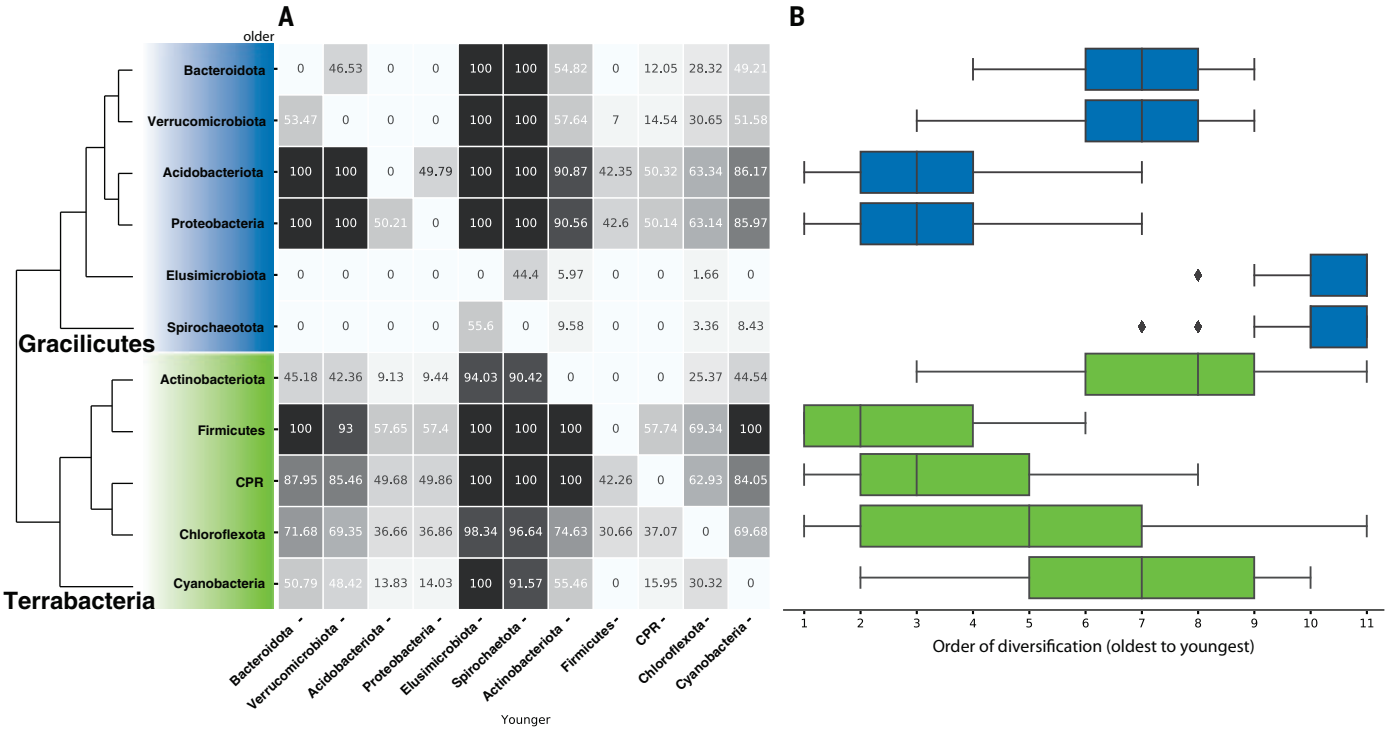


Fig. 2. Relative crown group ages of major bacterial phyla. Gene transfers that occurred during the diversification of Bacteria provide a record of the temporal sequence of events. We used the information provided by directional (donor-to-recipient) patterns of gene transfer to infer the relative ages of bacterial crown groups, focusing on phyla represented by at least five genomes in both of our datasets. To summarize this time information, we sampled 1000 time orders that were fully compatible with the constraints recovered from both datasets. (**A**) Pairwise relative ages of phyla. The proportion of sampled time orders in which each phylum on the *x* axis was recovered as younger than each phylum on the *y* axis. (**B**) Relative age distributions of major phyla. For each sampled time order, we ranked the phyla from oldest (1) to youngest (11) and plotted the distribution of the ranks. The crown group radiations of Firmicutes, CPR, Proteobacteria, and Acidobacteriota appear to be the oldest among sampled phyla, while those of Elusimicrobiota and Spirochaetota are the youngest.

mean verticality estimated to be 64% in the focal and 68% in the secondary dataset.

Genome-wide reconciliation of gene trees with the species tree demonstrates that the optimal rooted species tree provides an apt summary of much of bacterial evolutionary history, even for the deepest branches of the tree (*50*). From the gene's eye view, gene families evolve neither entirely vertically nor horizontally; core genes are occasionally transferred, and even frequently exchanged genes contribute useful vertical signal; for example, the median number of genes that evolve vertically on a branch of the species tree is 998.92 in the focal analysis (table S6), far greater than the number of genes that have been concatenated at the level of all Bacteria. From the perspective of the genome, constituent genes have different ages (or residence times), corresponding to the time at which they originated or were most recently acquired by gene transfer, within the resolution of our taxonomic sampling.

This analysis indicates that, on average, 82% of all genes from adequately represented phyla (five or more genomes) were most recently acquired after the diversification of that phylum, although all genomes retain a smaller proportion (10 to 27%) of genes that have descended vertically from the stem lineage of their phylum or even earlier (Fig. 3C). There are two explanations for this distribution of gene persistence times: (i) de novo gene origination within phyla (that is, lineage-specific gene families) and (ii) the cumulative impact of gene transfer, which curtails gene persistence times when looking back from the present day even though most transmissions are vertical.

Ancestral proteome of the last bacterial common ancestor

Reconciliation analyses not only allow us to infer the acquisition of genes across the tree but also to estimate the metabolic potential of the last bacterial common ancestor (LBCA). We built a second, smaller set of COG-based gene families better suited for functional annotation and reconciled their gene trees with the

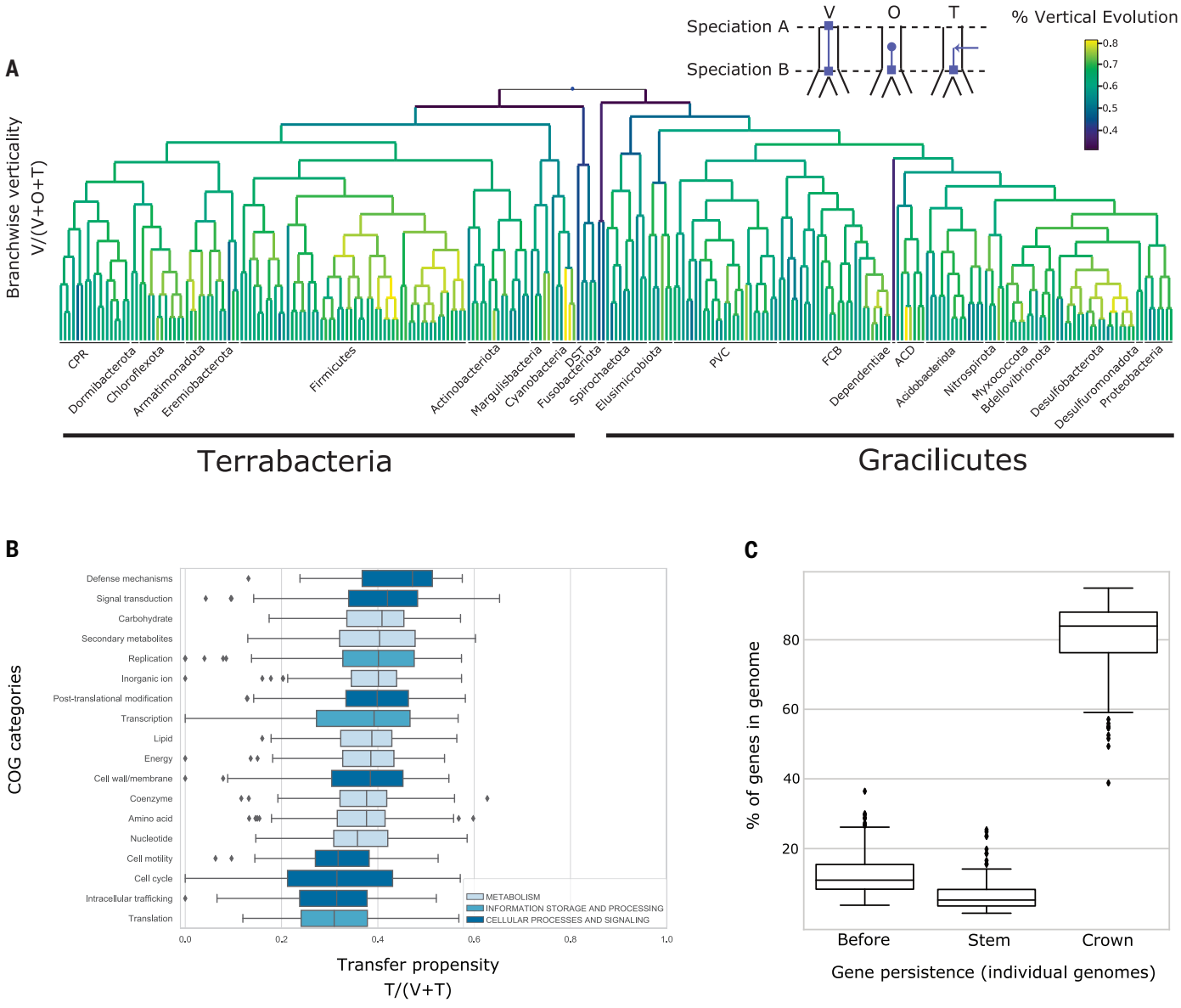


Fig. 3. The verticality of bacterial genome evolution. (**A**) The rooted bacterial species tree (Fig. 1), with branches colored according to verticality: the fraction of genes at the bottom of a branch that descend vertically from the top of that branch (see inset; V, vertical; O, origination; T, transfer into a branch) (*36*). Node heights reflect relative time order consistent with highly supported gene transfers (Fig. 2). (**B**) Transfer propensity by COG functional category; that is, the proportion of gene tree branches that are horizontal T/(V+T) for COG gene families. Genes involved in information processing, particularly translation (J), show the lowest transfer propensity (median: 0.31), while genes involved in cell defense mechanisms (V, such as genes involved in antibiotic defense and biosynthesis) are most frequently transferred (median transfer propensity: 0.47). (**C**) From the genome's eye view, this combination of vertical and horizontal processes gives rise to a distribution of gene persistences (residence times), reflecting the point in evolutionary history [within the Crown group, on the Stem, or earlier (Before)] at which the gene was most recently acquired. Across all phyla examined, 82% of genes on sampled genomes were most recently acquired since the crown group radiation of that phylum. Lineage acronyms are as in Fig. 1.

species tree (*36*). In the following reconstruction, we indicate when gene content inferences differ between roots (*36*). Posterior probabilities (PPs) for genes directly relevant to our reconstruction are provided in table S7, and all of the pathways we discuss below were confirmed in our analysis of the secondary dataset (*36*). From the root placement and estimated rates of gene family extinction in the focal analysis (*1*), we predict that LBCA encoded 1293 to 2143 COG family members, the majority of which (median estimates: 65 to 69.5%; 95% confidence interval: 57 to 82%) survived to be sampled in at least one present-day genome. On the basis of the relationship between COG family members and genome size for extant Bacteria (Pearson's correlation coefficient = 0.96, $P = 8 \times 10^{-153}$), we estimate the genome size of LBCA to be 2.7 ± 0.4 Mb (SE) for root 1 of the focal analysis (Fusobacteriota with Terrabacteria) (Fig. 1B), 2.6 ± 0.4 Mb for root 2 (Fusobacteriota with Gracilicutes), and 1.6 ± 0.5 Mb for root 3 (Fusobacteriota root). Under all three roots, the trend in genome size evolution from LBCA to modern taxa is an ongoing moderate increase through time in estimated COG family complements and genome sizes. The most notable departure from this trend is a reduction in genome size of between 0.47 and 0.56 Mb on the CPR stem lineage after

divergence from their common ancestor with Chloroflexota and Dormibacterota (fig. S16). COG families lost on the CPR stem include components of the electron transport chain, carbon metabolism, flagellar biosynthesis and motor switch proteins, amino acid biosynthesis, the Clp protease subunit ClpX, and RNA polymerase sigma factor-54 (table S8), consistent with their absence in extant CPR (*18*).

The inferred ancestral gene set for LBCA includes most components of the modern bacterial transcription, translation, and DNA replication systems (table S7). This gene set also includes an FtsZ-based cell division machinery and pathways for signal transduction, membrane transport, and secretion (Fig. 4) (*36*). Further, we identified proteins involved in bacterial phospholipid biosynthesis, suggesting that LBCA had bacterial-type ester-lipid membranes (Fig. 4). We also identified most of the proteins required for flagella and pili synthesis and those for quorum sensing, suggesting that LBCA was motile (*51*, *52*). Given that bacterial genes are typically maintained by strong purifying selection (*53*), these findings imply that LBCA lived in an environment in which dispersal, chemotaxis, and surface attachment were advantageous.

Moderate support for the presence of the shape-determining proteins MreB (PP = 0.9, 0.88, and 0.73 for roots 1 to 3, respectively, as depicted in Fig. 1B), MreC (PP = 0.82/0.79/0.57), and MreD (PP = 0.86/0.83/0.63) at the root suggests that LBCA was a rod-shaped cell (*52*). We also obtained high root PPs for proteins mediating outer cell envelope biosynthesis, including lipopolysaccharides (LPSs), from which we infer that LBCA had a double membrane with an LPS layer (*36*). Consistent with this inference, there was strong support for the flagellar subunits FlgH, FlgI, and FgA, which anchor flagella in diderm membranes (*54*), and for the type IV pilus subunit PilQ, which among extant bacteria is specific to diderms (*54*, *55*). Altogether, this supports hypotheses (*9*) in which LBCA was a diderm (*54*–*56*) and argues against scenarios in which the Gram-negative double membrane originated by endosymbiosis between monoderms [single-membraned bacteria (*10*)] or via the arrest of sporulation (*57*) in a spore-forming monoderm ancestor. Thus, diderm-to-monoderm transitions must have occurred subsequently on multiple occasions within Bacteria (*54*–*56*).

We recovered components of several core pathways for carbohydrate metabolism with high posterior support, including glycolysis, the tricarboxylic acid (TCA) cycle, and the pentose phosphate pathway (Fig. 4, figs. S17 and S18, and table S7) (*36*). Modern bacteria fix carbon using several different pathways, including the Calvin cycle, the 3-hydroxypropionate bicycle and variations thereof, the reductive glycine pathway (*58*), the Wood-Ljungdahl pathway (WLP), and the reverse TCA cycle, of which the latter two have been suggested to

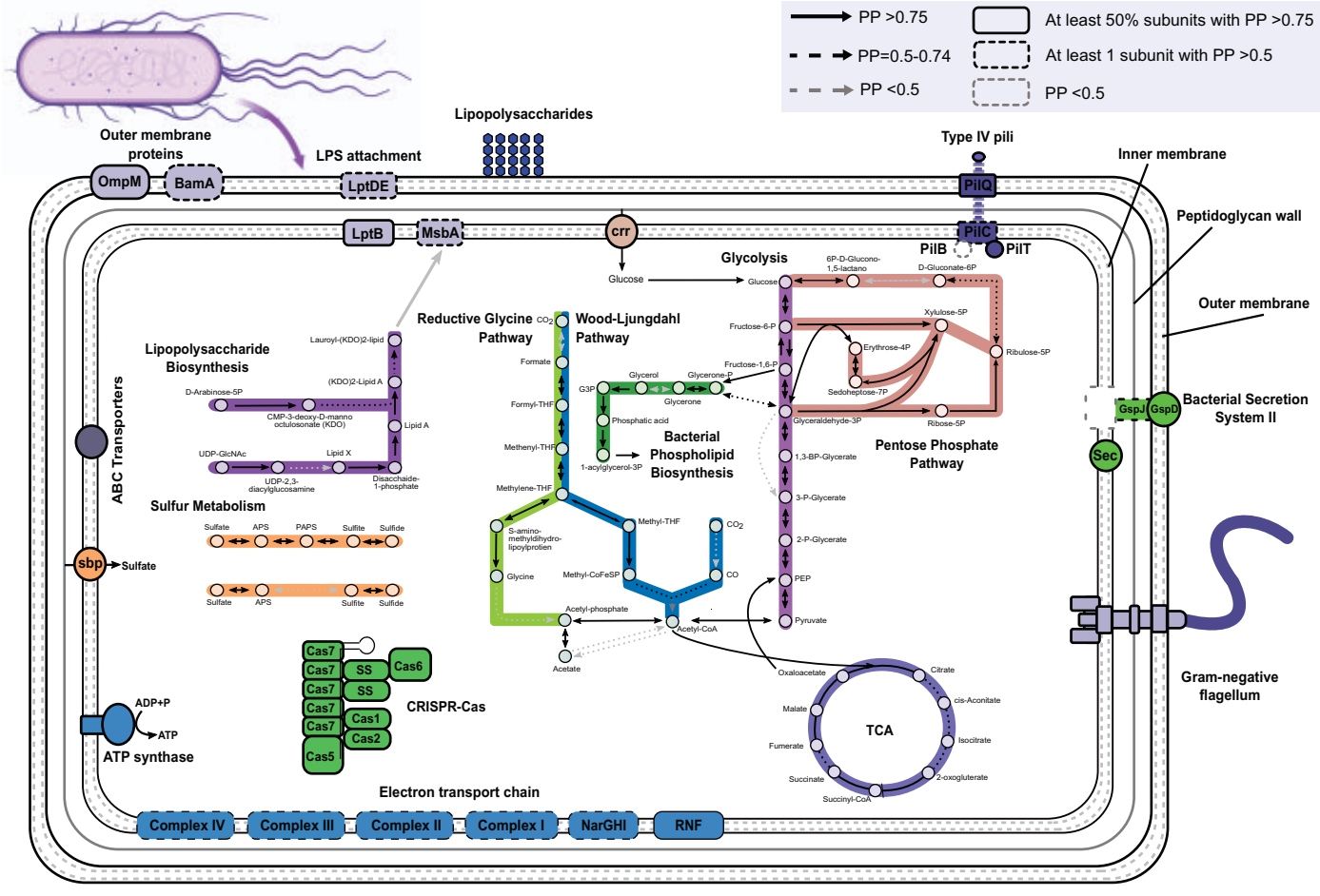


Fig. 4. Ancestral reconstruction of the last bacterial common ancestor (LBCA). The reconstruction is based on genes that could be mapped to at least one branch within the root region with a PP > 0.5 (figs. S17 and S18) (*36*). The presence of a gene within a pathway is indicated as shown in the key. Our analyses suggest that LBCA was a rod-shaped, motile, flagellated double-membraned cell. We recover strong support for central carbon pathways, including glycolysis, the TCA cycle, and the pentose phosphate pathway. We did not find unequivocal evidence for the presence of a carbon fixation pathway, but we did find moderate support for components of both the WLP and the reverse TCA cycle. Although not depicted here, our analyses suggest that the machinery for core information processing and quorum sensing was also present in LBCA (table S7).

have emerged early in the history of life (*41*, *59*–*63*). Of these, we identified several enzymes of the TCA cycle and the reductive glycine pathway, although we did not recover the key enzymes of either pathway, and the directionality of the recovered enzymes is difficult to assess (*64*) (Fig. 4 and figs. S17 and S18). Furthermore, we identified several enzymes of the methyl branch of the WLP for acetate biosynthesis and components of a putative Rhodobacter nitrogen-fixing (RNF) complex (Fig. 4 and figs. S17 and S18), which together may indicate that LBCA was capable of acetogenic growth (*36*, *65*). However, the key enzyme of the WLP, the carbon monoxide dehydrogenase/acetyl coenzyme A synthase complex (*41*), had only moderate root support (PP = 0.5 to 0.75) for two subunits and low support (PP < 0.5) for other subunits. Thus, while our analyses support the antiquity of components of the WLP, acetogenesis, the TCA cycle, and several other core metabolic pathways, they do not confidently establish the combination of pathways used by LBCA (*36*).

Finally, our reconstruction also indicated high posterior support for elements of an adaptive immune CRISPR-Cas system (*66*, *67*), including the universally conserved Cas endonuclease, Cas1 (PP = 0.96/0.93/0.89), essential for spacer acquisition and insertion into CRISPR cassettes (*68*, *69*). Among other roles, CRISPR systems are crucial in antiviral defense and are activated in response to viral exposure (*70*); therefore, these findings are consistent with hypotheses suggesting that LBCA was already coevolving with parasitic replicators such as bacteriophages and plasmids (*71*, *72*).

Vertical and horizontal evolution are complementary

Here, we have used reconciliation methods to model both the vertical and horizontal components of bacterial evolution. These components are complementary, illuminating different facets of bacterial evolution, and we show that the horizontal component can be used to root and orient the vertical tree. Our analyses root the Bacteria between two major clades, the Terrabacteria and Gracilicutes, in contrast to recent outgroup-rooted analyses that place the root on the CPR branch. Instead, we predict that CPR evolved from a common ancestor with the Chloroflexota and Dormibacterota by reductive evolution. We infer that the last bacterial common ancestor was a fully fledged free-living diderm cell with an LPS layer, a multimeric flagellum, and a type III CRISPR-Cas system.

Phylogenetic models are necessarily simplified, and there is much work to be done to better capture the full heterogeneity of the evolutionary process in the reconciliation framework, from varying diversification rates to endosymbioses. With increased sampling and improved methods, reconciliation analyses should be able to probe still deeper into the early evolutionary history of life on Earth.

Methods summary

Phylogenetics

We used two alternative approaches to assemble representative sets of bacterial genomes. In the focal analysis, we sampled 265 genomes evenly from across the GTDB taxonomy (*13*). In the secondary analysis, we sampled 341 genomes according to the diversity of major bacterial lineages reported in a previous study (*11*). We used the OMA (*73*) algorithm to identify candidate single-copy orthologs and manually inspected initial single gene trees to identify a set of 62 congruent phylogenetic markers. Sequences were aligned using MAFFT 7.453 (*74*) and trimmed using BMGE 1.12 (*75*) with the BLOSUM30 matrix. Unrooted species trees were inferred from a concatenation of the 62 markers under the LG+C60+R8+F model in IQ-TREE 1.6.10 (*76*), which was the best-fitting model according to the BIC (*37*). To perform outgroup rooting analyses, we searched the genomes of 148 Archaea for orthologs of the 62-marker gene set and identified a subset of 29 genes with congruent single-gene phylogenies. AU tests (*28*) were performed in IQ-TREE.

Gene tree-species tree reconciliation

To infer gene families, we performed all-versus-all DIAMOND (*77*) searches among the input protein sets and clustered the results using the MCL algorithm (*38*) with an inflation parameter of 1.2. Gene clusters were aligned and trimmed as described above, and bootstrap distributions inferred under the best-fitting model in IQ-TREE. We used ALEml_undated (*31*) to perform gene tree-species tree reconciliation. The relative ages of bacterial crown groups were estimated with MaxTiC (*43*) using only those transfer-based age constraints that were recovered in both the focal and secondary datasets. Estimates of gene family and lineage verticality were averaged over the reconciliations obtained in the focal analysis when rooting on each of the three candidate branches in the root region.

Simulations and sensitivity analyses

To evaluate ALE performance, we simulated gene family evolution using Zombi (*78*) combined with rejection sampling to obtain sets of simulated gene families similar to the real data in terms of inferred DTL events. We then compared simulated and inferred numbers of events under a range of conditions (*36*). To evaluate the robustness of root inferences, we ordered gene families by decreasing DTL rates, rate ratios, and a range of other proxies for lack of informativeness and potential for introducing bias (*36*) and compared the likelihoods of competing root hypotheses as increasing proportions of gene families were excluded from the calculation.

Ancestral metabolic reconstruction

We inferred COG gene families by assigning the sequences on each sampled genome to gene families from the COG ontology (*39*) using eggNOG-mapper 2 (*79*), which were then used to perform gene tree-species tree reconciliation. Root origination probabilities for each of the 23 COG functional categories were inferred by maximizing the total reconciliation likelihood over all gene families in that category. These category-specific probabilities for origination at the root were then used to estimate the PP that each gene family was present at the root of the tree. To infer the gene family content and metabolic repertoire of LBCA, functional annotations of protein sequences were obtained and assigned to COG families present at the root. The LBCA proteome was reconstructed taking into account the respective PPs for key gene families and metabolic pathways. A detailed account of all analyses is provided in the supplementary methods (*36*).

REFERENCES AND NOTES

1. T. A. Williams *et al.*, Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4602–E4611 (2017). doi: 10.1073/pnas.1618463114; pmid: 28533395
2. J. R. Brown, W. F. Doolittle, Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2441–2445 (1995). doi: 10.1073/pnas.92.7.2441; pmid: 7708661
3. J. P. Gogarten *et al.*, Evolution of the vacuolar H+-ATPase: Implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6661–6665 (1989). doi: 10.1073/pnas.86.17.6661; pmid: 2528146
4. N. Iwabe, K. Kuma, M. Hasegawa, S. Osawa, T. Miyata, Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9355–9359 (1989). doi: 10.1073/pnas.86.23.9355; pmid: 2531898
5. O. Zhaxybayeva, P. Lapierre, J. P. Gogarten, Ancient gene duplications and the root(s) of the tree of life. *Protoplasma* **227**, 53–64 (2005). doi: 10.1007/s00709-005-0135-1; pmid: 16389494
6. F. U. Battistuzzi, S. B. Hedges, A major clade of prokaryotes with ancient adaptations to life on land. *Mol. Biol. Evol.* **26**, 335–343 (2009). doi: 10.1093/molbev/msn247; pmid: 18988685
7. M. Bocchetta, S. Gribaldo, A. Sanangelantoni, P. Cammarano, Phylogenetic depth of the bacterial genera *Aquifex* and *Thermotoga* inferred from analysis of ribosomal protein, elongation factor, and RNA polymerase subunit sequences. *J. Mol. Evol.* **50**, 366–380 (2000). doi: 10.1007/s002399910040; pmid: 10795828
8. C. Brochier, H. Philippe, A non-hyperthermophilic ancestor for Bacteria. *Nature* **417**, 244 (2002). doi: 10.1038/417244a; pmid: 12015592
9. T. Cavalier-Smith, Rooting the tree of life by transition analyses. *Biol. Direct* **1**, 19 (2006). doi: 10.1186/1745-6150-1-19; pmid: 16834776
10. J. A. Lake, Evidence for an early prokaryotic endosymbiosis. *Nature* **460**, 967–971 (2009). doi: 10.1038/nature08183; pmid: 19693078
11. L. A. Hug *et al.*, A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016). doi: 10.1038/nmicrobiol.2016.48; pmid: 27572647
12. S. Mukherjee *et al.*, 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat. Biotechnol.* **35**, 676–683 (2017). doi: 10.1038/nbt.3886; pmid: 28604660
13. D. H. Parks *et al.*, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life.

Nat. Biotechnol. **36**, 996–1004 (2018). doi: 10.1038/nbt.4229; pmid: 30148503
14. D. H. Parks *et al.*, Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542 (2017). doi: 10.1038/s41564-017-0012-7; pmid: 28894102
15. C. T. Brown *et al.*, Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208–211 (2015). doi: 10.1038/nature14486; pmid: 26083755
16. Q. Zhu *et al.*, Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nat. Commun.* **10**, 5477 (2019). doi: 10.1038/s41467-019-13443-4; pmid: 31792218
17. C. J. Castelle, J. F. Banfield, Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell* **172**, 1181–1197 (2018). doi: 10.1016/j.cell.2018.02.016; pmid: 29522741
18. C. J. Castelle *et al.*, Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat. Rev. Microbiol.* **16**, 629–645 (2018). doi: 10.1038/s41579-018-0076-2; pmid: 30181663
19. J. P. Beam *et al.*, Ancestral absence of electron transport chains in Patescibacteria and DPANN. *Front. Microbiol.* **11**, 1848 (2020). doi: 10.3389/fmicb.2020.01848; pmid: 33013724
20. C. J. Castelle *et al.*, Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr. Biol.* **25**, 690–701 (2015). doi: 10.1016/j.cub.2015.01.014; pmid: 25702576
21. R. Méheust, D. Burstein, C. J. Castelle, J. F. Banfield, The distinction of CPR bacteria from other bacteria based on protein family content. *Nat. Commun.* **10**, 4173 (2019). doi: 10.1038/s41467-019-12171-z; pmid: 31519891
22. A. Graybeal, Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* **47**, 9–17 (1998). doi: 10.1080/106351598260996; pmid: 12064243
23. S. M. Hedtke, T. M. Townsend, D. M. Hillis, Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Syst. Biol.* **55**, 522–529 (2006). doi: 10.1080/10635150600697358; pmid: 16861214
24. J. Bergsten, A review of long-branch attraction. *Cladistics* **21**, 163–193 (2005). doi: 10.1111/j.1096-0031.2005.00059.x
25. R. Gouy, D. Baurain, H. Philippe, Rooting the tree of life: The phylogenetic jury is still out. *Philos. Trans. R. Soc. London Ser. B* **370**, 20140329 (2015). doi: 10.1098/rstb.2014.0329; pmid: 26323760
26. J. A. Lake, R. G. Skophammer, C. W. Herbold, J. A. Servin, Genome beginnings: Rooting the tree of life. *Philos. Trans. R. Soc. London Ser. B* **364**, 2177–2185 (2009). doi: 10.1098/rstb.2009.0035; pmid: 19571238
27. R. G. Skophammer, J. A. Servin, C. W. Herbold, J. A. Lake, Evidence for a Gram-positive, eubacterial root of the tree of life. *Mol. Biol. Evol.* **24**, 1761–1768 (2007). doi: 10.1093/molbev/msm096; pmid: 17513883
28. H. Shimodaira, An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508 (2002). doi: 10.1080/10635150290069913; pmid: 12079646
29. G. J. Szöllosi, B. Boussau, S. S. Abby, E. Tannier, V. Daubin, Phylogenetic modeling of lateral gene transfer reconstructs the pattern and relative timing of speciations. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17513–17518 (2012). doi: 10.1073/pnas.1202997109; pmid: 23043116
30. L. A. David, E. J. Alm, Rapid evolutionary innovation during an Archaean genetic expansion. *Nature* **469**, 93–96 (2011). doi: 10.1038/nature09649; pmid: 21170026
31. G. J. Szöllősi, W. Rosikiewicz, B. Boussau, E. Tannier, V. Daubin, Efficient exploration of the space of reconciled gene trees. *Syst. Biol.* **62**, 901–912 (2013). doi: 10.1093/sysbio/syt054; pmid: 23925510
32. L. A. Katz, J. R. Grant, L. W. Parfrey, J. G. Burleigh, Turning the crown upside down: Gene tree parsimony roots the eukaryotic tree of life. *Syst. Biol.* **61**, 653–660 (2012). doi: 10.1093/sysbio/sys026; pmid: 22334342
33. D. M. Emms, S. Kelly, STRIDE: Species Tree Root Inference from Gene Duplication Events. *Mol. Biol. Evol.* **34**, 3267–3278 (2017). doi: 10.1093/molbev/msx259; pmid: 29029342
34. A. Zwaenepoel, Y. Van de Peer, Inference of ancient whole-genome duplications and the evolution of gene duplication and loss rates. *Mol. Biol. Evol.* **36**, 1384–1404 (2019). doi: 10.1093/molbev/msz088; pmid: 31004147
35. S. Q. Le, O. Gascuel, An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008). doi: 10.1093/molbev/msn067; pmid: 18367465
36. See supplementary materials.
37. D. Posada, T. R. Buckley, Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* **53**, 793–808 (2004). doi: 10.1080/10635150490522304; pmid: 15545256
38. A. J. Enright, S. Van Dongen, C. A. Ouzounis, An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002). doi: 10.1093/nar/30.7.1575; pmid: 11917018
39. M. Y. Galperin, D. M. Kristensen, K. S. Makarova, Y. I. Wolf, E. V. Koonin, Microbial genome analysis: The COG approach. *Brief. Bioinform.* **20**, 1063–1070 (2019). doi: 10.1093/bib/bbx117; pmid: 28968633
40. K. Raymann, C. Brochier-Armanet, S. Gribaldo, The two-domain tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6670–6675 (2015). doi: 10.1073/pnas.1420858112; pmid: 25964353
41. P. S. Adam, G. Borrel, S. Gribaldo, Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1166–E1173 (2018). doi: 10.1073/pnas.1716667115; pmid: 29358391
42. A. A. Davín *et al.*, Gene transfers can date the tree of life. *Nat. Ecol. Evol.* **2**, 904–909 (2018). doi: 10.1038/s41559-018-0525-3; pmid: 29610471
43. C. Chauve *et al.*, MaxTiC: fast ranking of a phylogenetic tree by maximum time consistency with lateral gene transfers. bioRxiv 127548 [Preprint]. 6 October 2017. https://doi.org/10.1101/127548.
44. T. Dagan, W. Martin, Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 870–875 (2007). doi: 10.1073/pnas.0606318104; pmid: 17213324
45. W. F. Doolittle, Phylogenetic classification and the universal tree. *Science* **284**, 2124–2128 (1999). doi: 10.1126/science.284.5423.2124; pmid: 10381871
46. W. F. Doolittle, E. Bapteste, Pattern pluralism and the Tree of Life hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2043–2049 (2007). doi: 10.1073/pnas.0610699104; pmid: 17261804
47. T. Dagan, W. Martin, The tree of one percent. *Genome Biol.* **7**, 118 (2006). doi: 10.1186/gb-2006-7-10-118; pmid: 17081279
48. D. Alvarez-Ponce, P. Lopez, E. Bapteste, J. O. McInerney, Gene similarity networks provide tools for understanding eukaryote origins and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E1594–E1603 (2013). doi: 10.1073/pnas.1211371110; pmid: 23576716
49. R. Jain, M. C. Rivera, J. A. Lake, Horizontal gene transfer among genomes: The complexity hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3801–3806 (1999). doi: 10.1073/pnas.96.7.3801; pmid: 10097118
50. P. Puigbò, Y. I. Wolf, E. V. Koonin, The tree and net components of prokaryote evolution. *Genome Biol. Evol.* **2**, 745–756 (2010). doi: 10.1093/gbe/evq062; pmid: 20889655
51. R. Liu, H. Ochman, Stepwise formation of the bacterial flagellar system. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7116–7121 (2007). doi: 10.1073/pnas.0700266104; pmid: 17438286
52. F. El Baidouri, C. Venditti, S. Suzuki, A. Meade, S. Humphries, Phenotypic reconstruction of the last universal common ancestor reveals a complex cell. bioRxiv 2020.08.20.260398 [Preprint]. 21 August 2020. https://doi.org/10.1101/2020.08.20.26039.
53. I. Sela, Y. I. Wolf, E. V. Koonin, Theory of prokaryotic genome evolution. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11399–11407 (2016). doi: 10.1073/pnas.1614083113; pmid: 27702904
54. L. C. Antunes *et al.*, Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. *eLife* **5**, e14589 (2016). doi: 10.7554/eLife.14589; pmid: 27580370
55. D. Megrian, N. Taib, J. Witwinowski, C. Beloin, S. Gribaldo, One or two membranes? Diderm Firmicutes challenge the Gram-positive/Gram-negative divide. *Mol. Microbiol.* **113**, 659–671 (2020). doi: 10.1111/mmi.14469; pmid: 31975449
56. N. Taib *et al.*, Genome-wide analysis of the Firmicutes illuminates the diderm/monoderm transition. *Nat. Ecol. Evol.* **4**, 1661–1672 (2020). doi: 10.1038/s41559-020-01299-7; pmid: 33077930
57. E. I. Tocheva, D. R. Ortega, G. J. Jensen, Sporulation, bacterial cell envelopes and the origin of life. *Nat. Rev. Microbiol.* **14**, 535–542 (2016). doi: 10.1038/nrmicro.2016.85; pmid: 28232669
58. I. Sánchez-Andrea *et al.*, The reductive glycine pathway allows autotrophic growth of *Desulfovibrio desulfuricans*. *Nat. Commun.* **11**, 5090 (2020). doi: 10.1038/s41467-020-18906-7; pmid: 33037220
59. M. C. Weiss *et al.*, The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116 (2016). doi: 10.1038/nmicrobiol.2016.116; pmid: 27562259
60. F. L. Sousa, W. F. Martin, Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta Bioenerg.* **1837**, 964–981 (2014). doi: 10.1016/j.bbabio.2014.02.001; pmid: 24513196
61. F. L. Sousa, S. Nelson-Sathi, W. F. Martin, One step beyond a ribosome: The ancient anaerobic core. *Biochim. Biophys. Acta Bioenerg.* **1857**, 1027–1038 (2016). doi: 10.1016/j.bbabio.2016.04.284; pmid: 27150504
62. G. Borrel, P. S. Adam, S. Gribaldo, Methanogenesis and the Wood–Ljungdahl pathway: An ancient, versatile, and fragile association. *Genome Biol. Evol.* **8**, 1706–1711 (2016). doi: 10.1093/gbe/evw114; pmid: 27189979
63. G. Fuchs, Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658 (2011). doi: 10.1146/annurev-micro-090110-102801; pmid: 21740227
64. T. Nunoura *et al.*, A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* **359**, 559–563 (2018). doi: 10.1126/science.aao3407; pmid: 29420286
65. K. Schuchmann, V. Müller, Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**, 809–821 (2014). doi: 10.1038/nrmicro3365; pmid: 25383604
66. K. S. Makarova *et al.*, Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477 (2011). doi: 10.1038/nrmicro2577; pmid: 21552286
67. E. V. Koonin, K. S. Makarova, Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc. London Ser. B* **374**, 20180087 (2019). doi: 10.1098/rstb.2018.0087; pmid: 30905284
68. J. K. Nuñez *et al.*, Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **21**, 528–534 (2014). doi: 10.1038/nsmb.2820; pmid: 24793649
69. K. S. Makarova *et al.*, An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015). doi: 10.1038/nrmicro3569; pmid: 26411297
70. R. Barrangou *et al.*, CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007). doi: 10.1126/science.1138140; pmid: 17379808
71. E. V. Koonin, The origins of cellular life. *Antonie van Leeuwenhoek* **106**, 27–41 (2014). doi: 10.1007/s10482-014-0169-5; pmid: 24756907
72. M. Krupovic, V. V. Dolja, E. V. Koonin, Origin of viruses: Primordial replicators recruiting capsids from hosts. *Nat. Rev. Microbiol.* **17**, 449–458 (2019). doi: 10.1038/s41579-019-0205-6; pmid: 31142823
73. A. C. J. Roth, G. H. Gonnet, C. Dessimoz, Algorithm of OMA for large-scale orthology inference. *BMC Bioinformatics* **9**, 518 (2008). doi: 10.1186/1471-2105-9-518; pmid: 19055798
74. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi: 10.1093/molbev/mst010; pmid: 23329690
75. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010). doi: 10.1186/1471-2148-10-210; pmid: 20626897
76. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015). doi: 10.1093/molbev/msu300; pmid: 25371430
77. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015). doi: 10.1038/nmeth.3176; pmid: 25402007
78. A. A. Davín, T. Tricou, E. Tannier, D. M. de Vienne, G. J. Szöllősi, Zombi: A phylogenetic simulator of trees, genomes and sequences that accounts for dead linages. *Bioinformatics* **36**, 1286–1288 (2020). doi: 10.1093/bioinformatics/btz710; pmid: 31566657
79. J. Huerta-Cepas *et al.*, Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper.

Mol. Biol. Evol. **34**, 2115–2122 (2017). doi: 10.1093/molbev/msx148; pmid: 28460117
80. G. Coleman *et al.*, Extended Data for A rooted phylogeny resolves early bacterial evolution, Version 9, Figshare (2020); https://doi.org/10.6084/m9.figshare.12651074.v9.

ACKNOWLEDGMENTS

Funding: G.A.C. is supported by a Royal Society Research Grant to T.A.W. T.A.W. is supported by a Royal Society University Research Fellowship and NERC grant NE/P00251X/1. G.J.Sz. received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program under grant agreement 714774 and grant GINOP-2.3.2.-15-2016-00057. A.S. is supported by the Swedish Research Council (VR starting grant 2016-03559 to A.S.) and the NWO-I foundation of the Netherlands Organisation for Scientific Research (WISE fellowship to A.S.). A.A.D. and P.H. are supported by an Australian Research Council Laureate Fellowship (grant FL150100038).
Author contributions: The project was conceived of by T.A.W., G.J.Sz., P.H., A.S., G.A.C., and A.A.D. G.A.C., A.A.D., T.A.W., L.L.Sz., and G.J.Sz. performed phylogenomic analyses. G.J.Sz. developed new analytical methods. T.A.M., G.A.C., A.A.D., and A.S. performed metabolic annotations and reconstructions. All authors contributed to interpretation and writing. **Competing interests:** The authors have no competing interests to declare. **Data and materials availability:** All data and code are provided online at Figshare (*80*). New methods are described in detail in the supplementary methods (*36*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/eabe0511/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S13
References (*81*–*95*)
MDAR Reproducibility Checklist

28 July 2020; resubmitted 5 November 2020
Accepted 1 April 2021
10.1126/science.abe0511



Pushing the Boundaries of Knowledge

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

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

RESEARCH ARTICLE SUMMARY

GENETICS

Environmental robustness of the global yeast genetic interaction network

Michael Costanzo†, Jing Hou†, Vincent Messier†, Justin Nelson†, Mahfuzur Rahman†, Benjamin VanderSluis, Wen Wang, Carles Pons, Catherine Ross, Matej Ušaj, Bryan-Joseph San Luis, Emira Shuteriqi, Elizabeth N. Koch, Patrick Aloy, Chad L. Myers*, Charles Boone*, Brenda Andrews*

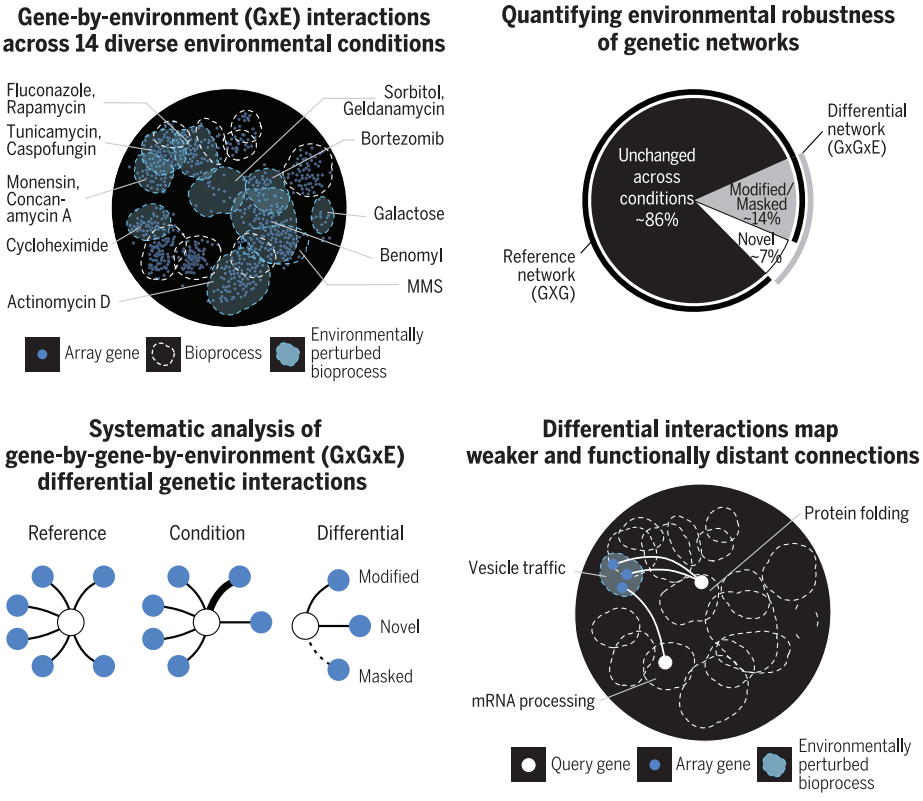
INTRODUCTION: Genetic interactions are identified when variants in different genes combine to generate an unusual phenotype compared with the expected combined effect of the corresponding individual variants. For example, a synthetic lethal genetic interaction occurs when two mutations, neither of which is lethal on their own, combine to generate a lethal double-mutant phenotype. Although there are millions of possible gene-gene combinations for any eukaryotic cell, only a rare subset of gene pairs will display a genetic interaction. Digenic, or gene-by-gene (GxG), interactions appear to underlie key aspects of biology, including the relationship between genotype and phenotype. Environmental conditions can modulate the phenotype associated with genetic variants, giving rise to gene-by-environment (GxE) interactions, when a single variant phenotype is modified, or gene-by-gene-by-environment (GxGxE) interactions when a genetic interaction is changed.

RATIONALE: A global genetic interaction network has been mapped for the budding yeast *Saccharomyces cerevisiae*, identifying thousands of connections that often occur between functionally related genes. Because the global genetic network was mapped in a specific reference condition, the potential for different environmental conditions to rewire the network remains unclear. Automated yeast genetics, combined with knowledge of a reference map, enables quantification of the extent to which new environmental conditions either modulate known genetic interactions or generate novel genetic interactions to influence the genetic landscape of a cell.

RESULTS: We tested ~4000 yeast single mutants for GxE interactions across 14 diverse environments, including an alternative carbon source, osmotic and genotoxic stress, and treatment with 11 bioactive compounds targeting distinct yeast bioprocesses. To quantify GxGxE interactions, we constructed ~30,000 different double mutants, involving genes annotated to all major yeast bioprocesses, and we scored them for genetic interactions. The plasticity of the network is revealed by differential genetic interactions, which occur when a genetic interaction observed in a particular condition deviates from that scored in the control reference network.

Although ~10,000 differential interactions were discovered across all 14 conditions, we observed ~60% fewer differential interactions per condition as compared with genetic interactions in the reference condition, indicating that GxGxE interactions are rare relative to GxG interactions. On average, a single environmental perturbation modulated ~14% of the reference genetic interactions and revealed a smaller subset of ~7% novel differential interactions. Whereas GxG genetic interactions tend to connect pairs of genes that share a close functional relationship, novel differential GxGxE interactions mediate weaker connections between gene pairs with diverse roles.

CONCLUSION: Our general findings reveal how environmental conditions modulate the yeast global genetic interaction network, allowing us to assess the plasticity of genetic networks and the extent to which mapping genetic interactions in different environments can expand a reference network. Although different environments have the potential to reveal novel interactions and uncover previously unidentified but weaker functional connections between genes, the vast majority of genetic interactions remained unchanged in different conditions, suggesting that the global yeast genetic interaction network is largely robust to environmental perturbation. ■



Systematic analysis of environmental impact on the global yeast genetic interaction network. (Top left) Mapping GxE interactions and (bottom left) GxGxE differential interactions reveals (top right) the environmental robustness of the global yeast genetic interaction network, (bottom right) highlighting new and distant functional connections associated with novel differential interactions.

The list of author affiliations is available in the full article online.
†These authors contributed equally to this work
***Corresponding author. Email: chadm@umn.edu (C.L.M.); charlie.boone@utoronto.ca (C.B.); brenda.andrews@utoronto.ca (B.A.)**
Cite this article as M. Costanzo *et al.*, *Science* 372, eabf8424 (2021). DOI: 10.1126/science.abf8424

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf8424

RESEARCH ARTICLE

GENETICS

Environmental robustness of the global yeast genetic interaction network

Michael Costanzo[1]†, Jing Hou[1]†, Vincent Messier[1]†, Justin Nelson[2,3]†, Mahfuzur Rahman[2,3]†, Benjamin VanderSluis[2], Wen Wang[2], Carles Pons[4], Catherine Ross[1], Matej Ušaj[1], Bryan-Joseph San Luis[1], Emira Shuteriqi[1], Elizabeth N. Koch[2], Patrick Aloy[4,5], Chad L. Myers[2,3]*, Charles Boone[1,6,7]*, Brenda Andrews[1,6]*

Phenotypes associated with genetic variants can be altered by interactions with other genetic variants (GxG), with the environment (GxE), or both (GxGxE). Yeast genetic interactions have been mapped on a global scale, but the environmental influence on the plasticity of genetic networks has not been examined systematically. To assess environmental rewiring of genetic networks, we examined 14 diverse conditions and scored 30,000 functionally representative yeast gene pairs for dynamic, differential interactions. Different conditions revealed novel differential interactions, which often uncovered functional connections between distantly related gene pairs. However, the majority of observed genetic interactions remained unchanged in different conditions, suggesting that the global yeast genetic interaction network is robust to environmental perturbation and captures the fundamental functional architecture of a eukaryotic cell.

Genetic interactions are identified when two or more different gene variants combine to cause an unusual phenotype that deviates from a model on the basis of the combined effects of the corresponding single-variant phenotypes (*1*). Digenic or gene-by-gene (GxG) interactions appear to underlie diverse and fundamental aspects of biology, including the relationship between genotype and phenotype, the evolution of sexual reproduction, and speciation (*2*–*6*). The phenotype associated with genetic variants can also be modulated by environmental factors—including growth conditions, age, cell type, or microbe exposure—giving rise to gene-by-environment (GxE) interactions or, if a genetic interaction is modified, gene-by-gene-by-environment (GxGxE) interactions. Eukaryotic genomes contain thousands of different genes and millions of possible genetic interactions and function in many environmental contexts, meaning that our ability to explore the extent to which genetic interactions contribute to trait heritability is a major combinatorial and statistical challenge (*3*).

Efforts to address this challenge have involved the systematic mapping of genetic interactions in accessible model systems, primarily the budding yeast *Saccharomyces cerevisiae*. Large-scale application of an automated genetic approach, synthetic genetic array (SGA) analysis, enabled the majority of all possible gene pairs to be tested for negative and positive genetic interactions (*7*, *8*). A negative genetic interaction corresponds to a synthetic lethal or sick interaction when a double mutant shows a fitness defect greater than the expected effect of the combined single-mutant fitness phenotypes. Conversely, a positive genetic interaction is scored in a double mutant that grows better than expected (*9*). Genetic suppression represents an extreme type of positive interaction, in which the double-mutant fitness phenotype is greater than that of the least fit single mutant (*10*). The resultant global digenic interaction network comprises ~350,000 positive and ~550,000 negative interactions, which tend to connect pairs of functionally related genes (*11*, *12*). Genes encoding members of the same biological pathway or protein complex often share similar patterns or profiles of negative and positive genetic interactions, and a global network clusters genes with similar genetic interaction profiles together into a hierarchy of organized modules, corresponding to protein complexes or pathways, biological processes, and cellular compartments, revealing the functional architecture of a eukaryotic cell (Fig. 1A) (*11*, *12*).

The reference yeast genetic network was mapped in a single genetic background in a specific reference condition. The phenotypes of both single- and double-mutant cells can be modulated by the environment, which leads to GxE and more complex GxGxE interactions. For example, subsets of functionally biased yeast genes have been surveyed in response to several different environmental conditions, and condition-specific genetic interactions have been identified (*13*–*18*). However, given the focus so far on specific bioprocesses, such as DNA repair, the plasticity of genetic networks across a broad array of environmental conditions remains unclear. We systematically assessed GxE and GxGxE effects in the context of the structure and topology of the global yeast genetic interaction network. This powerful model system allows us to assess the extent to which the environment modulates genetic interactions to influence the overall genetic landscape of a cell and provides insights into the utility of genetic network reference maps.

Single-mutant fitness and gene-environment interactions

Quantitative analysis of digenic interactions depends on the accurate assessment of single- and double-mutant phenotypes (*19*). Therefore, to accurately measure the effect of environment on the yeast genetic network, a baseline of quantitative measurements of condition-specific effects on single-mutant fitness (GxE interactions) is required. To identify GxE interactions in the context of a SGA screen, we subjected an array of single mutants to SGA analysis using a neutral (control) "query" locus and assessed colony size after growth in different conditions. We examined 14 diverse conditions, including an alternative carbon source, osmotic stress, genotoxic stress, and 11 bioactive compounds that target distinct yeast biological processes within the global genetic interaction profile similarity network (Fig. 1, A and B). Relative fitness measurements were obtained for 3704 viable deletion mutant strains and 782 temperature-sensitive (TS) alleles, corresponding to 553 essential genes, both in the reference SGA condition and in each of the 14 diverse conditions (data file 1) (*20*–*23*).

To identify GxE interactions from our single-mutant fitness measurements, we calculated a differential fitness score for each mutant by measuring the difference between mutant fitness in the reference condition versus each of the 14 conditions (fig. S1A) (*23*). More than half of all the genes that we examined (59%) were associated with a significant differential single-mutant fitness defect ($P < 0.05$, differential fitness score < −0.08) (Methods summary) (*23*) in at least one condition (fig. S1B). The union of all genes that displayed fitness defects in each condition revealed that six to eight conditions identified the majority of distinct genes associated with condition-specific

[1]The Donnelly Centre, University of Toronto, 160 College Street, Toronto, ON M5S 3E1, Canada. [2]Department of Computer Science and Engineering, University of Minnesota–Twin Cities, 200 Union Street, Minneapolis, MN 55455, USA. [3]Program in Biomedical Informatics and Computational Biology, University of Minnesota–Twin Cities, 200 Union Street, Minneapolis, MN 55455, USA. [4]Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute for Science and Technology, Barcelona, Spain. [5]Institució Catalana de Recerca I Estudis Avaçats (ICREA), Barcelona, Spain. [6]Department of Molecular Genetics, University of Toronto, 160 College Street, Toronto, ON M5S 3E1, Canada. [7]RIKEN Center for Sustainable Resource Science, Wako, Saitama, Japan.
†These authors contributed equally to this work.
*Corresponding author. Email: chadm@umn.edu (C.L.M.); charlie.boone@utoronto.ca (C.B.); brenda.andrews@utoronto.ca (B.A.)

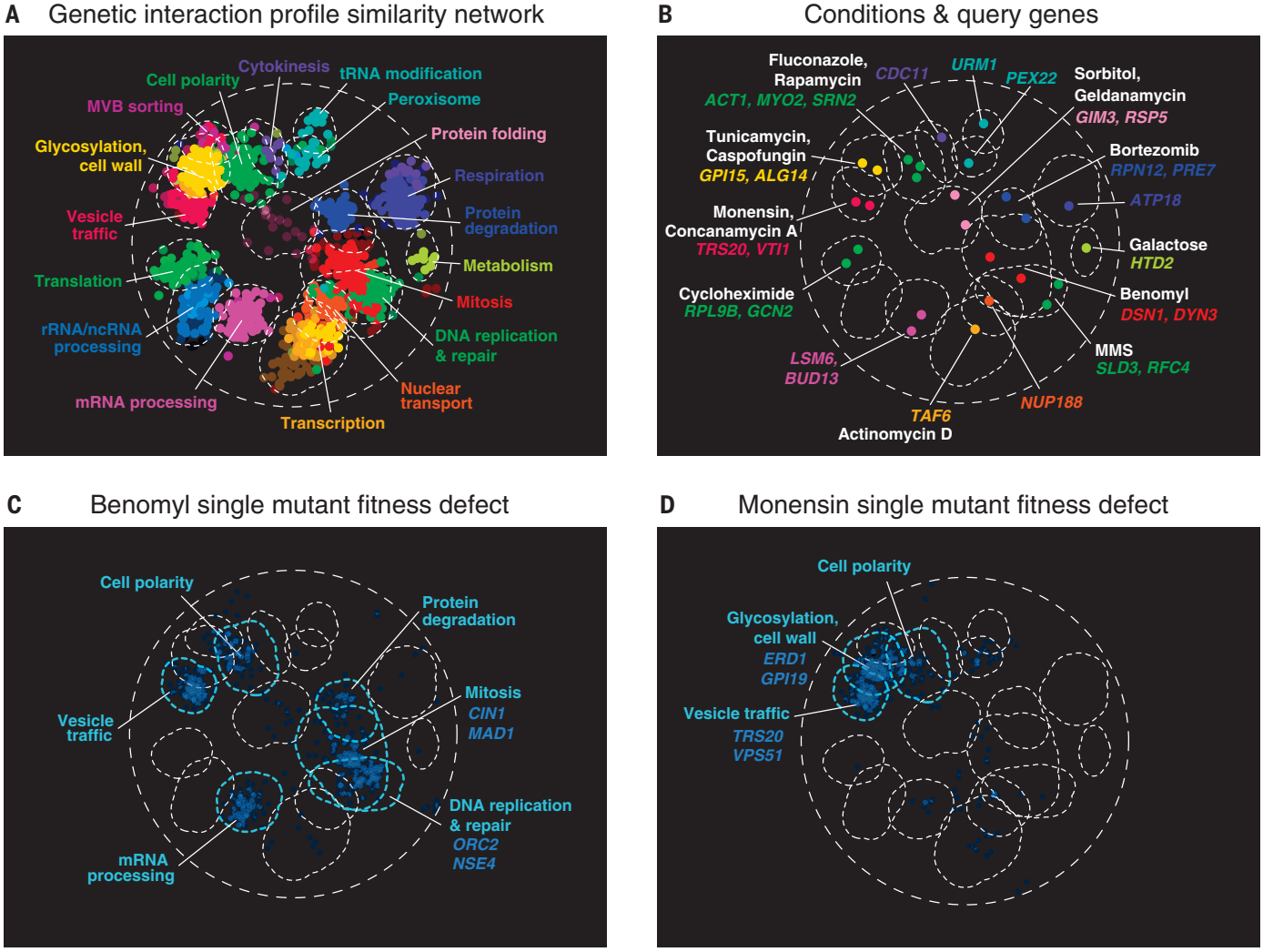


Fig. 1. A differential single-mutant fitness catalog. (**A**) A global genetic profile similarity network encompassing most nonessential and essential yeast genes (*11*). The similarity network was annotated by using SAFE (*54*), identifying network regions enriched for similar GO biological process terms, which are color-coded. (**B**) Conditions and query genes selected for this study. Location of each query gene on the global genetic interaction profile similarity network is indicated. Bioprocesses targeted by selected bioactive compounds are also shown. (**C**) Regions of the global similarity network significantly enriched for genes exhibiting differential single-mutant fitness defects in Benomyl. (**D**) Regions of the global similarity network significantly enriched for genes exhibiting differential single-mutant fitness defects in Monensin. For both (C) and (D), regions of the global similarity network significantly enriched for genes exhibiting negative differential fitness defects were mapped by using SAFE. Examples of genes located in the most enriched regions are indicated in blue.

growth defects (fig. S1C). Additional conditions captured relatively few new genes, perhaps because different conditions can lead to a similar stress response (*24*). Thus, our set of 14 conditions affects a wide range of functionally diverse genes and should be useful for exploring how different environments affect overall structure and topology of the global genetic interaction network.

The total number of strains with detectable differential fitness defects in a single condition ranged from ~250 (4.5%) to ~1200 (22%), with the largest number of condition-specific fitness defects observed in the presence of Benomyl, a microtubule depolymerizing agent (fig. S1D). Stronger negative differential fitness scores were associated with genes encoding known targets of compounds (fig. S1A). More generally, genes with differential fitness defects in a particular growth condition were often involved in a specific function perturbed by the corresponding growth condition (fig. S2 and data file 2) (*20*). For example, in Benomyl, numerous genes involved in mitosis or DNA replication and repair were enriched for differential fitness defects (Fig. 1C and data file 2) (*20*). A number of genes involved in mRNA processing also displayed fitness defects in Benomyl, which may reflect that genes encoding α-tubulin, *TUB1*, and *TUB3* are among the relatively few yeast genes with introns (Fig. 1C and data file 2) (*20*, *25*). By contrast, the differential fitness defects in Monensin, an intracellular traffic inhibitor (*26*), were most enriched among genes involved in vesicle trafficking, glycosylation, and cell wall biosynthesis (Fig. 1D and data file 2) (*20*).

Some mutants (1083) had positive differential fitness scores, indicating better growth relative to a wild-type strain in a particular condition compared with the corresponding mutant growth relative to wild-type in the reference condition (fig. S1, A, B, and D). For example, although all strains, including wild type, grew more slowly in the presence of the proteasome inhibitor Bortezomib, strains carrying TS alleles of essential genes were enriched among the set of mutant strains that showed positive differential fitness scores in this condition ($P < 10^{-27}$, Fisher's exact test) (fig. S3A). This finding mirrors a previous

observation from analysis of the global genetic interaction network in which hypomorphic alleles of proteasome genes show positive genetic interactions, which likely reflects reduced degradation of the products of TS alleles, providing a growth advantage (*11*). Similarly, approximately sixfold more TS alleles showed positive differential fitness, relative to those with a negative differential fitness, when grown in sorbitol, which may reflect a general effect of increased intracellular osmolarity ($P < 10^{-13}$, Fisher's exact test) (fig. S3B) (*27*). By contrast, TS alleles of essential genes were depleted for positive differential fitness effects in certain conditions, including growth in the presence of the *HSP90* inhibitor Geldanamycin ($P < 10^{-4}$, Fisher's exact test) (fig. S3A). This suggests that the fitness of different TS mutants is dependent on *HSP90* activity, which is consistent with the known role of *HSP90* as a phenotypic capacitor (*11*, *28*–*30*).

Last, we compiled a list of ~350 dynamic, "condition-responsive" genes, corresponding to the top ~5% of all tested genes based on the number of conditions in which they showed either a significant positive differential fitness score in two or more conditions or a significant negative differential fitness score in four or more conditions (data file 1) (*20*). Consistent with our findings, condition-responsive genes identified in our assay were also often sensitive to multiple chemicals when assayed as homozygous gene deletion mutants (2.7-fold, $P < 10^{-12}$, Fisher's exact test) (fig. S3C) (*31*). Moreover, condition-responsive genes were more likely to be essential or to have a significant growth phenotype as a deletion mutant in the reference condition, were often highly connected on the global yeast genetic interaction network, and often encoded highly conserved, multifunctional proteins (fig. S3D).

Mapping genetic interactions across different environments

Our systematic assessment of GxE interactions involving single mutants provides a foundation for comprehensive analysis of how genetic interactions are modulated by environment (GxGxE interactions). To do this, we took advantage of a diagnostic array used previously to survey complex, trigenic interaction networks (data file 1) (*6*, *20*). This array consists of ~1200 strains, comprising ~1000 nonessential gene deletion mutants and ~200 TS alleles of essential genes, spanning ~20% of all yeast genes and representing the functional breadth of the yeast genome (data file 1) (*6*, *20*). Mutant strains included on the diagnostic array exhibited a range of fitness defects and are generally representative of the genome-wide distribution of differential single-mutant fitness effects (data file 1) (*20*). For condition-specific genetic interaction analysis, we selected 26 query genes, each of which shows a substantial number of genetic interactions in the global network (*11*) and a functionally coherent genetic interaction profile that localizes the query gene to a specific biological process cluster on the reference genetic interaction similarity network (Fig. 1B).

We crossed the 26 query strains to the diagnostic mutant array, resulting in ~30,000 double-mutant strains, which were each assessed for environmental modulation of genetic interactions in all 14 conditions. To measure condition-dependent genetic interactions, every double-mutant array generated from a single-query gene was screened three times. One copy was grown in the standard SGA condition, whereas the two other copies were each grown in different conditional media (fig. S4). This configuration provided a matched reference control for every condition, which facilitated normalization of systematic experimental artifacts and improved accuracy of condition-specific genetic interaction measurements (*32*). The entire screening pipeline was repeated twice, providing two independent biological replicates for every query gene in every test condition, each one with four separate double-mutant colonies (eight tests of each double mutant per condition) and 14 independent biological replicates of each query screen in the standard reference SGA condition (*23*).

Negative and positive genetic interactions were quantified as previously described (*19*). We identified an average of ~2100 negative and ~1400 positive genetic interactions on the basis of 14 biological replicate screens performed in the standard reference condition, at an intermediate confidence threshold (fig. S5A and data file 3) (*20*, *23*). A similar number of gene pairs exhibited a genetic interaction in any single test condition, with an average of ~2150 negative and ~1500 positive genetic interactions identified per condition (fig. S5A). Our genetic interaction measurements were reproducible, and those interactions identified in the reference condition in this study overlapped substantially with interactions derived from our previous genome-wide study (fig. S5B) (*11*). Comparison of the overlap of genetic interaction profiles between a specific condition screen and its matched control with that seen with an unmatched control suggested that comparison with a matched reference provides the necessary sensitivity to detect rare condition-specific interactions (fig. S5C, tables S1 and S2, and data file 3) (*20*), which is consistent with previous studies (*32*).

Quantifying and classifying differential interactions

To discover differential or GxGxE interactions in our scored data, we quantified the difference between genetic interaction scores derived from each conditional screen and its corresponding matched reference (fig. S4) (*23*). We scored a differential negative interaction when a genetic interaction score observed in a particular condition was less than the genetic interaction score measured in the corresponding reference control. Conversely, a genetic interaction that was stronger in a given condition as compared with the matched reference control was scored as a differential positive interaction. Analysis of biological replicates confirmed the reproducibility of differential interaction scores (fig. S5D and tables S1 and S2) (*23*).

The Benomyl screen showed the largest number of differential, condition-specific fitness defects and genetic interactions, providing a rich context for developing a general scheme for defining various classes of differential interactions. We measured Benomyl differential interactions as described above, adopting an intermediate confidence threshold ($|\epsilon| > 0.08$, $P < 0.05$) (Methods summary and Fig. 2A, left) (*23*). In total, we identified 1367 differential interactions in Benomyl, which was ~60% fewer than the 3845 genetic interactions identified in the matched reference condition (Fig. 2B). Thus, the majority of the genetic interactions observed in the reference condition were also observed in the presence of Benomyl. To determine whether the prevalence of Benomyl differential interactions derived from the diagnostic array was generalizable to the whole genome, we also screened the same set of 26 SGA query mutant strains against the complete collection of nonessential gene deletion mutants and essential gene TS alleles, in the absence or presence of Benomyl (fig. S6, A and B, and data file 3) (*20*, *23*). This analysis also revealed ~65% fewer differential interactions relative to genetic interactions, indicating that trends observed with the diagnostic array accurately reflect an unbiased genome-scale analysis (fig. S6C).

In an effort to discover what new functional information might be associated with differential interactions, we divided them into four classes: "reversed," "modified," "masked," and "novel" (data file 3) (*20*, *23*). These classes of differential interactions can be depicted schematically for those with either negative (Fig. 2A, right) or positive (fig. S7A) scores. The reversed class was the rarest and included gene pairs that showed significant but opposite genetic interactions in a condition versus the matched reference, accounting for less than 1% of all Benomyl differential interactions (Fig. 2B). Often these differentials involved at least one relatively weak interaction, which means that they are likely more prone to false positives and false negatives. Thus, the sign of a fitness-based genetic interaction does not frequently change in an altered

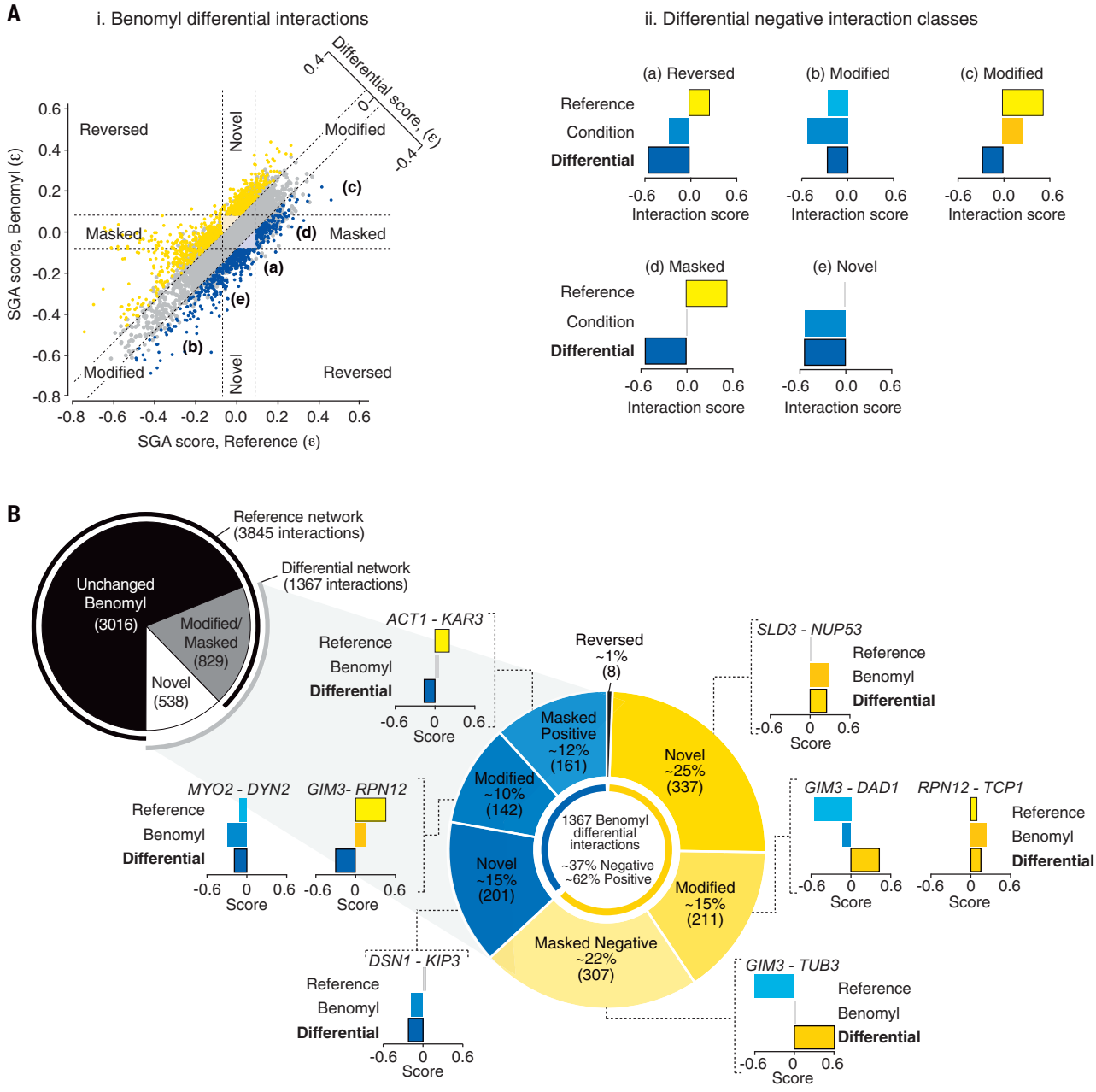


Fig. 2. Classification of differential interactions. (**A**) (i) Scatter plot of genetic interaction scores between the reference condition (*x* axis) and Benomyl-treated screens (*y* axis). Gene pairs with significant differential negative (blue) or differential positive (yellow) interactions are indicated. Schematic examples of differential negative interactions from each class as indicated by a letter, (a) to (e), on the scatter plot are shown in (ii). (ii) Schematic illustration of different classes of differential negative interactions. Specific classes of differential negative interactions are indicated and colored in shades of blue. Specific classes of differential positive interactions are indicated and colored in shades of yellow. (**B**) The number of genetic and differential interactions identified from SGA screens performed in Benomyl and in the reference condition. The fraction of significant reference condition genetic interactions that do not exhibit a significant differential interaction are shown in black. The fraction of differential interactions is shown in gray and white. (Inset) The outer black and gray rings indicate the size of the reference and differential interaction networks mapped in the reference and Benomyl conditions, respectively. The larger chart summarizes the total number of negative and positive Benomyl differential genetic interactions identified at an intermediate score threshold. The fraction of negative and positive interactions classified as reversed, novel, modified or masked is indicated. Specific classes of differential negative interactions are indicated and colored in shades of blue. Specific classes of differential positive interactions are indicated and colored in shades of yellow. Examples of gene pairs in each interaction subclass along with their corresponding reference, condition, and differential genetic interaction scores are shown.

environment. The remaining classes were roughly equivalent in size and accounted for the majority (>99%) of all Benomyl differential interactions (Fig. 2B and data file 3) (*20*).

We identified 538 (538 of 1367; ~39%) novel differential interactions, which were not observed as genetic interactions in the reference network. The remaining (829 of 1367; ~61%) differential interactions were either modified or masked, meaning that they were scored as genetic interactions in a particular condition as well as in the reference control but differed

in relative strength, and thus were modulated by the environment. Genome-wide analysis revealed similar fractions of modified, masked, and novel Benomyl differential interactions (fig. S6D).

Analysis of specific examples of modified differential interactions from the Benomyl screens revealed that modified differential interactions could arise in two ways (Fig. 2A and fig. S7A). First, a genetic interaction detected in the reference network may be exacerbated in a specific condition, resulting in a significant differential score. For example, a negative genetic interaction between *MYO2* and *DYN2* was stronger in Benomyl, reflecting the role of these two motor proteins in nuclear positioning and spindle orientation, especially when microtubule function is compromised (Fig. 2B). In another example, a positive genetic interaction between the proteasome gene *RPN12* and *TCP1*, which encodes an essential subunit of the chaperonin-containing T-complex involved in tubulin folding (Fig. 2B), was enhanced in the presence of Benomyl. In this case, compromising proteasome function may impair the degradation of the *TCP1* TS allele product, which is particularly relevant in the presence of Benomyl.

Second, a negative differential score may arise when a positive genetic interaction in the reference network is weakened in a specific condition (Fig. 2A and fig. S7A). For example, a positive genetic interaction between *GIM3* and *RPN12*, which encode subunits of the Prefoldin chaperone complex and the 19*S* proteasome, respectively, was weaker in the presence of Benomyl, resulting in a negative differential score (Fig. 2B). Similarly, a positive differential score can result from a condition-specific change in magnitude of a negative genetic interaction. For example, in the reference condition, *GIM3*, which is involved in tubulin folding, showed a relatively strong negative genetic interaction with *DAD1*, which is an essential kinetochore gene (Fig. 2B). In Benomyl, the *GIM3-DAD1* negative genetic interaction was significant but much weaker, resulting in a differential positive interaction (Fig. 2B), which may reflect that Benomyl perturbs the microtubule cytoskeleton in a manner that encompasses the cell physiology associated with the *GIM3-DAD1* interaction.

We also categorized a specific subset of modified differential interactions, which we call masked, in which positive or negative genetic interactions were only identified in the reference condition (Fig. 2A and fig. S7A). For example, whereas a negative genetic interaction was scored in the reference condition for *GIM3* and *TUB3*, which encodes α-tubulin, no genetic interaction was observed in the presence of Benomyl, leading to a *GIM3-TUB3* differential positive interaction (Fig. 2B). Positive genetic interactions scored in the reference condition can also be masked. For example, a positive genetic interaction between the yeast actin gene *ACT1* and the microtubule motor encoding gene *KAR3* was no longer detectable in Benomyl, leading to an *ACT1-KAR3* differential negative interaction (Fig. 2B).

Although relatively rare, novel differential interactions are noteworthy because they are revealed only in a specific condition, but not in the reference condition, and thus should reflect functional links between genes that are driven by condition-dependent cellular physiology (Fig. 2A and fig. S7A). For example, we identified a novel negative differential interaction between *DSN1* and *KIP3* in Benomyl that highlights a functional link between the kinetochore and a microtubule motor protein involved in spindle assembly (Fig. 2B).

Multiple environments, differential interactions, and the reference genetic network

We applied our scoring and classification system for differential genetic interactions, described above, to all 14 conditions surveyed. On average, less than 3% of all gene pairs tested in any given condition showed a differential interaction compared with ~13% of all gene pairs that exhibited a genetic interaction in the reference condition, which is consistent with the percent of overlapping gene pairs that show a genetic interaction in the context of a genome-wide study (fig. S7, B and C, and data file 3) (*11*, *20*). Comparison of negative interactions, in particular, revealed a notable difference in which ~1% of tested gene pairs exhibited a significant differential negative interaction, relative to ~8% of all gene pairs that showed a negative genetic interaction in any single condition tested (Fig. 3A, fig. S7C, and data file 3) (*20*). Two-thirds of all differential interactions were classified as modified or masked because they overlapped a genetic interaction in the reference condition (Fig. 3B). On average, modified and masked differential interactions from a single condition accounted for ~14% of all genetic interactions in the reference network (Fig. 3B). Depending on the condition, we estimate that between ~5 and 24% of genetic interactions detected in a reference genetic network can be modulated in a different environment.

Novel differential interactions provide a direct estimate of the additional functional information that different environments can contribute to a genetic network. On average, when compared with a matched reference control, ~7% of genetic interactions identified in a single condition were classified as novel differential interactions (Fig. 3B and data file 3) (*20*). Thus, the vast majority (~93%) of the genetic interactions mapped in different environmental conditions were also on the reference network. Novel differential negative interactions were also significantly weaker in magnitude as compared with genetic interactions measured in the reference condition (Fig. 3C). Detailed analysis of previously reported differential interaction networks (*14*, *15*, *17*) confirmed that differential interactions—in particular, novel differential interactions—were much less abundant than genetic interactions (fig. S7D).

It is critical to account for the false negative rate associated with high-throughput genetic interaction screens because failure to detect a true reference condition genetic interaction can be mistakenly classified as a novel differential interaction. We applied a Markov chain Monte Carlo (MCMC) modeling approach to generate a robust consensus set of negative and positive reference condition genetic interactions for each of the 26 SGA query genes based on the collection of 14 independent, biological replicate screens (*23*). The resultant consensus genetic interaction profiles were used as a gold standard to estimate false discovery and false negative rates at defined confidence thresholds (tables S1 and S2 and data file 4) (*20*). From analysis of reference condition replicate screens, we estimated false negative rates of 39 and 52% for negative and positive genetic interactions, respectively (table S1) (*23*). In the case of Benomyl, we mapped 538 novel differential interactions when compared with a reference condition network derived from a single, matched control (Fig. 3D and fig. S7C). However, a more rigorous comparison of Benomyl genetic interactions versus the MCMC consensus reference genetic interaction profile identified 454 novel differential interactions or ~15% fewer interactions in comparison with a matched control screen (Fig. 3D). A third comparison using a saturated reference network—on the basis of a reference network derived from the union of all the control screens (fig. S7E), which is less prone to false-negative interactions—identified 319 novel differential interactions, or ~40% fewer interactions as compared with those identified by using a single matched reference control (Fig. 3D) (*23*). The number of novel differential interactions could decrease by as much as ~60% depending on the condition and whether the reference was based on a single matched control, a consensus network control (*23*), or the union of reference control screens (Fig. 3D and data file 4) (*20*). Our analysis highlights the importance of a robust reference network, replicate screens, and rigorous estimates of false-negative rates for comparative study of genetic interactions in an alternative environmental condition.

Properties of differential interactions

Comparing all 14 conditions with their matched controls identified a combined total of ~10,000 differential interactions, the majority (61%) of

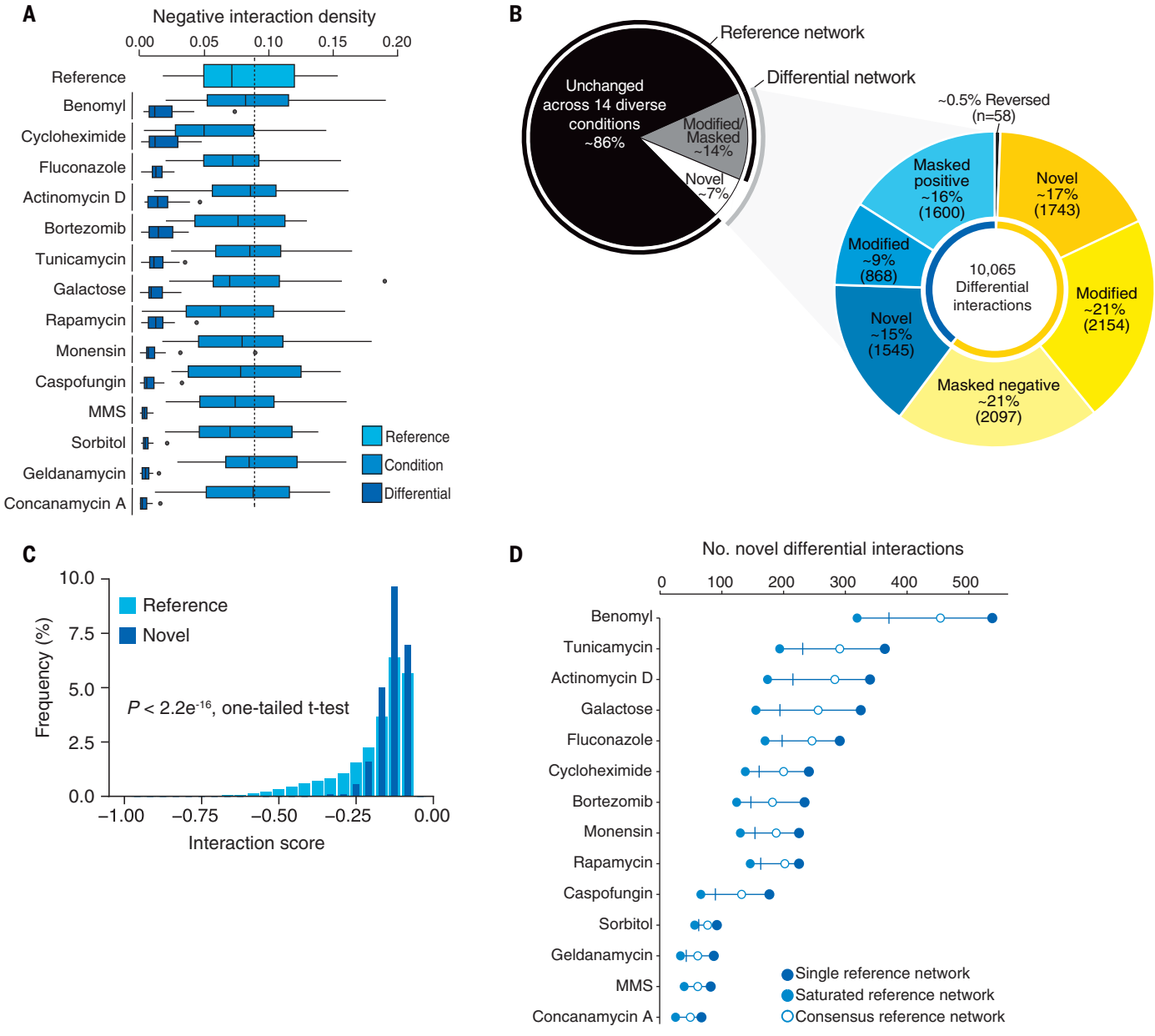


Fig. 3. The relative contribution of reference genetic interactions and differential interactions to the yeast genetic network. (**A**) Box plots showing the distribution of negative genetic and differential negative interaction density (total number of interactions per total gene pairs tested) per condition and per query mutant screened. The dotted line indicates the average genetic interaction density for the same set of array genes in the global genetic interaction network (*11*). (**B**) The average fraction of genetic and differential interactions identified from SGA screens performed in the reference and one additional condition. The fraction of reference condition genetic interactions that do not exhibit a significant differential interaction are shown in black. Modified and masked differential interactions, which overlap with interactions identified in the reference condition, and novel differential interactions are indicated. The outer black and gray rings indicate the average size of the reference and differential interaction networks mapped for one additional condition, respectively. The colored diagram summarizes the total number of negative and positive differential genetic interactions identified at an intermediate score threshold from analysis of 14 different test conditions. The fraction of negative (shades of blue) and positive (shades of yellow) interactions classified as reversed, novel, modified, or masked is indicated. (**C**) Distribution of negative genetic interaction scores (light blue) and novel differential negative interaction scores (dark blue). (**D**) The number of novel differential interactions identified per condition by using different reference genetic interaction networks. Reference networks were defined by exhaustively sampling all possible combinations from 1 to 14 biological replicate screens. The maximum number of novel differential interactions based on comparison with a single matched reference network for each condition is shown in dark blue. The minimum number of novel differential interactions based on comparison with a saturated reference network that combines all 14 replicates (union) is shown in light blue. The number of novel interactions based on comparison with a consensus MCMC-derived reference network (*23*) is also indicated (open circles). Vertical bars indicate the median number of novel differential interactions across all possible reference network combinations.

which were specific to a single growth condition, with differential negative interactions being less prevalent and relatively weaker than differential positive interactions (Fig. 3B and fig. S7F). Genes that were highly connected hubs on the global genetic network were also more likely to be hubs on a differential network because the average number of differential interactions for an individual array gene, across all 14 conditions, was significantly correlated to the interaction degree in the global genetic network [Pearson correlation, correlation coefficient (r) = ~0.7, $P < 10^{-16}$] (fig. S8, A and B). The number of novel differential interactions associated with each query mutant examined in this study was also correlated to interaction degree in the global genetic network (Pearson correlation, r = ~0.5, $P < 0.006$) (fig. S8C), suggesting that the frequency of novel differential interactions observed by using a diagnostic set of genes should reflect the genome-wide prevalence of novel differential interactions. Consistent with these observations, gene features associated with hubs on the reference network were shared with differential network hubs (fig. S8D) (*11*). Notably, high novel differential interaction degree was associated with genes whose loss of function resulted in a single-mutant fitness defect in the reference condition (fig. S8E), and genes with condition-dependent fitness defects had proportionally more novel differential interactions than those of genes lacking condition-dependent fitness defects (fig. S8F). Like genetic interaction hubs, genes with many novel differential interactions were associated with multiple Gene Ontology (GO) annotations and lower dN/dS [the ratio of the number of nonsynonymous substitutions per nonsynonymous site (dN) to the number of synonymous substitutions per synonymous site (dS)], suggesting that they are more functionally pleiotropic and tend to be under stronger evolutionary constraints (fig. S8, D and E). Conversely, transcript levels of high differential interaction degree genes did not vary substantially across different environments or genetic backgrounds, suggesting that

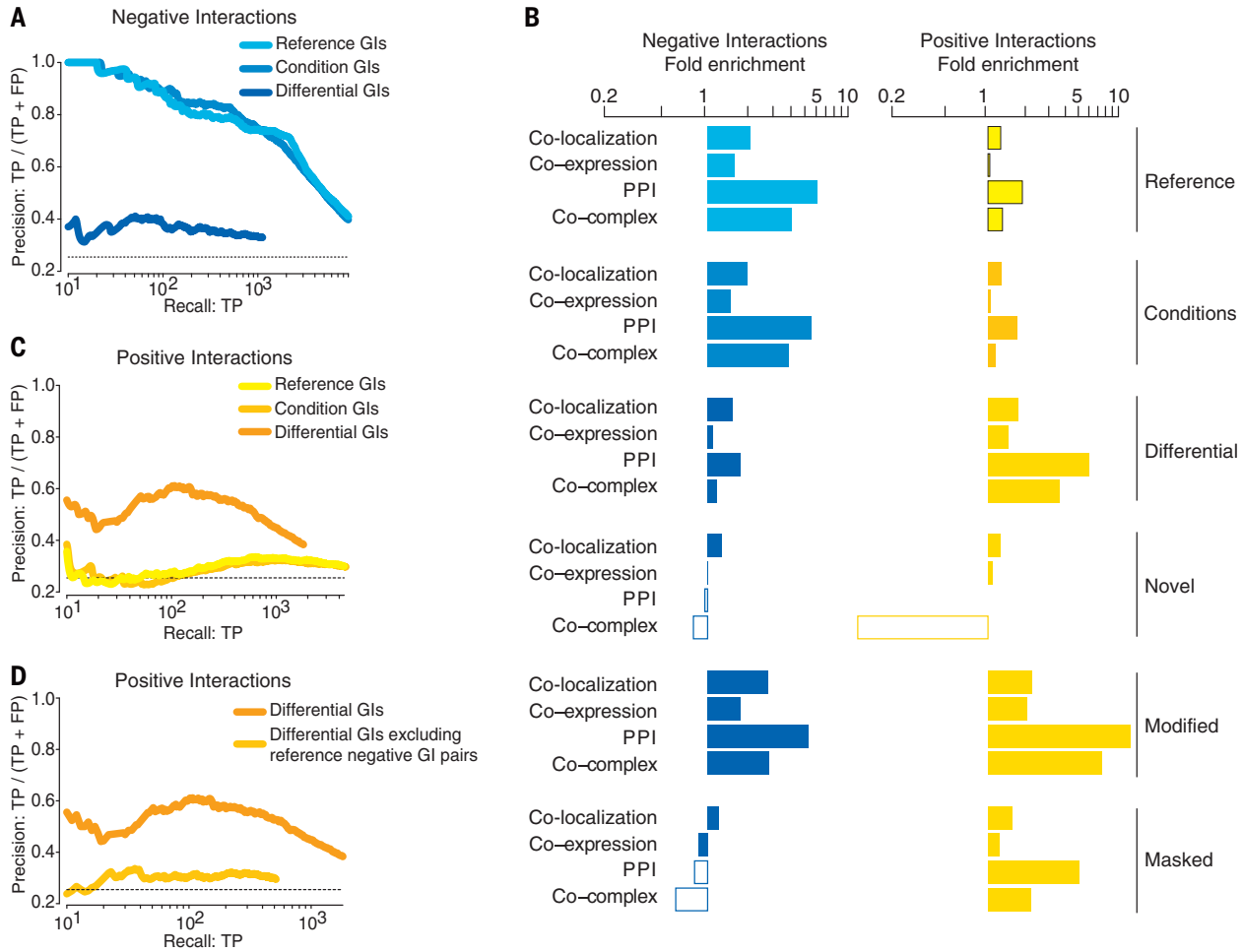


Fig. 4. Functional evaluation of differential interactions. (**A**) Plots of precision versus recall for negative genetic and differential interactions, as determined by our genetic interaction score ($|\epsilon| > 0.08$, $P < 0.05$). True-positive (TP) interactions were defined as those involving gene pairs co-annotated to a gold standard set of GO terms, as defined elsewhere (*55*). False-positive (FP) interactions were defined as those involving pairs of genes annotated to different GO terms, as defined elsewhere (*55*). The background precision at which true positives are randomly identified is indicated by the dotted line. The precision and recall values were calculated as previously described (*19*). (**B**) Fold enrichment for negative (blue) and positive (yellow) genetic and differential interactions among colocalized, coexpressed, physically interacting, or co-complexed gene pairs or their encoded proteins were calculated for genetic interactions identified in the reference condition, 14 conditions, and each differential interaction class. Comparisons based on fewer than 10 overlapping gene pairs are indicated (open boxes). (**C** and **D**) Plots of precision versus recall (number of TP) for positive genetic and differential interactions, as determined by our genetic interaction score ($|\epsilon| > 0.08$, $P < 0.05$). True-positive interactions were defined as those involving gene pairs co-annotated to a gold standard set of GO terms, as defined elsewhere (*55*). False-positive (FP) interactions were defined as those involving pairs of genes annotated to different GO terms, as defined elsewhere (*55*). The background precision at which true positives are randomly identified is indicated with the dotted line. The precision and recall values were calculated as described (*19*).

environmentally responsive gene expression patterns are not generally predictive of differential interactions (fig. S8D).

Negative genetic interactions measured in any condition tended to connect functionally related gene pairs and overlapped with other types of molecular interaction networks (Fig. 4, A and B). Differential negative interactions also identified functionally related gene pairs, but to a lesser extent (Fig. 4, A and B). Closer examination revealed that most of the functional signals associated with differential negative interactions were captured by the modified class, whereas those belonging to the novel or masked categories did not overlap substantially with other molecular interaction datasets (Fig. 4B). Thus, negative genetic interactions that occurred in a condition-specific manner (novel differential negative interaction) (Fig. 2A) and positive genetic interactions that are masked in a particular condition (masked differential negative interaction) (Fig. 2A) often involve gene pairs with unrelated functional annotations.

Consistent with previous observations (*11*), positive genetic interactions identified in either the reference or alternative conditions were less functionally informative than negative genetic interactions (Fig. 4, A and C). However, the complete set of differential positive interactions appeared to connect functionally related genes more often than positive genetic interactions measured in either the reference or individual test conditions (Fig. 4C). Modified and masked differential positive interactions were the most predictive of functionally related gene pairs, often connecting members of the same protein complex and overlapping substantially with protein-protein interactions (Fig. 4B). Further analysis revealed that the functional signal associated with the masked and modified classes was largely attributable to gene pairs that displayed negative genetic interactions in the reference condition but were either weaker (modified) or no longer detectable (masked) in a particular condition (Fig. 4D). This may reflect a particular environmental perturbation that mimics the cellular physiology associated with a double-mutant strain grown in the reference condition, obscuring a phenotype associated with a genetic interaction. Thus, although modified and masked differential positive interactions tend to connect functionally related gene pairs, this same information is captured by negative genetic interactions identified in the reference condition. Genome-wide analysis revealed similar functional trends associated with Benomyl differential negative and positive interactions (fig. S6, E and F).

To further explore the functional information associated with condition-specific interactions, we grouped together array genes that belong to the same biological process cluster represented on the global genetic profile similarity network (Fig. 1A) and measured how often each query gene showed either genetic or novel differential interactions with each functional group (Fig. 5 and data file 5) (*11*, *20*, *23*). Query and array genes within the same bioprocess cluster were often connected by negative genetic interactions in the reference condition (~3.6-fold enrichment within bioprocess) (Fig. 5A, on diagonal). For example, in the reference condition, the *VTI1* query gene, which encodes an essential SNAP [soluble *N*-ethylmaleimide–sensitive factor (NSF) attachment protein] receptor (v-SNARE) that is involved in multiple protein sorting pathways (*33*–*35*) and located in the vesicle traffic bioprocess cluster, showed strong enrichment for negative genetic interactions with functionally related array genes located in the same vesicle traffic bioprocess cluster (Fig. 5B).

By contrast, novel differential negative interactions did not connect query and array genes annotated to the same biological process (Fig. 5C). Although *VTI1* showed the most novel differential interactions of any query gene tested (data file 3) (*20*), most of these interactions did not involve other vesicle traffic–related genes. Instead, we observed modest but significant enrichment for novel differential negative interactions between the *VTI1* query gene and array genes with roles in cell polarity and nuclear transport (Fig. 5D). The vast majority (1489 of 1553, ~96%) of all novel differential negative interactions identified by using a matched reference control connected pairs of genes located in different bioprocess clusters on the global similarity network. These results further indicated that condition-specific genetic interactions do not identify gene pairs with a close functional relationship in the same general bioprocess but rather have the potential to uncover weaker functional associations between distinct biological processes.

Differential interactions capture distant but coherent functional relationships

We next explored the functional distribution of novel differential interactions across each of the 14 different conditions. In particular, we tested whether array genes annotated to the same function showed more novel differential interactions in a particular condition (Fig. 6). Array genes annotated to specific biological processes were enriched for novel differential negative interactions in response to a specific condition that perturbs the same bioprocess (~2.3 fold enrichment within bioprocess) (Fig. 6A, on-diagonal). For example, consistent with previous observations (*14*, *15*), we found that array genes in the DNA replication and repair bioprocess cluster were enriched for novel differential negative interactions in the presence of methyl methanesulfonate (MMS) (3.9-fold) (Fig. 6B). But as shown above (Fig. 5), these novel differential interactions did not involve related query genes with roles in DNA replication and repair. The enrichment observed among DNA replication and repair array genes was predominantly driven by MMS-specific novel differential negative interactions with the *MYO2* query gene, a type V myosin motor involved in actin-based vesicle transport and spindle orientation, and with the *VTI1* and *TRS20* query genes, which are involved in vesicle transport, highlighting a functional link between DNA replication and vesicle trafficking (Fig. 6B) (*33*–*35*). In another example, vesicle traffic genes were enriched for novel differential negative interaction in response to Monensin (2.5-fold) (Fig. 6A). In this case, Monensin-specific negative interactions connected vesicle traffic array genes to the *RSP5* query gene, which encodes a ubiquitin ligase involved in multivesicular body sorting, the heat shock response, endocytosis, and ribosome assembly (*36*), as well as the *LSM6* query gene, which has a general role in RNA processing (Fig. 6B) (*37*, *38*). Novel differential positive interactions were also not enriched among functionally related genes, and they were less informative of gene function than were novel differential negative interactions (fig. S9A and data file 5) (*20*).

A similar trend was observed in genome-scale Benomyl screens, in which novel differential negative interactions specifically connected array genes with mitosis-related roles to functionally diverse query genes, such as *NUP188*, which encodes a nuclear pore component (*39*), and *GPI15*, involved in glycosylphosphatidylinositol (GPI) anchor biosynthesis (fig. S9B and data file 5) (*20*, *40*, *41*). Thus, environment may sensitize genes with roles in a specific bioprocess to negative genetic interactions with functionally diverse query genes.

The majority of differential interactions overlap a genetic interaction identified in the reference condition, which often connect functionally related genes within the same biological process. The environmental rewiring of genetic networks is driven by rare condition-specific and relatively weak novel differential interactions, which identify new connections between genes with diverse functions in different biological processes (Fig. 7).

Discussion

We surveyed a set of functionally diverse yeast genes and different environments, quantifying how the growth phenotypes associated with different single-gene mutations were modulated by the environment (GxE) and how the growth phenotypes associated with different genetic interactions (GxG) respond dynamically to a particular condition to generate differential interactions (GxGxE). Our general findings reveal how environmental conditions modulate the yeast global genetic interaction

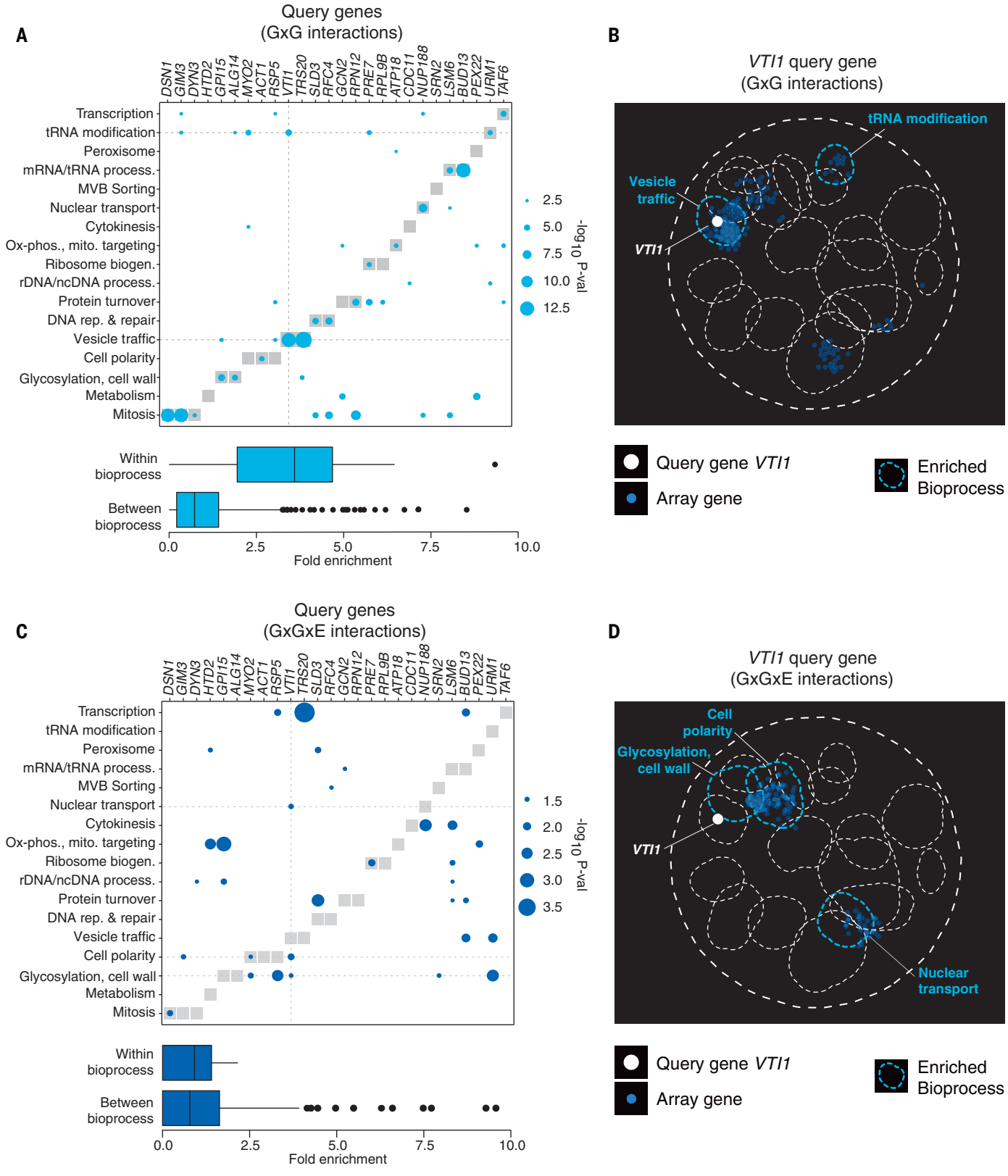


Fig. 5. Functional distribution of query gene negative genetic and novel differential negative interactions. (**A**) Negative genetic interactions for each query gene (*x* axis) in the reference condition were tested for enrichment for array genes in each of the biological processes indicated (*y* axis). Node size reflects the statistical significance of enrichment, and the shade boxes along the diagonal indicate instances in which the query and array genes belong to the same biological process cluster on the global genetic interaction profile similarity network (*11*). Dotted lines indicate bioprocesses enriched for negative genetic interactions with the *VTI1* query gene. The average fold enrichment of negative genetic interactions within and between specific biological processes

is shown in the box plot. (**B**) Regions of the global genetic interaction profile similarity network significantly enriched for array genes exhibiting negative genetic interactions with the *VTI1* query gene in the reference condition were mapped by using SAFE (*54*). The functional regions of the global similarity network significantly enriched for interactions with *VTI1* are indicated with blue dotted lines. Array genes enriched for negative genetic interactions are shown in blue. The location of the *VTI1* query gene on the global similarity network is indicated by a white node. (**C**) Novel differential negative interactions for each query gene (*x* axis) across all 14 test conditions were tested for enrichment for array genes in each of the biological processes (*y* axis). Node size reflects the statistical significance of enrichment, and the shaded boxes along the diagonal indicate instances in which the query and array genes would have belonged to the same biological process cluster on the global genetic interaction profile similarity network (*11*). Dotted lines indicate bioprocesses enriched for novel differential negative interactions with the *VTI1* query gene. The average fold enrichment of novel differential negative interactions within and between specific biological processes is shown in the box plot. (**D**) Regions of the global genetic interaction profile similarity network significantly enriched for array genes exhibiting novel differential negative interactions with the *VTI1* query gene were mapped by using SAFE (*54*). The functional regions of the global similarity network significantly enriched for interactions with *VTI1* are indicated with blue dotted lines. Array genes enriched for novel differential negative interactions are shown in blue. The location of the *VTI1* query gene on the global similarity network is indicated with a white node.

network, allowing us to assess the plasticity of genetic networks and the extent to which mapping genetic interactions in different environments can expand a reference network and generate new functional information.

In general, genetic interactions, especially negative genetic interactions, are rich in functional information because they tend to connect genes that function within the same biological process (Fig. 7B) (*11*). Analogously, if an environmental perturbation affects a particular biological process, then genes with roles in the perturbed bioprocess tend to show differential sensitivity in the corresponding condition (Fig. 7A). Most genetic interactions are not modulated by the environment and remain detectable in both a given test condition and the reference control condition. However, subsets of genetic interactions can be modified by a particular condition, leading to a differential interaction (Fig. 7, C and D). Differential interactions are relatively rare because, on average, we observed approximately threefold more genetic interactions in the reference condition as compared with differential interactions detected in any other single environment.

Most differential interactions belonged to the modified or masked classes, which overlapped with a negative or positive genetic interaction in the reference genetic network. The majority (~70%) overlapped specifically with negative genetic interactions and thus were functionally informative, connecting genes belonging to the same general biological process (Fig. 7C). However, because they recapitulate connections that are captured in the global genetic network mapped in the reference condition, modified and masked differential interactions do not contribute new functional information. On the other hand, novel differential interactions correspond to condition-specific genetic interactions between genes that do not interact in the reference condition. The rare subclass of novel differential negative interactions does not typically include functionally related gene pairs. Instead, these interactions tend to connect groups of genes that are sensitive to a particular environmental perturbation to functionally distant genes (Fig. 7D). Thus, novel differential interactions highlight new connections between distinct cellular functions and, as a result, have the potential to expand the global genetic network.

Although detecting novel differential interactions promises to add new information to genetic networks, an accurate definition of novel differential interactions is challenging because it depends on the quality and comprehensiveness of the reference genetic network. For example, multiple independent replicate screens reduce the number of false-negative interactions observed in the reference condition but consequently also decrease the number of differential interactions classified as novel upon environmental perturbation (Fig. 3D). The relative fraction of novel differential interactions contributed to the reference network by a single environmental condition can vary widely when using a single, matched control (~7%) or the union of all available replicate controls (~1%) (data file 4) (*20*). Using a high-confidence consensus reference network (data file 4) (*20*), we estimate that screening one additional environmental condition contributes less than ~4% (191 of 5174) novel interactions relative to the reference network, suggesting that most genetic interactions are captured in a single condition. Hence, although differential interaction analysis in multiple diverse conditions reveals the subset of genetic interactions that are modulated upon environmental perturbations (modified or masked interactions), they can only marginally expand the size of the network (novel interactions), highlighting that the global genetic interaction network is generally robust to environmental perturbations.

Although our study involved the use of a diagnostic array of yeast genes, selected environmental conditions, and functionally diverse query genes, several observations suggest that our results capture the general resilience of the global yeast genetic interaction network to environmental perturbation. First, our genome-scale single-mutant fitness analyses suggested that the selected set of conditions and small molecules elicited widespread but distinct cell physiological effects. Second, a genome-wide comparison of genetic interactions measured in the absence and presence of Benomyl yielded a similar fraction of differential interactions seen by using a diagnostic mini-array of genes. Last, analysis of several independent interaction datasets—derived from SGA-based approaches by using various subsets of yeast genes and a number of different conditions, including various DNA damaging agents and stress-response conditions—confirmed that differential interactions are substantially less prevalent than the genetic interactions observed in the reference condition (fig. S7D) (*14*, *15*, *17*).

Consistent with our results, previous surveys of subsets of genes involved in a specific cellular function perturbed by a particular condition revealed an enrichment for differential interactions that often occurred between functionally distant gene pairs (*14*, *15*, *17*). Although included in previous studies, a subset of differential interactions, based on statistically insignificant genetic interaction scores, was omitted from our analysis (*23*). However, these particular differential interactions were not enriched among functionally related genes and did not appreciably increase the total number of differential interactions relative to the size of the reference genetic interaction network examined here or reported in other studies (figs. S7D and S10). Thus, in general, genetic interactions between the vast majority of genes and their corresponding functional modules (such as complexes and pathways) are not dynamic or rewired in response to environmental stimuli. Moreover, we found that the ~20% of yeast genes with relatively sparse genetic interaction profiles (*11*) are statistically depleted for condition-specific fitness defects ($P < 10^{-99}$, Fisher's exact test, one-tailed). The relationship between fitness and interaction degree (fig. S8) along with the strong correlation observed between interaction degree for a given gene in the global genetic and differential networks (fig. S8) further suggests that condition-specific genetic interactions will not appreciably increase the number of interactions associated with low-degree genes in the global reference genetic network.

Differential interactions reflect the phenotypic consequences of combining three independent perturbations, two genetic and one

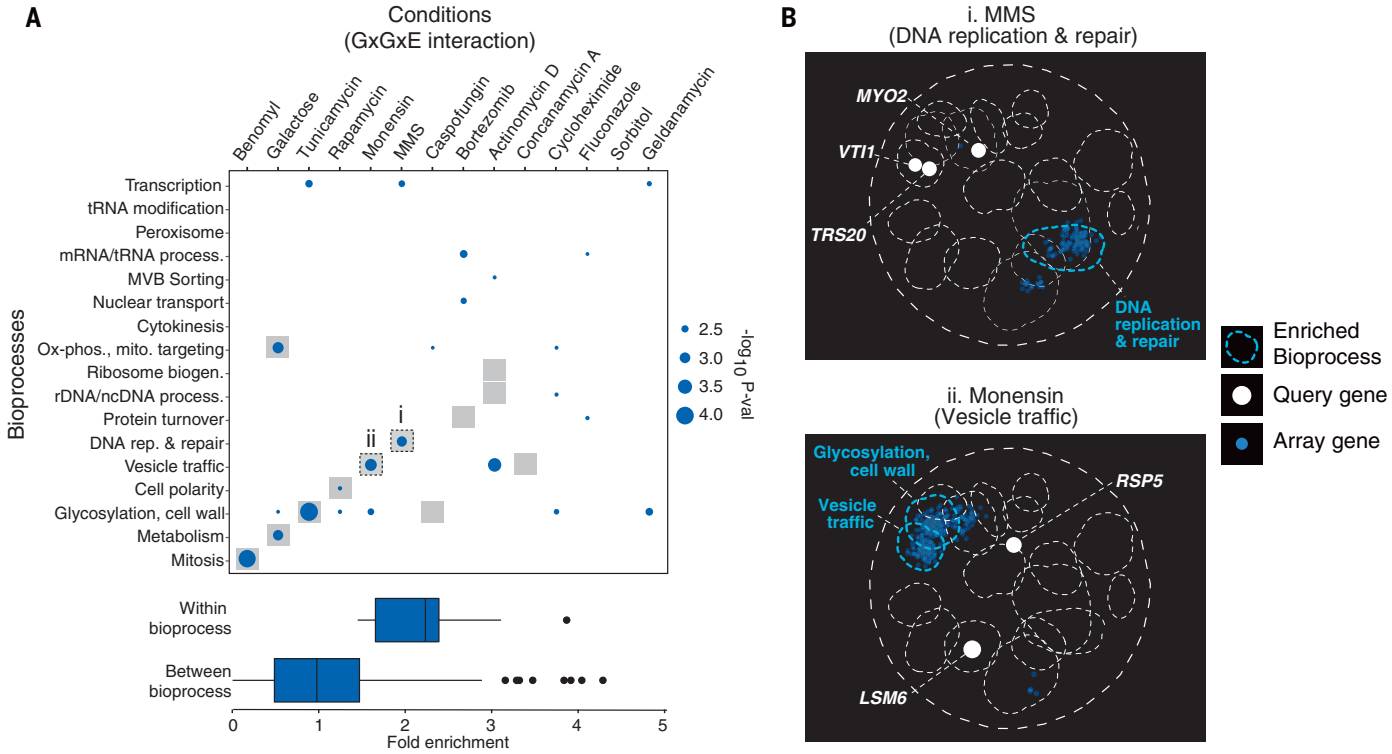


Fig. 6. Functional distribution of novel differential negative interactions across conditions. (**A**) Novel differential negative interactions from each of the 14 conditions (*x* axis) were tested for enrichment for array genes grouped according to the biological processes indicated (*y* axis). Node size reflects the statistical significance of enrichment, and the shaded boxes along the diagonal indicate the biological process targeted by a particular condition. The average fold enrichment of novel negative interactions within biological processes targeted by a specific condition and interactions enriched within biological processes unrelated to the condition are shown in the box plot. (**B**) Regions of the global similarity network significantly enriched for array genes exhibiting novel negative differential interactions in (i) MMS or (ii) Monensin were mapped by using SAFE (*54*). The functional region of the global similarity network targeted by Monensin and MMS are indicated with blue dotted lines. Array genes enriched for novel differential negative interactions are shown in blue. Query gene(s) responsible for enrichment of novel differential interactions with the indicated array genes are shown as white nodes.

environmental perturbation, and thus conceptually resemble trigenic interactions, which involve three independent genetic perturbations, especially if a condition involves a drug with a highly specific cellular target. As a result, differential and trigenic interaction networks share several properties in common: (i) As observed for differential interactions, the average trigenic interaction degree for a given gene was correlated to its connectivity in the global genetic network (*6*); (ii) like novel differential interactions, trigenic interactions occur at similar reduced frequencies relative to digenic interactions in a reference condition (fig. S7D); (iii) like modified and masked differentials, a substantial proportion of trigenic interactions overlapped with and exacerbated digenic interactions previously observed in the global digenic network (*6*); and (iv) as observed for novel differential interactions, novel trigenic interactions are also relatively rare and weaker but are highly coherent, in that they often involve array genes from the same biological process—however, they are also substantially more functionally diverse than digenic interactions (*6*).

GxE and GxGxE interactions involving the natural variation of different yeast strains can also be interpreted in the context of the global genetic interaction network and the properties of novel differential interactions. Using segregating populations of natural yeast isolates, recent surveys of GxG and GxGxE interactions between a deletion mutant and natural genetic variants showed that GxG interactions tend to connect genes within the same cellular function (*42*), whereas GxGxE interactions tend to be specific to a particular condition (*43*), and they often connected distantly related genes (*44*).

Quantitative trait loci (QTL) studies, focused on the fitness variation of different segregants derived from yeast crosses in different environments, showed that the majority of the phenotypic variation can be explained by single-locus effects (GxE), whereas epistatic genetic interactions, analogous to GxGxE interactions, are relatively rare (*4*, *45*) and tend to affect only individuals with extreme phenotypes (*5*). On the basis of our analyses, most GxGxE interactions involve genes with single-mutant differential effects (GxE) (fig. S8E), which should be analogous to single-locus effects identified in QTL analyses and contribute to both the additive and epistatic variance; that is, the low epistatic variance often seen in QTL analyses may be partly explained by the partial contribution of interacting loci to the additive variance component (*46*). As a result, the remaining epistatic effect may only be visible in individuals with the most extreme phenotypes (*5*). Given that most GxGxE interactions overlap a genetic interaction in the global genetic interaction network, applying principles and gene modules inferred from a global network should allow for the detection of coherent epistatic interactions with relatively small effects that cannot be recovered by using QTL analyses (*47*, *48*).

We conclude that the yeast genetic interaction map derived from a single reference condition is highly robust to environmental influences because most connections on the global genetic interaction network remain unchanged or unmodified in a new environment. Although each new condition has the potential to map a relatively small number of novel differential interactions, the global digenic network mapped in a single condition is

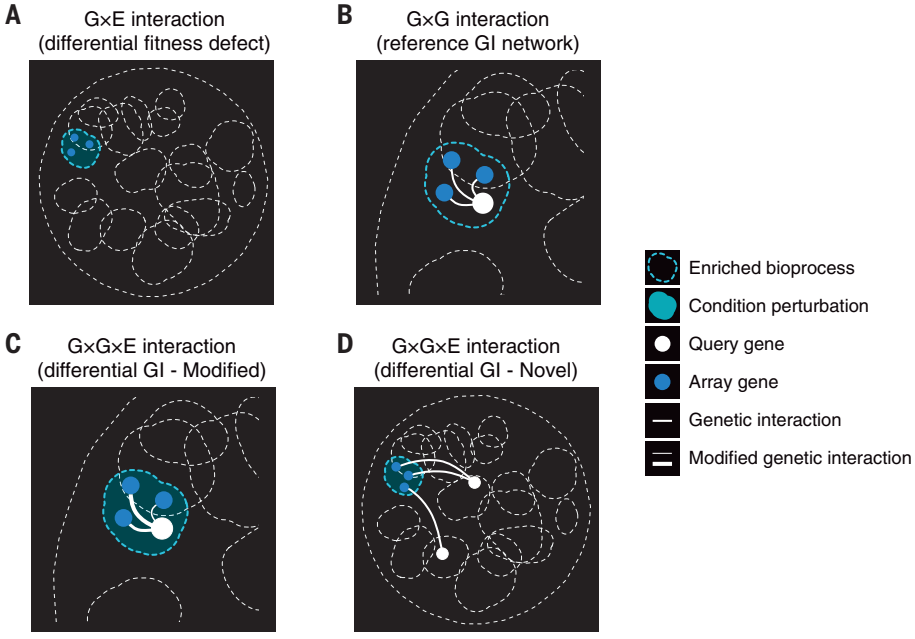


Fig. 7. Schematic summary of functional connections mediated by GxE, GxG, and GxGxE interactions. (**A**) GxE interactions. Genes with roles in the same bioprocess (illustrated with dotted outlines) are often sensitive to an environmental perturbation affecting that process (blue nodes, array genes; shaded blue area, condition perturbation). (**B**) GxG interactions. Genetic interactions (white edges) tend to connect array and query genes (white nodes) functioning in the same bioprocess. (**C**) GxGxE interactions (modified). Many genetic interactions are modulated by the environment (white edges), creating a modified differential genetic interaction that varies in magnitude when compared with the equivalent genetic interaction on the reference network. Modified GxGxE interactions often connect functionally related query-array gene pairs. (**D**) GxGxE interactions (novel). Novel gene interactions, which were not detected in the reference condition, often connect genes involved in a bioprocess perturbed by a specific condition to distant, functionally unrelated genes.

informative of differential interactions (fig. S8, A and B) and should serve as an accurate and representative reference network. Nonetheless, differential interactions can reveal new functional connections between distantly related genes and bioprocesses. Given the significant association between the differential single-mutant fitness defect and the frequency of differential interactions for a given gene, a logical and efficient strategy may involve mapping differential interactions for specific genes required for normal growth in an environment of interest (such as stress condition, drug treatment, or microbial exposure).

Although environment has a modest impact, cell type–specific gene regulation has the potential to substantially complicate genetic network analyses in more complex systems. However, genome-wide studies suggest that only ~8500 genes (~45 to 49%) are expressed in any given cancer cell line and that a "core set" of ~7700 genes are expressed in the majority of all cancer cell lines examined to date (*49*). Thus, a systematic survey of genetic interactions based on a single cell line grown in one environment should be sufficient to map a global reference genetic network that encompasses this core set of expressed genes and provide a basic scaffold for the genetic wiring of a human cell. As observed in yeast, a global genetic interaction network for a model human cell line should reveal a hierarchy of functional connections among genes (*1*, *11*), and expansion of this network to specific cell types would enable the deciphering of mechanisms that underlie cancer cell–specific genetic vulnerabilities (*50*, *51*). Ultimately, reference genetic networks mapped in human cells should facilitate interpretation of allelic combinations of genes underlying inherited traits (*3*, *47*, *48*).

Methods summary

Nonessential deletion and essential TS mutant arrays

The complete nonessential gene deletion and essential gene TS allele arrays were used for all single-mutant fitness and single-mutant differential fitness (GxE) analyses and are described elsewhere (*11*).

Diagnostic mutant array

Unless otherwise noted, genetic and differential interactions were based on analysis of query mutant strains crossed to a previously described, functionally representative diagnostic array comprising 1209 mutant strains (*6*). These included 1012 nonessential deletion mutant array (DMA) strains (labeled_dma#) and 197 essential TS allele array (TSA) mutant strains (labeled_tsa#). A complete list of diagnostic mutant strains used in this study is provided in data file 1 (*20*).

Query mutant strains

The set of 26 query mutant strains selected for this study were previously shown to have rich genetic interaction profiles that can recapitulate major bioprocess-enriched clusters on the genetic interaction profile similarity network and thus are representative of the global yeast genetic network (*11*, *52*). SGA query strain construction was conducted as described previously (*53*). A list of query mutant strains used in this study is provided in data file 1 (*20*).

Conditions

A list of 14 conditions and concentrations used in this study is provided in data file 1 (*20*).

SGA screening procedure

SGA experiments and selection steps were conducted as previously described (*11*, *53*), with the following modifications. To measure condition-dependent genetic interactions, every double-mutant array generated from a single-query SGA screen was copied three times. One copy was grown in the standard SGA reference condition, whereas the two other copies were each grown in different conditional media (fig. S4). This configuration provided a matched reference control for every condition, which facilitated normalization of systematic experimental artifacts and improved accuracy of condition-specific genetic interactions measurements, as described below. Because every double-mutant array could be replicated a maximum of three times, each of the 26 query mutant strains was independently screened against the diagnostic array seven times, allowing us to measure digenic interactions in 14 different test conditions (2 test conditions + 1 reference condition/screen × 7 screens/query strain) (fig. S4). The entire screening pipeline was repeated twice, resulting in independent biological replicates for each query screen in every test condition along with 14 biological replicates of each query screen in the standard, reference SGA condition. We also repeated screens for a subset of four query mutant strains (data file 3) (*20*), in which double-mutant arrays were copied two times and each copy was grown in the standard SGA reference condition to measure the reproducibility of genetic interactions derived from matched reference control screens (fig. S5C).

Single-mutant fitness scores

To derive accurate estimates of single-mutant fitness, we applied our colony size scoring method (*19*) to a set of control SGA screens, in which a query strain carrying a *natMX* marker inserted at a neutral genomic locus

was crossed to the *kanMX*-marked DMA (_dma#) and TSA (_tsa#) strain collections. Colony size measurements of SGA deletion and TS array mutant strains were based on an average of three replicate control screens conducted per each of 14 test conditions as well as the reference condition at 26°C. Colony size measurements were used to estimate single-mutant fitness in each condition as described previously (*19*), with the exception that bootstrapped means, instead of medians, across replicates were used in variance estimation and final fitness values. Because of technical reasons, single-mutant fitness for a small subset of mutant strains (0.15 to 1.0% of all strains) are not reported.

Single-mutant differential fitness scores

To obtain condition-specific fitness estimates, we computed the difference in colony size measured in a particular test condition versus the matched reference condition for each mutant. Single-mutant fitness and differential fitness estimates are provided in data file 1 (*20*). Because of technical reasons, single-mutant fitness for a small subset of mutant strains (0.15% to 1.0% of all strains) are not reported.

Genetic interaction score

To derive quantitative genetic interactions, we modeled colony size as a multiplicative combination of double-mutant fitness, time, and experimental factors as previously described (*11*, *19*). Briefly, for a double mutant carrying mutations of genes i and j, colony size C_{ij} can be expressed as $C_{ij} = f_{ij} \cdot t \cdot s_{ij} \cdot e$, where f_{ij} is the double mutant fitness, t is the incubation time, s_{ij} is the combination of all systematic factors, and e is log-normally distributed random noise. The double-mutant fitness f_{ij} can be further expressed as $f_{ij} = f_i f_j + \epsilon_{ij}$, where f_i and f_j represent the fitness of the two single mutants and ϵ_{ij} is a quantitative measure of the genetic interaction (genetic interaction score) between them.

Genetic interaction data corresponding to all tested gene pairs identified in the reference and 14 different conditions are provided in data file 3 (*20*). The data should be filtered before use. We suggest two different thresholds [intermediate ($P < 0.05$ and $|\epsilon| > 0.08$) and stringent confidence ($P < 0.05$ and $|\epsilon| > 0.12$)] that strike different balances between false negatives and false positives, as described in our previous studies (*11*, *12*). Although the majority of query mutant strains were screened in all test conditions, a smaller subset of query mutants could not be screened in all 14 conditions because of technical and data quality control reasons. In total, 22 of 26 of query mutants were screened in at least 12 test conditions, whereas five query mutants were screened for genetic interactions in less than 12 test conditions.

Differential interaction score

To score differential genetic interactions, genetic interaction scores derived from each condition were matched with a paired reference condition. Additionally, we used genetic interaction scores from the previously published reference genetic interaction network (*11*), termed "global" in the section below to normalize screen data before differential interaction scoring. The first step in scoring differential interactions was to apply a correction to each query, condition, and replicate screen that normalized the genetic interaction scores so that the corrected standard deviation for overlapping gene pairs matched the published genetic interaction network. We performed this per condition, per query, per replicate correction on reference (untreated) and condition (treated) scores as follows:

$$\text{correction}_{c,q,r} = \frac{\sigma\left(\epsilon_{\text{untreated},c,q,r}\right)}{\sigma\left(\epsilon_{\text{global},c,q}\right)}$$

$$\hat{\varepsilon}_{\text{untreated},c,q,r} = \text{correction}_{c,q,r} \times \epsilon_{\text{untreated},c,q,r}$$

where c is the condition, q is the query, r is the replicate, ε is the uncorrected epsilon score, σ is the calculated standard deviation of the reference condition (untreated) epsilon scores for that query, condition (treated) and replicate, or the global epsilon scores for that condition and query from our previous work (*11*), and $\hat{\epsilon}$ is the corrected epsilon score. This conforms the variance of interaction measurements in each condition to a fixed reference and facilitates comparisons of interaction density across different conditions. Each condition (treated) network has a paired reference (untreated) network, which is used to compute the correction factor above.

Three replicate scores were collected for every differential genetic interaction measurement. Some screens resulted in missing data for certain pairs because of data quality issues or strong fitness defects in the screened conditions. To mitigate the effect of missing data, we only considered interactions for which we had at least two replicate measurements. For interactions that had three replicates, we selected the two replicates with the highest correlation of corrected differential interaction scores for each query-condition pair.

We found that screening genetic interactions in compound stress conditions increased the variance in genetic interaction estimates. To mitigate this effect, we applied a variance stabilization correction to further correct conditional epsilon scores:

$$\text{variance stabilization}_c = \frac{\sigma\left(\left|\hat{\epsilon}_{\text{untreated},c,r=1} - \hat{\epsilon}_{\text{untreated},c,r=2}\right|\right)}{\sigma\left(\left|\hat{\epsilon}_{\text{treated},c,r=1} - \hat{\epsilon}_{\text{treated},c,r=2}\right|\right)}$$

$$\hat{\epsilon}_{\text{treated},c,q,r} = \text{variance stabilization}_c \times \hat{\epsilon}_{\text{treated},c,q,r}$$

where c is the condition, q is the query, r is the replicate, $\hat{\epsilon}$ is the corrected epsilon score, σ is the calculated standard deviation, and $\hat{\epsilon}$ is the variance-stabilized, corrected treated epsilon score. After applying these normalizations, the differential score is calculated for each of the paired replicates separately, as follows:

$$\text{differential score}_r = \hat{\epsilon}_{\text{treated},r} - \hat{\epsilon}_{\text{untreated},r}$$

A final differential score, final untreated score, and final treated score are calculated from the mean of the two replicate differential scores, corrected untreated epsilon scores, and variance-stabilized corrected treated epsilon score, respectively. Throughout these calculations, the standard deviations of the original measurements are propagated to derive an estimate of error on the final differential score. The resulting standard deviation is used to derive a P value for each differential interaction score, which was calculated as the two-sided probability of observing a more extreme score than the one measured, given a background normal distribution centered on zero with a standard deviation equal to the one observed.

Differential interaction data corresponding to all tested gene pairs are provided in data file 3 (*20*). For analysis of individual interactions, we recommend that the data first be filtered before further analysis by applying one of our recommended thresholds (such as an intermediate threshold, $P < 0.05$ and $|\epsilon| > 0.08$).

Classifying differential interactions

Differential interactions were categorized as either novel, modified, or masked by comparing the genetic interaction scores for a given double mutant measured in a particular condition and the matched reference control (Fig. 2). Two additional categories of differential interactions were excluded from our analysis. First, gene pairs that did not show a significant genetic interaction ($P < 0.05$ and $|\epsilon| > 0.08$) in either a given test condition or the matched reference condition were excluded from further analysis because we could not be confident of a significant interaction in either the reference or the test conditions, individually, even though the differential score was significant. Second, double mutants that exhibited a significant and extreme negative genetic interaction ($P < 0.05$ and $|\epsilon| < -0.3$) in both a condition and the matched reference control were considered synthetic lethal in both the reference and test conditions and also excluded from further analyses of differential interactions.

Evaluating functional relations captured by differential interactions

We explored the functional information associated with different types of genetic interactions based on spatial analysis of functional enrichment (SAFE) neighborhood enrichment analysis (*11*, *54*). Query and mini-array genes

were assigned to one of the 17 bioprocesses or neighborhoods on the global genetic interaction profile similarity map on the basis of SAFE analysis (*11*, *54*). For each query in each condition, we counted the number of interacting array genes that occurred in each of the 17 SAFE neighborhoods. For query-centric analyses (Fig. 5), we summed up interactions associated with a specific query gene across all 14 conditions and calculated the fold enrichment for each neighborhood with a one-sided Fisher's exact test, by comparing the number of interacting array genes assigned to a given neighborhood versus the total number of interactions with that particular query, using the fraction of array genes assigned to the same neighborhood across the mini-array as background. For condition-centric analyses (Fig. 6), we performed the same test by combining all queries within each tested condition. This analysis was performed for negative and positive genetic interactions (Figs. 5 and 6 and fig. S9B), novel negative differential interactions (Figs. 5 and 6 and fig. S9A) using the mini-array screens, and novel negative differential interactions using the genome-wide screen on Benomyl condition (fig. S9A).

A more detailed description of the experimental and computational analyses are provided as supplementary material.

REFERENCES AND NOTES

1. M. Costanzo *et al*., Global genetic networks and the genotype-to-phenotype relationship. *Cell* **177**, 85–100 (2019). doi: 10.1016/j.cell.2019.01.033; pmid: 30901552
2. P. C. Phillips, S. P. Otto, M. C. Whitlock, in *Epistasis and the Evolutionary Process*, J. B. Wolf, E. D. I. Brodie, M. J. Wade, Eds. (Oxford Univ. Press, 2000).
3. O. Zuk, E. Hechter, S. R. Sunyaev, E. S. Lander, The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1193–1198 (2012). doi: 10.1073/pnas.1119675109; pmid: 22223662
4. J. S. Bloom, I. M. Ehrenreich, W. T. Loo, T. L. Lite, L. Kruglyak, Finding the sources of missing heritability in a yeast cross. *Nature* **494**, 234–237 (2013). doi: 10.1038/nature11867; pmid: 23376951
5. S. K. Forsberg, J. S. Bloom, M. J. Sadhu, L. Kruglyak, Ö. Carlborg, Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat. Genet.* **49**, 497–503 (2017). doi: 10.1038/ng.3800; pmid: 28250458
6. E. Kuzmin *et al*., Systematic analysis of complex genetic interactions. *Science* **360**, eaao1729 (2018). doi: 10.1126/science.aao1729; pmid: 29674565
7. A. H. Tong *et al*., Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001). doi: 10.1126/science.1065810; pmid: 11743205
8. A. H. Tong *et al*., Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004). doi: 10.1126/science.1091317; pmid: 14764870
9. R. Mani, R. P. St Onge, J. L. Hartman 4th, G. Giaever, F. P. Roth, Defining genetic interaction. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3461–3466 (2008). doi: 10.1073/pnas.0712255105; pmid: 18305163
10. J. van Leeuwen *et al*., Exploring genetic suppression interactions on a global scale. *Science* **354**, aag0839 (2016). doi: 10.1126/science.aag0839; pmid: 27811238
11. M. Costanzo *et al*., A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, aaf1420 (2016). doi: 10.1126/science.aaf1420; pmid: 27708008
12. M. Costanzo *et al*., The genetic landscape of a cell. *Science* **327**, 425–431 (2010). doi: 10.1126/science.1180823; pmid: 20093466
13. R. P. St Onge *et al*., Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat. Genet.* **39**, 199–206 (2007). doi: 10.1038/ng1948; pmid: 17206143
14. S. Bandyopadhyay *et al*., Rewiring of genetic networks in response to DNA damage. *Science* **330**, 1385–1389 (2010). doi: 10.1126/science.1195618; pmid: 21127252
15. A. Guénolé *et al*., Dissection of DNA damage responses using multiconditional genetic interaction maps. *Mol. Cell* **49**, 346–358 (2013). doi: 10.1016/j.molcel.2012.11.023; pmid: 23273983
16. M. H. Kramer *et al*., Active interaction mapping reveals the hierarchical organization of autophagy. *Mol. Cell* **65**, 761–774. e5 (2017). doi: 10.1016/j.molcel.2016.12.024; pmid: 28132844
17. H. Martin *et al*., Differential genetic interactions of yeast stress response MAPK pathways. *Mol. Syst. Biol.* **11**, 800 (2015). doi: 10.15252/msb.20145606; pmid: 25888283
18. M. Jaffe *et al*., Improved discovery of genetic interactions using CRISPRiSeq across multiple environments. *Genome Res.* **29**, 668–681 (2019). doi: 10.1101/gr.246603.118; pmid: 30782640
19. A. Baryshnikova *et al*., Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. *Nat. Methods* **7**, 1017–1024 (2010). doi: 10.1038/nmeth.1534; pmid: 21076421
20. All data files are available from the DRYAD Digital Respository (doi: 10.5061/dryad.3r2280gfd) as well as https://boonelab.ccbr.utoronto.ca/condition_sga.
21. E. A. Winzeler *et al*., Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999). doi: 10.1126/science.285.5429.901; pmid: 10436161
22. Z. Li *et al*., Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat. Biotechnol.* **29**, 361–367 (2011). doi: 10.1038/nbt.1832; pmid: 21441928
23. Materials and methods are available as supplementary materials.
24. A. P. Gasch *et al*., Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**, 4241–4257 (2000). doi: 10.1091/mbc.11.12.4241; pmid: 11102521
25. T. A. Clark, C. W. Sugnet, M. Ares Jr., Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* **296**, 907–910 (2002). doi: 10.1126/science.1069415; pmid: 11988574
26. A. M. Tartakoff, Perturbation of vesicular traffic with the carboxylic ionophore monensin. *Cell* **32**, 1026–1028 (1983). doi: 10.1016/0092-8674(83)90286-6; pmid: 6340834
27. J. F. Back, D. Oakenfull, M. B. Smith, Increased thermal stability of proteins in the presence of sugars and polyols. *Biochemistry* **18**, 5191–5196 (1979). doi: 10.1021/bi00590a025; pmid: 497177
28. L. E. Cowen, S. Lindquist, Hsp90 potentiates the rapid evolution of new traits: Drug resistance in diverse fungi. *Science* **309**, 2185–2189 (2005). doi: 10.1126/science.1118370; pmid: 16195452
29. K. Rizzolo *et al*., Features of the chaperone cellular network revealed through systematic interaction mapping. *Cell Rep.* **20**, 2735–2748 (2017). doi: 10.1016/j.celrep.2017.08.074; pmid: 28903051
30. C. Queitsch, T. A. Sangster, S. Lindquist, Hsp90 as a capacitor of phenotypic variation. *Nature* **417**, 618–624 (2002). doi: 10.1038/nature749; pmid: 12050657
31. A. Y. Lee *et al*., Mapping the cellular response to small molecules using chemogenomic fitness signatures. *Science* **344**, 208–211 (2014). doi: 10.1126/science.1250217; pmid: 24723613
32. G. J. Bean, T. Ideker, Differential analysis of high-throughput quantitative genetic interaction data. *Genome Biol.* **13**, R123 (2012). doi: 10.1186/gb-2012-13-12-r123; pmid: 23268787
33. M. Dilcher, B. Köhler, G. F. von Mollard, Genetic interactions with the yeast Q-SNARE VTI1 reveal novel functions for the R-SNARE YKT6. *J. Biol. Chem.* **276**, 34537–34544 (2001). doi: 10.1074/jbc.M101551200; pmid: 11445562
34. J. J. Kim, Z. Lipatova, N. Segev, TRAPP Complexes in secretion and autophagy. *Front. Cell Dev. Biol.* **4**, 20 (2016). doi: 10.3389/fcell.2016.00020; pmid: 27066478
35. D. Pruyne, A. Legesse-Miller, L. Gao, Y. Dong, A. Bretscher, Mechanisms of polarized growth and organelle segregation in yeast. *Annu. Rev. Cell Dev. Biol.* **20**, 559–591 (2004). doi: 10.1146/annurev.cellbio.20.010403.103108; pmid: 15473852
36. D. Rotin, S. Kumar, Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **10**, 398–409 (2009). doi: 10.1038/nrm2690; pmid: 19436320
37. P. Kaliszewski, T. Zoładek, The role of Rsp5 ubiquitin ligase in regulation of diverse processes in yeast cells. *Acta Biochim. Pol.* **55**, 649–662 (2008). doi: 10.18388/abp.2008_3024; pmid: 19039336
38. J. D. Beggs, Lsm proteins and RNA processing. *Biochem. Soc. Trans.* **33**, 433–438 (2005). doi: 10.1042/BST0330433; pmid: 15916535
39. J. D. Aitchison, M. P. Rout, The yeast nuclear pore complex and transport through it. *Genetics* **190**, 855–883 (2012). doi: 10.1534/genetics.111.127803; pmid: 22419078
40. P. Orlean, Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall. *Genetics* **192**, 775–818 (2012). doi: 10.1534/genetics.112.144485; pmid: 23135325
41. C. K. Barlowe, E. A. Miller, Secretory protein biogenesis and traffic in the early secretory pathway. *Genetics* **193**, 383–410 (2013). doi: 10.1534/genetics.112.142810; pmid: 23396477
42. J. Hou, G. Tan, G. R. Fink, B. J. Andrews, C. Boone, Complex modifier landscape underlying genetic background effects. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5045–5054 (2019). doi: 10.1073/pnas.1820915116; pmid: 30804202
43. M. N. Mullis, T. Matsui, R. Schell, R. Foree, I. M. Ehrenreich, The complex underpinnings of genetic background effects. *Nat. Commun.* **9**, 3548 (2018). doi: 10.1038/s41467-018-06023-5; pmid: 30224702
44. R. Schell, M. N. Mullis, T. Matsui, R. Foree, I. M. Ehrenreich, Genetic architecture of a mutation's expressivity and penetrance. bioRxiv 024547 [Preprint] 5 April 2020. doi: 10.1101/2020.04.03.024547
45. J. S. Bloom *et al*., Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nat. Commun.* **6**, 8712 (2015). doi: 10.1038/ncomms9712; pmid: 26537231
46. W. Huang, T. F. Mackay, The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis. *PLOS Genet.* **12**, e1006421 (2016). doi: 10.1371/journal.pgen.1006421; pmid: 27812106
47. G. Fang *et al*., Discovering genetic interactions bridging pathways in genome-wide association studies. *Nat. Commun.* **10**, 4274 (2019). doi: 10.1038/s41467-019-12131-7; pmid: 31537791
48. W. Wang *et al*., Pathway-based discovery of genetic interactions in breast cancer. *PLOS Genet.* **13**, e1006973 (2017). doi: 10.1371/journal.pgen.1006973; pmid: 28957314
49. M. Ghandi *et al*., Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019). doi: 10.1038/s41586-019-1186-3; pmid: 31068700
50. R. M. Meyers *et al*., Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017). doi: 10.1038/ng.3984; pmid: 29083409
51. A. Tsherniak *et al*., Defining a cancer dependency map. *Cell* **170**, 564–576.e16 (2017). doi: 10.1016/j.cell.2017.06.010; pmid: 28753430
52. R. Deshpande *et al*., Efficient strategies for screening large-scale genetic interaction networks. bioRxiv 159632 [Preprint] 5 July 2017. doi: 10.1101/159632
53. E. Kuzmin *et al*., Synthetic genetic array analysis for global mapping of genetic networks in yeast. *Methods Mol. Biol.* **1205**, 143–168 (2014). doi: 10.1007/978-1-4939-1363-3_10; pmid: 25213244
54. A. Baryshnikova, Systematic Functional Annotation and Visualization of Biological Networks. *Cell Syst.* **2**, 412–421 (2016). doi: 10.1016/j.cels.2016.04.014; pmid: 27237738
55. C. L. Myers, D. R. Barrett, M. A. Hibbs, C. Huttenhower, O. G. Troyanskaya, Finding function: Evaluation methods for functional genomic data. *BMC Genomics* **7**, 187 (2006). doi: 10.1186/1471-2164-7-187; pmid: 16869964
56. M. Kofoed *et al*., An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. *G3 (Bethesda)* **5**, 1879–1887 (2015). doi: 10.1534/g3.115.019174; pmid: 26175450
57. A. Subramanian *et al*., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
58. G. Korotkevich, V. Sukhov, A. Sergushichev, Fast gene set enrichment analysis. *bioRxiv* 060012 [Preprint] 1 February 2019. doi: 10.1101/060012
59. M. Usaj *et al*., TheCellMap.org: A web-accessible database for visualizing and mining the global yeast genetic interaction network. *G3 (Bethesda)* **7**, 1539–1549 (2017). doi: 10.1534/g3.117.040220; pmid: 28325812
60. S. A. Hoose *et al*., A systematic analysis of cell cycle regulators in yeast reveals that most factors act independently of cell size

to control initiation of division. *PLOS Genet.* **8**, e1002590 (2012). doi: 10.1371/journal.pgen.1002590; pmid: 22438835
61. G. Giaever *et al*., Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002). doi: 10.1038/nature00935; pmid: 12140549
62. B. H. M. Meldal *et al*., Complex Portal 2018: Extended content and enhanced visualization tools for macromolecular complexes. *Nucleic Acids Res.* **47**, D550–D558 (2019). doi: 10.1093/nar/gky1001; pmid: 30357405
63. J. M. Cherry *et al*., Saccharomyces Genome Database: The genomics resource of budding yeast. *Nucleic Acids Res.* **40**, D700–D705 (2012). doi: 10.1093/nar/gkr1029; pmid: 22110037
64. N. J. Krogan *et al*., Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637–643 (2006). doi: 10.1038/nature04670; pmid: 16554755
65. A. C. Gavin *et al*., Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006). doi: 10.1038/nature04532; pmid: 16429126
66. K. Tarassov *et al*., An in vivo map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008). doi: 10.1126/science.1153878; pmid: 18467557
67. M. Babu *et al*., Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*. *Nature* **489**, 585–589 (2012). doi: 10.1038/nature11354; pmid: 22940862
68. E. N. Koch *et al*., Conserved rules govern genetic interaction degree across species. *Genome Biol.* **13**, R57 (2012). doi: 10.1186/gb-2012-13-7-r57; pmid: 22747640
69. C. Huttenhower, M. Hibbs, C. Myers, O. G. Troyanskaya, A scalable method for integration and functional analysis of multiple microarray datasets. *Bioinformatics* **22**, 2890–2897 (2006). doi: 10.1093/bioinformatics/btl492; pmid: 17005538
70. C. Stark *et al*., The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.* **39** (Database), D698–D704 (2011). doi: 10.1093/nar/gkq1116; pmid: 21071413
71. H. Yu *et al*., High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110 (2008). doi: 10.1126/science.1158684; pmid: 18719252
72. W. K. Huh *et al*., Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003). doi: 10.1038/nature02026; pmid: 14562095

ACKNOWLEDGMENTS

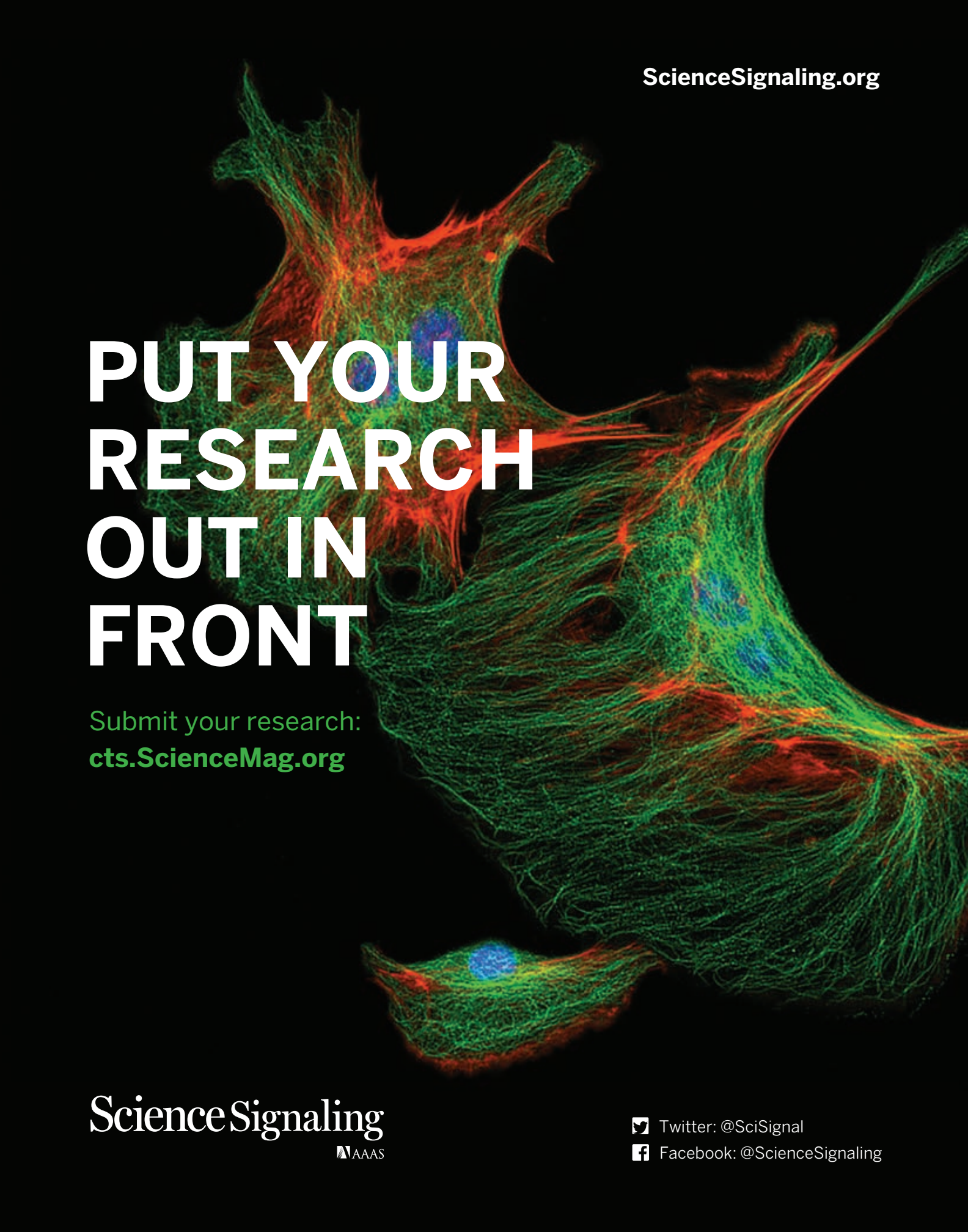
Funding: This work was primarily supported by the National Institutes of Health (R01HG005853) (to B.A., C.B., and C.L.M.), Canadian Institutes of Health Research (FDN-143264 and FDN-143265) (to C.B. and B.A.), National Institutes of Health (R01HG005084) (to C.L.M.), the National Science Foundation (DBI\0953881) (to C.L.M.), and a Ramon y Cajal fellowship (RYC-2017-22959) (to C.P.). Computing resources and data storage services were partially provided by the Minnesota Supercomputing Institute and the UMN Office of Information Technology, respectively. C.B. is a fellow of the Canadian Institute for Advanced Research (CIFAR). **Author contributions:** M.C., C.L.M., C.B., and B.A. conceived and coordinated the project. V.M., B.-J.S.L., and E.S. performed genetic interaction experiments. J.H., V.M., J.N., M.R., B.V.S., W.W., C.P., C.R., M.U., E.N.K., and P.A. analyzed the data. M.C., J.H., C.L.M., C.B., and B.A. prepared the manuscript. **Competing interests:** The authors declare no competing interests. **Data and material availability:** All data files associated with this study are available for download from the DRYAD Digital Repository (doi:10.5061/dryad.3r2280gfd) or from https://boonelab.ccbr.utoronto.ca/condition_sga.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/eabf8424/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*56*–*72*)

23 November 2020; accepted 30 March 2021
10.1126/science.abf8424

ScienceSignaling.org



PUT YOUR RESEARCH OUT IN FRONT

Submit your research:
cts.ScienceMag.org

Science Signaling
AAAS

Twitter: @SciSignal
Facebook: @ScienceSignaling

RESEARCH ARTICLE SUMMARY

PALEOGENOMICS

Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments

Benjamin Vernot*, Elena I. Zavala, Asier Gómez-Olivencia, Zenobia Jacobs, Viviane Slon, Fabrizio Mafessoni, Frédéric Romagné, Alice Pearson, Martin Petr, Nohemi Sala, Adrián Pablos, Arantza Aranburu, José María Bermúdez de Castro, Eudald Carbonell, Bo Li, Maciej T. Krajcarz, Andrey I. Krivoshapkin, Kseniya A. Kolobova, Maxim B. Kozlikin, Michael V. Shunkov, Anatoly P. Derevianko, Bence Viola, Steffi Grote, Elena Essel, David López Herráez, Sarah Nagel, Birgit Nickel, Julia Richter, Anna Schmidt, Benjamin Peter, Janet Kelso, Richard G. Roberts, Juan-Luis Arsuaga, Matthias Meyer*

INTRODUCTION: The study of hominin history has progressed through both archaeological and genetic insights. Although DNA sequencing from hominin skeletal remains allows the association of ancient populations with specific places in time and space, many archaeological sites lack associated hominin remains, limiting the scope of genetic analyses. Even when ancient hominin remains are found, they often do not cover the full time span of a site or sampling them for DNA may not be possible. The fossil record is particularly sparse for Pleistocene hominins, leaving large gaps in our understanding of the genetic histories of archaic and early modern humans.

RATIONALE: Recent work has demonstrated the feasibility of sequencing ancient mammalian mitochondrial DNA (mtDNA), including that of hominins, from Pleistocene cave sediments. However, mtDNA represents only the maternal lineage and thus provides limited data for the resolution of population relationships. It is therefore desirable to complement mtDNA analysis with the retrieval of nuclear DNA, but no strategies are in place to enrich hominin nuclear DNA from a background of related sequences from other mammals present in most sedimentary deposits. To close this gap, we developed a set of probes for hybridization capture that targets 1.6 million ancestry-informative positions in the hominin nuclear genome, specifically at loci with high mammalian sequence divergence. We then developed computational methods to deplete residual microbial and faunal DNA sequences, along with methods to account for such non-hominin DNA in population genetic analyses.

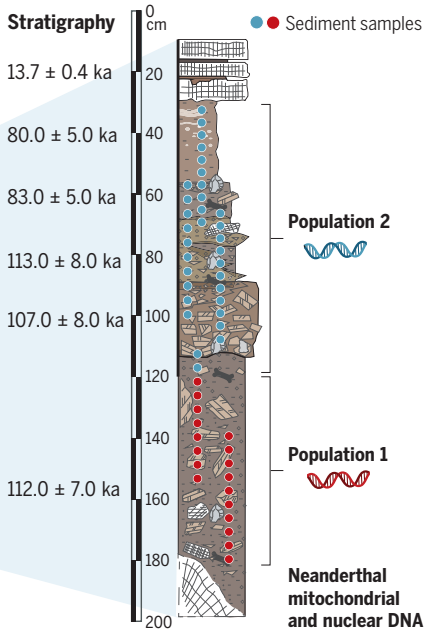
RESULTS: We applied these methods to explore the history of Neanderthal populations in western Europe and southern Siberia using sediment samples from three Pleistocene caves: Galería de las Estatuas, a site in northern Spain with 40 thousand years of Neanderthal occupation but that is genetically unexplored, and Chagyrskaya and Denisova Caves, which have previously yielded high-coverage genomes of two Neanderthals and one Denisovan hominin. In total, we recovered Neanderthal or Denisovan mtDNA from >60 sediment samples and nuclear DNA from 30 of these. For Chagyrskaya and Denisova Caves, our phylogenetic results from sediment DNA were consistent with previously published results from skeletal remains, confirming the accuracy of our approach. At Galería de las Estatuas, we recovered Neanderthal DNA from layers spanning nearly the entire stratigraphy, and identified a population turnover ~100,000 years ago accompanied by a loss of mtDNA diversity. By incorporating genetic data from previously published skeletal samples, we associated this turnover with two putative radiations in Neanderthal history.

CONCLUSION: We developed methods for the effective retrieval and analysis of ancient hominin nuclear DNA from sediments and used them to uncover previously unknown events in Neanderthal history. This work demonstrates that detailed genetic analyses are now possible for many more archaeological sites than previously thought, with DNA from abundant sediments allowing dense time-series studies that are independent of the fossil record. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: benjamin_vernot@eva.mpg.de (B.V.); mmeyer@eva.mpg.de (M.M.)
Cite this article as B. Vernot *et al.*, *Science* **372**, eabf1667 (2021). DOI: 10.1126/science.abf1667

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf1667





Sediments from Pleistocene caves contain hominin mitochondrial and nuclear DNA that can be enriched, sequenced, and analyzed to reveal the genetic histories of past occupants even in the absence of their skeletal remains. Shown is a view of pit I at the Galería de las Estatuas, Spain, and stratigraphic column with ages in thousands of years (ka).

PHOTO CREDIT : PANTOJA-PÉREZ.NTTF

RESEARCH ARTICLE

PALEOGENOMICS

Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments

Benjamin Vernot[1]*, Elena I. Zavala[1], Asier Gómez-Olivencia[2,3,4], Zenobia Jacobs[5,6], Viviane Slon[1,7,8], Fabrizio Mafessoni[1], Frédéric Romagné[1], Alice Pearson[1], Martin Petr[1], Nohemi Sala[4,9], Adrián Pablos[4,9], Arantza Aranburu[2,3], José María Bermúdez de Castro[9], Eudald Carbonell[10,11], Bo Li[5,6], Maciej T. Krajcarz[12], Andrey I. Krivoshapkin[13,14], Kseniya A. Kolobova[13], Maxim B. Kozlikin[13], Michael V. Shunkov[13], Anatoly P. Derevianko[13], Bence Viola[15], Steffi Grote[1], Elena Essel[1], David López Herráez[1], Sarah Nagel[1], Birgit Nickel[1], Julia Richter[1], Anna Schmidt[1], Benjamin Peter[1], Janet Kelso[1], Richard G. Roberts[5,6], Juan-Luis Arsuaga[4,16], Matthias Meyer[1]*

Bones and teeth are important sources of Pleistocene hominin DNA, but are rarely recovered at archaeological sites. Mitochondrial DNA (mtDNA) has been retrieved from cave sediments but provides limited value for studying population relationships. We therefore developed methods for the enrichment and analysis of nuclear DNA from sediments and applied them to cave deposits in western Europe and southern Siberia dated to between 200,000 and 50,000 years ago. We detected a population replacement in northern Spain about 100,000 years ago, which was accompanied by a turnover of mtDNA. We also identified two radiation events in Neanderthal history during the early part of the Late Pleistocene. Our work lays the ground for studying the population history of ancient hominins from trace amounts of nuclear DNA in sediments.

The analysis of ancient DNA from Pleistocene hominins has greatly enhanced our understanding of the evolutionary history of archaic humans and their interactions with early modern humans. To date, complete or partial nuclear genome sequences have been recovered from the skeletal remains of 23 archaic hominin individuals: 18 Neanderthals from 14 sites across Eurasia (mostly in Europe), four Denisovans, and the offspring of a Neanderthal mother and a Denisovan father (*Denisova 11*) (*1*) recovered from Denisova Cave in the Altai Mountains of southern Siberia. Although numerous Paleolithic sites have been excavated, relatively few have yielded skeletal remains of hominins, which are often concentrated in one or a few strata. Attempts to reconstruct the genetic history of archaic hominins are therefore constrained by an uneven temporal and spatial sampling, limited largely by the availability of specimens.

In 2017, it was found that hominin mitochondrial DNA (mtDNA) can be recovered from Pleistocene sediments (*2*), indicating that it may be possible to overcome the dependency on the scarce fossil record in the quest for hominin DNA. However, mtDNA only carries information about the maternal lineage and does not always reflect the complete population history [e.g., (*3*)]. Nuclear DNA contains far more information, but its retrieval from sediments presents substantial challenges: It is present in fewer copies than mtDNA, and many loci are not informative for population genetic analyses. Additionally, by far most mammalian DNA in sediments is non-hominin, which may be difficult to distinguish from hominin DNA because of sequence homology at many loci. These characteristics, as well as the dominance of microbial DNA (*2*), hamper attempts to retrieve nuclear sequences in sufficient number and quality for population genetic analyses by simple shotgun sequencing. To overcome these challenges, we set out to retrieve hominin nuclear genomic sequences from sediments using hybridization capture to target regions in the nuclear genome with high mammalian sequence diversity, and used these sequences to explore the history of Neanderthal populations in western Europe and southern Siberia.

Archaeological sites

We focused our analyses on sediments from three Paleolithic sites. Denisova Cave (*4*) and Chagyrskaya Cave (*5*), both located in the Altai Mountains (Fig. 1A), were included for their known mtDNA preservation in sediments (*2*) and to enable comparisons with three high-coverage nuclear genomes generated previously from bones from these sites: *Denisova 5* [the *Altai Neanderthal* toe bone, dated to 90.9 to 130.0 thousand years (ka) ago (*6*, *7*)], *Denisova 3* [a Denisovan finger bone, 51.6 to 76.2 ka ago (*6*, *8*)], and *Chagyrskaya 8* [a Neanderthal finger bone, 49.0 to 92 ka ago (*5*, *9*)]. All age ranges include the 95% confidence interval (CI) of the dating method(s). Whereas Denisova Cave has evidence for at least 250 millennia of archaic human occupation (*4*), the Neanderthal-bearing deposits at Chagyrskaya Cave (layers 5 and 6; Fig. 1B and fig. S1) accumulated in <10 millennia (*5*).

The third site, Galería de las Estatuas (hereafter "Estatuas"), is part of the Atapuerca archaeo-paleontological complex in northern Spain (Fig. 1C and fig. S2). Almost 500 stone artifacts with clear Mousterian affinities, combined with single-grain optical dating of the associated sediments, indicate Neanderthal occupation from at least 113 ± 8 to 70 ± 5 ka ago [total uncertainty at 1σ; (*10*, *11*)], yet only a single Neanderthal foot phalanx has been recovered (*12*) (fig. S3). Initial screening of the Estatuas sediments indicated the presence of ancient mammalian mtDNA, including that of hominins [figs. S4 to S6 and tables S10 and S11 (*13*)]. Analysis of sediment DNA may therefore be the only viable approach for reconstructing the population genetics of the occupants of this site during a time period not currently well represented in the genetic record of European Neanderthals.

For Denisova Cave, we retrieved nuclear DNA from three existing sediment samples with hominin mtDNA preservation, from: layers 11.4 and 15 in the East Chamber and layer 14.3 in the Main Chamber (*2*, *4*). At Chagyrskaya Cave and Estatuas, we extensively sampled across the Paleolithic layers, collecting 76 samples from two pits at Estatuas and 73 samples from Chagyrskaya Cave [Fig. 1, B and C, and figs. S4 and S5 (*13*)], and targeted both hominin mtDNA and nuclear DNA.

Retrieval of hominin mtDNA

We enriched hominin mtDNA from Chagyrskaya Cave and Estatuas samples using protocols for automated DNA extraction (*14*), library preparation (*15*), and hybridization capture

[1]Department of Evolutionary Genetics, Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany. [2]Departamento de Geología, Facultad de Ciencia y Tecnología, Universidad del País Vasco-Euskal Herriko Unibertsitatea (UPV/EHU), Leioa, Spain. [3]Sociedad de Ciencias Aranzadi, Donostia-San Sebastián, Spain. [4]Centro Mixto UCM-ISCIII de Evolución y Comportamiento Humanos, Madrid, Spain. [5]Centre for Archaeological Science, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, New South Wales, Australia. [6]Australian Research Council (ARC) Centre of Excellence for Australian Biodiversity and Heritage, University of Wollongong, Wollongong, New South Wales, Australia. [7]Department of Anatomy and Anthropology and Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [8]The Dan David Center for Human Evolution and Biohistory Research, Tel Aviv University, 6997801 Tel Aviv, Israel. [9]Centro Nacional de Investigación Sobre la Evolución Humana (CENIEH), Burgos, Spain. [10]Institut Català de Paleoecologia Humana i Evolució Social (IPHES), Universitat Rovira i Virgili, Tarragona, Spain. [11]Àrea de Prehistòria, Universitat Rovira i Virgili, Tarragona, Spain. [12]Institute of Geological Sciences, Polish Academy of Sciences, Warszawa, Poland. [13]Institute of Archaeology and Ethnography, Russian Academy of Sciences, Novosibirsk, Russia. [14]Novosibirsk State University, Novosibirsk, Russia. [15]Department of Anthropology, University of Toronto, Toronto, Ontario, Canada. [16]Departamento de Paleontología, Facultad Ciencias Geológicas, Universidad Complutense de Madrid, Madrid, Spain.
*Corresponding author. Email: benjamin_vernot@eva.mpg.de (B.V.); mmeyer@eva.mpg.de (M.M.)

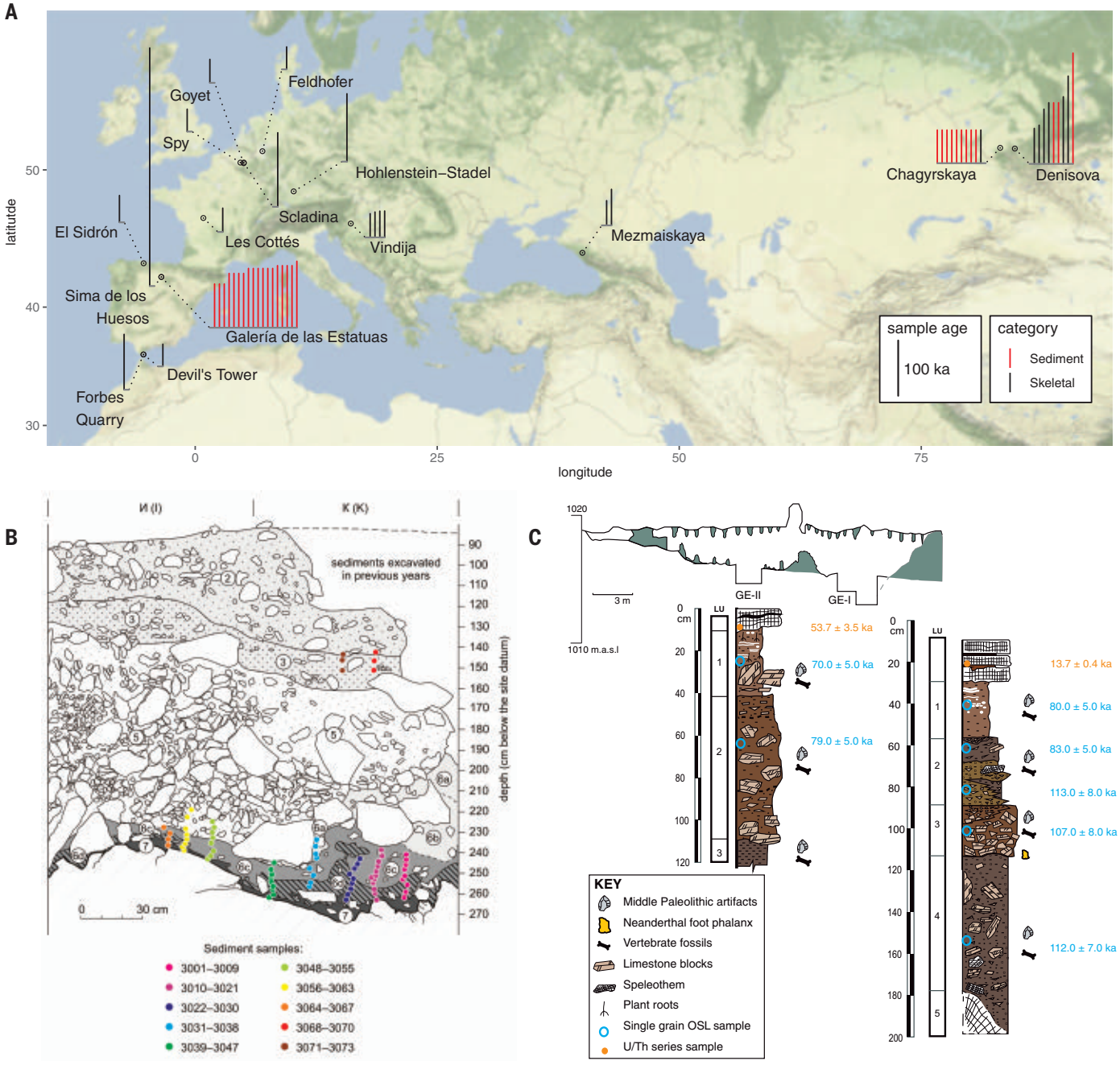


Fig. 1. Samples and stratigraphies. (**A**) Geographic locations of all skeletal (black) and sediment (red) samples with >0.0001× coverage. Each bar represents one sample; height is the age of the sample (ka, thousands of years). (**B**) Stratigraphic profile of the sampled section at Chagyrskaya Cave. Sediment samples were collected in 10 vertical columns and numbered from bottom to top. Colored circles denote individual sample locations. (**C**) Cross section of Estatuas showing the locations of the two test pits, GE-I and GE-II, and detailed stratigraphic columns showing the lithostratigraphic units (LUs) and ages obtained.

(*2*). To maximize the number of libraries containing sufficient amounts of hominin DNA for analysis, multiple subsamples were taken from some of the Estatuas samples and several libraries produced from some subsamples, for a total of 369 libraries. After assigning sequences to mammalian families using MEGAN (*16*), we found that 74% of samples (n = 54) from Chagyrskaya and 56% (n = 43) from Estatuas yielded at least one library containing hominin mtDNA fragments with significantly elevated frequencies of cytosine (C) to thymine (T) substitutions at their 5′ and 3′ ends, compatible with the presence of deaminated ancient DNA [binomial CI ≥ 10% at both ends; table S10 (*13*)]. Of the 223 libraries containing ancient hominin mtDNA, 182 (82%) yielded sufficient fragments to allow their assignment to a hominin group based on "diagnostic" positions in the mtDNA genome that are derived in one hominin group (modern humans, Neanderthals, Denisovans, or the Sima de los Huesos hominins) and ancestral in the others (*3*, *13*). All such assignments were to Neanderthal mtDNA, consistent with archaeological evidence for the presence of Neanderthals in all layers, with exception of the upper portions of layer 7 in Chagyrskaya Cave, which is archaeologically sterile and dated to >300 ka ago. The detection of Neanderthal

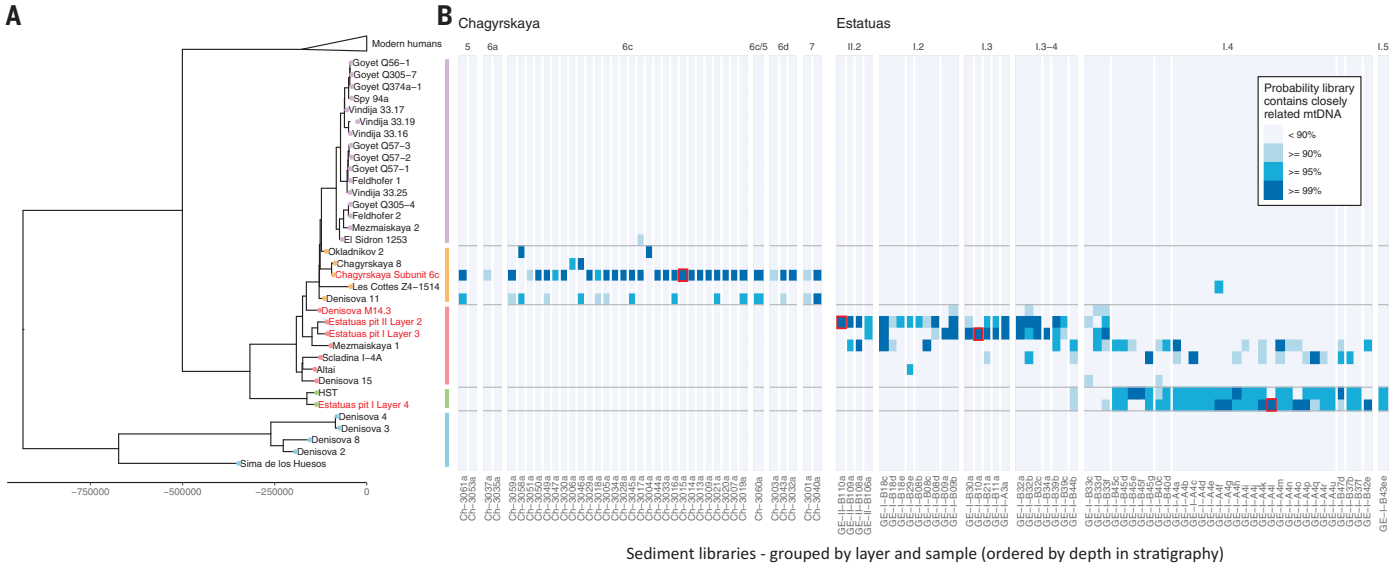


Fig. 2. mtDNA from sediments. (**A**) Mitochondrial phylogeny, including five haplotypes inferred from sediment samples (red labels). Five mtDNA groupings are labeled (colored dots and vertical bars). (**B**) Probabilistic phylogenetic placement of mtDNA from 97 sediment subsamples from 63 sediment samples from Chagyrskaya Cave (left) and Estatuas (right). Samples are divided by white lines. Subsamples from which mtDNA haplotypes were inferred are denoted with red boxes.

mtDNA near the top of layer 7 may be due to its being a former living floor and/or a consequence of post-depositional mixing with sediments from the overlying subunit 6c, resulting in subunit 6d (*5*, *13*).

Fourteen samples produced libraries with high coverage of the Neanderthal mtDNA genome (>17-fold) and point estimates of present-day human contamination <10% (*13*). Four of these (from Chagyrskaya subunit 6c, Estatuas pit II/layer 2, and Estatuas pit I/layers 3 and 4) appeared to contain a single mitochondrial sequence on the basis of the consistency of nucleotides observed at each position [fig. S7 and table S3 (*13*)]. These were used for generating near-complete consensus sequences and building a phylogenetic tree with BEAST2 (*17*), along with previously published hominin mtDNA sequences derived from skeletal remains or from present-day humans and a sequence reconstructed from the layer 14.3 sediment sample from the Main Chamber in Denisova Cave (*2*) [Fig. 2A, fig. S8, and table S4 (*13*)].

Most notably, we found that the consensus mtDNA genome from Estatuas pit I/layer 4 was most similar to the mtDNA of the ~120 ka old Neanderthal from Hohlenstein-Stadel (*HST*), Germany, which falls basal to all other known Neanderthal mitochondrial genomes (*18*, *19*). The Estatuas pit I/layer 4 sediments (dated to 112 ± 7 ka ago) are broadly contemporaneous with *HST*. The mtDNA genomes from Estatuas pit II/layer 2 (79 ± 5 ka ago) and pit I/layer 3 (107 ± 8 ka ago) group with mtDNA from the 60 to 70 ka old *Mezmaiskaya 1* individual from the northern Caucasus (*7*, *20*), whereas the sequence for Chagyrskaya subunit 6c groups with the mtDNA from *Chagyrskaya 8* (*9*). We note that ages estimated from the branch lengths of the sediment mtDNA sequences in the tree, such as 136 ka ago (95% CI: 75 to 200 ka ago) for the Estatuas pit I/layer 4 sample, match previously published ages for the respective sites and layers [table S5 (*13*)].

To further investigate mitochondrial diversity in the sediments from Chagyrskaya Cave and Estatuas, we developed a method to probabilistically place libraries with as few as 250 ancient fragments in the known Neanderthal mtDNA diversity [figs S9 to S12 (*13*)]. This method allowed phylogenetic assignments of mtDNA from 38 libraries from Chagyrskaya Cave and 59 from Estatuas, spanning most layers sampled in each cave (Fig. 2B). We found that Estatuas pit I/layers 4 and 5 contain both *HST*-like and non–*HST*-like Neanderthal mtDNA, often in the same subsample, with the latter largely grouping with *Mezmaiskaya 1* and *Scladina I-4a*, an ~130 ka old Neanderthal from western Europe (*19*, *21*). *HST*-like mtDNA then disappears from the upper layers of Estatuas, leaving mtDNA predominantly related to the *Mezmaiskaya 1*–like consensus sequences from those layers (Fig. 2B). Simulated mixing of DNA from the upper layers and pit I/layer 4 does not generate the observation of *Mezmaiskaya 1*– and *Scladina I-4a*–like DNA in pit I/layer 4, indicating true heterogeneity of mtDNA in the lower layers, consistent with the previously observed integrity of the Estatuas stratigraphy [fig. S13 (*11*, *13*)]. In Chagyrskaya Cave, we found remarkable homogeneity: All samples from layers 5 to 7 grouped with the subunit 6c consensus sequence, *Chagyrskaya 8* or *Okladnikov 2* (also from the Altai region) (*22*, *23*), with occasional support for *Denisova 11*–like sequences.

Nuclear DNA enrichment method

To extend the study of hominin DNA from sediments to the nuclear genome, we designed a probe set targeting 1.6 million informative single nucleotide polymorphisms (SNPs) in the nuclear genome and enriched for DNA fragments overlapping these sites through hybridization capture (*2*, *13*). We used several techniques to both reduce and measure the extent of mismapping of non-hominin faunal DNA, and evaluated these measures using simulated ancient brown bear (*Ursus arctos*) DNA, with fragment sizes and deamination profiles taken from an ancient DNA library (*13*). We found that faunal misalignment was substantially reduced in regions of high mammalian diversity (e.g., Fig. 3, A versus B), by at least 48-fold where the human genome differed by eight or more base pairs from at least one non-primate genome (in a 52–base pair region centered on a target SNP; Fig. 3C). We therefore restricted our design to SNPs in these regions. We additionally assigned each DNA fragment to National Center for Biotechnology Information (NCBI) taxonomy using the metagenomics software Kraken (*24*) and restricted our analyses to fragments classified as primate (*13*). This metagenomic filtering step substantially enriches for hominin DNA, reducing the alignment of simulated ancient bear DNA to the human reference genome by a factor of 140, versus 1.8-fold for Neanderthal DNA (*13*). However, these approaches may not eliminate all faunal misalignment, particularly if the DNA originates from a species that is

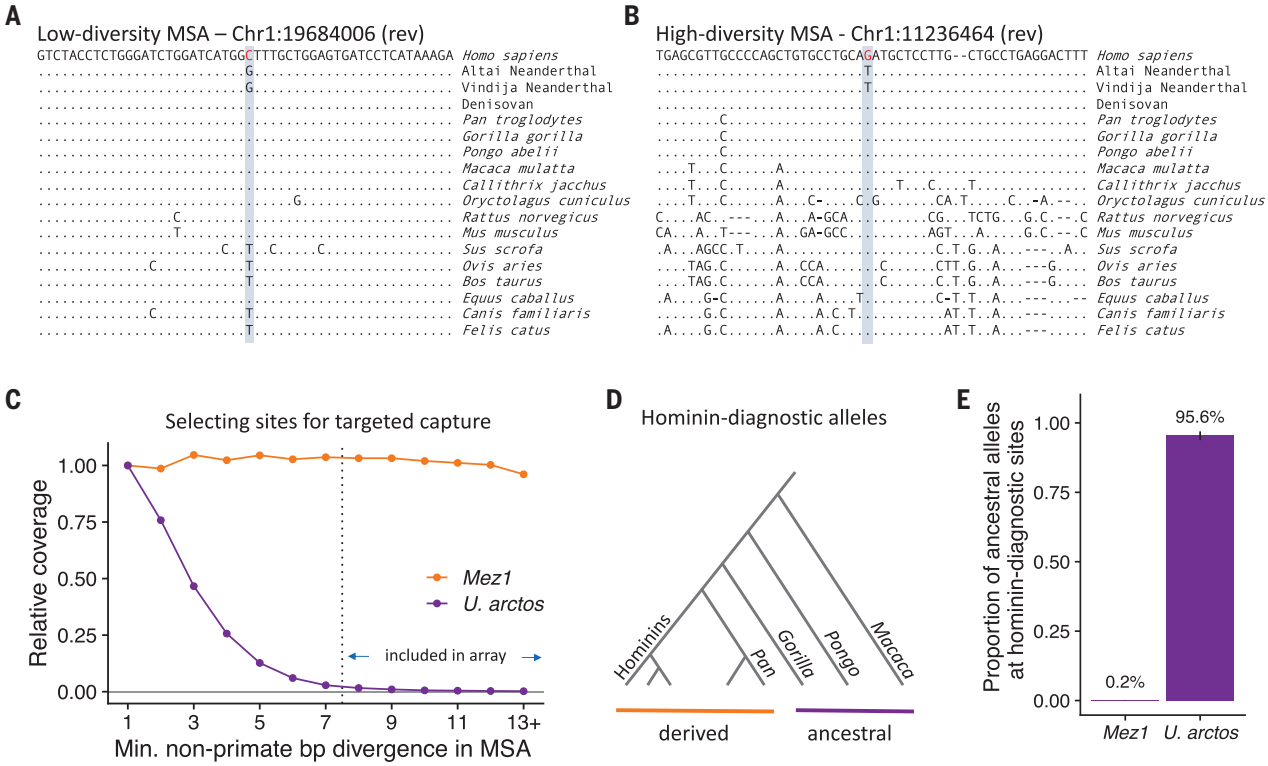


Fig. 3. Nuclear DNA capture design. (**A**) Multiple sequence alignment (MSA) of 15 mammalian species, plus the sequences of two Neanderthals, and one Denisovan. This region has low sequence diversity between hominins and non-primate mammals. (**B**) High-diversity MSA. (**C**) Faunal misalignment (purple) in the human genome as a function of mammalian diversity (the minimum base pair divergence between *Homo sapiens* and nine non-primate mammalian sequences) compared with the alignment of DNA fragments from the *Mezmaiskaya 1* Neanderthal (orange; *Mez1*). The *y*-axis is relative coverage compared with the first bin. (**D**) Ascertainment of hominin-diagnostic alleles; ancestral alleles are indicative of faunal misalignment. (**E**) Proportion of ancestral alleles at hominin-diagnostic sites in *Mezmaiskaya 1* and simulated *U. arctos* ancient DNA.

not well represented in the NCBI taxonomy. Therefore, we included in our probe design 98,887 "hominin diagnostic" sites that are fixed derived in hominins, chimpanzees, and bonobos (Fig. 3D). At these sites, ancestral alleles in the captured fragments are strongly indicative of the presence of non-hominin mammalian DNA [0.2% ancestral in Neanderthal compared with 95.6% in *U. arctos* DNA; Fig. 3E (*13*)] and can be used to estimate faunal misalignment proportions.

Nuclear DNA recovery and sexing

We first applied these methods to previously prepared libraries from three Denisova Cave sediments, two with Neanderthal mtDNA and one with Denisovan mtDNA (*2*). All showed significant levels of C-to-T substitutions in DNA fragments overlapping our targeted SNPs, consistent with the presence of ancient nuclear DNA [fig. S14 and table S13 (*13*)]. One sample showed evidence of moderate (~15%) non-hominin faunal misalignment before metagenomic filtering, but filtering reduced this to <1% (table S13). After filtering, 1764, 27,923, and 162,508 DNA fragments (459, 8698, and 42,103 with evidence of deamination, respectively) were retained from the three samples, representing up to ~0.1× coverage of our targeted sites.

Having confirmed the presence of ancient hominin nuclear DNA, we next investigated whether we could assign each sample to a hominin group. By examining deaminated DNA fragments at sites where the high-coverage Denisovan and *Altai Neanderthal* genomes are homozygous and differ from each other, we found that in the two samples containing Neanderthal mtDNA, ~90% of DNA fragments carried the Neanderthal derived state, versus 2% carrying the Denisovan derived state (Fig. 4A, top left, red points). By contrast, the nuclear DNA in the sample containing Denisovan mtDNA carried the Denisovan derived allele in 65% of cases but no Neanderthal derived alleles (Fig. 4A, top left, blue point). These results are consistent with those obtained from low-coverage Neanderthal and Denisovan genomes from skeletal remains [Fig. 4A, bottom right, red and blue points (*1*, *13*, *19*, *25*–*27*)], suggesting that the nuclear DNA in the three sediment samples is either of Neanderthal or Denisovan origin, but not both.

We next captured nuclear DNA from Chagyrskaya Cave and Estatuas sediment samples. Twenty-nine samples yielded at least one library with significant evidence for deamination and <5% faunal misalignment [tables S12 and S13 (*13*)]. Four libraries showed evidence for >5% faunal misalignment and were excluded from further analyses, highlighting the importance of per-library misalignment estimates [table S14 (*13*)]. In total, we retrieved nuclear DNA from Chagyrskaya Cave subunits 6a to 6d and 7 [the latter likely intrusive from subunits 6c and 6d (*13*)], and Estatuas pit II/layer 2 and pit I/layers 2 to 5. Recovery of hominin DNA was lower than for the Denisova Cave samples, with our best libraries yielding 134,497 fragments (33,594 with evidence of deamination) at target sites for Chagyrskaya Cave and 47,667 for Estatuas (16,678 deaminated). In a plot of Neanderthal versus Denisovan alleles, these samples clearly contained Neanderthal nuclear DNA and clustered together closely with Neanderthal skeletal samples (Fig. 4A, top right and bottom left). Equivalent plots considering sites that differ among the *Altai Neanderthal*, *Vindija 33.19*, or *Chagyrskaya 8* genomes lack the resolution to resolve their relationship to these Neanderthal genomes, given the small amounts of data per sample at these SNPs [figs. S15 and S16 (*13*)].

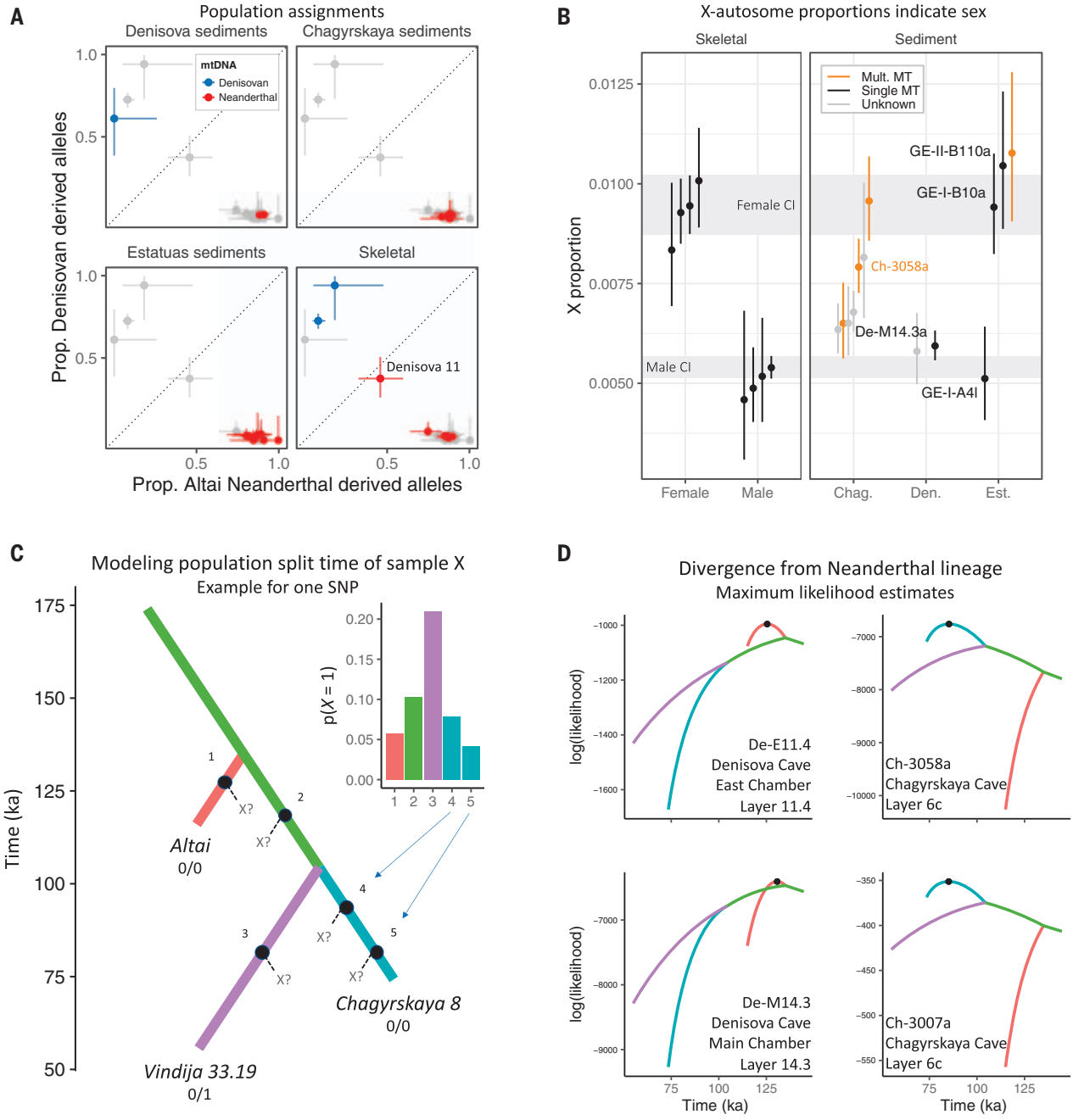


Fig. 4. Sediment nuclear DNA. (**A**) Neanderthal versus Denisovan alleles place sediment (top row and bottom left) and skeletal (bottom right) samples into broad population groups. All samples are shown in gray in all plots. CI, 95% binomial confidence interval. (**B**) X-autosome proportions for skeletal and sediment samples. Male and female CI bands (gray) denote male and female skeletal samples with the narrowest CI, respectively. Evidence of single (black) or multiple (orange) mtDNA haplotypes was taken from table S3. All samples are labeled in fig. S17. (**C**) Modeling $p(X = 1)$, the probability that a sediment sample X carries a derived allele at an example SNP in the genome, where *Vindija* is heterozygous (0/1) and *Chagyrskaya* and *Altai* are homozygous ancestral (0/0). $p(X = 1)$ depends on the time at which sample X diverges from the Neanderthal phylogeny [black dots are hypothetical split times; inset shows $p(X = 1)$ for each split time]. (**D**) Likelihood surfaces (lines) and maximum likelihood estimates of branching times for four sediment samples (black dots).

For skeletal specimens, the relative proportions of X and autosomal DNA have been used to determine sex (*3*). We applied this approach to sediment subsamples with <10% present-day human contamination and deaminated DNA fragments covering at least 5000 sites, and found that all such Denisova Cave and Estatuas samples showed X/autosome proportions consistent with hominin DNA originating primarily from a single sex (three male and three female; Fig. 4B and fig. S17). By contrast, most samples from Chagyrskaya Cave fell between the expected male and female proportions, suggesting that they contained DNA from multiple individuals of different sexes (Fig. 4B). All four libraries for which we identified a single mtDNA sequence had X/autosome proportions consistent with the DNA originating from a single sex, suggesting that they may contain DNA from individual Neanderthals (Fig. 4B, black-labeled sediment

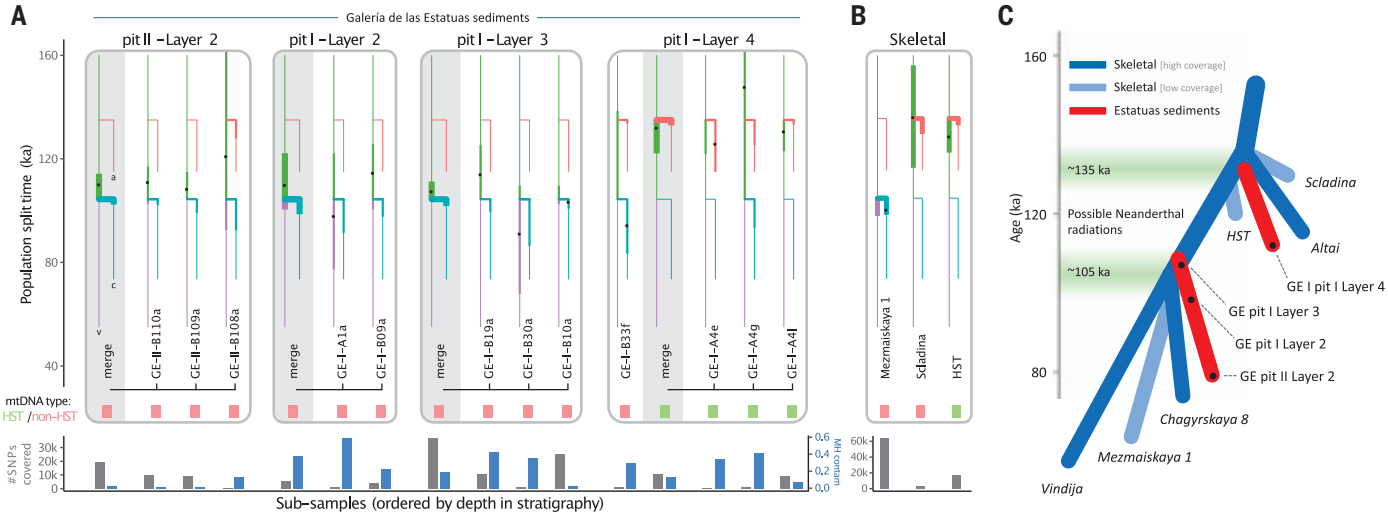


Fig. 5. Sediment samples placed on Neanderthal phylogeny. (**A**) Population split estimates for sediments across four layers of Estatuas (black dot is the maximum likelihood estimate; thick lines are 95% block bootstrap CI; CI clipped at 160 ka for GE-II-B108a and GE-I-A4g) plotted on the Neanderthal population tree with *Vindija 33.19* (v), *Chagyrskaya 8* (c), and the *Altai Neanderthal* (a). For each layer, subsamples were merged to produce per-layer estimates (gray blocks, "merge"). For pit I/layer 4, only samples with *HST*-like mtDNA (green squares) were merged. Bottom row shows the estimated number of Neanderthal SNPs used in the branch-time analysis (gray) and modern human contamination estimates (blue). (**B**) Same as (A) but for three skeletal samples. For *Mezmaiskaya 1*, a 1× coverage genome was down-sampled to ~60,000 informative reads. (**C**) Clustered split times in the Neanderthal phylogeny suggest successive radiations of Neanderthal populations ~105 and 135 ka ago.

points), although we cannot exclude the presence of identical mtDNA from multiple individuals of the same sex.

Nuclear phylogenetic analysis

To place each sample on the larger archaic phylogeny despite limited data, we developed a maximum likelihood framework that, for a sample *X*, co-estimates the point at which *X* branched from the archaic hominin tree—defined by the three Neanderthal and one Denisovan high-coverage genomes (e.g., Neanderthal phylogeny in Fig. 4C)—along with the proportion of fragments deriving from non-hominin faunal misalignment and present-day human contamination. This method makes use of the fact that all sites that are polymorphic in archaic hominins are informative for their population histories. For example, at a site that is heterozygous in one archaic genome but homozygous ancestral in the others, the probability of observing a derived allele in a sample *X* [$p(X = \text{der})$] varies based on the point at which *X* diverged from the overall tree (Fig. 4C, black points and bar plot). These probabilities are obtained from coalescent simulations for which effective population sizes and split times are inferred from the respective high-coverage genomes (*13*). Misalignment and contamination proportions are estimated independently for deaminated and nondeaminated fragments, allowing all fragments to be used in the analysis (*13*). Although some sediment samples may represent single individuals, this method operates on allele frequency expectations, making it equally applicable to samples representing single or multiple individuals from a population. When applied to previously published low-coverage genomes from skeletal samples, our method infers population split times and contamination proportions consistent with previous estimates (*13*). In a power analysis, we estimated accurate population split times in down-sampled low-coverage Neanderthal genomes (*25*) with up to 70% present-day human contamination [mean absolute error 10 to 16 ka with 500 Neanderthal DNA fragments; 4 to 7 ka with 4000 fragments; figs. S18 to S22 (*13*)] and accurately infer present-day human contamination proportions of up to 90% [fig. S20 (*13*)].

Applying this method to sediment libraries from Denisova and Chagyrskaya Caves, we found results consistent with previously published ancient DNA from skeletal elements. Specifically, we found that the two Neanderthal samples from Denisova Cave fall on the lineage leading to the *Altai Neanderthal* individual (Fig. 4D, De-E11.4 and De-M14.3, and table S14). This result is consistent with the first of these samples originating from the same layer in the East Chamber (layer 11.4) as the *Altai Neanderthal* and the second sample from a contemporaneous layer (layer 14.3) in the Main Chamber (deposited ~105 to 120 and 97 to 112 ka ago, respectively) (*4*). The Denisovan sample from layer 15 in the East Chamber (dated to ~200 ka ago) (*4*) falls on the Denisovan lineage, consistent with its mtDNA [fig. S23 (*2*)]. The sediment samples from all layers at Chagyrskaya Cave fall on the *Chagyrskaya 8* lineage (Fig. 4D, Ch-3058a and Ch-3007a, and fig. S24), consistent with a short-lived occupation of the site by Neanderthals associated with a distinctive Middle Paleolithic toolkit (*5*).

For Estatuas, where less nuclear DNA was recovered from the sediment samples and no previous genetic data exist from skeletal remains, we estimated population split times for individual subsamples with at least 500 Neanderthal DNA fragments and <70% present-day human contamination. We also merged samples to obtain per-layer estimates, with ~5000 to 36,000 fragments per layer (Fig. 5A). Samples from pit II/layer 2, pit I/layer 2, and pit I/layer 3 (*10*) diverged from the Neanderthal tree ~100 to 115 ka ago (block bootstrap 95% CIs: 102 to 114, 99 to 122, and 102 to 112 ka ago, respectively; Fig. 5A), similar to the split times of *Vindija 33.19*, *Chagyrskaya 8*, and *Mezmaiskaya 1* from each other (~104 ka ago; Fig. 5, B and C). The time of deposition of pit I/layer 3 [107 ± 8 ka ago; (*10*, *13*)] is indistinguishable from this date of divergence, suggesting that the Neanderthals from pit I/layer 3 were closely related to the ancestors of *Vindija 33.19* and *Chagyrskaya 8*. We were unable to determine whether the aforementioned layers represent a repeated occupation of the cave by the same Neanderthal population, but the split times and the mtDNA data (Fig. 2B) are consistent with this hypothesis.

In contrast, sediments from pit I/layer 4 carrying the *HST*-like mtDNA diverged from the Neanderthal tree ~122 to 135 ka ago (Fig. 5A, block bootstrap 95% CIs). This divergence

time is similar to that of the *HST* Neanderthal itself, as well as of the *Scladina Neanderthal* and the *Altai Neanderthal* (Fig. 5, B and C). The latter two carry the more common non–*HST*-like Neanderthal mtDNA type (Fig. 5B), consistent with mitochondrial diversity in the ancestral Neanderthal population, which was also observed in pit I/layer 4 (Fig. 2B). GE-I-B33f, the only sample of the common Neanderthal mtDNA type from pit I/layer 4 that yielded nuclear DNA, was collected from near the boundary with layer 3 and produced a divergence time similar to those of the layer 3 samples, albeit with a large CI (83 to 139 ka ago) because of the small dataset (867 ancient hominin DNA fragments). Taken together, these observations suggest that a population replacement occurred at Estatuas toward the end of the time of deposition of layer 4, which was accompanied by a loss of mtDNA diversity. Similar results were obtained when using only deaminated fragments (fig. S25), highlighting the robustness of the method to the effects of present-day human contamination.

Discussion

The apparent clustering of branching times suggests two distinct radiations of Neanderthal populations: *Mezmaiskaya 1*, *Vindija 33.19*, *Chagyrskaya 8*, and Estatuas pit II/layer 2 and pit I/layers 2 and 3 diverged from each other ~100 to 115 ka ago, whereas the *Altai*, *HST*, *Scladina*, and Estatuas pit I/layer 4 Neanderthals and the lineage leading to *Vindija 33.19* and *Chagyrskaya 8* diverged from each other ~135 ka ago (Fig. 5C). These radiation events therefore occurred during the early part of the Late Pleistocene and may be associated with changes in climate and environmental conditions during the last interglacial. In addition, it has been noted that the typical Neanderthal morphology evolved in several stages (*28*, *29*), with the last stage, the "classic" Neanderthals, appearing ~100 ka ago. Despite the uncertainty in dating these events, it seems plausible that the latter transition could be linked to the younger population radiation that we detected. However, determining whether such factors played a key role in the population dynamics of Neanderthals and other Pleistocene hominins (*30*, *31*) would require time-series data from additional sites and more precise estimates of the timing of these genetic events of interest. The methods presented here open the possibility of obtaining such data independently of the fossil record, limited only by biochemical constraints on long-term DNA preservation.

Our results also show that the recovery of hominin DNA from sediment may not be limited to population samples, as DNA that putatively derived from individual Neanderthals (i.e., sediment samples with a single mtDNA sequence and sex) was identified in sediment samples from all three study sites. This observation suggests that it may be possible in the future to also assess heterogeneity in the genetic composition of past populations based on the analysis of sediment DNA. In light of the substantial variation in the quantity of hominin DNA observed among sediment samples taken in close proximity (and within single samples), and considering the low abundance of hominin DNA versus non-hominin faunal DNA, it seems unlikely that the analysis of hominin DNA from Pleistocene sediments is substantially affected, if at all, by leaching of DNA through archaeological layers. However, the presence of Neanderthal DNA in layer 7 of Chagyrskaya Cave highlights the need to evaluate evidence for post-depositional mixing of sediments when assigning DNA sequences from sediment to specific layers, as is common practice when interpreting finds of artifacts, skeletal remains, and other archaeological materials. Finally, our work also highlights the value of high-coverage archaic human genomes, even if generated only in small numbers, as scaffolding for defining the past genetic landscape onto which less complete genome-wide sequence data from sediments and bones can be mapped.

Materials and Methods

Sediment samples were taken from the exposed sections of Chagyrskaya Cave and Estatuas, and subsamples with weights between 21 and 128 mg were used for DNA extraction (*14*) and single-stranded DNA library preparation (*15*). All libraries were enriched for hominin mtDNA using hybridization capture (*32*), and a subset for mammalian mtDNA (*33*) to evaluate the preservation of ancient faunal DNA. Based on the content of ancient hominin mtDNA, and aiming to recover hominin nuclear DNA from all relevant layers, libraries from this study and a previous study on DNA preservation in the sediments of Denisova Cave (*2*) were selected for hybridization capture using probes targeting phylogenetically informative positions in the nuclear genome. Mitochondrial sequences were identified on the biological family level using a previously established analysis pipeline (*2*), which includes an evaluation of the presence of deamination patterns typical for ancient DNA and assignments of the Hominidae component to specific hominid groups. Further assignments to specific branches of the hominin mtDNA tree were performed using a method based on the software kallisto (*13*, *34*). Hominin nuclear DNA sequences were identified, and phylogenetic analyses were performed as summarized in fig. S26 and described in full detail in (*13*), which also contains details on sample collection, sample preparation, and data processing. Software and scripts were written in Python (*35*) and R (*36*), and are publicly available as described in (*13*). Plots were created with ggplot2 (*37*), cowplot (*38*), RColorBrewer (*39*), and ggmap (*40*).

REFERENCES AND NOTES

1. V. Slon *et al*., The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature* **561**, 113–116 (2018). doi: 10.1038/s41586-018-0455-x; pmid: 30135579
2. V. Slon *et al*., Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605–608 (2017). doi: 10.1126/science.aam9695; pmid: 28450384
3. M. Meyer *et al*., Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* **531**, 504–507 (2016). doi: 10.1038/nature17405; pmid: 26976447
4. Z. Jacobs *et al*., Timing of archaic hominin occupation of Denisova Cave in southern Siberia. *Nature* **565**, 594–599 (2019). doi: 10.1038/s41586-018-0843-2; pmid: 30700870
5. K. A. Kolobova *et al*., Archaeological evidence for two separate dispersals of Neanderthals into southern Siberia. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2879–2885 (2020). doi: 10.1073/pnas.1918047117; pmid: 31988114
6. K. Douka *et al*., Age estimates for hominin fossils and the onset of the Upper Palaeolithic at Denisova Cave. *Nature* **565**, 640–644 (2019). doi: 10.1038/s41586-018-0870-z; pmid: 30700871
7. K. Prüfer *et al*., The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014). doi: 10.1038/nature12886; pmid: 24352235
8. M. Meyer *et al*., A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012). doi: 10.1126/science.1224344; pmid: 22936568
9. F. Mafessoni *et al*., A high-coverage Neandertal genome from Chagyrskaya Cave. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15132–15136 (2020). doi: 10.1073/pnas.2004944117; pmid: 32546518
10. M. Demuro, L. J. Arnold, A. Aranburu, A. Gómez-Olivencia, J.-L. Arsuaga, Single-grain OSL dating of the Middle Palaeolithic site of Galería de las Estatuas, Atapuerca (Burgos, Spain). *Quat. Geochronol.* **49**, 254–261 (2019). doi: 10.1016/j.quageo.2018.02.006
11. J. L. Arsuaga *et al*., Evidence of paleoecological changes and Mousterian occupations at the Galería de las Estatuas site, Sierra de Atapuerca, northern Iberian plateau, Spain. *Quat. Res.* **88**, 345–367 (2017). doi: 10.1017/qua.2017.46
12. A. Pablos, A. Gómez-Olivencia, J. L. Arsuaga, A Neandertal foot phalanx from the Galería de las Estatuas site (Sierra de Atapuerca, Spain). *Am. J. Phys. Anthropol.* **168**, 222–228 (2019). doi: 10.1002/ajpa.23729; pmid: 30351468
13. Materials and methods are available as supplementary materials.
14. N. Rohland, I. Glocke, A. Aximu-Petri, M. Meyer, Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* **13**, 2447–2461 (2018). doi: 10.1038/s41596-018-0050-5; pmid: 30323185
15. M.-T. Gansauge, A. Aximu-Petri, S. Nagel, M. Meyer, Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020). doi: 10.1038/s41596-020-0338-0; pmid: 32612278
16. D. H. Huson *et al*., MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Comput. Biol.* **12**, e1004957 (2016). doi: 10.1371/journal.pcbi.1004957; pmid: 27327495
17. R. Bouckaert *et al*., BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* **10**, e1003537 (2014). doi: 10.1371/journal.pcbi.1003537; pmid: 24722319
18. C. Posth *et al*., Deeply divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neanderthals. *Nat. Commun.* **8**, 16046 (2017). doi: 10.1038/ncomms16046; pmid: 28675384
19. S. Peyrégne *et al*., Nuclear DNA from two early Neandertals reveals 80,000 years of genetic continuity in Europe. *Sci. Adv.* **5**, eaaw5873 (2019). doi: 10.1126/sciadv.aaw5873; pmid: 31249872
20. A. R. Skinner *et al*., ESR dating at Mezmaiskaya Cave, Russia. *Appl. Radiat. Isot.* **62**, 219–224 (2005). doi: 10.1016/j.apradiso.2004.08.008; pmid: 15607452
21. M. Toussaint, D. Bonjean, *The Scladina I-4A Juvenile Neandertal, Andenne, Belgium: Palaeoanthropology and Context* (Université de Liège, 2014).

22. P. Skoglund *et al*., Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2229–2234 (2014). doi: 10.1073/pnas.1318934111; pmid: 24469802
23. K. A. Kolobova, A. V. Shalagina, V. P. Chabai, S. V. Markin, A. I. Krivoshapkin, Signification des technologies bifaciales au Paléolithique moyen des montagnes de l'Altaï. *Anthropologie* **123**, 276–288 (2019). doi: 10.1016/j.anthro.2019.06.004
24. D. E. Wood, S. L. Salzberg, Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014). doi: 10.1186/gb-2014-15-3-r46; pmid: 24580807
25. M. Hajdinjak *et al*., Reconstructing the genetic history of late Neanderthals. *Nature* **555**, 652–656 (2018). doi: 10.1038/nature26151; pmid: 29562232
26. K. Prüfer *et al*., A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017). doi: 10.1126/science.aao1887; pmid: 28982794
27. S. Sawyer *et al*., Nuclear and mitochondrial DNA sequences from two Denisovan individuals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15696–15700 (2015). doi: 10.1073/pnas.1519905112; pmid: 26630009
28. D. Dean, J.-J. Hublin, R. Holloway, R. Ziegler, On the phylogenetic position of the pre-Neandertal specimen from Reilingen, Germany. *J. Hum. Evol.* **34**, 485–508 (1998). doi: 10.1006/jhev.1998.0214; pmid: 9614635
29. J. L. Arsuaga *et al*., Neandertal roots: Cranial and chronological evidence from Sima de los Huesos. *Science* **344**, 1358–1363 (2014). doi: 10.1126/science.1253958; pmid: 24948730
30. R. Potts, Evolution and environmental change in early human prehistory. *Annu. Rev. Anthropol.* **41**, 151–167 (2012). doi: 10.1146/annurev-anthro-092611-145754
31. J.-J. Hublin, "Climatic changes, paleogeography, and the evolution of the Neandertals," in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, O. Bar-Yosef, Eds. (Springer, 1998), pp. 295–310.
32. Q. Fu *et al*., DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013). doi: 10.1073/pnas.1221359110; pmid: 23341637
33. V. Slon *et al*., Mammalian mitochondrial capture, a tool for rapid screening of DNA preservation in faunal and undiagnostic remains, and its application to Middle Pleistocene specimens from Qesem Cave (Israel). *Quat. Int.* **398**, 210–218 (2016). doi: 10.1016/j.quaint.2015.03.039
34. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016). doi: 10.1038/nbt.3519; pmid: 27043002
35. G. Van Rossum, F. L. Drake Jr., "Python tutorial" (Centrum voor Wiskunde en Informatica, 1995).
36. R Core Team, "R: A language and environment for statistical computing" (R Foundation for Statistical Computing, 2019); https://www.R-project.org/.
37. H. Wickham, "ggplot2: Elegant graphics for data analysis" (Tidyverse, 2016); https://ggplot2.tidyverse.org.
38. C. O. Wilke, "cowplot: Streamlined plot theme and plot annotations for "ggplot2" (R package version 1.1.1, 2019); https://CRAN.R-project.org/package=cowplot
39. E. Neuwirth, "RColorBrewer: ColorBrewer palettes" (R package 1.1-2, 2014); https://CRAN.R-project.org/package=RColorBrewer.
40. D. Kahle, H. Wickham, ggmap: Spatial Visualization with ggplot2. *R J.* **5**, 144–161 (2013). doi: 10.32614/RJ-2013-014
41. B. Vernot, "bvernot/branch_estim: v0.1-beta" (Zenodo, 2021); https://zenodo.org/record/4534664.
42. B. Vernot, "bvernot/make_sed_capture_targets: v0.1-beta" (Zenodo, 2021); https://zenodo.org/record/4534734.
43. F. Romagne, "mpieva/nuclear_sediment_pipeline: metagen_utils v1.21" (Zenodo, 2021); https://zenodo.org/record/4468181.
44. J. Visagie, "mpieva/biohazard-tools: Biohazard Tools v0.2" (Zenodo, 2021); https://zenodo.org/record/4616582.
45. K. Prüfer, "Alignments of modern and archaic human genomes with primate outgroup genomes in tab-separated-values format" (EDMOND, 2021); https://dx.doi.org/10.17617/3.5h.

ACKNOWLEDGMENTS

We thank S. Pääbo for supporting the project and commenting on the manuscript; Moritz C. Meyer and K. O'Gorman for help with sampling; B. Schellbach and A. Weihmann for help in the laboratory; S. Peyregne for data assistance; L. Jáuregui for organizational support; and the Galería de las Estatuas, Chagyrskaya Cave, and Denisova Cave excavation and research teams for providing the foundations that made this project possible. **Funding:** This work was supported by the Max Planck Society, the European Research Council (grant 694707 to Svante Pääbo) and the Australian Research Council (fellowships FL130100116 to R.G.R., FT140100384 to B.L., and FT150100138 to Z.J.). The archaeological field studies were funded by the Russian Foundation for Basic Research (project 20-29-01011 to M.V.S. and A.P.D.), the Russian Science Foundation (project 19-48-04107 to A.I.K.), and the Junta de Castilla y León and Fundación Atapuerca. The geological study in Chagyrskaya Cave was supported by the National Science Center, Poland (project 2018/29/B/ST10/00906 to M.T.K.). A.G.-O. and N.S. were supported by Ramón y Cajal (RYC-2017-22558) and Juan de la Cierva-Incorporación (IJCI-2017-32804) fellowships, respectively. B.Vi.'s research was supported by an Insight grant from the Social Sciences and Humanities Research Council (Canada). Further support was provided by the FEDER/Ministerio de Ciencia e Innovación-Agencia Estatal de Investigación (projects PGC2018-093925-B-C31, PGC2018-093925-B-C32, and PGC2018-093925-B-C33). **Author contributions:** B.Ve. and M.M. conceived the study. A.G.-O., Z.J., N.S., A.Pa., A.A., J.M.B.d.C., E.C., B.L., M.T.K., A.I.K., K.A.K., M.B.K., M.V.S., A.P.D., B.Vi., D.L.H., R.G.R., J-L.A., and M.M. sampled sediments and/or provided archaeological context and interpretation of results. E.I.Z., E.E., S.N., B.N., J.R., and A.S. performed aDNA extraction and sequencing. B.Ve., E.I.Z., V.S., F.M., F.R., A.Pe., M.P., S.G., B.P., J.K., and M.M. analyzed data. B.Ve., E.I.Z., A.G.-O., Z.J., V.S., R.G.R., J-L.A., and M.M. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are presented in the paper and/or the supplementary materials. Sequencing data have been deposited in the ENA, reference number PRJEB42656. Software is archived at Zenodo (*41*–*44*). Primate alignments are available at EDMOND (*45*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/eabf1667/suppl/DC1
Materials and Methods
Figs. S1 to S26
Tables S1 to S14
References (*46*–*103*)
MDAR Reproducibility Checklist

9 October 2020; accepted 31 March 2021
Published online 15 April 2021
10.1126/science.abf1667

RESEARCH ARTICLE SUMMARY

NEURODEVELOPMENT

The coding and long noncoding single-cell atlas of the developing human fetal striatum

Vittoria Dickinson Bocchi, Paola Conforti, Elena Vezzoli, Dario Besusso, Claudio Cappadona, Tiziana Lischetti, Maura Galimberti, Valeria Ranzani, Raoul J. P. Bonnal, Marco De Simone, Grazisa Rossetti, Xiaoling He, Kenji Kamimoto, Ira Espuny-Camacho, Andrea Faedo, Federica Gervasoni, Romina Vuono, Samantha A. Morris, Jian Chen, Dan Felsenfeld, Giulio Pavesi, Roger A. Barker, Massimiliano Pagani*, Elena Cattaneo*

INTRODUCTION: The striatum modulates distinct characteristics of human social behavior and is an area affected in many neurological diseases. We created a comprehensive single-cell atlas of this area during early human fetal development, considering both protein-coding transcripts and long intergenic noncoding RNAs (lincRNAs).

RATIONALE: Understanding of the molecular mechanisms that define human striatal development has been limited by the scarcity of relevant fetal tissue and the use of only a limited panel of protein-coding genes in most gene identification studies. We created a cell-specific molecular atlas of the lateral ganglionic eminence (LGE), the striatal primordium. Our first goal was to develop a catalog of de novo identified lincRNAs of this area using bulk RNA sequencing. This catalog should help to clarify the specific characteristics of human development because lincRNAs exhibit accelerated evolution, are highly cell-specific, and are required for brain development. Our second goal was to understand how the medium spiny neurons (MSNs), the principal cell types in the striatum, differentiate and diversify, and which genes act as master regulators of fate determination. MSNs diversify into D1 and D2 types, so named for their expression of one of the two variants of the human dopamine receptor (D1 and D2). We used single-cell RNA sequencing to infer the developmental landscape of MSNs and to define and validate fate markers.

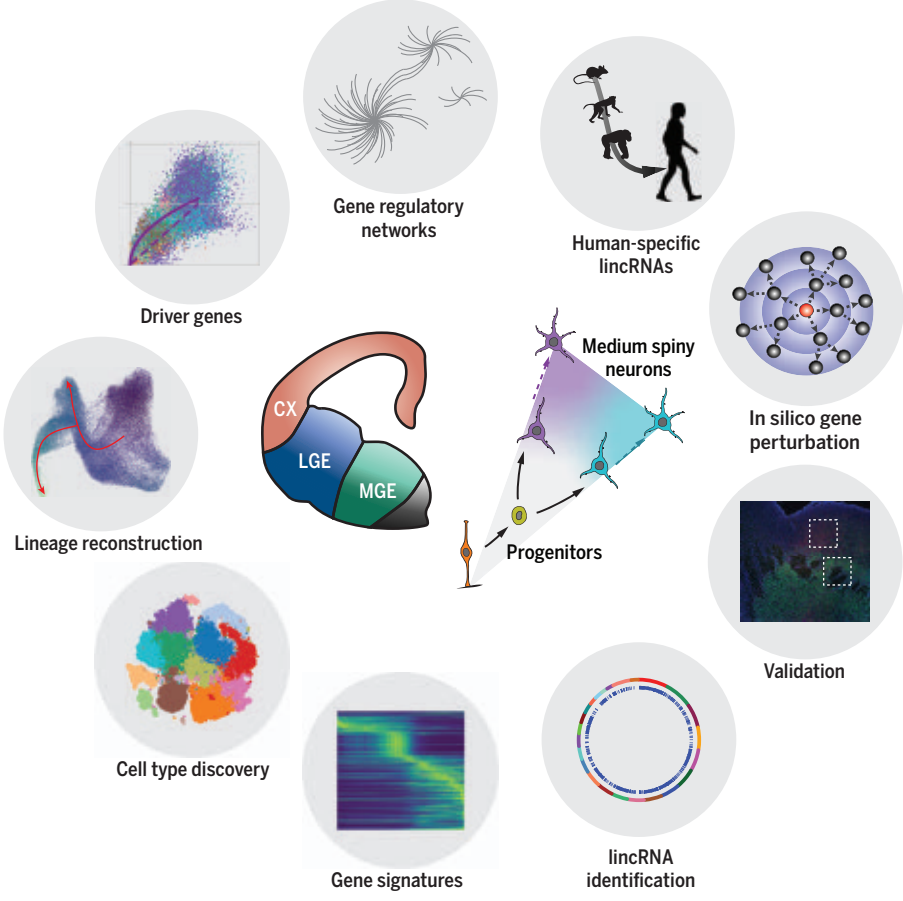


ILLUSTRATION: TIZIANA LISCHETTI

The molecular blueprint of striatal development. Combined bulk and single-cell RNA sequencing of the human fetal striatum reveals cell states together with their key coding and lincRNA fate determinants, and defines the developmental hierarchies underlying lineage commitment in medium spiny neurons. CX, neocortex; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence.

RESULTS: Bulk RNA sequencing enabled the annotation of 1116 novel lincRNAs of different areas of the human developing telencephalon, and we found that these lincRNAs are less conserved among species than those previously identified in the adult brain. Bulk measurements enabled us to pinpoint the distinctive signature of the striatum relative to surrounding areas, and we determined that huntingtin (HTT) is a specific upstream regulator of this region. We then profiled 96,789 single cells of the LGE, based on both coding RNAs and the newly identified lincRNAs. This enabled us to uncover the transcriptional profiles of 15 different cell states that included lincRNAs that were gained throughout evolution. We found that a common progenitor generates both D1- and D2-MSNs and that this progenitor is distinct from the progenitor of interneurons. We also discovered a postmitotic precursor cell state for both D1- and D2-MSNs, which falls within a continuum of key fate determinants. Finally, we identified a panel of gene regulatory networks that define D1- and D2-MSNs, and we showed that in silico knockout of the transcription factors governing these networks could cause the arrest of both MSN lineages, blockage of a specific MSN class, or switching between MSN fates.

CONCLUSION: Our findings reveal the differentiation hierarchies that govern human striatal development. We anticipate that the set of transcription factors and lincRNAs identified in this study will be leveraged to recreate MSN differentiation in vitro and that these cells can then be used for cell replacement therapies in Huntington's disease (HD). Furthermore, we expect that this atlas will guide investigations of the developmental components related to HD. Finally, we foresee that our lincRNA catalog will contribute to understanding the additional layer of fine-tuning mechanisms present in the human striatum but not in other species. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: elena.cattaneo@unimi.it (E.C.); massimiliano.pagani@unimi.it (M.P.)
Cite this article as V. D. Bocchi *et al.*, *Science* **372**, eabf5759 (2021). DOI: 10.1126/science.abf5759

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf5759

RESEARCH ARTICLE

NEURODEVELOPMENT

The coding and long noncoding single-cell atlas of the developing human fetal striatum

Vittoria Dickinson Bocchi[1,2], Paola Conforti[1,2], Elena Vezzoli[1,2]†, Dario Besusso[1,2], Claudio Cappadona[1,2]‡, Tiziana Lischetti[1,2], Maura Galimberti[1,2], Valeria Ranzani[2], Raoul J. P. Bonnal[2]§, Marco De Simone[2]¶, Grazisa Rossetti[2]§, Xiaoling He[3], Kenji Kamimoto[4,5,6], Ira Espuny-Camacho[1,2]#, Andrea Faedo[1,2]**, Federica Gervasoni[2,7]§, Romina Vuono[3]††, Samantha A. Morris[4,5,6], Jian Chen[8], Dan Felsenfeld[8], Giulio Pavesi[1], Roger A. Barker[3], Massimiliano Pagani[2,7]§*, Elena Cattaneo[1,2]*

Deciphering how the human striatum develops is necessary for understanding the diseases that affect this region. To decode the transcriptional modules that regulate this structure during development, we compiled a catalog of 1116 long intergenic noncoding RNAs (lincRNAs) identified de novo and then profiled 96,789 single cells from the early human fetal striatum. We found that D1 and D2 medium spiny neurons (D1- and D2-MSNs) arise from a common progenitor and that lineage commitment is established during the postmitotic transition, across a pre-MSN phase that exhibits a continuous spectrum of fate determinants. We then uncovered cell type–specific gene regulatory networks that we validated through in silico perturbation. Finally, we identified human-specific lincRNAs that contribute to the phylogenetic divergence of this structure in humans. This work delineates the cellular hierarchies governing MSN lineage commitment.

The striatum, a subcortical structure made up of the caudate nucleus and putamen, is important in motor control and learning, procedural behavior, cognition, and emotional and motivational responses. Many of these functions are associated with uniquely human abilities including the development of speech and language (*1*). In adults, the striatum is primarily composed of medium spiny neurons (MSNs) carrying either D1-type or D2-type dopamine receptors interspersed with aspiny interneurons (*2*). This organization does not appear to vary among species (*3*). However, this apparent homogeneity in cellular composition masks functional complexity of human-specific striatal circuits.

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, Milan, Italy. [2]INGM, Istituto Nazionale Genetica Molecolare, Milan, Italy. [3]WT-MRC Cambridge Stem Cell Institute and Department of Clinical Neuroscience, University of Cambridge, Cambridge, UK. [4]Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110, USA. [5]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA. [6]Center of Regenerative Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA. [7]Dipartimento di Biotecnologie Mediche e Medicina Traslazionale, Università degli Studi di Milano, Milan, Italy. [8]CHDI Management/CHDI Foundation, New York, NY, USA. †Present address: Department of Biomedical Sciences for Health, Università degli Studi di Milano, 20133 Milan, Italy. ‡Present address: Department of Biomedical Sciences, Humanitas University, Pieve Emanuele 20090 Milan, Italy. §Present address: IFOM-FIRC Institute of Molecular Oncology, Milan, Italy. ¶Present address: Department of Radiation Oncology, Cedars-Sinai Medical Center, Los Angeles, CA, USA. #Present address: Université de Liège, GIGA Stem Cells, Quartier hôpital 15, B-4000 Liège, Belgium. **Present address: Axxam, OpenZone, 20091 Bresso, Milan, Italy. ††Present address: Medway School of Pharmacy, University of Kent, Chatham, Kent, UK.
*Corresponding author. Email: elena.cattaneo@unimi.it (E.C.); massimiliano.pagani@unimi.it (M.P.)

During development, MSNs are derived from progenitors in the lateral ganglionic eminence (LGE) of the telencephalon (*4*). Most studies of this area have concentrated on rodent models and the coding signature of limited cell types. Long noncoding RNAs (lncRNAs) have a higher tissue specificity than mRNAs (*5*) and a diversity that correlates more closely with brain complexity than protein-coding genes (*6*). LncRNAs may, therefore, better discriminate cell states during striatal development in humans.

In this study, we first conducted bulk RNA-seq to identify novel long intergenic noncoding RNAs (lincRNAs) expressed in the developing human LGE that we then leveraged to define the spatial coding and noncoding map for this area that discriminates it from the surrounding neocortex (CX) and medial ganglionic eminence (MGE). We then performed single-cell RNA-seq (scRNA-seq) to characterize cell diversity in the LGE, identify LGE-specific coding and noncoding cell states, define cell type–specific gene regulatory networks and decipher pivotal bifurcation points in the formation of human MSNs. Finally, we performed in silico perturbations by knocking out and overexpressing key transcription factors (TFs) of the MSN lineage, together with immunohistochemistry and fluorescence in situ hybridization (FISH) analysis to validate our findings.

lincRNA discovery and scRNA-seq profiling of the human LGE

To identify unannotated putative lincRNAs expressed during human striatal development, we dissected the CX, LGE, and MGE from human embryos between 7 and 20 post-conceptional weeks (pcw), with at least two or three biological replicates for each developmental stage (table S1). We then profiled by RNA-seq the transcriptome of each area (Fig. 1A) and set up a computational pipeline (fig. S1A) to create a corresponding catalog of lincRNAs (fig. S1, B to E). Because our bulk RNA-seq protocol was unstranded, and because genic lncRNAs are difficult to correctly predict without strand information discriminating them from overlapping genes, we decided to include only a reliable set of intergenic lncRNAs (lincRNAs). To better understand the relationship between lincRNAs identified in other tissues and our newly identified lincRNAs, we integrated two catalogs derived from different human tissues and cell types (*7*, *8*), plus a lincRNA catalog of the developing human CX (*9*), into this study.

We detected 1116 novel lincRNA loci (Fig. 1B and table S2), among which the highest number was found in the developing LGE and the lowest in the CX (fig. S1F), probably because of the greater extent of data available from the CX. We found that lincRNAs had an average of 2.27 exons; this was also the case for all lincRNA catalogs integrated into this study, with the exception of the FANTOM catalog, which used CAGE-seq instead of classic RNA-seq and showed fewer exons (fig. S1G). For each sampled pcw, we then defined a signature of protein-coding genes and lincRNAs that were uniquely expressed in either the LGE, the CX, or the MGE (Fig. 1C, fig. S2A, and table S3). Gene Ontology (GO) analysis in the LGE between 7 and 11 pcw revealed an enrichment for terms related to neural differentiation and to forebrain and subpallium development (fig. S2B and table S4), whereas at 20 pcw this changed to regulation of synaptic plasticity and neurotransmitter secretion, reflecting the more mature state of the 20-pcw striatum (fig. S2C and table S4).

Pathway enrichment analysis (IPA, Ingenuity Systems) revealed huntingtin (HTT) as the most significant upstream regulator ($P < 0.001$) of the LGE at all time points considered (Fig. 1D, fig. S2D, and table S5). This suggests that HTT may have an important role in human striatal development.

To explore how these protein-coding genes and lincRNAs define specific cell states during development of the human striatum, we surveyed 96,789 high-quality single cells from the LGE between 7 and 11 pcw (Fig. 1A and table S1). We were able to discriminate 15 clusters and their transcriptional signatures (Fig. 1E and table S6) that were then classified according to canonical markers (Fig. 1F). Biological replicates were well distributed in each detected cluster (fig. S3A), and all time points considered between 7 and 11 pcw covered the identified clusters (fig. S3, B and C), indicating that during this stage of early development the same cell types are present. Both D1- and D2-MSN subtypes were

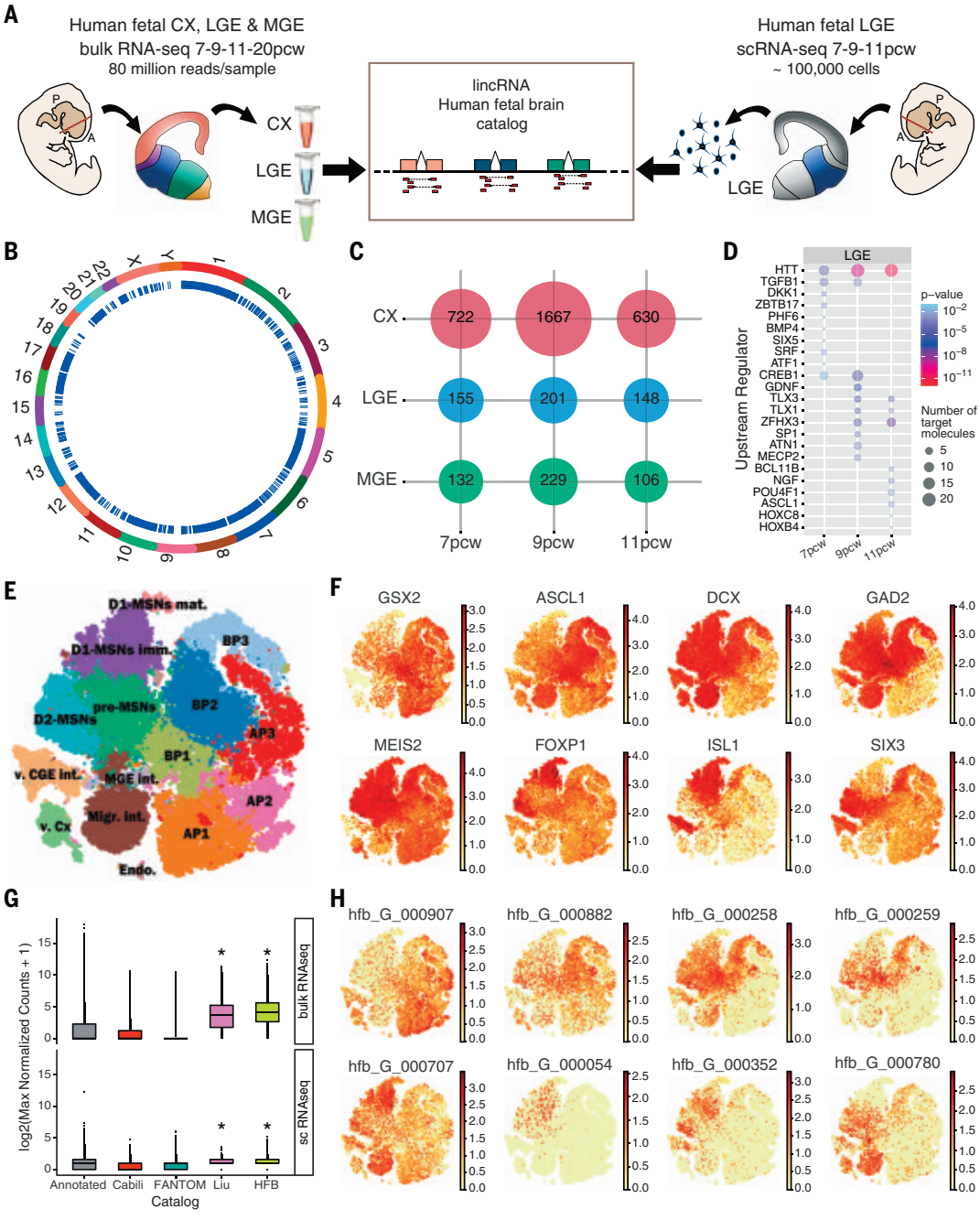


Fig. 1. The coding and noncoding transcriptional landscape of the developing human striatum. (**A**) Schematic representation of the experimental design. CX, neocortex; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; pcw, post-conceptional weeks; A, anterior; P, posterior. (**B**) Circular plot showing the genomic location of 1116 novel lincRNAs. Outer ring, chromosomes; inner ring, loci of newly identified lincRNAs. (**C**) Bubble matrix showing the number of uniquely expressed protein-coding genes and lincRNAs per area and per pcw from bulk RNA-seq data. (**D**) Upstream regulators of the bulk LGE-specific signature between 7 and 11 pcw. (**E**) t-SNE (t-distributed stochastic neighbor embedding) plot of 96,789 single cells from the LGE between 7 and 11 pcw, color-coded by cell type. AP, apical progenitors; BP, basal progenitors; pre-MSNs, precursor medium spiny neurons; D1-MSNs imm., immature D1 medium spiny neurons; D1-MSNs mat., mature D1 medium spiny neurons; D2-MSNs, D2 medium spiny neurons; MGE int., MGE interneurons; Migr. int., migrating interneurons; v.CGE int., ventral caudal ganglionic eminence interneurons; v.Cx, ventral neocortical neurons; Endo., endothelial cells. (**F**) Gene expression levels of markers for early progenitors (*GSX2*), intermediate progenitors (*ASCL1*), neurons (*DCX*), GABAergic neurons (*GAD2*), LGE-lineage cells (*MEIS2*), general MSNs (*FOXP1*), D1-MSNs (*ISL1*), and D2-MSNs (*SIX3*). (**G**) Boxplot showing distribution of maximal normalized expression of lincRNAs from different catalogs in bulk and single-cell data. *$P < 2 \times 10^{-16}$ (pairwise comparisons: Liu (lincRNA catalog of the developing CX) versus FANTOM (lincRNA catalog of different adult tissues and cell types), Cabili (lincRNA catalog of different adult tissues and cell types), and annotated lincRNAs; HFB (human fetal brain; catalog of lincRNAs identified de novo in this study) versus FANTOM, Cabili, and annotated lincRNAs; Wilcoxon test with Bonferroni correction). Boxplots show the first, second, and third quartiles. Lines summarize values within 1.5 times the first and third interquartile ranges. Points beyond the whiskers are outliers. (**H**) Gene expression levels of highly specific lincRNAs identified de novo in this study.

present in all time points considered (fig. S3C) and were characterized by known markers such as *ISL1* in D1-MSNs and *SIX3* in D2-MSNs (Fig. 1F and fig. S4, A to C).

GO analysis revealed that apical progenitors (APs) are associated with terms such as cell adhesion and glial cell differentiation, mirroring the proliferating nature of this cell type. Basal progenitors (BPs) express genes related to mRNA splicing, RNA binding, and forebrain development, which suggests that a posttranscriptional control network coordinates and primes these cells for their final cell division and fate acquisition. Genes identified in the pre-MSN phase are enriched in GO terms linked to nervous system development and axon guidance, whereas D1- and D2-MSNs show a more mature gene signature linked to synapse organization and chemical synaptic transmission (fig. S4D and table S7).

Marginally contaminating populations of cells included *NKX2.1*-, *LHX6*-, and *LHX8*-expressing MGE interneurons, migrating interneurons, ventral neocortical cells, ventral caudal ganglionic eminence (CGE) cells, and endothelial cells (fig. S5, A to C). By subclustering the migrating interneurons, we were able to observe dorsal LGE (dLGE) interneurons expressing *SP8* and *COUP-TFII* (*NR2F2*) together with *PAX6*; we also observed MGE migrating interneurons

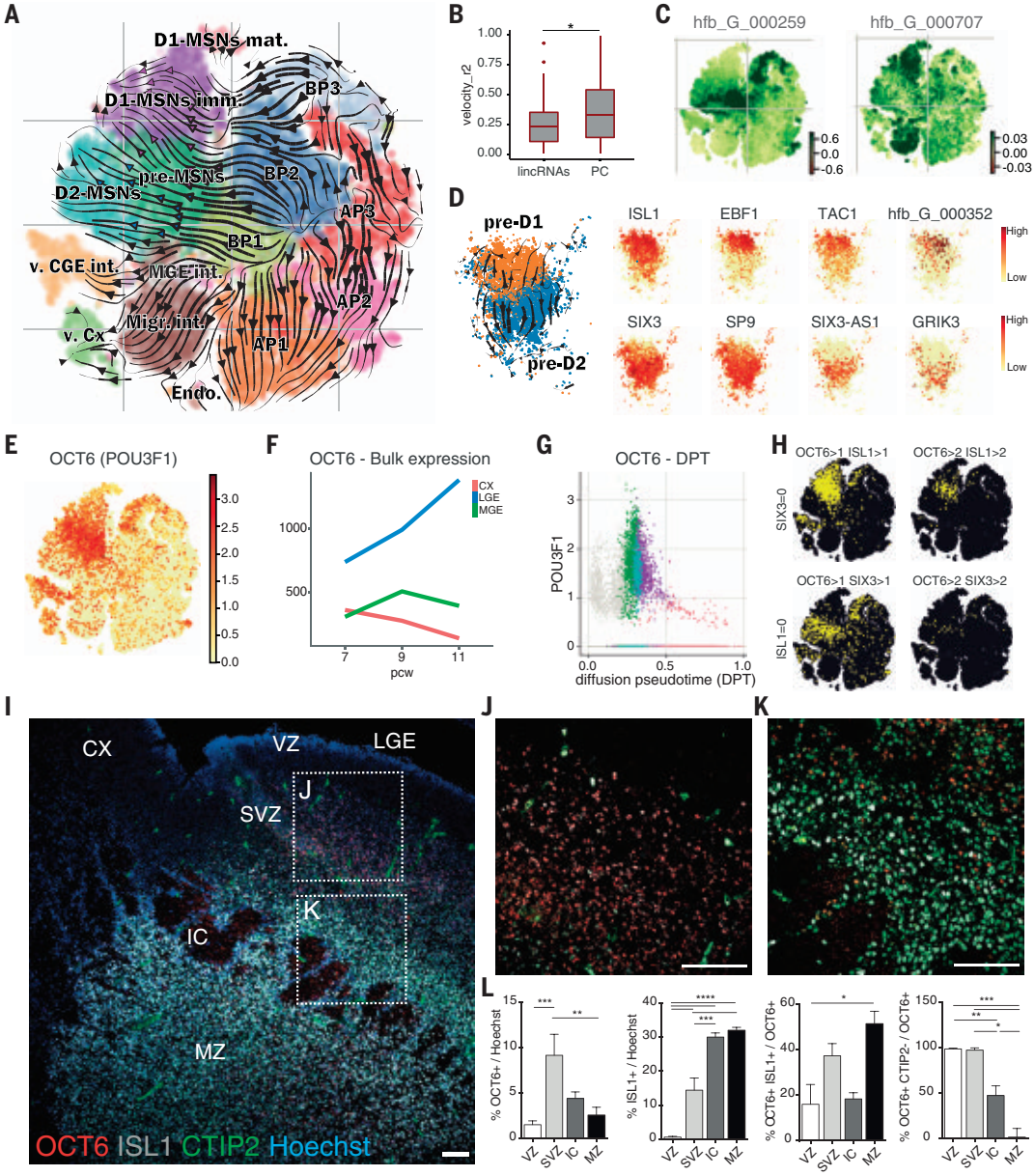


Fig. 2. Maturation and differentiation trajectory of MSNs. (**A**) Velocity estimates projected onto a two-dimensional t-SNE plot of the LGE dataset. Arrowheads in the pre-MSN cluster are colored according to commitment to either the D1 (purple) or D2 (cyan) MSN fate. (**B**) Boxplot showing splicing kinetics of lincRNAs and protein-coding genes. *P < 0.05 (Wilcoxon test with Bonferroni correction). (**C**) t-SNE plot showing velocities of each gene. Positive velocities (dark green) indicate that a gene is up-regulated, which occurs for cells that show higher abundance of unspliced mRNA for that gene than expected. (**D**) Subclusters within the pre-MSN cluster and the velocity vector fields, together with expression levels of canonical D1-MSN (top row) and D2-MSN (bottom row) markers. (**E** to **G**) Single-cell (E) and bulk (F) expression levels, together with single-cell expression levels (cells are colored according to Louvain cluster) plotted against pseudotime (G), of the pre-MSN specific marker *OCT6* (*POU3F1*). (**H**) t-SNE plot showing cells coexpressing *OCT6*/*ISL1* and *OCT6*/*SIX3* with different thresholds of gene expression (coexpressing cells are shown in yellow). (**I**) OCT6, ISL1, and CTIP2 staining of a telencephalic coronal hemisection at 9 pcw. IC, internal capsule. Scale bar, 100 μm. (**J** and **K**) Magnification (40×) of the SVZ and IC sections marked in (I). (**L**) Automatic quantification of the percentage of cells positive for OCT6, ISL1, and CTIP2 with the NIS software on confocal images at 40× magnification. N = 3 to 6 fields for each zone: VZ, SVZ, IC, and MZ from two or three coronal slices of one fetus; two or three z stacks of each field are pre-mediated. Statistics were performed with Prism: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001 (one-way analysis of variance, Bonferroni posttest).

expressing *NXK2.1* and *LHX6* that were presumably passing through the subventricular zone (SVZ) of the LGE at this developmental stage (because the same population did not express *LHX8*) (fig. S5, C and D). As expected, *ERBB4* was found in both types of migrating interneurons (fig. S5D). The findings that these two populations cluster together despite their different sites of origin and that subclustering was required to reveal their diversity suggest that their interneuron migratory class signature is stronger than their lineage of origin signature. We did not observe any markers of oligodendrocytes, astrocytes, or microglia, indicating that these cells appear later in normal human striatal development, and found no evidence of corridor cells that have probably migrated to the globus pallidus by 7 pcw in humans.

As previously reported (*7*), in our study bulk expression levels of lincRNAs are lower than those of protein-coding genes, and our single-cell data also show differential expression for these two biotypes (fig. S5E). However, when comparing the different catalogs, we observed that lincRNAs identified in the developing human CX (*9*) and those identified in this study (HFB) displayed higher transcript levels, both in bulk and single-cell data, relative to lincRNAs identified in different adult tissues (Fig. 1G). This suggests that these lincRNAs are associated with the development of the human telencephalon. Overall, lincRNAs are more specific than protein-coding genes both in bulk and single-cell measurements (fig. S5F). We found that these lincRNAs characterize different cell states of the LGE lineage; for example, the lincRNA *hfb_G_000907* is associated with APs, whereas *hfb_G_000258* characterizes pre-MSNs and D2-MSNs (Fig. 1H).

MSN differentiation passes through a pre-MSN cell state

To infer how fate decisions occur, we defined the cell trajectories that give rise to MSNs

using velocyto (*10*, *11*). The velocity field map (Fig. 2A) reflects the dynamics of MSN differentiation as these cells transition from APs to D1- and D2-MSNs, passing through a pre-MSN phase. This model shows that in the pre-MSN phase, there is already a separation between the D1 and D2 states (Fig. 2A). We found that lincRNAs display reduced splicing kinetics relative to protein-coding genes (Fig. 2B), reflecting their inefficient splicing (*12*). However, within the putative driver genes (table S8), we identified a number of lincRNAs that may guide MSN differentiation. These include *hfb_G_000259*, which shows pronounced velocities in BPs, pre-MSNs, and D2-MSNs, and *hfb_G_000707*, which contributes to the D1-MSN transition (Fig. 2C).

Overall, our data suggest that the same progenitors give rise to both D1- and D2-MSNs. Subclustering the AP and BP groups confirmed the absence of lineage commitment at these stages (fig. S6, A and B). This was further proved by calculating the connectivity between each cluster of the LGE lineage by means of partition-based graph abstraction (PAGA), which showed that progenitors are highly interconnected and converge to give rise to D1- and D2-MSNs through the pre-MSN phase (fig. S6C). To explore the potential temporal relationships between cell states in the LGE lineage, we used a measure of graph distance (diffusion pseudotime, DPT) (fig. S6, D and E). We observed that cell states and DPT recapitulate known temporal dynamics of neural maturation, with mitotic APs of the ventricular zone (VZ) having the lowest DPT score and mitotic BPs of the SVZ having an intermediate DPT score (fig. S6E). These findings suggest that the latter cells represent a more mature progenitor state, as reflected by the expression of *HES6*, which is found in cells committing to neural differentiation.

Pre-MSNs together with D1- and D2-MSNs were classified as postmitotic (fig. S6F) and showed the highest DPT score, which peaked in cells classified as mature D1-MSNs (fig. S6E). This cell population was the only one to express the more mature neuronal cytoskeleton marker neurofilament-M (*NEFM*) (fig. S6G), and most cells of this cluster are found at the most advanced pcw analyzed (fig. S3C). This temporal signature was confirmed using FISH (fig. S6H). Finally, at the cluster level, we found that splicing kinetics and differentiation accelerates substantially after cell cycle exit (especially in pre-MSNs and D2 neurons), maintaining pace during D2 production but slowing down during D1 production (fig. S6I). This finding is consistent with a model in which immature differentiating cells (D2-MSNs) are less transcriptionally stable and reveal greater splicing dynamics than terminally differentiated cells (D1-MSNs) (*13*). Overall, these findings suggest that both MSN subtypes are present at this early stage of development but that D1-MSNs mature and reach equilibrium at a faster rate than D2-MSNs.

We then focused on pre-MSNs because this cell state escaped previous identification. Subclustering of pre-MSNs revealed that these cells exhibit either a D1- or D2-MSN blueprint with enrichment of cell-specific markers of each subtype (Fig. 2D). Although subclustering revealed two discrete cell states, markers such as *SIX3*, *SP9*, and *ISL1* showed a gradient of opposing expression, with both markers found in both subclusters at different expression levels. We therefore tested different thresholds of *SIX3* and *ISL1* coexpression in the entire LGE lineage to test how the transcripts behaved throughout differentiation. We found that high levels of *SIX3* and low levels of *ISL1* are present in the D2-MSN lineage at the pre-MSN stage, and that coexpression is lost at terminal D2-MSN fates (fig. S7A). The opposite trend was seen for high *ISL1* and low *SIX3* in the D1 MSN lineage (fig. S7A). High levels of both transcripts were instead found in a very small percentage of cells (~7%), and this was also reflected at the protein level (fig. S4, B and C). This suggests that the pre-MSN phase spans a transcriptional continuum rather than being a discrete cell state.

We then identified *OCT6* (*POU3F1*), a member of the POU-III subfamily, as a specific marker of this pre-MSN cell state (Fig. 2E), as well as being LGE-specific (Fig. 2F) and exhibiting transient expression (Fig. 2G). When looking at its expression at the pre-MSN stage, we also found that this gene is more enriched in pre-D1 MSNs and appears to follow an expression gradient (fig. S7B). Examining coexpression levels of this transcript with candidate D1- and D2-MSN markers, we observed that low levels of *OCT6-ISL1* characterize the D1 lineage, whereas high levels define the pre-D1 state (Fig. 2H). Low levels of *OCT6-SIX3* were also found in D2-MSNs; however, very few positive cells were found with high expression of both markers (Fig. 2H). To confirm these findings, we performed systematic immunohistochemistry analysis of these markers at 9 pcw. We observed specific OCT6 staining in the SVZ of the LGE (Fig. 2, I to K), with none in the CX, MGE, and CGE (fig. S7C). We also found that OCT6-positive cells are postmitotic neurons, as only 2% are double-positive for the BP marker ASCL1 (fig. S7, D to G) and more than 80% of OCT6 cells do not express the proliferative marker Ki67 in the SVZ (fig. S7, H and I). Forty percent of OCT6 cells were positive for ISL1, and they were all negative for CTIP2, a marker of mature MSNs, showing that *OCT6* is a key gene of pre–D1-MSNs in the SVZ (Fig. 2, J to L). For pre-D2 MSNs, we found ~10% of OCT6 and SIX3–coexpressing cells in the SVZ (fig. S8, A to D). This suggests that the low OCT6 expression level characterizing this cell state translates into only a few cells detectable by immunohistochemistry. This low-to-high gradient is also observed with immunohistochemistry, where SIX3 cells were enriched dorsally and faded moving ventrally, whereas OCT6 showed a reverse ventral-to-dorsal enrichment (fig. S8A). These patterns of OCT6-ISL1 and OCT6-SIX3 coexpression in the SVZ were also confirmed at 11 pcw (fig. S8, E to H). To confirm the pre-D2 state, we looked at how ASCL1, SIX3, and CTIP2 behave in the SVZ (fig. S8I). We found that most of the SIX3-expressing cells in the SVZ lack expression of both ASCL1 and CTIP2 (fig. S8, J to L), confirming the presence of the pre-D2 state in the SVZ.

Our findings support a model in which D1- and D2-MSNs derive from a common progenitor; cells become postmitotic as they pass through a primed precursor phase that exhibits heterogeneity in the expression of key MSN regulators, and specific subtypes of MSNs are gradually established. If this model is correct, fate determination in MSNs does not occur as an instantaneous shift in cell fate, but rather is a smooth and continuous process that falls within two gradients of expression.

Gene regulatory networks of the MSN lineage

The human LGE domain has been the most elusive area of the basal ganglia to characterize. To bridge this gap, we applied SCENIC (*14*) to our scRNA-seq data and reconstructed the gene regulatory networks and combinatorial codes of TFs that define the different cell states in this domain (Fig. 3A and table S9). Overall, gene regulatory network analysis revealed that distinct regulons classify different axes of MSN development, with D1- and D2-MSNs sharing most of the active TFs (Fig. 3B). We found a number of TFs that were not previously associated with LGE development. These include OTX1 for APs, VAX1 for BPs, POU2F2 for both types of MSNs, NANOG in D2-MSNs, and FOXO1 for D1-MSNs.

We then used SCENIC coexpression networks to infer potential relationships between our de novo identified lincRNAs and our cell type–specific TFs (fig. S9A and table S10), because it has been shown that lincRNA loci contain many conserved TF binding sites (*12*), hence their relationship may be functionally relevant. We found that *hfb_G_000907*, a specific lincRNA of APs, is strongly linked to *SOX* genes, and that *hfg_G_000296*, another lincRNA specific for APs, shows a high connection with a PDZ-LIM domain family protein called PDLIM5. The BP lincRNA *hfb_G_00882* shows a correlation with *RARA*, an important mediator of retinoid signaling. For MSNs, we found common lincRNAs between D1- and D2-MSN subtypes that are in the same network with universal MSN signatures such as *ZNF467* and *POU2F2*. *hfb_G_000494*, a D1-MSN–specific

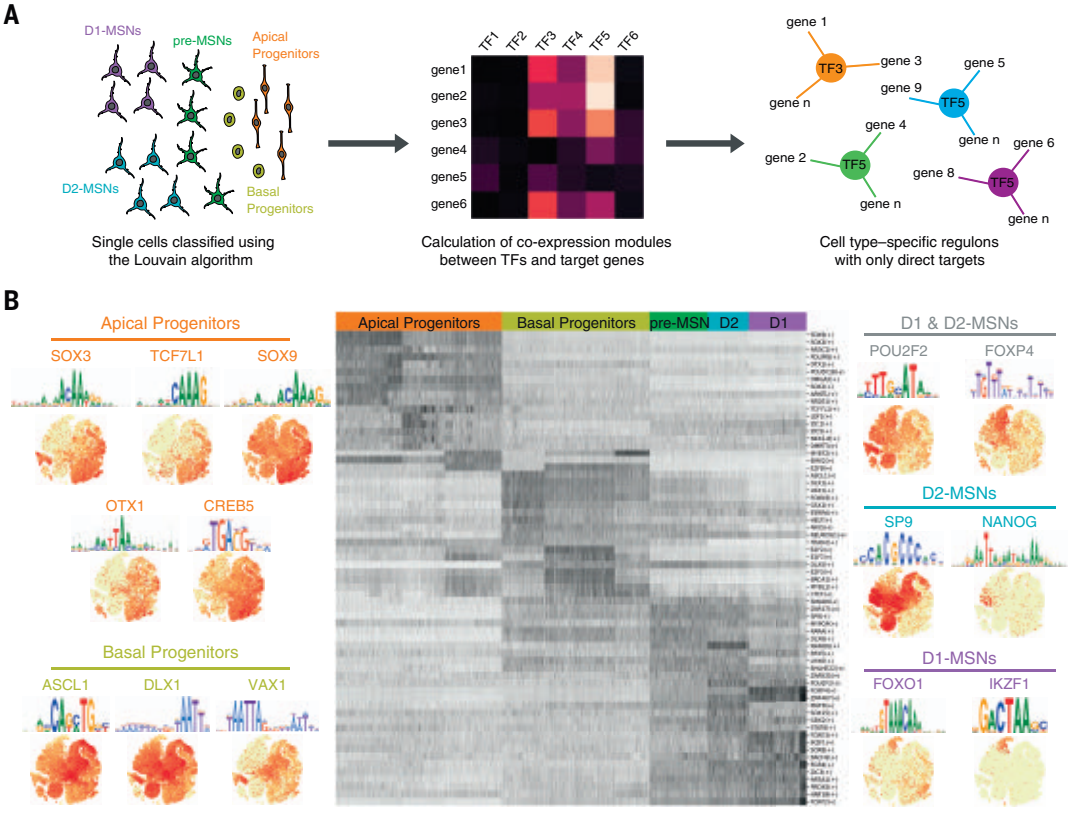


Fig. 3. The cell-specific gene regulatory networks of the LGE lineage. (**A**) Schematic overview of the SCENIC computational approach used to infer cell type–specific gene regulatory networks. (**B**) SCENIC regulon activity matrix showing the top active transcription factors in each cell class. For each cell type, specific transcription factors and their associated motif and expression patterns are shown in t-SNE plots.

lincRNA, shows a connection with *SOX8* and *IKZF1*, two players of the D1-MSN class, whereas many of the D2-MSN–specific lincRNAs are linked to *SP9* and *MAFB*. Overall, these observations reveal insights into how this panel of lincRNAs can be regulated.

LGE and MGE progenitors are transcriptionally distinct

Given the limited information available on APs of the LGE, we then focused our attention on this particular progenitor domain. We found that SOX3 and TCF7L1 (both inhibitors of posteriorizing WNT signals) together with SOX9 (which is induced by SHH signaling) are active TFs that may ventralize and maintain this early progenitor phase (Fig. 4A). Other TFs of this progenitor domain include OTX1, which has been predominantly associated with the CX, and CREB5, which has a role in neural progenitor differentiation. These regulators are interlinked by shared target genes and are likely important for the control of cell fate decisions within the AP domain. The *TCF7L1* network does not share target genes with the other TFs; however, it does regulate genes such as *NKX2.1* and *OTX2* that play a role in maintaining a regional ventral identity (Fig. 4A).

We then investigated whether transcriptional diversity exists at early developmental stages between APs of the LGE and MGE. We combined our set of LGE-specific marker genes (that are not expressed in the CX and MGE) identified in the bulk dataset with the signatures of the different cell states identified in the scRNA-seq dataset (Fig. 4B). We identified 199 genes that are LGE- and cell state–specific (Fig. 4B and table S11); within these, 31 were specific for APs of the LGE, and for two of these candidates—the TF gene *SALL3* and its topologically adjacent *LINC01896* (Fig. 4, C and D)—we confirmed by FISH that they are expressed in APs of the LGE (and CGE) and not in the MGE (Fig. 4E and fig. S10, A to D). Our results propose a model where the neuronal fate of APs in the VZ of the developing striatum is already restricted by a specific topographical and transcriptional program that subsequently drives LGE-specific (versus MGE-specific) cell fates.

Gene network interference reveals key MSN fate determinants

Within the striatum, D1- and D2-MSNs contribute to the direct and indirect pathways, respectively (*15*). We found that both types of MSNs are shaped by a shared set of gene regulatory networks governed by specific MSN "master" TFs (Fig. 5A and fig. S11, A and B). To validate these observations and predict their role in cell lineage determination, we used CellOracle (*16*) to perturb, in silico, these key TFs and their gene regulatory networks. We tested CellOracle on *SP9*, a functionally validated gene in MSN fate determination (*17*). We observed that *SP9* knockout causes a block in BP differentiation, together with a specific arrest in D2-MSN generation, resulting in a shift toward D1-MSN fates (fig. S11C). When we tested *SP9* overexpression, the opposite trend was observed, with a D1- to D2-MSN shift (fig. S11C). This recapitulates what is known about SP9 during striatal development (*17*) and confirms the ability of CellOracle to identify functional TFs. We then tested ZNF467, a zinc finger protein whose function has not yet been characterized in any tissue. We found that its knockout halts MSN differentiation at the progenitor stage, whereas it promotes MSN specification when overexpressed (Fig. 5B). This suggests that ZNF467 may be an important TF for the entire MSN lineage. We then tested *OCT6* and found that knockout of this gene causes an inhibition of the pre-MSN phase, with cells reverting back to progenitor domains and also a stimulation of the transition from pre-MSN to mature MSN states (Fig. 5C). This is in line with the transient nature of *OCT6* expression and may reflect the result of removing this gene at different stages of differentiation. Overexpressing *OCT6* has the opposite effect, as it stalls differentiation in the pre-MSN phase (Fig. 5C). These data confirm the probable function of this gene in the pre-MSN phase, especially in pre-D1 MSNs because we detect OCT6 protein translation mainly in the D1 lineage (Fig. 2, I to L).

We then tested key TFs that may drive the D1-D2 MSN bifurcation point. We found that

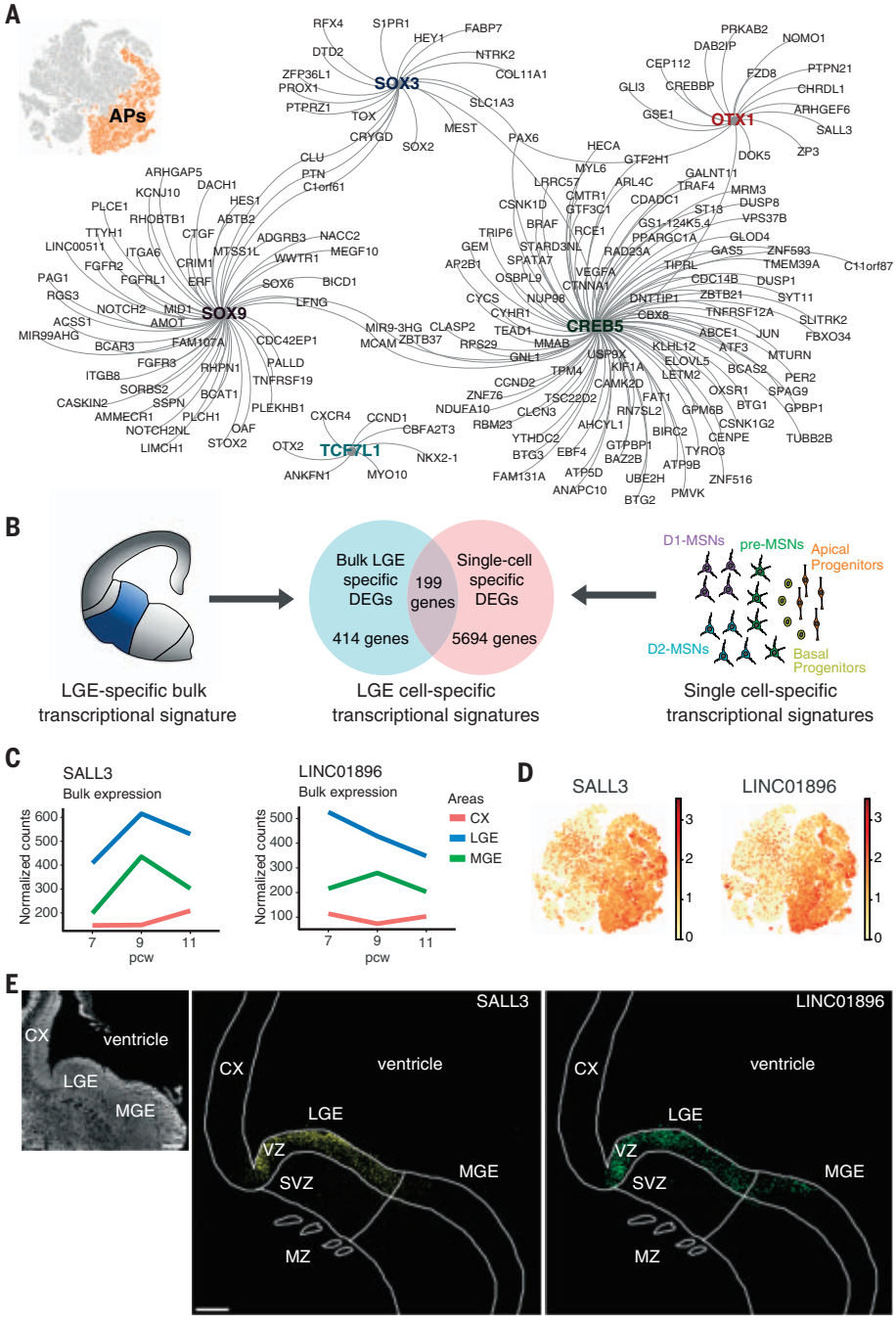


Fig. 4. The emergence of specific apical progenitors within the LGE. (**A**) Specific active regulons in APs. (**B**) Schematic overview of the computational approach used to infer LGE and cell type–specific transcriptional signatures from bulk and scRNA-seq data. DEGs, differentially expressed genes. (**C**) Bulk expression levels of the LGE-specific genes *SALL3* and *LINC01896*. (**D**) Single-cell expression levels (t-SNE plots) of the AP-specific genes *SALL3* and *LINC01896* (*RP11-849I19.1*) in the LGE lineage. (**E**) FISH validation of *SALL3* and *LINC01896* on a telencephalic coronal hemisection at 9 pcw. Upper leftmost panel, DAPI signal. Scale bar, 500 μm. Right panels, probe signals. Scale bar, 200 μm.

IKZF1, an essential gene in the generation of D2-MSNs in mice (*18*), is instead associated to D1-MSNs in humans and therefore may have an opposing role during human development. Perturbation of this gene caused a specific obstruction in D1-MSN maturation, whereas its overexpression led to increased D1 production (Fig. 5D). Finally, we tested MAFB, a specific TF of the D2 lineage that has previously been associated with the survival of MGE-derived cortical interneurons (*19*). Knockout of *MAFB* causes an arrest in BP differentiation and an interruption in D2 production that shifts to the D1 lineage (Fig. 5E), mimicking what is seen with *SP9* knockout. Instead, overexpression of *MAFB* triggers BP maturation and conversion from D1- to D2-MSNs as well as promoting conversion from pre-MSNs to interneurons (Fig. 5E), supporting the role of this gene in promoting this cell class.

Because most studies in humans have generated a general MSN-specific or a D1-MSN–specific signature (*20*), we next decided to characterize a highly specific D2-MSN signature. From our combined bulk and scRNA-seq data (Fig. 4B), we defined 13 specific LGE and D2-MSN genes (table S11). Among them, we validated *LINC01305* and *KCNA5* (fig. S12A), which are specific for the mantle zone (MZ) of the LGE and are not found in the MGE or CGE (fig. S12, B to E).

Finally, to reveal different populations within each major class of MSNs, we performed subclustering of the two classes of MSNs. This revealed two subclusters in both D1- and D2-MSNs (Fig. 5F), and we tested whether these subclusters were already specified in the patch compartment, one of the two major areas (the other being the matrix) that characterize the organization of the striatum, which is formed during early development in mouse models (*21*). A previous single-cell study on postnatal mouse striatal cells (*22*) showed that *Pdyn* and *Tshz1* are specific patch markers. Here, we found that one of the clusters in the D1 population expresses *PDYN* and low levels of *TSHZ1*, whereas in D2-MSNs the opposite trend is observed (Fig. 5F). This suggests that in humans these two markers define MSN lineage–specified patch cells. *PDYN* expression in the MZ confirms that this organization is present at 9 pcw (Fig. 5G). Overall, our data suggest that human MSNs are already compartmentalized in subtype-specific patch regions at this stage of early development.

Human-specific lincRNAs show striatal specificity

We then decided to investigate how lincRNAs define striatal evolution and development. Using liftOver (*23*) and TransMap (*24*), we identified lincRNAs that map across different species, have conserved sequence identity, and are also expressed in the other species considered (Fig. 6A).

Conservation scores showed that lincRNAs identified in the brain exhibit a higher score in primates than in mice, which suggests that they are less conserved in lower mammals (Fig. 6B). However, in all the species considered, conservation scores increased for lincRNAs expressed in the adult brain relative to the developing brain (Fig. 6B). This suggests that the transcriptional complexity of the primate brain, in terms of lincRNAs, is mainly established during fetal development and that in the

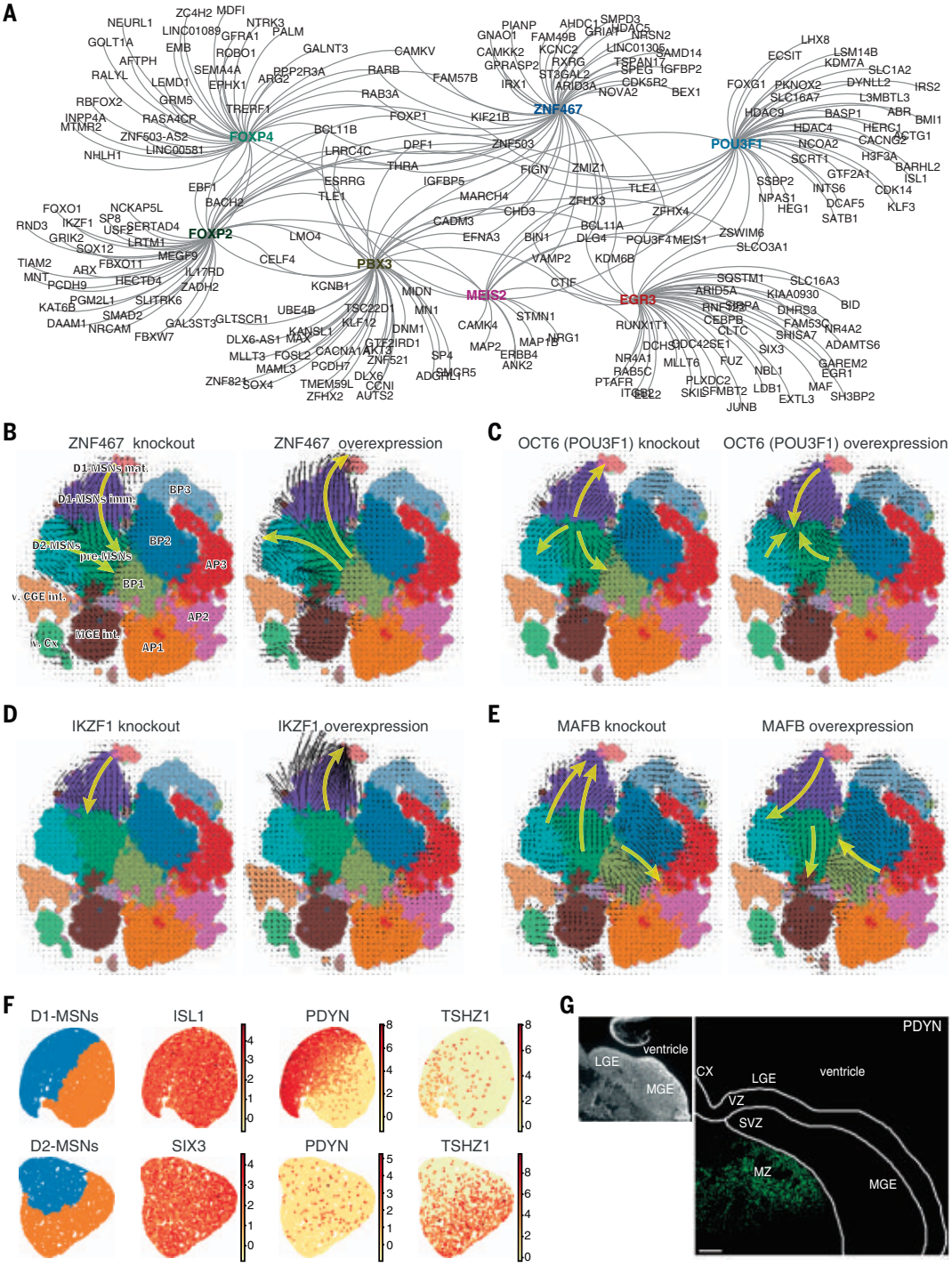


Fig. 5. In silico perturbation of MSN-specific gene regulatory networks. (**A**) Shared active regulons that define D1- and D2-MSNs. (**B** to **E**) CellOracle simulation of knockout and overexpression of the two key MSN markers *ZNF467* (B) and *OCT6* (C); *IKZF1*, a D1-MSN–specific transcription factor (D); and *MAFB*, a D2-MSN–specific transcription factor (E). The effect of the perturbation is shown with projections of cell state transition vectors on each cell's t-SNE plot. Yellow arrows were manually added to represent overall directionality. (**F**) Subclusters of D1- and D2-MSNs, together with expression levels of two specific patch markers (*PDYN* and *TSHZ1*) and canonical D1 (*ISL1*) and D2 (*SIX3*) MSN markers. (**G**) FISH validation of *PDYN* on a telencephalic coronal hemisection shows exclusive expression in the MZ at 9 pcw. Left panel, DAPI signal. Scale bar, 500 μm. Right panel, probe signal. Scale bar, 200 μm.

adult brain this biotype probably possesses more conserved functions among different species.

We further explored two human-specific lincRNAs that were not conserved in the other species considered and that were explicitly expressed in MSNs of the developing human LGE (Fig. 6, C and D). We initially used a guilt-by-association approach (*25*) to investigate the potential function of protein-coding genes that highly correlate with this set of lincRNAs, in order to predict the latter's potential role in MSN specification. We inferred that the MSN-specific lincRNA *hfb_G_000053* could be involved in brain development and neuronal differentiation (Fig. 6E and table S12) and that the mature D1-MSN–specific lincRNA *hfb_G_000494* could function in synaptic organization (Fig. 6F and table S12). When we validated these unannotated lincRNAs by reverse transcription polymerase chain reaction and FISH, we found a specific enrichment of these lincRNAs in the MZ of the developing human striatum relative to the MGE and CGE (Fig. 6, G and H, and fig. S13, A to E).

These findings suggest that these two lincRNAs have a role in the development of human MSNs and may provide insight into how the striatal architecture has evolved to accommodate the expanded human behavioral repertoire.

Discussion

Recent studies of the developing human fetal brain have been crucial in efforts to decipher

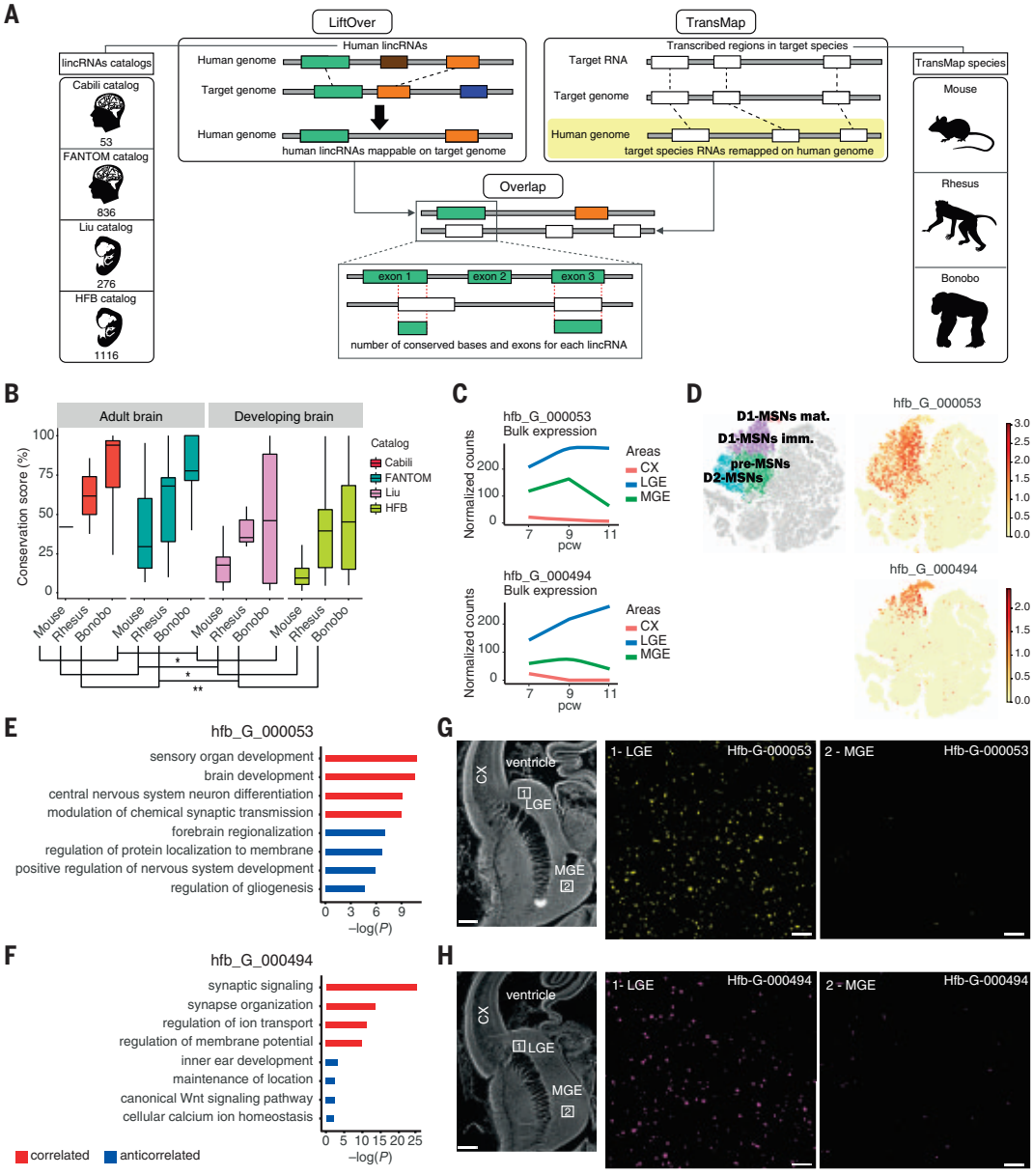


Fig. 6. The human-specific long noncoding signature of MSNs. (**A**) Computational pipeline developed to establish the conservation score of each lincRNA. (**B**) Boxplot showing distribution of conservation scores for different lincRNA classes during development and in adult tissues shows significant differences between lincRNAs identified in the developing brain (Liu, HFB) and those identified in the adult brain (Cabili, FANTOM): *$P < 0.05$, **$P < 0.002$ (Wilcoxon rank-sum test). No differences were observed between the two adult catalogs (Cabili versus FANTOM) or between the two catalogs of the developing brain (Liu versus HFB). (**C**) Bulk expression levels of the human-specific lincRNAs *hfb_G_000053* and *hfb_G_000494*. (**D**) Single-cell expression levels of *hfb_G_000053* and *hfb_G_000494*. (**E** and **F**) Enriched Gene Ontology terms related to genes that have a high correlation or anticorrelation pattern ($P < 0.05$ and $r > 0.6$ or $r < -0.6$) with *hfb_G_000053* (E) and *hfb_G_000494* (F). (**G** and **H**) FISH validation of *hfb_G_000053* (G) and *hfb_G_000494* (H) expression on a telencephalic coronal hemisection at 11 pcw. Leftmost panel, DAPI signal. Scale bar, 500 μm. Right panel, probe signals. Scale bar, 100 μm.

the basis of human brain formation, function, and evolution (*26*–*29*). In this study, we created a high-resolution single-cell map of the early development of the human striatum from a coding and long noncoding perspective. This atlas will pave the way for future integrations with chromatin accessibility data and single-cell proteomics data that will help to resolve the relationship among gene regulation, transcription, and protein production during lineage determination in MSNs, especially at the pre-MSN phase identified in this study.

The main limitation of this work is the high degree of sparsity of the single-cell data. Future protocols with increased sensitivity of RNA detection may reveal new and rare cell types and will probably enable a clearer separation of the pre-MSN phase. This study was also limited by the restricted time window (early fetal development) and brain region (LGE) examined. Studies of both earlier and later developmental time points combined with single-cell measurements of the CX, MGE, and CGE will enhance our understanding of lineage establishment and diversification.

Nonetheless, this study has set the foundation for understanding human striatal development at unprecedented granularity. We foresee that key TFs and lincRNAs defined in this study will be leveraged in vitro to generate authentic MSNs that can serve as suitable donor preparations for future clinical trials in cell replacement therapies. This dataset will also be critical to understand the neurodevelopmental component of Huntington's disease (HD) during striatal development, given the recent evidence of alterations in normal cortical development in human HD fetal samples (*30*). Finally, we also expect that the newly identified human-specific lincRNAs will enable us to understand the underpinnings that characterize human striatum function.

REFERENCES AND NOTES

1. M. A. Raghanti *et al.*, Human-specific increase of dopaminergic innervation in a striatal region associated with speech and language: A comparative analysis of the primate basal ganglia. *J. Comp. Neurol.* **524**, 2117–2129 (2016). doi: 10.1002/cne.23937; pmid: 26715195
2. C. R. Gerfen, The neostriatal mosaic: Multiple levels of compartmental organization. *Trends Neurosci.* **15**, 133–139 (1992). doi: 10.1016/0166-2236(92)90355-C; pmid: 1374971
3. S. Grillner, B. Robertson, M. Stephenson-Jones, The evolutionary origin of the vertebrate basal ganglia and its role in action selection. *J. Physiol.* **591**, 5425–5431 (2013). doi: 10.1113/jphysiol.2012.246660; pmid: 23318875

4. J. Stiles, T. L. Jernigan, The basics of brain development. *Neuropsychol. Rev.* **20**, 327–348 (2010). doi: 10.1007/s11065-010-9148-4; pmid: 21042938
5. M. E. Dinger, K. C. Pang, T. R. Mercer, J. S. Mattick, Differentiating protein-coding and noncoding RNA: Challenges and ambiguities. *PLOS Comput. Biol.* **4**, e1000176 (2008). doi: 10.1371/journal.pcbi.1000176; pmid: 19043537
6. S. Djebali *et al*., Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012). doi: 10.1038/nature11233; pmid: 22955620
7. M. N. Cabili *et al*., Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927 (2011). doi: 10.1101/gad.17446611; pmid: 21890647
8. C. C. Hon *et al*., An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **543**, 199–204 (2017). doi: 10.1038/nature21374; pmid: 28241135
9. S. J. Liu *et al*., Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* **17**, 67 (2016). doi: 10.1186/s13059-016-0932-1; pmid: 27081004
10. G. La Manno *et al*., RNA velocity of single cells. *Nature* **560**, 494–498 (2018). doi: 10.1038/s41586-018-0414-6; pmid: 30089906
11. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020). doi: 10.1038/s41587-020-0591-3; pmid: 32747759
12. M. Melé *et al*., Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. *Genome Res.* **27**, 27–37 (2017). doi: 10.1101/gr.214205.116; pmid: 27927715
13. J. Q. Wu *et al*., Dynamic transcriptomes during neural differentiation of human embryonic stem cells revealed by short, long, and paired-end sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5254–5259 (2010). doi: 10.1073/pnas.0914114107; pmid: 20194744
14. S. Aibar *et al*., SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017). doi: 10.1038/nmeth.4463; pmid: 28991892
15. D. J. Surmeier, W. J. Song, Z. Yan, Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J. Neurosci.* **16**, 6579–6591 (1996). doi: 10.1523/JNEUROSCI.16-20-06579.1996; pmid: 8815934
16. K. Kamimoto, C. M. Hoffmann, S. A. Morris, CellOracle: Dissecting cell identity via network inference and in silico gene perturbation. bioRxiv [preprint]. 17 February 2020. pmid: 947416
17. Q. Zhang *et al*., The Zinc Finger Transcription Factor Sp9 Is Required for the Development of Striatopallidal Projection Neurons. *Cell Rep.* **16**, 1431–1444 (2016). doi: 10.1016/j.celrep.2016.06.090; pmid: 27452460
18. R. Martín-Ibáñez *et al*., Ikaros-1 couples cell cycle arrest of late striatal precursors with neurogenesis of enkephalinergic neurons. *J. Comp. Neurol.* **518**, 329–351 (2010). doi: 10.1002/cne.22215; pmid: 19950118
19. E. L. L. Pai *et al*., *Maf* and *Mafb* control mouse pallial interneuron fate and maturation through neuropsychiatric disease gene regulation. *eLife* **9**, e54903 (2020). doi: 10.7554/eLife.54903; pmid: 32452758
20. M. Onorati *et al*., Molecular and functional definition of the developing human striatum. *Nat. Neurosci.* **17**, 1804–1815 (2014). doi: 10.1038/nn.3860; pmid: 25383901
21. S. M. Kelly *et al*., Radial Glial Lineage Progression and Differential Intermediate Progenitor Amplification Underlie Striatal Compartments and Circuit Organization. *Neuron* **99**, 345–361.e4 (2018). doi: 10.1016/j.neuron.2018.06.021; pmid: 30017396
22. A. Saunders *et al*., Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* **174**, 1015–1030.e16 (2018). doi: 10.1016/j.cell.2018.07.028; pmid: 30096299
23. R. M. Kuhn, D. Haussler, W. J. Kent, The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013). doi: 10.1093/bib/bbs038; pmid: 22908213
24. J. Zhu *et al*., Comparative genomics search for losses of long-established genes on the human lineage. *PLOS Comput. Biol.* **3**, e247 (2007). doi: 10.1371/journal.pcbi.0030247; pmid: 18085818
25. M. Guttman *et al*., Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009). doi: 10.1038/nature07672; pmid: 19182780
26. T. J. Nowakowski *et al*., Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017). doi: 10.1093/bib/bbs038; pmid: 22908213
27. G. La Manno *et al*., Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. *Cell* **167**, 566–580.e19 (2016). doi: 10.1016/j.cell.2016.09.027; pmid: 27716510
28. X. Fan *et al*., Single-cell transcriptome analysis reveals cell lineage specification in temporal-spatial patterns in human cortical development. *Sci. Adv.* **6**, eaaz2978 (2020). doi: 10.1126/sciadv.aaz2978; pmid: 32923614
29. S. Zhong *et al*., A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018). doi: 10.1038/nature25980; pmid: 29539641
30. M. Barnat *et al*., Huntington's disease alters human neurodevelopment. *Science* **369**, 787–793 (2020). doi: 10.1093/bib/bbs038; pmid: 22908213

ACKNOWLEDGMENTS

We thank C. Cordiglieri and A. Fasciani of the INGM Imaging Facility (Istituto Nazionale Genetica Molecolare–INGM, Milan, Italy) for scientific and technical assistance, and S. Noble of the CHDI Foundation for help with manuscript editing. **Funding:** This work was conducted with collaboration and funding from CHDI Foundation (JSC A11103), a nonprofit biomedical research organization exclusively dedicated to developing therapeutics that will substantially improve the life of individuals affected by HD, to E.C. This study was also supported by the European Union–funded projects NeurostemcellRepair (FP7, GA no. 602278, 2013-17) and NSC-Reconstruct (H2020, GA no. 874758, 2020-23) to E.C. and R.A.B. The fetal tissue collection in Cambridge is supported by NIHR funding of the Biomedical Research Centre and by the Wellcome Trust 203151/Z/16/Z. **Author contributions:** V.D.B. and E.C. conceived the study; V.D.B., P.C., E.V., D.B., and E.C. designed the experiments; V.D.B. performed computational analysis with contributions from C.C., T.L., V.R., R.J.P.B., F.G., and M.P.; C.C., V.D.B., and G.P. performed the conservation analysis study; P.C. performed the bulk and single-cell experiments with contributions from D.B., I.E.-C., A.F., M.D.S., and G.R.; E.V. performed FISH analysis; M.G. performed immunohistochemistry analysis; V.R., R.J.P.B., F.G., M.D.S., G.R., and M.P. contributed the INGM single-cell and bioinformatics platforms; X.H., R.V., and R.A.B. collected the fetal samples and performed the dissection; K.K. and S.A.M. performed in silico perturbation analysis with V.D.B. and R.J.P.B.; J.C., D.F., M.P., and E.C. contributed to data elaboration and presentation. All authors contributed to interpretation of the results. Figures were assembled by V.D.B. together with T.L. The manuscript was written by V.D.B., revised by D.B. and E.C., and edited and proofread by all authors. E.C. proposed the research program, secured the funding, established the collaborations, and coordinated the study. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All bulk and scRNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress/) under accession no. E-MTAB-8893 (bulk RNA-seq) and E-MTAB-8894 (scRNA-seq). All other data are present in the main paper or the supplement.

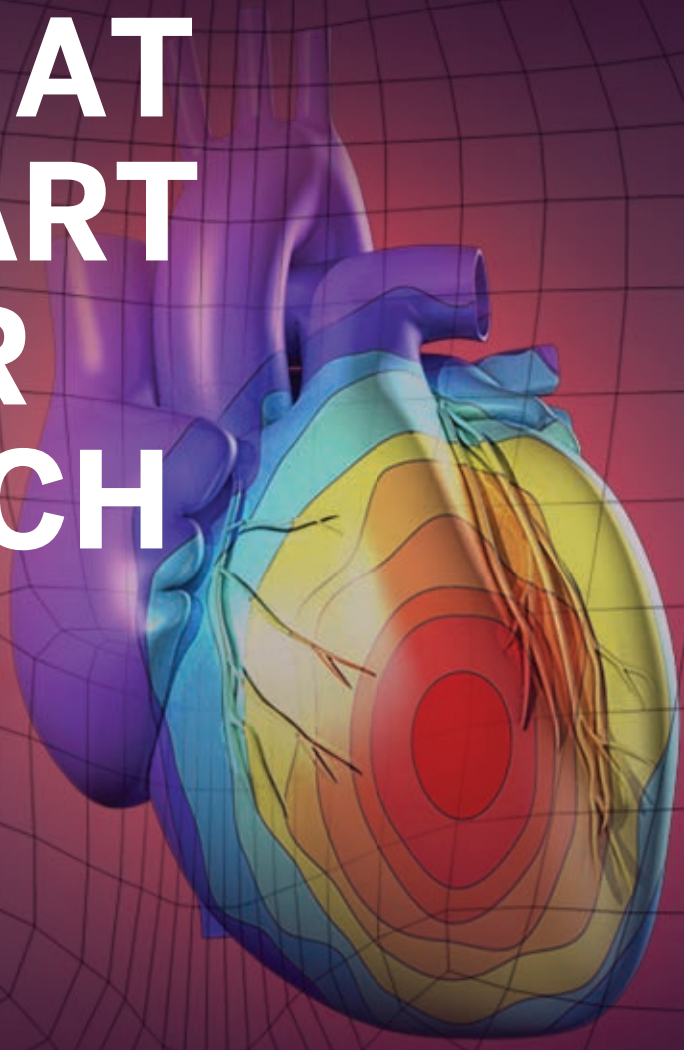
SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/eabf5759/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S12
References (*31*–*48*)

4 November 2020; accepted 29 March 2021
10.1126/science.abf5759

ScienceTranslationalMedicine.org

PUT HUMAN HEALTH AT THE HEART OF YOUR RESEARCH



Submit your research:
cts.ScienceMag.org

Science Translational Medicine
AAAS

Twitter: @ScienceTM
Facebook: @ScienceTranslationalMedicine

RESEARCH ARTICLES

SEX DETERMINATION

Sex chromosome transformation and the origin of a male-specific X chromosome in the creeping vole

Matthew B. Couger[1]†, Scott W. Roy[2,3]†, Noelle Anderson[3], Landen Gozashti[4], Stacy Pirro[5], Lindsay S. Millward[6], Michelle Kim[7], Duncan Kilburn[7], Kelvin J. Liu[7], Todd M. Wilson[8], Clinton W. Epps[6], Laurie Dizney[9], Luis A. Ruedas[10], Polly Campbell[11]*

The mammalian sex chromosome system (XX female/XY male) is ancient and highly conserved. The sex chromosome karyotype of the creeping vole (*Microtus oregoni*) represents a long-standing anomaly, with an X chromosome that is unpaired in females (X0) and exclusively maternally transmitted. We produced a highly contiguous male genome assembly, together with short-read genomes and transcriptomes for both sexes. We show that *M. oregoni* has lost an independently segregating Y chromosome and that the male-specific sex chromosome is a second X chromosome that is largely homologous to the maternally transmitted X. Both maternally inherited and male-specific sex chromosomes carry fragments of the ancestral Y chromosome. Consequences of this recently transformed sex chromosome system include Y-like degeneration and gene amplification on the male-specific X, expression of ancestral Y-linked genes in females, and X inactivation of the male-specific chromosome in male somatic cells. The genome of *M. oregoni* elucidates the processes that shape the gene content and dosage of mammalian sex chromosomes and exemplifies a rare case of plasticity in an ancient sex chromosome system.

Sex chromosomes have arisen many times in evolution. A new pair of sex chromosomes generally originates when an autosomal locus acquires a sex-determining function (*1*). What happens next differs substantially across taxa. At one extreme, turnover of sex chromosomes can be rapid and even polymorphic within species [e.g., frogs (*2*, *3*) and fishes (*4*, *5*)]; at the other, the sex chromosome systems of taxa such as mammals, birds, and *Drosophila* show marked evolutionary stability (*6*).

More than 150 million years of evolution have shaped the identity and dosage of genes on mammalian sex chromosomes since they evolved from an autosomal pair in the ancestor of eutherian (placental) and metatherian (marsupial) mammals (*7*–*9*). Yet the defining features of the X and Y chromosomes arose very early in their evolution and are highly conserved across extant species. On the eutherian Y chromosome, fewer than 5% of genes survived the massive degeneration that followed the progressive suppression of recombination with its once homologous partner, the X chromosome (*10*, *11*). Between 9 and 16 of these ancient genes persist in any single lineage. Of the 10 such genes present in rodents, 5 (*Sry*, *Ddx3y*, *Usp9y*, *Uty*, and *Zfy*) are common to all eutherian species with sequenced Y chromosomes (*8*, *11*, *12*). Surviving Y-linked genes are not specifically expressed in the testes, nor are their functions exclusive to spermatogenesis. Instead, their expression profile is broad, and functions include fundamental cellular processes such as transcription, translation, and chromatin modification (*8*, *11*).

In contrast to the gene-depleted Y chromosome, more than 90% of ancestral genes survive on mammalian X chromosomes (*13*, *14*). Most of these genes are subject to dosage compensation by female X inactivation, the transcriptional silencing of one X chromosome in female somatic cells that balances the dosage of most X-linked genes between XY males and XX females. Although the initial driver of the evolution of X inactivation remains controversial (*15*), it is clear that silencing of X-linked genes evolved in parallel with early gene loss on the Y (*16*). Because a critical minority of X-linked genes escape silencing, two X chromosomes are essential for female-typical development, whereas more than one X chromosome leads to anomalous development in males (*17*, *18*). In this study, we show that these ancient and highly conserved properties of the X and Y chromosomes can be lost, swapped between chromosomes, and even reversed between the sexes, all on a short evolutionary time scale.

The creeping vole (*Microtus oregoni*) is one of very few mammals with an atypical sex chromosome system (*19*, *20*). Fifty years after Susumo Ohno described the cytological details of this system (*21*), it remains an unsolved puzzle in sex chromosome evolution. Figure 1 depicts the notable features of the system, highlighting differences between standard XY systems (Fig. 1A) and the *M. oregoni* system as defined by Ohno (Fig. 1B). Like other therian mammals, *M. oregoni* has two sex chromosomes, termed X and Y by Ohno (though shown here to have more complex chimeric origins). However, the number of sex chromosomes is different between diploid germline and somatic cells (*21*). Even more unusual, the distribution of sex chromosomes across cell types is reversed between females and males. Adopting Ohno's designations, females have paired sex chromosomes in the germ line ($2n$ = 18, XX) but not in the soma ($2n$ = 17, X0), whereas males have paired sex chromosomes in somatic cells ($2n$ = 18, XY) but a single unpaired chromosome in the germ line ($2n$ = 17, Y0), the latter of which leads to production of gametes bearing either a Y chromosome or no sex chromosome (*21*, *22*). Ohno demonstrated that the two X chromosomes in female meiosis are one and the same, the products of mitotic nondisjunction after DNA replication in the germ line (*22*, *23*). Because the X chromosome was never observed in the male germ line, Ohno inferred that a similar process of X nondisjunction produced XXY and Y0 germ cells, with the former rapidly eliminated by apoptosis (*21*, *22*). The reason the latter (which would be expected to lack essential X-linked genes) would preferentially survive remained unknown.

Ohno's observations imply that the *M. oregoni* X chromosome is effectively nonrecombining, is hemizygous in the somatic cells of both sexes, and can respond only to selection on allelic transmission in females. Moreover, these pronounced changes in X chromosome biology are evolutionarily recent: Whereas the diversification of *Microtus* [~70 species in 2 million years (*24*)] is characterized by dynamic karyotype evolution, all other species are XX/XY (*25*). The value of *M. oregoni* as a rare natural experiment in mammalian sex chromosome evolution has received surprisingly little attention (*26*), and the genetics of this system are completely unstudied. Here we present a chromosome-scale assembly of the *M. oregoni* genome. Analysis of this genome shows a number of unusual features (schematically summarized in Fig. 1C), demonstrating that the distribution of sex chromosomes and sex-linked genes between males and females is even more surprising than Ohno supposed.

Genome reorganization and Y-derived genes in females

We first used read cloud genome assembly (*27*) with Hi-C chromosome scaffolding (*28*, *29*) to sequence and assemble a male *M. oregoni*

[1]Department of Thoracic Surgery, Brigham and Women's Hospital, Boston MA, 02115, USA. [2]Department of Biology, San Francisco State University, San Francisco, CA 94117, USA. [3]Department of Molecular and Cell Biology, University of California, Merced, Merced, CA 95343, USA. [4]Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA. [5]Iridian Genomes, Inc., Bethesda, MD 20817, USA. [6]Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR 97330, USA. [7]Circulomics Inc., Baltimore, MD 21202, USA. [8]US Forest Service, PNW Research Station, Corvallis, OR 97331, USA. [9]Department of Biology, University of Portland, Portland, OR 97203, USA. [10]Department of Biology and Museum of Natural History, Portland State University, Portland, OR 97207, USA. [11]Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, CA 92521, USA.
†These authors contributed equally to this work.
*Corresponding author. Email: polly.campbell@ucr.edu

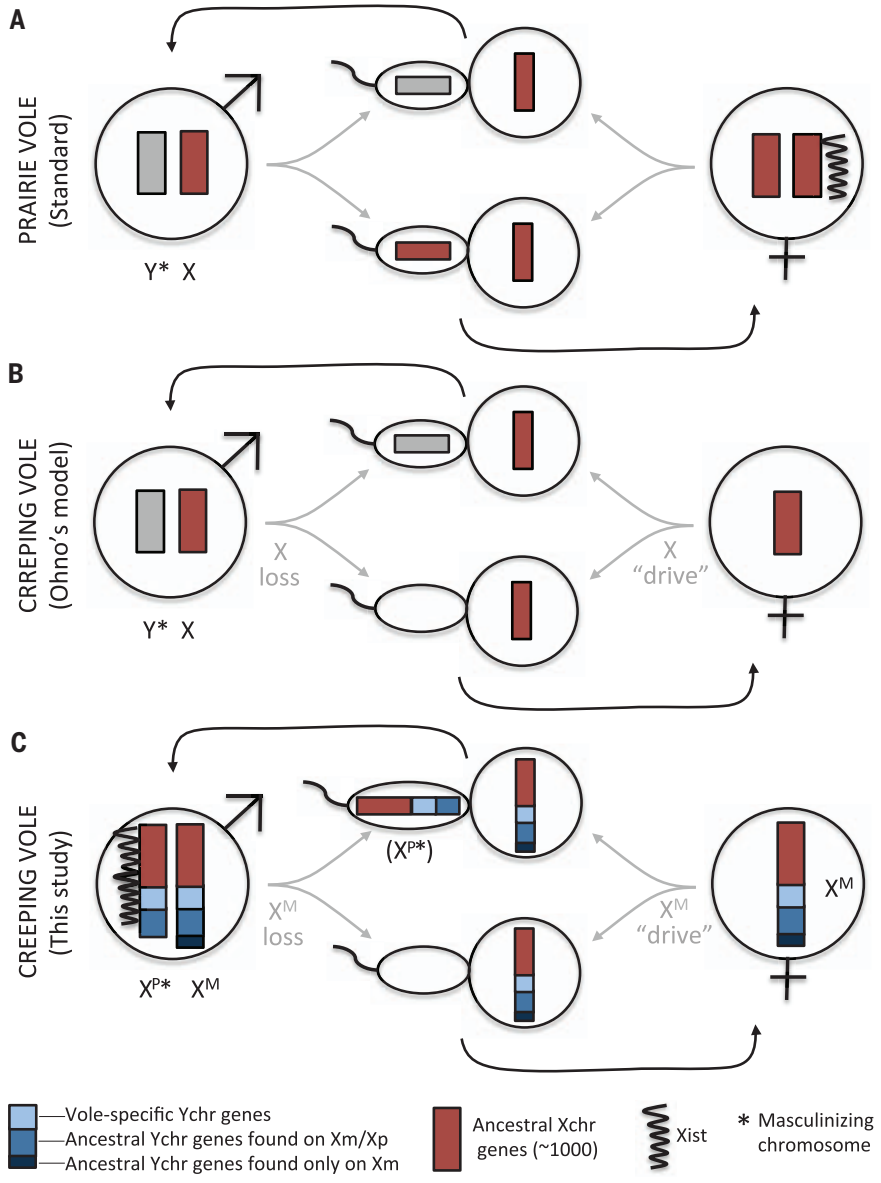


Fig. 1. Atypical chromosome structures and cycles of the creeping vole (*Microtus oregoni*). (**A**) Standard XY chromosome cycle as seen in the prairie vole (*M. ochrogaster*). (**B**) Atypical chromosome cycles of *M. oregoni*, showing chromosomal loss in spermatogenesis and chromosomal "drive" (transmission to 100% of oocytes). Ohno referred to male-specific and sex-shared sex chromosomes as Y and X, respectively (*21*). (**C**) Summary of the present results, showing X-Y hybrid structures of the two sex chromosomes (which we call X^P and X^M), X^M-specific ancestral Y chromosomal genes, and silencing of the X^P by *Xist* in males. Ychr, Y chromosome; Xchr, X chromosome.

somatic genome. This assembly comprised 10 scaffolds, a number consistent with the diploid karyotype as defined by Ohno (*21*) ($2n$ = 18, XY), suggesting that most scaffolds represented complete chromosomes (tables S1 and S2). We produced high-quality gene models by using de novo assembled and genome-aligned RNA sequencing (RNA-seq) reads for multitissue male and female *M. oregoni* transcriptomes, identified genes with one-to-one homologs in the prairie vole (*Microtus ochrogaster*) female genome assembly (Mi_och 1.0), and compared their distribution across chromosomes. Nine of the scaffolds contained one-to-one gene homologs originating from multiple *M. ochrogaster* autosomes. Homology searches against the *M. ochrogaster*, mouse, and rat genomes unequivocally identified the remaining scaffold as homologous to the X chromosome.

Notably, the assembly did not contain a scaffold homologous to other mammalian Y chromosomes. Instead, eight ancestrally Y-linked genes—*Ddx3y*, *Eif2s3y*, *Kdm5d*, *Sry*, *Ube1y*, *Usp9y*, *Uty*, and *Zfy*—were placed on the X chromosome scaffold. Blast searches for additional Y-linked genes recovered *Rbmy*. Thus, of the 10 rodent ancestral Y-linked genes, only *Tspy1* was missing. *Tspy1* could not be recovered from related species [*Microtus agrestis* (field vole) and *Myodes glareolus* (bank vole)], suggesting ancestral loss. We used polymerase chain reaction (PCR) to confirm the presence of these Y-derived genes in additional male *M. oregoni* samples, including females as a control. Unexpectedly, all nine genes amplified in both sexes. By contrast, amplification was limited to males in the sister species *Microtus longicaudus* (long-tailed vole) (*30*) and in *M. ochrogaster* (Fig. 2). Thus, recent genome reorganization has exposed female *M. oregoni* to Y-derived genes that have been evolving independently in males for up to 150 million years.

Two X chromosomes in males

To better understand the relationship between sex and sex chromosome identity, we produced short-read assemblies (table S3) for somatic tissues from additional *M. oregoni* males and females (n = 2 animals per sex). On the basis of karyotypic studies (*21*, *22*), somatic cells for both males (XY) and females (X0) have a single X chromosome and should therefore both have an X-linked read depth that is half that of the autosomes. As expected, females exhibited an ~50% reduction in X-linked relative to autosome-linked read depth, comparable to that in the XY male *M. agrestis* genome (Fig. 3A). By contrast, X-linked read depth in *M. oregoni* males was indistinguishable from autosomal read depth and was similar to that found in the XX female *M. ochrogaster* genome (Fig. 3A). The presence of two X chromosome haplotypes in males was further supported by the observation of many heterozygous X-linked single-nucleotide polymorphisms (SNPs) in males but not in females (Table 1).

In most mammals, X chromosome expression is balanced between the sexes by the silencing of one X in female somatic cells. Initiation of this process relies on *Xist* noncoding RNA, which is highly expressed from the inactive X chromosome in XX females (*31*). Notably, we found a sex-reversed pattern in *M. oregoni*, with *Xist* detected in male but not female transcriptomes. We confirmed this result with reverse transcription (RT)–PCR in different individuals sampled from multiple populations (n = 10 per sex) (Fig. 3B). Inclusion of *M. longicaudus* confirmed that this reversal of sex-specific expression arose in the *M. oregoni* lineage (Fig. 3B).

The doubling of X chromosome haplotypes and the high frequency of X-linked heterozygous sites shared across males can be reconciled with the *M. oregoni* karyotype if the chromosome called the Y by Ohno is largely derived from a complete copy of the ancestral X chromosome. Male-limited *Xist* expression provides strong support for this inference. Hereafter, we refer to these X-derived chromosomes according



Ancestral Y-linked genes	*M. oregoni*		*M. longicaudus*		*M. ochrogaster*	
	♂	♀	♂	♀	♂	♀
Eif2s3y	✓	✓	✓	✗	✓	✗
Zfy	✓	✓	✓	✗	✓	✗
Usp9y	✓	✓	✓	✗	✓	✗
Ddx3y	✓	✓	✓	✗	✓	✗
Uty	✓	✓	✓	✗	✓	✗
Ube1y	✓	✓	✓	✗	✓	✗
Kdm5d	✓	✓	✓	✗	✓	✗
Rbmy	✓	✓	✓	✗	✓	✗
Sry	✓	✓	✓	✗	✓	✗

Fig. 2. Distribution of ancestral Y-linked genes between males and females in *M. oregoni* and closely related species. The nine genes were recovered from both male (n = 3) and female (n = 2) genomes in *M. oregoni*. PCR results showed that Y genes amplified in all male and female *M. oregoni* but only in male *M. longicaudus* (n = 4 per sex) and *M. ochrogaster* (n = 3 per sex).

to their parental origin: X^M is present in the somatic cells of both sexes but is maternally transmitted; X^P is male specific and is paternally transmitted. We designate the latter as a second X rather than a Y because it is recognized as such by the cellular machinery of male *M. oregoni*.

A haplotype-aware long-read assembly of the *M. oregoni* genome

We next sought to understand the longer-range structure of the X^P and X^M chromosomes. Previous karyotypic analysis has shown that X^M contains two large blocks of constitutive heterochromatin that are not apparent on X^P (*25*, *32*), resulting in a size difference between the two sex chromosomes that explains Ohno's designation of X^P as the Y chromosome (*21*). Because our original male genome was based on short reads, we could not accurately resolve this repetitive content. With the goal of resolving repetitive structures and accurately discriminating X^M and X^P alleles, we produced an additional male somatic genome by using high-coverage Hifi PacBio reads, ultralong Nanopore reads, and Bionano optical mapping. We produced PacBio circular consensus sequencing (CCS) libraries and assembled the genome by using a haplotype-aware assembler (*33*). To better resolve ancestral X- and Y-derived genes on X^M and X^P, we used a similar approach (CCS plus Bionano mapping) to produce a male *M. ochrogaster* genome.

Our *M. oregoni* long-read assembly was a marked improvement on our short-read–based assembly (tables S4 to S7). The autosomal primary contigs had an N50 statistic value of 42 Mb, whereas the sex chromosomal primary contigs had an N50 of 7.74 Mb. The more fragmented nature of X^M and X^P contigs was likely due to the large amount of repetitive content on both chromosomes (see below).

Separate assembly and comparison of candidate X^P and X^M haplotypes

To distinguish X^M and X^P contigs within the X-derived parts of the chromosomes, we turned to the SNPs that we had identified from the two male and two female short-read genomes. Sites where X^P and X^M differ are expected to lead to cases in which females show the same single SNP (X^M), whereas males are heterozygous for the same two variants (X^P and X^M). Indeed, 91.6% (2648 of 2892) of observed X-linked SNPs showed exactly this pattern (Table 1). For each contig, we identified sites containing either putative X^P or X^M variants. Different sites on the same contig showed clear correspondence, with 91.3% of the contigs (63 of 69) harboring multiple sites showing either all X^P or all X^M variants. This pattern supports the inference that these SNPs reflect allelic differences between X^P and X^M and allowed us to parse individual contigs between the two chromosomes. Additional candidate X^M and X^P contigs were then identified by searching for pairs of regions with multigene homology to the *M. ochrogaster* X chromosome. This yielded partial candidate assemblies of the X-derived portions of X^M and X^P, spanning a total of 127.5 and 82.5 Mb, respectively. Blast searches of the X^P and X^M haplotypes confirmed that most genes within this region were homologous to X-linked genes on the *M. ochrogaster* and mouse X chromosomes, indicating a dearth of gene movement to the X-derived parts of these chromosomes (Table 2). Figure 3C shows the aligned candidate X^M (yellow) and X^P (green) haplotypes. The colinearity between the two haplotypes is clearly visible from the many parallel arcs connecting homologous X^M and X^P sequences (dark lines).

Transposable element proliferation on X^P and X^M

Scrutiny of X chromosomal haplotypes revealed a significantly elevated density of transposable elements. We focused on the five most commonly found transposable elements, as judged by number of Blast hits. These included multiple L1_Mur2 family long interspersed nuclear elements as well as MYSERV6 elements. We identified putative transposase genes and assessed the fraction of each contig constituted by these genes. Relative to autosomes, putative transposase genes made up a larger percentage of X chromosomal sequences (4.1 and 5.0% of X^M and X^P sequences, respectively, compared with 1.3 to 2.0% for the autosomes; $P < 0.001$ by bootstrapping; Fig. 3D). Notably, elevated transposable element content was broadly distributed across X^P and X^M (outer histogram in Fig. 3C).

Complex divergent structures of the Y-derived regions on X^M and X^P

The presence of Y-derived genes in both sexes and the replacement of the ancestral Y chromosome by a second X-like chromosome in males could be explained by Y-to-X gene movement followed by Y chromosome loss. At least one other *Microtus* species carries pseudogenized copies of *Sry* on the X chromosome (*34*), and Y-to-X translocations preceded loss of the Y chromosome in two more distantly related vole species (*35*). In this case, we would expect to find Y-derived genes integrated into an otherwise X-derived chromosome. Alternatively, chromosomal breakage could have resulted in X-Y fusion, in which instance we would expect to detect a distinct region of Y-derived gene content. Given the high rate of karyotype evolution in *Microtus* and the substantially reduced diploid number in *M. oregoni* compared with close relatives (*25*), this scenario is also plausible.

We found some support for an X-Y fusion in both X^P and X^M sequences, albeit with distinct structural patterns and differences in gene order, identity, and copy number between the two chromosomes (Fig. 4). In particular, a probable break point was captured on a

6.5-Mb X^P scaffold, in which a repetitive array of nine ancestral X-linked genes was followed by an interval containing ancestral Y genes as well as genes found within an interval on the *M. ochrogaster* Y (Fig. 4A). The scaffold ended with >1.5 Mb of ampliconic content comprising 143 copies of an autosomal retrogene, *Rpl21*, with 3 copies of *Sry* embedded in the second half of the array. All *Rpl21* copies shared the same 8–base pair deletion in coding sequence and subsequent premature stop, indicating pseudogenization before amplification. X^M contigs with putative homology to this scaffold were highly fragmented and did not contain the ancestral Y genes or the *Rpl21* array (fig.

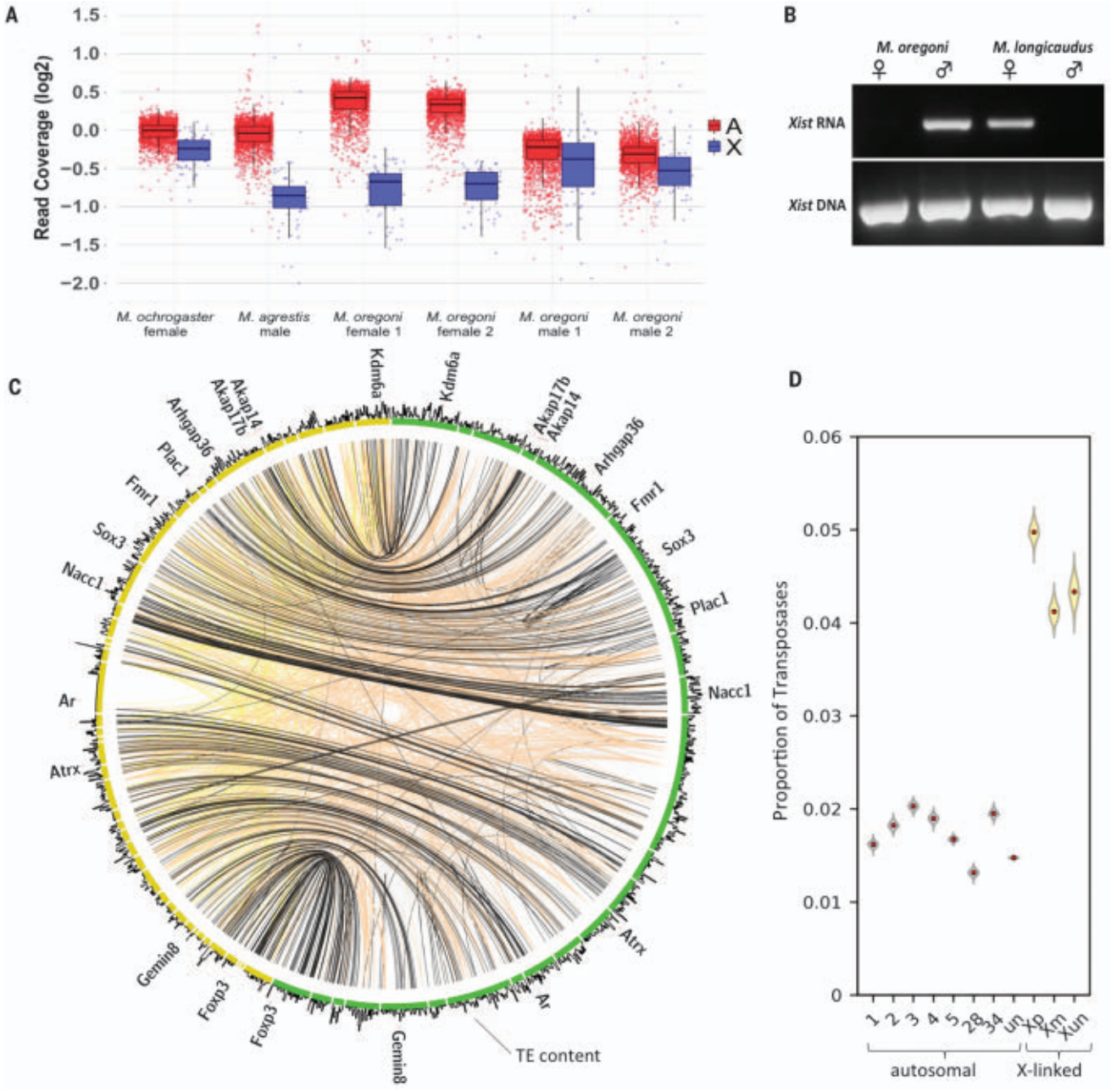


Fig. 3. Separate X^P and X^M haplotypes in male *M. oregoni*. (**A**) Doubled X chromosomal DNA sequencing read coverage in *M. oregoni* males suggests two X chromosomes. Read coverage comparisons are shown for autosomal (red) and X-linked (blue) genes for males and females of three species of *Microtus*. Related species show the expected reduction of X-linked coverage for males (*M. agrestis*) but not females (*M. ochrogaster*), whereas *M. oregoni* shows the reversed pattern. (**B**) *Xist* expression is sex-reversed in *M. oregoni*. *Xist* is expressed in male but not female *M. oregoni* (RT-PCR, replicated in n = 10 per sex); female-limited expression in *M. longicaudus* is typical of other eutherian mammals (RT-PCR, n = 1 per sex). *Xist* DNA amplifies in both sexes in both species (PCR, n = 1 per sex per species). (**C**) Circos plot of assembled X^P (green) and X^M (yellow) chromosomes. Black linkages denote homologies between low–copy number genes, showing clear conservation of synteny; colored linkages represent homologies for genes with higher copy numbers, showing more complex patterns of gene amplification (yellow and green show within-chromosome homologies; orange shows between-chromosome homologies). The histogram plot around the outside of the circle indicates transposase density. (**D**) Higher transposable element density on X chromosomes. The fraction of contig sequence covered by transposase genes is shown. Numbers denote the seven autosomal scaffolds that are nearly the length of chromosomes. un, unscaffolded autosomal sequence; Xun, unassigned X-derived sequence. Violin plots show variation over bootstrap replicates.

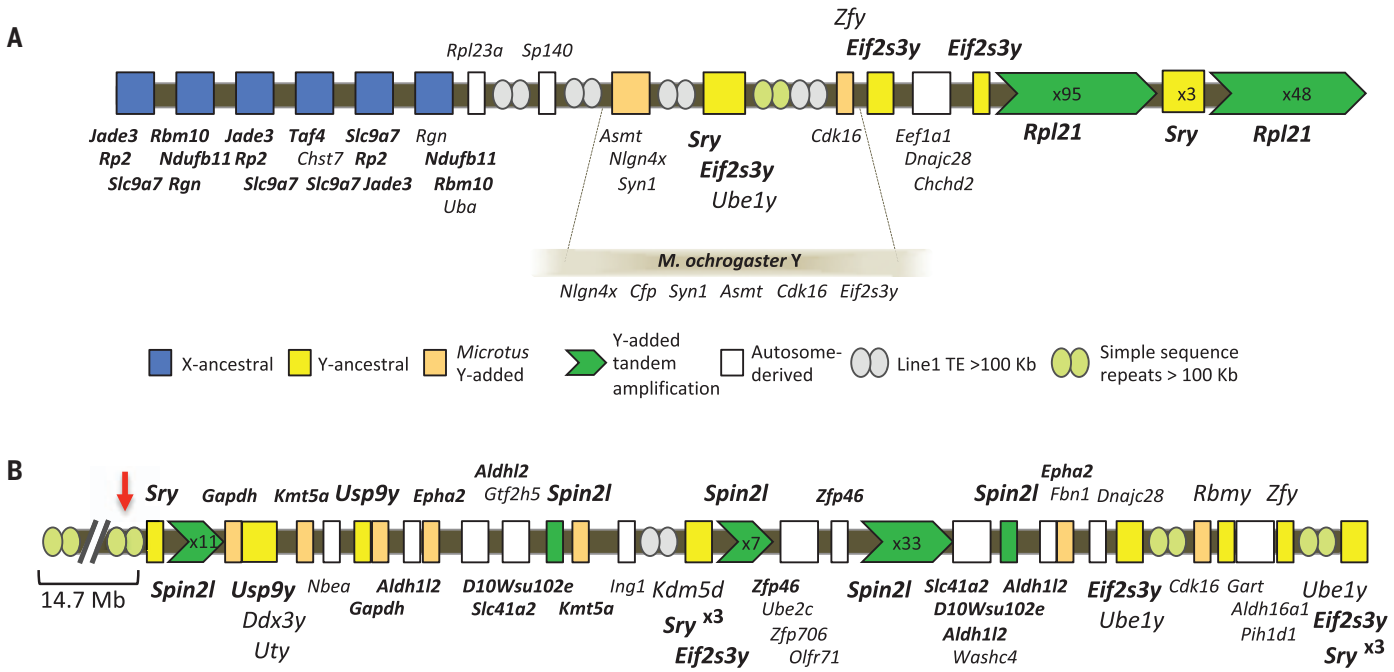


Fig. 4. Complex structure of Y-derived intervals on X^P and X^M. Cartoons of Y-derived scaffolds showing gene order and distributions of transposable elements and simple sequence repeats on (**A**) X^P and (**B**) X^M. Multicopy genes and genes amplified in tandem arrays are in bold; ancestral Y genes are in a larger font. *Microtus* Y-added genes are found on the *M. ochrogaster* Y chromosome but are not reported from other sequenced Y chromosomes. These include *Spin2l*, a single-copy autosomal gene with two translocated copies on the *M. ochrogaster* Y. The X^P scaffold is 6.5 Mb; the dashed lines mark an interval with putative homology to an interval on the *M. ochrogaster* Y chromosome. The X^M scaffold is 20.6 Mb; the red arrow indicates the start of a 14.7-Mb simple sequence repeat expansion. See also fig. S1. TE, transposable element.

S1A). Instead, all nine ancestral Y genes were concentrated on a large X^M scaffold that included 53 copies of a member of the spindlin gene family homologous to *Spin2l* in *M. ochrogaster*, as well as a 14.7-Mb span of simple sequence repeats that likely represents one of the heterochromatic regions that are specific to X^M (Fig. 4B). We found more than 100 Mb of smaller, currently unscaffolded contigs composed exclusively of this simple sequence repeat. We speculate that this large tract of heterochromatin forms a boundary between Y- and X-derived genes on X^M. We identified an X^P contig with putative homology to the genic interval shown in Fig. 4B that contained a subset of the same genes but was not colinear (fig. S1B). Notably, three ancestral Y genes (*Kdm5d*, *Rbmy*, and *Uty*) placed on this X^M scaffold were missing from all X^P contigs, suggesting absence from the X^P.

Taken together, our reconstructions of Y-derived intervals on X^M and X^P suggest an initial breakage-and-fusion event during male meiosis. This initial event was likely followed by a period during which both ancestral and newly formed sex chromosomes segregated together through the germ line. Additional breakage-and-fusion events and nonhomologous recombination before exclusion of X^M from male meiosis are non–mutually exclusive explanations for the partially shared gene content and structural differences between X^M and X^P. The excess of transposable elements in these Y-derived intervals (Fig. 4), and throughout X^M and X^P (Fig. 3, C and D), suggests a candidate mechanism for inferred breakage events.

The fate of Y-derived genes

We next studied the ancestral Y-linked genes. Given that these genes have a long history of being male specific, and that the X^M is transmitted only through females, we expected that inactivating mutations in the ancestral Y genes would be enriched on X^M relative to X^P. Using subassemblies of contigs from male and female short-read genomes, together with the male long-read genome and sex-specific transcriptomes, we reconstructed complete exon-intron structures for *Eif2s3y*, *Ddx3y*, *Kdm5d*, and *Uty* (Table 3). *Ddx3y* and *Kdm5d* were expressed in both sexes; *Uty*, *Ube1y*, and *Sry* were specific to male transcriptomes (Table 3). Notably, evidence for pseudogenization was not chromosome specific. For example, a nonsense mutation in *Ube1y* exon 15 was present in all reads within both male and female genome samples, consistent with presence on both X^M and X^P. Similarly, we found three *Ddx3y* transcripts (two female, one male) with different retained introns, all of which produced frameshifts and premature stop codons. Thus, we found no evidence for differential survival of ancestral Y genes in male versus female *M. oregoni*.

Both gene model–based annotation of Y-derived genes on X^M and X^P (Fig. 4) and read depth–based comparisons between female and male genomes (see below) provided evidence for copy number differences between the sex chromosomes. Given the primary function of *Sry* in initiating the male developmental pathway, the presence of multiple copies of *Sry* on a maternally transmitted X chromosome is particularly surprising. To validate and extend this result, we assayed *Sry* copy number in *M. oregoni* sampled from two populations and in *M. ochrogaster*, a species with multiple *Sry* copies on the Y (*36*) but none on the X. Results confirmed multiple copies of *Sry* in *M. oregoni* females, and the relative copy number for both populations was comparable to that of male *M. ochrogaster* (Fig. 5A). Notably, *Sry* copy number was highly elevated in male *M. oregoni*, with between-population differences suggestive of ongoing amplification of this gene on X^P.

To better understand the distribution of functional *Sry* copies between males and females, we retrieved and aligned all *Sry* sequences from contigs robustly assigned to X^M or X^P in our long-read male *M. oregoni* genome. We recovered 23 copies of *Sry* from this genome, including six pseudogenes, all assigned to X^P. The remainder had intact coding regions with seven copies assigned to X^M and 10 to X^P. All but two of the intact copies differed by at least one nonsynonymous substitution, with a minimum identity of 81.6% at the protein level. Phylogenetic analysis of

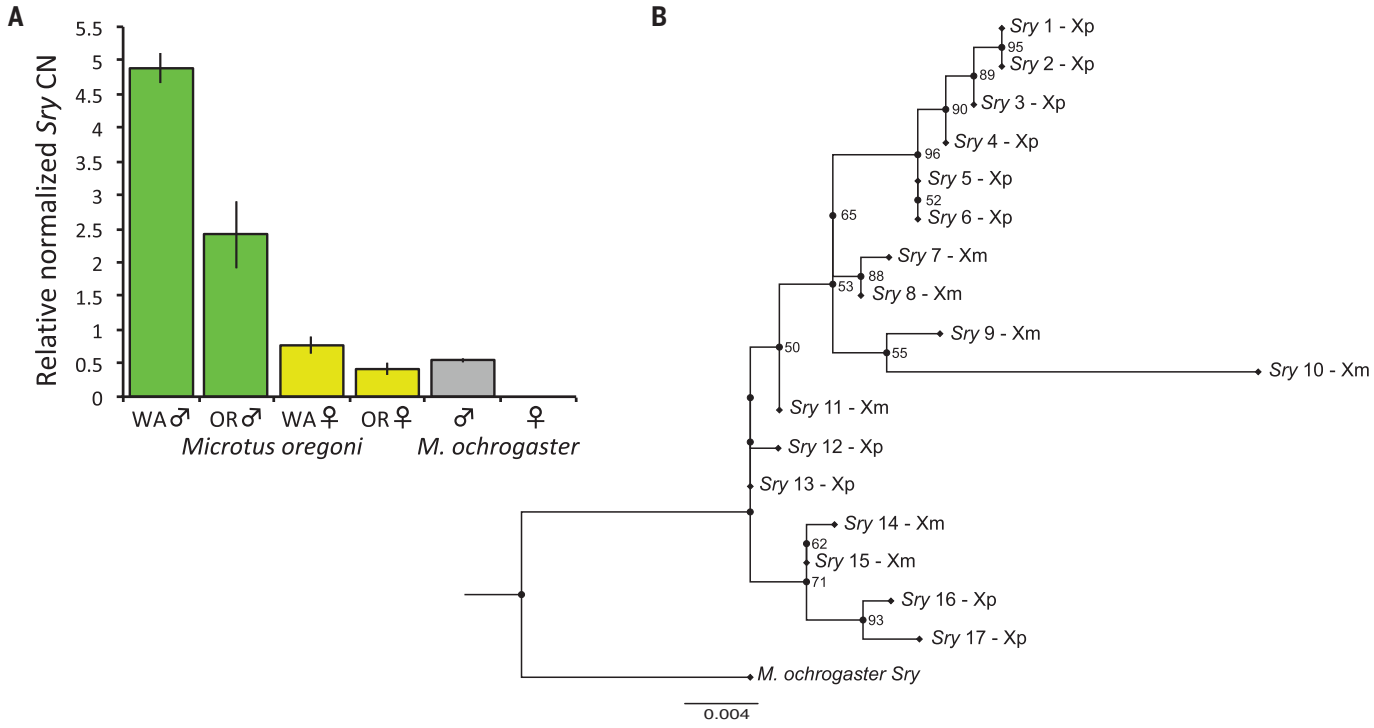


Fig. 5. Amplification and diversification of *Sry* in *M. oregoni*. (**A**) Relative copy number (CN) assay for *Sry*, showing differences between *M. oregoni* males from northern Washington (WA, *n* = 7) and Oregon (OR, *n* = 4), equivalent CN in female *M. oregoni* (WA, *n* = 12; OR, *n* = 3) and male *M. ochrogaster* (*n* = 3), and no signal in female *M. ochrogaster* (*n* = 3). Error bars indicate SEM. (**B**) Phylogenetic tree of 17 functional *Sry* copies on X^M and X^P in a single male *M. oregoni*, with *Sry* from *M. ochrogaster* as the outgroup. Bootstrap values were generated from 10,000 pseudoreplicates; values under 50 are not shown.

intact *Sry* copies suggests rapid diversification through amplification in the *M. oregoni* lineage, with some evidence for divergence between sequences on the two chromosomes (Fig. 5B).

Gene amplification across the X^P and X^M chromosomes

Copy number amplifications are a common feature of nonrecombining chromosomes, and amplification is thought to promote gene longevity on mammalian Y chromosomes (*11*). We therefore asked whether the copy number amplifications observed in the Y-derived genomic regions represented a more general feature of the X^P and X^M chromosomes. Given the likely incompleteness of the X^P and X^M haplotypes, we first evaluated gene amplification at a chromosome scale by mapping the short-read data from two female and two male *M. oregoni* to the (largely unamplified) *M. ochrogaster* short-read–based assembly (Mi_och 1.0) and assessing read coverage. Read coverage in *M. oregoni* females, and a single *M. agrestis* male, closely tracked that of the female *M. ochrogaster* genome (Fig. 6A). By contrast, *M. oregoni* males showed extended regions of amplification ranging from ~100 kb to ~1.5 Mb, with an estimated two- to eightfold elevation in read coverage. No similar amplifications were found on other chromosomes in *M. oregoni*. Comparisons across all four *M. oregoni* short-read genomes confirmed the lack of amplification in both females and showed largely indistinguishable patterns of amplification in the two males. A single observed region of amplification relative to *M. ochrogaster* (Fig. 6A) was shared between males and females, suggesting a shared X^P and X^M amplification.

We also compared read coverage for individual genes (Fig. 6B and table S8). These results showed the expected pattern for *M. oregoni*, with most ancestral X-linked genes (blue dots) showing male/female coverage ratios of ~2 for most X-derived genes (blue line), ~1 for putatively X^M-specific genes (*Kdm5d*, *Rbmy*, and *Uty*), and >2 for *Sry*. This analysis also suggested that X^P copy number is roughly equal to that on X^M for two additional ancestral Y genes (*Zfy* and *Eif2s3y*) and is higher for three genes (*Ube1y*, *Usp9y*, and *Ddx3y*). Consistent with the results in Fig. 4A, we also found evidence for elevated X^P copy number for many X-linked genes (dots above the blue line).

Increased protein sequence evolution on X^P and X^M chromosomes

Another common feature of nonrecombining chromosomes is accelerated rates of nonsynonymous change, generally attributed to reduced selective efficiency. We studied the candidate X^M-X^P differences (Table 1), which constitute most observed SNPs. We directionalized those changes as either X^M or X^P mutational changes. Consistent with inefficient selection due to lack of recombination, putative X^P changes had a highly elevated ratio of nonsynonymous to synonymous changes (dN/dS = 0.63 compared with 0.21 for fixed interspecific difference with *M. ochrogaster*; $P < 10^{-5}$). Moreover, the normalized incidence of SNPs leading to premature stop codons was 40 times that for species differences (0.46 versus 0.012). Rates of nonsynonymous changes and premature stop codons were not significantly different for various classes of genes that might be expected to experience differential selection on male- versus female-transmitted chromosomes (e.g., genes involved in gametogenesis and sex-biased genes; table S9). However, the pattern of putative fixed changes on X^M was very different from that on X^P: dN/dS was only moderately elevated relative to interspecific differences (dN/dS = 0.37), and no premature stop codons were observed.

Somatic inactivation of the X^P chromosome

Finally, given our finding of male-limited *Xist* expression in *M. oregoni*, we were interested in understanding X chromosome expression in XX males. In eutherian mammals, female somatic cells randomly silence one X chromosome in early development, a choice that is preserved in daughter cells. Using the 2648 putative X^P/X^M SNP differences, we quantified

allelic representation in RNA-seq datasets from somatic tissues (seminal vesicles, heart, brain, and liver) collected from two male *M. oregoni*. Notably, across seven of eight datasets, 99.4% of informative reads supported exclusive expression of X^M alleles, including ≥80% reads for each of the 134 genes individually (data S1). Thus, X chromosome inactivation occurs but is nonrandom in male *M. oregoni*. The four informative SNPs within the Y-derived chromosomal region also showed X^M-specific expression. This suggests that transcriptional silencing extends across much or all of the physical X^P chromosome and provides additional justification for referring to it as an X rather than a Y.

Notably, the seminal vesicle transcriptome was the only dataset with a few exceptions to X^P silencing. The expression of four genes (*Mbtps2*, *Prkx*, *LOC101999900*, and *Spin4*) was strongly X^P biased, potentially indicating local reactivation of the X^P in association with male reproduction. Consistent with X inactivation mediated by the expression of *Xist* from the otherwise silenced X, only X^P *Xist* variants were found in the transcriptomes (100% of reads supporting the X^P variant for all 66 SNP sites).

Nonrandom X chromosome inactivation in a eutherian mammal is surprising but could be explained in *M. oregoni* by *Xist* loss of function on X^M. We therefore evaluated the genomic *Xist* alleles on X^P and X^M but found no obvious functional sequence differences between the haplotypes in transcribed regions. However, comparing the *Xist* minimal promoter sequence with those of the mouse and *Microtus levis* (southern vole), we identified two X^M-specific point mutations in conserved *Xist* promoter elements (*37*, *38*) with likely functional effects. One mutation is located in conserved element VI, which contains a binding motif for a CTF/NF1 transcription factor (Fig. 7). Mutations in CTF/NF1 binding sites have been shown to decrease transcriptional activity in other promoters (*39*, *40*). Most notably, we found an A>G mutation at the transcription initiation site in conserved element I, only present in the X^M *Xist* allele (Fig. 7). In an in vitro assay using synthetic promoters with directed mutations in this transcriptional initiator element, introduction of a construct containing the same A>G mutation reduced transcriptional activity by 97% and reduced initiation-related transcription factor binding (*41*). Thus, the few nucleotide differences in the X^M *Xist* promoter may provide a functional explanation for nonrandom X chromosome inactivation in male *M. oregoni*, likely mediated at the level of *Xist* transcription initiation on X^M.

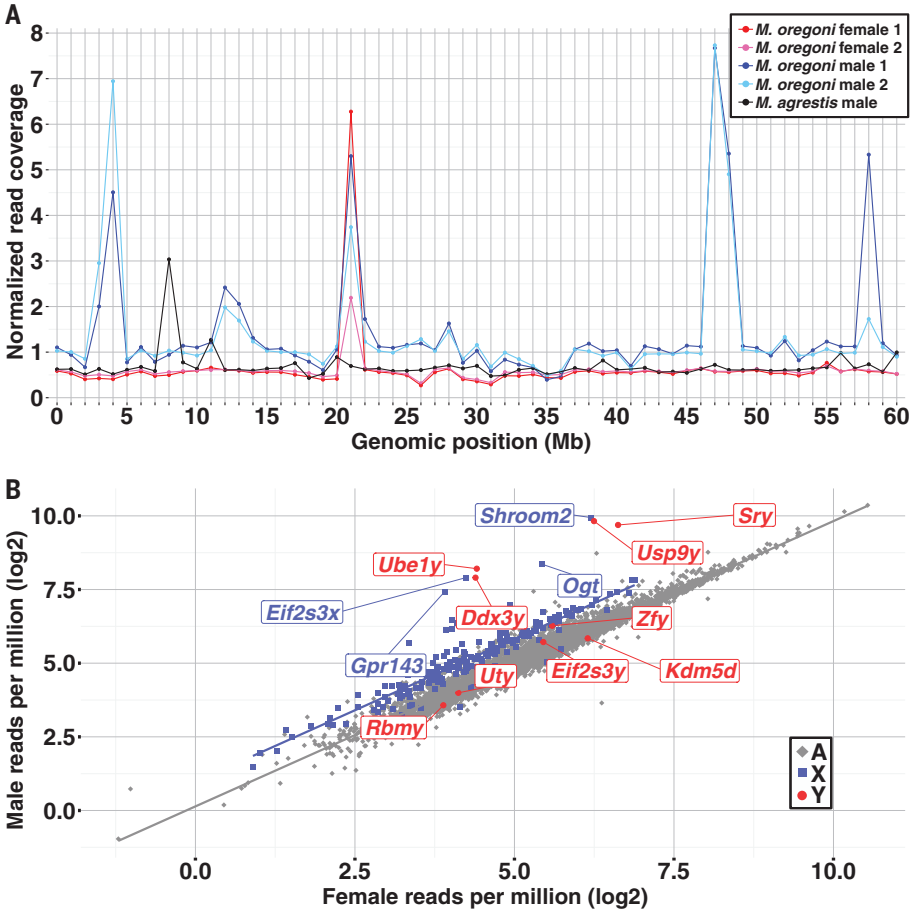


Fig. 6. Amplification of X^P- and X^M-linked genes. (**A**) Amplified regions on the male-specific *M. oregoni* X chromosome. DNA read density is shown in 1-Mb windows across the X chromosome of *M. ochrogaster*. Two *M. oregoni* males are shown in shades of blue, two females in shades of red, and male *M. agrestis* in black. Read density is normalized to the genome-wide average density for each species and to the read density of female *M. ochrogaster* for each window. (**B**) Read density for autosomal (A, gray), X-linked (X, blue), and ancestral Y chromosome (Y, red) genes in *M. oregoni*. Ancestral Y-linked genes and X-linked genes with elevated copy numbers in males are labeled, and regression lines for autosomal and X-linked genes (excluding outliers) are shown.

Discussion

We report substantial sex chromosome turnover in the creeping vole, including (i) presence of two X-like chromosomes in males; (ii) *Xist*-mediated silencing of the male-specific X chromosome in male somatic tissues; and (iii) a complete suite of ancestral Y-derived genes on the maternally transmitted X chromosome (summarized in Fig. 1C). Sex chromosomes in this species have undergone radical restructuring that includes homogenization of gene content between males and females, scrambling of gene order, and divergent molecular evolution. Both the maternally transmitted (X^M) and the male-limited (X^P) chromosomes carry full sets of X chromosome genes joined to core ancestral Y-linked genes. The presence of multiple functional copies of *Sry* on both sex chromosomes raises the question of why X^M0 individuals do not develop as males. It will be interesting to learn how masculinization is avoided. Even more surprising, at least two ancestral Y genes are expressed in females, three may be exclusively transmitted through females, and none show evidence of X^M-specific degeneration. Given that X chromosome genes with surviving Y homologs are overrepresented among the X inactivation escape genes common to mice and humans (*42*, *43*) and are disproportionately likely to be haploinsufficient in humans (*44*), it is possible that the functional Y gene copies retained on X^M compensate for reduced dosage of their X-linked homologs in X0 females. On the other hand, independent loss of most Y-linked genes in two rodent species in which both sexes are X0 [*Ellobius lutescens* (*35*) and *Tokudaia osimensis* (*45*)] suggests that such dosage constraints are surmountable. Alternatively, ancestral Y-linked gene copies on the X^M could be essential for somatic functions in males, particularly given X^P silencing.

In species with ancient, heteromorphic sex chromosomes (e.g., *Drosophila* and mammals),

the rare events that give rise to a new sex chromosome often involve chromosomal fusion (*46*, *47*). Indeed, the male-specific sex chromosome in *M. oregoni* was originally interpreted as an X-Y fusion (*48*, *49*), a hypothesis dismissed by Ohno (*21*). Features of the genome both support and complicate this hypothesis. Concentration of Y-derived genes within specific chromosomal regions suggests the possibility of an X-Y fusion, and we find a point of contact between X- and Y-derived genes on X^P. However, other features, including structural differences between Y-related regions on the X^M and X^P chromosomes and the apparent absence of some Y-derived genes from the X^P, point to a more complex history.

Perhaps the most likely scenario involves initial creation of the proto-X^P through an X-Y fusion during male meiosis. The Y-derived sequence could then have been transferred from the proto-X^P chromosome to a "standard" X chromosome through nonhomologous recombination during spermatogenesis. However, we cannot exclude the possibility that the X^P and X^M independently acquired sets of Y-linked genes directly from a more typical Y chromosome.

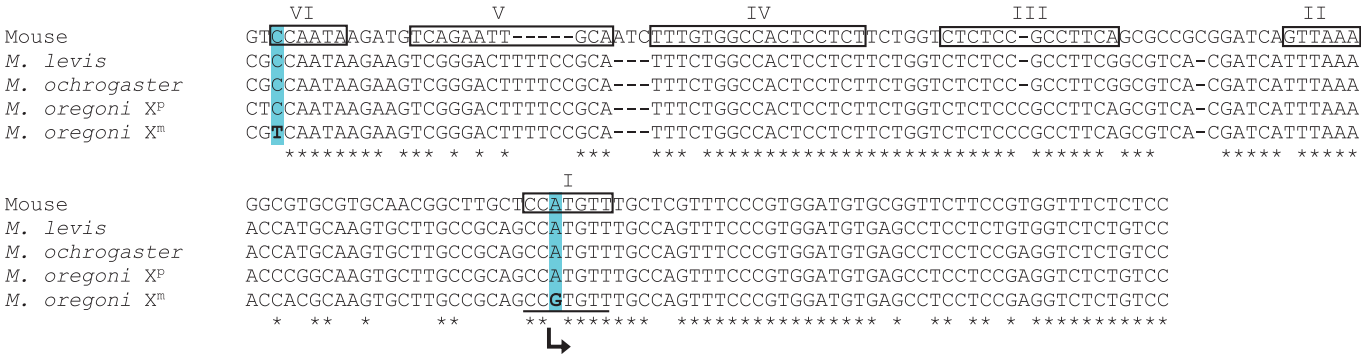


Fig. 7. Multiple mutations in core promoter elements of *Xist* on X^M. The sequence alignment of the *Xist* minimal promoter shows disruptions in conserved promoter elements I to VI (boxes). The underline and arrow represent the transcription start sequence and site, respectively, in element I. Blue shading indicates conserved sites at which X^M allele mutations (bold) potentially affect transcription of *Xist* in *M. oregoni* males. Asterisks indicate fully conserved bases.

Table 1. X-linked polymorphism types and rates, and measures of evolutionary rate for *M. oregoni* males and females and between *M. oregoni* and *M. ochrogaster*. The second column indicates whether *M. ochrogaster* exhibits the sex-specific or shared sequence. Synon., synonymous; Nonsynon., nonsynonymous; CI, confidence interval; –, not applicable.

M. oregoni sequences	*M. ochrogaster* sequence	Candidate explanation	Sites	% of Intraspecific SNPs	Synon.	Nonsynon.	Stop	dS (95% CI)	dN/dS (95% CI)	dSTOP/dS (95% CI)
Monomorphic	Alternative	Species difference	9145	–	5385	3748	12	0.024 (0.024–0.025)	0.21 (0.21–0.22)	0.01 (7×10^{-3}–0.02)
Male-specific (both males heterozygous)	Shared	X^P mutation	2030	70.2%	649	1324	57	2.9×10^{-3} (2.7×10^{-3}–3.2×10^{-3})	0.63 (0.57–0.69)	0.46 (0.35–0.61)
Male-specific (both males heterozygous)	Male-specific	X^M mutation	618	21.4%	280	338	0	1.3×10^{-3} (1.1×10^{-3}–1.4×10^{-3})	0.37 (0.32–0.44)	0.0 (0.0–0.15)
Singleton in female	Shared	X^M mutation	20	0.7%	11	9	0	5×10^{-5} (2×10^{-5}–8×10^{-5})	0.252 (0.11–0.6)	0.0 (0.0–3.85)
Singleton in male (heterozygous)	Shared	X^M or X^P mutation	124	4.3%	36	84	4	1.6×10^{-4} (1.1×10^{-4}–2.2×10^{-4})	0.718 (0.48–1.1)	0.580 (0.24–1.71)
Various	–	Various	100	3.5%	33	64	3	1.5×10^{-4} (1×10^{-4}–2×10^{-4})	0.597 (0.39–0.90)	0.474 (0.18–1.64)

Table 2. Conservation of X-derived gene content on the *M. oregoni* sex chromosomes, X^M and X^P. "Total genes" denotes the subset of X-derived protein coding genes that could be assigned with confidence to X^M and X^P.

	M. oregoni	*M. ochrogaster*		Mouse	
	Total genes	X	Autosome	X	Autosome
X^M	272	225	47	187	41
X^P	335	264	71	215	65

Table 3. Ancestral Y-linked gene reconstruction and expression in male and female *M. oregoni*. ?, not all exons found; RI, retained introns in all transcripts; stop, one or more nonsense mutations.

Gene	Intact in genomes		Detected in transcriptomes		Functional copy	
	♂	♀	♂	♀	♂	♀
Eif2s3y	✓	✓	X	X	✓	✓
Zfy	?	?	X	X	?	?
Usp9y	?	?	X	X	?	?
Ddx3y	✓	✓	✓	✓	RI	RI
Uty	✓	✓	✓	X	✓	✓
Ube1y	?	?	✓	X	RI, stop	stop
Kdm5d	✓	✓	✓	✓	✓	✓
Rbmy	✓	✓	X	X	✓	✓
Sry	✓	✓	✓	X	✓	✓

Notably, whereas X^P exhibits the hallmarks of degradation of nonrecombining chromosomes (*46*, *50*), X^M appears somewhat less transformed, with less-pronounced increase in dN/dS, lower transposable element density, and less genomic amplification. This pattern could be explained if recombination cessation of the X^P occurred first—for instance, during an intermediate stage in which the X^P segregated like a Y chromosome in a standard XY system. However, more effective purifying selection on the X^M could also be due to lack of expression of X^P-linked genes and/or a higher mutation rate on X^P due to male-driven mutation (*51*–*53*). Indeed, the existence of nonrandom X inactivation raises the question of why the X^P remains largely intact. Given that X-linked genes are highly expressed during the early stages of mammalian spermatogenesis (*7*, *54*), we speculate that large-scale deletions on X^P are opposed by selection on germline functions in male *M. oregoni*.

Collectively, these results provide insight into the essential and expendable properties of mammalian sex chromosomes, demonstrating that ancestrally male-specific genes can be accommodated in female genomes, while supporting the proposed dosage sensitivity of ancient X-Y gene pairs (*8*, *11*). The *M. oregoni* sex chromosome karyotype has intrigued evolutionary biologists for decades (*21*, *26*, *55*–*57*). Our resolution of this puzzle reveals a pattern of sex chromosome transformation that was previously unknown in mammals and lays the foundation for using *M. oregoni* as a model system for sex chromosome evolution and function.

REFERENCES AND NOTES

1. J. J. Bull, *Evolution of Sex Determining Mechanisms* (Benjamin/Cummings, 1983).
2. Á. S. Roco *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4752–E4761 (2015).
3. N. Rodrigues, Y. Vuille, A. Brelsford, J. Merilä, N. Perrin, *Heredity* **117**, 25–32 (2016).
4. J. E. Mank, J. C. Avise, *Sex Dev.* **3**, 60–67 (2009).
5. J. Kitano, C. L. Peichel, *Environ. Biol. Fishes* **94**, 549–558 (2012).
6. D. Bachtrog *et al.*, *PLOS Biol.* **12**, e1001899 (2014).
7. L. Potrzebowski *et al.*, *PLOS Biol.* **6**, e80 (2008).
8. D. Cortez *et al.*, *Nature* **508**, 488–493 (2014).
9. R. A. Close, M. Friedman, G. T. Lloyd, R. B. J. Benson, *Curr. Biol.* **25**, 2137–2142 (2015).
10. J. F. Hughes *et al.*, *Nature* **483**, 82–86 (2012).
11. D. W. Bellott *et al.*, *Nature* **508**, 494–499 (2014).
12. B. M. Skinner *et al.*, *Genome Res.* **26**, 130–139 (2016).
13. D. W. Bellott *et al.*, *Nature* **466**, 612–616 (2010).
14. J. L. Mueller *et al.*, *Nat. Genet.* **45**, 1083–1087 (2013).
15. F. Lin, K. Xing, J. Zhang, X. He, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11752–11757 (2012).
16. E. Pessia, T. Makino, M. Bailly-Bechet, A. McLysaght, G. A. B. Marais, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5346–5351 (2012).
17. X. Deng, J. B. Berletch, D. K. Nguyen, C. M. Disteche, *Nat. Rev. Genet.* **15**, 367–378 (2014).
18. F. Printzlau, J. Wolstencroft, D. H. Skuse, *J. Neurosci. Res.* **95**, 311–319 (2017).
19. R. Matthey, *Experientia* **12**, 337–338 (1956).
20. R. Jiménez, F. J. Barrionuevo, M. Burgos, *Sex Dev.* **7**, 147–162 (2013).
21. S. Ohno, J. Jainchill, C. Stenius, *Cytogenetics* **2**, *232*–*239* (1963).
22. S. Ohno, C. Stenius, L. Christian, in *Chromosomes Today*, C. D. Darlington, K. R. Lewis, Eds. (Oliver and Boyd, 1966), vol. 1, pp. 182–187.
23. S. Ohno, in *Congenital Malformations, Papers and Discussion Presented at the 2nd International Conference* (New York, 1964), pp. 40–42.
24. G. G. Musser, M. D. Carleton, in *Mammal Species of the World: A Taxonomic and Geographic Reference*, D. E. Wilson, D. M. Reeder, Eds. (Smithsonian Institution Press, 2005), pp. 955–1188.
25. W. S. Modi, *Syst. Biol.* **36**, 109–136 (1987).
26. B. Charlesworth, N. D. Dempsey, *Heredity* **86**, 387–394 (2001).
27. N. Spies *et al.*, *Nat. Methods* **14**, 915–920 (2017).
28. E. Lieberman-Aiden *et al.*, *Science* **326**, 289–293 (2009).
29. J. N. Burton *et al.*, *Nat. Biotechnol.* **31**, 1119–1125 (2013).
30. S. Fink, M. C. Fischer, L. Excoffier, G. Heckel, *Syst. Biol.* **59**, 548–572 (2010).
31. I. Okamoto, A. P. Otte, C. D. Allis, D. Reinberg, E. Heard, *Science* **303**, 644–649 (2004).
32. B. L. Libbus, L. A. Johnson, *Cytogenet. Cell Genet.* **47**, 181–184 (1988).
33. S. Nurk *et al.*, *Genome Res.* **30**, 1291–1305 (2020).
34. J. A. Marchal, M. J. Acosta, M. Bullejos, R. Díaz de la Guardia, A. Sánchez, *Genomics* **91**, 142–151 (2008).
35. E. Mulugeta *et al.*, *Genome Res.* **26**, 1202–1210 (2016).
36. M. J. Acosta *et al.*, *Sex Dev.* **4**, 336–347 (2010).
37. S. A. Sheardown, A. E. T. Newall, D. P. Norris, S. Rastan, N. Brockdorff, *Gene* **203**, 159–168 (1997).
38. T. B. Nesterova *et al.*, *Genome Res.* **11**, 833–849 (2001).
39. B. W. Konicek, X.-L. Xia, M. A. Harrington, *DNA Cell Biol.* **14**, 961–969 (1995).
40. J. M. Giger, P. W. Bodell, K. M. Baldwin, F. Haddad, *Exp. Physiol.* **94**, 1163–1173 (2009).
41. R. Javahery, A. Khachi, K. Lo, B. Zenzie-Gregory, S. T. Smale, *Mol. Cell. Biol.* **14**, 116–127 (1994).
42. F. Yang, T. Babak, J. Shendure, C. M. Disteche, *Genome Res.* **20**, 614–622 (2010).
43. B. P. Balaton, C. J. Brown, *Trends Genet.* **32**, 348–359 (2016).
44. J. F. Hughes, D. C. Page, *Annu. Rev. Genet.* **49**, 507–527 (2015).
45. A. Kuroiwa, Y. Ishiguchi, F. Yamada, A. Shintaro, Y. Matsuda, *Chromosoma* **119**, 519–526 (2010).
46. D. Bachtrog, *Nat. Rev. Genet.* **14**, 113–124 (2013).
47. C. Murata, Y. Kuroki, I. Imoto, A. Kuroiwa, *Chromosome Res.* **24**, 407–419 (2016).
48. M. J. D. White, *Proc. Zool. Calcutta, Mookerjee Memor. Vol.* 113–114 (1957).
49. R. Matthey, *Experientia* **14**, 240–241 (1958).
50. B. Charlesworth, D. Charlesworth, *Philos. Trans. R. Soc. London Ser. B* **355**, 1563–1572 (2000).
51. J. F. Crow, *Nat. Rev. Genet.* **1**, 40–47 (2000).
52. J. Shendure, J. M. Akey, *Science* **349**, 1478–1483 (2015).
53. M.-C. Grégoire *et al.*, *Cell. Mol. Life Sci.* **75**, 2859–2872 (2018).
54. S. H. Namekawa *et al.*, *Curr. Biol.* **16**, 660–667 (2006).
55. E. Lifschytz, D. L. Lindsley, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 182–186 (1972).
56. L. D. Hurst, A. Pomiankowski, *Genetics* **128**, 841–858 (1991).
57. C.-I. Wu, A. W. Davis, *Am. Nat.* **142**, 187–212 (1993).

ACKNOWLEDGMENTS

We are grateful to T. Manning for advice and assistance with field work and to W. Timp for use of the PromethION instrument. We thank P. Blood, R. Acosta, and other members of the Pittsburgh Supercomputing Center for assistance in implementing assembly programs on Bridges. We thank J. Schafer at the Oklahoma State University High Performance Computing Center for implementation of software and computational jobs on the cluster. P.C. thanks J. M. Campbell for help translating R. Matthey and dedicates this work to her memory. **Funding:** This work was funded by National Science Foundation IOS 1558109 and start-up funds from the University of California, Riverside (P.C.), National Science Foundation MCB 1616878 (S.W.R.), and National Institutes of Health R44GM134994 (K.J.L.). **Author contributions:** Conceptualization: P.C., M.B.C., and S.W.R.; Field collections: P.C., L.S.M., T.M.W., C.W.E., L.D., and L.A.R.; Assembly, annotation, and SNP calling: M.B.C.; High–molecular weight DNA extraction and Nanopore data production: M.K., D.K., and K.J.L.; Illumina sequencing: S.P.; Data analysis and validations: S.W.R., M.B.C., N.A., L.G., and P.C.; Writing: P.C., S.W.R., M.B.C., and N.A. **Competing interests:** K.J.L., M.K., and D.K. are employees of Circulomics Inc. K.J.L. owns stock in Circulomics. K.J.L. and D.K. have awarded and pending patents related to the methods used. **Data and materials availability:** Raw data are deposited in the NCBI Sequence Read Archive (Bioproject IDs PRJNA361280 and PRJNA505851). The genome assembly is available at NCBI (accession number JAGKIF000000000). Custom scripts are available at https://github.com/lgozasht/Microtus-sex-chromosome-evolution.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/592/suppl/DC1
Materials and Methods
Fig. S1
Tables S1 to S13
References (*58*–*91*)
Data S1

22 January 2021; accepted 7 April 2021
10.1126/science.abg7019

PALEONTOLOGY

The early origin of a birdlike inner ear and the evolution of dinosaurian movement and vocalization

Michael Hanson[1,2], Eva A. Hoffman[3], Mark A. Norell[3], Bhart-Anjan S. Bhullar[1,2]*

Reptiles, including birds, exhibit a range of behaviorally relevant adaptations that are reflected in changes to the structure of the inner ear. These adaptations include the capacity for flight and sensitivity to high-frequency sound. We used three-dimensional morphometric analyses of a large sample of extant and extinct reptiles to investigate inner ear correlates of locomotor ability and hearing acuity. Statistical analyses revealed three vestibular morphotypes, best explained by three locomotor categories—quadrupeds, bipeds and simple fliers (including bipedal nonavialan dinosaurs), and high-maneuverability fliers. Troodontids fall with *Archaeopteryx* among the extant low-maneuverability fliers. Analyses of cochlear shape revealed a single instance of elongation, on the stem of Archosauria. We suggest that this transformation coincided with the origin of both high-pitched juvenile location, alarm, and hatching-synchronization calls and adult responses to them.

Birds (Aves) are the most functionally diverse fliers (*1*) and, save for humans and cetaceans, perhaps the most sophisticated vocal communicators among vertebrates (*2*, *3*). The early history of the acquisition of these locomotor and vocal abilities is obscure. Attempts to reconstruct the history of avian flight have relied primarily on the musculoskeletal anatomy of the forelimb (*4*) and the presence of asymmetric feathers (*5*). Putatively lift-capable wings, necessary for aerial locomotion, first appeared in some form in the last common ancestor of either Pennaraptora or Paraves (*4*); however, their presence alone is insufficient to demonstrate the sophisticated control of three-dimensional (3D) locomotion required for birdlike flight. Attempts to reconstruct the origin of complex vocalization that rely strictly on the functional anatomy of the vocal apparatus are hindered by a lack of preservation; a well-preserved vocal apparatus has been recovered from a single extinct reptile—a member of crown Aves, whose ancestral vocal abilities were never in question (*6*).

There is another approach to investigating these transitions that does not require the preservation of the organs that produce the physical action in question. This approach involves examining the structure of the sensory organs that, together with the brain, monitor and mediate the action (*7*–*9*). In the cases of both flight and communication, the components of the inner ear are potentially useful to infer behavior with this method. The inner ear is a multipart sensory structure—situated deep within the skull and surrounded by the neurocranium (Fig. 1)—whose form and function are thought to be intimately linked to behavior (*7*, *10*). It comprises two mechanosensory organs: (i) the vestibular labyrinth, composed of the three semicircular canals, which detect angular movement and acceleration, and the utriculus and sacculus, which respond to linear motion and (ii) the cochlea, composed of the lagena or cochlear duct and the perilymphatic duct, which detects sound in the form of aerial vibrations transmitted (in crown reptiles) from the middle ear (Fig. 1 and figs. S1 to S6) (*10*, *11*). These organs can be understood as two evolutionary modules subject to independent selective pressures related to locomotion and hearing. A geometric morphometric test for modularity supports this assumption (*12*) (table S7). Thus, the form of the vestibular system (semicircular canals and associated structures) is often taken as indicative of locomotor ability and habits (*11*, *13*), and the form of the cochlea is taken as a proxy for hearing range and sensitivity (*7*).

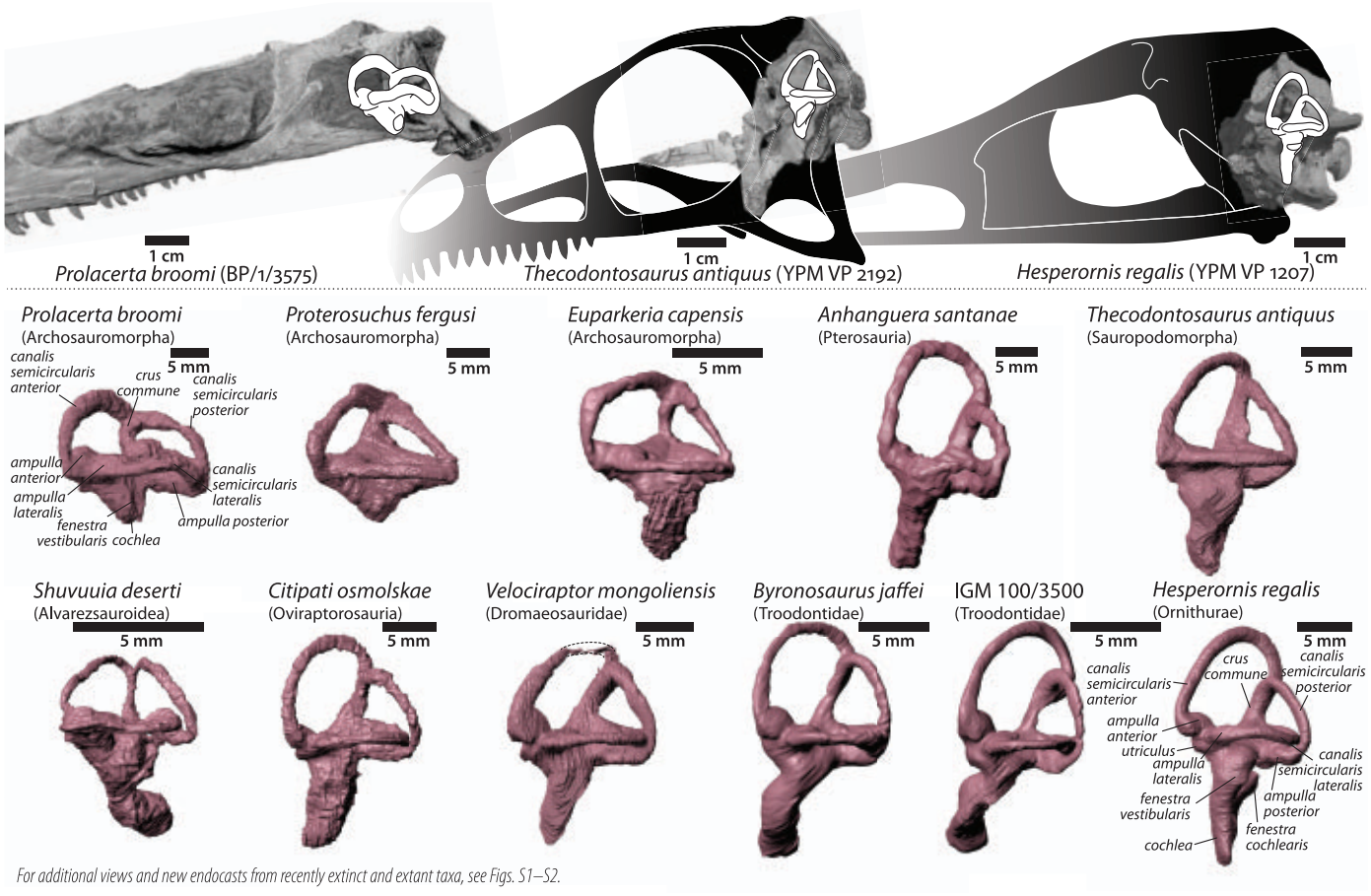
Reptiles are enormously varied in their locomotor habits, ranging from limbless, aquatic, and quadrupedal lepidosaurs to aquatic, bipedal, and flying archosaurs (*1*, *7*, *11*). Acoustic behaviors can be difficult to detect (*14*), but lepidosaurs are generally limited in vocal use and complexity, whereas archosaurs as a whole produce a wide array of vocalizations, especially for intraspecies communication (*15*). These vocalizations become particularly complex in birds (*16*). Vocal communication in reptiles typically requires the ability to detect high-frequency far-field sounds (*7*). A growing body of research is identifying morphology-behavior correspondences in both extant and extinct reptiles (*11*, *13*) as well as correspondences between hearing sensitivity and cochlear geometry among extant reptiles (*7*). However, past research is limited in taxonomic and/or anatomical scope; no all-reptile (including birds) analyses have been conducted, and few if any fossil taxa have been included.

[1]Department of Earth and Planetary Sciences, Yale University, New Haven, CT, USA. [2]Peabody Museum of Natural History, Yale University, New Haven, CT, USA. [3]Division of Paleontology, American Museum of Natural History, New York, NY, USA.
*Corresponding author. Email: bhart-anjan.bhullar@yale.edu

Reptile inner ear geometry, modularity, and allometry

We conducted a comprehensive geometric morphometric analysis of 38 landmarks [on the basis of conserved anatomy and on previously identified homologous landmarks in previous analyses (*7*, *11*, *17*, *18*)] (fig. S6). We placed the landmarks on 3D digital endocasts extracted from micro–computed tomography (micro-CT) scans of the inner ears of 124 reptilian taxa, including four non-archosaurian archosauromorphs and seven nonavialan dinosaurs. Data entirely new to this study (fig. S1) (*19*) include the complete virtual inner ears of the early-diverging archosauromorph *Prolacerta*; the near-crown archosauromorph *Euparkeria*; the pterosaur *Anhanguera*; the early-diverging sauropodomorph *Thecodontosaurus*; and a range of pennaraptoran dinosaurs including the alvarezsauroid *Shuvuuia*, the oviraptorosaur *Citipati*, the dromaeosaur *Velociraptor*, two troodontids, and the avialan *Hesperornis*. We also extracted six new endocasts from extant and recently extinct taxa (fig. S2). We chose not to use semilandmarks in recognition of the potential for spurious signal or noise to be introduced [(*20*) but see (*21*)]. We recognize, however, that many researchers do use semilandmarks, including researchers who recently performed a large-scale analysis of avian ears (*22*). To ensure that our results would be comparable, we performed analyses of only the taxa from that work and recovered a nearly identical distribution of ears in morphospace (figs. S30 to 33). In general, the inner ear complex is thought to be bilaterally symmetrical across Vertebrata (*10*), and this has recently been confirmed in archosaurs (*18*). We nonetheless verified symmetry in a phylogenetically broad subset of our sample (figs. S74 to S77). This subsample included an owl (*Tyto alba*), whose inner ears we found to be as symmetrical as those of other taxa, despite the well-known asymmetry in the outer and middle ears in Strigiformes (*23*).

Our morphological dataset enabled us to reconstruct ancestral inner ear architecture at key nodes from Reptilia to Neoaves (Figs. 1 to 5 and figs. S1 to S5 and S27 to S29). We mined behavioral data on locomotion, diet, and social behavior from a comprehensive set of sources (*12*) (table S2). This allowed us to infer the behavior of extinct taxa on the basis of their position in morphospace relative to known functional and behavioral clusters, corroborated by unweighted pair-group method with arithmetic mean (UPGMA) analyses (figs. S71 to S73). We tested these associations using phylogenetic analyses of variance (ANOVAs) (Tables 1 and 2; see also tables S4 and S5 for Goodall's *F*) and canonical variates and



For additional views and new endocasts from recently extinct and extant taxa, see Figs. S1–S2.

Fig. 1. Selected archosauromorph inner ear endocasts. (**Top**) Cranial micro-CT scans and inner ear locations of selected fossil taxa new to this study. (**Bottom**) Inner ear endocasts of fossil taxa new to or newly segmented for this study. All images are original.

discriminant function analyses (figs. S60 to S70; for behavioral data and classifiers, see Figs. 2 and 3, figs. S7 to S26, and table S2) (*12*).

Although we analyzed whole-ear shape for our primary analyses, we also tested for and found significant modularity between the vestibular system and the cochlea (covariance ratio, 0.874; effect size, −3.384; $P = 0.002$ for the phylogenetic modularity test) (*12*). Therefore, we performed secondary and supplementary analyses on the vestibular and cochlear regions separately, using procrustes alignments derived from the whole-ear dataset and from the regional datasets alone (*12*) (tables S6, S7, S9, and S10 and figs. S7 to S10, S27 to S29, S34 to S39, and S44 to S51). We additionally analyzed the entire dataset to determine the effects of phylogeny, body size, and neurocranial anatomy on ear shape (vestibular, cochlear, and combined). As expected, whole ears, vestibular systems, and cochleae showed significant phylogenetic signal (table S6), both across all reptiles and within many more exclusive clades. To account for this signal, we conducted phylogenetic significance tests and analyses in addition to traditional nonphylogenetic analyses (*12*).

We tested for allometric scaling by plotting both centroid size and log-transformed body mass against whole-ear, semicircular canal, and cochlea shapes (*12*) (figs. S23 to S26; body-mass estimates were available for only a subset of fossil taxa). The results indicate a low correlation between size and shape [coefficient of determination (R^2) of 0.037 for centroid size versus whole ear and 0.22 for mass versus whole ear], which suggests weak allometric influence on ear anatomy. Clusters identified in the principal components analyses (PCAs) are broadly distinct and seem to be explained by locomotor and behavioral factors as discussed below. The functional signal we identify therefore appears to be largely independent of allometry: It is stronger than the allometric signal and it holds within as well as among the various clusters responsible for the appearance of allometric scaling. This stands in contrast to the conclusions of a recent study of bird semicircular canals (*22*), in which the allometric signal was found to be much stronger than the functional. The data analyzed in this study consisted of semicircular canal curvature and intercanal angle; the authors did not include the additional static vestibular landmarks that we used, but instead primarily relied on a large number of sliding semilandmarks along a midline curvature defined with static landmarks solely at the start and endpoints of the semicircular canals. Additional regions of the vestibular system (e.g., the ampullae or utricular regions) were not investigated, nor was the cochlea a subject of the study, because its focus was on locomotion. We suspect that a combination of broader phylogenetic scope and increased coverage of inner ear shape in our study, with static landmarks to better capture more homologous or partially homologous regions, contributes to these differences, in addition to our consideration of previously untested functional and behavioral variables using several statistical methods.

We considered the possibility that skull shape constrains ear shape to some extent, as has been suggested in some recent work (*22*). We argue instead that the relationship is likely one in which ear shape helps to determine the shape of the skull, especially the surrounding neurocranium. Embryonically, inner ears are formed long before skeletal tissues differentiate (*24*, *25*). They are placodal derivatives and their morphogenesis is substantially under

the control of a poorly understood set of intrinsic factors (*26*). The inner ears are, in turn, known to generate paracrine signals that affect skeletogenic mesenchyme (*26*, *27*), and it is clear that they also have substantial mechanical influence on the developing chondrocranium (*24*, *27*); indeed, much of the braincase develops as a sheath for the labyrinth (*28*). We hypothesized that skull shape would be correlated with ear shape owing to the ear's function as a morphogenetic signaling center. We focused on the posterior neurocranium, which is the part of the skull that develops in association with the inner ear. As predicted, phylogenetic two-block partial least-squares tests yielded modest correlations between skull shape and both vestibular and cochlear shapes; the vestibular effect size was considerably larger than the cochlear (*12*) (table S7 and figs. S44 to S51). Neurocranium-shape clusters are fewer and less well-separated than ear-shape clusters, which are best explained by locomotor and vocal function, as described below. This disparity is consistent with the idea that the ear is just one of multiple influences on the development of the skull. For example, it is known that other neurosensory structures, including the brain itself, play a role in patterning the skull (*29*, *30*).

Reconstructed ancestral inner ears

Reconstruction of ancestral inner ear shapes revealed that the inner ears of early-diverging archosauromorphs closely resemble those of lepidosaurs and turtles, and of the inferred ancestral reptile, in having anteroposteriorly broad vestibular systems and short cochleae (Fig. 5 and figs. S4 and S5) (*31*). Crown-clade archosaurs have taller and more acutely arched anterior semicircular canals, long cochleae, and reduced saccular and vestibular fenestrae (Figs. 2 to 5 and figs. S1 to S5). A tall anterior semicircular canal evolved early within Avemetatarsalia, a trend that continued further in late-diverging paravians (Figs. 1, 2, and 5 and fig. S4). Pterosaurs convergently developed a paravian-like condition, as has been observed previously in qualitative analyses [(*11*, *32*, *33*) and comments on pterosaur ear evolution in (*12*)]. For comments on several interesting features of individual inner ear endocasts, see the supplementary materials (*12*).

Vestibular system

Our analyses of the entire ear and of the isolated vestibular system (Figs. 2 and 3, Table 1, table S4, and figs. S19 to S22 and S60 to S67) (*12*) show overall division of extant taxa into three distinct, largely nonoverlapping vestibular morphology clusters, also appearing in UPGMA analysis (figs. S71 to S73). We examined the structure of the data by testing the explanatory power of a series of behavioral hypotheses (Table 1, table S4, and figs. S60 to S67). Phylogenetic ANOVAs revealed significant differences among taxa with different locomotor modes. Quadrupeds, limbless taxa, bipeds, and fliers all had significantly different vestibular shapes, with an especially strong signal for flying versus flightless taxa across the entire sample. There was also a strong signal for aquatic and semiaquatic habitats. Taxa that exhibit aerial predation, a behavior strongly associated with high-maneuverability or sophisticated flight, were significantly different from taxa that do not. Taxa in nonlocomotor behavioral and ecological categories, on the other hand, did not have significantly different vestibular shapes. Canonical variate and discriminant function analyses of whole-ear and vestibular-system geometries showed very strong predictive power for locomotor mode including flight (99.2 to 100% discrimination of confusion matrices; 86.4 to 98.4% discrimination of jackknifed confusion matrices), for

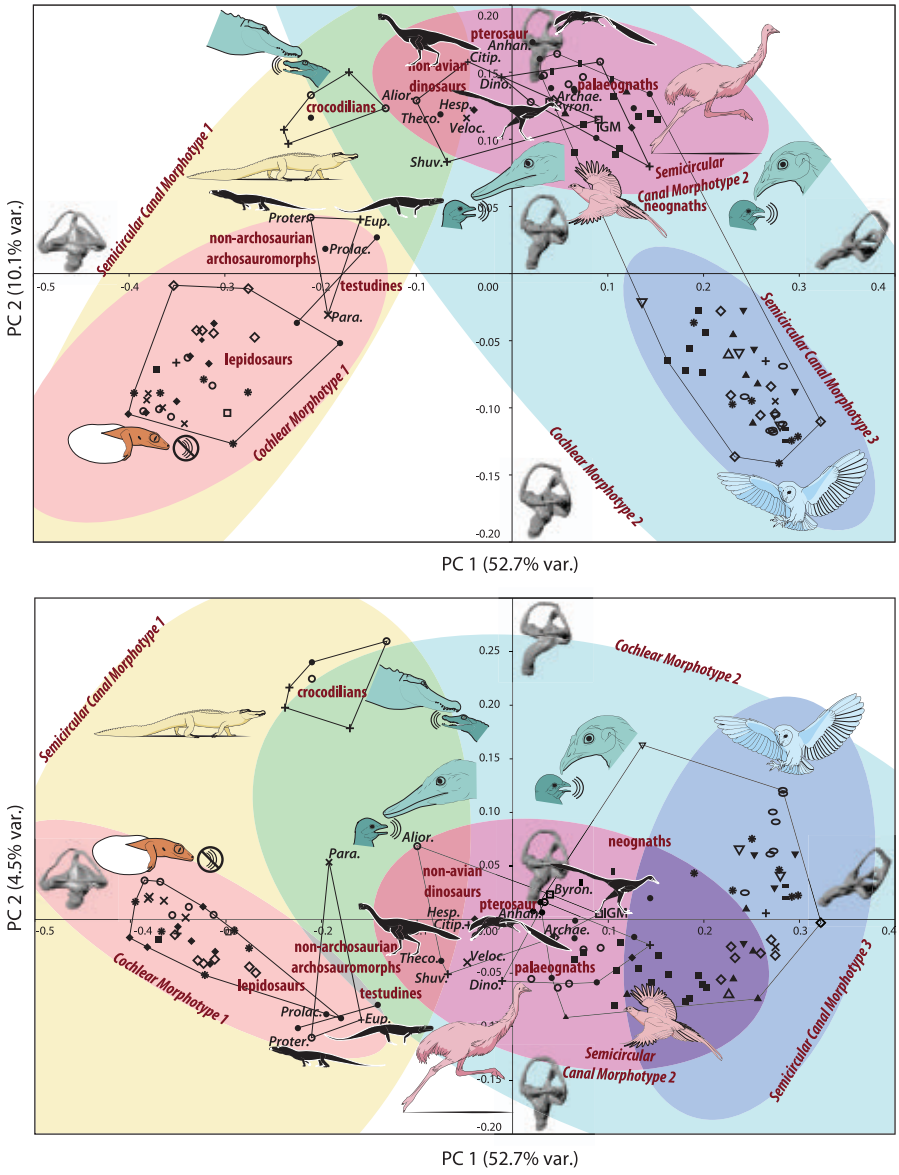


Fig. 2. PCA plots of reptilian inner ears and phylogenetic distribution of functional morphotypes. (**Top**) PCA of principal component 1 (PC1) to PC2. (**Bottom**) PCA of PC1 to PC3. Colored zones represent 99.7% confidence intervals based on behavioral data, and they also correspond to the three vestibular system morphotypes and two cochlear morphotypes discussed here [see Figs. 3 and 4 and (*12*)]. Abbreviations indicate the following taxa: Anhan., *Anhanguera santanae*; Archae., *Archaeopteryx lithographica*; Byron., *Byronosaurus jaffei*; Citip., *Citipati osmolskae*; Dino., *Dinornis robustus*; Eup., *Euparkeria capensis*; Hesp., *Hesperornis regalis*; IGM, unnamed troodontid IGM 100/3500; Para., *Parasuchus angustifrons*; Prolac., *Prolacerta broomi*; Proter., *Proterosuchus fergusi*; Shuv., *Shuvuuia deserti*; Theco., *Thecodontosaurus antiquus*; Veloc., *Velociraptor mongoliensis*. All images are original.

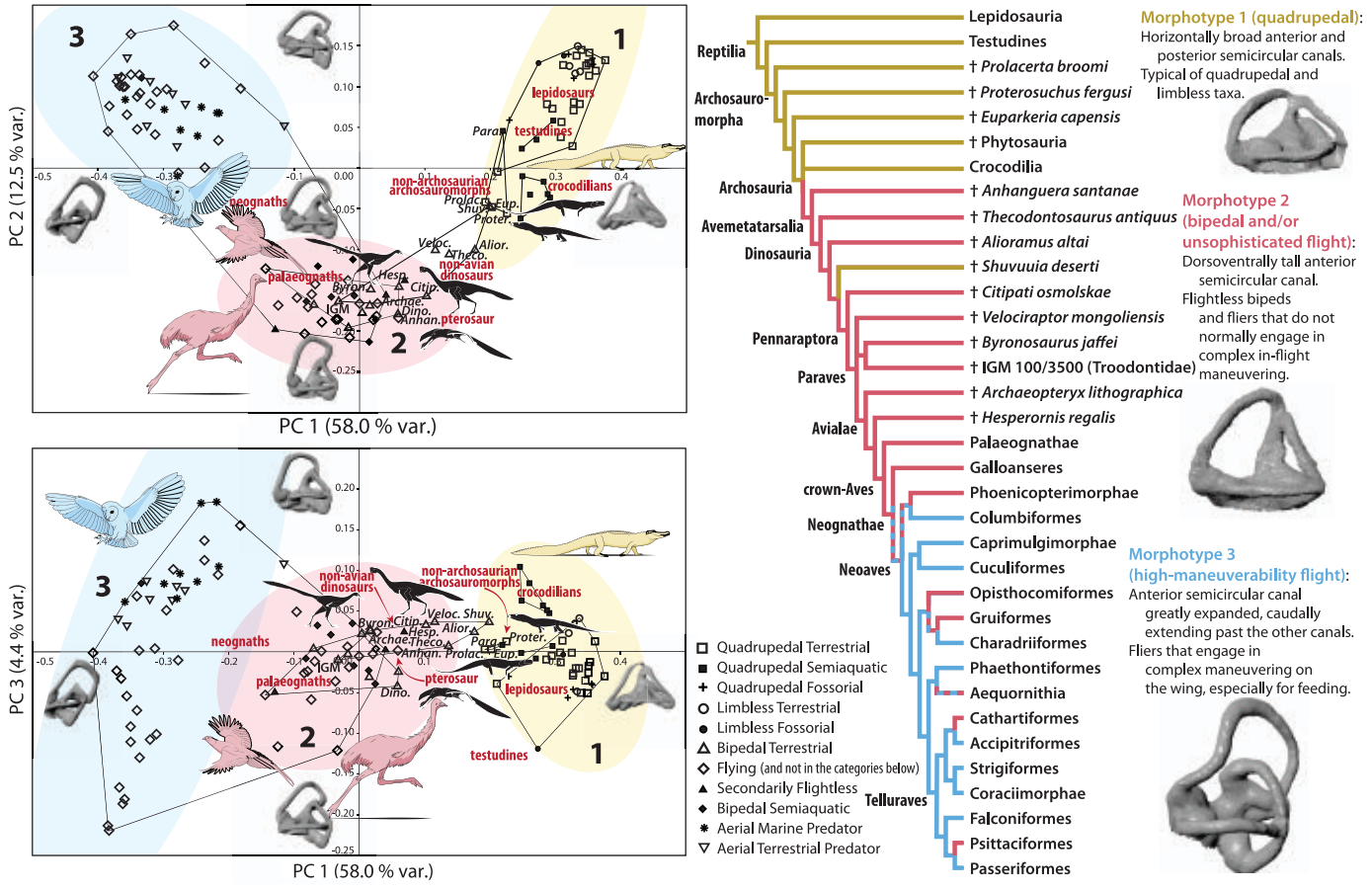


Fig. 3. PCA plots of reptile vestibular systems and phylogenetic distribution of functional morphotypes. (**Top left**) PCA of PC1 to PC2. (**Bottom left**) PCA of PC1 to PC3. Colored zones represent 99.7% confidence intervals based on behavioral data (*12*). (**Right**) Distribution of inner ear morphotypes (far right) on a tree of reptiles (right). Abbreviations are as in Fig. 2. All images are original.

aerial predation (97.6 to 99.2% and 81.3 to 84.6%), and for semiaquatic habits (99.2 and 83.2%) (figs. S60 to S67). Finally, we tested the correlation of whole-ear and vestibular shape with length ratios of forelimb elements, which have been used widely as proxies for locomotor style, particularly flight (*22*) (figs. S40 to S43). The strongest relationship (R^2 = 0.32) was with upper arm–to–lower arm ratio (humerus-to-radius), which is distinctively low in flighted taxa (*12*, *22*).

In sum, the three major PCA clusters in our analyses are best described as being distinguished by locomotor mode—one cluster comprising quadrupedal and limbless taxa (lepidosaurs and crocodilians); another comprising bipedal taxa, secondarily flightless taxa, and simple fliers; and a third consisting exclusively of taxa that engage in high-maneuverability flight behaviors, particularly related to feeding [for details of locomotor categories, see (*12*)]. This cluster of sophisticated fliers encompasses all aerial predators, including active raptorial species, aerial piscivores, and aerial insectivores. The ears of sophisticated fliers have a caudally enlarged anterior semicircular canal, where the majority of the canal is caudal to the *crus commune* and the canal often extends far caudal to the caudalmost extant of the posterior and lateral semicircular canals (Figs. 2, 3, and 5 and figs. S1 to S6 and S19 to S22).

All members of the high-maneuverability flight cluster are neoavian birds (Figs. 2 and 3 and figs. S19 to S22, S27 to S29, and S60 to S67), though some neoavian lineages have reverted to more characteristically non-neoavian flight behavior, and these cluster with flightless bipeds and less-sophisticated fliers along several axes. Reversals to an ancestral avian morphotype and behavior in Neoaves occur in a number of reluctant-flier lineages: grebes and flamingoes (Phoenicopterimorphae); loons and flightless penguins (early-diverging Aequornithes); the hoatzin (Opisthocomiformes); burst-adapted parrots (Psittaciformes); and, among raptorial bird clades, New World vultures (Cathartiformes), which are soaring scavengers that do not engage in aerial predation. The frequency of morphological reversals in disparate neoavian lineages with similar flight characteristics, along with statistically significant behavioral signals (Figs. 2 and 3, figs. S37 to S39 and S60 to S67, and table S4), demonstrates that neither phylogeny nor the shape of the surrounding skull or brain structures provides a complete explanation for the derived neoavian anterior semicircular canal. [In fact, the brains of flightless lineages, unlike the ears, are not predictably different from those of their flighted sister taxa (*8*, *34*).] Notably, morphological reversals also apparently occurred among some lineages of secondarily quadrupedal nonavian dinosaurs, which returned to the dorsoventrally short, rostrocaudally long vestibular system associated with quadrupedality (*31*, *35*, *36*).

Early-diverging archosauromorph taxa (*Prolacerta*, *Proterosuchus*, *Euparkeria*, and phytosaurs) fall, as expected, within the extant quadrupedal cluster (Figs. 2 and 3 and figs. S19 to S22) (*37*, *38*). Non-pennaraptoran dinosaurs plot in morphospace between extant quadrupeds and simple fliers. We consider this intermediate region to be indicative of nonflying bipedal locomotion, perhaps with a reasonable degree of 3D agility, turning, and

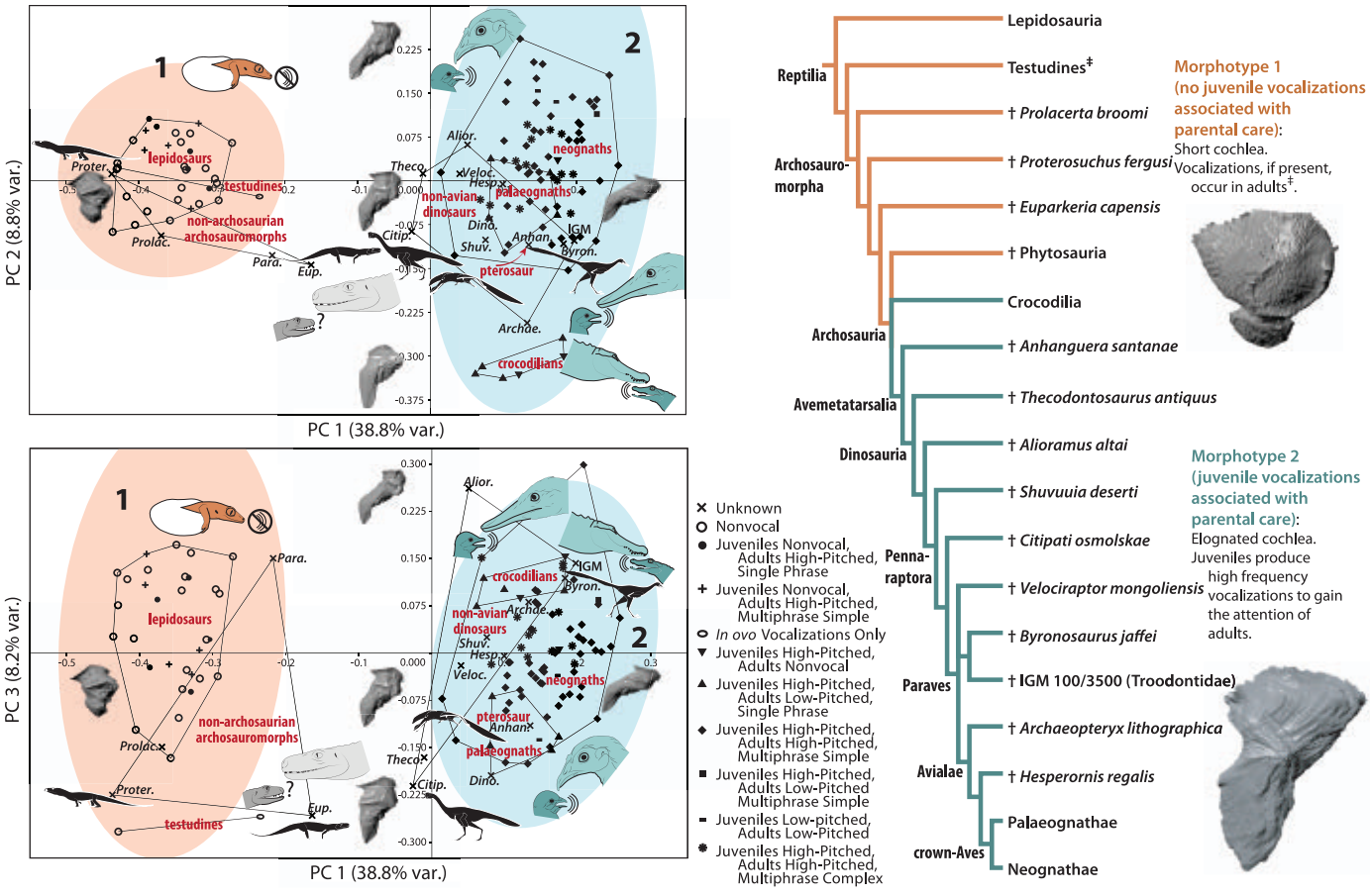


Fig. 4. PCA plots of reptile cochleae and phylogenetic distribution of functional morphotypes. (**Top left**) PCA of PC1 to PC2. (**Bottom left**) PCA of PC1 to PC3. Colored zones represent 99.7% confidence intervals based on behavioral data (*12*). (**Right**) Distribution of cochlea morphotypes (far right) on a tree of reptiles (right). Abbreviations are as in Fig. 2. All images are original.

acceleration—of the head, neck, or the entire body—and novel adaptations pertaining to the vestibulo-ocular and vestibulo-collic reflexes stabilizing the eyes and head during rapid and complex movements relative to the external environment (*8*). Notably, many small saurischian dinosaurs, aside from birds, had dexterous hands and feet and were of a size suited to climbing (*39*–*42*). *Citipati* and *Velociraptor*, representatives of early-diverging lineages within Pennaraptora—a clade distinguished by forelimbs bearing pennaceous feathers—fall just outside the simple-flier cluster (Figs. 2 and 3 and figs. S19 to S22 and S64 to S70) (*43*).

The whole-ear and vestibular-system morphologies of the pterosaur *Anhanguera* [(*11*) and comments on pterosaur inner ears in (*12*)], the Mesozoic avialans *Archaeopteryx* and *Hesperornis* (Fig. 1), and the troodontids *Byronosaurus* and IGM 100/3500 (an unnamed Cretaceous Mongolian troodontid) not only fall within the 99.7% confidence zone of bipedal taxa and simple fliers but also fall on the margin of (in the case of *Anhanguera*) or within the convex hull delineated by extant flying-bird species Figs. 2 and 3). However, discriminant function analyses identify them as flightless (figs. S60 to S63). In these pterosaurian, troodontid, and avialan taxa, the anterior semicircular canal is greatly enlarged dorsally compared with that of other avemetatarsalians and begins to extend caudal to the *crus commune* (Figs. 1 to 3 and 5 and figs. S1, S2, S5, and S19 to S22), a trend carried further in sophisticated fliers. This dorsal and caudal expansion may reflect greater specialization to locomotion in a highly 3D environment—an adaptation to either arboreality or early forms of volant activity. The notably birdlike morphology of the troodontid inner ear supports recent phylogenetic hypotheses that posit that at least some "troodontids" are not part of a clade that is sister to Dromaeosauridae as is the prevailing consensus (Deinonychosauria), but are rather part of a monophyletic or paraphyletic assemblage closer to Avialae (*44*).

The alvarezsauroid *Shuvuuia* (Figs. 2 and 3 and fig. S1) represents an exceptional reversal to the ancestral reptilian condition wherein the semicircular canals are approximately equal in dorsoventral height, and the vestibular system is rostrocaudally longer than dorsoventrally tall. The fossil specimen studied (IGM 100/977) is well preserved in three dimensions, and the braincase is uncrushed. *Shuvuuia* nests among taxa that otherwise have the dorsoventrally tall vestibular system that is typical of dinosaurs, and our ancestral shape reconstructions (Figs. 2, 3, and 5 and figs. S4, S5, and S19 to S22) indicate that the taller-than-long vestibular system is ancestral for saurischian dinosaurs. The condition seen in *Shuvuuia* is, therefore, best explained as a marked reversal among theropods, and it may be related to a suite of other adaptations to an insectivorous lifestyle (*45*–*47*). Despite being bipedal with long legs suitable for running, alvarezsauroids might have been more restricted to a more low-lying, if not fossorial, lifestyle than other theropods, owing either to the ground-dwelling habits of their eusocial insect prey or to an inability to climb with their highly modified forelimbs and comparatively uncurved pedal unguals (*48*). Outside crown birds (*49*–*51*), specialized

Table 1. Results of phylogenetic ANOVAS (1000 iterations) for behavioral signal in the semicircular canals. In the significance column, two asterisks indicate a very significant *P* value <0.01, a single asterisk indicates a significant *P* value <0.05, a period indicates a *P* value <0.1, and an empty cell indicates a *P* value >0.1. See also table S5 for additional analyses and table S2 for behavioral categories. R^2, coefficient of determination; *F*, Goodall's *F*.

Clade	R^2	*F*	*P* value	Significance
		Locomotion (all)		
Reptilia	0.02218	1.3837	0.095	.
		Locomotion (limbless versus quadrupedal)		
Lepidosauria	0.07072	2.2069	0.031	*
		Locomotion (bipedal versus quadrupedal plus limbless)		
Reptilia	0.00967	1.2006	0.227	
Archosauromorpha	0.01216	1.1078	0.261	
Archosauria	0.01127	0.98	0.361	
Avemetatarsalia	0.01446	1.174	0.225	
		Locomotion (all types): flight		
Reptilia	0.01594	2.0318	0.034	*
		Locomotion (bipedal-nonbipedal): flight		
Reptilia	0.01594	2.0218	0.035	*
Archosauromorpha	0.02002	1.8699	0.047	*
Archosauria	0.01481	1.3129	0.202	
		Locomotion (flying versus flightless)		
Reptilia	0.01985	2.4913	0.005	**
Archosauromorpha	0.02495	2.3034	0.022	*
Archosauria	0.02608	2.3032	0.028	*
Avemetatarsalia	0.02454	2.0126	0.032	*
		Locomotion (bipedal flying versus bipedal flightless)		
Dinosauria	0.02375	1.9223	0.041	*
Maniraptora	0.02433	1.92	0.044	*
Avialae	0.02379	1.7546	0.062	.
Aves	0.02387	1.7116	0.082	.
Palaeognathae	0.40336	3.3803	0.061	.
Neognathae	0.0171	1.096	0.336	
		Aerial predation (general presence or absence)		
Reptilia	0.05365	3.458	0.001	**
Archosauromorpha	0.06747	3.2195	0.001	**
Archosauria	0.06863	3.1316	0.001	**
Avemetatarsalia	0.06898	2.9264	0.001	**
Dinosauria	0.07139	2.9983	0.001	**
Maniraptora	0.07342	3.011	0.002	**
Avialae	0.07334	2.8096	0.001	**
Aves	0.06272	4.6841	0.001	**
Neognathae	0.06937	4.6958	0.001	**
Neoaves	0.07245	4.3741	0.001	**
		Semiaquatic lifestyle		
Reptilia	0.02794	3.536	0.001	**
Lepidosauria	0.01686	0.4974	0.867	
Archosauromorpha	0.04281	4.025	0.001	**
Archosauria	0.04606	4.1527	0.001	**
Avemetatarsalia	0.04892	4.115	0.001	**
Dinosauria	0.04964	4.1267	0.002	**
Maniraptora	0.05087	4.1273	0.003	**
Avialae	0.05612	4.2805	0.001	**
Aves	0.06021	4.4847	0.001	**
Neognathae	0.06822	4.6123	0.001	**
Neoaves	0.08806	5.4073	0.001	**
		Aerial predation (all types)		
Neoaves	0.11498	2.3385	0.003	**
		Aerial predation and habitat (absent versus terrestrial versus marine)		
Neoaves	0.09616	2.9257	0.003	**
		Terrestrial aerial predation prey (absent versus vertebrates versus insectivorous)		
Neoaves	0.04442	1.2785	0.192	
		Insectivory		
Neoaves	0.02912	1.6799	0.096	.
		Piscivory		
Neoaves	0.01362	0.7732	0.649	
		Active vertebrate predation		
Neoaves	0.00845	0.4771	0.936	

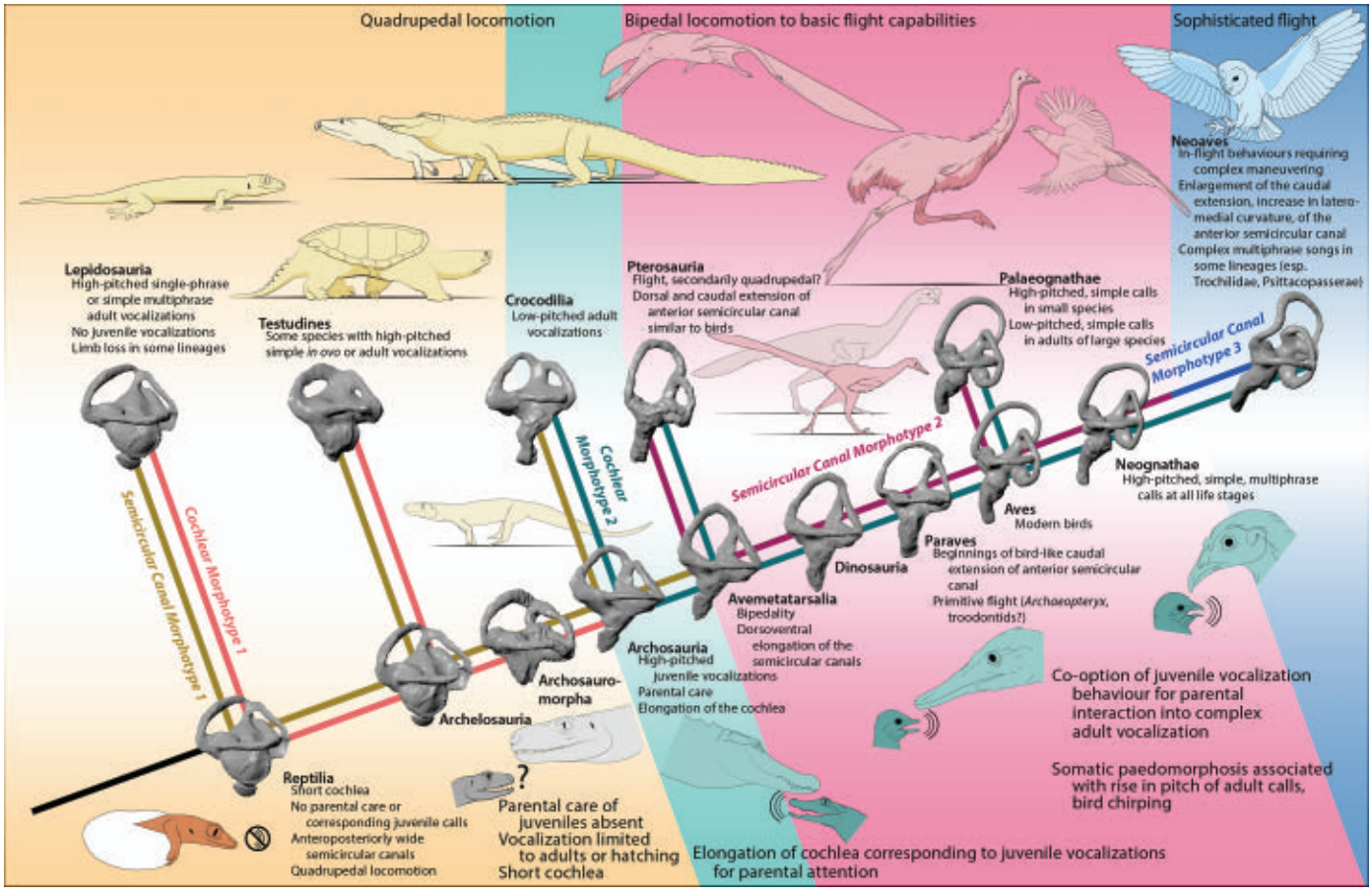


Fig. 5. Tree of reptiles showing changes in inner ear morphologies and inferred function. All images are original.

myrmecophagous tetrapod species within omnivorous, carnivorous, or otherwise non–eusocial insect–consuming clades also tend to have altered (often depressed) metabolic rates and modified movement coordination compared with their relatives (*52*, *53*).

Cochlear system

Extant reptile cochleae (Figs. 2 and 4 and figs. S19 to S22 and S68 to S73) fall into two distinct, nonoverlapping clusters identifiable in both the whole-ear and cochlea-only analyses—one consisting of taxa with short cochleae (i.e., lepidosaurs and turtles) and the other consisting exclusively of archosaurs, all of which have elongated cochleae. The positions of fossils indicate that cochlear elongation occurred only once in the evolutionary history of reptiles, among the more crownward archosauromorphs, with *Euparkeria* and the phytosaur *Parasuchus* being intermediate in morphology between early archosauromorphs and crown archosaurs. The cochleae of the troodontids and of *Shuvuuia* are exceptionally elongated with distal inflections, apparently convergent modifications that may be indicators of specialized sensation and behavior; however, these qualitatively distinctive characteristics do not have a strong effect on the overall shape analysis [fig. S1 and extended anatomical discussion in (*12*)].

Walsh *et al.* (*7*) have demonstrated that an increase in cochlear length is associated with greater hearing sensitivity, especially at high frequencies. To infer hearing sensitivity of the fossil taxa in our dataset, we interpolated their cochlear-length measurements into basicranial length–scaled and log-transformed plots of Walsh *et al.* (*7*). These results [see (*12*): interpolation into Walsh *et al.* 2009; figs. S52 to S57] indicate that most Mesozoic saurischian dinosaurs studied, and the small archosauromorph *Euparkeria*, likely had an optimal hearing frequency between 3500 and 5500 Hz and best mean hearing frequencies between 2250 and 3250 Hz, as compared with lower-pitched optimal frequencies between 1000 and 3000 Hz [comparable to extant lepidosaurs and crocodilians (*10*)] in early-diverging archosauromorphs. Optimal hearing frequencies between 3500 and 5500 Hz are in the range of extant neognathous birds (*7*). The large theropod dinosaur *Alioramus* was convergent on early archosauromorphs in having a lower-pitch optimal frequency range.

Various explanations have been proposed for the evolution of increased hearing sensitivity in vertebrates, among them prey localization (*16*), predator avoidance (*54*), and intraspecies communication (*15*). As with the vestibular system, we tested for significant differences in whole-ear and cochlear shape associated with a range of behavioral characters using phylogenetic ANOVAs (Table 2 and table S5). Dietary characteristics failed to show a significant signal. This is in keeping with the fact that there is no indication that a change in prey type or size occurred at the same time as the cochlear elongation event in archosauromorphs, as both stemward and crownward taxa are inferred to have been generalized predators on small vertebrates (*55*). We did not test for a correlation between cochlear length and predation pressure because this pressure is difficult to score effectively for extant taxa and is effectively impossible to estimate for extinct taxa, but we see little reason that early archosauromorphs, beginning with *Euparkeria*, should have faced greater predation risk than their predecessors or than contemporary lepidosauromorphs. We did find a weakly significant signal for nocturnality within Reptilia

Table 2. Results of phylogenetic ANOVAS (1000 iterations) for behavioral signal in the cochleae. As in Table 1, two asterisks indicate a very significant *P* value <0.01, a single asterisk indicates a significant *P* value <0.05, and an empty cell indicates a *P* value >0.1. See also table S6 for additional analyses and table S2 for vocalization quality categories.

Clade	R^2	*F*	*P* value	Significance
		Parental care		
Reptilia	0.0186	2.0473	0.042	*
Archosauria	0.01799	1.3737	0.179	
Aves	0.01924	1.3535	0.166	
Palaeognathae	0.23171	1.2064	0.4005	
Neognathae	0.01498	0.9582	0.341	
Neoaves	0.01448	0.8225	0.418	
		Nocturnal-diurnal		
Reptilia	0.01904	2.0965	0.012	*
Lepidosauria	0.10674	3.4653	0.002	**
Archosauria	0.01811	1.383	0.153	
Aves	0.01381	0.9665	0.444	
Palaeognathae	0.23171	1.2064	0.4025	
Neognathae	0.01621	1.0379	0.401	
Neoaves	0.01748	0.9963	0.431	
		Adult vocalizations, presence or absence		
Reptilia	0.00869	0.947	0.509	
Lepidosauria	0.03867	1.1667	0.268	
Archosauria	0.00656	0.4955	0.912	
Aves	0.00775	0.5387	0.846	
Neognathae	0.00825	0.5239	0.854	
Neoaves	0.00883	0.4987	0.848	
		Adult vocalizations, pitch		
Reptilia	0.01826	0.9953	0.475	
Lepidosauria	0.03867	1.1667	0.268	
Archosauria	0.01839	0.6934	0.84	
Aves	0.02259	0.786	0.686	
Palaeognathae	0.1832	0.8971	0.6165	
Neognathae	0.02727	0.869	0.524	
Neoaves	0.02915	0.8257	0.591	
		Adult vocalizations, complexity		
Reptilia	0.02537	0.9198	0.601	
Lepidosauria	0.05874	0.8736	0.629	
Archosauria	0.03987	1.0105	0.424	
Aves	0.04314	1.007	0.412	
Palaeognathae	0.21798	1.1149	0.3115	
Neognathae	0.0336	1.0777	0.336	
Neoaves	0.03437	0.9789	0.387	
		Juvenile vocalizations (all high-frequency), presence or absence		
Reptilia	0.03483	3.8969	0.001	**

and a strongly significant signal within Lepidosauria, which may relate to the differing sensory demands of nighttime activity. However, there was no significant activity-pattern signal in other clades. With regard to communication, presence, pitch, and complexity of adult vocalization did not show significant signal. However, the presence of high-pitched juvenile vocalization showed a strong signal (P = 0.001), and parental care also showed some signal for all reptiles (P = 0.042). Canonical variate and discriminant function analyses on cochlear geometry showed very strong predictive power for juvenile vocalization (99.2% discrimination of confusion matrices; 92.0% discrimination of jackknifed confusion matrices) and parental care (97.6 and 85.6%, respectively) (figs. S68 to S70).

The major vocal behavior differentiating archosaurs from other reptiles is the use of juvenile calls or chirps to gain the attention of, and provide spatial information to, parents or other adult caregivers (*56*, *57*). Unlike adult vocalizations, juvenile vocalizations are universally high pitched (*7*). The association of juvenile calls with parental care is borne out by statistically significant cochlear signal for both behavioral phenomena. It should be noted that the evolution of altriciality is distinct from the evolution of juvenile vocalization and parental care. According to our ancestral state reconstructions (fig. S59), ancestral archosaurs had precocial young; altriciality is restricted to some neoavian clades (*58*). Notably, many archosaurian juvenile calls are location calls (*7*), which are primarily useful for precocial offspring capable of traveling independently.

We note that an instance of cochlear elongation in mammals, on the stem of Theria, is also associated with the evolution of juvenile vocalization. Monotremes have a long but uncoiled cochlea, which appears to be the ancestral condition for Mammalia, if not Mammaliaformes (*59*). Therian mammals further elongate their cochleae into coiled

structures (*59*). Both monotremes and therians vocalize as adults; however, as juveniles, monotremes are not known to vocalize (*60*–*63*). It is in therians that high-pitched juvenile chirps to gain and maintain the attention of adults first occur (*62*), along with the therian coiled cochlea (*59*).

As the correlation of elongated cochleae with high-pitched juvenile calls seems to relate to parental care, the presence of intermediately elongated cochleae in some later-diverging archosauromorphs, such as *Euparkeria* and perhaps phytosaurs, indicates that archosaurian parental care and associated vocal behavior may have originated outside crown-group archosaurs in these later-diverging archosauromorphs (Fig. 4). Ancestral archosaurs can be inferred to have had low-pitched, closed-mouth adult vocalizations like those of modern crocodilians and some large flightless birds (*64*) [see also (*12*) and table S2]. As in modern crocodilians and non-neoavian birds, the precocial offspring of these ancestral archosaurs may have used a set of high-pitched calls to gain the attention of adults (*56*, *57*). The complex high- and low-frequency vocalizations of adult archosaurs, and adult sensitivity to wide ranges of sound frequencies not typically produced by other conspecific adults (*7*), may be a subsequent evolutionary novelty enabled by the greater hearing sensitivity brought about by cochlear elongation, which evolved in tandem with parental behaviors. Autapomorphic low-frequency hearing in *Alioramus* could reflect anomalously low-pitched juvenile calls or possibly a derived lack of parental care as seen in megapodiid and apterygid birds (*49*, *50*, *51*, *58*). Although we found little cochlear signal for diet, it is also possible that low-frequency optimal hearing assisted in the detection of sounds made by large prey [(*65*) but see (*66*)].

Discussion

Our analyses demonstrate that inner ear variation is structured in predictable ways across large evolutionary distances, and that this variation is correlated with locomotor behavior and hearing acuity in those taxa for which such data are available. The markedly avian semicircular canals of troodontids (Figs. 2 and 3 and fig. S1) suggest that these dinosaurs had locomotor control comparable to that of pterosaurs, *Archaeopteryx*, and early-diverging crown birds (Figs. 1 to 3 and 5). Many neoavian birds have a modified semicircular canal structure associated with high-maneuverability flight, aerial feeding, and predation that suggests these behaviors originated in this clade, but this group also exhibits reversals in vestibular-system structure corresponding to derived losses of these behaviors. The elongated cochleae of archosaurs and relatives such as *Euparkeria* suggest that certain complex social behaviors such as parental care may have originated as early as the Triassic and before crocodilians diverged from birds (Figs. 2, 4, and 5). Juvenile vocalizations may have been an innovation that brought about greater aural sensitivity in adult archosaurs, permitting them to respond to a wider repertoire of social calls than observed in other reptiles (*15*). High-pitched calling in adult birds may even be a paedomorphic retention of juvenile dinosaurian vocal capabilities.

REFERENCES AND NOTES

1. J. J. Videler, *Avian Flight* (Oxford Univ. Press, 2006).
2. E. D. Jarvis, *J. Ornithol.* **148**, 35–44 (2007).
3. P. L. Tyack, *Phil. Trans. R. Soc. B* **375**, 20180406 (2020).
4. D. F. A. E. Voeten *et al*., *Nat. Commun.* **9**, 923 (2018).
5. X. Wang, H. K. Tang, J. A. Clarke, *Biol. Lett.* **15**, 20190622 (2019).
6. J. A. Clarke *et al*., *Nature* **538**, 502–505 (2016).
7. S. A. Walsh, P. M. Barrett, A. C. Milner, G. Manley, L. M. Witmer, *Proc. R. Soc. B* **276**, 1355–1360 (2009).
8. S. A. Walsh *et al*., *PLOS ONE* **8**, e67176 (2013).
9. C. Grohé, Z. J. Tseng, R. Lebrun, R. Boistel, J. J. Flynn, *J. Anat.* **228**, 366–383 (2016).
10. E. G. Wever, *The Reptile Ear* (Princeton Univ. Press, 1978).
11. L. M. Witmer, S. Chatterjee, J. Franzosa, T. Rowe, *Nature* **425**, 950–953 (2003).
12. Materials and methods are available as supplementary materials.
13. J. M. Neenan *et al*., *Curr. Biol.* **27**, 3852–3858.e3 (2017).
14. B. A. Young, N. Mathevon, Y. Tang, in *Insights from Comparative Hearing Research*, vol. 49 of *Springer Handbook of Auditory Research*, C. Köppl, G. A. Manley, A. N. Popper, R. R. Fay, Eds. (Springer, 2013), pp. 323–346.
15. A. L. Vergne, M. B. Pritz, N. Mathevon, *Biol. Rev.* **84**, 391–411 (2009).
16. O. Gleich, F. P. Fischer, C. Köppl, G. A. Manley, in *Evolution of the Vertebrate Auditory System*, vol. 22 of *Springer Handbook of Auditory Research*, G. A. Manley, R. R. Fay, A. N. Popper, Eds. (Springer, 2004), pp. 224–255.
17. A. Palci, M. N. Hutchinson, M. W. Caldwell, M. S. Y. Lee, *R. Soc. Open Sci.* **4**, 170685 (2017).
18. D. G. Cerio, L. M. Witmer, *PeerJ* **7**, e7355 (2019).
19. B.-A. Bhullar, Reptilian ear 3D morphometric data, Dryad, Dataset (2021); https://doi.org/10.5061/dryad.866t1g1q4.
20. C. P. Klingenberg, *Evol. Biol.* **35**, 186–190 (2008).
21. C. Bardua, R. N. Felice, A. Watanabe, A. C. Fabre, A. Goswami, *Integr. Org. Biol.* **1**, obz016 (2019).
22. R. B. J. Benson, E. Starmer-Jones, R. A. Close, S. A. Walsh, *J. Anat.* **231**, 990–1018 (2017).
23. M. Krings, L. Rosskamp, H. Wagner, *Zoology* **126**, 82–88 (2018).
24. V. Y. Knowlton, *J. Morphol.* **121**, 179–207 (1967).
25. J. P. Bissonnette, D. M. Fekete, *J. Comp. Neurol.* **368**, 620–630 (1996).
26. L. P. Sanford *et al*., *Development* **124**, 2659–2670 (1997).
27. L. M. Yang, D. M. Ornitz, *Dev. Dyn.* **248**, 88–97 (2019).
28. E. S. Goodrich, *Studies on the Structure and Development of Vertebrates* (Dover Publications, ed. 2, 1958).
29. J. T. Richtsmeier, K. Flaherty, *Acta Neuropathol.* **125**, 469–489 (2013).
30. M. Fabbri *et al*., *Nat. Ecol. Evol.* **1**, 1543–1550 (2017).
31. G. Sobral, J. Müller, in *Evolution of the Vertebrate Ear*, vol. 59 of *Springer Handbook of Auditory Research*, J. A. Clack, R. R. Fay, A. N. Popper, Eds. (Springer, 2016), pp. 285–326.
32. S. A. Walsh, Z.-X. Luo, P. M. Barrett, in *Insights from Comparative Hearing Research*, vol. 49 of *Springer Handbook of Auditory Research*, C. Köppl, G. A. Manley, A. N. Popper, R. R. Fay, Eds. (Springer, 2013), pp. 227–261.
33. L. Codorniú, A. P. Carabajal, D. Pol, D. Unwin, O. W. M. Rauhut, *PeerJ* **4**, e2311 (2016).
34. M. E. L. Gold, A. Watanabe, *BMC Evol. Biol.* **18**, 190 (2018).
35. A. P. Carabajal, *Anat. Rec.* **295**, 2141–2156 (2012).
36. L. G. Leahey, R. E. Molnar, K. Carpenter, L. M. Witmer, S. W. Salisbury, *PeerJ* **3**, e1475 (2015).
37. G. Sobral *et al*., *R. Soc. Open Sci.* **3**, 160072 (2016).
38. E. E. Brown, R. J. Butler, M. D. Ezcurra, B. A.-S. Bhullar, S. Lautenschlager, *Palaeontology* **63**, 255–282 (2020).
39. M. J. Benton, R. Schouten, E. J. A. Drewitt, P. Viegas, *Proc. Geol. Assoc.* **123**, 210–225 (2012).
40. P. C. Sereno, R. N. Martínez, O. A. Alcober, *J. Vertebr. Paleontol.* **32**, 83–179 (2012).
41. S. H. Burch, *J. Anat.* **225**, 271–297 (2014).
42. M. S. Y. Lee, A. Cau, D. Naish, G. J. Dyke, *Science* **345**, 562–566 (2014).
43. A. M. Balanoff, G. S. Bever, M. A. Norell, *PLOS ONE* **9**, e113559 (2014).
44. F. L. Agnolin, M. J. Motta, F. Brissón Egli, G. Lo Coco, F. E. Novas, *Front. Earth Sci.* **6**, 252 (2019).
45. P. Senter, *Paleobiology* **31**, 373–381 (2005).
46. N. R. Longrich, P. J. Currie, *Cretac. Res.* **30**, 239–252 (2009).
47. J. N. Choiniere *et al*., *Science* **327**, 571–574 (2010).
48. T. R. Holtz Jr., *Curr. Biol.* **28**, R941–R943 (2018).
49. J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, G. Kirwan, Eds., Handbook of the Birds of the World Alive (Lynx Edicions, 2020); www.hbw.com.
50. P. G. Rodenwald, B. K. Keeney, Eds., Birds of the World (Cornell Lab of Ornithology, 2020); https://birdsoftheworld.org.
51. In 2020, the contents of (*49*) were moved to (*50*), but the citation of (*49*) reflects the situation at the time when this reference was accessed for use in this work.
52. T. A. Vaughan, J. M. Ryan, N. J. Czaplewski, *Mammalogy* (Jones & Bartlett Learning, 2011).
53. G. J. Clemente, G. G. Thompson, P. C. Withers, D. Lloyd, *Aust. J. Zool.* **52**, 487–503 (2004).
54. G. A. Manley, R. R. Fay, A. N. Popper, Eds., *Evolution of the Vertebrate Auditory System*, vol. 22 of *Springer Handbook of Auditory Research* (Springer, 2004).
55. S. J. Nesbitt *et al*., *Nature* **544**, 484–487 (2017).
56. R. B. Brua, in *Avian Incubation: Behaviour, Environment, and Evolution*, D. C. Deeming, Ed. (Oxford Univ. Press, 2002), pp. 88–99.
57. P. Senter, *J. Herpetol.* **42**, 354–360 (2008).
58. J. M. Starck, R. E. Ricklefs, in *Avian Growth and Development: Evolution within the Altricial-Precocial Spectrum*, J. M. Starck, R. E. Ricklefs, Eds. (Oxford Univ. Press, 1998), pp. 3–30.
59. Z.-X. Luo, I. Ruf, J. A. Schultz, T. Martin, *Proc. R. Soc. B* **278**, 28–34 (2011).
60. G. Bennett, *Trans. Zool. Soc. Lond.* **1**, 229–258 (1835).
61. B. H. Brattstrom, *J. Mammal.* **54**, 50–70 (1973).
62. J. F. Eisenberg, L. R. Collins, C. Wemmer, *Z. Tierpsychol.* **37**, 379–399 (1975).
63. T. Fitch, in *Encyclopedia of Language and Linguistics*, K. Brown, Ed. (Elsevier, ed. 2, 2006) pp. 115–121.
64. T. Riede, C. M. Eliason, E. H. Miller, F. Goller, J. A. Clarke, *Evolution* **70**, 1734–1746 (2016).
65. P. E. Brown, P. M. Narins, A. D. Grinnell, *Amer. Neurosci. Abstracts* **10**, 1150 (1984).
66. G. Koay, K. S. Bitter, H. E. Heffner, R. S. Heffner, *Hear. Res.* **171**, 96–102 (2002).

ACKNOWLEDGMENTS

We thank B. Rubidge, S. Walsh, and P. Barrett for access to specimens and data. J. A. Gauthier, S. Nesbitt, and R. Benson provided considerable insight during the early stages of this work. **Funding:** B.-A.S.B. and M.H. were funded by the Department of Earth and Planetary Sciences at Yale University and the Yale Institute for Biospheric Studies. M.A.N. and E.A.H. were funded by the American Museum of Natural History and the National Science Foundation. **Author contributions:** M.H. and B.-A.S.B. conceived the project. M.H., E.A.H., and B.-A.S.B. performed CT scanning and processed data. M.A.N. collected and directed the conservation and preparation of several of the specimens scanned and analyzed. M.H., E.A.H., M.A.N., and B.-A.S.B. planned and discussed analyses. M.H. performed analyses. All authors wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All 3D data used in the analyses herein and all additional input and output files used in and produced by the analyses are archived with Data Dryad (*19*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/601/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S77
Tables S1 to S10
References (*67*–*198*)
MDAR Reproducibility Checklist

4 March 2020; accepted 9 March 2021
10.1126/science.abb4305

REPORTS

PALEONTOLOGY

Evolution of vision and hearing modalities in theropod dinosaurs

Jonah N. Choiniere[1]†, James M. Neenan[1,2]†, Lars Schmitz[3,4], David P. Ford[1], Kimberley E. J. Chapelle[1,5], Amy M. Balanoff[5,6], Justin S. Sipla[7], Justin A. Georgi[8], Stig A. Walsh[9,10], Mark A. Norell[5], Xing Xu[11,12], James M. Clark[13], Roger B. J. Benson[1,14]*

Owls and nightbirds are nocturnal hunters of active prey that combine visual and hearing adaptations to overcome limits on sensory performance in low light. Such sensory innovations are unknown in nonavialan theropod dinosaurs and are poorly characterized on the line that leads to birds. We investigate morphofunctional proxies of vision and hearing in living and extinct theropods and demonstrate deep evolutionary divergences of sensory modalities. Nocturnal predation evolved early in the nonavialan lineage Alvarezsauroidea, signaled by extreme low-light vision and increases in hearing sensitivity. The Late Cretaceous alvarezsauroid *Shuvuuia deserti* had even further specialized hearing acuity, rivaling that of today's barn owl. This combination of sensory adaptations evolved independently in dinosaurs long before the modern bird radiation and provides a notable example of convergence between dinosaurs and mammals.

Sensory specializations are some of the most distinctive vertebrate innovations (*1*, *2*) and are common as adaptations to low-light activity in birds and mammals. Facultative nocturnal behaviors of living birds, such as tactile probing and dabbling in some water birds, require limited sensory anatomical changes (*3*). However, most nocturnal birds have conspicuous modifications of the visual system, and specialized nocturnal foragers of active prey combine adaptations of both vision and hearing (*4*). Such adaptations enable distinctive foraging strategies such as the use of precise sound localization and low-light (scotopic) vision in the barn owl (*Tyto alba*), which can hunt in complete darkness (*5*). These sensory adaptations leave clear skeletal signatures that should be evident in fossils. Nevertheless, sensory evolution in birds and their theropod stem lineage is poorly understood [but see, e.g., (*6*–*9*)]. This is a substantial shortcoming in our understanding of dinosaurian biology and of the structure of Mesozoic ecosystems.

To evaluate the evolution of vison and hearing in extinct theropods, we studied skeletal proxies for two sensory systems: the scleral ossicle ring (hereafter "scleral ring") of the eye for vision and the endosseous cochlear duct of the bony labyrinth for hearing. The scleral ring is embedded in the eyeball surface in numerous living and extinct amniotes. Nocturnal species typically have wider ring apertures, reflecting larger pupil sizes that increase light sensitivity (*9*). The endosseous cochlear duct is intimately linked to hearing performance (e.g., sensitivity and frequency range) because it houses the basilar papilla or cochlea (*10*). The elongate mammalian cochlea has been an area of continued evolutionary innovation [e.g., (*11*, *12*)], demonstrating the

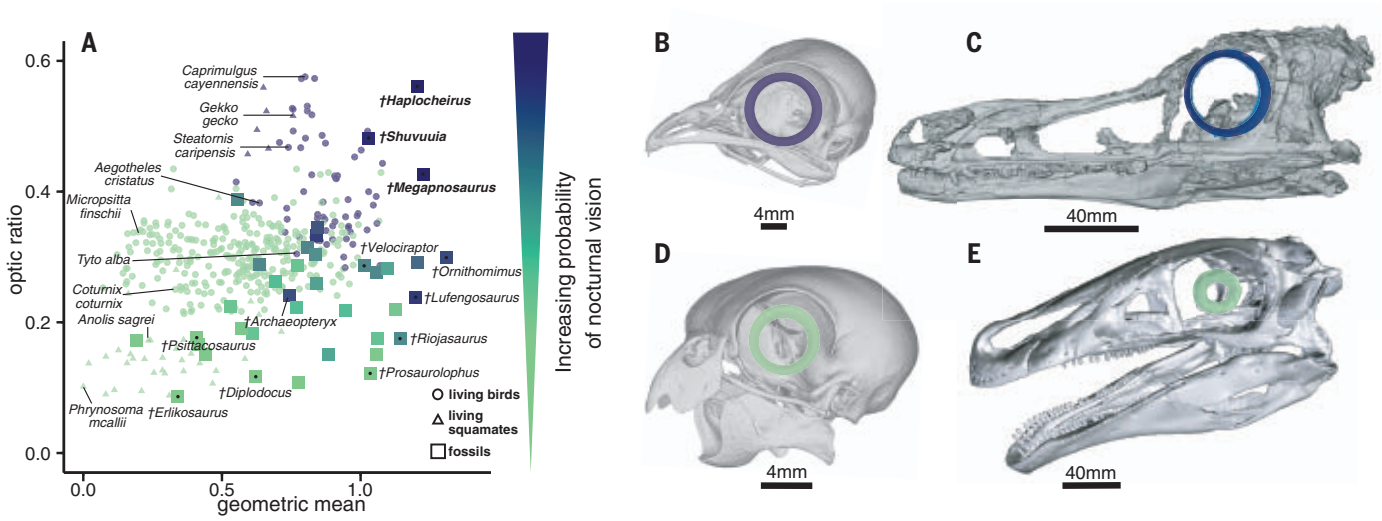


Fig. 1. Optic ratio by geometric mean of optic measurements for living birds, squamates, and nonavialan theropods. (**A**) Scatterplot of optic ratio versus geometric mean of orbit and scleral ring measurements for extant birds and reptiles and extinct species. See data S3. (**B** to **E**) Scleral ossicle ring anatomy in (B) *Aegotheles cristatus*, (C) *H. sollers*, (D) *Micropsitta finchii*, and (E) *Erlikosaurus andrewsi* (*29*). Blue indicates nocturnality, green indicates non-nocturnality, and color gradient indicates the posterior probability of nocturnality (data S3) (*19*). † indicates extinct species.

[1]Evolutionary Studies Institute, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg 2000, South Africa. [2]Oxford University Museum of Natural History, University of Oxford, Parks Road, Oxford OX1 3PW, UK. [3]W.M. Keck Science Department, Claremont McKenna, Scripps, and Pitzer Colleges, 925 N Mills Ave., Claremont, CA 91711, USA. [4]Dinosaur Institute, Los Angeles County Museum of Natural History, 900 Exposition Boulevard, Los Angeles, CA 90007, USA. [5]Division of Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA. [6]Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 North Charles St., Baltimore, MD 21218, USA. [7]Department of Anatomy and Cell Biology, Carver College of Medicine, University of Iowa, 51 Newton Road, 100 Bowen Science Building, Iowa City, IA 52242, USA. [8]Department of Anatomy, College of Graduate Studies, Midwestern University, 19555 N 59th Avenue, Glendale, AZ 85308, USA. [9]Department of Natural Sciences, National Museums Scotland, Chambers Street, Edinburgh EH1 1JF, UK. [10]School of Geosciences, University of Edinburgh, Grant Institute, Hutton Road, Edinburgh EH9 3FE, UK. [11]Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, 142 Xizhimenwai Street, Beijing 100044, China. [12]Chinese Academy of Sciences Center for Excellence in Life and Paleoenvironment, Beijing 100044, China. [13]Department of Biological Sciences, The George Washington University, 2029 G St. NW, Washington, DC 20052, USA. [14]Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK.

†These authors contributed equally to this work.

*Corresponding author. Email: roger.benson@earth.ox.ac.uk

importance of hearing in their foraging strategies. Bird species that rely on auditory cues for foraging, such as owls, have an elongate endosseous cochlear duct, alongside other anatomical modifications (*13*).

Our analyses show visual and auditory adaptations to nocturnality in alvarezsauroids, an enigmatic theropod lineage that spanned the Late Jurassic to Late Cretaceous. Early branching alvarezsauroids, such as *Haplocheirus sollers*, retain generalized theropod features that suggest a relatively unspecialized predatory ecology (*14*). However, late-branching alvarezsauroids such as *Shuvuuia deserti* have a curious mixture of derived traits, including birdlike skulls, cursorial hindlimbs, and short, functionally monodactyl forelimbs interpreted as adaptations for scratch digging (*15*, *16*). These features are the source of continued speculation and debate about alvarezsauroid paleoecology (*15*–*18*).

Our digital reconstructions of the scleral rings of the alvarezsauroids *Haplocheirus* and *Shuvuuia* show a proportionally large eye with an extremely wide aperture (Fig. 1 and figs. S1 to S5) (*19*). Phylogenetic flexible discriminant analysis of the scleral ring and orbit morphology, extended from previous analyses (*9*), has a mean accuracy of 92.0% for the classification of extant species as nocturnal or non-nocturnal. Among theropods, the highest posterior probabilities for nocturnal vision ($pp_{nocturnal}$) are for *Haplocheirus* ($pp_{nocturnal} > 0.99$), *Shuvuuia* ($pp_{nocturnal} > 0.87$), and the coelophysoid *Megapnosaurus* ($pp_{nocturnal} > 0.99$) (Fig. 1 and table S1). These have morphologies similar to those of birds that have specialized low-light visual systems (Fig. 1) but are different from those of other theropods, including many for which nocturnality was previously inferred at a much less stringent probability threshold ($pp_{nocturnal} > pp_{diurnal}$ and $pp_{cathemeral}$) (*6*).

Micro–computed tomography (CT) reveals anatomical specialization of the endosseous cochlear duct in *Shuvuuia*, similar to the barn owl (*T. alba*), with large duct diameters and a proximodistally elongate morphology, curving posteromedially under the brain cavity. These anatomies are unlike those of almost all other theropods. Cochlear duct elongation imposes spatial constraints on the labyrinth in both taxa (Figs. 2 and 3), which have low, broad semicircular canals compared with their evolutionary relatives (*20*) (Fig. 3B). Both taxa also show two well-defined laminae that extend along the length of the endosseous cochlear duct, medially (primary bony lamina) and laterally (secondary bony lamina) (Fig. 2). Similar laminae in mammals are attachment sites for the basilar membrane, which supports the basilar papilla.

Comparison of endosseous cochlear duct length to braincase height across 88 extant

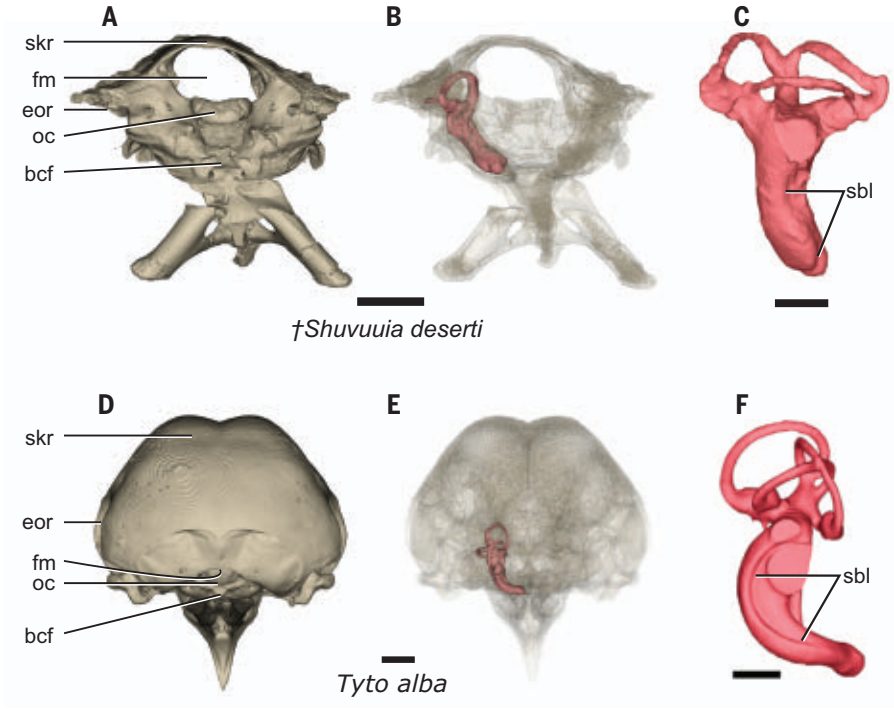


Fig. 2. Comparative anatomy of the endosseous cochlear duct of *S. deserti* (uncrushed adult specimen IGM100/1304) and the extant barn owl *T. alba*. (**A** to **C**) *S. deserti* uncrushed adult specimen (IGM100/1304). (**D** to **F**) Extant barn owl *T. alba*. (A and D) Posterior views of braincases showing external anatomy. (B and E) Transparent CT renderings of braincase in posterior view, showing endosseous labyrinths. (C and F) Digital reconstructions of the endosseous labyrinths in lateral view. Scale bars, 5 mm (B and E) and 2.5 mm (C and F). bcf, braincase floor; eor, external otic region; fm, foramen magnum; oc, occipital condyle; sbl, secondary bony lamina; skr, skull roof. † indicates extinct species.

birds corroborates our interpretation of relative duct elongation as related to nocturnality and other foraging traits (Fig. 3, table S2, and figs. S6 to S10) (*19*). Owls (Strigiformes); large-bodied, nocturnal Strisores (e.g., owlet nightjars, pootoos, frogmouths, and the echolocating oilbird); and some other taxa have moderately or greatly elongated ducts (Fig. 3), whereas vocal learners do not (fig. S6 and table S3). This suggests that duct elongation is an adaptation for auditory foraging, contradicting the hypothesis that it evolved to facilitate intraspecific communication (*10*). Woodpeckers, which use auditory foraging of concealed insects, show limited to moderate duct elongation (table S3). Some water-associated birds, including gannets, kingfishers, and cormorants, also show moderate duct elongation that deserves further investigation (table S3). Nightjars, which are low-light predators of aerial prey in generally open environments, show only limited duct elongation (table S3).

Elongation of the endosseous cochlear duct also occurs among Mesozoic theropods, including some predatory groups (Troodontidae, Tyrannosauridae, and Dromaeosauridae); in the secondarily herbivorous Therizinosauria; and in alvarezsauroids, including the Jurassic to Early Cretaceous taxa *Haplocheirus* and *Xiyunykus* (Fig. 3). The barn owl (*T. alba*), *Shuvuuia* , and an undescribed troodontid [Mongolian Institute for Geology (IGM) 100/1126] are outliers that demonstrate auditory specialization, even compared with other owls, strisorans, and Mesozoic theropods (Fig. 2) (*20*, *21*). Proportional reduction of the ensosseous cochlear duct occurs in oviraptorosaurs and ornithomimosaurs, which are hypothesized to be herbivores or omnivores (*22*, *23*).

The proportionally large eyes and scleral ring apertures of *Haplocheirus* and *Shuvuuia* indicate nocturnal visual adaptation. These represent some of the earliest and latest alvarezsauroids, suggesting that nocturnal visual capabilities were widespread in that group. The early appearance and phylogenetic retention of nocturnal visual adaptations in alvarezsauroids contrasts with inferences of equivocal or diurnal activity patterns in most other theropods, including early birds (Avialae) (Fig. 4). This distribution suggests a deep evolutionary divergence of activity patterns among theropods, with alvarezsauroids becoming nocturnal

visual specialists (Fig. 4). The hearing-related anatomy of *Shuvuuia*, comprising an exceptionally elongate endosseous cochlear duct and the presence of a secondary bony lamina, is highly comparable to that of the barn owl among extant birds (Fig. 2) and corroborates the hypothesis of specialized nocturnal foraging in alvarezsauroids. This mode of life is rare in birds (*4*), and we also find that it was rare in nonavialan theropods, contrasting with the relatively high incidence of nocturnal foraging in mammals (*24*).

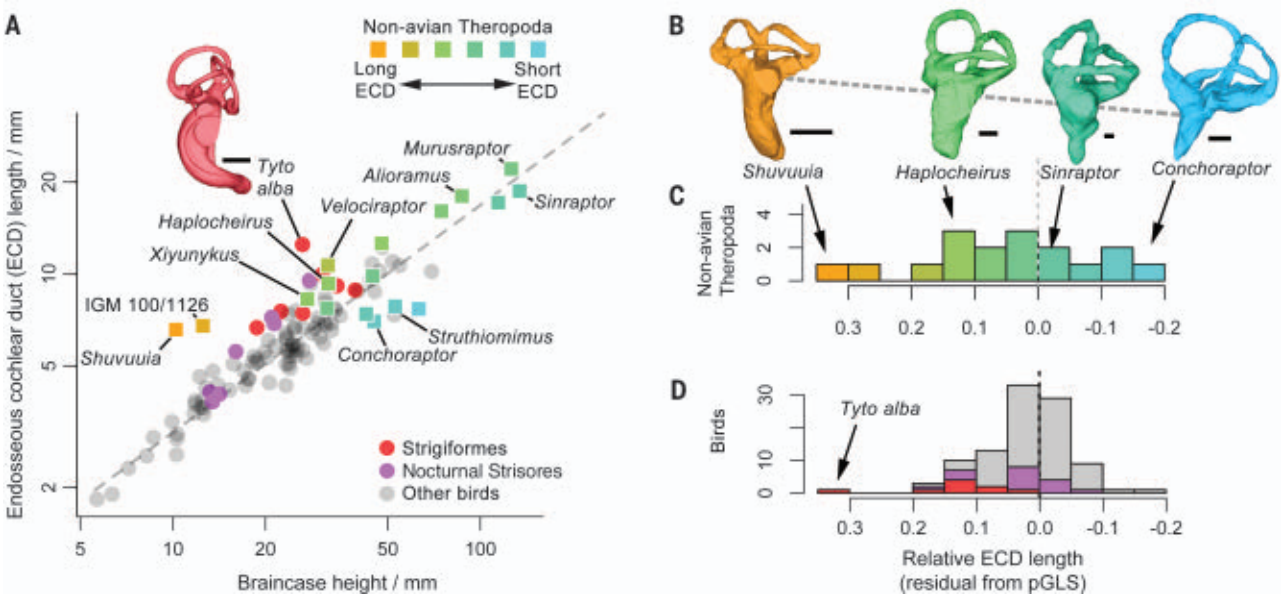


Fig. 3. Endosseous cochlear duct (ECD) length in a sample of birds and nonavialan theropod dinosaurs. (**A**) ECD length versus braincase height for extant birds [regression line $\log_{10}$(ECD length) = −0.265 + 0.744 × $\log_{10}$(braincase height); $P < 0.0001$; λ (phylogenetic signal) = 0.89; $N = 88$]. Residuals indicate relative duct length, indicated with a color gradient for nonavialan theropods. (**B**) Endosseous labyrinths of selected nonavialan theropods colored according to residual duct length. The gray line follows the fenestra ovalis. (**C** and **D**) Histograms showing relative duct lengths among sampled extant birds (C) and nonavialan theropods (D). Scale bars, 2.5 mm. pGLS, phylogenetic generalized least squares regression.

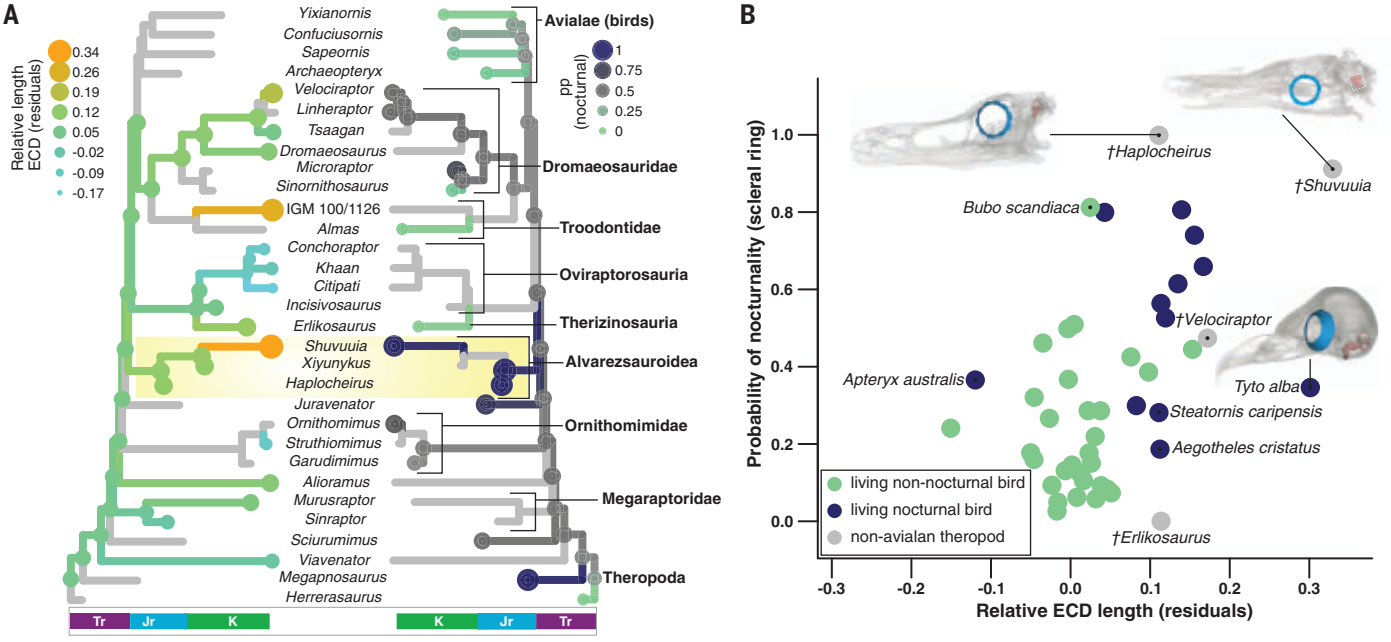


Fig. 4. ECD length residuals and $pp_{nocturnal}$ for selected theropods. (**A**) Phylogeny of study sample mapping duct-length residuals and posterior probability of nocturnality in Mesozoic theropods. (**B**) $pp_{nocturnal}$ versus relative ECD length (residuals) for extant and extinct theropods. Skulls (clockwise): *H. sollers* [Institute of Vertebrate Paleontology and Paleoanthropology (IVPP) V15988], *S. deserti* (IGM100/0977; labyrinth from IGM100/1304), and *T. alba*. See also figs. S8 and S9. † indicates extinct species.

Even moderate elongation of the endosseous cochlear duct is rare in our broad sample of extant birds (Fig. 3D). It occurs most frequently (though not exclusively) in owls and large-bodied nightbirds such as frogmouths, owlet nightjars, and pootoos (table S3). These are predominantly nocturnal predators of highly mobile and often cryptic, ground-based, and near-perch prey (*3*). Many of these taxa forage for prey items that were also available to extinct theropod dinosaurs (including terrestrial invertebrates and small vertebrates) by ground running and pouncing from a low perch (e.g., frogmouths) or descent from quartering flight (e.g., owls) (*3*). Increased hearing ability in these groups is an adaptation that improves the chances of locating and successfully striking prey in low light (*5*). For example, the specialized hearing of barn owls, which have highly elongated duct lengths, even compared with other strigiforms, enables them to catch prey with high accuracy in complete darkness on the basis of acoustic cues alone (*25*, *26*).

In contrast to its scarcity in extant birds, moderate elongation of the endosseous cochlear duct is common in our sample of nonavialan theropods, occurring in all alvarezsauroids; the hypercarnivorous Tyrannosauridae, Dromaeosauridae, and Troodontidae; and the secondarily herbivorous Therizinosauria (Fig. 3). Theropods, therefore, evolved increased hearing ability more frequently than in modern birds, although Oviraptoridae and Ornithomimidae have shorter ducts (Fig. 3) and may have relied on other senses.

Previous hypotheses of alvarezsauroid ecology (*15*, *16*, *27*) noted a distinctive combination of ecomorphological traits in geologically younger taxa such as *Shuvuuia*. These traits include dental reduction and a slender mandible, combined with fossorial adaptations of its forelimbs, and have been used to support hypotheses of specialized predation on colonial insects (*15*, *16*). Our study additionally indicates low-light vision and specialized auditory capabilities, traits that are widespread among mammals but rare in dinosaurs, including birds. Myrmecophagous mammals have previously been suggested as analogs for alvarezsauroids (*15*, *16*). However, we note that nocturnality, excellent hearing, and digging forelimbs occur much more widely in mammals that span a broad range of ecologies. Moreover, dental or mandibular reduction in mammals may not be a functional analog of dental or mandibular reduction in birds and other dinosaurs; extant birds with proportionally small crania or mandibles, such as galliforms, can have wide-ranging diets, and birds that consume small vertebrates can subdue their prey without the use of teeth. We therefore suggest that derived alvarezsauroids such as *Shuvuuia* could have consumed a wide range of nocturnal invertebrates and potentially small vertebrates, and that fossorially adapted forelimbs had versatile functions, allowing excavation of shallow-burrowing or crevice-inhabiting animals from the substrate or vegetation.

Combined visual and auditory specializations for nocturnality evolved independently in mammals, birds, and, as reported in this study, nonavialan dinosaurs, providing an example of convergent sensory evolution in vertebrates. The sensory paleoecology of dinosaurs remains poorly understood in general. Nevertheless, our findings provide information on deep evolutionary divergences of activity patterns among nonavialan theropods and strong evidence for nocturnal specialization through 95 million years of alvarezsauroid evolution. Many living animals are active at night, but nocturnal communities remain poorly studied both today (*28*) and in the past. Identifying specialized night foragers such as *Shuvuuia* highlights the occurrence of diel partitioning among predators in Mesozoic terrestrial ecosystems. It indicates that richly sampled paleoecosystems such as the Djadokhta Formation hosted previously unrecognized nocturnal and diurnal subcommunities, and expands our understanding of the structure of past ecosystems and of the ecological traits of theropod dinosaurs.

REFERENCES AND NOTES

1. J. Müller, C. Bickelmann, G. Sobral, *Annu. Rev. Earth Planet. Sci.* **46**, 495–519 (2018).
2. G. Schlosser, *Integr. Comp. Biol.* **58**, 301–316 (2018).
3. G. Martin, *Birds by Night* (A&C Black, 2010).
4. G. R. Martin, *J. Ornithol.* **153** (S1), 23–48 (2012).
5. R. S. Payne, W. Drury, *Nat. Hist.* **67**, 316–323 (1958).
6. L. Schmitz, R. Motani, *Science* **332**, 705–708 (2011).
7. D. K. Zelenitsky, F. Therrien, R. C. Ridgely, A. R. McGee, L. M. Witmer, *Proc. Biol. Sci.* **278**, 3625–3634 (2011).
8. Y. Wu, H. Wang, *Proc. Biol. Sci.* **286**, 20182185 (2019).
9. L. Schmitz, R. Motani, *Vision Res.* **50**, 936–946 (2010).
10. S. A. Walsh, P. M. Barrett, A. C. Milner, G. Manley, L. M. Witmer, *Proc. Biol. Sci.* **276**, 1355–1360 (2009).
11. E. G. Ekdale, *J. Anat.* **228**, 324–337 (2016).
12. Z.-X. Luo, J. A. Schultz, E. G. Ekdale, in *Evolution of the Vertebrate Ear*, J. A. Clack, R. R. Fay, A. N. Popper, Eds. (Springer, 2016), pp. 139–174.
13. O. Gleich, U. Langemann, *Hear. Res.* **273**, 80–88 (2011).
14. J. N. Choiniere *et al*., *Science* **327**, 571–574 (2010).
15. P. Senter, *Paleobiology* **31**, 373–381 (2005).
16. N. R. Longrich, P. J. Currie, *Cretac. Res.* **30**, 239–252 (2009).
17. F. L. Agnolin, J. E. Powell, F. E. Novas, M. Kundrat, *Cretac. Res.* **35**, 33–56 (2012).
18. J. N. Choiniere, J. M. Clark, M. A. Norell, X. Xu, *Am. Mus. Novit.* **3816**, 1–44 (2014).
19. Materials and methods are available as supplementary materials.
20. R. B. J. Benson, E. Starmer-Jones, R. A. Close, S. A. Walsh, *J. Anat.* **231**, 990–1018 (2017).
21. O. Gleich, R. J. Dooling, G. A. Manley, *Naturwissenschaften* **92**, 595–598 (2005).
22. L. E. Zanno, P. J. Makovicky, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 232–237 (2011).
23. P. M. Barrett, *Annu. Rev. Earth Planet. Sci.* **42**, 207–230 (2014).
24. K. E. Jones *et al*., *Ecology* **90**, 2648–2648 (2009).
25. R. S. Payne, *J. Exp. Biol.* **54**, 535–573 (1971).
26. M. Konishi, *Am. Sci.* **61**, 414–424 (1973).
27. A. Perle, L. M. Chiappe, R. Barsbold, J. M. Clark, M. Norell, *Am. Mus. Novit.* **3105**, 1–29 (1994).
28. K. J. Gaston, *Am. Nat.* **193**, 481–502 (2019).
29. S. Lautenschlager, L. M. Witmer, P. Altangerel, L. E. Zanno, E. J. Rayfield, *J. Vertebr. Paleontol.* **34**, 1263–1291 (2014).
30. R. Benson, L. Schmitz, J. Choiniere, J. Neenan, K. Chapelle, Supplementary models, data and scripts for Choiniere et al 2021, OSF, 27 April 2021; https://doi.org/10.17605/OSF.IO/TEQ73.

ACKNOWLEDGMENTS

We acknowledge the Radiology Department of the Stony Brook University Hospital, S. Lautenschlager, A. Paulina-Carabajal, L. Filippi, P. Currie, and P. Barrett for scan data (NERC NE/E008380/1); J. White, J. Cooper, J. Hinshaw, M. Lowe, and B. Marks for access to extant bird specimens; Volume Graphics, Thermo Fisher Scientific (Amira), and Materialise (Mimics) for facilitating home licenses; and C. Carr for discussions of hearing physiology. **Funding:** Funding was received from the European Union's Horizon 2020 research and innovation program 2014–2018, starting grant (R.B.J.B., 677774); Leverhulme Trust (J.M.N., ECF-2017-360); NERC (S.A.W., NE/H012176/1); National Research Foundation of South Africa (J.N.C.; 118994, 98800, and 95449); Kalbfleish Fellowship through the Richard Gilder Graduate School of the American Museum of Natural History (J.N.C. and K.E.J.C.); the National Natural Science Foundation of China (X.X., 41688103); Strategic Priority, Research Program of the Chinese Academy of Sciences (X.X., XDB18000000); National Science Foundation of the USA (A.M.B., M.A.N., and J.M.C.; DEB 1801224 and EAR 1636753); Paleontological Scientific Trust (J.N.C. and K.E.J.C.); DSI/NSF Center of Excellence Paleosciences (J.N.C. and K.E.J.C.); Jurassic Foundation (J.S.S. and J.A.G.); and Institute for Global/Local Action & Study, Pitzer College (L.S.). **Author contributions:** Conceptualization: J.N.C., J.M.N., R.B.J.B., and L.S.; Methodology: J.N.C., J.M.N., L.S., D.P.F., K.E.J.C., and R.B.J.B.; Software: J.M.N., L.S., and R.B.J.B.; Validation: J.N.C., L.S., and R.B.J.B.; Formal analysis: R.B.J.B. and L.S.; Investigation: J.N.C., J.M.N., L.S., A.M.B., J.S.S., J.A.G., S.A.W., and R.B.J.B.; Resources: J.N.C., J.M.N., L.S., A.M.B., J.S.S., J.A.G., S.A.W., M.A.N., X.X., J.M.C., and R.B.J.B.; Data Curation: J.N.C., L.S., and R.B.J.B.; Writing – original draft preparation: J.N.C., J.M.N., and R.B.J.B.; Writing – review and editing: all authors; Visualization: J.N.C., J.M.N., L.S., D.P.F., K.E.J.C., and R.B.J.B.; Supervision: J.N.C., M.A.N., X.X., J.M.C., and R.B.J.B.; Project administration: J.N.C., J.M.N., and R.B.J.B.; Funding acquisition: J.N.C., J.M.N., L.S., M.A.N., X.X., J.M.C., and R.B.J.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data described in this paper, including measurements, models, and provenance information for fossil specimens, are available in the supplementary materials and at OSF (*30*). Models used in this study are also available at the following Morphosource links. For new models used in this study (N = 33 living birds; 17 nonavialan theropods): www.morphosource.org/projects/0000C1096. For previously published models (N = 55) (*20*) used in this study: www.morphosource.org/Detail/ProjectDetail/Show/project_id/377.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/610/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S3
Data S1 to S12
References (*31*–*47*)
MDAR Reproducibility Checklist

16 September 2020; accepted 12 March 2021
10.1126/science.abe7941

MATERIALS SCIENCE

Reversible fusion and fission of graphene oxide–based fibers

Dan Chang[1], Jingran Liu[2], Bo Fang[1], Zhen Xu[1], Zheng Li[1]*, Yilun Liu[2]*, Laurence Brassart[3]†, Fan Guo[1], Weiwei Gao[1], Chao Gao[1]*

Stimuli-responsive fusion and fission are widely observed in both bio-organizations and artificial molecular assemblies. However, the design of a system with structure and property persistence during repeated fusion and fission remains challenging. We show reversible fusion and fission of wet-spun graphene oxide (GO) fibers, in which a number of macroscopic fibers can fuse into a thicker one and can also separate into original individual fibers under stimulation of solvents. The dynamic geometrical deformation of GO fiber shells, caused by solvent evaporation and infiltration, is the key to the reversible fusion-fission cycles. This principle is extended to implement flexible transitions between complex fiber assemblies and the inclusion or expulsion of guest compounds.

Fusion and fission behaviors have been extensively studied in biology, chemical engineering, and theoretical physics to understand cellular processes, develop morphological events of artificial assemblies, and create multimetallic compounds. Fusion and fission of lipid/surfactant/small organic molecules/polymer micelles and vesicles are usually triggered by either introducing salts, surfactants, ions, oxidants, and reductants or applying ultraviolet and visible light to change the interactions inside bilayer membranes, as well as by dissolving additives such as saccharides to alter the osmotic pressure between the aqueous interior of vesicles and the bulk solution (*1*–*6*). Similar behaviors of metal particles and clusters are also induced by heat or cluster deposition (*7*–*9*). Although progress has been made on fusion-fission of artificial vesicles and nanoparticles, reversible fusion and fission are still difficult to realize, mainly because of the irreversible physical or chemical changes at the interface between individual assemblies. The exploration of reversible and controllable fusion and fission would inspire the development of stimuli-responsive materials, which show promise in deriving dynamic transformable systems and structural materials with customized fibrous substructures. The recyclability of the assembled structures is also a beneficial attribute.

We propose a solvent-triggered topography-regulation strategy to implement reversible fusion and fission. Graphene oxide (GO) fiber was selected as a model because of the nature of GO sheets, including two-dimensional (2D) topology, abundant chemical moieties, superflexibility, and self-adhesion capability (*10*–*13*). After swelling (*14*), the wet-spun GO fiber features a shell (an outermost skin) that confines the movement of internal GO sheets and shows a solvent-triggered large volume change and elastic deformation capability. Upon stimuli of water and polar organic solvents, the topography of fiber shells reversibly switches between a wrinkled, tubular state and a spread, cylindrical state through swelling and deswelling, thus causing a transient fiber interface and leading to cyclic self-fusion and self-fission of an arbitrary number of GO fibers (Fig. 1A). In each cycle, the number, size, composition, structure, and properties of GO fibers are recovered after fission, exhibiting the precise reversibility of fusion and fission.

Continuous GO fibers were prepared following a wet-spinning protocol (fig. S1) (*15*, *16*). In Fig. 1B, we demonstrate the reversible fusion and fission involving up to 100 GO fibers (movie S1). For simplicity, the fused GO fiber assembled by n individual fibers is denoted as FuF-n, and the corresponding fissured, individual dried GO fiber is denoted as FiF-n. A typical fusion process requires three steps, including solvent swelling of fibers in a bundle, drawing the fiber bundle out of the solvent, and air drying. The fission steps comprise reswelling of the fused fiber in a solvent, splitting of swelled fiber bundles, and drying of the fissured fibers separately (see the supplementary materials).

In situ optical microscopy and ex situ scanning electron microscopy revealed the water-induced fusion and fission processes (Fig. 1, C to F; figs. S2 and S3; and movies S2 and S3). The swelling ratio, defined as the ratio of swelled fiber diameter to raw fiber diameter, is used to quantify the extent of swelling, which is influenced by solvent type, the chemical nature of the GO fibers, and the soaking time (fig. S4) (*17*). The equilibrium-swelled GO fibers randomly immersed in water exhibited a swelling ratio of 541% with a core shell structure, where the core (pore size of 2 to 5 μm after freeze drying; inner wall thickness of 28 ± 8 nm) was enclosed by a skin-like shell (80 ± 18 nm thick) of densely packed GO sheets that oriented along the circumference (Fig. 1, C1 and D1, and figs. S5 and S6). This different arrangement of GO sheets within the outer layer and inner fiber is a featured structure of wet-spun fibers that originates from the aligning effect during fiber extrusion and is enhanced by double-diffusive convection in the coagulation bath (*18*). The thickness of both the shell and inner wall depends on the swelling of fibers, which relates to the quantity of absorbed solvent. Driven by the surface tension of water, these swelled fibers collected together and deformed into a cylindrical group spontaneously as they were drawn out from the liquid. During air drying (0 to 40 min), the swelled fibers bonded together accompanied by adaptive crumpling of fiber shells, exhibiting a volume shrinkage of 98% (Fig. 1, C1 to C4 and D1 to D4). Meanwhile, the interlayer spacing decreased from >2.21 to 0.84 nm according to in situ x-ray diffraction characterizations (fig. S7A). Consequently (40 min), the resulted FuF-100 featured compact packing of GO platelets (interlayer spacing of 0.84 nm; density of 1.51 g cm^{-3}) with a tensile strength of 281 MPa (fig. S2, G to I). GO fibers also fused while not drying under tension despite the decrease in density and tensile strength compared with those drying under tension (fig. S8).

Fission of FuF-100 started with a homogeneous swelling when it was resoaked in water, and small gaps emerged at the interfiber interfaces as the swelling persisted. Subsequently (150 s), a rapid gap propagation, along with volume expansion of the whole fiber assembly, led to the complete fission into 100 individual quasicylindrical fibers. The fissured fibers sustained in water for a long time rather than dissolving into pieces despite a swelling ratio of 538%, because of the protective effect of the outer shell. During the whole process, solvent infiltration caused expansion between GO sheets, with interlayer spacing restoring from 0.84 nm to >2.21 nm (fig. S7B). After drying separately, the interlayer spacing between GO sheets (0.84 nm), density (1.54 g cm^{-3}), and tensile strength (259 MPa) of FiFs-100 were close to those of FuF-100 (fig. S2). In situ optical microscopy and polarized optical microscopy observations on the cross section of two GO fibers during their fusion and fission further verified the above description (movie S3).

A GO fiber was labeled with fluorescent 1,1,2-triphenyl-2-(4-bromomethylphenyl) ethylene (TPE-Br) and silicon (Si) nanoparticles. Tracking on the fusion and fission of the labeled fiber and a pristine fiber under fluorescence microscope revealed no substance exchange across the

[1]MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Department of Polymer Science and Engineering, Key Laboratory of Adsorption and Separation Materials and Technologies of Zhejiang Province, Zhejiang University, Hangzhou 310027, China. [2]State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China. [3]Department of Materials Science and Engineering, Monash University, Clayton, Victoria 3800, Australia.
*Corresponding author. Email: lizheng_zju@zju.edu.cn (Z.L.); yilunliu@mail.xjtu.edu.cn (Y.L.); chaogao@zju.edu.cn (C.G.)
†Present address: Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK.

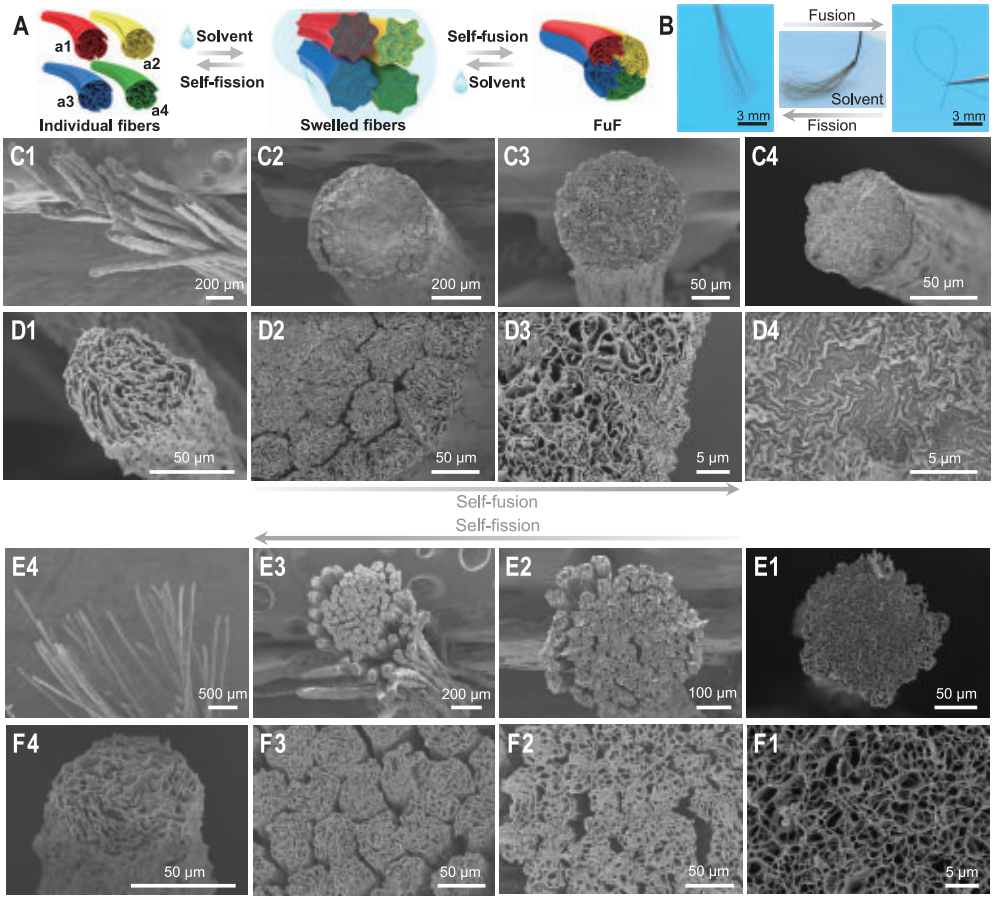


Fig. 1. Reversible fusion and fission of GO fibers. (**A**) Schematic of solvent-triggered precisely reversible self-fusion and self-fission of GO fibers, in which the individual fibers such as a1 to a4 are adaptively deformed by solvent swelling to form a thicker fiber FuF. The dried FuF is swelled in a solvent such as water to gradually fissure into original thinner fibers a1 to a4 without the exchange of GO sheets among them. (**B**) Photographs of reversible transition between ~100 GO fibers (left) and a single FuF-100 (right) through swelling-assisted (middle) fusion and fission. (**C** to **F**) Sequential scanning electron microscopy images of the water-induced fusion process (from left to right) of ~100 GO fibers and the reverse fission process (from right to left) of the FuF-100. (D) and (F) show corresponding enlarged local regions in (C) and (E), respectively. (C1) to (C4) and (D1) to (D4) correspond to 0, 25, 30, and 40 min in elapsed time, respectively. (E1) to (E4) and (F1) to (F4) correspond to 3, 75, 150, and 180 s in elapsed time, respectively. The time of fusion is set as zero when the GO fibers in a bundle are swelled in water just before being drawn out, as is the fission when the dried FuF-100 is resoaked in water.

fiber interface (Fig. 2, A to D; fig. S9; and movie S4). Energy-dispersive spectroscopy (EDS) analysis performed on a neat GO fiber and another composite Si/GO fiber (Fig. 2, E to L, and figs. S10 and S11) showed that fusion was finally accomplished by interlocking of the synergistically crumpled shells, and fission was induced by topographical recovery of shells with opposite normal vectors in geometry on both sides of the bonding interface (Fig. 2, K and L).

On the basis of the above characterizations, fusion and fission were found to be conducted by reversible crumpling and spreading of fiber shells while deswelling and reswelling (Fig. 3A and figs. S12 and S13). Here, the GO fiber shell is defined as the outermost dense layer of a saturated swelled GO fiber, characterized by macroscale cylindrical configuration, microscale ripples, and nanoscale closely organized GO platelets (fig. S13). The shell plays a major role in the reversible fusion-fission process because it is the boundary contacting with adjacent fibers, which provides interfiber bonding and debonding and protects inner fiber GO sheets from diffusion. A swelled GO fiber maintained a nearly constant perimeter during deswelling and reswelling courses, indicating that only elastic deformation occurred at the shell (Fig. 3B). The shell crumpled and spread because of wrinkling and unfolding of the ripple microstructures, with the curvature radius of the ripples varying in the range of ~10 nm to 83 μm (Fig. 3C and fig. S13, A1 to B4). At the same time, the inner fiber GO sheets deformed accordingly because they were interconnected in the swelled gel fiber. The reversible fusion-fission capability is only found in wet-spun GO fibers. Although dry-spun GO fibers can be fused, they failed to fissure, mainly because of the insufficiency of a dense and protective shell (fig. S14) (*19*). The topographical and volumetric evolution of fibers during the fusion process is driven by the surface tension of the solvent (*20*) and the Laplace pressure difference (P_c) (*21*). The attractive stress P_c facilitated further bonding of fiber shells through noncovalent interactions [π–π interaction and hydrogen bonding (*22*)], as well as subsequent wrinkling of GO sheets and densifying of the whole fiber bundle (fig. S15A). During the fusion process, the solvent-responsive fiber shell acts as an elastic barrier, preventing the sheet interdiffusion across the transient interface. Herein, the fusion degree, proposed to evaluate the fusion status (see the supplementary materials for the calculation), increased from 0% (unfused) to 100% (fused) upon swelling and traversed a hemifused region (swelling ratio from 165 to 358%) (fig. S16).

Conversely, fission was attributed to the cylindrical geometry–driven detachment between fiber shells (fig. S17A). As the FuF was soaked in good solvents for GO, the solvent infiltration weakened the adhesion strength (σ_{ad}*) between individual fibers (*23*). The tendency of cylindrical geometry reversion at the interface contributes to the interfacial detachability and compels the attached fibers to separate from each other. This repulsion is estimated by the net stress (denoted as fission stress σ_{fis}) of elastic tensile stress σ_{e1} and swelling pressure p_s along the inward-pointing normal direction of the curved microelement (see the supplementary materials). This hypothesis was further confirmed by finite element analysis, which showed that σ_{fis} was generated by the curved geometry of the shell when the swelling ratio of individual fibers surpassed a sufficiently high value (475 and 521% for separation of point 1 and 2, respectively) (fig. S17, B to E, and movie S5). Further swelling led to the rising σ_{fis}, which consequently reached a peak value equal to σ_{ad}* and triggered a sudden drop of σ_{fis} to 0 MPa. In fission experiments, the fission capability of FuF-2 was initiated when the average swelling ratio of individual fibers traversed a critical range from 310 to 419%, either by applying polarity-enhanced solvent or by reducing the pretreating temperature

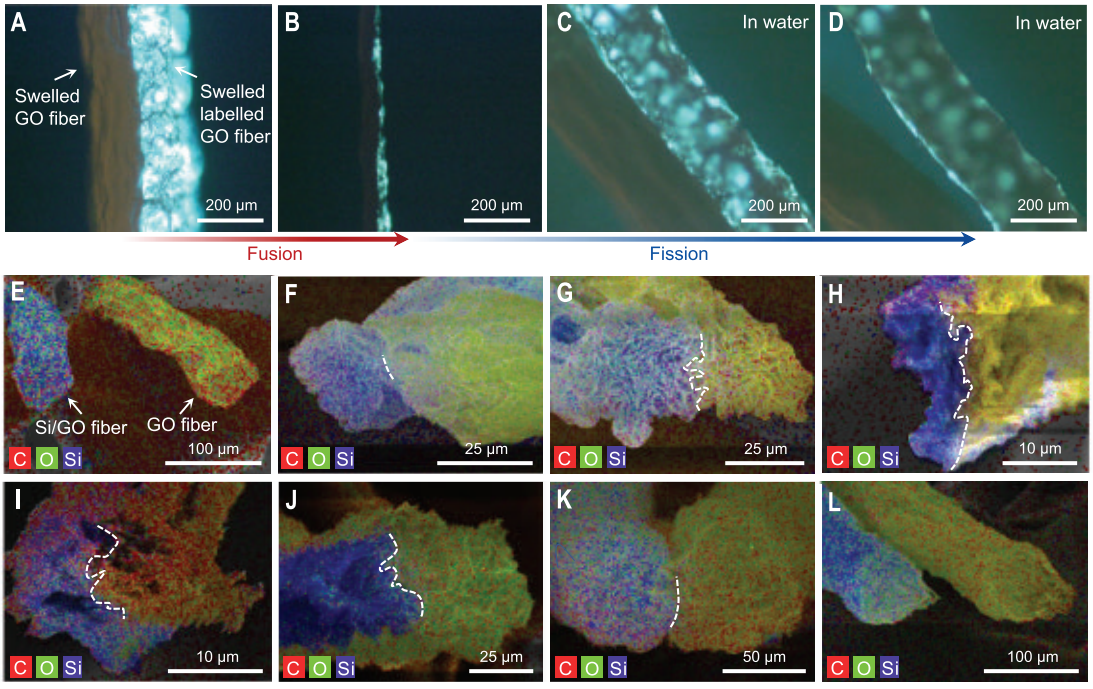


Fig. 2. In situ fluorescence observation and ex situ EDS analysis of a fusion-fission cycle. (**A** to **D**) Sequential fluorescent micrographs showing water-triggered fusion (A and B) and fission (B to D) of a neat GO fiber and a fluorescent TPE-Br–labeled GO fiber. A clear interface between the two fibers can be seen, demonstrating no exchange of GO sheets between them while fusing and fissuring. (**E** to **L**) Sequential overlapped elemental mapping images of carbon (red), oxygen (green), and silicon (blue) elements, respectively, showing the water-induced fusion (top) and fission (bottom) procedures involving a neat GO fiber and a Si/GO fiber. Dashed lines indicate the interface between the fibers, which suggests interlocking of the synergistically crumpled shells after fusion and the corresponding topographical recovery after fission.

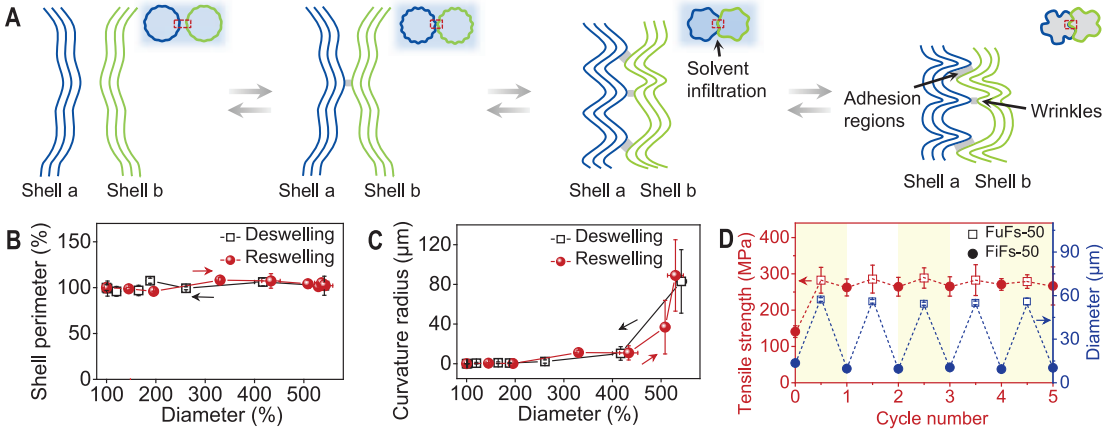


Fig. 3. Dynamic topographical deformation mechanism for reversible fusion and fission. (**A**) Schematic depicting the topographical crumpling of shells at the fiber interface that leads to self-fusion and reversible recovery that induces self-fission. (**B**) Shell perimeter of the water-swelled single GO fibers as a function of fiber diameter during the deswelling and reswelling courses. Corresponding values are plotted relative to that at the initial swelling state. (**C**) Average tip radius of curvature of a unit ripple in different fiber diameter during the deswelling and reswelling procedures of the water-swelled single GO fibers. The diameter of raw dry GO fiber in (B) and (C) is 52 μm. (**D**) Tensile strength (red) and corresponding diameter (blue) of FuF-50 and FiF-50 as a function of cycle number. The mechanical properties remained constant during several cycles.

on the FuF-2 (fig. S18A). Additionally, in contrast to GO fibers, GO belts featuring a distinct flat topography after swelling were able to be fused together but failed to fissure because of the absence of an arc conformation of GO sheets stored in wrinkles, which therefore lack a driving force σ_{fis} for self-fission (fig. S19).

The morphology and structure of both FiF-50 and FuF-50 were recovered throughout multiple fusion-fission cycles while exhibiting a relatively constant tensile strength of 281 and 259 MPa, respectively (Fig. 3D, fig. S20, and movie S6). After several cycles, the interlayer spacing of GO sheets in fibers remained unchanged at the same fusion or fission time (fig. S21). Additionally, there was no apparent decline in the tensile and compressive strength of FuFs when the fiber diameter increased from 23 to 78 μm (by increasing the number of individual GO fibers engaged in fusion from 10 to 100) (fig. S22). The values were stable at ~282 and 129 MPa, respectively. After thermal reduction at 1000°C, the FuF-100 exhibited a tensile strength of 597 MPa with a diameter of 58 μm. The consistency of the mechanical performances reflects that the stacking order within individual GO fibers remained intact. One advantage of thick FuFs is that firmly combined fibers are stronger than either the as-spun thick GO fibers or the unfused yarns with separated fibers (figs. S22B and S2I). Therefore, the affordable force on the thick FuFs is higher, which is better for structural materials that may exert mechanical superiority in engineering fields. Fluorescence tracking and EDS analysis on the labeled GO fibers further showed the cycles of reversible fusion and fission (figs. S23 and S24). Investigation on bundles composed of one Si/GO fiber and 99 neat GO fibers showed that the relative atomic content of Si in the Si/GO fiber remained constant whether

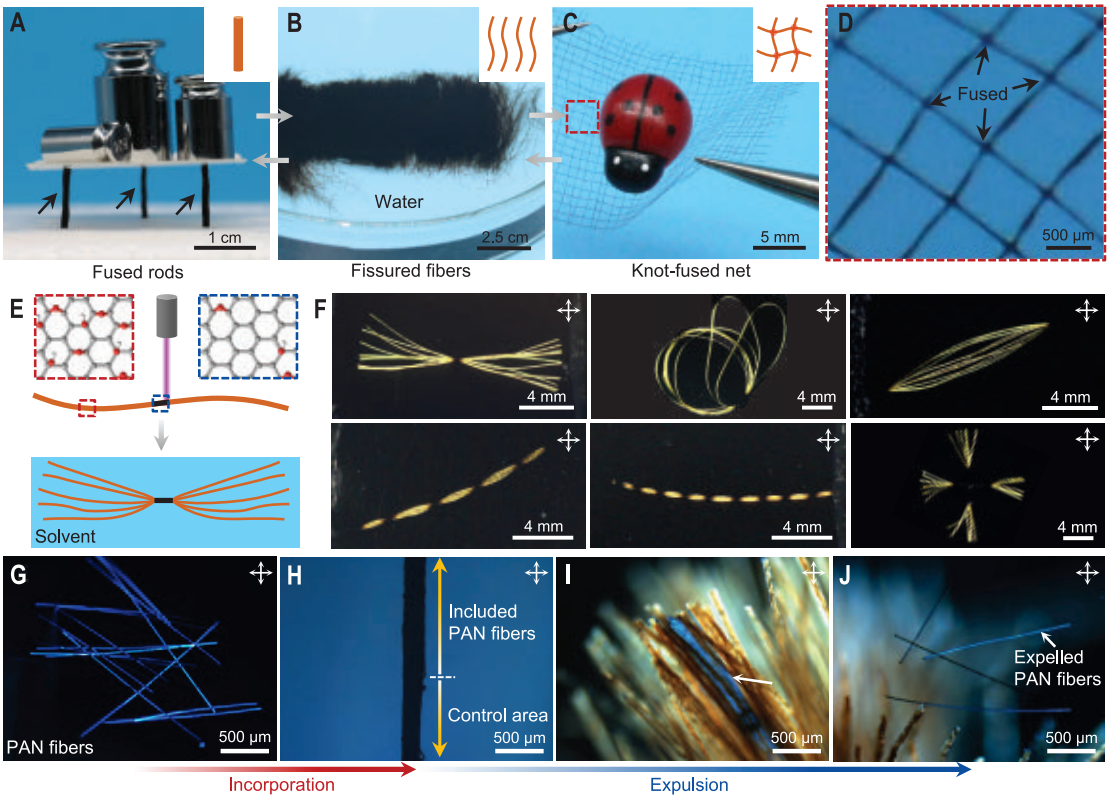


Fig. 4. Reversible fusion-fission promises controllable transformation. (**A** to **D**) Photographs of reversible transitions between a 3D, stiff fused GO rod (A) and a 2D, flexible knot–fused GO net (C and D) through fission (B) and refusion. (**E** and **F**) Illustration and polarized optical photographs showing the programmable 1D and 2D architectures of the locally laser-reduced FuF-10 under a water/isopropanol (8:2 v/v) mixture, including star, wristband, and multiblock wires with 3, 11, and 19 blocks and a dendrite with 40 arms. In the top of panel (E), the gray, red, and white spheres denote carbon, oxygen, and hydrogen atoms, respectively. (**G** to **J**) Polarized optical microscopy images showing the entrapment of polyacrylonitrile (PAN) staple fibers into a FuF by fusion (G and H) and the reversible expulsion by fission (H to J).

the Si-labeled fiber was within the fused fiber surrounded by neat GO fibers or at the fissured state (fig. S24, C to E; see the supplementary materials for the calculation).

Figure 4 demonstrates concepts for potential application of the fusion-fission behavior shown here. First, flexible transformation between diverse fiber–based assembled structures becomes possible. This would allow adaptive application of GO fiber–based systems in different scenarios with specific performance needs. For instance, the GO fiber assemblies were demonstrated to transform reversibly between a 3D stiff rod and a 2D flexible net through fission and refusion (Fig. 4, A to D; fig. S25A; and movie S7). In this case, as many as 13,500 fibers with micrometer-scale diameter and centimeter-scale length were fused into one 1.2-mm-thick rod that was rigid and strong enough to support 680 times its weight. Alternatively, transitions between a 1D fused GO fiber and various 1D and 2D sophisticated fiber assemblies were also conducted by localized fission and fusion (Fig. 4, E and F, and movie S8). The second application is that, through fusion and fission, a GO fiber bundle would be able to implement functions of including and expelling of guest objects, which may show promise in dynamic systems for controllable delivery. Various guest objects of different materials, sizes, and shapes, such as polyacrylonitrile chopped fibers, polystyrene microspheres, and glass beads in submillimeter scale, were absorbed into a FuF during fusion and then expelled during fission (Fig. 4, G to J; fig. S26; and movie S9). The third application is endowing ordinary fibers with reversible fusion and fission properties through GO coating. Conventional polymer, metal, and ceramic fibers were endowed with the reversible fusion-fission capability by simply coating a GO outer layer (figs. S27 and S28 and movie S10). This extension further expands the coverage of corresponding application fields.

Precisely reversible fusion and fission were induced here by the recoverable geometrical deformation of GO fibers. This enabled a dynamic feature in the fiber-assembled system so that transformation among structures and responsive actuation were realized. The concept was further expanded to conventional fibers through GO coatings while providing the potential for recyclability. The featured fusion-fission behavior constitutes a versatile strategy for the design of functional responsive materials.

REFERENCES AND NOTES

1. F. Menger, N. Balachander, *J. Am. Chem. Soc.* **114**, 5862–5863 (1992).
2. Y. Zhou, D. Yan, *J. Am. Chem. Soc.* **127**, 10468–10469 (2005).
3. Y. Zhou, D. Yan, *Angew. Chem. Int. Ed.* **44**, 3223–3226 (2005).
4. J. C. Shillcock, R. Lipowsky, *Nat. Mater.* **4**, 225–228 (2005).
5. H. Cai, G. Jiang, Z. Shen, X. Fan, *Soft Matter* **9**, 11398–11404 (2013).
6. Y. Rharbi, M. Karrouch, P. Richardson, *Langmuir* **30**, 7947–7952 (2014).
7. Y. Yao *et al.*, *Science* **359**, 1489–1494 (2018).
8. P. Jensen, *Rev. Mod. Phys.* **71**, 1695–1735 (1999).
9. C. C. Wang *et al.*, *Nano Res.* **8**, 2143–2151 (2015).
10. J. Kim, L. J. Cote, J. Huang, *Acc. Chem. Res.* **45**, 1356–1364 (2012).
11. S. Presolski, M. Pumera, *Angew. Chem. Int. Ed.* **57**, 16713–16715 (2018).
12. P. Poulin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11088–11093 (2016).
13. C. Luo, C. N. Yeh, J. M. L. Baltazar, C. L. Tsai, J. Huang, *Adv. Mater.* **30**, e1706229 (2018).
14. Z. Li, Z. Xu, Y. Liu, R. Wang, C. Gao, *Nat. Commun.* **7**, 13684 (2016).
15. Z. Xu, C. Gao, *Nat. Commun.* **2**, 571 (2011).
16. B. Fang, D. Chang, Z. Xu, C. Gao, *Adv. Mater.* **32**, e1902664 (2020).
17. M. Tanaka *et al.*, *J. Phys. Condens. Matter* **17**, S649–S663 (2005).
18. Z. Xu, C. Gao, *Acc. Chem. Res.* **47**, 1267–1276 (2014).
19. Q. Tian *et al.*, *Nanoscale* **9**, 12335–12342 (2017).
20. X. Li *et al.*, *Langmuir* **27**, 12164–12171 (2011).
21. D. Chandra, S. Yang, *Acc. Chem. Res.* **43**, 1080–1091 (2010).
22. H. Cheng *et al.*, *Adv. Funct. Mater.* **27**, 1703096 (2017).
23. N. V. Medhekar, A. Ramasubramaniam, R. S. Ruoff, V. B. Shenoy, *ACS Nano* **4**, 2300–2306 (2010).

ACKNOWLEDGMENTS

We thank N. Zheng in the State Key Laboratory of Chemical Engineering at Zhejiang University for assistance with performing EDS analysis and SSRF for x-ray diffraction measurements. **Funding:** This work was funded by the National Key R&D Program of China (grant 2016YFA0200200), the National Natural Science Foundation of China (grants 52090030, 51533008, 51703194, 11890674, and 51873191), and the Hundred Talents Program of Zhejiang University (grant 188020*194231701/113). **Author contributions:** D.C., Z.L., and C.G. conceived the project. D.C., Z.L., and B.F. designed the experiments. D.C. conducted the experiments. J.L. and Y.L. conducted the finite element analysis and mechanical analysis. L.B. provided the modeling code in the finite element analysis. D.C., Z.L., Z.X., B.F., and C.G. analyzed the results. D.C., Z.L., B.F., and C.G. wrote the manuscript with input from all authors. C.G. supervised the study. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/614/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S28
References (*24*–*38*)
Movies S1 to S10

10 March 2020; accepted 16 March 2021
10.1126/science.abb6640

SOLAR CELLS

Interfacial toughening with self-assembled monolayers enhances perovskite solar cell reliability

Zhenghong Dai, Srinivas K. Yadavalli, Min Chen, Ali Abbaspourtamijani, Yue Qi, Nitin P. Padture*

Iodine-terminated self-assembled monolayer (I-SAM) was used in perovskite solar cells (PSCs) to achieve a 50% increase of adhesion toughness at the interface between the electron transport layer (ETL) and the halide perovskite thin film to enhance mechanical reliability. Treatment with I-SAM also increased the power conversion efficiency from 20.2% to 21.4%, reduced hysteresis, and improved operational stability with a projected T80 (time to 80% initial efficiency retained) increasing from ~700 hours to 4000 hours under 1-sun illumination and with continuous maximum power point tracking. Operational stability–tested PSC without SAMs revealed extensive irreversible morphological degradation at the ETL/perovskite interface, including voids formation and delamination, whereas PSCs with I-SAM exhibited minimal damage accumulation. This difference was attributed to a combination of a decrease in hydroxyl groups at the interface and the higher interfacial toughness.

The promise of low cost, high power conversion efficiency (PCE), and versatility has driven research efforts on perovskite solar cells (PSCs) (*1*–*3*). The low formation energies of metal halide perovskite (MHP) light absorbers that enable solution processing at or near room temperature (*4*) also make them unstable (*5*, *6*). Research on improving PSC operational stability (*5*, *6*) together with scalability (*7*) has made steady progress, but PSCs will also need to be mechanically reliable if they are to operate efficiently for decades (*6*, *8*–*11*). Enhancing the mechanical reliability of PSCs is particularly challenging because the low formation energies of MHPs result in inherently poor mechanical properties, meaning that they are compliant (low Young's modulus, E), soft (low hardness, H), and brittle (low toughness, G_C) (*9*). For example, the prototypical MHP, methylammonium lead triiodide ($CH_3NH_3PbI_3$, or $MAPbI_3$), has E ~ 17.8 GPa, H ~ 0.6 GPa, and G_C ~ 2.7 $J{\cdot}m^{-2}$ (cohesion), as measured by nanoindentation of single crystals (*9*). Furthermore, interfaces between the MHP thin film and the adjacent functional layers in the planar PSC multilayer stack are even more brittle, with $G_C < 1.5$ $J{\cdot}m^{-2}$ (adhesion) typically (*8*, *10*, *12*, *13*), making them prone to premature delamination.

The sources of internal and external mechanical stresses in PSCs that can drive fracture include (*9*, *11*) (i) coefficient of thermal expansion (CTE) mismatch between the layers; (ii) in-service thermal excursions; (iii) in-service damage accumulation; and (iv) deformation during manufacturing, installation, maintenance, and service (such as bending, stretching, and twisting). In tandem photovoltaics (PVs) incorporating PSCs, the CTE-induced internal stresses are expected to be even higher owing to the additional layers, and in the case of flexible single-junction PSCs, the externally applied stresses are typically more severe (*9*, *11*). Although the ultimate in-service delamination failure of devices depends on many factors, G_C of the weakest interface—akin to the weakest link in a chain—is perhaps the most important metric that determines the mechanical reliability of multilayer devices (*14*).

There have been a few attempts to enhance the G_C of the weakest interface in PSCs, with varying degrees of success, using approaches such as adding interfacial layers (*10*, *15*–*18*), scaffolding (*19*), interpenetrating interfaces (*20*), introducing additives (*13*, *21*), and grain coarsening (*12*). In this study, we demonstrate the substantial toughening of the brittle interface between the MHP thin film and the underlying SnO_2 electron transport layer (ETL) using an iodine-terminated self-assembled monolayer (I-SAM) that acts as a "molecular glue." This processing step not only increased PCE (up to 21.44%) and reduced hysteresis but also improved operational stability under 1-sun maximum power point (MPP) continuous operation [projected T80 (time to 80% initial PCE retained) up to ~4000 hours]. The characterization of the operational stability–tested PSCs reveals that the I-SAM helped preserve the mechanical integrity of the ETL/MHP interface, and what is called operational stability in previous PSC studies is closely intertwined with mechanical reliability. We note that SAMs have been used to toughen interfaces in other types of devices while improving other functional properties, for example, thermal conductivity (*22*).

We chose the mixed-composition MHP, $Cs_{0.05}(FA_{0.85}MA_{0.15})_{0.95}Pb(I_{0.85}Br_{0.15})_3$ with 4 mol % excess PbI_2 [FA = formamidinium or $HC(NH_2)_2^+$], optimized for high PCE and stability (*23*) and SnO_2 as the ETL, as it provides a more favorable energy-level alignment with this MHP and minimizes the photocatalytic degradation of the MHP compared with TiO_2. The 3-iodopropyl trimethoxysilane [$Si(OCH_3)_3(CH_2)_3I$] I-SAM, with $-Si(OCH_3)_3$ anchor group, –I terminal group, and alkyl $(CH_2)_n$ chain (n = 3) linker, was chosen on the basis of the following considerations. First, the surface of the SnO_2 ETL, as in most oxides (*24*), is covered with adsorbed hydroxyl (–OH) groups that get trapped during deposition of the MHP thin film on top in n-i-p "regular" planar PSCs and degrade performance (*25*, *26*). Trialkoxysilane SAMs are known to self-assemble and cross-link readily on such oxide surfaces by the silanization process (*27*) and greatly reduce the number of surface –OH groups as they create anchoring O–Si bonds. Second, assuming a "brushlike" SAM morphology due to the short alkyl chain length used here, the –I terminal group is expected to form electrostatic bonds to the MHP thin film on top (*28*). A control H-SAM with a –H terminal group [$Si(OCH_3)_3(CH_2)_3H$] was used in separate experiments to isolate the effects of the –I terminal group. Third, the short length of the alkyl chain (n = 3) makes the SAM-coated surface sufficiently lyophilic for the subsequent solution deposition of the MHP thin film. Finally, SAMs in general (*29*, *30*), and trialkoxysilane-based SAMs in particular (*31*, *32*), increase PCE and reduce hysteresis in PSCs through improved extraction of photocarriers, reduced charge accumulation at interfaces, and passivation of interfacial charge traps.

We measured the adhesion toughness, G_C, of the ETL/MHP interface using the "sandwich" double-cantilever beam (DCB) delamination method (Fig. 1A) (*8*, *10*, *12*, *14*), as described in the supplementary materials (SM) and fig. S1. The SnO_2 ETL was deposited on indium tin oxide (ITO)–coated glass substrates using the method of Jiang *et al.* (*33*). The SAM was deposited on the SnO_2 surface by dip-coating at room temperature using the procedure described in the SM. We note that SAM deposition is sensitive to several experimental parameters, including water content, solvent used, solution age, deposition time, temperature, etc. (*27*). The presence of Si on the H-SAM–coated ETL surface and both Si and I on the I-SAM–coated ETL surface was confirmed with x-ray photoelectron spectroscopy (XPS) (fig. S2). The MHP thin film (~500-nm thickness) was then deposited through a variation of the solvent-engineering method (*34*). X-ray diffraction (XRD) patterns of the MHP thin films without SAMs and ones with H-SAM or I-SAM underneath (fig. S3) showed no difference in the thin-film MHP phase. Similarly, the top-surface scanning electron microscopy (SEM) images in fig. S4 showed no discernible difference in the MHP thin-film microstructure; the average apparent grain sizes of the MHP thin films are between ~330 and ~370 nm.

School of Engineering, Brown University, Providence, RI 02912, USA.
*Corresponding author. Email: nitin_padture@brown.edu

The G_C of the ETL/MHP interface (Fig. 1B and table S1) was enhanced by ~50% with I-SAM (1.91 ± 0.48 to 2.83 ± 0.35 J·m^{-2}). For H-SAM, the mechanical bonding with MHP was weak, resulting in G_C decreasing to 1.72 ± 0.54 J·m^{-2}. Delamination failure occurred at the ETL/MHP interface in each case; SEM images of the mating fracture surfaces are shown in fig. S5. For the I-SAM case, the SEM image of the SnO_2 ETL side fracture surface showed a few adhered smaller MHP grains due to occasional intergranular fracture.

Density functional theory (DFT) calculations were performed to study the adhesion provided by the interfacial bonding between the MHP surface and idealized brushlike SAMs (see SM for details) (*35*–*42*), where we used α-$FAPbI_3$ to represent the MHP. Aligned butane [$H(CH_2)_4H$] and I-terminated butane [$(H(CH_2)_4I$] molecules were used to represent H-SAM and I-SAM, respectively, because the Si-based anchor group in the SAMs was unlikely to influence the bonding at the other end of the alkyl chains. The adhesion of α-$FAPbI_3$ (001) surfaces, of different terminations, with $H(CH_2)_4I$ was found to be about twice that with $H(CH_2)_4H$ (table S2 and fig. S6). The strongest adhesion was achieved at the PbI_2-terminated α-$FAPbI_3$ (001) surface with two aligned $H(CH_2)_4I$ molecules per unit cell (near-ideal packing density of 20 Å^2 per molecule). The interfacial bonding with I-SAM has the characteristics of the halogen bond, where a low–electron density region on a covalently bonded halogen atom (mainly in the heavier halogens) forms attractive interaction with electron-rich sites (*43*). The electrophilic regions associated with the halogen atoms on $H(CH_2)_4I$ gained electrons from both the undercoordinated Pb^{2+} (electron donor) and fully coordinated I^- (nucleophile) on the PbI_2-terminated α-$FAPbI_3$ (001) surface (Fig. 1D). Halogen bonding changes with electrophilic order: I > Br > Cl > F, which increases with polarizability and decreases with electronegativity. A similar interface using $H(CH_2)_4F$ molecules at the other extreme of the electrophilic order confirms this trend (table S2). Naturally, no such bond is expected at the interface with $H(CH_2)_4H$ (Fig. 1C). The lack of electron transfer and the longer I···H and Pb···H distances indicated the absence of chemical bonding (see fig. S6 and table S2 for other α-$FAPbI_3$ surface terminations and different packing densities). Although I-SAM in the experiments may not be as idealized (i.e., brushlike, full-coverage), the DFT calculations support the hypothesis that I-termination enhances the bonding with the MHP surface substantially and that it is preferred over other halogen terminations. The latter was further confirmed experimentally using 3-bromopropyl trimethoxysilane [$Si(OCH_3)_3(CH_2)_3Br$] Br-SAM, where an average G_C of

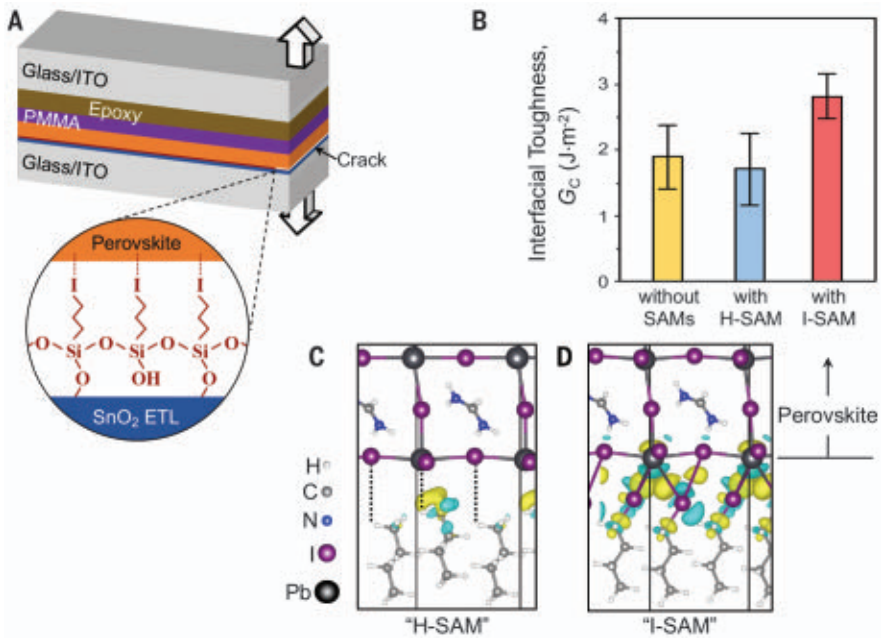


Fig. 1. Mechanical behavior of the ETL/MHP interface. (**A**) Schematic illustration of the sandwich DCB specimen for toughness testing (not to scale). PMMA, poly(methyl methacrylate). Inset shows magnified schematic illustration of idealized I-SAM. (**B**) Toughness of ETL/MHP interface without SAMs and ones with H-SAM or I-SAM. The histograms and error bars represent average and standard deviation, respectively, of 12 specimens each (see table S1 for the data). (**C** and **D**) Charge transfer density difference plots from DFT calculations. Bonding between PbI_2-terminated α-$FAPbI_3$ (001) surface and (C) $H(CH_2)_3H$ ("H-SAM") or (D) $H(CH_2)_3I$ ("I-SAM"). The yellow and blue colors indicate electron gain or loss above 0.0047 e·Å^{-3}, respectively. Dashed lines across the interface indicate no bonding.

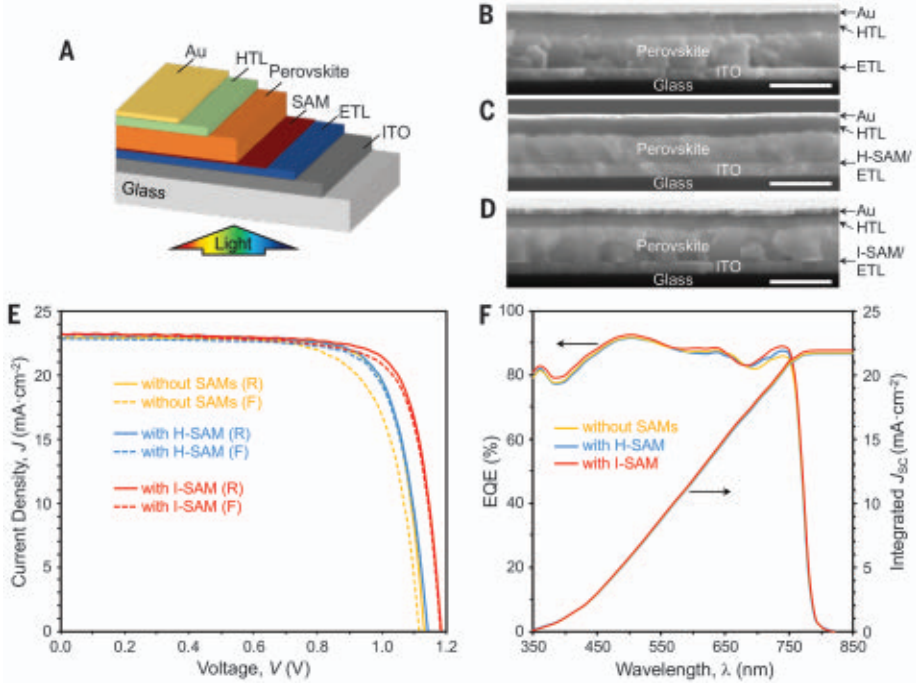


Fig. 2. Structure and performance of PSCs. (**A**) Schematic illustration (not to scale) of the n-i-p regular planar PSCs with SAMs (SAMs absent in control PSC). HTL, hole transport layer. Cross-sectional SEM images of as-fabricated PSCs: (**B**) without SAMs, (**C**) with H-SAM, and (**D**) with I-SAM. Scale bars, 1 μm. (**E**) *J*-*V* responses, in reverse (R) and forward (F) scans, of champion PSCs without SAMs and ones with H-SAM or I-SAM. (See Table 1 for PV performance parameters.) (**F**) EQE spectra and integrated J_{SC} of the champion PSCs without SAMs and ones with H-SAM or I-SAM.

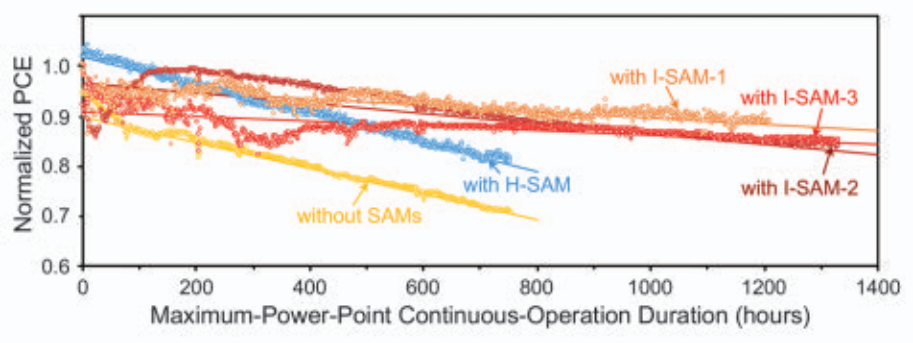


Fig. 3. Operational stability of PSCs. Normalized PCE of PSCs without SAMs and ones with H-SAM or I-SAM (three devices) as a function of time under the following conditions: 1-sun continuous illumination, MPP tracking, unencapsulated, flowing N_2 atmosphere, and room temperature. PCE was recorded about every hour. The lines are linear fits to the data after initial burn-in and nonmonotonic behavior, where the *y* intercept and the slope are used to estimate and project the T80 duration.

Table 1. PV performance parameters. Values extracted from the *J-V* data, in reverse (R) and forward (F) scans, of champion PSCs without SAMs and ones with H-SAM or I-SAM.

PSC	Scan	V_{OC} (V)	J_{SC} (mA·cm^{-2})	FF	PCE (%)
Without SAMs	R	1.131	23.02	0.774	20.15
	F	1.117	22.99	0.721	18.52
With H-SAM	R	1.143	23.18	0.762	20.19
	F	1.142	22.86	0.760	19.84
With I-SAM	R	1.185	23.26	0.778	21.44
	F	1.182	23.22	0.759	20.83

2.08 ± 0.31 J·m^{-2} was measured (table S1), which falls in between the H-SAM and I-SAM cases.

We fabricated PSCs with planar n-i-p regular architecture (Fig. 2A) for different SAMs (see SM for device fabrication and testing details). Cross-sectional SEM images in Fig. 2, B to D, showed no discernible differences. The current density (*J*)–voltage (*V*) response of the "champion" (i.e., best-performing) PSCs without SAMs and the ones with H-SAM or I-SAM are presented in Fig. 2E, and the corresponding PV performance parameters are listed in Table 1. The I-SAM increased open-circuit voltage (V_{OC}), and both PSCs with SAMs showed lower hysteresis indices (1.8% H-SAM and 2.9% I-SAM versus 8.8% without SAMs). The short-circuit current density (J_{SC}) values for all PSCs compared favorably with the respective values derived from the external quantum efficiency (EQE) spectra of these PSCs in Fig. 2F. The stabilized PCE output at MPP of these three PSCs is presented in fig. S7.

Statistics for the PV performance parameters J_{SC}, V_{OC}, fill factor (FF), and PCE for PSCs (20 of each type) are presented in fig. S8, with an average PCE of 19.04 ± 0.49% for PSCs without SAMs, 19.04 ± 0.48% for PSCs with H-SAM, and 20.20 ± 0.62% for PSCs with I-SAM, confirming overall increased performance of the latter and also demonstrating reproducibility. The increased PCE in PSCs with I-SAM was largely the result of the increased V_{OC}. In this context, the estimated trap filled limit voltages (V_{TFL}) for electron-only transport devices shown in fig. S9 were used to estimate trap densities (n_{Trap}) of 5.21 × 10^{15} cm^{-3} for the device without SAMs, 1.98 × 10^{15} cm^{-3} for the device with H-SAM, 1.36 × 10^{15} cm^{-3} for the device with I-SAM. We attributed the increased V_{OC} to the passivation of charge traps by I-SAM that improved extraction of electrons. The reduced hysteresis with both SAMs was attributed to the reduced charge accumulation with fewer surface –OH groups. Similar increases in PSC performance with different types of SAMs have been demonstrated in several studies (*25*, *28*–*32*).

In Fig. 3, we present the operational stability of PSCs under continuous 1-sun illumination with MPP tracking. In all the tested PSCs, there was initial burn-in instability with a decline, rise, or both in PCE followed by linear steady-state degradation, as has been seen previously for single-junction PSCs (*44*). Following a method proposed by Khenkin *et al.* (*44*), we estimated a T80 of ~692 hours for PSCs without SAMs and ~714 hours for PSCs with H-SAM of (see SM for details). The steady-state degradation of these PSCs appeared to occur at about the same rate. The extrapolated T80 durations for the three PSCs with I-SAM, assuming linear degradation, were estimated at ~3006, ~1896, and ~3921 hours. The normalized PV performance parameters (J_{SC}, V_{OC}, FF, and PCE) of the PSCs from the *J-V* responses measured during operational stability testing are presented in fig. S10. [Note that PCE derived from the *J-V* response was typically higher than that measured by MPP tracking (*44*, *45*).] The performance degradation in the PSC without SAMs was in all three parameters, J_{SC}, V_{OC}, and FF. Table S3 presents PV performance parameters of the PSCs with SAMs, before and after the completion of operational stability testing, showing decay primarily in V_{OC} and FF.

The operational stability–tested PSCs were characterized by cross-sectioning one half of each of these devices and deliberately delaminating the other half. Cross-sectional SEM images of the PSC without SAMs (Fig. 4, A to C) showed three types of irreversible morphological degradation features at the interface compared with the corresponding "before" SEM image in Fig. 2B: small voids, large voids, and delamination. In the PSC with H-SAM (Fig. 4D), although voids were not apparent, interfacial delamination was observed. In contrast, such degradation features were not seen in the PSC with I-SAM in the SEM image in Fig. 4E, which looked similar to the corresponding "before" SEM image in Fig. 2D.

Figure 4, F to I, shows SEM images of the perovskite bottom side fracture surfaces of the same delaminated PSCs. Consistent with the cross-sectional observations, the perovskite bottom side fracture surface of the PSC without SAMs showed small voids, large voids (Fig. 4F), and delamination (Fig. 4G). Light-contrast degradation-product particles were also seen in Fig. 4, F and G, which were absent in the pristine interface (fig. S5B). In the PSC with H-SAM (Fig. 4H), delamination and some voids were observed. In comparison, degradation features in the PSCs with I-SAM were relatively unremarkable (Fig. 4I), which is,

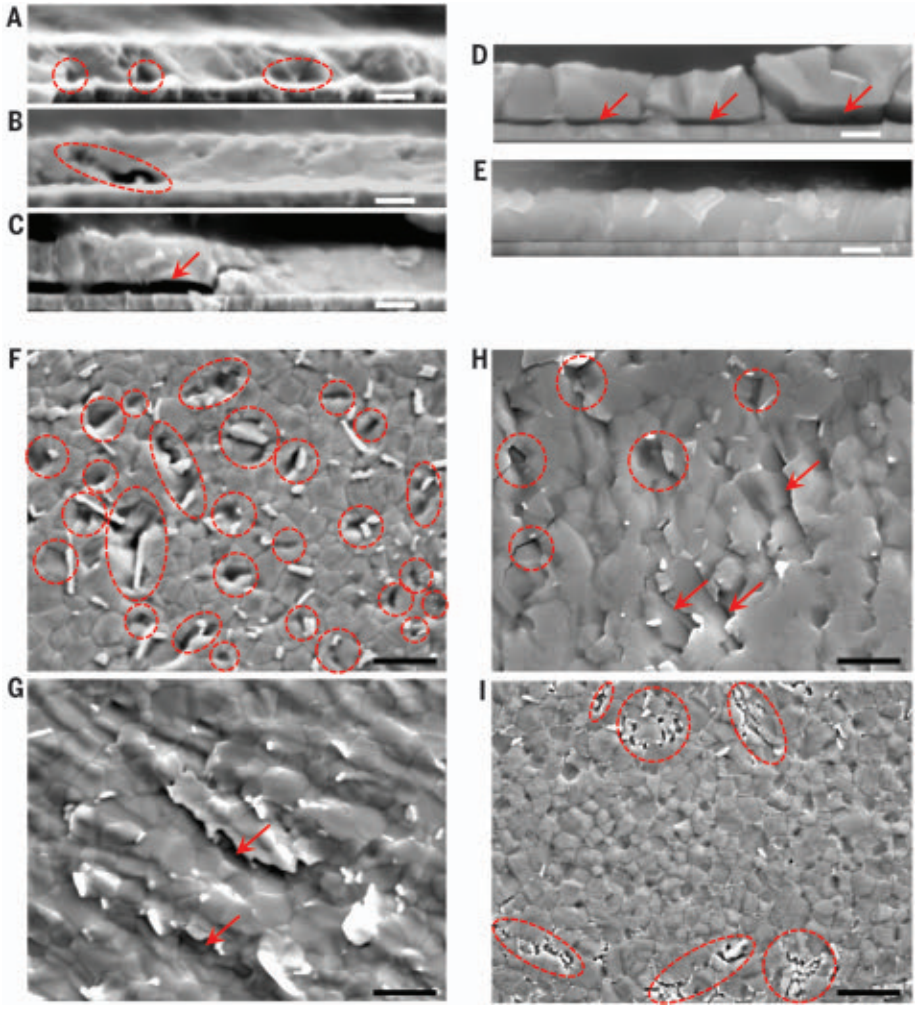


Fig. 4. Characterization of operational stability–tested PSCs. (**A** to **C**) Cross-sectional SEM images of the PSC without SAMs tested for 757 hours showing morphological degradation at the ETL/MHP interface: (A) small voids (dashed circles/ovals), (B) large voids (dashed oval), and (C) delamination (arrow). (**D**) Cross-sectional SEM image of the PSC with H-SAM tested for 754 hours showing delamination (arrows). (**E**) Cross-sectional SEM image of a PSC with I-SAM tested for 1331 hours with intact ETL/MHP interface. (**F** and **G**) SEM images of fracture surface (perovskite bottom side) of the PSC without SAMs showing: (F) small and large voids (dashed circles/ovals) and (G) interfacial delamination (arrows). (**H**) Corresponding SEM image of the PSC with H-SAM showing voids (dashed circles/ovals) and delamination (arrows). (**I**) Corresponding SEM image of the PSC with I-SAM showing minimal morphological degradation (dashed circles/ovals). Scale bars, 0.5 μm [(A) to (E)] and 1 μm [(F) to (I)].

once again, consistent with the corresponding cross-sectional SEM image (Fig. 4E).

We attributed the reduced severity of irreversible degradation of the ETL/MHP interface with I-SAM, and the attendant decay in the PSC performance, to the following effects of the I-SAM. First, both H-SAM and I-SAM should reduce the prevalence of –OH groups at the SnO_2 ETL surface. This results in the reduction of the concentration of positively charged point defects (*25*), which should slow down the ion migration–induced formation of interfacial voids that can serve as sites for photocarrier recombination (*46*) and nucleation of interfacial delamination cracks. The enhanced G_C of the ETL/MHP interface with I-SAM should impede propagation of these incipient interfacial delamination cracks that would block electric current (*46*). To isolate the effect of light, we exposed the DCB specimens without SAMs and the ones with H-SAM or I-SAM to 1-sun (light-emitting diode) continuous illumination in air (~35°C, ~40% relative humidity) from the bottom for 120 or 240 hours and then tested them to measure the G_C. The results in fig. S11 and table S4 show that after 240-hour exposure there was 14 and 17% loss in the average G_C in PSCs without SAMs and those with H-SAM, respectively, but only 5% for those with I-SAM. This suggested that the light-induced reduction in G_C in specimens without SAMs and ones with H-SAM, both with initial lower G_C (Fig. 1B), also contributed to the degradation of their respective PSCs.

We chose the most brittle interface (ETL/MHP) in PSCs for SAM-induced toughening for maximum impact. When sufficiently toughened, delamination failure will shift to the next weakest interface in the PSC multilayer stack, and so on (*9*, *10*), so future efforts will need to focus on toughening the other PSC interfaces using SAMs. Room-temperature solution deposition of SAMs is a low-cost approach that is amenable to scale-up for batch and/or continuous manufacturing of a wide variety of PSCs (rigid, flexible, and tandem). Eventually, when all the interfaces are sufficiently toughened and optimized, the ultimate failure of PSCs will be determined by fracture within the individual layers. Toughening of the individual layers themselves, for example, by incorporating reinforcements as in ceramic or polymer nanocomposites, while simultaneously improving their functional properties, will be the next challenge.

REFERENCES AND NOTES

1. N.-G. Park, M. Grätzel, T. Miyasaka, *Organic-Inorganic Halide Perovskite Photovoltaics: From Fundamentals to Device Architectures* (Springer, 2016).
2. H. J. Snaith, *Nat. Mater.* **17**, 372–376 (2018).
3. A. K. Jena, A. Kulkarni, T. Miyasaka, *Chem. Rev.* **119**, 3036–3103 (2019).
4. W. A. Dunlap-Shohl, Y. Zhou, N. P. Padture, D. B. Mitzi, *Chem. Rev.* **119**, 3193–3295 (2019).
5. Y. Rong *et al.*, *Science* **361**, eaat8235 (2018).
6. S. P. Dunfield *et al.*, *Adv. Energy Mater.* **10**, 1904054 (2020).
7. Z. Li *et al.*, *Nat. Rev. Mater.* **3**, 18017 (2018).
8. N. Rolston *et al.*, *Extreme Mech. Lett.* **9**, 353–358 (2016).
9. C. Ramirez, S. K. Yadavalli, H. F. Garces, Y. Zhou, N. P. Padture, *Scr. Mater.* **150**, 36–41 (2018).
10. N. Rolston *et al.*, *Adv. Energy Mater.* **8**, 1702116 (2018).
11. S. K. Yadavalli, Z. Dai, H. Zhou, Y. Zhou, N. P. Padture, *Acta Mater.* **187**, 112–121 (2020).
12. Z. Dai *et al.*, *Scr. Mater.* **185**, 47–50 (2020).
13. M. Gutwald *et al.*, *Sol. Energy Mater. Sol. Cells* **209**, 110433 (2020).
14. J.-H. Kim *et al.*, *MRS Bull.* **42**, 115–123 (2017).
15. J. H. Yun *et al.*, *J. Mater. Chem. A* **3**, 22176–22182 (2015).
16. B. L. Watson *et al.*, *ACS Appl. Mater. Interfaces* **8**, 25896–25904 (2016).
17. S. Jeong, I. Lee, T.-S. Kim, J.-Y. Lee, *Adv. Mater. Interfaces* **7**, 2001425 (2020).
18. S. Song *et al.*, *Adv. Energy Mater.* **11**, 2003382 (2021).
19. B. L. Watson, N. Rolston, A. D. Printz, R. H. Dauskardt, *Energy Environ. Sci.* **10**, 2500–2508 (2017).
20. Q. Dong *et al.*, *Nat. Commun.* **12**, 973 (2021).
21. J. Tong *et al.*, *Matter* **4**, 1365–1376 (2021).
22. P. J. O'Brien *et al.*, *Nat. Mater.* **12**, 118–122 (2013).
23. M. Saliba *et al.*, *Energy Environ. Sci.* **9**, 1989–1997 (2016).
24. C. Noguera, *Physics and Chemistry at Oxide Surfaces* (Cambridge Univ. Press, 1996).
25. G. Tumen-Ulzii *et al.*, *Commun. Mater.* **1**, 31 (2020).
26. Z. Liu *et al.*, *Nat. Energy* **5**, 596–604 (2020).
27. A. Ulman, *Chem. Rev.* **96**, 1533–1554 (1996).
28. A. Abate *et al.*, *Nano Lett.* **14**, 3247–3254 (2014).
29. P. Schulz, D. Cahen, A. Kahn, *Chem. Rev.* **119**, 3349–3417 (2019).
30. S. Y. Kim, S. J. Cho, S. E. Byeon, X. He, H. Y. Yoon, *Adv. Energy Mater.* **10**, 2002606 (2020).
31. L. Liu *et al.*, *J. Am. Chem. Soc.* **137**, 1790–1793 (2015).
32. G. Yang *et al.*, *J. Mater. Chem. A* **5**, 1658–1666 (2017).
33. Q. Jiang *et al.*, *Nat. Energy* **2**, 16177 (2017).
34. N. J. Jeon *et al.*, *Nat. Mater.* **13**, 897–903 (2014).
35. G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **6**, 15–50 (1996).
36. G. Kresse, J. Furthmüller, *Phys. Rev. B* **54**, 11169–11186 (1996).

37. J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
38. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
39. A. A. Tamijani, A. Salam, M. P. de Lara-Castells, *J. Phys. Chem. C* **120**, 18126–18139 (2016).
40. M. T. Weller, O. J. Weber, J. M. Frost, A. Walsh, *J. Phys. Chem. Lett.* **6**, 3209–3212 (2015).
41. J. Liu *et al.*, *Phys. Chem. Chem. Phys.* **22**, 19718–19724 (2020).
42. Y. Qi, L. G. Hector, *Phys. Rev. B* **69**, 235401 (2004).
43. G. Cavallo *et al.*, *Chem. Rev.* **116**, 2478–2601 (2016).
44. M. V. Khenkin *et al.*, *Nat. Energy* **5**, 35–49 (2020).
45. M. Saliba, M. Stolterfoht, C. M. Wolff, D. Neher, A. Abate, *Joule* **2**, 1019–1027 (2018).
46. D. Bi *et al.*, *Nat. Commun.* **9**, 4482 (2018).

ACKNOWLEDGMENTS

Experimental assistance from M. Hu, N. Pellet, and Y. Zhou is gratefully acknowledged. **Funding:** The work of Z.D., S.K.Y., M.C., and N.P.P was supported by the Office of Naval Research (grants N00014-17-1-2232 and N00014-20-1-2574) and the National Science Foundation (NSF; grant 1538893). The work of A.A. and Y.Q. was supported by NSF Center for Synthetic Organic Electrochemistry (grant 2002158). **Author contributions:** N.P.P. conceived of, designed, and supervised the research. Z.D., S.K.Y., and M.C. fabricated the thin films and the devices. Z.D. characterized and tested the thin films and the devices. A.A and Y.Q. performed the DFT calculations. N.P.P. and Z.D. interpreted the data and wrote the manuscript, with contributions from the other authors. **Competing interests:** The authors declare no competing interests. Provisional application for a patent has been filed. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/618/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 to S4
References

20 November 2020; resubmitted 25 February 2021
Accepted 6 April 2021
10.1126/science.abf5602

QUANTUM SYSTEMS

Direct observation of deterministic macroscopic entanglement

Shlomi Kotler[1,2]*†, Gabriel A. Peterson[1,2], Ezad Shojaee[1,2], Florent Lecocq[1,2], Katarina Cicak[1], Alex Kwiatkowski[1,2], Shawn Geller[1,2], Scott Glancy[1], Emanuel Knill[1,3], Raymond W. Simmonds[1], José Aumentado[1], John D. Teufel[1]

Quantum entanglement of mechanical systems emerges when distinct objects move with such a high degree of correlation that they can no longer be described separately. Although quantum mechanics presumably applies to objects of all sizes, directly observing entanglement becomes challenging as masses increase, requiring measurement and control with a vanishingly small error. Here, using pulsed electromechanics, we deterministically entangle two mechanical drumheads with masses of 70 picograms. Through nearly quantum-limited measurements of the position and momentum quadratures of both drums, we perform quantum state tomography and thereby directly observe entanglement. Such entangled macroscopic systems are poised to serve in fundamental tests of quantum mechanics, enable sensing beyond the standard quantum limit, and function as long-lived nodes of future quantum networks.

The idea that motion has a nonclassical nature dates back to the early days of quantum mechanics. One of the first triumphs of the theory was explaining the emission and absorption spectra of atoms by quantizing the motion of their electrons. Quantum mechanics is not limited to the atomic scale; in principle, it extends to all objects of all sizes. We expect that quantum behavior of macroscopic systems will enhance our ability to build more powerful sensing, communication, processing, and storage devices (*1*).

Many future applications of quantum technology rely heavily on entanglement; that is, on the ability to generate strong quantum correlations between separate objects. For entanglement to be useful, it must be prepared efficiently, followed by measurement and control with precision that is inversely proportional to the square root of the masses of the objects involved. The task becomes more difficult in the presence of noise, especially given that larger objects tend to interact more strongly with noisy environments and that the measurement process also introduces noise (Fig. 1A). The communication or processing protocol in which the entanglement might be used limits the amount of noise allowed before the entanglement is rendered useless.

Entanglement of mechanical motion was first demonstrated with two trapped atomic ions (*2*). It was generated deterministically and measured directly with high fidelity, and was therefore available as a resource that could be used for further processing. Taking the same level of quantum control and measurement from the atomic scale to macroscopic engineered objects then remained an outstanding challenge. Important experimental milestones toward this goal have been reached using optical photons in a probabilistic scheme (*3*) or microwave radiation with indirect inference (*4*).

Here we strongly and deterministically entangle two massive mechanical oscillators and directly observe their state. Our technology allows for on-demand reproducible entanglement generation. For direct observation of the entangled state, we implement a near quantum–limited measurement of the position and momentum quadratures of both mechanical oscillators in every realization of the experiment. By repeating these measurements, we completely characterize their joint covariance matrix. This tomography demonstrates clear evidence of continuous-variable (CV) entanglement (*5*) in the measurement signals, without noise subtraction.

The entangled state measured here manifests strong correlations between seemingly disparate systems. The two-oscillator system can be characterized by the first and second moments of the dimensionless quadratures of motion X_j and P_j ($j = 1, 2$), which satisfy the canonical commutation relations $[X_j, P_j] = i$ and are related to the underlying mechanical positions and momenta (*6*). After entanglement generation, the X-quadrature of each harmonic oscillator is drawn from a Gaussian probability distribution with a variance that is large compared to its zero-point fluctuations. However, when compared against each other, X_1 and X_2 are highly correlated, and similarly, P_1 and P_2 are anticorrelated. These features, however notable, could be consistent with classical correlations. To verify that the correlations originate from entanglement, we use the Simon-Duan criterion (*7*–*9*), calculated from the covariance matrix C of $\vec{S} = (X_1, P_1, X_2, P_2)$. Element j, k of the 4×4 covariance matrix is $C_{jk} = \frac{1}{2}\langle (S_k - \langle S_k \rangle)(S_j - \langle S_j \rangle) + (S_j - \langle S_j \rangle)(S_k - \langle S_k \rangle) \rangle$ where $\langle \ldots \rangle$ denotes expectation value. The smallest symplectic eigenvalue ν of the partially transposed covariance matrix quantifies the entanglement of the system (*9*, *10*) [see (*11*) for explicit expressions]. The two-oscillator state is entangled if $\nu < \frac{1}{2}$, where the zero-point fluctuations have variance $\frac{1}{2}$.

To extract the full covariance matrix with minimal assumptions, we measure all four quadratures of motion describing the two oscillators in each single experiment. Our method is analogous to a heterodyne measurement of electromagnetic radiation and allows more accurate covariance matrix estimation than its homodyne counterpart (*12*). Crucially, this concurrent measurement improves substantially if it is efficient. The inefficiency of our

[1]National Institute of Standards and Technology, Boulder, CO 80305, USA. [2]Department of Physics, University of Colorado, Boulder, CO 80309, USA. [3]Center for Theory of Quantum Matter, University of Colorado, Boulder, CO 80309, USA.
*Corresponding author. Email: shlomi.kotler@mail.huji.ac.il
†Present address: Department of Applied Physics, The Hebrew University of Jerusalem, Jerusalem, 9190401, Israel.

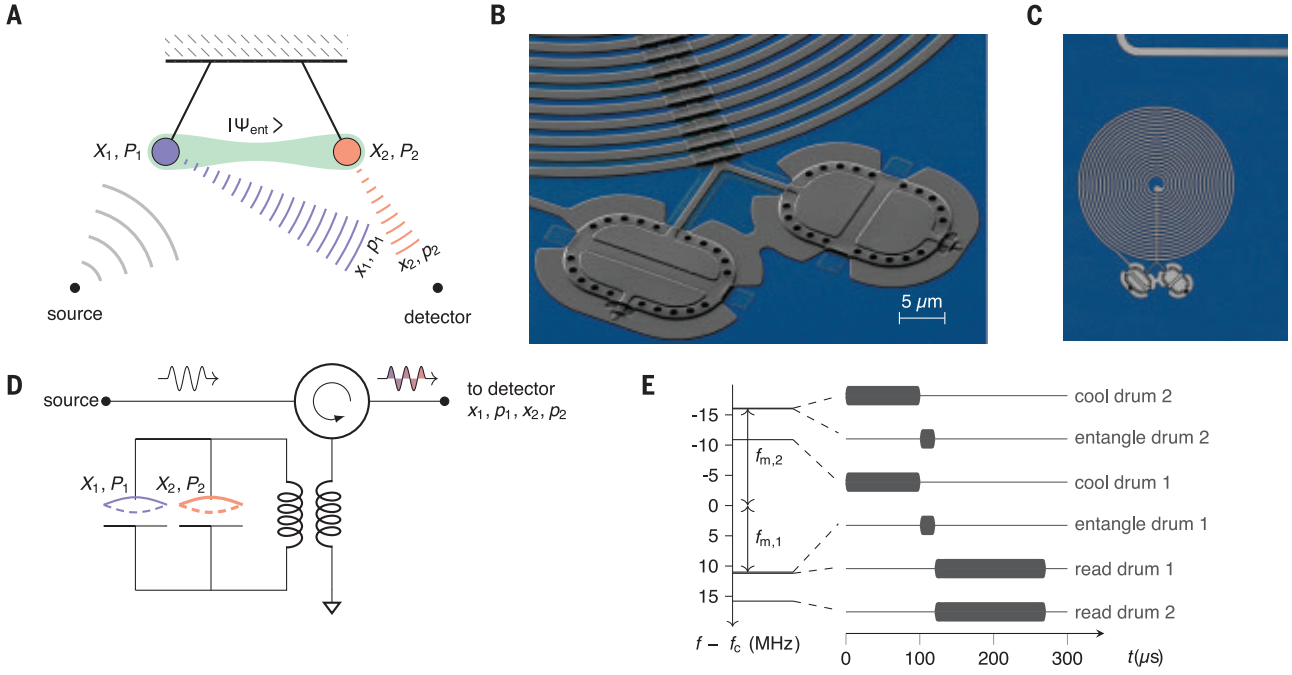


Fig. 1. Experiment overview. (**A**) Concept. Two mechanical harmonic oscillators (pendulums), characterized by their respective quadratures of motion X_1, P_1, X_2, P_2, are placed in an entangled state $|\psi_{ent}\rangle$. Electromagnetic radiation Doppler-shifts as it reflects off the moving pendulums, carrying information mixed with noise: x_1, p_1, x_2, p_2, due to inevitable loss effects. (**B**) Scanning electron micrograph (false color) of a device similar to the one used in this study. Two aluminum drums are suspended above a sapphire substrate, resulting in well-defined harmonic modes in the direction perpendicular to the substrate with frequencies $f_{m,1} = 10.9$ MHz (left drum) and $f_{m,2} = 15.9$ MHz (right drum). Each drum forms the top plate of a capacitor, along with the bottom plate, which is fixed to the substrate. (**C**) Device optical image (false color). A spiral inductor shunts the parallel capacitance of the two drums. Together they form a microwave cavity at a frequency of $f_c = 6.0806$ GHz. An input line (top right) inductively couples to the microwave cavity. (**D**) Circuit schematics. Mechanical motion (dashed) modulates the frequency of the microwave cavity. Therefore, an incoming pulse is Doppler-shifted as it reflects off the cavity, encoding information about the drums' quadratures of motion. Incoming and reflected pulses are separated using a circulator. (**E**) Experiment sequence. The carrier frequency of the incoming pulses determines the nature of their interaction with the drums. State initialization is achieved by sideband cooling each drum mode close to its ground state followed by a short entangling pulse. A readout pulse imprints an amplified record of the mechanical states onto the reflected microwave pulse.

measurement apparatus can be modeled as an effective beam splitter (*13*–*15*), where each variable describing the motion $S_j \in \{X_1, P_1, X_2, P_2\}$ becomes mixed with vacuum noise such that $s_j = \sqrt{\eta_j}\, S_j + \sqrt{1-\eta_j}\, \xi_j$, where η_j is the efficiency of the measurement of oscillator j, ξ_j is a Gaussian random variable with zero mean and vacuum variance $\langle \xi_j^2 \rangle = \frac{1}{2}$, and ξ_j and ξ_k are independently distributed for $j \neq k$. Therefore, following calibration of the measurement chain, we have direct access to the measured variables s_j and their minimal symplectic eigenvalue ν_{meas}. For the states measured here, $\nu_{meas} < \frac{1}{2}$ implies that $\nu < \frac{1}{2}$ and vice versa (*11*). This mutual relation takes an even simpler form if in addition the state is symmetric with respect to exchanging the roles of the two oscillators and undergoes symmetric loss: $\nu_{meas} = \eta\nu + (1-\eta)\frac{1}{2}$. In both cases, if efficiencies are low, ν_{meas} will approach $\frac{1}{2}$, and it will be more difficult to certify that $\nu < \frac{1}{2}$ with high confidence. Moreover, quantum information protocols, such as teleportation and entanglement swapping, require a high measurement efficiency, as demonstrated for light fields (*16*, *17*).

Our two mechanical oscillators are made of lithographically patterned thin-film aluminum that forms drum-like membranes (*18*), each with a mass of ≈70 pg, suspended above a sapphire substrate (Fig. 1B). We use the $f_{m,1} = 10.9$ MHz mode of the left drum and the $f_{m,2} = 15.9$ MHz mode of the right drum. We manipulate and measure the motion of the drums using electromechanics (*9*). The drums are embedded into a single microwave resonator, known as the "cavity," whose resonance frequency, centered at $f_c = 6.0806$ GHz, shifts according to the drums' motion (Fig. 1, C and D). A microwave pulse, reflected off the cavity, imparts forces on the drums and encodes the amplitudes of their quadratures of motion into Doppler-shifted sidebands. The carrier frequency of the incoming microwave pulse determines whether the drums are cooled, entangled, or measured (Fig. 1E). To entangle the two drums, we irradiate the cavity with two pulses simultaneously. One pulse has a carrier frequency $f_c + f_{m,1}$. This results in a two-mode squeezing (TMS) interaction that entangles the cavity with drum 1, by generating correlated photon-phonon pairs: photons at a frequency f_c and phonons at $f_{m,1}$ (*14*). The other pulse has a carrier frequency $f_c - f_{m,2}$. This results in a beam-splitter (BS) interaction that swaps cavity photons at a frequency f_c with phonons in drum 2 at a frequency $f_{m,2}$ (*13*). If the TMS and BS interactions were applied separately, the former would energize drum 1 and the latter would cool drum 2 (*11*). However, because we apply the TMS and BS pulses simultaneously, energy flows to both drums. Thus, the cavity mediates the interaction between the drums in a manner that is similar to other theoretical proposals (*19*–*27*). As a result, strong correlations form between the quadratures of motion.

Measurement is performed by amplifying a reflected microwave readout pulse after it has interacted with the drums. Typical microwave measurement efficiencies are ~0.01 even when using the best commercially available low-noise amplifiers. Here we achieve higher effective efficiency by using the TMS interactions native to the device as a preamplifier, on the basis of techniques that were developed for single-drum readout (*14*, *15*, *28*, *29*). We

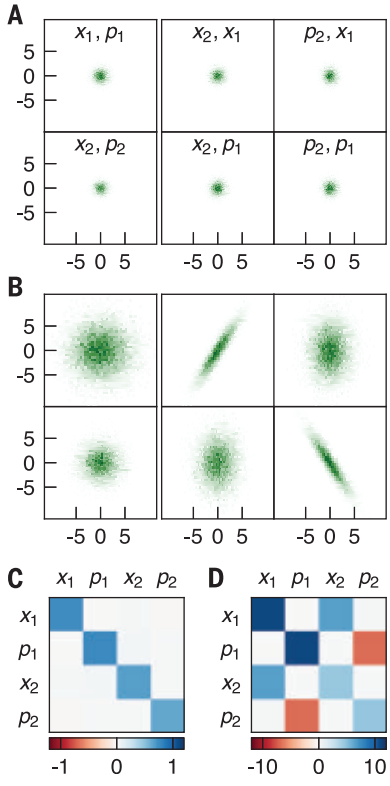


Fig. 2. Tomography of two mechanical oscillators. (**A**) Histograms of a sideband-cooled state of two drums. Each experiment records the system variables $\vec{s} = (x_1, p_1, x_2, p_2)$ concurrently. Variables are scaled to dimensionless units according to the canonical commutation relation $[x_j, p_j] = i$ for $j = 1, 2$, so the vacuum state has variance $\frac{1}{2}$. Panel with legend s_j, s_k corresponds to a correlation of s_j along the x axis and s_k along the y axis, quantified using a normalized 2D histogram of 10,000 repetitions of the experiment. The drums' individual variances are $V_1 = 0.75(1)$ and $V_2 = 0.63(1)$. (**B**) Histograms of an entangled state of two drums. After sideband cooling, a 16.8-μs entangling pulse generates x_1, x_2 correlation and p_1, p_2 anticorrelation. Because the entanglement pulse pumps energy into the two-drum system, each drum's individual variance grows from its ground-state cooled value to $V_1 = 10.9(1)$ and $V_2 = 4.63(5)$, respectively. (**C**) Covariance matrix of the data in (A). (**D**) Covariance matrix of the data in (B). Correlations and anticorrelations are apparent in the off-diagonal elements.

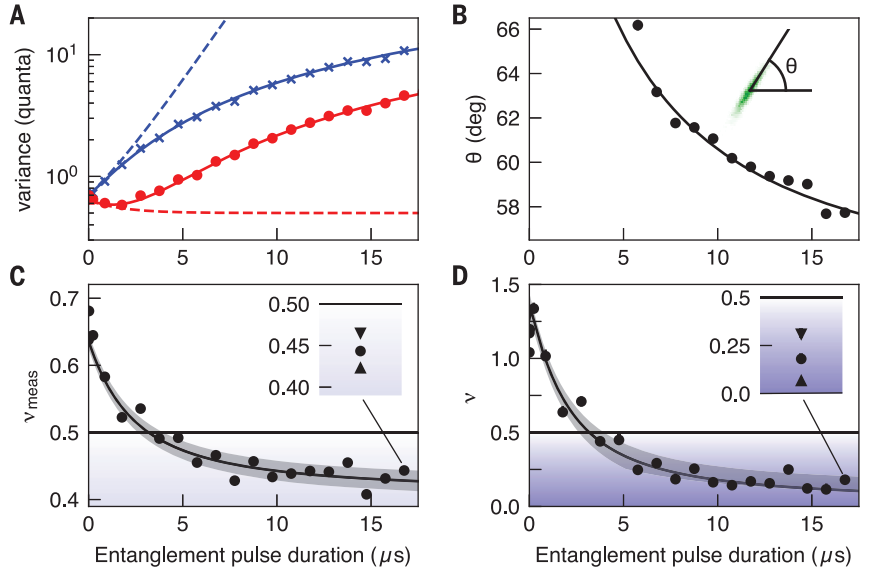


Fig. 3. Entanglement of two drums versus pulse duration. (**A**) Measured variances of individual drums. Entangling is composed of an energizing pulse for drum 1 and a cooling pulse for drum 2, applied simultaneously. The dashed lines show a theoretical prediction of the drums' individual variances if energizing and cooling were employed separately. Blue and red marks are the measured variances V_1 and V_2 for drums 1 and 2, respectively. Solid lines in all panels show theory with parameters obtained by fitting to an independent data set and without further adjustment (*11*). (**B**) Angle of the x_1, x_2 correlation that determines the squeezed and anti-squeezed joint quadratures of the bipartite system. (**C**) Entanglement in the measured variables, after loss, quantified by ν_{meas}. Points below $\frac{1}{2}$ indicate entanglement of the two drums. Statistical error bars, quantified by 1-σ bias-corrected bootstrapping confidence intervals, are smaller than the markers for most points. Shaded gray area corresponds to a 1-σ uncertainty region in the location of the black theoretical prediction curve caused by measurement efficiency uncertainties. Inset shows the systematic uncertainty (±1-σ) of the last measured point in the main graph (circle), indicated by the upper and lower points (triangles). The last point attains $\nu_{meas} = 0.44^{+0.004}_{-0.004}(\text{stat})^{+0.022}_{-0.021}(\text{sys})$. (**D**) Entanglement in the mechanical variables, prior to loss, quantified by ν. Uncertainties and inset plot are similar to those in (C). The last point attains $\nu = 0.18^{+0.03}_{-0.02}(\text{stat})^{+0.13}_{-0.11}(\text{sys})$.

extend these methods, using frequency multiplexing, and improve our measurement efficiencies by more than an order of magnitude: $\eta_1 = 0.26(2)$, and $\eta_2 = 0.153(3)$. Ultimately, the Heisenberg uncertainty principle prevents these efficiencies from exceeding $\frac{1}{2}$ because they quantify a concurrent measurement of both quadratures of motion (*6*).

Figure 2 shows tomography of the two-drum system, as characterized by its covariance matrix, for different protocols. First, we prepare a fiducial cold state by applying a pulse sequence of ground-state cooling followed by readout, rendering a single concurrent measurement of x_1, p_1, x_2, p_2. Figure 2A shows the experimental distribution of the measured variables for 10,000 repetitions of the experiment. The two-dimensional (2D) histograms show no correlation between any of the measured variables. This is reaffirmed by the covariance matrix of the state, which is diagonal to a good approximation (Fig. 2C). The magnitude of the diagonal elements correspond to nearly ground-state variances of $V_1 = 0.75(1)$ and $V_2 = 0.63(1)$ for drum 1 and 2, respectively, where $V_j = \frac{1}{2}\left(\langle (x_j - \langle x_j \rangle)^2 + (p_j - \langle p_j \rangle)^2 \rangle\right)$ for $j = 1, 2$. We now turn to a pulse sequence of ground-state cooling, entanglement, and readout. Figure 2B exhibits all the expected features of a highly correlated state. First, the x, p histogram for each drum is consistent with a Gaussian distribution of large variance. Drum 1, which undergoes a TMS interaction, has a bigger variance ($V_1 = 10.9(1)$) than that of drum 2 ($V_2 = 4.63(5)$), which undergoes a BS interaction. Second, a clear signature of drum-drum interaction is demonstrated by the correlation of x_1, x_2 and the anticorrelation of p_1, p_2. The covariance matrix of the measured variables in Fig. 2D displays a dominant diagonal and four off-diagonal elements $C_{1,3} \approx -C_{2,4}$ and $C_{3,1} \approx -C_{4,2}$. Indeed, these clear correlations are directly observable in the measured variables. Delineating them from classical correlations requires an application of the Simon-Duan criteria.

Evolution of the two-drum state is shown for different entangling pulse durations (Fig. 3), all of which are kept shorter than 100 μs to avoid the thermal decoherence of both drums (*11*). First, we focus on the individual variances of each drum (Fig. 3A). At short times < 1 μs, the drums show no evidence of interaction. Drum 2 cools while drum 1 becomes energized, the same behavior that would have been expected if the drums did not interact with one another. The dashed lines in Fig. 3A extrapolate this noninteracting behavior. Entangling pulse durations of > 1 μs result in both drums deviating from this independent evolution; their variances now grow together in time

with similar rates. Second, recall that in Fig. 2C, the histogram plot exhibited a correlation between x_1 and x_2. Theory predicts that the angle that the elliptical x_1, x_2 distribution's major axis makes with the horizontal will evolve as shown by the solid theory line in Fig. 3B. Third, we focus on the measured variable ν_{meas}, shown in Fig. 3C. Because our cooling of the drums is imperfect, their initial state contains some residual thermal motion in addition to the quantum fluctuations. The entangling operation must overcome this classical noise before the drums can be truly entangled. This is why, for pulse durations shorter than ~4 μs, $\nu_{meas} > \frac{1}{2}$, which strongly suggests that the drums are only classically correlated. For longer pulse durations, true quantum behavior, inconsistent with classical correlation, is observed and ν_{meas} crosses below $\frac{1}{2}$, indicating entanglement. At 16.8–μs interaction time, we observe $\nu_{meas} = 0.44^{+0.004}_{-0.004}(\text{stat})^{+0.022}_{-0.021}(\text{sys})$, where "sys" indicates systematic uncertainty dominated by uncertainty in the measurement efficiencies, and "stat" indicates statistical uncertainty estimated using bootstrapping (*11*). This is a direct measurement of entanglement for a bipartite system of macroscopic objects. It quantifies the amount of entanglement left in the system after noise processes have intervened during the measurement, and is therefore important for future quantum information applications. The amount of entanglement prior to the effect of noise can be estimated as well. To that end, we use a semidefinite program to find the closest (in l_2 distance) physically realizable covariance matrix that, after an interaction with noise, is consistent with the data. From that we estimate the entanglement criterion ν, which is shown in Fig. 3D. Indeed, we see more entanglement for the same interaction time, with $\nu = 0.18^{+0.03}_{-0.02}(\text{stat})^{+0.13}_{-0.11}(\text{sys})$. Such a level of entanglement might be useful for the exploration of mesoscopic Einstein-Podolsky-Rosen nonlocality (*30*) and fundamental tests of quantum mechanics (*27*).

Our results demonstrate pulsed, time-domain control of three important building blocks for CV quantum information processing and quantum communication: state initialization, entanglement, and measurement. Pulsed control played a key role. It allowed optimization of each piece separately and improved our measurement efficiency by more than an order of magnitude compared to traditional steady-state operation. As a result, we generated a highly entangled state of two macroscopic mechanical oscillators, surpassing the entanglement threshold by $4.43^{+0.6}_{-0.7}(\text{stat})^{+4.4}_{-2.3}(\text{sys})$ dB. Most excitingly, we observe entanglement directly in the measured variables. This is relevant to future applications that require decisions based on measurement outcomes. We therefore expect the methods described here to serve as a stepping stone for teleportation and entanglement swapping of states of massive objects. This would enable hybrid quantum networks, in which mechanics entangles with microwave fields (*14*) or with spin systems (*31*), or is used as an intermediary to entangle radiation (*32*, *33*).

REFERENCES AND NOTES

1. A. H. Safavi-Naeini, D. Van Thourhout, R. Baets, R. Van Laer, *Optica* **6**, 213 (2019).
2. J. D. Jost *et al*., *Nature* **459**, 683–685 (2009).
3. R. Riedinger *et al*., *Nature* **556**, 473–477 (2018).
4. C. F. Ockeloen-Korppi *et al*., *Nature* **556**, 478–482 (2018).
5. S. L. Braunstein, P. van Loock, *Rev. Mod. Phys.* **77**, 513–577 (2005).
6. C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, M. Zimmermann, *Rev. Mod. Phys.* **52**, 341–392 (1980).
7. R. Simon, *Phys. Rev. Lett.* **84**, 2726–2729 (2000).
8. L.-M. Duan, G. Giedke, J. I. Cirac, P. Zoller, *Phys. Rev. Lett.* **84**, 2722–2725 (2000).
9. W. P. Bowen, G. J. Milburn, *Quantum Optomechanics* (CRC Press, 2015).
10. A. Serafini, F. Illuminati, S. D. Siena, *J. Phys. At. Mol. Opt. Phys.* **37**, L21–L28 (2003).
11. See supplementary materials.
12. Y. S. Teo *et al*., *Phys. Rev. A* **95**, 042322 (2017).
13. T. A. Palomaki, J. W. Harlow, J. D. Teufel, R. W. Simmonds, K. W. Lehnert, *Nature* **495**, 210–214 (2013).
14. T. A. Palomaki, J. D. Teufel, R. W. Simmonds, K. W. Lehnert, *Science* **342**, 710–713 (2013).
15. R. D. Delaney, A. P. Reed, R. W. Andrews, K. W. Lehnert, *Phys. Rev. Lett.* **123**, 183603 (2019).
16. A. Furusawa *et al*., *Science* **282**, 706–709 (1998).
17. X. Jia *et al*., *Phys. Rev. Lett.* **93**, 250503 (2004).
18. J. D. Teufel *et al*., *Nature* **471**, 204–208 (2011).
19. S. Mancini, V. Giovannetti, D. Vitali, P. Tombesi, *Phys. Rev. Lett.* **88**, 120401 (2002).
20. M. Pinard *et al*., *Europhys. Lett.* **72**, 747–753 (2005).
21. M. J. Hartmann, M. B. Plenio, *Phys. Rev. Lett.* **101**, 200503 (2008).
22. S. G. Hofer, W. Wieczorek, M. Aspelmeyer, K. Hammerer, *Phys. Rev. A* **84**, 052327 (2011).
23. H. Tan, G. Li, P. Meystre, *Phys. Rev. A* **87**, 033829 (2013).
24. Y.-D. Wang, A. A. Clerk, *Phys. Rev. Lett.* **110**, 253601 (2013).
25. J. Li, I. M. Haghighi, N. Malossi, S. Zippilli, D. Vitali, *New J. Phys.* **17**, 103037 (2015).
26. L. F. Buchmann, D. M. Stamper-Kurn, *Phys. Rev. A* **92**, 013851 (2015).
27. R. Schnabel, *Phys. Rev. A* **92**, 012126 (2015).
28. F. Lecocq, J. D. Teufel, J. Aumentado, R. W. Simmonds, *Nat. Phys.* **11**, 635–639 (2015).
29. A. P. Reed *et al*., *Nat. Phys.* **13**, 1163–1167 (2017).
30. M. D. Reid, Q. Y. He, *Phys. Rev. Lett.* **123**, 120402 (2019).
31. R. A. Thomas *et al*., *Nat. Phys.* **17**, 228–233 (2021).
32. S. Barzanjeh *et al*., *Nature* **570**, 480–483 (2019).
33. J. Chen, M. Rossi, D. Mason, A. Schliesser, *Nat. Commun.* **11**, 943 (2020).

ACKNOWLEDGMENTS

We thank B. Katz and D. Ben-Zvi for feedback and insight on data taking and analysis. We thank B. Hauer and A. Sirois for their careful reading of the manuscript. We thank K. Lehnert and R. Delaney for useful discussions on measurement efficiency. We thank N. Kotler for consulting on data and concept visualization. **Funding:** At the time this work was performed, S.K., E.S. F.L., A.K., and S.Geller were supported as Associates in the Professional Research Experience Program (PREP) operated jointly by NIST and the University of Colorado Boulder under Award no. 70NANB18H006 from the U.S. Department of Commerce. **Author contributions:** J.D.T. and S.K. designed the experiment. S.K. and G.A.P. fabricated the device. F.L. and K.C. supervised device fabrication. S.K. performed the measurements. F.L., R.W.S., and J.A. advised on measurement techniques. S.K., E.S., A.K., and S. Geller developed the theory and wrote, analyzed, and tested the data analysis code. S.K. and E.S. analyzed the results. S.Glancy and E.K. supervised theory work and data analysis. S.K., G.A.P., F.L., K.C., R.W.S., J.A., and J.D.T. developed the experimental infrastructure necessary to conduct the experiment. J.D.T. supervised the work. All authors provided experimental suggestions, discussed the results, and contributed to the writing of the manuscript. **Competing interests:** The authors declare no competing interests. S.K. is also affiliated with Qedma Quantum Computing Ltd. G.A.P. is also affiliated with PsiQuantum. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. This is a contribution of the National Institute of Standards and Technology, not subject to U.S. copyright.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/622/suppl/DC1
Supplementary Text
Figs. S1 to S4
Tables S1 to S3
References (*34*–*42*)

16 October 2020; accepted 4 March 2021
10.1126/science.abf2998

QUANTUM SYSTEMS

Quantum mechanics–free subsystem with mechanical oscillators

Laure Mercier de Lépinay[1], Caspar F. Ockeloen-Korppi[1]†, Matthew J. Woolley[2], Mika A. Sillanpää[1]*

Quantum mechanics sets a limit for the precision of continuous measurement of the position of an oscillator. We show how it is possible to measure an oscillator without quantum back-action of the measurement by constructing one effective oscillator from two physical oscillators. We realize such a quantum mechanics–free subsystem using two micromechanical oscillators, and show the measurements of two collective quadratures while evading the quantum back-action by 8 decibels on both of them, obtaining a total noise within a factor of 2 of the full quantum limit. This facilitates the detection of weak forces and the generation and measurement of nonclassical motional states of the oscillators. Moreover, we directly verify the quantum entanglement of the two oscillators by measuring the Duan quantity 1.4 decibels below the separability bound.

Measuring a quantum mechanical system without disturbing it is not possible. Consider an oscillator of angular frequency ω_0 and dimensionless position $x(t)$ and momentum $p(t)$. The measurement of $x(t)$ causes $p(t)$ to suffer a disturbance called quantum back-action (QBA). The disturbance of momentum dynamically leads to a disturbance of position and therefore a fundamental limit on continuous position measurement. The oscillator's position can be written $x(t) = X \sin(\omega_0 t) + P \cos(\omega_0 t)$

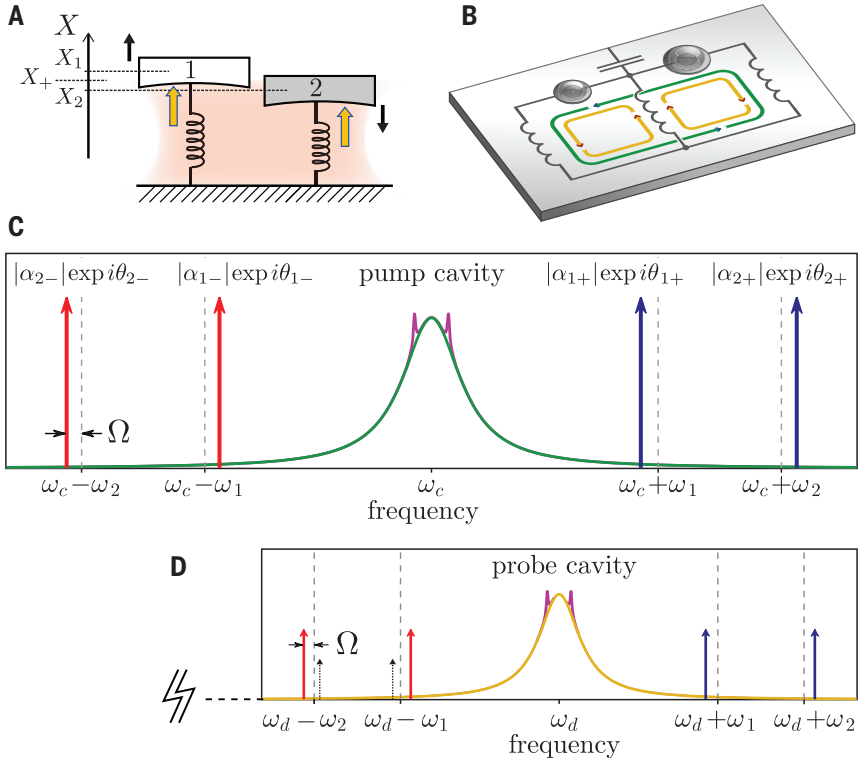


Fig. 1. Four-tone quantum back-action–evading measurement. (**A**) Illustration of a quantum mechanics–free subsystem (QMFS) constructed from an effective positive-mass oscillator (1, white) and an effective negative-mass oscillator (2, shaded) coupled to a cavity. Under the same force applied on both oscillators (wide arrows), such as a fluctuating radiation-pressure force, both momenta $P_{1,2}$ are changed by the same amount, but oscillator 2 is displaced in the opposite direction to oscillator 1 as a result of its negative mass. The sum momentum $P_+ = (P_1 + P_2)/\sqrt{2}$ is therefore dynamically coupled to the difference position $X_- = (X_1 - X_2)/\sqrt{2}$, just like the two quadratures of a single oscillator. However, unlike a standard oscillator's quadratures, these commute; as a result, $\{X_-, P_+\}$ span a QMFS, and similarly for $\{X_+, P_-\}$. (**B**) Physical realization with two Al drumhead oscillators coupled to two microwave resonances in a superconducting microcircuit. The arrows denote currents in the two modes (see fig. S1). (**C**) An optomechanical cavity (frequency ω_c) is pumped by four coherent tones at frequencies close to the red and blue sidebands of two mechanical oscillators having frequencies ω_1 and ω_2, detuned by Ω. (**D**) The oscillators are also coupled to an auxiliary (probe) cavity that is used for tomographic monitoring with four probe tones, as well as for sideband cooling (dashed black arrows). In (C) and (D), the set of small peaks centered on cavity resonance denote the sideband processes, spaced by 2Ω.

where, in a quantum mechanical framework, the quadratures of the motion X and P are noncommuting conjugate observables that cannot be known simultaneously with arbitrarily high precision: If X is measured, then P is subject to back-action, and vice versa.

In a back-action–evading (BAE) measurement strategy (*1*), a probe couples to only one quadrature of the oscillator's motion, say X. The back-action associated with this measurement disturbs the P quadrature, but the disturbance is not fed back to the measured X quadrature. Therefore, the X quadrature can be measured without any fundamental limit, at the expense of lost information on the P quadrature (*2*–*6*).

Remarkably, by constructing an effective oscillator from two physical oscillators, one can monitor both quadratures of an oscillator without any fundamental limit. This involves pairing of effective positive and negative mass oscillators, as illustrated in Fig. 1A (*7*–*10*), such that QBA is coherently canceled for two of their collective quadratures. This approach creates a quantum mechanics–free subsystem (QMFS) (*7*, *8*). This possibility is particularly crucial in several ultrasensitive measurements, exceeding conventional quantum limits, and in studies of nonclassical states of motion (*11*).

The principle of the scheme has been demonstrated in experiments using optomechanics and spin-mechanics hybrid systems (*4*, *5*, *12*). Here, we present direct evidence of the absence of QBA on the collective dynamics of a pair of mechanical oscillators. Our system consists of two aluminum membrane micromechanical oscillators (*13*) optomechanically coupled to two microwave cavity modes (Fig. 1B). One of the modes, called the pump cavity mode (Fig. 1C), is characterized by the frequency $\omega_c/2\pi$ and damping rate $\kappa/2\pi$ and is described by the annihilation and creation operators a, $a^\dagger$. The mechanical oscillators, labeled 1 and 2, have frequencies $\omega_1/2\pi$ and $\omega_2/2\pi$ and damping rates $\gamma_1/2\pi$ and $\gamma_2/2\pi$. The average damping rate is $\gamma \equiv (\gamma_1 + \gamma_2)/2$. The creation and annihilation operators for phonons are denoted b_j, $b_j^\dagger$, with j = 1, 2. Additionally, both oscillators are coupled to a probe cavity (Fig. 1D) of frequency $\omega_d/2\pi$ and damping rate $\kappa_d/2\pi$.

To create the effective interaction that realizes the QMFS, we drive the cavity with four coherent pump tones (*14*). The angular frequencies of the tones are $\omega_{1\pm} = \omega_c \pm (\omega_1 - \Omega)$ and $\omega_{2\pm} = \omega_c \pm (\omega_2 + \Omega)$ (Fig. 1C). Here, $\Omega > 0$ is a detuning from respective motional sidebands. We describe the system in a frame oscillating at ω_c for the cavity field, and $(\omega_1 - \Omega)$ and $(\omega_2 + \Omega)$ for the two oscillators. In this frame, the mechanical quadratures are $X_j = \left(b_j^\dagger + b_j\right)/\sqrt{2}$ and $P_j = i\left(b_j^\dagger - b_j\right)/\sqrt{2}$ for j = 1, 2, and the collective mechanical quadratures are $X_\pm = \frac{1}{\sqrt{2}}(X_1 \pm X_2)$ and $P_\pm = \frac{1}{\sqrt{2}}(P_1 \pm P_2)$.

With strong driving tones, the radiation-pressure interaction is linearized (*15*). Each pump tone gives rise to an effective complex-valued optomechanical coupling strength $|G_{j\pm}| \exp(i\theta_{j\pm})$ associated with the tone of frequency $\omega_{j\pm}/2\pi$. Initially, we choose all amplitudes $|G_{j\pm}| \equiv G$ to be equal, and define the cooperativity $C = 4G^2/(\kappa\gamma)$. We assume that $\omega_j >> \kappa$ and $|\omega_1 - \omega_2| >> \kappa$, and write the Hamiltonian as the sum of an uncoupled part and a coupling term: $H = H_0 + H_c$. The uncoupled part $H_0/\hbar = (\Omega/2)(X_1^2 + P_1^2 - X_2^2 - P_2^2)$, where $\hbar$ is the Planck constant divided by 2π, attributes a negative mass to oscillator 2, as is required to generate a QMFS (see Fig. 1A). In terms of collective quadratures (see eq. S4),

$$\frac{H_0}{\hbar} = \Omega(X_+X_- + P_+P_-)$$
$$\frac{H_c}{\hbar} = \frac{G}{2}a(A_-X_- + A_+X_+ + B_-P_- + B_+P_+) + \text{h.c.} \quad (1)$$

The coefficients $A_\pm$ and $B_\pm$ are functions of the pump tones' phases. Evolution under H_0 couples only pairs of commuting mechanical quadratures, so that each subsystem $\{X_-, P_+\}$ and $\{X_+, P_-\}$ can constitute a QMFS (*8*).

[1]QTF Centre of Excellence, Department of Applied Physics, Aalto University, FI-00076 Aalto, Finland. [2]School of Engineering and Information Technology, UNSW Canberra, Canberra, ACT 2600, Australia.
†Present address: IQM Finland Oy, Keilaranta 19 02150 Espoo, Finland.
*Corresponding author. Email: mika.sillanpaa@aalto.fi

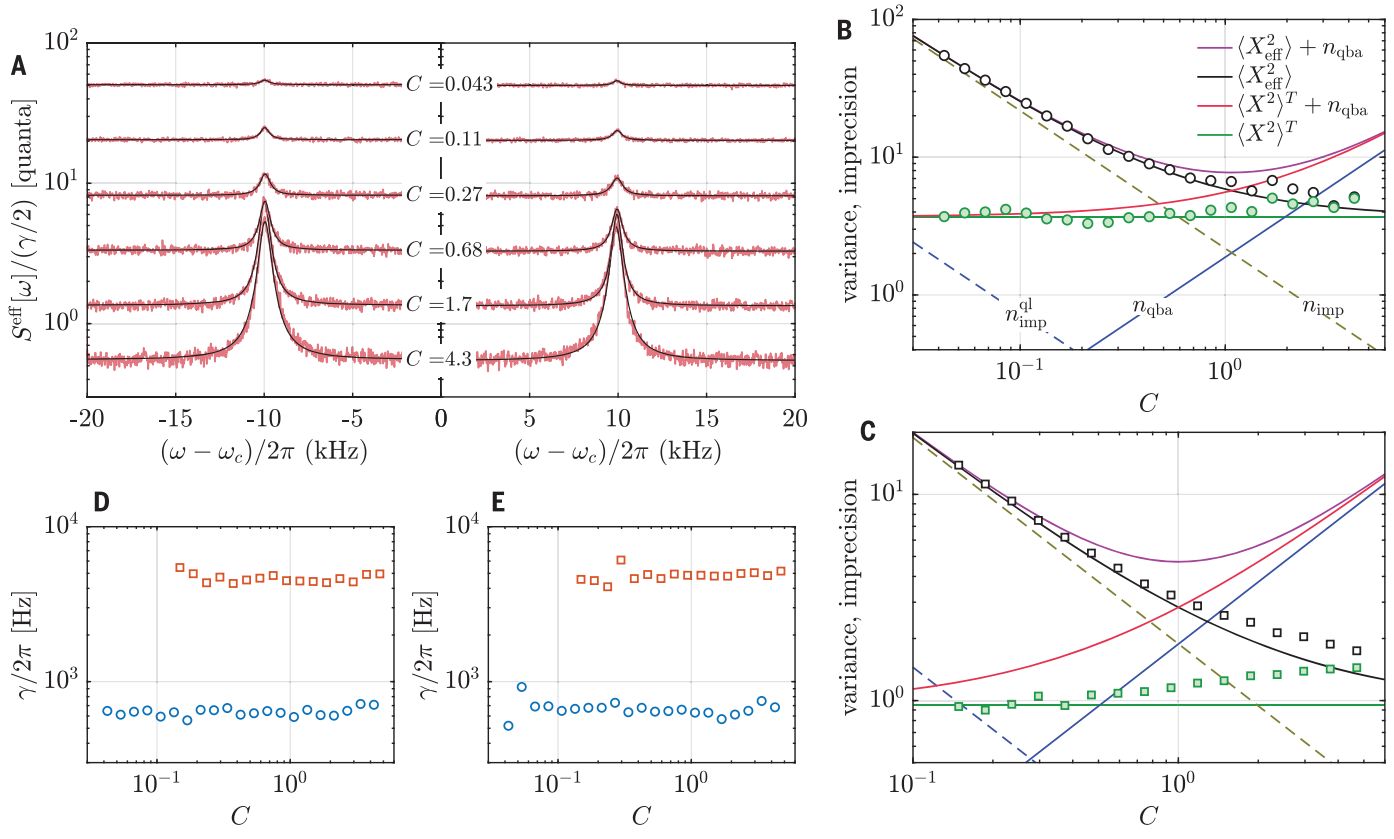


Fig. 2. Imprecision close to the full quantum limit via two-mode BAE monitoring. (**A**) The pump cavity output spectrum when measuring the X_+ quadrature shows two peaks whose signal-to-noise level improves as the cooperativity C is increased. (**B**) Mechanical noise of the measurement in (A) showing different contributions. Green circles, $\langle X^2\rangle$; black circles, $\langle X^2_{\rm eff}\rangle$. The theoretical lines are labeled. $N^{\rm ql}_{\rm imp}$ indicates the imprecision with $n_{\rm amp}$ = 0. (**C**) As (B), but with stronger sideband cooling. Coloring corresponds to that used in (B). (**D** and **E**) The linewidths corresponding to the left and right peaks in (A) (circles), and similarly for the measurement in (C) (rectangles). Pump detuning is $\Omega/2\pi$ = 10 kHz in (A) and (B), $\Omega/2\pi$ = 200 kHz in (C).

For certain combinations of $\theta_{j\pm}$ (fig. S6), a BAE measurement of any particular quadrature can be achieved. For example, $\theta_{1-} = \theta_{2+} = 0$ and $\theta_{1+} = \theta_{2-} = \phi$ realizes

$$\frac{H_c}{\hbar} = 2\sqrt{2}GX_c^{\phi}\left(X_+\cos\frac{\phi}{2} + P_-\sin\frac{\phi}{2}\right) \quad (2)$$

which couples any linear combination $X_+{}^{\phi} \equiv X_+\cos(\phi/2) + P_-\sin(\phi/2)$ of quadratures X_+ and P_- to a certain quadrature of the cavity field

$$X_c^{\phi} = \frac{a\exp\left(-i\frac{\phi}{2}\right) + a^{\dagger}\exp\left(i\frac{\phi}{2}\right)}{\sqrt{2}} \quad (3)$$

The cavity thus measures a given mechanical quadrature—for example, X_+ for $\phi = 0$. In this case there is back-action on the conjugate quadrature P_+ but, because the evolution of the subsystem $\{X_+, P_-\}$ is independent from P_+, the disturbance never leaks back to this subsystem, which remains an isolated QMFS.

Another case is that of $\theta_{1-} = \theta_{2-} = 0$ and $\theta_{1+} = \theta_{2+} \equiv \theta$, which measures an arbitrary linear combination of X_+ or P_+ depending on θ. Similarly, any linear combination of $X_\pm$ and $P_\pm$ can be measured with an appropriate choice of phases by detecting the field leaking out of the cavity, on which is imprinted the information on the mechanical quadrature (*4*, *16*).

Thermal and quantum fluctuations determine the different noise contributions of position measurement. The thermal occupation number of a single oscillator j in equilibrium with a bath of temperature T is $n_j{}^T = [\exp(\hbar\omega_j/k_BT) - 1]^{-1}$. The fluctuations of the collective quadratures, including the zero-point fluctuations, are described by their respective variances, e.g., $\langle X_+{}^2\rangle$, $\langle P_-{}^2\rangle$ for X_+, P_-. For simplicity, we do not explicitly write the quadrature label for the measured quadratures. If the two oscillators are in a thermal state, each collective quadrature's variance (*4*) is $\langle X^2\rangle = \langle X^2\rangle^T$, where $\langle X^2\rangle^T = \frac{1}{2}(n_1{}^T + n_2{}^T) + \frac{1}{2}$.

Even though QBA is associated with the shot noise of the pump tones (*17*–*19*), the variances of the measured quadrature and its pair in the QMFS are still given by $\langle X^2\rangle = \langle X^2\rangle^T$ thanks to BAE. An additional classical contribution (technical heating), however, makes $\langle X^2\rangle^T$ weakly power-dependent. The conjugate quadratures each receive QBA $n_{\rm qba} = 2C$ as well as a small classical contribution $n_{\rm cba} = 4Cn_c{}^T$ due to cavity thermal occupation $n_c{}^T$ (*4*, *16*), such that the variances of the quadratures orthogonal to the QMFS are $\langle X^2\rangle = \langle X^2\rangle^T + n_{\rm qba} + n_{\rm cba}$.

The measurement also suffers from imprecision noise that can be written (see eqs. S32 and S33) as an equivalent collective quadrature occupation $n_{\rm imp} \approx [1/(8C)](n_{\rm amp} + \frac{1}{2})$, here dominated by the number of noise quanta $n_{\rm amp}$ added by the amplifier. In the case of a continuous non-BAE measurement, the tradeoff between $n_{\rm imp}$ and $n_{\rm qba}$ defines the standard quantum limit (SQL) (*20*), where the added noise equals the zero-point motion noise. The SQL has been approached (*21*, *22*) and recently surpassed in an off-resonant case (*23*). We refer to the resonant-case SQL as the "full quantum limit" (FQL). Although thermal noise represents a severe obstacle to reaching the SQL, approaching the FQL remains even more challenging. In the BAE case, SQL and FQL can be achieved for the two concerned quadratures simultaneously, provided that their thermal noise does not dominate.

The output spectrum $S_{\rm out}(\omega)$ from the pump cavity consists of two peaks located at $\pm\Omega$ with respect to the cavity resonance frequency (Fig. 1C). With properly chosen pump phases, these peaks faithfully reproduce the spectrum of the

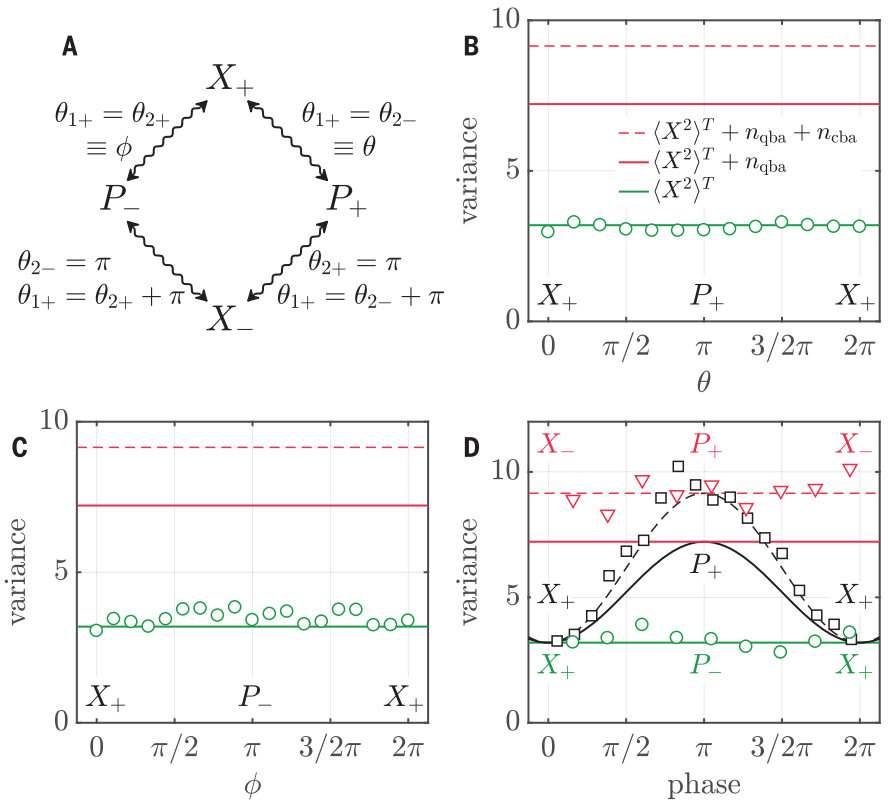


Fig. 3. Moving within the quantum mechanics–free subsystem. (**A**) Simultaneously shifting two phases in four-tone BAE as indicated allows the measurement of particular linear combinations of the collective quadratures. This holds for the pumping as well as for tomographic probing. The other unspecified phases in each quadrant of the figure equal zero. (**B**) Pump cavity signal when shifting between X_+ and P_+. The circles denote the measured $\langle X^2\rangle$. The theoretical lines depict the ideal QBA model, and they are colored similarly to the corresponding data points. (**C**) As (B), but shifting from X_+ to P_-. (**D**) Probe cavity signal while the pump cavity performs a strong BAE measurement of X_+. The phase sweep is executed as follows: ϕ (circles), θ (rectangles), or along the route $P_+ \Leftrightarrow X_-$ in (A) (triangles). These sweeps correspond to moving between different collective quadratures as labeled. The black solid line is the prediction including QBA only; the black dashed line also includes n_{cba}. In (D), the detuning is $\Omega/2\pi$ = 200 kHz; in other panels, $\Omega/2\pi$ = 10 kHz. The pump cooperativity is $C \approx 2.1$, $\langle X^2\rangle^T \approx 3.2$, and $\gamma/2\pi \approx 630$ Hz.

measured mechanical quadrature. After amplification, the recorded spectrum sits on top of a microwave noise background (n_{amp} + ½), which contributes to an effective equivalent mechanical occupation in the inferred mechanical spectrum $S^{eff}[\omega]$ (eq. S33). The effective mechanical variance can then be determined as $\langle X_{eff}^2\rangle \equiv \gamma S^{eff}[\pm\Omega]/2 = \langle X^2\rangle^T + n_{imp}$.

The measurements are carried out in a dry dilution refrigerator with a base temperature of 8 mK, where we find the mechanical oscillators to have the equilibrium thermal phonon numbers $n_1^{T,0} \approx 32$ and $n_2^{T,0} \approx 24$, which give $\langle X^2\rangle^{T,0} \approx 28$ corresponding to an effective bath temperature of 10.5 mK. The pump cavity has a frequency $\omega_c/2\pi \approx 4.98$ GHz and a damping rate $\kappa/2\pi \approx 1.58$ MHz, which is divided into internal $\kappa_I/2\pi \approx 130$ kHz and external $\kappa_E/2\pi \approx 1.45$ MHz damping. The mechanical frequencies are $\omega_1/2\pi \approx 6.692$ MHz and $\omega_2/2\pi \approx 9.032$ MHz, and the intrinsic damping rates are $\gamma_1^0/2\pi \approx 55$ Hz and $\gamma_2^0/2\pi \approx 84$ Hz. The probe cavity frequency is $\omega_d/2\pi \approx 6.62$ GHz, and the damping rates are $\kappa_d/2\pi \approx 1.17$ MHz, $\kappa_I^d/2\pi \approx 350$ kHz, and $\kappa_E^d/2\pi \approx 820$ kHz.

The microwave tones are generated from separate phase-locked oscillators (figs. S2 and S3). We calibrate the cooperativity of each red-detuned tone individually through its sideband cooling effect. Single-mode BAE measurements (*2*) are used to match the cooperativities of the blue-detuned tones. Additionally, we apply two cooling tones (Fig. 1D) to the probe cavity in order to independently sideband-cool both oscillators. This broadens the mechanical linewidths up to $\gamma_1 \approx \gamma_2 = \gamma >> \gamma_1^0, \gamma_2^0$ and reduces the variance $\langle X^2\rangle^T$ far below the value $\langle X^2\rangle^{T,0}$ reached with cryogenic cooling alone.

To demonstrate BAE, we choose all pump phases to be equal to zero so as to measure X_+. In Fig. 2A, we display the measured noise spectra when initially $\langle X^2\rangle^T \approx 3.2$ ($\gamma/2\pi \approx 630$ Hz). From the spectra, we extract the effective imprecision noise n_{imp} as well as the mechanical variance $\langle X^2\rangle$. The latter is nearly independent of the measurement strength, as seen in Fig. 2B. At large cooperativities, it is clear that both $\langle X_{eff}^2\rangle$ and $\langle X^2\rangle$ stay well below n_{qba}, showing a nearly ideal BAE and thus the existence of a QMFS. We then intensify the sideband cooling down to $\langle X^2\rangle^T \approx 1.0$ ($\gamma/2\pi \approx 4.6$ kHz), obtaining the noise displayed in Fig. 2C. All measured noise figures are smaller, $\langle X_{eff}^2\rangle$ falling 7 dB below the QBA level at high cooperativity and $\langle X^2\rangle$ reaching 8 dB below. The effective noise is less than a factor of 2 from FQL ($\langle X_{eff}^2\rangle = 1$). This exceeds the best values reported thus far, achieved in cold-atom optomechanics (*18*). Reaching the FQL is here precluded by thermal noise, which is enhanced by technical heating of the oscillators (*4*, *11*). The linewidth of the peaks is not affected by the measurement (Fig. 2, D and E).

To demonstrate the ability to rotate the QMFS, we now change two phases synchronously. This operation is represented in Fig. 3A. In Fig. 3, B and C, we show the measured noise while sweeping the quadrature probed by the probe cavity between X_+ and P_+, or X_+ and P_-, respectively. Because QBA is always directed to an unmeasured quadrature, the measured fluctuations remain at the same back-action–free level. To examine the impact of back-action on the conjugate quadratures, we apply another set of four BAE tones, enabling realization of a tomography of the mechanical state under a pump-cavity BAE measurement through the probe cavity (Fig. 1D). The tomography measures the variance of a generalized collective quadrature (e.g., X_+^ϕ). The tomographic angles (e.g., ϕ) for the probe-cavity BAE measurement are now determined by the pump-cavity BAE setup.

The variance of the generalized quadrature is $\langle (X_+^\phi)^2\rangle = \langle X_+^2\rangle \cos^2(\phi/2) + \langle P_-^2\rangle \sin^2(\phi/2) + \langle X_+P_-\rangle \sin\phi$, including a cross-correlation term $\langle X_+P_-\rangle$. Although $\langle X_+P_-\rangle$ is not directly accessible, the modeling (see supplementary materials) shows that it is negligible. The amplitude of the tomography tones is chosen to be much smaller (by ~ −15 dB) than those driving the pump cavity; thus, we can neglect QBA exerted by the probe cavity in comparison to that from the pump cavity. Calibrations similar to those for the pump cavity allow for a readout of the fluctuations in all the collective quadratures (figs. S4 and S5). As seen in Fig. 3D, the tomography shows a reasonable agreement with predictions based on the assumption that the back-action consists of QBA only, without adjustable parameters. The additional heating of the conjugate quadratures is associated to a small $n_c^T \approx 0.23$.

Finally, we use the four-tone setup to create and detect quantum entanglement between the two mechanical oscillators (*11*, *24*). Entanglement for continuous-variable states can be characterized by inequalities involving $\langle X_+^2\rangle$

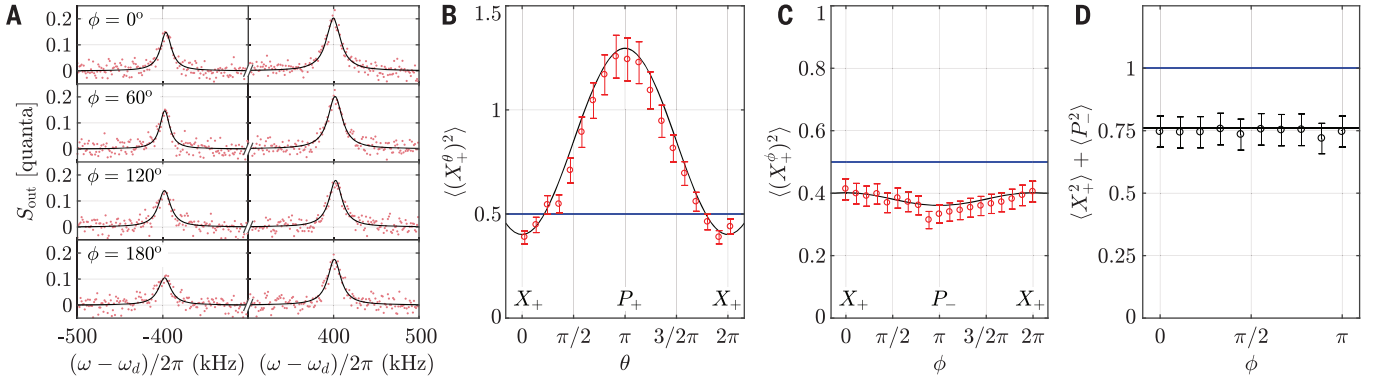


Fig. 4. Direct measurement of stabilized quantum entanglement of two mechanical oscillators. (**A**) Output spectrum peaks in tomography carried out through the probe cavity while scanning between the noiseless quadratures X_+ and P_-. At the different phase values $\phi = 0$ (top) to $\phi = \pi$ (bottom), the measurement probes the linear combinations $\{X_+\}$, $\{\frac{\sqrt{3}}{2}X_+ + \frac{1}{2}P_-\}$, $\{\frac{1}{2}X_+ + \frac{\sqrt{3}}{2}P_-\}$, and $\{P_-\}$. The solid lines are Lorentzian fits. (**B**) Calibrated variances of the system of the noiseless quadrature X_+ and the noisy P_+ swept by changing the probing phase. (**C**) Integrated variance from (A). (**D**) Duan quantity extracted from (C) by summing two generalized quadratures spaced by π. The horizontal axis is the starting phase in (C). The blue lines in (B) and (C) mark the vacuum fluctuation levels, and in (D), the blue line gives the separability criterion of the quantum state. The red-detuned effective coupling $|G_{j-}|/2\pi \approx 121$ kHz, and $|G_{j+}/G_{j-}| \approx 0.37$, corresponding to $C \approx 490$. In (B), the detuning $\Omega/2\pi = 100$ kHz, and in (C) and (D), $\Omega/2\pi = 400$ kHz. The black lines in (B) to (D) are theoretical predictions using $\langle X^2 \rangle^T = 54$. The error bars in (B) to (D) indicate 95% statistical confidence bounds.

and $\langle P_-^2 \rangle$ (*25*, *26*), and it can be stabilized (*11*), transient (*24*, *27*), or conditional on a measurement record (*28*). According to the Duan criterion (*26*, *29*), two oscillators are entangled if $\langle X_+^2 \rangle + \langle P_-^2 \rangle < 1$. In an earlier study (*11*), two-tone BAE tomography (*16*) allowed access to the conjugate quadratures X_+ and P_+ only, and P_- had to be inferred on the basis of other information. Using four-tone driving and four-tone BAE tomography, we can access both X_+ and P_- directly.

To generate entanglement, we modify the BAE setup by reducing the amplitudes of the blue tones applied to the pump cavity (*14*, *30*, *31*). In this experiment, we do not use cooling tones in the probe cavity and $\langle X^2 \rangle^T = \langle X^2 \rangle^{T,0}$. The probe cavity is used for tomography, providing the spectra shown in Fig. 4A whose integration gives the variances, as previously. Scanning the tomography between quadratures X_+ and P_+, we show in Fig. 4B a large contrast between $\langle P_+^2 \rangle$ and $\langle X_+^2 \rangle$, the latter reaching below the level of vacuum fluctuations. Changing the probe phase configuration to align the probe QMFS with the pump QMFS $\{X_+, P_-\}$, we demonstrate in Fig. 4C that the variances inside this subspace are almost independent on the tomographic angle, as hinted by the modest variation of peak amplitudes in Fig. 4A, and that all measured variances remain clearly below the level of vacuum fluctuations. Finally, to assess the Duan criterion, we show in Fig. 4D the sum of variances of two orthogonal quadratures X_+^ϕ and $X_+^{\phi+\pi}$ taken for different values of ϕ. The sum satisfies the Duan criterion by a 1.4-dB margin, constituting a direct and robust measurement of the entanglement of two massive mechanical oscillators.

We have demonstrated the monitoring of two quadratures of an effective oscillator without quantum back-action disturbance to the oscillator, which, according to a common paradigm, is not possible. This allows for a complete characterization of a weak classical force driving an oscillator and will have practical relevance when cavity optomechanical techniques become available for sensitive measurements in the quantum regime at room temperature. The directly demonstrated entanglement of two massive oscillators is a further tool to reduce intrinsic noise in such measurements (fig. S7). Combined with squeezing of probe electromagnetic fields and phase-sensitive amplification, it could allow for noiseless monitoring of weak external forces.

REFERENCES AND NOTES

1. V. B. Braginsky, Y. I. Vorontsov, K. S. Thorne, *Science* **209**, 547–557 (1980).
2. J. Suh *et al.*, *Science* **344**, 1262–1265 (2014).
3. C. U. Lei *et al.*, *Phys. Rev. Lett.* **117**, 100801 (2016).
4. C. F. Ockeloen-Korppi *et al.*, *Phys. Rev. Lett.* **117**, 140401 (2016).
5. C. B. Møller *et al.*, *Nature* **547**, 191–195 (2017).
6. I. Shomroni, L. Qiu, D. Malz, A. Nunnenkamp, T. J. Kippenberg, *Nat. Commun.* **10**, 2086 (2019).
7. M. Tsang, C. M. Caves, *Phys. Rev. Lett.* **105**, 123601 (2010).
8. M. Tsang, C. M. Caves, *Phys. Rev. X* **2**, 031016 (2012).
9. K. Zhang, P. Meystre, W. Zhang, *Phys. Rev. A* **88**, 043632 (2013).
10. A. Motazedifard, F. Bemani, M. H. Naderi, R. Roknizadeh, D. Vitali, *New J. Phys.* **18**, 073040 (2016).
11. C. F. Ockeloen-Korppi *et al.*, *Nature* **556**, 478–482 (2018).
12. T. Caniard, P. Verlot, T. Briant, P.-F. Cohadon, A. Heidmann, *Phys. Rev. Lett.* **99**, 110801 (2007).
13. J. D. Teufel *et al.*, *Nature* **471**, 204–208 (2011).
14. M. J. Woolley, A. A. Clerk, *Phys. Rev. A* **89**, 063805 (2014).
15. M. Aspelmeyer, T. J. Kippenberg, F. Marquardt, *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
16. M. J. Woolley, A. A. Clerk, *Phys. Rev. A* **87**, 063846 (2013).
17. T. P. Purdy, R. W. Peterson, C. A. Regal, *Science* **339**, 801–804 (2013).
18. S. Schreppler *et al.*, *Science* **344**, 1486–1489 (2014).
19. J. D. Teufel, F. Lecocq, R. W. Simmonds, *Phys. Rev. Lett.* **116**, 013602 (2016).
20. A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, R. J. Schoelkopf, *Rev. Mod. Phys.* **82**, 1155–1208 (2010).
21. J. D. Teufel, T. Donner, M. A. Castellanos-Beltran, J. W. Harlow, K. W. Lehnert, *Nat. Nanotechnol.* **4**, 820–823 (2009).
22. G. Anetsberger *et al.*, *Phys. Rev. A* **82**, 061804 (2010).
23. D. Mason, J. Chen, M. Rossi, Y. Tsaturyan, A. Schliesser, *Nat. Phys.* **15**, 745–749 (2019).
24. R. Riedinger *et al.*, *Nature* **556**, 473–477 (2018).
25. S. M. Tan, *Phys. Rev. A* **60**, 2752–2758 (1999).
26. L.-M. Duan, G. Giedke, J. I. Cirac, P. Zoller, *Phys. Rev. Lett.* **84**, 2722–2725 (2000).
27. S. Kotler *et al.*, *Science* **372**, 622 (2021).
28. R. A. Thomas *et al.*, *Nat. Phys.* **17**, 228–233 (2020).
29. M. D. Reid *et al.*, *Rev. Mod. Phys.* **81**, 1727–1751 (2009).
30. Y.-D. Wang, A. A. Clerk, *Phys. Rev. Lett.* **110**, 253601 (2013).
31. H. Tan, G. Li, P. Meystre, *Phys. Rev. A* **87**, 033829 (2013).

ACKNOWLEDGMENTS

We acknowledge the facilities and technical support of Otaniemi research infrastructure for Micro and Nanotechnologies (OtaNano), part of the European Microkelvin Platform. **Funding:** Supported by ARC Centre for Engineered Quantum Systems (CE170100009) and AFOSR FA 2386-18-1-4026 through the Asian Office of Aerospace Research and Development (AOARD) (M.J.W.) and by the Academy of Finland (contracts 308290, 307757, 312057), the European Research Council (615755-CAVITYQPD), and the Aalto Centre for Quantum Engineering. The work was performed as part of the Academy of Finland Centre of Excellence program (project 336810). We acknowledge funding from the European Union's Horizon 2020 research and innovation program under grant agreement 732894 (FETPRO HOT). **Author contributions:** M.A.S. supervised the project and was involved in all subsequent stages. L.M.d.L. and C.F.O.-K. designed and fabricated the devices. L.M.d.L. carried out most of the experiments. M.J.W. suggested the experiment and developed the theory. All authors participated in the writing of the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/625/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
References (*32*, *33*)

4 November 2020; accepted 1 March 2021
10.1126/science.abf5389

FERROELECTRICS

Reversible oxygen migration and phase transitions in hafnia-based ferroelectric devices

Pavan Nukala[1,2]*, Majid Ahmadi[1,3], Yingfen Wei[1,3], Sytze de Graaf[1], Evgenios Stylianidis[1,4], Tuhin Chakrabortty[2], Sylvia Matzen[5], Henny W. Zandbergen[6], Alexander Björling[7], Dan Mannix[8,9,10], Dina Carbone[7], Bart Kooi[1,3], Beatriz Noheda[1,3]*

Unconventional ferroelectricity exhibited by hafnia-based thin films—robust at nanoscale sizes—presents tremendous opportunities in nanoelectronics. However, the exact nature of polarization switching remains controversial. We investigated a $La_{0.67}Sr_{0.33}MnO_3/Hf_{0.5}Zr_{0.5}O_2$ capacitor interfaced with various top electrodes while performing in situ electrical biasing using atomic-resolution microscopy with direct oxygen imaging as well as with synchrotron nanobeam diffraction. When the top electrode is oxygen reactive, we observe reversible oxygen vacancy migration with electrodes as the source and sink of oxygen and the dielectric layer acting as a fast conduit at millisecond time scales. With nonreactive top electrodes and at longer time scales (seconds), the dielectric layer also acts as an oxygen source and sink. Our results show that ferroelectricity in hafnia-based thin films is unmistakably intertwined with oxygen voltammetry.

The discovery of silicon-compatible nanoferroelectricity in hafnia-based thin films (*1*) has triggered a vast amount of fundamental research and has rejuvenated interest in ferroelectric materials in microelectronics, for both low-power, nonvolatile memory and logic devices. This ferroelectricity is robust even at film thicknesses as low as 1 nm (*2*–*4*), a situation that was believed to be impossible on the basis of the classical understanding of ferroelectricity. The spontaneous polarization observed in these films is ascribed generally to a metastable polar orthorhombic (o-) phase ($Pca2_1$) (*5*). A higher-energy rhombohedral (r-) phase (*R3m*/*R3*) has been reported for epitaxial growth of $Hf_{0.5}Zr_{0.5}O_2$ (HZO) on $La_{0.67}Sr_{0.33}MnO_3$ (LSMO)–buffered perovskite and on trigonal substrates (*6*). Remnant polarization (P_r) values as high as 35 μC/cm^2 were measured (*7*) on HZO. The pronounced effects of particle size reduction, surface effects, dopants, oxygen vacancies ($\ddot{V}_o$), epitaxial strain, and residual stresses at the nanoscale have been investigated as possible reasons to stabilize these otherwise metastable phases in thin films (*5*, *8*, *9*). On the other end of the size range, films as thick as 1 μm (*10*) and bulk samples (*11*) are ferroelectric, with the stabilization resulting from dopant and defect chemistry (*10*, *12*).

Armed with an understanding of the virgin state polarization, research is being conducted on the mechanism of polarization switching. The dynamics of this process, through the lens of a nucleation-limited switching model and pointing out the negligible role of domain growth (or domain wall motion), have been studied for doped hafnia films grown in the o-phase (*13*–*19*). Flat phonon bands and localized dipoles in half unit cells in the o-phase have been postulated as an intrinsic reason for switching without forming domain walls in these systems (*20*).

Hafnia-based and zirconia-based materials are also an important class of resistive memory devices and oxygen conductors that exhibit memristive hysteresis driven by $\ddot{V}_o$ conduction and redox reactions (*21*, *22*). More generally, in devices of thin-film ferroelectric oxides such as tunnel junctions, both $\ddot{V}_o$ migration and polarization switching lead to hysteresis (*23*–*28*). Understanding whether these effects are synergetic or independent is crucial to achieve device control (*29*–*33*). In tunnel junctions of HZO thin films on LSMO-buffered $SrTiO_3$ (STO), Wei *et al.* (*23*) observed a divergence of the tunnel electroresistance (TER) from 100% to 10^6% upon device cycling, which the authors explained as a possible transition from polarization switching to $\ddot{V}_o$ migration–assisted switching. This observation suggested that the two mechanisms are independent. Sulzbach *et al.* (*24*) have also reported a similar divergence in the TER as a function of the applied voltage in HZO layers before breakdown. On the other hand, the electric polarization in hafnia has been theoretically proposed to originate from oxygen vacancies through electrostrictive effects, which strongly suggests the extrinsic nature of the polarization switching (*29*). Direct structural observations during polarization switching can potentially resolve these controversies (*34*).

We report operando atomic-scale electron microscopy investigations of LSMO/HZO/LSMO capacitor stacks grown on conducting (Nb-doped) STO substrate under an electric field (*35*). LSMO is a standard choice of bottom electrode in complex oxide devices (*23*–*26*, *31*), so our conclusions are relevant for understanding a wider class of devices. We performed in situ biasing measurements while employing two scanning transmission electron microscopy (STEM) imaging modes—high-angle annular dark field (HAADF)–STEM and integrated differential phase contrast (iDPC)–STEM. With iDPC-STEM, we recently imaged H atoms next to Ti metal atoms (*36*), demonstrating that this is a robust atomic-resolution imaging technique for simultaneously measuring heavy and light elements. By directly imaging oxygen, we provide evidence of the reversible and hysteretic migration of $\ddot{V}_o$ from the bottom to the top electrode through the HZO layer. Associated with such migration, we show that $\ddot{V}_o$ induce phase transitions in the LSMO (bottom electrode) and HZO layers. Additionally, through operando x-ray diffraction, ex situ microscopy, and transport measurements on devices with both noble (Au) and oxygen reactive top electrodes (LSMO and TiN), we show that oxygen voltammetry is also found at short time scales (millisecond and less). We believe these observations clearly show that polarization switching and oxygen voltammetry are not independent.

We begin by showing the evolution of the epitaxial LSMO layer (bottom electrode), with the voltage applied to the top electrode and keeping the bottom electrode at 0 V (fig. S1) (*22*). Our iDPC-STEM image of the virgin state (Fig. 1A and fig. S2A) shows the antiphase octahedral δ tilts present in the LSMO perovskite structure (*22*). Mn-O-Mn bond angles, as we measured in various regions, are between 165° and 176° (*37*). Upon increasing the bias to 2 V, a noticeable displacement of Mn columns away from the center of the oxygen octahedra (Fig. 1B) appears throughout the film (barring the first three monolayers at the interface with Nb:STO). These displacements are randomly oriented with a mean value of 18.6 pm and standard deviation of 10.2 pm (fig. S3), which indicates a transformation from an MnO_6 octahedral toward an MnO_5 square pyramidal coordination (*35*). Thus at 2 V, the LSMO film contains a combination of MnO_5 and MnO_6 polyhedra. Although this structural feature was not previously observed for LSMO, Brownmillerite [(BM), oxygen-deficient perovskite] phases are reported to exhibit MnO_5 square pyramids in the parent

[1]Zernike Institute of Advanced Materials, University of Groningen, 9747 AG Groningen, Netherlands. [2]Centre for Nano Science and Engineering, Indian Institute of Science, Bengaluru, 560012, India. [3]CogniGron (Groningen Cognitive Systems and Materials Center), University of Groningen, 9747 AG Groningen, Netherlands. [4]Department of Physics and Astronomy, University College London, London WC1E 6BT, UK. [5]Center for Nanoscience and Nanotechnology, Paris-Saclay University, CNRS, 91120 Palaiseau, France. [6]Kavli Institute of Nanoscience, Faculty of Applied Sciences, Delft University of Technology, 2628 CJ Delft, Netherlands. [7]MAX IV Laboratory, Lund University, SE-221 00 Lund, Sweden. [8]University Grenoble Alpes, CNRS, Institut Néel, 38042 Grenoble, France. [9]European Spallation Source, SE-221 00 Lund, Sweden. [10]Department of Chemistry, Aarhus University, DK-8000 Aarhus, Denmark.
*Corresponding author. Email: pnukala@iisc.ac.in (P.N.); b.noheda@rug.nl (B.N.)

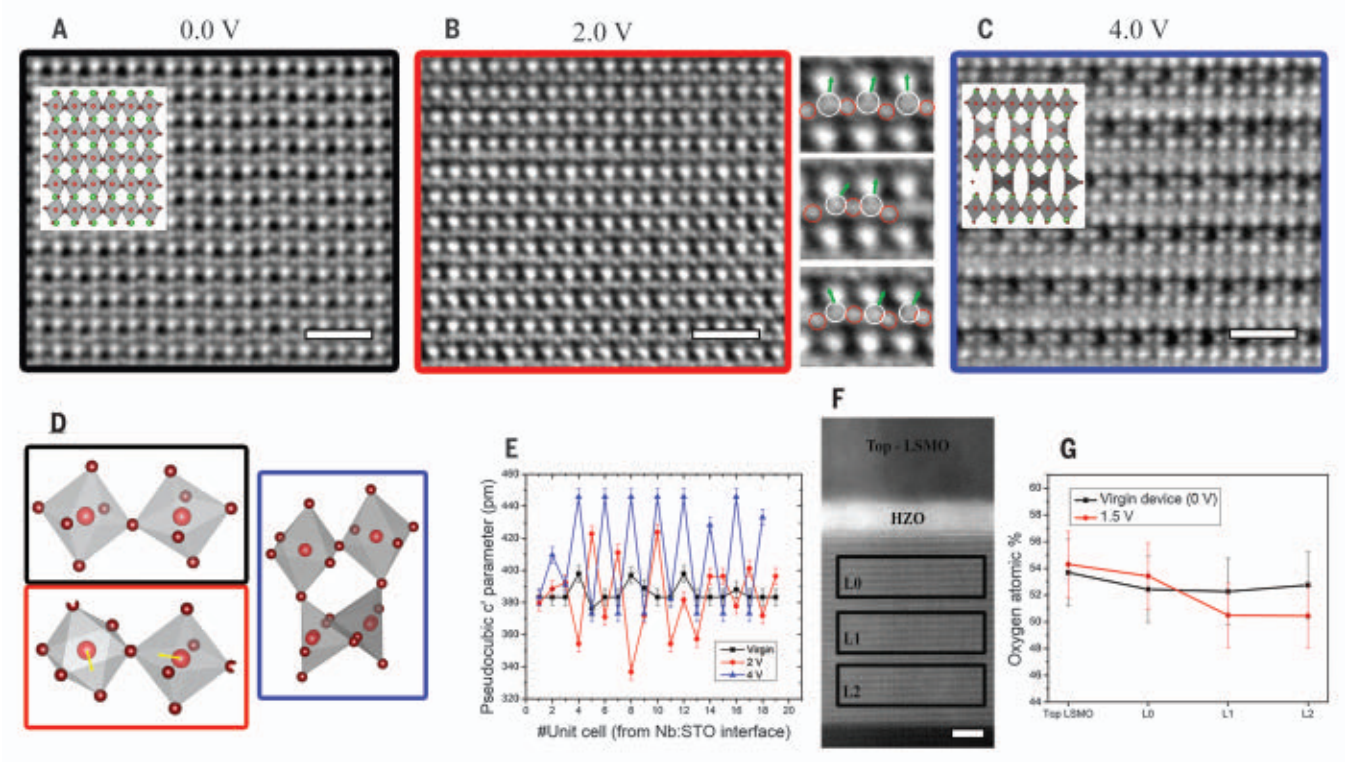


Fig. 1. Deoxygenation of bottom electrode LSMO layer with increasing positive bias. (**A**) iDPC-STEM image of a representative region of the bottom LSMO layer in the virgin state, viewed along the [110] zone axis, exactly matching the perovskite structure. Schematic in the inset shows MnO_6 octahedra and their antiphase tilts, clearly imaged in (A). In the inset, La/Sr is green, Mn is red, and O is brown. (**B**) iDPC-STEM image at V = 2 V. Panels on the right show various unit cells, illustrating Mn columns (circled in white) and their displacements (marked by green arrows) away from the center of an octahedron. Oxygen columns are marked in red circles. (**C**) iDPC-STEM at 4 V. BM LSMO (zone axis: *a*; schematic in inset) denoted by alternating MnO_4 tetrahedra and MnO_6 octahedra along *c*′. (**D**) Schematic showing the evolution of an MnO_6 octahedra in the virgin state (enclosed in black box), toward MnO_5 square pyramids at 2 V (enclosed in red box), to alternating MnO_4 tetrahedra and MnO_6 octahedra at 4 V (enclosed in blue box). (**E**) Plot of variation of *c*′ (La-La distance) parameter from the STO interface in perovskite (black), BM-precursor (red), and BM phases (blue). (**F** and **G**) Overview image of LSMO/HZO/LSMO capacitor (F) with regions marked where oxygen content was quantified from EDS (G) at 0 V and 1.5 V. Scale bars, 1 nm in (A) to (C) and 5 nm in (F).

compound, $SrMnO_3$ (*38*). We refer to this MnO_5-MnO_6 combination as a BM-precursor phase (*39*). As for the first three monolayers, an important feature is the exaggerated antiphase δ tilts, with Mn-O-Mn bond angles of 143° to 146° that are not typical of perovskite structures (fig. S2, B to D).

Upon increasing the biasing voltage to 4 V, LSMO converts into a well-studied BM phase (*40*, *41*) except for the first few monolayers near the interface with the substrate, which transform to the BM-precursor phase (fig. S4A). Transformation from perovskite to the BM phase occurs through $\ddot{V}_o$ ordering in every alternate Mn-O plane along the *c*′-axis (Fig. 1C), transforming the Mn coordination from octahedral or square pyramidal to tetrahedral. Back-to-back MnO_4 tetrahedra in [1-10] alternate with MnO_6 octahedra along *c*′, indicating the BM phase (Fig. 1C). This transformation is hysteretic and nonvolatile, with LSMO remaining in the BM phase even when the external bias is removed (fig. S4) (*35*).

The multiple-step transformations (Fig. 1D) from MnO_6 octahedra (virgin state) toward square pyramids plus octahedra (2 V), to alternating octahedra and tetrahedra (4 V) also correlate to the variation of the pseudocubic lattice parameter along the electric-field direction (called *c*′). We determined the *c*′ values at various bias voltages for the first 20 monolayers in LSMO starting from the Nb:STO interface (Fig. 1E). In the virgin state, we measured *c*′ to be 384 (±5) pm with $\ddot{V}_o$ disorder–induced expansion in some planes (*35*). At 2 V, *c*′ oscillates with values between 335 and 425 pm, without any superstructure. At 4 V, except for the first few monolayers, *c*′ alternates between 375 and 445 pm, doubling the lattice periodicity. Energy-dispersive spectroscopy (EDS) (*35*) reveals a clear gradient of oxygen concentration in the bottom electrode compared with the virgin state, even at a low bias of 1.5 V, with more $\ddot{V}_o$ occurring closer to the Nb:STO interface (Fig. 1, F and G).

Notably, the BM phase can be reoxygenated when negative voltages are applied to the top electrode. We show that this takes place for biases as low as –1 V (fig. S5A). The hysteretic BM phase clearly begins to reoxygenate in iDPC-STEM images of the same field of view at 0 and –1.3 V (Fig. 2, A and B), with the appearance of extra oxygen columns at –1.3 V in the Mn-O planes that were oxygen deficient at 0 V (fig. S5A). Upon ramping the bias to –3 V, the entire layer converts to the BM-precursor phase (Fig. 2C), which is retained also when the bias is removed (Fig. 2D), as we confirmed by the corresponding disorderly *c*′ variation (fig. S5B).

To address the time scales of the processes associated with de- and reoxygenation of LSMO layers, we followed the dynamics through HAADF-STEM image acquisition, after poling at –4 V (transforming LSMO completely back to the starting perovskite phase). Upon increasing the bias to 3 V (Fig. 2E), the initial perovskite phase changes to the BM-precursor phase in 60 s and then to the BM phase within 120 s. These changes are indicated by the variations we observe in the *c*′ parameter. By applying

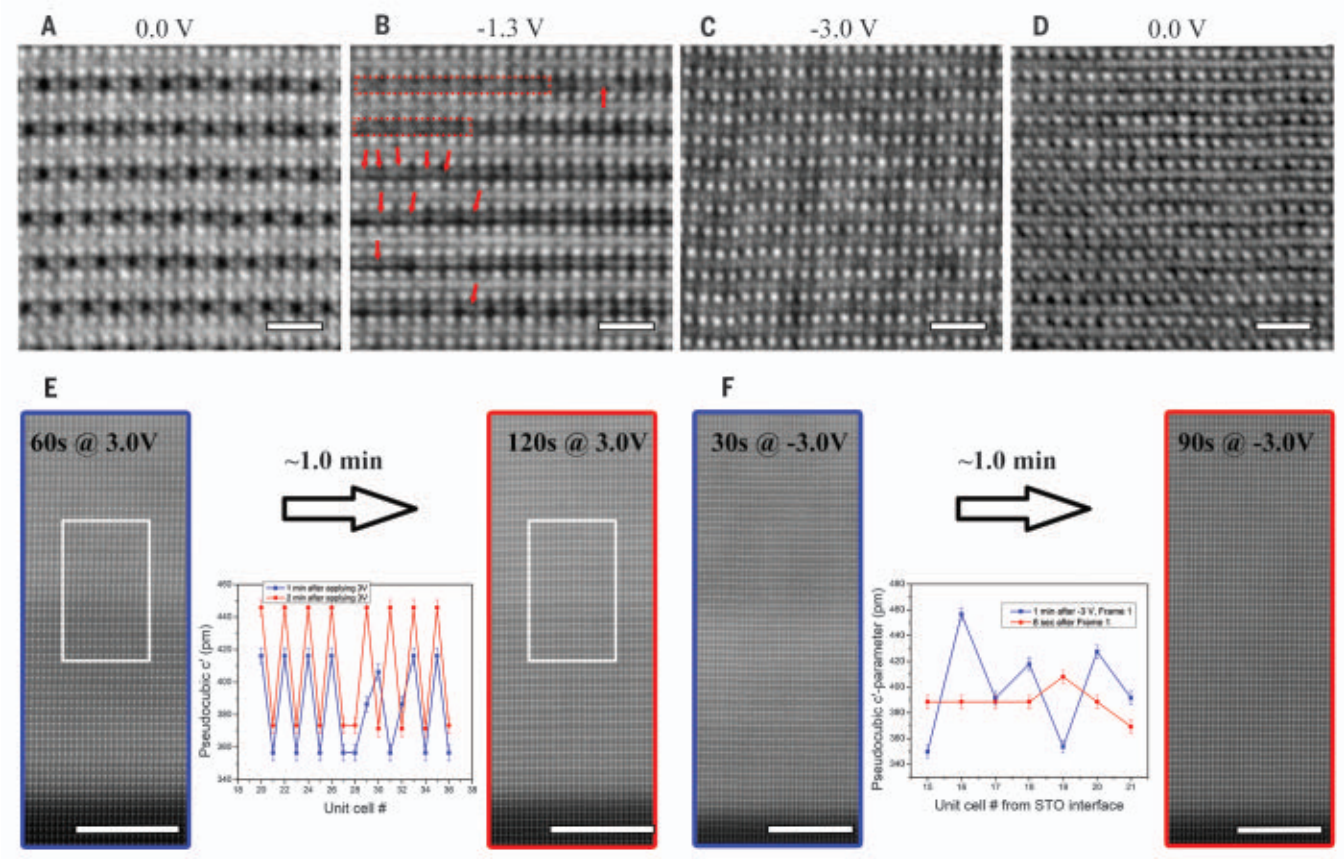


Fig. 2. Oxygenation of bottom electrode LSMO layer with increasing negative bias. (**A** and **B**) iDPC-STEM images of a region in the same field of view at 0 (A) and −1.3 V (B). At −1.3 V, oxygen columns start to appear (marked by red arrows) in positions where there were none in the BM phase at 0 V. (**C** and **D**) BM phase transforms to BM-precursor phase at −3 V (C) and is retained so at 0 V (D). (**E**) Dynamics are recorded through HAADF-STEM imaging within 120 s of ramping from 0 to 3 V from a starting perovskite phase. A BM-precursor phase is imaged at 60 s, and a BM phase is imaged at 120 s. (**F**) Upon changing the bias to −3 V, a BM phase is recorded at 30 s that changes to a perovskite phase by 90 s (for disappearance of the superstructure spots in FFT, see fig. S4). The intermediate BM-precursor phase recorded at 60 s converts to a perovskite-like phase in ~6 s, as can be seen by the variation in the *c*′ parameter, which gives an idea about the time scales of $\ddot{V}_o$ migration. Scale bars, 1 nm in (A) to (D) and 5 nm in (E) and (F).

−3 V, a complete transformation from a BM phase back to perovskite phase occurs within 90 s in the region [Fig. 2F; fast Fourier transform (FFT) in fig. S5C]. We then follow the same region for 10 s using a faster HAADF-STEM image series (1.2 s per frame). From the *c*′ parameter variations, we conclude that the BM-precursor phase transformed to the perovskite phase (Fig. 2F, center panel). Thus, although the complete transformation from perovskite to BM and back takes about a couple of minutes at 3 and −3 V, $\ddot{V}_o$ migration and partial phase-transitions already start occurring in time scales of seconds at these voltages. At 2 V, however, the partial transition to the BM-precursor phase itself takes 3 to 4 hours. The change in kinetics with voltage is consistent with the ultra-nonlinear voltage-time dilemma typically observed in oxide-resistive memories (*42*). Thus, we expect that at higher voltages these mechanisms will occur at very short time scales.

We also monitored the structural evolution in the HZO (6 nm) layer under the application of bias. From the multislice iDPC-STEM image simulations for HZO [see also (*43*)] in the r-phase (*R3m*) with [111] out-of-plane (Fig. 3A, inset), we recognize the (001) planes (at ~55° with respect to the [111] direction) by cationic (Hf/Zr) columns surrounded by two oxygen columns on either side of them. In the virgin state, our experimental images perfectly match the r-phase simulations. We followed the evolution of a supercell (Fig. 3A) in this grain upon application of bias along the out-of-plane [111] direction. We show the displacement of $\ddot{V}_o$ for this supercell occurs (*35*) (fig. S6A) with respect to the 0-V configuration in Fig. 3B. While $\ddot{V}_o$ migrate toward the bottom electrode with increasing bias (Fig. 3B), they also gather some in-plane displacement (Fig. 3B, inset).

At 4 V, the same grain transforms into a combination of multiple grains (Fig. 3C). Upon inspecting various regions in the film, we found that the majority of the grains have changed their structure from r-phase to the more thermodynamically stable o- and monoclinic (m-) phases (Fig. 3D and fig. S6C). The o-phase is commonly observed in ferroelectric HZO layers grown by various methods (*5*), whereas the r-phase is only observed under specific growth procedures and conditions (*6*). Our observations on HZO point to the r-phase being stabilized under slightly oxygen-deficient conditions. Replenishment of oxygen in the HZO layer under bias (originating from the bottom LSMO layer) transforms it into more stoichiometric m- or o-phases. The $\ddot{V}_o$ in the HZO layer, and thus the r-phase, is restored (by reverse migration) upon applying a bias of −3 V, as can be seen from the perfect match of the experimental iDPC-STEM of two representative domains (180° rotated from each other) (Fig. 3E) with the multislice image simulations (Fig. 3A, inset).

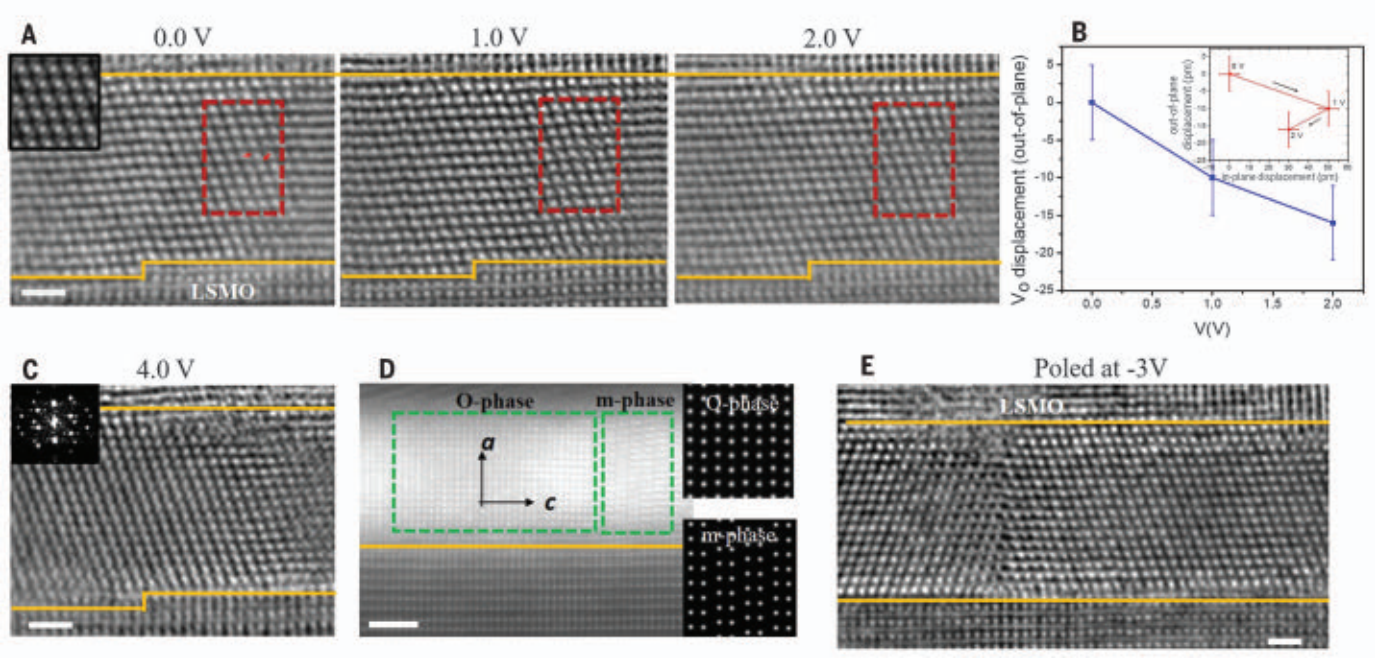


Fig. 3. Oxygenation and deoxygenation of HZO and associated phase transitions. (**A**) Evolution of an r-phase HZO grain while oxygenating under positive bias followed through iDPC-STEM images, where both cations and oxygen columns are displayed. The image of the virgin state at 0 V shows two red arrows pointing to two oxygen columns in the (100) planes neighboring a cationic column. Multislice iDPC-STEM image simulations of the r-phase (*R3m* symmetry) in the inset shows the good match with the observations. (**B**) Out-of-plane displacement of $\ddot{V}_o$ (in picometers) with external bias, in the marked supercell (red box) with respect to the positions in (A). Negative values indicate displacement toward bottom electrode. $\ddot{V}_o$ shows both in-plane and out-of-plane (toward bottom electrode) components (inset). (**C**) A new grain nucleates in the same region at 4 V, giving rise to a polycrystalline nature (FFT shown in inset). (**D**) Another region in the HZO film back at 0 V showing o-phase and m-phase (with multislice simulations of both in the insets). Note the change of orientation from [111] to [100]. (**E**) iDPC-STEM image of domains (mutually rotated by 180° about [111]) in the r-phase [to be compared with the simulation in the inset of figure 3A from (*36*)], which is retained when poled at −3 V (imaged at 0 V). Scale bars, 1 nm in (A), (C), and (E) and 2 nm in (D). Interfaces between the HZO and the top and bottom LSMO layers are marked in orange. In (D), only the HZO/bottom electrode interface is shown.

The m- and o-phases can be reversibly obtained again in subsequent cycles of positive bias (fig. S6D).

To elucidate the effects of device cycling with submillisecond pulses (*35*), we present results on tunnel junction Co/HZO (2 nm)/LSMO//STO devices switched between low-resistance state (LRS) and high-resistance state (HRS) at large voltages of ±6 V. The devices showed an increasing TER (*35*) from 100% (stage A) to 10^6% (stage B) upon cycling ~100 times (*23*, *44*). An iDPC-STEM image from a selected region in the LSMO (bottom electrode) layer in the LRS (stage B) (Fig. 4A) shows a clear perovskite structure. In the HRS (stage B), however (Fig. 4B), Co inhomogeneously oxidizes exhibiting coexisting Co-rich and CoO_x-rich regions as revealed from the EDS analysis (fig. S7). This leaves oxygen-deficient LSMO in the BM-precursor phase. Thus, progression from stage A to stage B results in the gradual increase of CoO_x regions, an accumulated effect of oxygen voltammetry, which is reflected electrically in diverging resistance values.

To disentangle the short-term field effects on our devices from accumulated effects (*35*), we report ex situ structure-property correlation results on ferroelectric capacitor stacks of LSMO (or TiN)/HZO (7 nm)/LSMO//STO cycled at 1 kHz fewer than 10 times at 5.5 V. The intrinsic P_r of our virgin devices in r-phase obtained from atomic displacements (*35*) (fig. S5B) is <9 μC/cm^2, which is very small compared with the 35 μC/cm^2 that were measured from polarization-voltage (*P-V*) loops at room temperature (*7*). This discrepancy is already an indication that most of the switching charge is intertwined with extrinsic factors. From the *P-V* hysteresis loops at various temperatures (Fig. 4C), we observed that, contrary to what is expected in classical ferroelectrics, the P_r increases with increasing temperature in the range from 150 to 300 K (*22*). This observation is in line with the polarization switching being correlated to the thermally activated oxygen-migration mechanism. These devices were prepared in a "down-polarized" configuration (5.5 V, 1 kHz) and imaged. The iDPC-STEM images clearly reveal the oxygen-deficient BM-precursor phase in the bottom LSMO layer close to the HZO interface (Fig. 4D) as also confirmed by corresponding disorderly oscillations in the *c*′ parameter (Fig. 4D, inset). Close to the STO interface, we found exaggerated oxygen octahedral tilts with Mn-O-Mn bond angles <146° (marked in fig. S8A), which is the same interfacial feature observed during in situ STEM DC (direct current) testing at 2 V (fig. S2B). These observations also suggest that such extreme tilts initiate the transformation from perovskite to BM-precursor phase. These incipient oxygen migration–induced topotactic structural transitions result from the discharge of $\ddot{V}_o$ in the electrodes. This thermally activated process leads to a built-in field at lower temperatures, which decreases as the temperature rises (fig. S8, B and C). The HZO layer, however, remains in an r-phase (Fig. 4E), revealing that it just acts as a conduit of oxygen between the source and sink (both reactive) electrodes in the short term.

We conducted comparative operando synchrotron nanobeam diffraction on Au/HZO/LSMO and LSMO/HZO/LSMO capacitors to answer the question of whether oxygen voltammetry can still occur if HZO is interfaced with a nonreactive electrode in the short term (*35*) (figs. S9 and S10). Both devices showed ferroelectric switching peaks (Fig. 4C and fig. S10), and we cycled them at 100 Hz in a

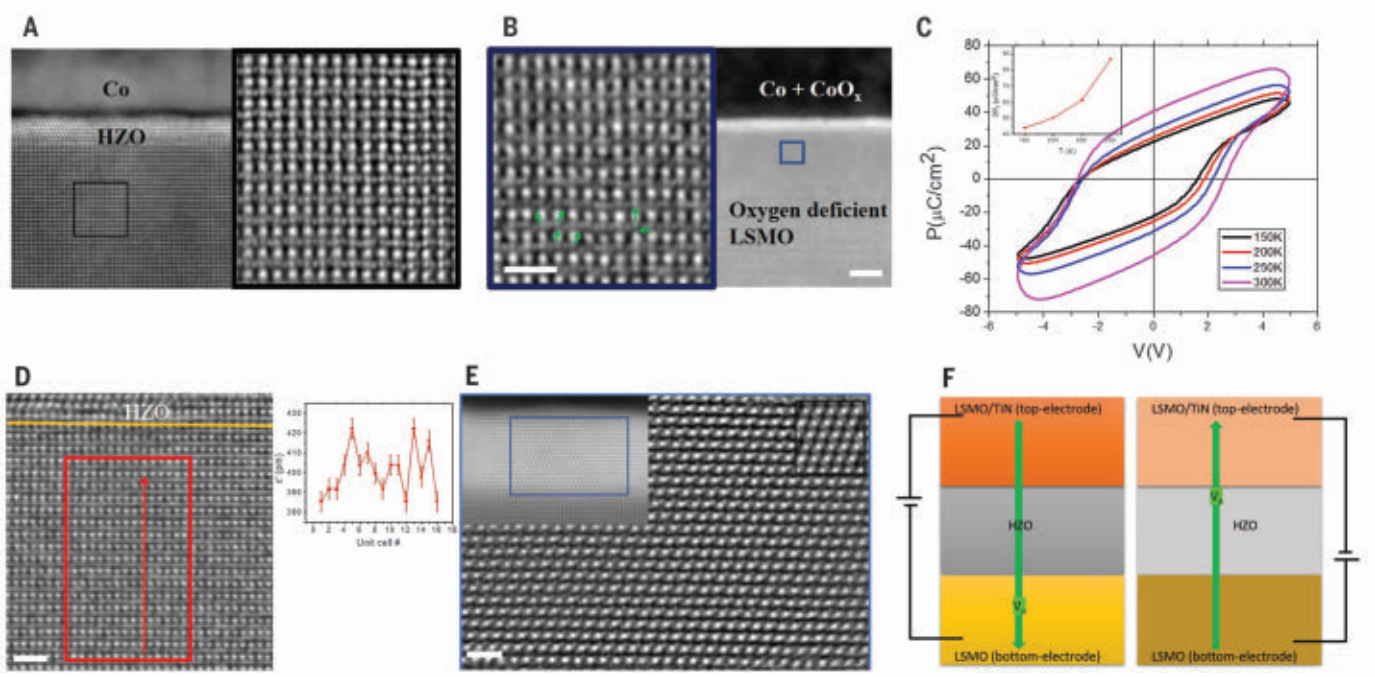


Fig. 4. Cycled tunnel junctions and ferroelectric capacitors. (**A**) Co/HZO (2 nm)/LSMO tunnel junction in LRS (500 μs, 100 s of cycles poled at −6 V). A HAADF-STEM image is shown on the left (scale bar, 5 nm), and an iDPC-STEM image of a selected region in the LSMO is shown on the right (scale bar, 1 nm). (**B**) Ex situ imaging of tunnel junction in HRS (500 μs, 100 s of cycles, poled at 6 V). (Right) HAADF-STEM (and corresponding information from EDS) shows CoOx/HZO/LSMO stack (scale bar, 5 nm). (Left) iDPC-STEM of a region in LSMO shows BM-precursor phase (scale bar, 1 nm). Selected Mn column displacements toward a square pyramidal geometry are marked by green arrows. (**C**) Temperature-dependent *P-V* loops obtained from dynamic hysteresis measurements at 1 kHz in LSMO/HZO (7 nm)/LSMO (30 nm) capacitors. Inset shows the corresponding $2P_r$ versus *T* (temperature) curve (see also fig. S8). (**D**) iDPC-STEM image of LSMO close to the HZO interface (marked in orange) in TiN/HZO (7 nm)/LSMO capacitors, prepared in a down-polarized state (at 5.5 V). BM-precursor phase is formed, confirmed by disorder in the *c*′ parameter (inset). (**E**) iDPC-STEM image of a down-polarized HZO layer, which is structurally still in r-phase. The left inset shows a HAADF-STEM image with the iDPC-STEM image presented for the region in the blue box, and the right inset shows a multislice simulation of the *R3m* phase. Scale bar, 1 nm in (D) and (E). (**F**) Schematic of the polarity-dependent oxygen voltammetry process in the metal-insulator-metal device structure. Oxygen-deficient phases in various layers are represented with lighter colors.

modest voltage range between –3.5 and 3.5 V for tens of thousands of cycles (fig. S10C). In LSMO/HZO/LSMO devices, the virgin state and cycled state do not show any substantial differences in the HZO lattice parameter, which indicates the persistence of the r-phase during cycling (fig. S10D). However, in the case of Au/HZO/LSMO, we show that an m-phase (with a Bragg peak at 2θ = 24.9°) appears after cycling and coexists with the r-phase (fig. S10, E and F). This reveals that the HZO layer itself is forced to act as source and sink of oxygen vacancies in the short term (*35*) when Au is used as a top electrode, which results in reversible structural phase transitions between the r- and m-phases. Further DC biasing results in an increase of m-phase fraction increasing positive bias, in agreement with the long-term in situ transmission electron microscopy (TEM) observations (fig. S11).

Using the model system of epitaxial HZO/LSMO/STO, we demonstrate that oxygen voltammetry and ferroelectric switching are intertwined (Fig. 4F) (*35*).With reactive top electrodes such as TiN (*45*), Co, and LSMO, the HZO layer acts as a mere conduit for reversible oxygen migration between the electrodes, with incipient topotactic transformations taking place in them. With a noble top electrode (Au), HZO instead itself acts as a sink and source of $\ddot{V}_o$. In the longer term, redox-based phase-transition effects that follow oxygen migration in both HZO and LSMO become more marked. Although we remain agnostic about the nature of coupling between voltammetry and polarization switching, recent predictions of electrochemical origin of ferroelectricity in hafnia-based compounds (*29*), as well as demonstrations of giant electrostriction in the sister compound Gd:CeO_2 (*46*), can provide important clues.

In the context of tunnel junction devices, perovskite manganites (*47*) are routinely used as back electrodes (*24*–*27*, *48*). Their crucial role as oxygen-conducting memristive layers (*41*) actively participating in charge transport is being recognized (*35*), with recent seminal demonstrations in ferroelectric tunnel junctions (*28*). Thus, our results have deep implications not just for hafnia-based ferroelectrics, but also for the wider field of oxide electronics.

REFERENCES AND NOTES

1. T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder, U. Böttger, *Appl. Phys. Lett.* **99**, 102903 (2011).
2. A. Chernikova *et al.*, *ACS Appl. Mater. Interfaces* **8**, 7232–7237 (2016).
3. X. Tian *et al.*, *Appl. Phys. Lett.* **112**, 102902 (2018).
4. S. S. Cheema *et al.*, *Nature* **580**, 478–482 (2020).
5. U. Schröder, C. S. Hwang, H. Funakubo, Eds., *Ferroelectricity in Doped Hafnium Oxide* (Woodhead Publishing, ed. 1, 2019).
6. P. Nukala *et al.*, *Ferroelectrics* **569**, 148–163 (2020).
7. Y. Wei *et al.*, *Nat. Mater.* **17**, 1095–1100 (2018).
8. R. Materlik, C. Kunneth, A. Kersch, *J. Appl. Phys.* **117**, 134109 (2015).
9. S. J. Kim, J. Mohan, S. R. Summerfelt, J. Kim, *JOM* **71**, 246–255 (2019).
10. T. Mimura, T. Shimizu, H. Funakubo, *Appl. Phys. Lett.* **115**, 032901 (2019).
11. X. Xu *et al.*, *Nat. Mater.* 10.1038/s41563-020-00897-x (2021).
12. Y. Zhou *et al.*, *Comput. Mater. Sci.* **167**, 143–150 (2019).
13. H. Mulaosmanovic *et al.*, *ACS Appl. Mater. Interfaces* **9**, 3792–3798 (2017).
14. M. Pešić *et al.*, *J. Comput. Electron.* **16**, 1236–1256 (2017).
15. N. Gong *et al.*, *Appl. Phys. Lett.* **112**, 262903 (2018).
16. P. Buragohain *et al.*, *Appl. Phys. Lett.* **112**, 222901 (2018).

17. S. Migita *et al.*, *Jpn. J. Appl. Phys.* **57**, 04FB01 (2018).
18. Y. H. Lee *et al.*, *Adv. Electron. Mater.* **5**, 180436 (2019).
19. T. Schenk *et al.*, *ACS Appl. Mater. Interfaces* **7**, 20224–20233 (2015).
20. H.-J. Lee *et al.*, *Science* **369**, 1343–1347 (2020).
21. C. Li *et al.*, *Nano Res.* **8**, 3571–3579 (2015).
22. S. U. Sharath *et al.*, *Adv. Funct. Mater.* **27**, 1700432 (2017).
23. Y. Wei *et al.*, *npj Quantum Mater.* **4**, 62 (2019).
24. M. C. Sulzbach *et al.*, *Adv. Electron. Mater.* **6**, 1900852 (2020).
25. M. Hambe *et al.*, *Adv. Funct. Mater.* **20**, 2436–2441 (2010).
26. D. Pantel, S. Goetze, D. Hesse, M. Alexe, *ACS Nano* **5**, 6032–6038 (2011).
27. S. Boyn *et al.*, *Nat. Commun.* **8**, 14736 (2017).
28. V. Rouco *et al.*, *Nat. Commun.* **11**, 658 (2020).
29. V. Mikheev *et al.*, *ACS Appl. Mater. Interfaces* **11**, 32108–32114 (2019).
30. M. D. Glinchuk *et al.*, *J. Alloys Compd.* **830**, 153628 (2020).
31. M. Qian, I. Fina, M. C. Sulzbach, F. Sánchez, J. Fontcuberta, *Adv. Electron. Mater.* **5**, 1800646 (2019).
32. C. Ferreyra *et al.*, *Phys. Rev. Appl.* **14**, 044045 (2020).
33. M. Pešić *et al.*, *Adv. Funct. Mater.* **26**, 4601–4612 (2016).
34. P. Gao *et al.*, *Nat. Commun.* **2**, 591 (2011).
35. Materials and methods are available as supplementary materials online.
36. S. de Graaf, J. Momand, C. Mitterbauer, S. Lazar, B. J. Kooi, *Sci. Adv.* **6**, eaay4312 (2020).
37. G. Sanchez-Santolino *et al.*, *Microsc. Microanal.* **20**, 825–831 (2014).
38. S. Stølen, E. Bakken, C. E. Mohn, *Phys. Chem. Chem. Phys.* **8**, 429–447 (2006).
39. P. Nukala *et al.*, *Appl. Phys. Lett.* **118**, 062901 (2021).
40. L. Cao *et al.*, *Adv. Mater.* **31**, 1806183 (2019).
41. L. Yao, S. Inkinen, S. van Dijken, *Nat. Commun.* **8**, 14544 (2017).
42. C. Baeumer *et al.*, *Nat. Commun.* **6**, 8610 (2015).
43. L. Bégon-Lours *et al.*, *Phys. Rev. Mater.* **4**, 043401 (2020).
44. Y. Wei *et al.*, *Phys. Rev. Appl.* **12**, 031001 (2019).
45. W. Hamouda *et al.*, *Appl. Phys. Lett.* **116**, 252903 (2020).
46. S. Santucci, H. Zhang, S. Sanna, N. Pyrds, V. Esposito, *J. Mater. Chem. A* **8**, 14023–14030 (2020).
47. A. Herpers *et al.*, *Adv. Mater.* **26**, 2730–2735 (2014).
48. G. Radaelli *et al.*, *Adv. Mater.* **27**, 2602–2607 (2015).
49. P. Nukala *et al.*, Dataset: Reversible oxygen migration and phase transitions in hafnia-based ferroelectric devices, version 1.0, Dataverse NL (2021); https://doi.org/10.34894/EIOXPT.

ACKNOWLEDGMENTS

Funding: P.N. acknowledges funding from European Union's Horizon 2020 research and innovation program under Marie Sklodowska-Curie grant agreement no. 794954 (nickname: FERHAZ). M.A., Y.W., B.K., and B.N. acknowledge the financial support of the CogniGron research center and the Ubbo Emmius Funds (University of Groningen). Y.W. and S.M. acknowledge a China Scholarship Council grant and a Van Gogh travel grant. We acknowledge MAX IV Laboratory for time on Beamline NanoMAX under proposal 20190954. Research conducted at MAX IV, a Swedish national user facility, is supported by the Swedish Research council under contract 2018-07152, the Swedish Governmental Agency for Innovation Systems under contract 2018-04969, and Formas under contract 2019-02496. S.M. acknowledges the French RENATECH network that supported the cleanroom activity. P.N. acknowledges all the discussions with P. Buragohain from the University of Nebraska. **Author contributions:** P.N., B.N., M.A., and D.C. conceived the idea. P.N. synthesized the samples through PLD. M.A. and P.N. devised the in situ biasing device preparation protocol using FIB and made the devices. P.N., M.A., and S.d.G. set up the electrical biasing system compatible with Themis-Z microscope. P.N., M.A., and S.d.G. carried out the in situ biasing experiments with timely help from H.W.Z. and B.K.. Y.W. and S.M. synthesized and fabricated tunnel junction and capacitor devices and did the ex situ electrical testing. P.N. and M.A. prepared the lamellae and carried out imaging experiments. P.N., M.A., Y.W., T.C., and S.d.G. analyzed the TEM and transport data. D.C., A.B., and D.M. carried out the experiments on the beamline. E.S., P.N., D.C., and A.B. carried out synchrotron data analysis. All the authors discussed the data. P.N. and B.N. cowrote the manuscript, which was read, edited, and approved by all the authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data are available from the University of Groningen data repository (*49*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/630/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
References (*50*–*53*)

22 October 2020; accepted 2 April 2021
Published online 15 April 2021
10.1126/science.abf3789

CORONAVIRUS

The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia

Martin Pavelka[1,2,3]*, Kevin Van-Zandvoort[4,5], Sam Abbott[4,5], Katharine Sherratt[4,5], Marek Majdan[6], CMMID COVID-19 working group†, Inštitút Zdravotných Analýz†, Pavol Jarčuška[7], Marek Krajčí[1], Stefan Flasche[4,5]‡, Sebastian Funk[4,5]‡

Slovakia conducted multiple rounds of population-wide rapid antigen testing for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in late 2020, combined with a period of additional contact restrictions. Observed prevalence decreased by 58% (95% confidence interval: 57 to 58%) within 1 week in the 45 counties that were subject to two rounds of mass testing, an estimate that remained robust when adjusting for multiple potential confounders. Adjusting for epidemic growth of 4.4% (1.1 to 6.9%) per day preceding the mass testing campaign, the estimated decrease in prevalence compared with a scenario of unmitigated growth was 70% (67 to 73%). Modeling indicated that this decrease could not be explained solely by infection control measures but required the addition of the isolation and quarantine of household members of those testing positive.

Nonpharmaceutical interventions have been extensively used worldwide to limit the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (*1*). These have included travel restrictions, mandating of face masks, closure of schools and nonessential businesses, and nationwide stay-at-home orders. All the measures were aimed at mitigating ill-health due to COVID-19 (*2*, *3*); however, they also place an unprecedented economic and social burden on the majority of uninfected people (*4*, *5*). Testing of reported symptomatic cases and tracing their contacts aims to provide a more targeted measure but, in many settings, has proven insufficient for containing transmission (*6*).

Mass testing campaigns are an alternative way to identify infectious individuals and allow the targeting of interventions without much added burden to those not infectious. However, the polymerase chain reaction (PCR) for the diagnosis of a SARS-CoV-2 infection is not suitable for mass use. Although laboratory capacities have been upscaled in record time, PCR testing remains expensive and often has turnaround times of more than 1 day, diminishing its utility (*7*). The PCR detection window also typically extends to the postinfectious period by detecting RNA fragments, hence identifying as infected those who are no longer infectious (*8*).

By contrast, rapid antigen tests are cheap and can be quickly produced in large quantities, offering results on site in 15 to 30 min without the need for a laboratory. They are less sensitive in detecting infections with low viral load that are less likely to transmit, but can detect over 70% of likely infectious cases. A recent observational study estimated the sensitivity of lateral flow devices in detecting infectious individuals to be as high as 83 to 91% (*9*). This makes mass testing a viable part of the portfolio of nonpharmaceutical interventions (*10*, *11*).

In October and November 2020, Slovakia used rapid antigen tests in a campaign that targeted the whole population to identify

[1]Slovak Ministry of Health, Bratislava, Slovakia. [2]Inštitút Zdravotných Analýz (Institute of Health Analyses), Bratislava, Slovakia. [3]Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK. [4]Department for Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. [5]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, UK. [6]Institute for Global Health and Epidemiology, Faculty of Health Sciences and Social Work, Trnava University, Trnava, Slovakia. [7]Faculty of Medicine, Pavol Jozef Šafárik University, Košice, Slovakia.
***Corresponding author. Email: martin.pavelka@health.gov.sk**
†CMMID COVID-19 working group and Inštitút Zdravotných Analýz authors and their affiliations are listed in the supplementary materials.
‡These authors contributed equally to this work.

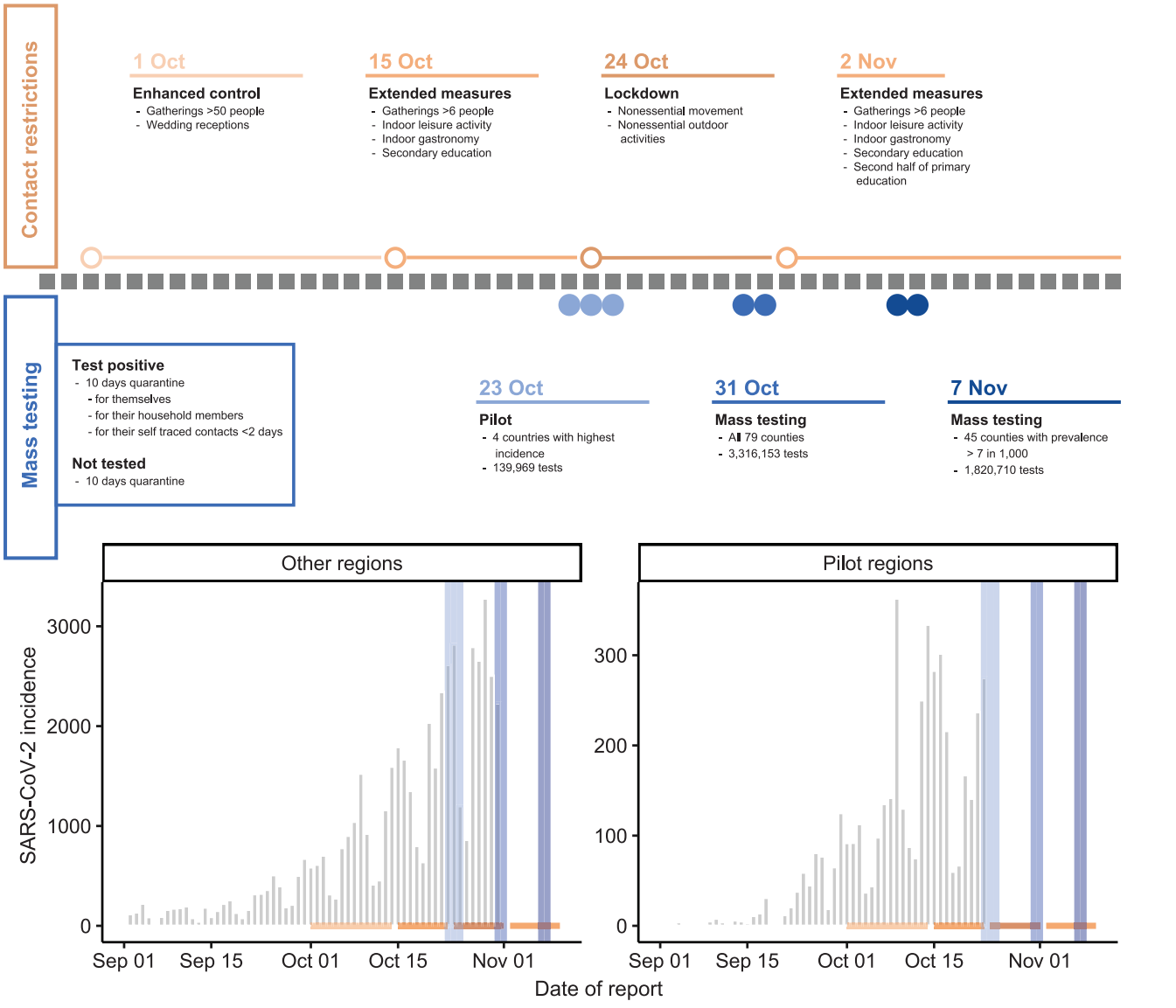


Fig. 1. Overview of interventions and pre–mass testing epidemiology. (**Top**) Description of timing and extent of national contact restriction in Slovakia (color intensity indicates intensity of the measures) and timing and extent of the mass testing campaigns. Open circles and lines in respective colors indicate the start and duration of the contact restrictions, and the blue solid circles indicate the days on which mass testing was conducted, although the highest turnout was usually on the first day. (**Left**) Box illustrating contact-reducing measures for those testing positive and those who chose not to be tested. (**Bottom**) SARS-CoV-2 infection incidence as reported by the Slovak Ministry of Health and collected through passive symptom–triggered PCR testing. Using the same color coding as at the top, contact interventions are indicated by horizontal lines, and mass testing campaigns are indicated by vertical lines. Data from the passive surveillance subsequent to the respective first mass testing campaign are omitted to clearly illustrate the trends in infection rates that led up to the mass testing and because mass testing is likely to have changed the sensitivity of the passive surveillance, thereby distorting the observation of infection trends that followed mass testing.

infectious cases at scale, rapidly reduce transmission, and thus allow easing of lockdown measures (*12*). A pilot took place between 23 and 25 October in the four most affected counties, followed by a round of national mass testing on 31 October and 1 November (round 1). High prevalence counties were again targeted with a subsequent round of testing on 7 and 8 November (round 2) (Fig. 1).

In total, 5,276,832 SD-Biosensor Standard Q rapid antigen tests were conducted by trained medical personnel during the mass testing campaigns, with 65% of the respective populations tested in the pilot, 66% in mass testing round 1 and 62% in round 2. This corresponded to 87, 83, and 84% of the age-eligible population (10 to 65 years and older adults in employment) in each round, respectively. It does not include residents who were quarantining at the time of the campaign or the 534,300 tests that were conducted on medical, military, and governmental personnel who were not included in geographical county data.

A total of 50,466 participants tested positive, indicating the presence of currently infectious SARS-CoV-2. The proportion of positive tests was 3.91% (range across counties: 3.12 to 4.84%) in the pilot, 1.01% (range: 0.13 to 3.22%) in

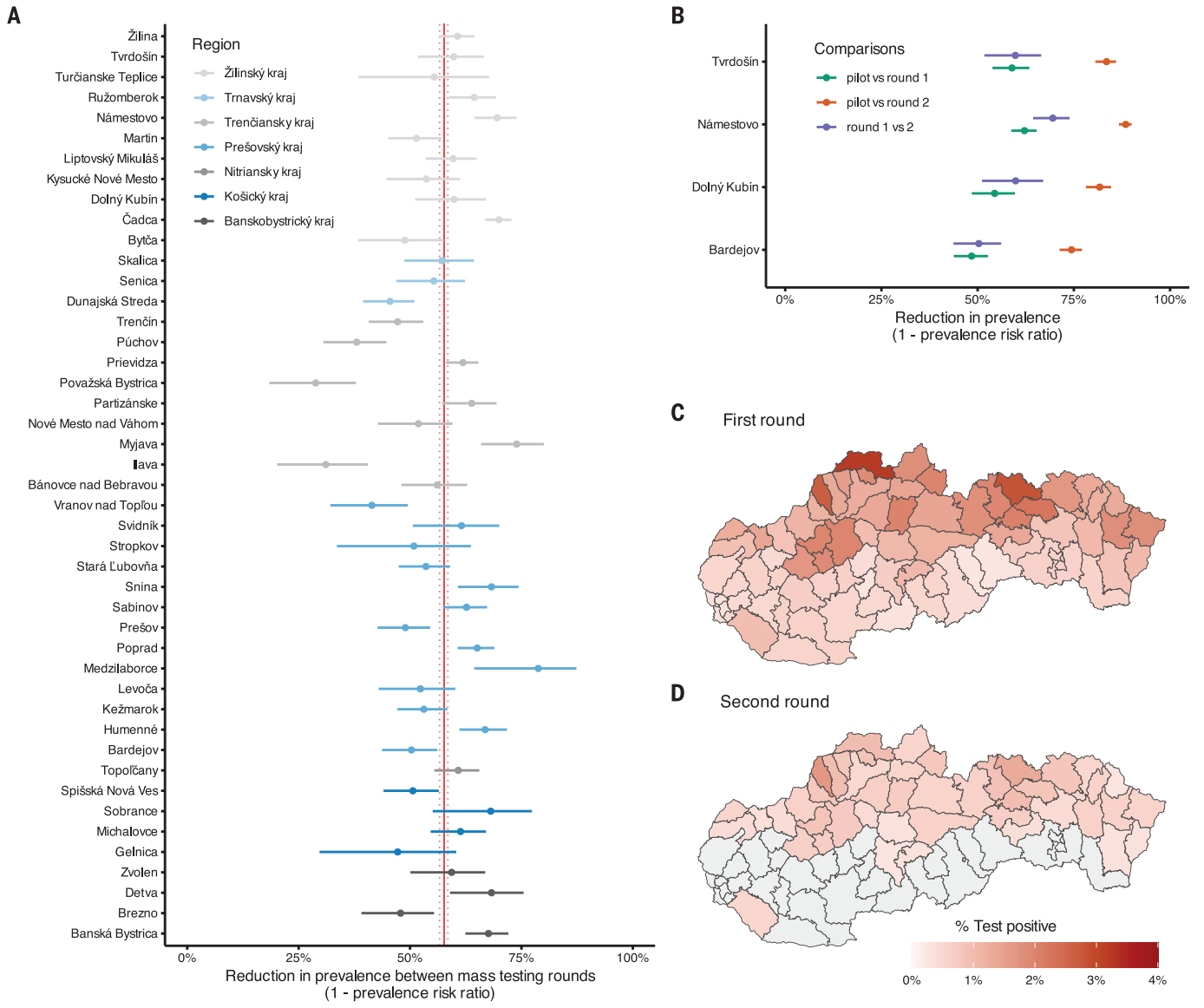


Fig. 2. The change in test positivity between mass testing campaigns. (**A**) Change in test positivity [1 − crude prevalence ratio (cPR)] observed from mass testing round 1 to round 2 in the 45 counties that were eligible for both rounds of mass testing. Counties are grouped and color coded into regions. The crude pooled estimate and its 95% confidence bounds are shown as red vertical lines. The confidence intervals were estimated using a normal approximation (Wald interval). (**B**) Change in test positivity (1 − cPR) observed from the pilot mass testing round to either the first (green) or the second (orange) national round and from the first to the second mass testing round (blue) in the four counties that were included in the pilot. The confidence intervals were estimated using a normal approximation (Wald interval). (**C** and **D**) County-level test positivity in the (C) first and (D) second round of mass testing. Gray areas indicate counties that were not part of the second round because their test positivity rate was less than 7 per 1000 and hence have no estimates.

round 1, and 0.62% (range: 0.28 to 1.65%) in round 2 (Fig. 2, C and D).

The potential for large numbers of false-positive tests has been a point of criticism for mass testing campaigns. Although multiple studies have found high specificity for the Biosensor test kit, they are not sufficiently powered to exclude specificity levels that at a population level would yield an overwhelming amount of false positives (*13*). From the low test-positive rates in some counties, we estimate with 95% certainty that the specificity of the SD Biosensor Standard Q antigen test exceeded 99.85%, and the occurrence false positives was therefore not of major concern in this study.

The counties with the highest prevalence were found in the Northern part of the country, whereas the two main Slovakian cities of Bratislava and Košice had some of the lowest observed prevalences (Fig. 1C). Reflecting this pattern, we found that high county-level prevalence was associated with a younger average population age and a lower population density (fig. S8). Given that prevalence varied at a much smaller than county scale (*14*), such associations may be clearer at the individual or community level, as observed in other countries.

In the four counties where the pilot was conducted, observed infection prevalence decreased by 56% [95% confidence interval (CI): 54 to 58%] between the pilot and round 1 of the mass testing campaign and a further 60% (95% CI: 56 to 63%) between rounds 1 and 2, totaling a decrease of 82% (95% CI: 81 to 83%) over 2 weeks. There was little heterogeneity between counties (Fig. 2B).

Among the 45 counties that were included in round 2 of the mass testing campaign, observed infection prevalence decreased by 58% (95% CI: 57 to 58%) in 1 week. Combining the pilot results with the ones from the two rounds of testing in 45 counties, each round of mass testing was estimated to have reduced observed infection prevalence by 56% (95% CI: 52 to 59%) when adjusted for attendance rates, reproduction number, and prevalence in previous rounds. The estimated reduction between rounds varied considerably by county, from

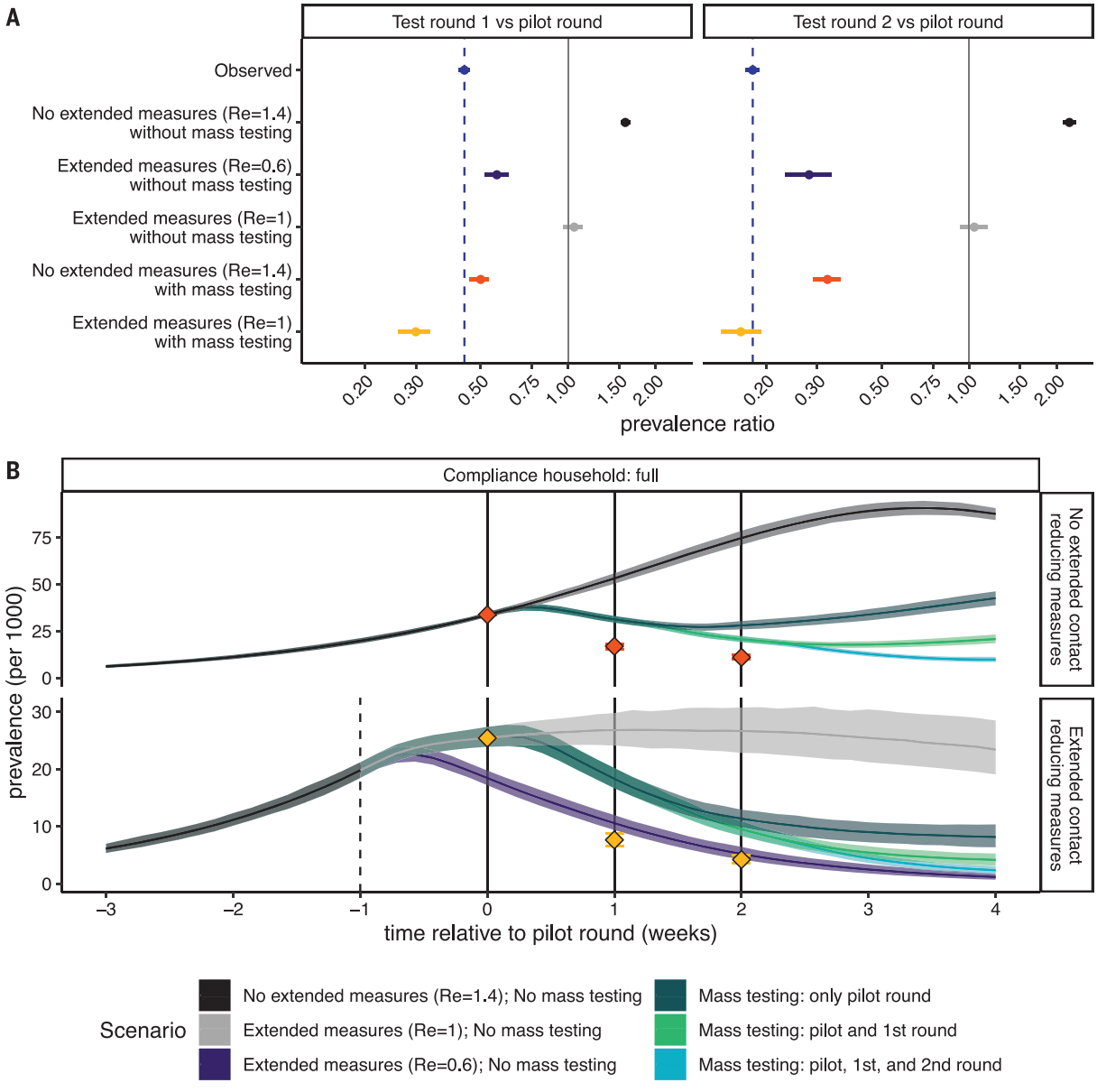


Fig. 3. Simulated relative effectiveness of the extended contact-reducing measures and the mass testing. (**A**) The change in observed prevalence of infectious nonquarantining individuals between 10 and 65 years of age as predicted by the microsimulation model. For comparison, the observed test-positivity rate is shown in blue. The facets show changes (left) from the pilot to the first round of mass testing and (right) from the pilot to the second round of mass testing. Shown scenarios compare the effect of (top to bottom) no additional interventions that limit the growth rate of reproduction number (Re) = 1.4, the extended contact-reduction measures drastically reducing the growth rate to Re = 0.6 and no mass testing being conducted, the extended contact-reduction measures reducing the growth rate to Re = 1.0 and no mass testing being conducted, no change in growth rate but mass testing, and the extended contact reduction measures reducing the growth rate to Re = 1 and mass testing. In scenarios without mass testing, we compared prevalence of infectious individuals on the same days as testing occurred in scenarios with mass testing. CIs around the modeled values in each scenario are calculated as the 2.5 and 97.5% percentiles across 500 model iterations, with the point estimate representing the median. The CI around the observed value is its binomial CI. (**B**) Simulated infection incidence of alternative intervention strategies. Simulations are aligned by the date of the first mass test [time (t) = 0]. The dashed line indicates the timing of the extended contact-reducing measures, and the solid lines indicate the timing of the mass testing campaigns. Colors indicate the simulations stratified into whether no mass testing or one, two, or three testing rounds were performed and the effectiveness of the extended contact-reduction measures on the growth rate. Red and yellow diamonds indicate the prevalence of infectiousness observed among the tested nonquarantining age-eligible population, corresponding to the scenarios in (A).

Table 1. Overview of county-specific test numbers and reductions for the 79 counties in Slovakia. R, median estimate of the reproduction number on 22 October, based on test-positive cases from syndromic surveillance up to 30 October and estimated by using a renewal process model on back-calculated estimates of infection incidence; %, proportion positive out of those attending mass testing.

				Pilot			Round 1			Round 2		
County	**Region**	**Population**	**R**	**Attendance**	**Positive**	**%**	**Attendance**	**Positive**	**%**	**Attendance**	**Positive**	**%**
Bánovce nad Bebravou	Trenčiansky kraj	36281.5	1.4				23264	457	1.96	22248	192	0.86
Banská Bystrica	Banskobystrický kraj	110828.5	1.2				64127	687	1.07	66544	231	0.35
Banská Štiavnica	Banskobystrický kraj	16086.0	0.7				11725	33	0.28			
Bardejov	Prešovský kraj	77771.0	0.7	48320	1569	3.25	44197	740	1.67	43983	366	0.83
Bratislava I	Bratislavský kraj	44798.0	1.2				29047	108	0.37			
Bratislava II	Bratislavský kraj	108139.0	1.2				80958	345	0.43			
Bratislava III	Bratislavský kraj	61418.0	1.2				49788	175	0.35			
Bratislava IV	Bratislavský kraj	93058.0	1.2				63857	81	0.13			
Bratislava V	Bratislavský kraj	141259.0	1.2				68139	268	0.39			
Brezno	Banskobystrický kraj	61449.5	1.4				37339	450	1.21	38515	242	0.63
Bytča	Žilinský kraj	30917.0	1.6				21419	328	1.53	20931	164	0.78
Čadca	Žilinský kraj	90080.0	1.0				53907	1736	3.22	52304	506	0.97
Detva	Banskobystrický kraj	32051.0	1.3				19704	211	1.07	23255	79	0.34
Dolný Kubín	Žilinský kraj	39456.5	1.0	29347	916	3.12	24251	345	1.42	24170	138	0.57
Dunajská Streda	Trnavský kraj	122358.0	1.3				87329	840	0.96	110083	577	0.52
Galanta	Trnavský kraj	94076.0	1.3				71243	349	0.49			
Gelnica	Košický kraj	31868.0	1.3				18331	131	0.71	19087	72	0.38
Hlohovec	Trnavský kraj	45012.5	1.4				28892	171	0.59			
Humenné	Prešovský kraj	61985.5	1.1				32962	598	1.81	32750	197	0.60
Ilava	Trenčiansky kraj	59187.5	1.4				37604	442	1.18	35931	291	0.81
Kežmarok	Prešovský kraj	75235.0	1.4				43959	845	1.92	43252	390	0.90
Komárno	Nitriansky kraj	101711.5	1.5				61268	343	0.56			
Košice - okolie	Košický kraj	129543.5	1.2				32849	196	0.60			
Košice I	Košický kraj	67513.0	1.2				39314	295	0.75			
Košice II	Košický kraj	82287.5	1.2				11109	41	0.37			
Košice III	Košický kraj	28748.5	1.2				26992	135	0.50			
Košice IV	Košický kraj	60126.0	1.2				80426	487	0.61			
Krupina	Banskobystrický kraj	22182.0	1.4				13388	66	0.49			
Kysucké Nové Mesto	Žilinský kraj	32914.0	1.6				20605	384	1.86	20491	177	0.86
Levice	Nitriansky kraj	110824.0	1.4				70155	375	0.53			
Levoča	Prešovský kraj	33702.0	1.0				18344	373	2.03	17747	172	0.97
Liptovský Mikuláš	Žilinský kraj	72260.5	1.2				47172	667	1.41	46827	267	0.57
Lučenec	Banskobystrický kraj	73466.0	1.0				40655	213	0.52			
Malacky	Bratislavský kraj	74323.0	1.3				54657	285	0.52			
Martin	Žilinský kraj	96338.0	1.5				56533	771	1.36	57513	381	0.66
Medzilaborce	Prešovský kraj	11841.5	1.1				6980	91	1.30	6142	17	0.28
Michalovce	Košický kraj	110705.0	1.0				58929	512	0.87	62790	211	0.34
Myjava	Trenčiansky kraj	26356.0	0.9				17753	249	1.40	18599	68	0.37
Námestovo	Žilinský kraj	62663.5	0.9	40052	1910	4.77	37029	668	1.80	37659	207	0.55
Nitra	Nitriansky kraj	161560.0	1.3				99175	674	0.68			
Nové Mesto nad Váhom	Trenčiansky kraj	62553.5	1.5				40829	363	0.89	46269	198	0.43
Nové Zámky	Nitriansky kraj	139004.5	1.3				79234	478	0.60			
Partizánske	Trenčiansky kraj	45596.5	1.5				26492	494	1.86	27585	186	0.67
Pezinok	Bratislavský kraj	65145.0	1.3				45801	240	0.52			
Piešťany	Trnavský kraj	62802.5	1.3				40122	183	0.46			
Poltár	Banskobystrický kraj	21471.0	2.0				12455	71	0.57			
Poprad	Prešovský kraj	104913.5	1.4				59072	1059	1.79	58098	364	0.63
Považská Bystrica	Trenčiansky kraj	62438.5	1.4				37822	505	1.34	36092	343	0.95
Prešov	Prešovský kraj	175609.5	1.0				84781	724	0.85	108271	472	0.44
Prievidza	Trenčiansky kraj	133979.5	1.3				76457	1497	1.96	77170	576	0.75
Púchov	Trenčiansky kraj	44309.5	1.3				29455	782	2.65	28017	461	1.65
Revúca	Banskobystrický kraj	39636.5	1.7				21419	58	0.27			
Rimavská Sobota	Banskobystrický kraj	84159.0	1.7				46872	197	0.42			
Rožňava	Košický kraj	62208.5	1.2				34307	100	0.29			

continued on next page

				Pilot			Round 1			Round 2		
County	Region	Population	R	Attendance	Positive	%	Attendance	Positive	%	Attendance	Positive	%
Ružomberok	Žilinský kraj	56702.0	1.6				34000	682	2.01	33056	236	0.71
Sabinov	Prešovský kraj	60518.5	1.4				35366	804	2.27	34757	295	0.85
Šaľa	Nitriansky kraj	51685.0	1.2				31993	199	0.62			
Senec	Bratislavský kraj	89832.0	1.4				66052	314	0.48			
Senica	Trnavský kraj	60446.0	1.2				40675	384	0.94	46000	194	0.42
Skalica	Trnavský kraj	47104.5	1.2				29223	368	1.26	31200	168	0.54
Snina	Prešovský kraj	36240.5	1.3				19122	345	1.80	19396	111	0.57
Sobrance	Košický kraj	22819.0	0.9				12986	135	1.04	12966	43	0.33
Spišská Nová Ves	Košický kraj	99765.0	1.3				54279	739	1.36	53712	361	0.67
Stará Ľubovňa	Prešovský kraj	53953.5	1.2				28749	805	2.80	27234	354	1.30
Stropkov	Prešovský kraj	20532.0	1.1				10494	125	1.19	10764	63	0.59
Svidník	Prešovský kraj	32564.0	1.1				16631	220	1.32	16705	85	0.51
Topoľčany	Nitriansky kraj	70131.5	1.4				44627	748	1.68	50253	330	0.66
Trebišov	Košický kraj	105353.0	0.9				68503	400	0.58			
Trenčín	Trenčiansky kraj	114523.0	1.2				73424	832	1.13	72546	434	0.60
Trnava	Trnavský kraj	132454.5	1.2				92215	557	0.60			
Turčianske Teplice	Žilinský kraj	15884.0	1.7				11287	112	0.99	12210	54	0.44
Tvrdošín	Žilinský kraj	36180.0	1.3	22250	1078	4.84	18541	369	1.99	20502	164	0.80
Veľký Krtíš	Banskobystrický kraj	43473.0	1.2				24652	76	0.31			
Vranov nad Topľou	Prešovský kraj	80766.5	1.4				43552	460	1.06	45424	281	0.62
Žarnovica	Banskobystrický kraj	26152.5	1.4				16272	105	0.65			
Žiar nad Hronom	Banskobystrický kraj	46861.5	0.8				26260	108	0.41			
Žilina	Žilinský kraj	158043.0	1.5				111155	1392	1.25	103898	512	0.49
Zlaté Moravce	Nitriansky kraj	40572.5	0.9				26180	156	0.60			
Zvolen	Banskobystrický kraj	68758.5	1.4				39422	276	0.70	47764	136	0.28

29% in county Považská Bystrica to 79% in county Medzilaborce, but although heterogeneous showed no regional pattern (Fig. 2A). Neither region, attendance rates, prevalence in round 1, nor the estimated growth rate before mass testing showed any significant impact on the observed county-specific reductions.

At the time of round 1 of the mass testing campaign, the incidence of confirmed cases reported through the syndromic surveillance system was rising in nonpilot counties, with an estimated infection growth rate of 4.4% (1.1% to 6.9%) per day. When adjusting for this growth trend, we estimated a self-adjusted prevalence ratio (saPR) of 0.30 (0.27 to 0.33). In the pilot counties, reported infection incidence showed signs of leveling in the week before the mass testing campaign, with an estimated infection growth rate of 1.3% (–7.4 to 7.8%), yielding a respective saPR of 0.31 (0.26 to 0.33).

Because we used the test positivity rate of the subsequent round to estimate the impact of the previous one, we were unable to observe the impact of the last round in each county and hence the full effect of the campaign.

However, we found that the reduction achieved per round of testing was 56% (52 to 59%), indicating that the 41 counties with two rounds of testing likely reduced infection prevalence by 81% (77 to 83%) within 2 weeks and that the four counties included into the pilot testing reduced infection prevalence by 91% (89 to 93%) within 3 weeks.

The observational nature of this study made it difficult to separate the effects of the mass testing campaigns from that of the other nonpharmaceutical interventions introduced over the same period that aimed to reduce contacts and mobility, although much less than during the spring lockdown (fig. S4). Nevertheless, a greater than 50% decline in infection prevalence within 1 week (or 80% in 2 weeks) is noteworthy, particularly while primary schools and workplaces were mostly open. For comparison, a month-long lockdown in November in the UK resulted in just a 30% decrease in prevalence (*15*). This, alongside the inability in December to control the rebounding spread of SARS-CoV-2 in Slovakia through even more stringent contact restrictions, indicates that the mass testing campaigns were responsible for a large share of case reduction in the previous months.

To further investigate the relationship between the reduction in prevalence, mass testing, and nonpharmaceutical interventions, we used a microsimulation model for fine-scale SARS-CoV-2 transmission in a representative county included in the pilot phase of the mass testing. Among the multiple intervention scenarios tested, only the scenario that assumed a substantial impact of both the additional contact reducing measures and the mass testing campaigns was able to generate reductions in test positivity rates between testing rounds that were similar to those observed (Fig. 3). Thus, the requirement for quarantine for the whole household after a positive test was essential for the combined effect of mass testing and contact reduction measures. The model predicted a prevalence ratio between the first two testing rounds of 0.30 (0.26 to 0.34) with household quarantine and 0.78 (0.72 to 0.84) without household quarantine.

Despite a reduction of more than 50% in test positivity between mass testing campaigns, standard syndromic surveillance did not report a rapid collapse in test-positive cases corresponding to drastic reductions in prevalence. This may be explained by a variety of reasons. Foremost, the national mass testing campaigns are likely to have a major disruptive effect on routine passive syndromic surveillance. Also, the ability of PCR to detect viral RNA well beyond the infectious period will partially

mask a sudden drop in infectious cases. In addition, starting mid-September incidence surveillance has been operating at capacity with long waiting lists for testing and stricter eligibility criteria, which reduced substantially in the period after mass testing and hence may have artificially reduced the observable change in these data. By contrast, data on hospital bed occupancy shows a sudden flattening from mid-November, indicating a sharp decrease in new admissions that is consistent with a sizable reduction in new infections when the mass testing campaigns occurred (fig. S6).

Executing a large-scale mass testing campaign comes with several challenges. The need to mobilize sufficient medical personnel to conduct the nasopharyngeal swabs proved to be a major obstacle. Also, the logistics of mobilizing large numbers of assisting army personnel and vast amounts of testing and personal protective equipment (PPE) material proved challenging. Some of the challenges could be overcome by using other rapid antigen tests with similarly high sensitivity but that are also licensed for use with nasal swabs (*16*, *17*). Nasal swabs can be self-administered and reduce demand on trained personnel and transmission risk in the process of sample collection and can even enable testing at home. Self-administered swabs are also less intrusive and can be better suited for children and mass testing at schools. However, these benefits must be weighed against the potential loss of sensitivity if self-administered swabs are not conducted appropriately (*18*). The details of the Slovak mass testing experience need to be studied carefully before considering potential replication elsewhere (*19*).

The combination of nationwide restrictions and mass testing with quarantining of household contacts of test positives rapidly reduced the prevalence of infectious residents in Slovakia. Although it was impossible to disentangle the precise contribution of control measures and mass testing, the latter is likely to have had a substantial effect in curbing the pandemic in Slovakia and may provide a valuable tool in future containment of SARS-CoV-2 elsewhere.

REFERENCES AND NOTES

1. Guidelines for the implementation of non-pharmaceutical interventions against COVID-19. *Eur. Cent. Dis. Prev. Control* (2020); www.ecdc.europa.eu/en/publications-data/covid-19-guidelines-non-pharmaceutical-interventions.
2. N. G. Davies *et al*., *Lancet Public Health* 10.1016/S2468-2667(20)30133-X (2020).
3. S. Flaxman *et al*., *Nature* **584**, 257–261 (2020).
4. Everyone Included, Social Impact of COVID-19–Dros. Inf. Serv. D; www.un.org/development/desa/dspd/everyone-included-covid-19.html.
5. Socio-economic impact of COVID-19. UNDP; www.undp.org/content/undp/en/home/coronavirus/socio-economic-impact-of-covid-19.html.
6. A. J. Kucharski *et al*., *Lancet Infect. Dis.* **20**, 1151–1160 (2020).
7. M. E. Kretzschmar *et al*., *Lancet Public Health* **5**, e452–e459 (2020).
8. R. Wölfel *et al*., *Nature* **581**, 465–469 (2020).
9. Oxford University, An observational study of SARS-CoV-2 infectivity by viral load and demographic factors and the utility lateral flow devices to prevent transmission (Univ. of Oxford, 2021); www.ox.ac.uk/news/2021-01-21-lateral-flow-devices-detect-most-infectious-covid-19-cases-and-could-allow-safer.
10. M. J. Mina, R. Parker, D. B. Larremore, *N. Engl. J. Med.* **383**, e120 (2020).
11. D. B. Larremore *et al*., *Sci. Adv.* **7**, eabd5393 (2021).
12. Som Zodpovedný; www.somzodpovedny.sk.
13. Z. Iglói *et al.*, medRxiv 20234104 [Preprint] 20 November 2020.
14. K. Boďová, R. Kollár, medRxiv 20248808 [Preprint] 26 December 2020.
15. S. Riley *et al.*, medRxiv 20239806 [Preprint] 2 December 2020.
16. Panbio COVID-19 Ag Rapid Test Device; www.globalpointofcare.abbott/en/product-details/panbio-covid-19-ag-antigen-test.html.
17. Sofia SARS Antigen FIA, Quidel; www.quidel.com/immunoassays/rapid-sars-tests/sofia-sars-antigen-fia.
18. E. Mahase, *BMJ* **371**, m4469 (2020).
19. K. Landokova, How Slovakia tested 3.6 million people for COVID-19 in a single weekend; www.bi.team/blogs/how-slovakia-tested-3-6-million-people-for-covid-19-in-a-single-weekend.
20. Koronavírus a Slovensko; https://korona.gov.sk.
21. R Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2019; www.R-project.org).
22. M. Pavelka *et al*., Data and R code accompanying the manuscript "The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia". Zenodo (2021); doi:10.5281/zenodo.4570939.

ACKNOWLEDGMENTS

We thank all health care workers, Slovak armed forces, and countless volunteers who helped with the execution of the mass testing campaign. Further, we thank all participants who contributed their time to help curb the pandemic and particularly those who had to quarantine as a result of their or their household's or contact's test result. **Funding:** This work was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society 208812/Z/17/Z (to S.Fl.); Wellcome Trust Senior Research Fellowship (to S.Fu., S.A., and K.S.); and Elrha's Research for Health in Humanitarian Crises (R2HC) Programme funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR) (K.V.-Z.). The following funding sources are acknowledged as providing funding for the working group authors: BBSRC LIDP BB/M009513/1 (to D.S.); Bill & Melinda Gates Foundation INV-001754 (to M.Q.); INV-003174 (to K.P., M.J., and Y.L.); OPP1157270 (to K.A.); NTD Modelling Consortium OPP1184344 (to C.A.B.P. and G.F.M.); OPP1180644 (to S.R.P.); OPP1183986 (to E.S.N.); DFID/Wellcome Trust Epidemic Preparedness Coronavirus research program 221303/Z/20/Z: (to C.A.B.P.); EDCTP2 RIA2020EF-2983-CSIGN (to H.P.G.); ERC #757699 (to M.Q.); The European Union's Horizon 2020 research and innovation program—project EpiPose 101003688 (to K.P., M.J., P.K., R.C.B., W.J.E., and Y.L.); The Global Challenges Research Fund (GCRF) project "RECAP" managed through RCUK and ESRC ES/P010873/1 (to A.G., C.I.J., and T.J.); HDR UK MR/S003975/1 (to R.M.E.); MRC MR/N013638/1 (to N.R.W.); Nakajima Foundation (to A.E.); The National Institute for Health Research NIHR 16/136/46 (to B.J.Q.); 16/137/109 (to B.J.Q., F.Y.S., M.J., and Y.L.); Health Protection Research Unit for Immunisation NIHR200929 (to N.G.D.); Health Protection Research Unit for Modelling Methodology HPRU-2012-10096 (to T.J.); NIHR200908 (to R.M.E.); NIHR200929 (to F.G.S. and M.J.); PR-OD-1017-20002 (to A.R. and W.J.E.); The Royal Society Dorothy Hodgkin Fellowship (to R.L.); RP\EA\180004 (to P.K.); UK DHSC/UK Aid/NIHR PR-OD-1017-20001 (to H.P.G.); UK MRC MC_PC_19065—Covid 19: Understanding the dynamics and drivers of the COVID-19 epidemic using real-time outbreak analytics (to A.G., N.G.D., R.M.E., S.C., T.J., W.J.E., and Y.L.); MR/P014658/1 (to G.M.K.); The UK Public Health Rapid Support Team funded by the United Kingdom Department of Health and Social Care (to T.J.); Wellcome Trust 206250/Z/17/Z (to A.J.K. and T.W.R.); 206471/Z/17/Z (to O.J.B.); 208812/Z/17/Z (to S.C.); 210758/Z/18/Z (to J.D.M., J.H., N.I.B., S.A., and S.R.M.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions: The following authors were part of the Centre for Mathematical Modelling of Infectious Disease (CMMID) COVID-19 Working Group. Each contributed in processing, cleaning, and interpretation of data; interpreted findings; contributed to the manuscript; and approved the work for publication: G. M. Knight, N. R. Waterlow, C. A. B. Pearson, F. Y. Sun, S. R. Procter, A. Showering, R. M. Eggo, Y.-W. Desmond Chan, E. S. Nightingale, D. Simons, O. Brady, B. J Quilty, P. Klepac, A. Gimma, H. P. Gibbs, W. J. Edmunds, A. J. Kucharski, S. Abbott, J. Williams, K. Prem, R. C. Barnard, T. Jombart, G. Medley, K. E. Atkins, S. Clifford, N. G. Davies, K. Abbas, M. Jit, T. W. Russell, F. G. Sandmann, D. C. Tully, J. D. Munday, A. M. Foss, A. Rosello, S. R. Meakin, J. Hellewell, C. J. Villabona-Arenas, C. I. Jarvis, R. Lowe, A. Endo, M. Quaife, N. I. Bosse, and Y. Liu. The following authors were part of the Inštitút Zdravotných Analýz. Each contributed in processing, cleaning, and interpretation of data; interpreted findings; contributed to the manuscript; and approved the work for publication: M. Mišík, K. Šufliarsky, A. Kluka, E. Kofira, P. Židuliak, S. Slobodníková, and J. Červený. Conceptualization: M.P., M.M., P.J., M.K., S.Fu., and S.Fl. Methodology: K.V.-Z., S.A., M.M., S.Fu., and S.Fl. Software: K.V.-Z., M.M., S.Fu., and S.Fl. Validation: K.V.-Z., S.A., K.S., P.J., M.K., S.Fu., and S.Fl. Formal analysis: M.P., K.V.-Z., S.A., K.S., M.M., S.Fu., and S.Fl. Investigation: M.P., K.V.-Z., P.J., S.Fu., and S.Fl. Resources: M.P., P.J., and M.K. Data curation: M.P., K.V.-Z., P.J., M.K., S.Fu., and S.Fl.. Writing, original draft: M.P., K.V.-Z., and S.Fl.. Writing, review and editing: M.P., K.S., M.M., P.J., M.K., S.Fu., and S.Fl. Visualization: M.P., K.V.-Z., S.Fu., and S.Fl. Supervision: P.J., M.K., S.Fu., and S.Fl.. Project administration: M.P., S.Fu., and S.Fl.. Funding acquisition: Table 1: M.P. and S.Fu.; Fig. 1: M.P., K.V.-Z., S.Fu., and S.Fl.; Fig. 2: M.P., K.V.-Z., S.Fu., and S.Fl.; Fig. 3: M.P., K.V.-Z., S.Fu., and S.Fl.; table S1: M.P., K.V.-Z., S.Fu., and S.Fl.; fig. S1: S.Fu.; fig. S2: K.V.-Z.; fig. S3: K.V.-Z.; fig. S4: S.Fu.; fig. S5: K.V.-Z.; fig. S6: M.P., S.Fu., and S.Fl.; fig. S7: M.P., S.Fu., and S.Fl.; fig. S8: M.P., S.Fu., and S.Fl.; table S2: S.Fu. and S.Fl.; table S3: M.P., S.Fu., and S.Fl.. S.Fl. and S.Fu. contributed equally to this work, and their order in the author list was determined with a coin toss. **Competing interests:** M.P. is employed as epidemiologist and public health data analyst at the Slovak Ministry of Health but had no involvement in the design of the mass testing campaigns. M.K. is the current Slovak Minister of Health and was involved in the design of the mass testing campaigns. S.Fu. advises the UK government as a member of the Scientific Pandemic Influenza Group on Modelling (SPI-M) in an unpaid role. All other authors declare that they have no conflicts of interest. **Data and materials availability:** Daily incidence of positive COVID-19 test reports and the results of the mass testing are available through governmental websites (*12*, *20*). All analyses were conducted in R (*21*) and can be found at (*22*).

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/372/6542/635/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 to S3
CMMID COVID-19 Working Group Authors List
Inštitút Zdravotných Analýz Authors List
References (*23*–*39*)
MDAR Reproducibility Checklist

1 December 2020; resubmitted 29 January 2021
Accepted 17 March 2021
Published online 23 March 2021
10.1126/science.abf9648

CORONAVIRUS

X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease

Sebastian Günther[1]*†, Patrick Y. A. Reinke[1]†, Yaiza Fernández-García[2], Julia Lieske[1], Thomas J. Lane[1], Helen M. Ginn[3], Faisal H. M. Koua[1], Christiane Ehrt[4], Wiebke Ewert[1], Dominik Oberthuer[1], Oleksandr Yefanov[1], Susanne Meier[5,6], Kristina Lorenzen[7], Boris Krichel[8], Janine-Denise Kopicki[8], Luca Gelisio[1], Wolfgang Brehm[1], Ilona Dunkel[9], Brandon Seychell[10], Henry Gieseler[5,6], Brenna Norton-Baker[11,12], Beatriz Escudero-Pérez[2], Martin Domaracky[1], Sofiane Saouane[13], Alexandra Tolstikova[1], Thomas A. White[1], Anna Hänle[1], Michael Groessler[1], Holger Fleckenstein[1], Fabian Trost[1], Marina Galchenkova[1], Yaroslav Gevorkov[1,14], Chufeng Li[1], Salah Awel[1], Ariana Peck[15], Miriam Barthelmess[1], Frank Schlünzen[1], P. Lourdu Xavier[1,11], Nadine Werner[16], Hina Andaleeb[16], Najeeb Ullah[16], Sven Falke[16], Vasundara Srinivasan[16], Bruno Alves França[16], Martin Schwinzer[16], Hévila Brognaro[16], Cromarte Rogers[5,6], Diogo Melo[5,6], Joanna J. Zaitseva-Doyle[5,6], Juraj Knoska[1], Gisel E. Peña-Murillo[1], Aida Rahmani Mashhour[1], Vincent Hennicke[1], Pontus Fischer[1], Johanna Hakanpää[13], Jan Meyer[13], Philip Gribbon[17], Bernhard Ellinger[17], Maria Kuzikov[17], Markus Wolf[17], Andrea R. Beccari[18], Gleb Bourenkov[19], David von Stetten[19], Guillaume Pompidor[19], Isabel Bento[19], Saravanan Panneerselvam[19], Ivars Karpics[19], Thomas R. Schneider[19], Maria Marta Garcia-Alai[19], Stephan Niebling[19], Christian Günther[19], Christina Schmidt[7], Robin Schubert[7], Huijong Han[7], Juliane Boger[20], Diana C. F. Monteiro[21], Linlin Zhang[20,22], Xinyuanyuan Sun[20,22], Jonathan Pletzer-Zelgert[4], Jan Wollenhaupt[23], Christian G. Feiler[23], Manfred S. Weiss[23], Eike-Christian Schulz[11], Pedram Mehrabi[11], Katarina Karničar[24,25], Aleksandra Usenik[24,25], Jure Loboda[24], Henning Tidow[5,26], Ashwin Chari[27], Rolf Hilgenfeld[20,22], Charlotte Uetrecht[8], Russell Cox[28], Andrea Zaliani[17], Tobias Beck[5,10], Matthias Rarey[4], Stephan Günther[2], Dusan Turk[24,25], Winfried Hinrichs[16,29], Henry N. Chapman[1,5,30], Arwen R. Pearson[5,6], Christian Betzel[5,16], Alke Meents[1]*

The coronavirus disease (COVID-19) caused by SARS-CoV-2 is creating tremendous human suffering. To date, no effective drug is available to directly treat the disease. In a search for a drug against COVID-19, we have performed a high-throughput x-ray crystallographic screen of two repurposing drug libraries against the SARS-CoV-2 main protease (M^{pro}), which is essential for viral replication. In contrast to commonly applied x-ray fragment screening experiments with molecules of low complexity, our screen tested already-approved drugs and drugs in clinical trials. From the three-dimensional protein structures, we identified 37 compounds that bind to M^{pro}. In subsequent cell-based viral reduction assays, one peptidomimetic and six nonpeptidic compounds showed antiviral activity at nontoxic concentrations. We identified two allosteric binding sites representing attractive targets for drug development against SARS-CoV-2.

Infection of host cells by SARS-CoV-2 is governed by the complex interplay of molecular factors from both the host and the virus (*1*, *2*). Coronaviruses are RNA viruses with a genome of approximately 30,000 nucleotides. The viral open reading frames are expressed as two overlapping large polyproteins which must be separated into functional subunits for replication and transcription activity (*1*). This proteolytic cleavage is primarily accomplished by the main protease (M^{pro}), also known as 3C-like protease $3CL^{pro}$ or nsp5. M^{pro} cleaves the viral polyprotein pp1ab at 11 distinct sites. The core cleavage motif is Leu-Gln↓(Ser/Ala/Gly) (*1*). M^{pro} possesses a chymotrypsin-like fold appended with a C-terminal helical domain and harbors a catalytic dyad comprised of Cys^{145} and His^{41} in its active site, which is formed by four major pockets that are labeled according to their position relative to the scissile bond of the substrate (Fig. 1) (*1*). The active site is located in a cleft between the two N-terminal domains of the three-domain structure of the monomer, whereas the C-terminal helical domain is involved in regulation and dimerization of the enzyme (Fig. 1A). Because of its central involvement in virus replication, M^{pro} is recognized as a prime target for antiviral drug discovery and compound screening activities aiming to identify and optimize drugs which can tackle coronavirus infections (*3*). Indeed, a number of recent publications confirm the potential of targeting M^{pro} for inhibition of virus replication (*1*, *2*, *4*).

In order to find drug candidates against SARS-CoV-2, we performed a large-scale x-ray crystallographic screen of M^{pro} against two repurposing libraries containing 5953 compounds from the Fraunhofer IME Repurposing Collection and the Safe-in-man library from Dompé Farmaceutici S.p.A. (*5*).

In contrast to crystallographic fragment screening experiments, compounds in repurposing libraries are chemically more complex (fig. S1A) (*6*, *7*). Thus, these compounds likely bind more specifically and with higher affinity (*8*). Because of the higher molecular weights, we performed cocrystallization experiments at a physiological pH of 7.5 instead of compound soaking into native crystals (*9*).

From the 5953 compounds in our screen, we obtained x-ray diffraction datasets for 2381 compounds, which we subjected to automated structure refinement followed by cluster analysis (*10*) and pan dataset density analysis (PanDDA) (*11*) (table S1). We observed additional electron density, indicating binding to M^{pro}, for 43 compounds, which were classified as hits, representing 37 distinct compounds (tables S1, S2, and S3). From these, the binding mode could be unambiguously determined for 29

[1]Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany. [2]Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Str. 74, 20359 Hamburg, Germany. [3]Diamond Light Source Ltd., Diamond House, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK. [4]Universität Hamburg, Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany. [5]Hamburg Centre for Ultrafast Imaging, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany. [6]Universität Hamburg, Institut für Nanostruktur- und Festkörperphysik, Luruper Chaussee 149, 22761 Hamburg, Germany. [7]European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany. [8]Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Martinistr. 52, 20251 Hamburg, Germany. [9]Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany. [10]Universität Hamburg, Department of Chemistry, Institute of Physical Chemistry, Grindelallee 117, 20146 Hamburg, Germany. [11]Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany. [12]Department of Chemistry, UC Irvine, Irvine, CA 92697-2025, USA. [13]Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607, Hamburg, Germany. [14]Vision Systems, Hamburg University of Technology, 21071 Hamburg, Germany. [15]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [16]Universität Hamburg, Department of Chemistry, Institute of Biochemistry and Molecular Biology and Laboratory for Structural Biology of Infection and Inflammation, c/o DESY, 22607 Hamburg, Germany. [17]Fraunhofer Institute for Translational Medicine and Pharmacology and Fraunhofer Cluster of Excellence for Immune Mediated Diseases, Schnackenburgallee 114, 22525 Hamburg, Germany. [18]Dompé Farmaceutici SpA, 67100 L'Aquila, Italy. [19]EMBL Outstation Hamburg, c/o DESY, Notkestr. 85, 22607 Hamburg, Germany. [20]Institute of Molecular Medicine, University of Lübeck, 23562 Lübeck, Germany. [21]Hauptmann Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, NY 14203, USA. [22]German Center for Infection Research, Hamburg-Lübeck-Borstel-Riems Site, University of Lübeck, 23562 Lübeck, Germany. [23]Helmholtz Zentrum Berlin, Macromolecular Crystallography, Albert-Einstein-Str. 15, 12489 Berlin, Germany. [24]Department of Biochemistry and Molecular and Structural Biology, Jozef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia. [25]Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Jamova 39, 1000 Ljubljana, Slovenia. [26]Universität Hamburg, Department of Chemistry, Institute of Biochemistry and Molecular Biology, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany. [27]Research Group for Structural Biochemistry and Mechanisms, Department of Structural Dynamics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. [28]Institute for Organic Chemistry and BMWZ, Leibniz University of Hannover, Schneiderberg 38, 30167 Hannover, Germany. [29]Universität Greifswald, Institute of Biochemistry, Felix-Hausdorff-Str. 4, 17489 Greifswald, Germany. [30]Universität Hamburg, Department of Physics, Luruper Chaussee 149, 22761 Hamburg, Germany.

†These authors contributed equally to this work.

*Corresponding author. Email: sebastian.guenther@desy.de (S.G.); alke.meents@desy.de (A.M.)

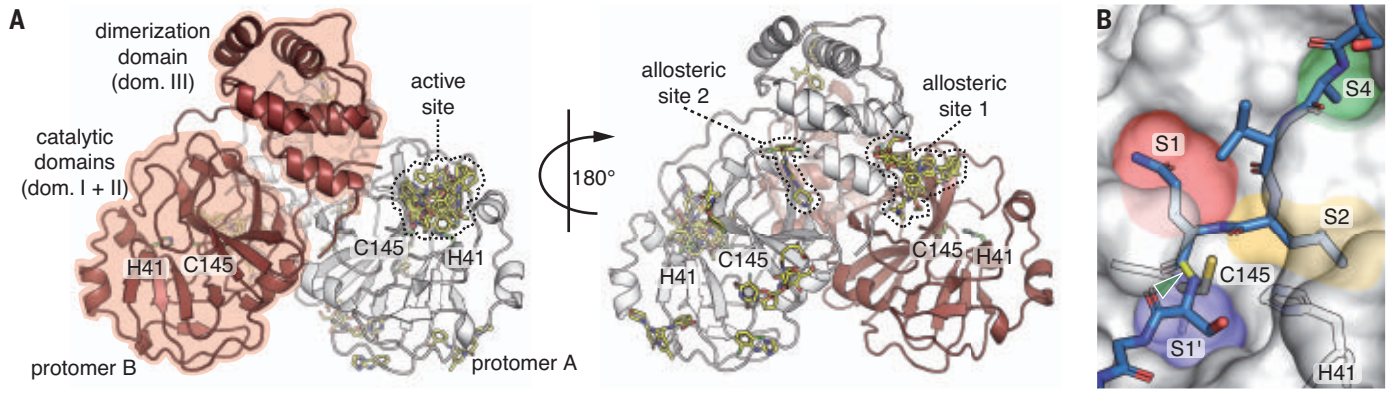


Fig. 1. The x-ray screening of drug-repurposing libraries reveals compound binding sites distributed across the complete M^{pro} surface. (**A**) Schematic drawing of M^{pro} dimer structure. Protomer A is shown in white, and protomer B is in red. For clarity, the 29 binding compounds (yellow sticks) are only depicted on one of the two protomers. Catalytic residues His^{41} (H41) and Cys^{145} (C145), the active site, and two allosteric drug binding sites are highlighted. (**B**) Close-up view of the active site with peptide substrate bound (blue sticks), modeled after SARS-CoV M^{pro} (PDB 2Q6G). The scissile bond is indicated in yellow and with the green arrowhead. Substrate binding pockets S1′, S1, S2, and S4 are indicated by colored regions.

molecules (Fig. 1A and table S4). The majority of hits were found in the active site of the enzyme. Of the 16 active site binders, six covalently bind as thioethers to Cys^{145}, one compound binds covalently as a thiohemiacetal to Cys^{145}, one is zinc-coordinated, and eight bind noncovalently. The remaining 13 compounds bind outside the active site at various locations (Fig. 1A).

Of the 43 hits from our x-ray screen, 37 compounds were available in quantities required for testing their antiviral activity against SARS-CoV-2 in cell assays (table S2). Nine compounds that reduced viral RNA (vRNA) replication by at least two orders of magnitude in Vero E6 cells (fig. S2) were further evaluated to determine the effective concentrations that reduced not only vRNA but also SARS-CoV-2 infectious particles by 50% (EC_{50}) (Fig. 2). Additionally, AT7519 and ifenprodil, which showed slightly lower vRNA level reduction, were included because of their distinct binding sites outside of the active site. From these 11, seven compounds (AT7519, calpeptin, ifenprodil, MUT056399, pelitinib, tolperisone, and triglycidyl isocyanurate) exhibited a ≥100-fold reduction in infectious particles in combination with either a selectivity index [SI; calculated as the 50% cytotoxic concentration (CC_{50}) divided by the EC_{50}] of >5 or no cytotoxicity in the tested concentration range and are considered antivirally active (table S5).

Here, we focus on a more detailed description of the 11 compounds analyzed in the secondary screen, which are grouped according to their different binding sites. The remaining hits are described in the supplementary text and figs. S3 to S5.

Tolperisone, 2-[β-(4-hydroxyphenyl)-ethylaminomethyl]-tetralone (HEAT), and isofloxythepin bind covalently to the active site. Tolperisone is antivirally active (EC_{50} = 19.17 μM) and shows no cytotoxicity (CC_{50} > 100 μM) (Fig. 2), whereas HEAT (EC_{50} = 24.05 μM, CC_{50} = 55.42 μM) and isofloxythepin (EC_{50} = 4.8 μM, CC_{50} = 17 μM) show unfavorable cytotoxicity. For all three compounds, only breakdown products are observed in the active site. Tolperisone and HEAT are β-aminoketones, but we only observe the part of the drug containing the ketone (2,4′-dimethylpropiophenone and 2-methyl-1-tetralone), whereas the remaining part with the amine group is missing. The breakdown product binds as a Michael acceptor to the thiol of Cys^{145}, independently confirmed for HEAT by mass spectrometry (fig. S6 and table S6). The decomposition of tolperisone and HEAT was detected in both the crystallization and cell culture conditions (fig. S7) and is reported to be pH dependent (*12*). The parent compounds can be regarded as prodrugs (*13*, *14*). In the x-ray structures the aromatic ring systems of tolperisone (Fig. 3A) and HEAT (Fig. 3B) protrude into the S1 pocket and form van der Waals contacts with the backbone of Phe^{140} and Leu^{141} and the side chain of Glu^{166}. In addition, the keto group accepts a hydrogen bond from the imidazole side chain of His^{163}. Tolperisone is used as a skeletal muscle relaxant (*15*). The x-ray structure suggests that isofloxythepin binds similarly as a fragment to Cys^{145} (Fig. 3C).

Triglycidyl isocyanurate has antiviral activity (EC_{50} = 30.02 μM, CC_{50} > 100 μM) and adopts covalent and noncovalent binding modes to the active site. In both modes, the compound's central ring sits on top of the catalytic dyad (His^{41}, Cys^{145}), and its three epoxypropyl substituents reach into subsites S1′, S1, and S2. The noncovalent binding mode is stabilized by hydrogen bonds to the main chain of Gly^{143} and Gln^{166} and to the side chain of His^{163}. In the covalently bound form, one oxirane ring is opened by nucleophilic attack of Cys^{145}, forming a thioether (Fig. 3D). Triglycidyl isocyanurate has been tested as an antitumor agent (*16*).

Calpeptin shows the highest antiviral activity in the screen (EC_{50} = 72 nM, CC_{50} > 100 μM). It binds covalently via its aldehyde group to Cys^{145}, forming a thiohemiacetal. This peptidomimetic inhibitor occupies substrate pockets S1 to S3, similar to the peptidomimetic inhibitors GC-376 (*17*, *18*), calpain inhibitors (*19*), N3 (*2*), and the α-ketoamide 13b (*1*). The peptidomimetic backbone forms hydrogen bonds to the main chain of His^{164} and Glu^{166}, whereas the norleucine side chain maintains van der Waals contacts with the backbone of Phe^{140}, Leu^{141}, and Asn^{142} (Fig. 3E). Calpeptin has known activity against SARS-CoV-2 M^{pro} in enzymatic assays (*17*). The structure is highly similar to the common protease inhibitor leupeptin (fig. S3A), which served as a positive control in our x-ray screen but was not tested further in antiviral assays. In silico docking experiments also suggested calpeptin as a possible M^{pro} binding molecule (table S7). Calpeptin also inhibits cathepsin L (*20*), and dual targeting of cathepsin L and M^{pro} is suggested as an attractive path for SARS-CoV-2 inhibition (*19*).

MUT056399 binds noncovalently to the active site (EC_{50} = 38.24 μM, CC_{50} > 100 μM). The diphenyl ether core of MUT056399 blocks access to the catalytic site, which consists of Cys^{145} and His^{41}. The terminal carboxamide group occupies pocket S1 and forms hydrogen bonds to the side chain of His^{163} and the backbone of Phe^{140} (Fig. 3F). The ethyl phenyl group of the molecule reaches deep into pocket S2, which is enlarged by a shift of the side chain of Met^{49} out of the substrate binding pocket. MUT056399 was developed as an antibacterial agent against multidrug-resistant *Staphylococcus aureus* strains (*21*).

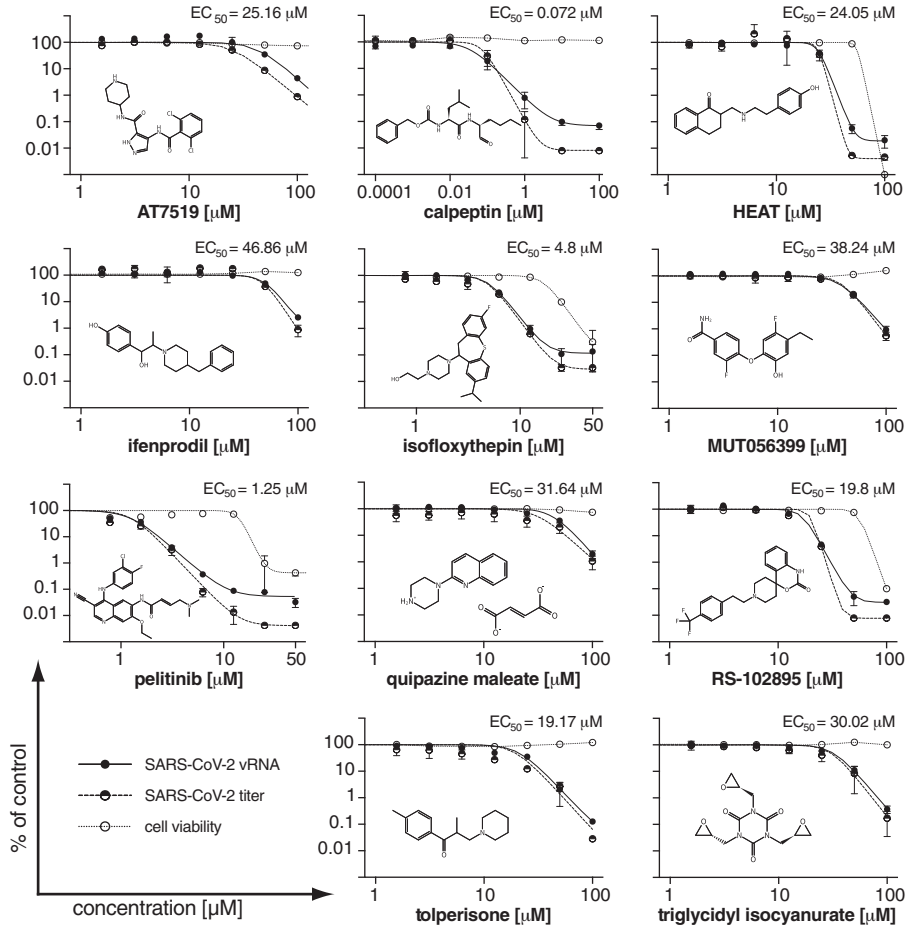


Fig. 2. Effect of selected compounds on SARS-CoV-2 replication in Vero E6 cells. The vRNA yield (solid circles), viral titers (half-solid circles), and cell viability (empty circles) were determined by reverse transcription–quantitative polymerase chain reaction, immunofocus assays, and the CCK-8 method, respectively. EC_{50} for the viral titer reduction is shown. Individual data points represent means ± SD from three independent replicates in one experiment.

Quipazine maleate showed moderate antiviral activity (EC_{50} = 31.64 μM, CC_{50} > 100 μM). In the x-ray structure, only the maleate counterion is observed covalently bound as a thioether (supplementary text and fig. S3B). Maleate is observed in structures of six other compounds showing no antiviral activity. The observed antiviral activity is thus likely caused by an off-target effect of quipazine.

In general, the enzymatic activity of M^{pro} relies on the architecture of the active site, which critically depends on the dimerization of the enzyme and the correct relative orientation of the subdomains. This could allow ligands that bind outside of the active site to affect activity. In fact, we identified two such allosteric binding sites of M^{pro}.

Five compounds of our x-ray screen bind in a hydrophobic pocket in the C-terminal dimerization domain (Fig. 4, A and B), located close to the oxyanion hole in pocket S1 of the substrate binding site. One of these showed strong antiviral activity (Fig. 2). Another compound binds between the catalytic and dimerization domains of M^{pro}.

Central to the first allosteric binding site is a hydrophobic pocket formed by Ile^{213}, Leu^{253}, Gln^{256}, Val^{297}, and Cys^{300} within the C-terminal dimerization domain (Fig. 4A). Pelitinib, ifenprodil, RS-102895, PD-168568, and tofogliflozin all exploit this site by inserting an aromatic moiety into this pocket.

Pelitinib shows the second highest antiviral activity in our screen (EC_{50} = 1.25 μM, CC_{50} = 13.96 μM). Its halogenated benzene ring binds to the hydrophobic groove in the helical domain, which becomes accessible by movement of the Gln^{256} side chain (Fig. 4A). The central 3-cyanoquinoline moiety interacts with the end of the C-terminal helix (Ser^{301}). The ethyl ether substituent pushes against Tyr^{118} and Asn^{142} (from loop 141–144 of the S1 pocket) of the opposing protomer within the native dimer. The integrity of this pocket is crucial for enzyme activity (*22*). Pelitinib is an amine-catalyzed Michael acceptor (*23*) and was developed as an anticancer agent to bind to a cysteine in the active site of the tyrosine kinase epidermal growth factor receptor inhibitor (*24*). However, from its observed binding position, it is impossible for it to reach into the active site, and no evidence for covalent binding to Cys^{145} is found in the electron density maps.

Ifenprodil and RS-102895 bind to the same hydrophobic pocket in the dimerization domain as pelitinib (Fig. 4B; fig. S4, A and B; and supplementary text). Only ifenprodil (EC_{50} = 46.86 μM, CC_{50} > 100 μM) shows moderate activity. RS-102895 (EC_{50} = 19.8 μM, CC_{50} = 54.98 μM) interacts, similar to pelitinib, with the second protomer by forming two hydrogen bonds to the side and main chains of Asn^{142}, whereas the other compounds exhibit weaker or no interaction with the second protomer. PD-168568 and tofogliflozin bind the same site but are inactive (Fig. 4B and fig. S4, C and D).

The second allosteric site is formed by the deep groove between the catalytic domains and the dimerization domain. AT7519 is the only compound in our screen that we identified bound to this site (Fig. 4C). Though it has only moderate activity, we discuss it here because this site may be a target. The chlorinated benzene ring is engaged in various van der Waals interactions to loop 107-110, Val^{202}, and Thr^{292}. The central pyrazole has van der Waals contacts to Ile^{249} and Phe^{294}, and its adjacent carbonyl group forms a hydrogen bond to the side chain of Gln^{110}. The terminal piperidine sits on top of Asn^{151} and forms hydrogen bonds to the carboxylate of Asp^{153}. This results in a displacement of loop 153-155, slightly narrowing the binding groove. The Cα atom of Tyr^{154} moves 2.8 Å, accompanied by a conformational change of Asp^{153} (Fig. 4D). This allows hydrogen bonding to the compound and the formation of a salt bridge to Arg^{298}. Arg^{298} is crucial for dimerization (*25*). The mutation $Arg^{298}Ala$ causes a reorientation of the dimerization domain relative to the catalytic domain, leading to changes in the oxyanion hole and destabilization of the S1 pocket by the N terminus. AT7519 was evaluated for treatment of human cancers (*26*). The potential of allosteric inhibition of M^{pro} through modulation of Arg^{298} has been independently demonstrated by mass spectrometry (*27*).

Our x-ray screen revealed 43 compounds binding to M^{pro}, with seven compounds showing antiviral activity against SARS-CoV-2. We present structural evidence for interaction of these compounds at active and allosteric sites of M^{pro}, although we cannot exclude that off-target effects played a role in the antiviral effect in cell culture, in particular for compounds with a low selectivity index. Conversely, an absence of antiviral activity of compounds binding clearly to M^{pro} in the crystal might be due to rapid metabolization in the cellular

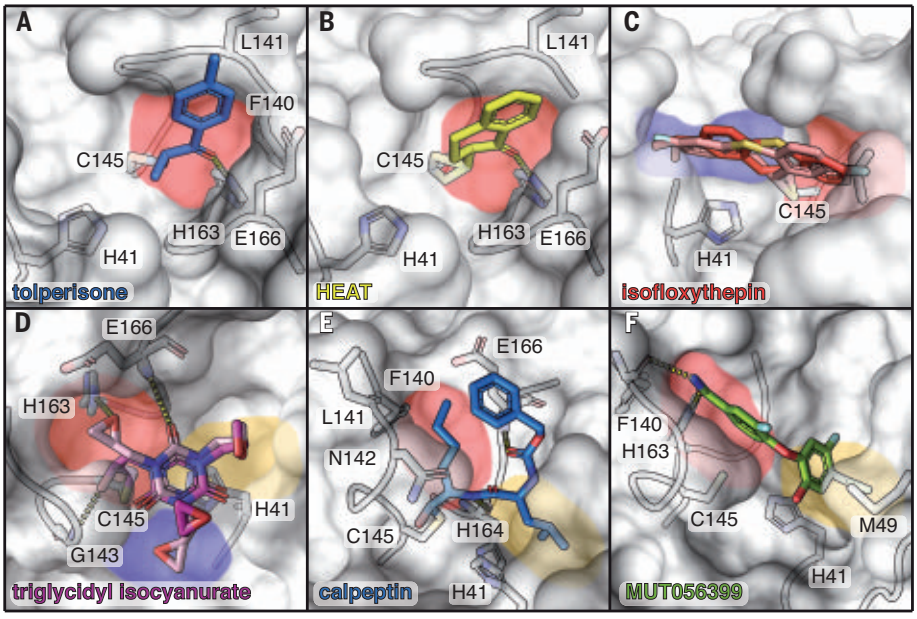


Fig. 3. Covalent and noncovalent binders in the active site of M^{pro}. Bound compounds are depicted as colored sticks, and the surface of M^{pro} is shown in gray with selected interacting residues shown as sticks. Substrate binding pockets are colored as in Fig. 1. Hydrogen bonds are depicted by dashed lines. (**A**) Tolperisone. (**B**) HEAT. (**C**) Isofloxythepin. (**D**) Triglycidyl isocyanurate. (**E**) Calpeptin. (**F**) MUT056399.

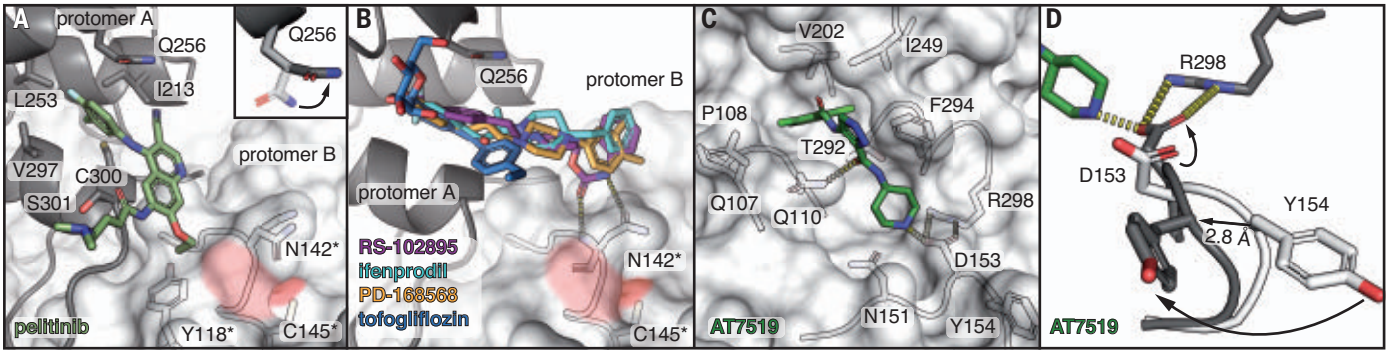


Fig. 4. Screening hits at allosteric sites of M^{pro}. (**A**) Close-up view of the binding site in the dimerization domain (protomer A, gray cartoon representation), close to the active site of the second protomer (protomer B, surface representation) in the native dimer. Residues forming the hydrophobic pocket are indicated. Pelitinib (dark green) binds to the C-terminal α-helix at Ser^{301} and pushes against Asn^{142} and the β-turn of the pocket S1 of protomer B (residues marked with an asterisk). The inset shows the conformational change of Gln^{256} (gray sticks) compared with the M^{pro} apo structure (white sticks). (**B**) RS-102895 (purple), ifenprodil (cyan), PD-168568 (orange), and tofogliflozin (blue) occupy the same binding pocket as pelitinib. (**C**) AT7519 occupies a deep cleft between the catalytic and dimerization domain of M^{pro}. (**D**) Conformational changes in the AT7519-bound M^{pro} structure (gray) compared with those in the apo structure (white).

environment. Calpeptin and pelitinib showed strong antiviral activity with low cytotoxicity and are suitable for preclinical evaluation. In any case, all hit compounds are valuable lead structures with potential for further drug development, especially because drug-repurposing libraries offer the advantage of proven bioactivity and cell permeability (*28*).

The most active compound, calpeptin, binds in the active site similar to other members of the large class of peptide-based inhibitors that bind as thiohemi-acetals or -ketals to M^{pro} (*29*). In addition to this peptidomimetic inhibitor, we discovered several nonpeptidic inhibitors. Those compounds binding to the active site of M^{pro} contained new Michael acceptors based on β-aminoketones (tolperisone and HEAT). These compounds lead to the formation of thioethers and have not been described as prodrugs for viral proteases. We also identified a noncovalent binder, MUT056399, that blocked the active site. In addition to this common active site inhibition, we identified compounds that inhibit the enzyme through binding at two allosteric sites of M^{pro}.

The first allosteric site (dimerization domain) is in the direct vicinity of the S1 pocket of the adjacent monomer within the native dimer. The potential for antiviral inhibition through this site is demonstrated by pelitinib. The hydrophobic nature of the residues forming the main pocket is conserved in all human coronavirus M^{pro} (fig. S8). Consequently, potential drugs targeting this binding site may be effective against other coronaviruses. The potential of the second allosteric site as a druggable target is demonstrated by the observed moderate antiviral activity of AT7519.

REFERENCES AND NOTES

1. L. Zhang *et al.*, *Science* **368**, 409–412 (2020).
2. Z. Jin *et al.*, *Nature* **582**, 289–293 (2020).
3. R. Hilgenfeld, *FEBS J.* **281**, 4085–4096 (2014).
4. J. Qiao *et al.*, *Science* **371**, 1374–1378 (2021).
5. M. Kuzikov *et al.*, *ACS Pharmacol. Transl. Sci.* 10.1021/acsptsci.0c00216 (2021).
6. J. Wollenhaupt *et al.*, *Structure* **28**, 694–706.e5 (2020).
7. O. B. Cox *et al.*, *Chem. Sci.* **7**, 2322–2330 (2016).
8. M. M. Hann, A. R. Leach, G. Harper, *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
9. F. R. Ehrmann *et al.*, *PLOS ONE* **12**, e0175723 (2017).
10. H. M. Ginn, *Acta Crystallogr. D Biol. Crystallogr.* **76**, 1134–1144 (2020).
11. N. M. Pearce *et al.*, *Nat. Commun.* **8**, 15123 (2017).
12. A. L. Simplício, J. M. Clancy, J. F. Gilmer, *Int. J. Pharm.* **336**, 208–214 (2007).
13. J. Rautio, N. A. Meanwell, L. Di, M. J. Hageman, *Nat. Rev. Drug Discov.* **17**, 559–587 (2018).

14. M. Altmeyer *et al.*, *Bioorg. Med. Chem. Lett.* **24**, 5310–5314 (2014).
15. S. Quasthoff, C. Möckel, W. Zieglgänsberger, W. Schreibmayer, *CNS Neurosci. Ther.* **14**, 107–119 (2008).
16. M. Piccart *et al.*, *Eur. J. Cancer Clin. Oncol.* **17**, 1263–1266 (1981).
17. C. Ma *et al.*, *Cell Res.* **30**, 678–692 (2020).
18. W. Vuong *et al.*, *Nat. Commun.* **11**, 4282 (2020).
19. M. D. Sacco *et al.*, *Sci. Adv.* **6**, eabe0751 (2020).
20. T. Sasaki *et al.*, *J. Enzyme Inhib.* **3**, 195–201 (1990).
21. S. Escaich *et al.*, *Antimicrob. Agents Chemother.* **55**, 4692–4697 (2011).
22. J. Tan *et al.*, *J. Mol. Biol.* **354**, 25–40 (2005).
23. A. Wissner *et al.*, *J. Med. Chem.* **46**, 49–63 (2003).
24. C. Erlichman *et al.*, *J. Clin. Oncol.* **24**, 2252–2260 (2006).
25. J. Shi, J. Sivaraman, J. Song, *J. Virol.* **82**, 4620–4629 (2008).
26. P. G. Wyatt *et al.*, *J. Med. Chem.* **51**, 4986–4999 (2008).
27. T. J. El-Baba *et al.*, *Angew. Chem. Int. Ed.* **59**, 23544–23548 (2020).
28. S. Pushpakom *et al.*, *Nat. Rev. Drug Discov.* **18**, 41–58 (2019).
29. Y. Liu *et al.*, *Eur. J. Med. Chem.* **206**, 112711 (2020).

ACKNOWLEDGMENTS

We acknowledge Deutsches Elektronen-Synchrotron (DESY; Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities. Parts of this research were carried out at PETRA III at beamline P11. Further MX data were collected at beamline P13 and P14 operated by EMBL. We thank the DESY machine group, in particular M. Wunderlich, K. Heuck, A. Brinkmann, O. Goldbeck, J. Haar, T. Schulz, G. Priebe, M. Holz, B. Lemcke, K. Knaack, O. Seebauer, P. Willanzheimer, R. Jonas, and N. Engling. We thank T. Dietrich, S. Geile, F. Guicking, H. Noei, and T. Pakendorf from DESY and B. Di Fabrizio and S. Kühn from BNITM for assistance. This research was supported in part through the Maxwell computational resources operated at DESY. We acknowledge the use of the XBI biological sample preparation laboratory at European XFEL, enabled by the XBI User Consortium. **Funding:** We acknowledge financial support from the EXSCALATE4CoV EU-H2020 Emergency Project (101003551), the Cluster of Excellence "Advanced Imaging of Matter" of the Deutsche Forschungsgemeinschaft (DFG), EXC 2056 project ID 390715994, the Helmholtz Association Impulse and Networking funds (projects ExNet-0002 and InternLabs-0011 "HIR3X"), the Federal Ministry of Education and Research (BMBF) via projects 05K16GUA, 05K19GU4, 05K20BI1, 05K20FL1, 16GW0277, and 031B0405D, and the Joachim-Herz-Stiftung Hamburg (project Infecto-Physics). C.E. and M.R. acknowledge financial support from grant HIDSS-0002 DASHH (Data Science in Hamburg, HELMHOLTZ, Graduate School for the Structure of Matter). R.C. is supported by DFG grants INST 187/621-1 and INST 187/686-1. D.T. is supported by the Slovenian Research Agency (research program P1-0048, Infrastructural program IO-0048). B.S. was supported by an Exploration Grant from the Boehringer Ingelheim Foundation. The Heinrich Pette Institute, Leibniz Institute for Experimental Virology was supported by the Free and Hanseatic City of Hamburg and the Federal Ministry of Health. C.U. and B.K. were supported by EU Horizon 2020 ERC StG-2017 759661, BMBF RTK Struktur 01KI20391, BMBF Visavix 05K16BH1, and the Leibniz Association SAW-2014-HPI-4 grant. **Author contributions:** Se.G., P.Y.A.R., Y.F.-G., W.B., P.G., A.R.B., R.C., D.T., A.Z., H.N.C., A.R.P., C.B., and A.M. designed the research. Se.G., P.Y.A.R., T.J.L., W.H., H.N.C., A.R.P., C.B., and A.M. wrote the manuscript. Se.G., P.Y.A.R., J.L., F.H.M.K., S.M., W.B., I.D., B.S., H.Gie., B.N.-B., M.B., P.L.X., N.W., H.A., N.U., S.F., B.A.F., M.S., H.B., J.K., G.E.P.-M., A.R.M., P.G., V.H., P.F., M.W., E.-C.S., P.M., H.T., and T.B. participated in sample preparation. P.Y.A.R., performed crystallization experiments. Se.G., P.Y.A.R., J.L., T.J.L., O.Y., S.S., A.T., M.Gr., H.F., F.T., M.Ga., Y.G., C.L., S.A., A.P., G.B., D.v.S., G.P., T.R.S., I.B., and S.P. performed x-ray data collection. T.J.L., H.M.G., D.O., O.Y., L.G., M.D., T.A.W., F.S., C.R., D.M., J.J.Z.-D., I.K., C.S., R.S., H.H., and D.C.F.M. contributed to x-ray data management. Se.G., P.Y.A.R., J.L., T.J.L., H.M.G., F.H.M.K., W.E., D.O., A.H., V.S., J.H., J.M., J.B., J.W., C.G.F., M.S.W., A.C., D.T., W.H., and A.M. performed x-ray data analysis. K.L., B.K., C.U., and R.C. performed and analyzed mass spectrometry experiments. Y.F.-G., B.E.-P., and St.G. performed and analyzed antiviral activity assays. P.G., B.E., M.K., M.M.G.-A., S.N., C.G., L.Z., X.S., K.K., A.U., J.L., and R.H. performed and analyzed ligand binding studies and protein activity assays. C.E., J.P.-Z., and M.R. performed computational binding studies. **Competing interests:** M.R. is a stakeholder of BioSolveIT GmbH, licensor of the software HYDE. **Data and materials availability:** The coordinates and structure factors for all described crystal structures of SARS-CoV-2 M^{pro} in complex with compounds are deposited in the PDB with accession codes 6YNQ, 6YVF, 7A1U, 7ABU, 7ADW, 7AF0, 7AGA, 7AHA, 7AK4, 7AKU, 7AMJ, 7ANS, 7AOL, 7AP6, 7APH, 7AQE, 7AQI, 7AQJ, 7AR5, 7AR6, 7ARF, 7AVD, 7AWR, 7AWS, 7AWU, 7AWW, 7AX6, 7AXM, 7AXO, 7AY7, 7B83, and 7NEV. Code used in this analysis has been previously published (*10*). The code for forcing adherence to the Wilson distribution is included in the Vagabond refinement package (https://vagabond.hginn.co.uk/) under a GPLv3 license. Compounds from the Fraunhofer IME Repurposing Collection were obtained from the Fraunhofer Institute for Molecular Biology and Applied Ecology under a material transfer agreement. Compounds from the Safe-in-man Library were kindly provided by Dompé Farmaceutici S.p.A. Other materials are available from Se.G. or A.M. upon request. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/642/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S7
References (*30*–*54*)
MDAR Reproducibility Checklist

20 November 2020; accepted 29 March 2021
Published online 2 April 2021
10.1126/science.abf7945

MIGRATION

Extreme altitudes during diurnal flights in a nocturnal songbird migrant

Sissel Sjöberg[1,2]*, Gintaras Malmiga[3], Andreas Nord[1], Arne Andersson[1], Johan Bäckman[1], Maja Tarka[1], Mikkel Willemoes[1], Kasper Thorup[2], Bengt Hansson[1], Thomas Alerstam[1], Dennis Hasselquist[1]

Billions of nocturnally migrating songbirds fly across oceans and deserts on their annual journeys. Using multisensor data loggers, we show that great reed warblers (*Acrocephalus arundinaceus*) regularly prolong their otherwise strictly nocturnal flights into daytime when crossing the Mediterranean Sea and the Sahara Desert. Unexpectedly, when prolonging their flights, they climbed steeply at dawn, from a mean of 2394 meters above sea level to reach extreme cruising altitudes (mean 5367 and maximum 6267 meters above sea level) during daytime flights. This previously unknown behavior of using exceedingly high flight altitudes when migrating during daytime could be caused by diel variation in ambient temperature, winds, predation, vision range, and solar radiation. Our finding of this notable behavior provides new perspectives on constraints in bird flight and might help to explain the evolution of nocturnal migration.

Most songbird migrants are known to preferably fly at night throughout their entire migratory journeys. However, recent geolocator studies have shown that some individuals of several songbird species sometimes prolong nighttime flights into daytime, or even change to a nonstop flight strategy, when crossing the inhospitable Sahara Desert (*1*, *2*).

Current knowledge suggests that songbirds typically migrate below 2000 m above ground level (*3*). By selecting altitudes with favorable winds, birds can substantially increase the flight range and thus reduce energy expenditure and water loss (*4*, *5*). However, knowledge of how birds regulate altitude throughout their flights is limited because of previous difficulties in quantifying individual flight characteristics over longer distances. Recent studies have revealed a more complex altitudinal pattern than previously assumed (*6*, *7*), suggesting that migrants may change flight altitude to gain wind support and avoid warm ambient temperatures (*8*). Flights as high as 5000 m above sea level (m asl) or higher have been observed in songbirds crossing ecological barriers [e.g., (*3, 9*)] but have hitherto been interpreted as rare extreme cases.

To provide more insights into flight behavior of migrants crossing barriers, we equipped great reed warblers (*Acrocephalus arundinaceus*) with custom-made multisensor data loggers. The great reed warbler is a medium-sized songbird (~30 g) that performs solitary nocturnal migratory flights between breeding sites in Europe (in this case, Sweden) and wintering areas in tropical Africa (*10*). This migratory cycle includes crossings of barriers, i.e., the Mediterranean Sea and the Sahara Desert, twice a year. The data loggers recorded physical data on light period and sensor temperature and behavioral data on altitude [obtained from barometer data (*7*)] and activity [exact timing of continuous flapping flight and resting and movements on the ground based on accelerometer data (*11*)]. We obtained such data for 23 barrier crossings of 14 individuals (nine in autumn, 14 in spring). Each individual made at least one prolonged flight during the barrier crossing (i.e., flights starting in the nighttime

[1]Department of Biology, Lund University, Lund, Sweden. [2]Center for Macroecology, Evolution and Climate, Globe Institute; University of Copenhagen, Copenhagen, Denmark. [3]Nature Research Centre, Vilnius, Lithuania.
*Corresponding author. Email: sissel.sjoberg@biol.lu.se

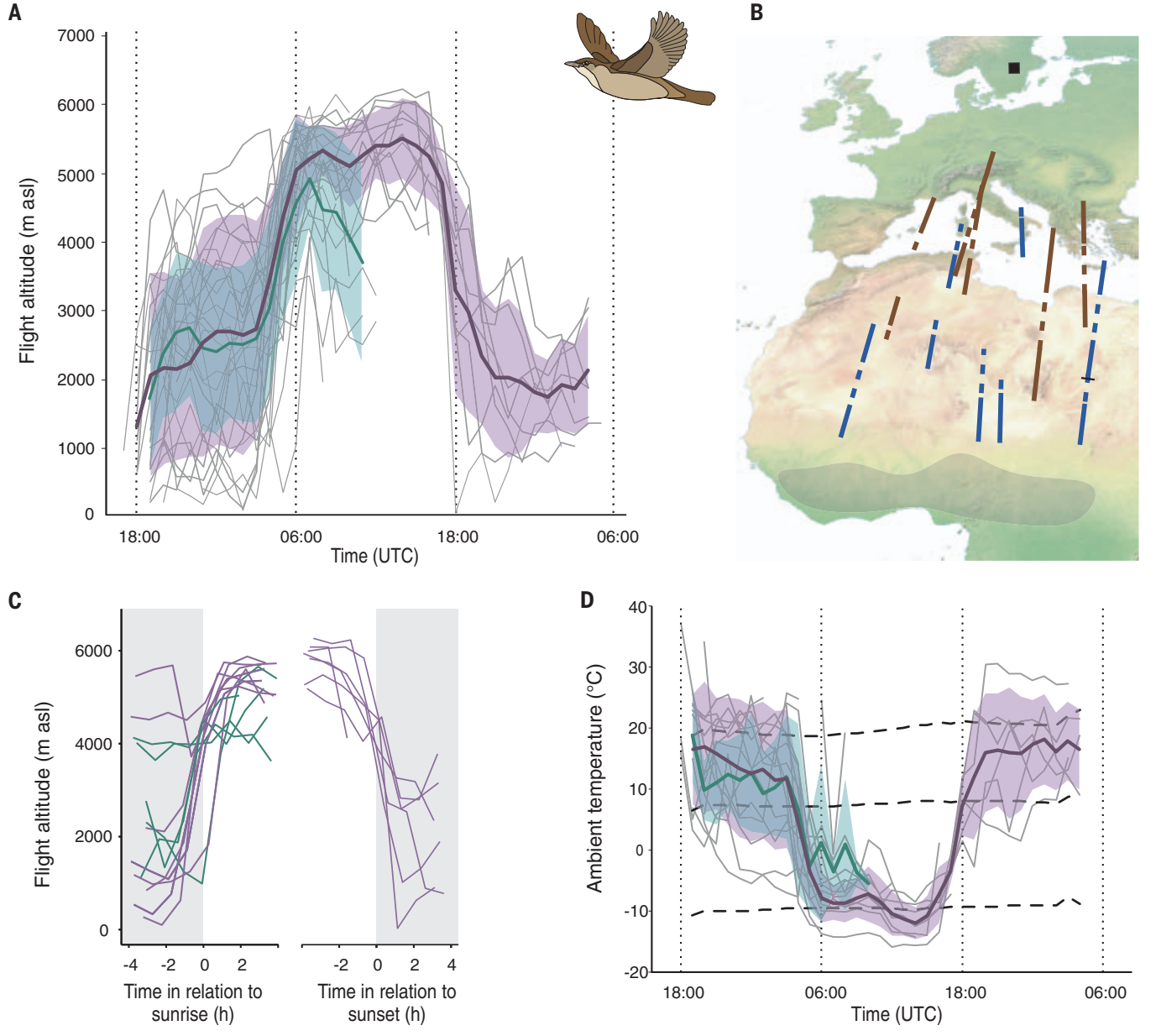


Fig. 1. Flight altitudes and ambient temperatures during prolonged migratory flights by great reed warblers over the Mediterranean Sea and Sahara Desert. (**A**) Altitude profiles in relation to the time of day and night for prolonged flights into the day in otherwise typically nocturnally migrating great reed warblers. Flight altitudes were significantly higher during daytime compared with nighttime hours [colored lines and areas shows the mean ± SD for fully (purple) and partly (green) prolonged flights; gray lines show individual flights; linear mixed-models test of difference in flight altitude between nocturnal and diurnal flights: df = 342.7, t = 25.3, P < 0.001; table S3]. (**B**) Approximate locations of the prolonged flights [n = 7 individuals and n = 14 flights; spring is shown in blue; autumn is shown in brown; diurnal part of flight indicated by broken lines; wintering areas shadowed; background map was made from Natural Earth, https://naturalearthdata.com, and positions of prolonged flights were generated using QGIS version 3.4.11 (*23*)]. Square shows the breeding site of the population (Lake Kvismaren, Sweden) where the data loggers were attached. (**C**) Ascents and descents in flight altitude in relation to sunset and sunrise for the 14 flights where locations were estimated (gray areas indicate night). (**D**) Ambient temperature profiles in relation to time of day for the prolonged flights [colored lines and areas shows the mean ± SD for fully (purple) and partly (green) prolonged flights; gray lines show individual flights; only flights with an estimated location are included]. The three dashed lines show ambient temperatures (~20°C, ~8°C, and ~−10°C) that would be expected if the fully prolonged flights were at a constant pressure level of 850, 700, and 500 hPa, respectively, during both night and day (corresponding to constant altitudes at 1457, 3011, and 5573 m asl, respectively).

and continuing during daytime after 6:30 Coordinated Universal Time). In total, we recorded 29 prolonged flights ranging from 12.1 to 34.2 hours (table S1). For 14 of these prolonged flights (involving seven individuals), it was also possible to determine the approximate geographic location of the associated flight sessions from light data (Fig. 1B and table S2), allowing us to extract relevant data on air temperature and geopotential height from the National Centers for Environmental Predictions (NCEP) data.

An unexpected pattern emerged when we analyzed the prolonged flights [which only occurred during barrier crossings (*1*)]. The

birds consistently made a substantial climb at dawn (Fig. 1A). Accordingly, whereas the overall nighttime mean flight altitude was relatively high (mean ± SD: 2394 ± 1270 m asl; N = 29 prolonged flights), cruising altitude in the daytime was consistently extremely high (mean ± SD: 5367 ± 589 m asl; N = 13 full day flights; maximum altitude range 5445 to 6267 m asl; Fig. 1A and table S1). The birds climbed and descended relatively steeply immediately before or at sunrise and at or immediately after sunset, respectively (Fig. 1C). Although the birds were flying at ambient temperatures of 13.8 ± 9.0°C at night (mean ± SD for times, altitudes, and locations of seven birds with available light data), the climb at dawn left them flying in much colder ambient temperatures of −9.3 ± 3.9°C during the day (Fig. 1D).

The explanations for the marked increase in flight altitude between night and day are not clear. We will briefly discuss five possible hypotheses related to diel variation in ambient temperature, winds, predation, vision range, and solar radiation. It has been proposed that birds reduce evaporative water loss and risk of dehydration by performing sustained migratory flights at altitudes with cool temperatures [below 15 to 20°C (*5*, *12*)]. Moreover, birds are expected to prefer flight altitudes with the most favorable winds (*4*, *5*, *8*). However, diel variation in ambient temperature and wind speed and/or direction is expected to be small above 1500 m asl in the troposphere [at pressure levels of 850 to 300 hPa; Fig. 1D; (*13, 14*)]. Therefore, for birds that fly above 1500 m asl without changing their flight altitude between night and day, only small differences in air temperature and wind vector would be expected between the diurnal and nocturnal parts of the flight.

Migrating birds are liable to predation on their flights across the Mediterranean Sea and the Sahara Desert, mainly from falcons that hunt from dawn and into the daytime hours. In particular, Eleonora's falcons (*Falco eleonora*), which breed in the Mediterranean region during late summer and early autumn and travel across the Sahara Desert during autumn and spring on their way to and from winter quarters in Madagascar (*15*), constitute a possible threat to songbird migrants. This may have triggered the great reed warblers to climb to altitudes above the falcons' hunting range in the daytime. How high this range may extend is presently unknown, but there are estimates of hunting altitudes up to 3500 m asl (*16*).

By climbing to high altitudes, migrants may extend their vision range when daylight provides full visibility of the landscape below. Because of the Earth's curvature, the theoretical range of vision to an object on ground level will approximately equal $\sqrt{2 \times h \times R}$, where h is the altitude of the observer and R is the radius of Earth (6371 km). Therefore, from altitudes of 0.5, 2, and 5 km, the theoretical range of vision will be ~80, ~160, and ~252 km, respectively. Landmarks protruding above the surface, such as mountains, will theoretically be visible at even longer ranges depending on their height. Even if atmospheric visibility conditions limit long-range vision to large and contrasting features, and the great reed warblers are nocturnal migrants well adapted for flying and navigating during the night, one cannot exclude the possibility that an improved overview of the landscape from higher altitude may be useful, e.g., when searching for suitable stopover and/or landing habitat.

Solar radiation can affect the heat balance of flying birds (*17*), potentially exposing them to the risk of overheating. Increased body temperature has been shown to limit flight duration in ducks (*18*), and birds in warm tropical regions more often fly with trailing legs when exposed to strong solar radiation, most likely to increase the rate of heat loss (*19*). Risk of hyperthermia caused by solar radiative heat gain may constrain flight activity during daylight in bats (*20*, *21*), which is further supported by elevated flight costs in terms of increased body temperature and metabolic rate in bats flying at day compared with at night (*22*). Thus, we propose the hypothesis that migrating birds climb to extremely high altitudes in the daytime (Fig. 1A) to reach much colder conditions (Fig. 1D) to mitigate the risk of solar radiation–generated heat stress, allowing for a faster crossing of vast ecological barriers. This hypothesis calls for future investigations of the thermal and behavioral consequences of solar radiation on songbird migration.

Here, we demonstrate an unexpected and pronounced pattern in which songbirds, when prolonging their nocturnal migratory flights into daytime, climb steeply at dawn to reach astonishingly high cruising altitudes of 5000 to 6300 m asl. Future studies of differences in the physical and biological conditions encountered by the birds during nocturnal versus diurnal flights are needed to reveal the causes of this behavior. This finding sheds new light on constraints on migratory flights and may help to explain why the overwhelming majority of nonsoaring long-distance migrant birds generally conduct their migratory flights during the night.

REFERENCES AND NOTES

1. G. Malmiga, M. Tarka, T. Alerstam, B. Hansson, D. Hasselquist, *J. Avian Biol.* **52**, e02549 (2021).
2. J. Ouwehand, C. Both, *Biol. Lett.* **12**, 20151060 (2016).
3. B. Bruderer, D. Peter, F. Korner-Nievergelt, *J. Ornithol.* **159**, 315–336 (2018).
4. F. Liechti, B. Bruderer, *J. Avian Biol.* **29**, 561–568 (1998).
5. F. Liechti, M. Klaassen, B. Bruderer, *Auk* **117**, 205–214 (2000).
6. M. S. Bowlin *et al*., *Auk* **132**, 808–816 (2015).
7. S. Sjöberg *et al*., *J. Avian Biol.* **49**, e01821 (2018).
8. N. R. Senner *et al*., *Proc. Biol. Sci.* **285**, 20180569 (2018).
9. F. Liechti *et al*., *Mov. Ecol.* **6**, 19 (2018).
10. J. Koleček *et al*., *J. Avian Biol.* **47**, 756–767 (2016).
11. J. Bäckman *et al*., *J. Avian Biol.* **48**, 309–319 (2017).
12. J. R. Torre-Bueno, *J. Exp. Biol.* **75**, 231–236 (1978).
13. J. M. Wallace, F. R. Hartranft, *Mon. Weather Rev.* **97**, 446–455 (1969).
14. D. J. Seidel, M. Free, J. Wang, *J. Geophys. Res. Atmos.* **110**, D09102 (2005).
15. U. Mellone, P. López-López, R. Limiñana, G. Piasevoli, V. Urios, *J. Avian Biol.* **44**, 417–426 (2013).
16. S. M. Xirouchakis, M. Panuccio, *J. Raptor Res.* **53**, 56 (2019).
17. B. O. Wolf, G. E. Walsberg, *Am. Zool.* **40**, 575–584 (2000).
18. M. Guillemette *et al*., *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150386 (2016).
19. D. M. Bryant, *Ibis* **125**, 313–323 (1983).
20. J. R. Speakman, E. Król, *J. Anim. Ecol.* **79**, 726–746 (2010).
21. J. Rydell, M. B. Fenton, E. Seamark, P. W. Webala, T. C. Michaelsen, *Can. J. Zool.* **98**, 149–156 (2020).
22. C. C. Voigt, D. Lewanzik, *Proc. Biol. Sci.* **278**, 2311–2317 (2011).
23. QGIS Development Team, "QGIS geographic information system" (QGIS, 2019); https://qgis.org.
24. S. Sjöberg, G. Malmiga, A. Nord, A. Andersson, J. Bäckman, M. Tarka, M. Willemoes, K. Thorup, B. Hansson, T. Alerstam, D. Hasselquist, Data for: Extreme altitudes during diurnal flights in a nocturnal songbird migrant, Dryad (2021); https://doi.org/10.5061/dryad.q2bvq83j9.

ACKNOWLEDGMENTS

We thank L. Aranda, M. Lapa, J. Roved, D. Gómez Blanco, A. Jara Navarro, V. Caballero, and the Kvismare Bird Observatory (report 197) for help with fieldwork; meteorologist B. Larsson for pointing out and explaining the potential importance of solar radiation for the heat balance of a flying bird at different altitudes; and four anonymous referees for valuable comments and suggestions. **Funding:** This project received funding from the European Union's Horizon 2020 research and innovation program (Marie Sklodowska Curie grant 751692 to S.S., ERC Advanced grant 742646 to D.H.), the Danish National Research Foundation (grant DNRF96 to the Center for Macroecology, Evolution and Climate), the Swedish Research Council [consolidator grant 621-2016-689 to B.H., research grants 2016-04391 and 2020-03976 to D.H., and Linnaeus grant 349-2007-8690 to the Center for Animal Movement research (CAnMove), Lund University], the Birgit and Hellmuth Hertz Foundation/Royal Physiographic Society of Lund (grant 2017-39034 to A.N.), and Lunds Djurskyddsfond. **Author contributions:** D.H., B.H., and M.T. conceived and designed the study. A.A. and J.B. designed and built the multisensor data loggers. G.M., B.H., M.T., and D.H. performed fieldwork and gathered the tracking data. S.S., D.H., and T.A. analyzed and interpreted the data with assistance from G.M., A.N., M.W., and K.T. S.S. wrote the draft of the manuscript with input from G.M., A.N., T.A., and D.H. All authors read, revised, and commented on the manuscript and approved the final version. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data supporting the results of this study are available from Dryad (24).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6542/646/suppl/DC1
Materials and Methods
Tables S1 to S3
References (*25*–*30*)

11 September 2020; accepted 25 March 2021
10.1126/science.abe7291



Differential Refractometer

The HK Series Differential Refractometer from Testa Analytical Solutions sets a benchmark for precise determination of specific refractive index increment (dn/dc)—an important parameter for polymer and protein molecular weight determination using light-scattering techniques. It is a flexible device that may be used in either static or dynamic mode. The new system offers several wavelength options enabling precise matching of your light-scattering detector laser operating wavelength. Where the dn/dc value of the sample is already known, the unit provides precise, sensitive concentration determination. The amount of sample used for this determination can also be easily recollected and therefore utilized for further investigations.

Testa Analytical Solutions
For info: +49-(0)-30-864-24076
www.testa-analytical.com/gpc-sec-chromatography.html

Antigens and Antibodies for Veterinary Diseases

The Native Antigen Company announces its new range of veterinary reagents, including over 90 high-quality native and recombinant antigens as well as monoclonal and polyclonal antibodies for a range of viral, bacterial, and protozoal veterinary diseases, all of which can be used in the development of diagnostics and vaccines. Our reagents have been developed for both livestock and companion animals, such as dogs, cats, and pigs, and include a range of clinically relevant pathogens, such as coronaviruses, *Borrelia*, Rift Valley fever virus, African swine fever virus, *Toxocara*, and heartworm. Suitable for a wide range of applications, these reagents can be used in research as well as the development of immunoassays for diagnostic and vaccine R&D.

The Native Antigen Company
For info: +44-(0)-1865-595230
thenativeantigencompany.com/veterinary-reagents

Display Antibody Discovery Platform

OXGENE announces the introduction of its self-labeling integral membrane (SLIM) protein discovery system. This novel method allows for the identification of antibodies against membrane proteins in their native configuration in mammalian cells, which could catalyze the discovery of new candidates for antibody-based therapies. In the SLIM display, a target membrane protein and a library of potential binders are expressed in the same mammalian cells. These then progress through multiple rounds of enrichment and purification to obtain a single clone expressing a single-chain variable fragment that can bind the target antigen. The SLIM display implements a specific strategy to reduce false positives and ensures that confirmed binders only target physiological epitopes. The platform also uses Chinese hamster ovary cells throughout discovery, so binders are far more likely to be compatible with downstream processing.

OXGENE
For info: +44-(0)-1865-415-107
www.oxgene.com/antibody

Reversed-Phase Column for HPLC and LC/MS Analysis

Use the Thermo Fisher Scientific MAbPac Capillary Reversed-Phase HPLC Column for high-sensitivity HPLC or LC/MS characterization of intact proteins in top-down proteomics applications where sample amount is critically limited. This versatile column is also ideal for peptide-mapping characterization, as its capillary format balances sensitivity with throughput. Additional applications are monoclonal antibodies (mAbs), fragments, variants, antibody drug conjugates, PEGylated proteins, and bispecific proteins. These columns have excellent resolution, long column lifetime, and low carryover. Proteomics, clinical peptides, and biopharmaceutical research laboratories can now achieve high-resolution, low-flow separations for intact proteins, mAbs, and peptides with minimal sample volume.

Thermo Fisher Scientific
For info: 800-955-6288
www.thermofisher.com/capLC

Cytosections for Antibody Screening

AMS Biotechnology (AMSBIO) has launched a comprehensive range of cytosections that offer a verified, reproducible, and renewable source of positive/negative controls where the expression of the target biomarker is confirmed for accuracy and specificity by an immunoassay. Researchers, antibody and assay developers, and immunohistochemistry (IHC) labs often have difficulty finding a reliable and verified source of positive/negative controls for their IHC experiments. AMSBIO's new cytosections range consists of high-quality FFPE sections of cell pellets of overexpressing targeted biomarkers. The expression of target biomarkers in these cytosections is carefully verified for accuracy and specificity using Western blot. Positive controls (cells transfected with the target protein) and negative controls (mock transfected cells) are available for genome-wide human and mouse genome targets. In addition to its standard range, AMSBIO also offers a custom cytosections service where cell-pellet composition and density can be designed according to the customer's needs.

AMS Biotechnology
For info: 617-945-5033
www.amsbio.com/cytosections

Temperature Gradient System for DART-MS

The ionRocket from BioChromato is a temperature-controlled heating device for direct thermal desorption and pyrolysis of samples prior to ionization and analysis by mass spectrometry. ionRocket generates a temperature gradient from ambient up to 600°C in just a few minutes. This enables compounds in your samples to be sublimated, vaporized, or pyrolyzed according to their volatility, and then introduced into the direct analysis in real time-mass spectrometry (DART-MS) gas stream. The system offers the option to weigh samples prior to analysis, resulting in reproducible MS intensity and allowing you to calculate calibration curves of analyte amount versus peak intensity. Data obtained from ionRocket yields another axis of data (time/temperature) beyond that obtained from normal DART-MS analysis.

BioChromato
For info: +81-(0)-466-23-8382
biochromato.com/analyticalinstruments/ionrocket

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.

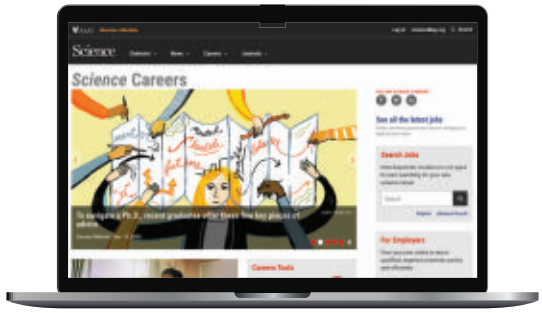
ScienceCareers online @sciencecareers.org



Science Careers helps you advance your career. Learn how!

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete an interactive, personalized career plan at "my IDP."
- Visit our Employer Profiles to learn more about prospective employers.
- Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org** today — all resources are free



SCIENCECAREERS.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

上海商學院
Shanghai Business School

Introduction to Shanghai Business School

Shanghai Business School, formerly East China Branch of Tax Administration School of the Central Government established in 1950, is a state-run application-oriented university offering undergraduate education, with the distinctive feature of business studies. Adhering to the school-running philosophy of "business-based and application-oriented" and the motto of "govern and benefit the people by being virtuous and learned", Shanghai Business School strives to cultivate high-quality application-oriented business talents with the sense of social responsibility, professionalism, practical abilities, innovative spirit, and international visions.

Shanghai Business School has now set up three campuses in Fengpu and Xuhui Districts and on Guoquan Road, and a school-running station on Fuzhou Road, with nearly 10,000 full-time students. It has established 10 secondary schools and 30 undergraduate disciplines. Among them, three disciplines, including business administration, hotel management, and e-commerce, are national first-class undergraduate majors construction sites; the major of finance is Shanghai municipal first-class undergraduate majors construction site; nine majors, including taxation, etc., are Shanghai municipal application-oriented undergraduate pilot majors; and four majors, including business administration, hotel management, e-commerce, and marketing, have passed ACBSP international certification.

The School has been continuously deepening the connotation of disciplinary construction, and formed a discipline layout with applied economics and business management as the main body and coordinated development of multiple disciplines. The applied economics discipline of the School is a Shanghai class-II plateau discipline; the business communication is a key discipline of Shanghai Municipal Education Commission; and the business administration discipline has been selected into the Shanghai first-class university disciplines (class B) cultivation program. In the past five years, the School has won many awards such as National Social Science Fund of China, the National Natural Science Foundation of China, the humanities and social sciences project of the national education authority, the horizontal scientific research project, and Shanghai policy-making consulting research achievement award.

The School focuses on national and local strategic demands and major industry demands to carry out business (commercial) applied research and consulting services, with its municipal scientific research bases and platforms including Shanghai Business Development Institute, Shanghai University Think Tank, Shanghai Social Science Innovation Research Base, Shanghai Key University Humanities and Social Science Research Base, and Shanghai University Knowledge Service Platform. In the past five years, 43 business decision-making consulting achievements of the School received instructions given by provincial and ministerial leaders, and the School presided over and participated in the formulation of logistics, tourism and other industry standards. The MOFCOM Training Base for International Business Officials(Shanghai) established by the School has trained an accumulative amount of more than 4,400 business officials from more than 100 countries (regions).

The School insists on opening school-running and deepening international exchanges and cooperation. It has successively carried out cooperation and exchanges with more than 90 universities and research institutes from the United States, the United Kingdom, Australia and other countries and regions. It has also initiated the establishment of the "Belt and Road" International Business Education Alliance, established the International High-end Think Tank Alliance, and set up the Central and Eastern Europe Overseas Education Base.

Forum time

Shanghai Business School plans to hold the 2nd International Youth Scholars "SBS" Cloud Forum in October.

Remuneration and benefits

Shanghai Business School provides competitive remuneration and benefits, and one-time talent introduction fees, which will be determined in the manner of one discussion for one case.

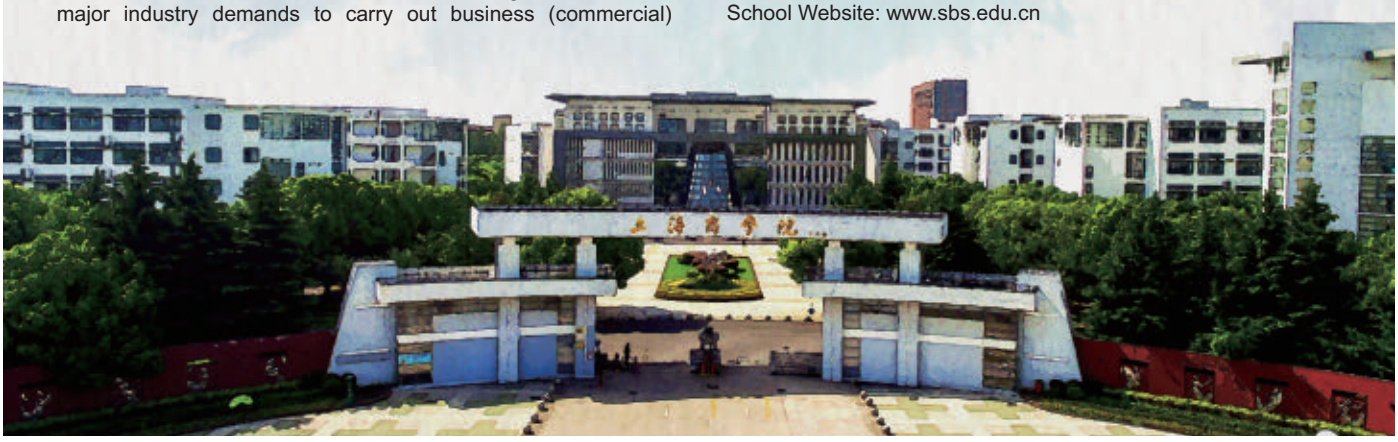
Contact information

Contact Person: Mr. Li

Tel: 021-67105340

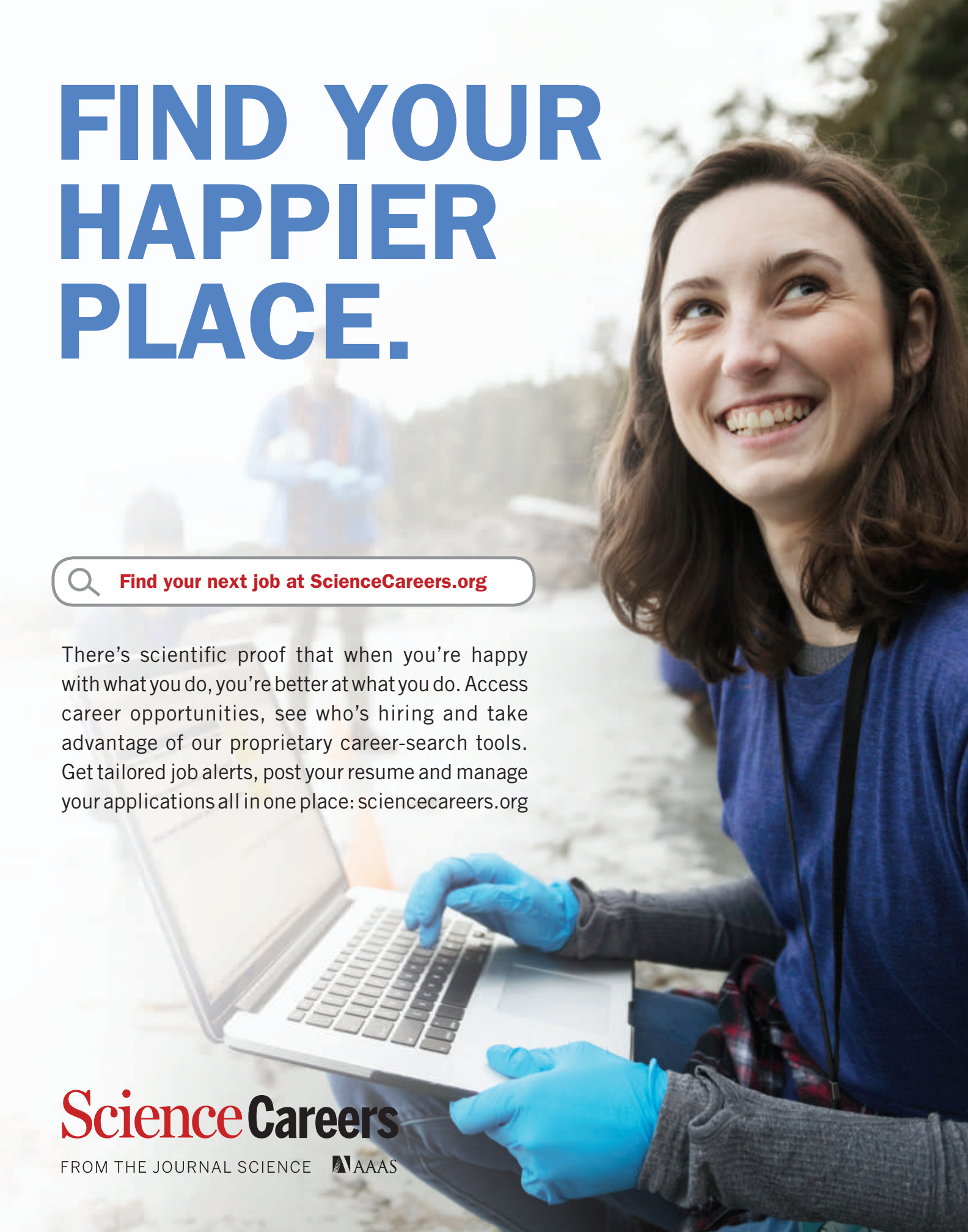
E-mail: sbsrsc@126.com

School Website: www.sbs.edu.cn



ScienceCareers online @sciencecareers.org
OPPORTUNITIES IN CHINA

FIND YOUR HAPPIER PLACE.



Find your next job at ScienceCareers.org

There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: sciencecareers.org

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

LSU Health NEW ORLEANS
School of Medicine

WILLIAM & SARAH JANE PELON ENDOWED CHAIR DEPARTMENT OF MICROBIOLOGY, IMMUNOLOGY & PARASITOLOGY

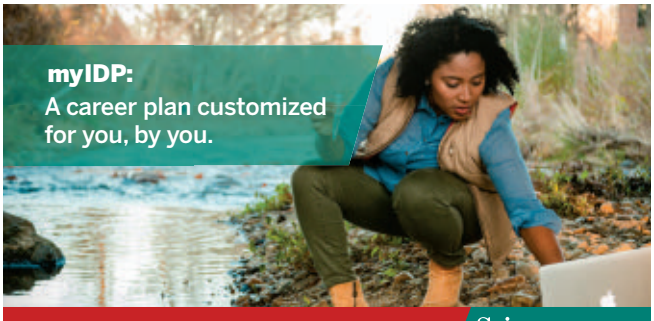
LSU Health Sciences Center School of Medicine is initiating a search for a senior researcher to join the Department of Microbiology, Immunology and Parasitology. Candidates will be considered for appointment as Professor on the tenure track and should have a strong record of research accomplishment, lead an active, nationally funded research program, and have a commitment to developing collaborative translational research programs. This position is associated with the endowed William and Sarah Jane Pelon Chair, designed to enhance the scholarly productivity of the incumbent.

The ideal candidate will have a Ph.D. and/or M.D., demonstrated team-building ability, and a strong track record of developing translational research from basic scientific observations. Expertise in all areas of host/pathogen interaction will be considered, but special consideration will be given to those complementing existing core departmental strengths in HIV, HIV-related infections, and sexually transmitted infections.

LSU Health Sciences Center School of Medicine offers a highly interactive and collegial environment, with a strong history of collaborative research programs and state-of-the art infrastructure, including core laboratories in genomics, proteomics, bioinformatics, imaging, and flow cytometry. Excellent opportunities exist for interaction with clinical departments, and with research Centers of Excellence in Vaccine Development, Cancer, Alcohol and Drug Abuse, Cardiovascular Biology, and Neuroscience. Anticipated duties and responsibilities will include sustaining an exceptional research program, mentoring graduate students and post-doctoral fellows, and participating in departmental and school graduate and undergraduate teaching programs. The institution offers competitive start-up packages and salaries.

Qualified applicants with a substantial record of scientific achievement should send a single PDF document containing their curriculum vitae including details of publications, previous and current research funding, teaching experience, a statement of research plans, and the names of at least three referees through the following link located on the LSU Career website: **https://lsuhsc.peopleadmin.com/postings/6731**

The LSUHSC School of Medicine in New Orleans encourages women and minority candidates to submit applications for this position. LSUHSC is an Equal Opportunity Employer for females, minorities, individuals with disabilities and protected veterans.



For your career in science, there's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

ScienceCareers AAAS In partnership with:

FASEB Federation of American Societies for Experimental Biology
University of Massachusetts Medical School UMASS
UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND

EPFL

FACULTY POSITION IN SUSTAINABLE URBAN SYSTEMS AND CLIMATE MITIGATION STRATEGIES

AT THE ECOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE (EPFL)

EPFL's School of Architecture, Civil and Environmental Engineering (ENAC) invites applications for a faculty position in Sustainable Urban Systems and Climate Mitigation Strategies. The position is open at the level of **Assistant Professor** (Tenure Track).

Climate change poses threats to the sustainable development of cities and their related rural areas. Climate change-induced warming is 50-150% greater in urban areas than in non-urban areas. It leads to the formation of urban heat islands, which can reduce the quality of life. Furthermore, about 80% of the CO2 emissions are produced in urban territories. These two phenomena make cities the place for high-impact measures with regard to climate mitigation and sustainable development of the built environment.

The School of Architecture, Civil and Environmental Engineering at EPFL aims to address climate change and sustainable urbanization challenges. We encourage applications from individuals with a strong profile in urban climate mitigation and sustainable development of urban territories. The professor will contribute to re-thinking and design of new climate mitigation technologies and strategies that target urban systems. Potential research areas include, but are not limited to, data-driven lifecycle assessment for urban systems and management of territorial metabolism, urban water system design, circular economy for urban systems, dynamic urban metabolism, heat islands and CO2 emissions mitigation. Candidates are expected to have excellent skills in digital tools and related methodologies, and will contribute ENAC's interdisciplinary research, e.g., through the CLIMACT centre or the ENAC cluster "Sustainable Territories". He/she is expected to contribute to interdisciplinary teaching projects with colleagues from Architecture, Civil and Environmental Engineering as well as other schools of EPFL. The professor will be attached to the Institute of Environmental Engineering and the Institute of Architecture and the City.

EPFL is an internationally leading institution in environmental engineering as well as urban and territorial design. With its main campus located in Lausanne and its developing campuses in neighbouring cantons in Switzerland, EPFL is a growing and well-funded institution fostering excellence and diversity. It is well equipped with experimental and computational infrastructure, and offers a fertile environment for research collaborations between different disciplines. The EPFL environment is multilingual and multicultural, with English serving as a common interface. EPFL offers internationally competitive start-up resources, salaries, and benefits.

The following documents are requested in PDF format: cover letter that includes a statement of motivation, curriculum vitae, publications list, research vision, statement of teaching interests, and the names and addresses, including emails, of at least three references (may be contacted at a later stage). Applications should be uploaded to the EPFL recruitment web site:

https://facultyrecruiting.epfl.ch/position/31391745

Formal evaluation of the applications will begin on **June 6, 2021** and the search will continue until the position is filled.

Further enquiries should be made to:
Prof. Claudia R. Binder, ENAC Dean or
Prof. D. Andrew Barry, Chair of the Search Committee

e-mail: **SearchSUS-CMS@epfl.ch**

For additional information on EPFL, please consult:
http://www.epfl.ch or http://enac.epfl.ch

EPFL is an equal opportunity employer and a family-friendly university. It is committed to increasing the diversity of its faculty, and strongly encourages women to apply.

By **Anne Crecelius**

From professor to patient X

I walked into the classroom feeling nervous. It wasn't my first time teaching undergraduate students about human endocrine physiology. I knew the material well. But today's lecture was different. I pulled up slides depicting a hypothetical cancer patient and told them, "Patient X had a biopsy that detected invasive carcinoma in her breast." I described the many months of chemotherapy, surgery, and radiation treatments she went through before going into remission. Then I taught the students about the hormonal therapy she was prescribed—drugs her doctor hoped would limit the growth of any remaining cancer cells and prevent a recurrence. On the final slide, I showed them a picture of me on my last day of chemotherapy. "I am patient X," I revealed.

When I started my faculty position, I never would have dreamed of giving a lecture that delved into my personal medical history. I made a point of presenting myself as a consummate professional to my students and colleagues. I was only 28 years old—fresh out of my Ph.D. program—and I wanted to be respected as a professor.

But that all changed in the second year of my faculty position, when I was diagnosed with an aggressive form of breast cancer. I continued to work while I underwent treatment, which meant that signs of my chemotherapy, such as hair loss, were clearly visible. I told undergraduate students in my classes about my condition, in part to cut short curious stares but also to warn them that my treatment could interfere with my teaching.

I appreciated the compassion and care those students showed me—putting together care packages, leaving cards, praying for my recovery. I was taken aback by their maturity and empathy—and reminded that they are people, too. They may have a mother or grandmother who suffered from breast cancer. They may have religious beliefs that call them to pray.

When my hair grew back, I didn't want to return to being the same professor I was before. I wanted my students to see me as a person first and a professor second, and I hoped my story would show them that the science they were learning had relevance in the real world. So 1 year after my diagnosis, I developed my lecture about patient X.

For nearly 5 years, I gave that lecture, with the positive news that I was in remission: "The power of medicine!" I'd tell them. The nerves never went away—I found it emotionally trying to talk about my body and my disease—but I persisted because the lecture was clearly making an impact. Each time, a few students would stay after class or send me an email, making it clear they appreciated my honesty.



"I wanted my students to see me as a person first and a professor second."

Then, in March 2020—just as most people's lives were upended by the COVID-19 pandemic—I got a call from my oncologist. My cancer was back, and it had spread. I was now facing a diagnosis of stage 4 metastatic disease. It was hard news to hear, but I decided to continue my teaching, research, and service activities because I felt the same physically as I did before. I also thought my work would be a welcome distraction from the stress of doctor's visits, tests, and scans.

It was time to update patient X's story. At first, I added details about new lines of hormonal treatment and an unfulfilling ending of "time will tell." More recently, I added somewhat more optimistic news: Patient X's third-line treatment of infusion chemotherapy seems to be working a bit. The metastatic nodules in her lungs have decreased in size and number, and she might have access to new treatment options in the future.

My diagnosis has likely shortened the number of years I will have to build and leave a legacy as a scientist, mentor, and teacher. Yet, it has also given me a powerful tool. It helps me connect with my students on a personal level, tap into their empathy, and show them why the material they're learning matters. It isn't easy to be vulnerable with my students. But I believe my personal story is one of the most important lessons I teach all semester. ■

Anne Crecelius is an associate professor at the University of Dayton.

ILLUSTRATION: ROBERT NEUBECKER

NOMIS & Science
YOUNG EXPLORER AWARD



Research at the intersection of the social and life sciences

Unconventional. Interdisciplinary. Bold.

The newly launched **NOMIS & *Science* Young Explorer Award** recognizes bold, early-career researchers with an M.D., Ph.D., or M.D./Ph.D. who ask fundamental questions at the intersection of the social and life sciences. It is awarded to scientists who conduct research with an enthusiasm that catalyzes cross-disciplinary collaboration and who take risks to creatively address relevant and exciting questions.

Award winners will receive a cash prize of up to $15,000, and their essay submissions will be published in *Science*. They will also be invited to share their ideas with leading scientists in their respective fields at an award ceremony and to present their research to the NOMIS Board of Directors for potential future funding.

Apply by May 15, 2021
at www.sciencemag.org/nomis

Science
AAAS

NOMIS
FOUNDATION
Creating the Spark

AAAS AWARDS & PRIZES

HONORING EXCELLENCE IN SCIENCE

AAAS DAVID AND BETTY HAMBURG AWARD FOR SCIENCE DIPLOMACY

The **AAAS David and Betty Hamburg Award for Science Diplomacy** recognizes an individual or a limited number of individuals working together in the scientific and engineering or foreign affairs communities who are making an outstanding contribution to furthering science diplomacy.

This year, the award has been renamed for David and Betty Hamburg to recognize their unparalleled commitment to the significant role of science diplomacy to advance science, human rights, peace, and cooperation.

- The monetary prize that the award winner(s) receives has been doubled to a total of $10,000.
- We accept group nominations.
- The award is open to all regardless of nationality or citizenship.
- We accept self-nominations.

SUBMIT A NOMINATION TODAY
Deadline June 30, 2021

RECOGNIZING PAST AWARD WINNERS

2019: Group Award Winner – SESAME

Five architects of SESAME were recognized for their dedication and commitment to building SESAME as a hub for research and understanding in the Middle East.



Christopher Llewellyn Smith



Eliezer Rabinovici



Zehra Sayers



Herwig Schopper



Khaled Touka

To learn more, visit
aaas.org/awards/science-diplomacy/about

AAAS gratefully acknowledges the Carnegie Corporation of New York for their generous support to launch the AAAS David and Betty Hamburg Award for Science Diplomacy and the individuals and foundations whose contributions have begun an endowment that will allow us to sustain it in perpetuity.

AAAS
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Photo Credits: Christopher Llewellyn Smith: Oxford University / Eliezer Rabinovici: Michael Yakirevich Zehra Sayers: Zehra Sayers / Herwig Schopper: Herwig Schopper / Khaled Touka: Khaled Touka